Supplemental Material 1:

Results overview full sample

Here, we report an overview over all major statistical results reported in the main text and the supplementary material. These results are based on all participants in both samples. For details on how mixed-effects regression were performed, see the Methods section of the main text.

	DV	IV	Sample	b	SE	$\chi^2(1)$	p
Task performance	Go/ NoGo	Required action	1	1.075	0.097	53.191	< .001
			2	1.265	0.091	89.190	< .001
Effect of stake valence and magnitude on action (i.e., Pavlovian bias)	Go/ NoGo	Stake difference	1	0.117	0.027	15.320	< .001
			2	0.092	0.031	7.916	.005
		Reward stake	1	0.135	0.028	20.791	< .001
			2	0.081	0.027	8.151	.004
		Punishment stake	1	-0.051	0.026	3.301	.069
			2	-0.063	0.028	4.707	.030
	RT	Stake difference	1	-0.041	0.015	7.323	.007
			2	-0.025	0.011	6.313	.012
		Reward stake	1	-0.028	0.014	3.983	.046
			2	-0.012	0.010	0.031	.861
		Punishment stake	1	0.034	0.017	4.012	.045
			2	0.029	0.011	7.311	.006
Effect of attention on action (Go/ NoGo and Go RTs)	Go/ NoGo	Dwell time difference	1	0.132	0.034	12.203	< .001
			2	0.192	0.032	28.443	< .001
		Dwell time ratio	1	0.140	0.031	15.331	< .001
			2	0.221	0.039	27.528	< .001
		Reward dwell time	1	0.035	0.034	0.945	.331
			2	0.069	0.031	4.617	.032
		Punishment dwell time	1	-0.185	0.037	18.042	< .001
			2	-0.278	0.041	35.080	< .001
		First fixation on rewards	1	-0.053	0.025	4.495	.034
			2	-0.059	0.022	7.164	.007
	RT	Dwell time difference	1	-0.036	0.026	1.900	.168
			2	-0.030	0.012	4.533	.033
		Dwell time ratio	1	-0.032	0.026	1.489	.222
			2	-0.030	0.014	4.429	.035
		Reward dwell time	1	-0.034	0.027	1.619	.203
			2	0.013	0.015	0.757	.384
		Punishment dwell time	1	0.027	0.028	0.939	.333
			2	0.039	0.013	7.668	.006
		First fixation on rewards	1	-0.010	0.016	0.255	.613
			2	0.008	0.011	0.461	.497
Effect of required action on attention (first fixation and dwell time)	First fixation	Required action	1	0.113	0.035	13.915	< .001
			2	0.090	0.028	7.882	.005
		Q-value difference	1	0.091	0.033	8.044	.005
			2	0.132	0.039	9.171	.002
	Dwell time diff.	Required action[1]	1	0.030	0.010	4.711	.030
			2	0.032	0.008	13.791	< .001
		Q-value difference[1]	1	0.026	0.010	4.361	.037
			2	0.039	0.008	24.823	< .001
	Dwell time ratio	Required action[1]	1	0.026	0.009	6.896	.009
			2	0.030	0.007	15.364	< .001
		Q-value difference[1]	1	0.016	0.011	0.951	.329
			2	0.036	0.007	13.231	< .001

[1]Controlling for first fixation and the stake difference. All effects are significant with required action/ Q-value difference as sole predictor.

Supplemental Material 2:

Results overview: Participants not significantly above chance excluded

We report an overview over all major statistical results as reported in the main text and the supplementary material, but excluding the five (seven) participants in Sample 1 (2) that did not perform significantly above chance level, i.e., did not learn the task. For details on how mixed-effects regression were performed, see the Methods section of the main text. These analyses led to the same conclusions as the analyses based on the full samples reported in S01.

	DV	IV	Sample	b	SE	$\chi^2(1)$	p
Task performance	Response	Required action	1	1.230	0.076	68.376	< .001
			2	1.422	0.077	111.816	< .001
Effect of stake valence and magnitude on action (i.e., Pavlovian bias)	Response	Stake difference	1	0.130	0.030	14.830	< .001
			2	0.092	0.035	6.434	.011
		Reward stake	1	0.146	0.029	21.802	< .001
			2	0.078	0.030	6.072	.014
		Punishment stake	1	-0.058	0.030	3.543	.060
			2	-0.066	0.031	4.209	.040
	RT	Stake difference	1	-0.045	0.016	8.068	.005
			2	-0.031	0.013	5.828	.016
		Reward stake	1	-0.036	0.016	4.887	.027
			2	-0.015	0.011	1.208	.272
		Punishment stake	1	0.029	0.016	3.123	.077
			2	0.034	0.012	7.560	.006
Effect of attention on action (Go/ NoGo and Go RTs)	Response	Dwell time difference	1	0.142	0.037	10.442	.001
			2	0.205	0.032	30.129	< .001
		Dwell time ratio	1	0.144	0.035	12.762	< .001
			2	0.237	0.040	27.436	< .001
		Reward dwell time	1	0.033	0.040	0.593	.441
			2	0.078	0.033	5.158	.023
		Punishment dwell time	1	-0.202	0.038	19.051	< .001
			2	-0.301	0.043	35.949	< .001
		First fixation on rewards	1	-0.060	0.027	4.410	.036
			2	-0.064	0.023	7.490	.006
	RT	Dwell time difference	1	-0.009	0.026	0.122	.727
			2	-0.029	0.013	4.557	.033
		Dwell time ratio	1	-0.014	0.024	0.335	.551
			2	-0.025	0.013	3.731	.053
		Reward dwell time	1	-0.005	0.027	0.042	.838
			2	-0.016	0.016	0.977	.323
		Punishment dwell time	1	0.012	0.029	0.165	.685
			2	0.031	0.014	5.175	.023
		First fixation on rewards	1	-0.003	0.018	0.023	.881
			2	0.009	0.012	0.478	.490
Effect of required action on attention (first fixation and dwell time)	First fixation	Required action	1	0.106	0.034	9.417	.002
			2	0.097	0.032	6.955	.008
		Q-value difference	1	0.091	0.036	6.892	.009
			2	0.135	0.043	7.689	.006
	Dwell time diff.	Required action[1]	1	0.037	0.011	9.913	.002
			2	0.034	0.010	11.465	< .001
		Q-value difference[1]	1	0.029	0.012	4.140	.042
			2	0.040	0.008	22.650	< .001
	Dwell time ratio	Required action[1]	1	0.035	0.010	15.359	< .001
			2	0.032	0.008	14.013	< .001
		Q-value difference[1]	1	0.020	0.012	2.090	.148

	2	0.032	0.008	11.304	< .001

[1]Controlling for first fixation and the stake difference. All effects are significant with required action/ Q-value difference as sole predictor.

Supplemental Material 3:

Effects of stake magnitudes on responses and reaction times

Given that stake differences (reward minus punishment stake) affected both Go/ NoGo responses and reaction times, we additionally tested for separate effects of the reward and punishment stake magnitude on responses and reaction times using in mixed-effects logistic regressions (for Go/ NoGo responses) and linear regressions (for reaction times). We coded reward and punishment stake magnitudes as separate regressors (instead of as their difference).

The effect of reward stake magnitude on responses was significant in both samples (Sample 1: $b = 0.14$, 95% CI [0.08 0.19], $\chi^2(1) = 20.79$, $p < .001$; Sample 2: $b = 0.08$, 95% CI [0.03 0.13], $\chi^2(1) = 8.15$, $p = .004$; Fig. S03A), while the effect of punishment stake magnitude was only significant in Sample 2 (Sample 1: $b = -0.05$, 95% CI [-0.10 0.001], $\chi^2(1) = 3.30$, $p = .069$; Sample 2: $b = -0.06$, 95% CI [-0.12 -0.01], $\chi^2(1) = 4.71$, $p = .030$; Fig. S03B). In contrast, for RTs, higher reward stake magnitude predicted faster responses only in Sample 1 (Sample 1: $b = -0.03$, 95% CI [-0.06 -0.001], $\chi^2(1) = 3.98$, $p =.046$; Sample 2: $b = -0.01$, 95% CI [-0.03 0.01], $\chi^2(1) = 0.03$, $p = .861$; Fig. S03C), while higher punishment stake magnitude consistently predicted slower responses (Sample 1: $b = 0.03$, 95% CI [0.001 0.07], $\chi^2(1) = 4.01$, $p = .045$; Sample 2: $b = 0.03$, 95% CI [0.01 0.05], $\chi^2(1) = 7.31$, $p = .006$; Fig. S03D). Note that RTs are only available for Go responses; hence, the amount of data (and resulting statistical power) are lower compared to the Go/ NoGo response data.

In conclusion, effects of stake magnitude on driving Pavlovian biases reported in the main manuscript were driven by variations in both the reward and the punishment stake. These effects resemble effects of Pavlovian biases reported before, but in this study emerged in a graded fashion, i.e., more and faster Go responding the larger the reward stake was, and less and slower Go responding the larger the punishment stake was.

In addition, we tested whether the effect of stake difference on responses (i.e., the Pavlovian bias) became weaker over time. For this purpose, we used mixed-effects logistic regression models including stake difference, time, and their interaction. As time, we either used a) trial number across the whole task (1–264), b) trial number within each block (1–88), c) cue repetition number (1–22), or d) block number (1–3). A significant interaction would indicate that the Pavlovian bias changes over time. However, we did not find any significant interaction, neither a) for trial number across the task (Study 1: b = -0.04, 95% CI [-0.10 0.02], $\chi2(1) = 2.55$, p = .110; Study 2: b = -0.03, 95% CI [-0.07 0.01], $\chi2(1) = 1.49$, p = .222) nor b) for trial number within blocks (Study 1: b = -0.02, 95% CI [-0.08 0.04], $\chi2(1) = 0.61$, p = .433; Study 2: b = -0.02, 95% CI [-0.06 0.02], $\chi2(1) = 1.11$, p = .293), nor c) by cue repetition number (Study 1: b = -0.02, 95% CI [-0.07 0.04], $\chi2(1) = 0.27$, p = .597; Study 2: b = -0.02, 95% CI [-0.07 0.02], $\chi2(1) = 0.89$, p = .345), nor d) for block number (Study 1: b = -0.03, 95% CI [-0.09 0.03], $\chi2(1) = 1.14$, p = .286; Study 2: b = -0.02, 95% CI [-0.06 0.02], $\chi2(1) = 0.56$, p = .455). Numerically (but not significantly), the bias got weaker with time, which is to be expected given that people make less errors over time, while errors are necessary to detect the presence of a Pavlovian bias. In sum, we found no evidence that the Pavlovian bias vanishes over time.

Of note, in our pre-registration, we mentioned under “exploratory analyses” that we would fit reinforcement-learning drift diffusion models (RL-DDMs) to jointly analyze the effects of stakes/ dwell times on choices and RTs. We decided to not report the results from these models because data simulated from them was markedly different from the empirical data. We suspect that DDMs cannot capture data from this task due to i) the tight response deadline (600 ms), leading to overall fast (but regularly incorrect) responses while preventing late responses, and ii) the absence of RTs for the NoGo responses, which can be computationally dealt with, but which implies a lack of constraint on the parameters (especially the starting point bias term). Lastly, enforcing a strict response threshold is not possible in the DDM framework. Potentially, evidence accumulation frameworks in which the response thresholds decrease and eventually become zero at the response deadline might be able to accommodate such data, but likelihood functions for such models are not readily available. We encourage other researchers to reanalyze this data with more suitable modeling frameworks that might arise in the future.

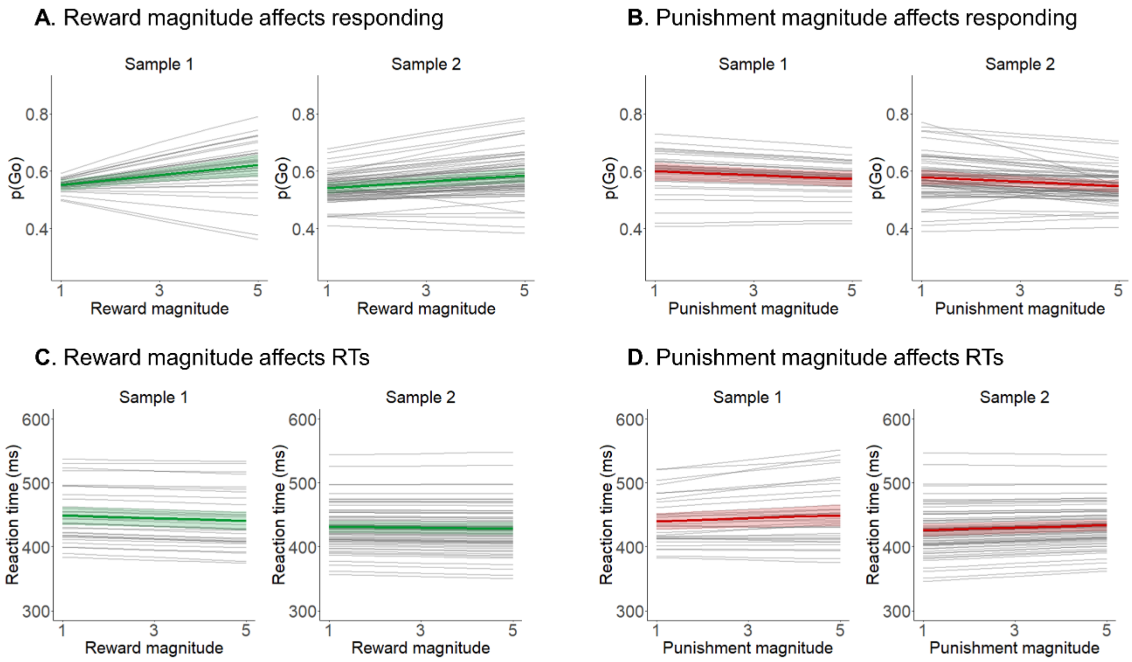

Figure SI03. Effect of stake magnitudes on responses and reaction times. A higher reward stake magnitude led to a higher proportion of Go responses (**A**; significant in both studies), while a higher punishment stake magnitude led to a lower proportion of Go responses (**B**; only significant in Study 2). Similarly, a higher reward stake magnitude tended to speed up reaction times (**C**; significant only in Study 1), while a higher punishment stake magnitude tended to slow down reaction times (**D**; significant in both studies).

Supplemental Material 4:

Effect of action plans on attentional measures

As our first key prediction, we tested whether attention allocation to reward and punishment stake was affected by action requirements. For this purpose, we regressed attention measures (first fixation and dwell time difference) on participants' trial-by-trial action plans (required action and Q-value differences) using mixed-effects logistic (first fixation) and linear (dwell time difference) regression. Results are reported in the main text as well as in S01. First fixations were more likely on rewards when a Go action was required/ Q-values favored Go over NoGo. Similarly, participants looked overall longer at the reward (compared to the punishment) stake when a Go action was required/ Q-values favored Go over NoGo. Taken together, all these results suggest that attention to rewards/ punishments was synchronized to participants' action plans.

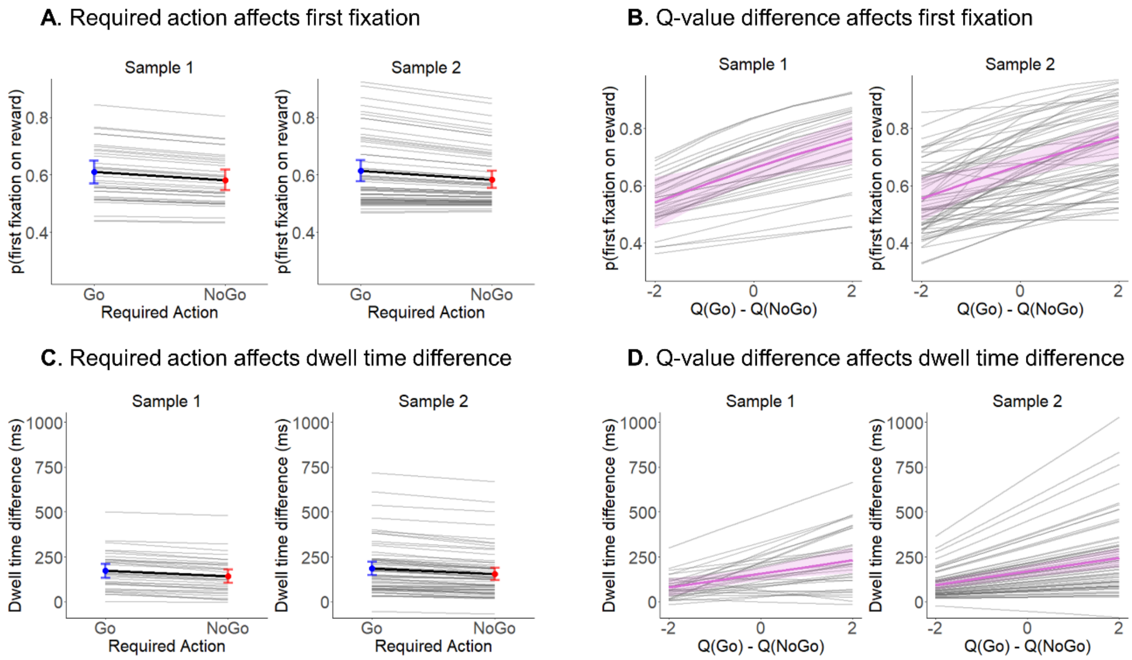

Figure SI04. Effect of action plans on attention measures. Action requirements, i.e., whether participants should make a Go or a NoGo response based on the cue they see, biases participants' attention during the stakes phase: A Go compared to a NoGo requirements led to a higher proportion of first fixations on the reward stake (**A**) and longer dwell time on rewards (compared to punishments) (**C**). The same finding was obtained when fitting a Rescorla-Wagner model to participants' responses and using the Q-values based on responses from past trials to predict what participants should do on the current trial (**B** and **D**).

Supplemental Material 5:

Effect of dwell times on responses and reaction times

Given that dwell time differences (reward minus punishment dwell times) affected both Go/ NoGo responses and reaction times, we additionally tested for separate effects of reward and punishment dwell times (instead of the difference in dwell times) on responses and reaction times using in mixed-effects logistic (for Go/ NoGo responses) and linear (for reaction times) regressions. Dwell time on rewards predicted a higher proportion of Go responses significantly only in Sample 2 (Sample 1: $b = 0.04$, 95% CI [-0.03 0.10], $\chi^2(1) = 0.95$, $p = .331$; Sample 2: $b = 0.07$, 95% CI [0.01 0.13], $\chi^2(1) = 4.62$, $p = .032$; Fig. S04A). Dwell time on punishments significantly predicted a lower proportion of Go responses in both samples (Sample 1: $b = -0.19$, 95% CI [-0.26 -0.11], $\chi^2(1) = 18.04$, $p < .001$; Sample 2: $b = -0.28$, 95% CI [-0.36 -0.20], $\chi^2(1) = 35.08$, $p < .001$; Fig. S04B). Reward dwell time did not significantly predict RTs in either sample (Sample 1: $b = -0.03$, 95% CI [-0.09 0.02], $\chi^2(1) = 1.62$, $p = .203$; Sample 2: $b = -0.01$, 95% CI [-0.04 0.02], $\chi^2(1) = 0.76$, $p = .384$; Fig. S04C), but punishment dwell time predicted slower RTs in Sample 2 (Sample 1: $b = 0.03$, 95% CI [-0.03 0.08], $\chi^2(1) = 0.94$, $p = .333$; Sample 2: $b = 0.04$, 95% CI [0.01 0.07], $\chi^2(1) = 7.67$, $p = .006$; Fig. S04D). Note that RTs are only available for Go responses; hence, the amount of data (and resulting statistical power) are lower compared to Go/ NoGo response data.

Interestingly, stake magnitudes and dwell times exerted highly similar effects on both responses and reaction times, with higher reward stake magnitude as well as more attention to them increased Go responding and speeded responses, while higher punishment stake magnitude as well as more attention to them decreased Go responding and slowed responses. Given that stake magnitudes and dwell times exerted such highly similar effects, one might expect them to operate through the same underlying mechanism. One consequence following from such a shared architecture is that the effects might influence each other, predicting an interaction effect. We hence performed exploratory analyses testing for such an interaction effect, reflecting whether effects of longer vs. shorter attention to the reward (punishment) stake were amplified when participants saw many vs. few potential rewards (punishments) or vice versa. The interaction between the stake difference and the dwell time difference on responses was not significant in either study (Sample 1: $b = -0.008$, 95% CI [-0.06 0.04], $\chi^2(1) =$

0.10, p = .755; Sample 2: b = -0.04, 95% CI [-0.08 0.002], $\chi^2(1) = 3.42$, p = .064), and neither was the case for RTs (Sample 1: b = 0.04, 95% CI [-0.01 0.08], $\chi^2(1) = 2.25$, p = .133; Sample 2: b = -0.003, 95% CI [-0.02 0.02], $\chi^2(1) = 0.03$, p = .856), providing no evidence for attention amplifying effects of stake magnitudes or vice versa.

In conclusion, longer dwell time on rewards led to more and faster responding while longer dwell time on punishments led to less and slower responding. However, effects on reaction times were only significant in the punishment domain. We did not find evidence for an interaction between stake magnitudes and dwell times, yielding no conclusive evidence whether both effects rely on the same underlying mechanism or not.

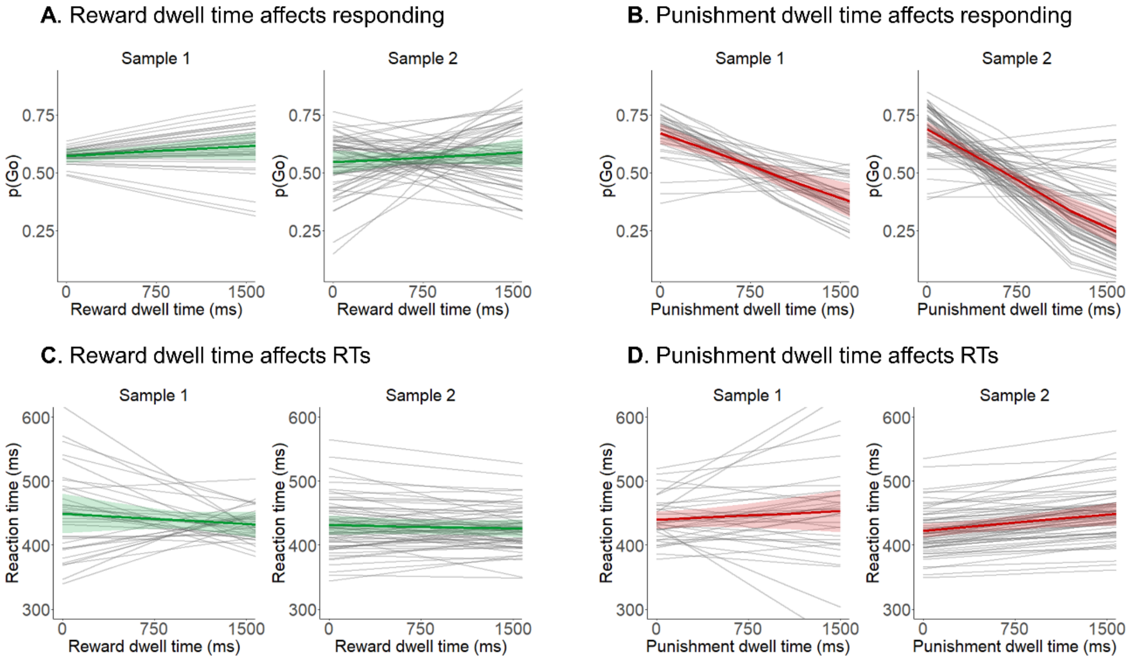

Figure SI05. Effect of dwell times on responses and reaction times. Higher absolute dwell time on rewards led to a higher proportion of Go responses (**A**; only significant in Study 2), while higher absolute dwell time on punishments led to a lower proportion of Go responses (**B**; significant in both studies). Similarly, higher dwell time on rewards tended to speed up reaction times (**C**; though not significant in either study), and higher dwell time on punishment tended to slow down reaction times (**D**; only significant in Study 2).

Supplemental Material 6:

Supplementary online study manipulating attention to reward and punishment stakes

In the results from our eye-tracking studies reported in the main text, we observed an effect of (manipulated) action requirements on eye-gaze (first fixation and dwell time) and an effect of (measured) eye-gaze on the ultimate response. Given that both action requirements and eye-gaze predicted the ultimate response, on might wonder whether the link between eye-gaze and the ultimate response was spurious, induced by action plans as a "common cause" (an instance of the "third variable problem"). Note that all analyses regressing responses onto dwell time reported in the main text controlled for the action plans. In addition, we tested for a causal effect of attention to reward/ punishment information on responses in a separate online study in which we manipulated attention. This study was performed as a thesis project for Bachelor students at the beginning of the COVID-19 pandemic.

Participants and Exclusion Criteria

We collected data from 34 participants (M_{age} = 22.4, SD_{age} = 2.1, range 19–27; 18 female, 31 right-handed). Data collection and analyses were pre-registered (https://osf.io/kzdhm). Data was collected under a stopping rule of N = 55 as maximal sample size or May 10, 2020 as final data collection date (set by financial/ time constraints). As pre-registered, we conducted all analyses in two ways, once including all participants and once excluding participants who a) guessed the research hypotheses (zero participants) or b) did not significantly perform above chance (based on a per-participant logistic regression with response as dependent and required action as independent variable, with $p < .05$ as cut-off; three participants). Both ways led to identical conclusions.

We recruited participants via the SONA Radboud Research Participation System of Radboud University. Participants were required to be at least 18 years old, understand English at a sufficient level (self-reported), not be color-blind, perform the experiment on a PC with a keyboard (no phones or tablets) and complete the study within a maximum of 90 minutes (i.e., 1.5 times the expected completion time). The experiment was administered via the Gorilla platform (Anwyl-Irvine, Dalmaijer, Hodges, & Evershed, 2020). After providing informed consent and demographic information on age, gender, and handedness, participants completed the “reversed-dot-probe” version of the Motivational Go/ NoGo Task for 30-40 minutes (see below). Afterwards, they filled out the brief (13-item) version of the Self-Control Scale (SCS) (Tangney, Baumeister, & Boone, 2004) and the Behavioral Activation/ Behavioral Inhibition System Scales (BIS/BAS) (Carver & White, 1994). Additionally, participants completed two vignettes (measuring omission bias) in which they rated the experienced regret and responsibility of two football coaches who won/ lost a match, afterwards changed/ kept their match plan, and then lost the next game (adapted from (Zeelenberg, van den Bos, van Dijk, & Pieters, 2002). Finally, participants performed a debriefing questionnaire asking them to a) guess the hypotheses of the experiment, b) report any (non-instructed) strategies they used, and c) guess whether the additional instructions helped them perform the task better. Participants were then debriefed about the purposes of the study. In compensation for participation, participants received 1 hour of course credit. Furthermore, participants with at least 60% accuracy in the Go/ NoGo task received tickets (proportional to performance) for a lottery featuring two 20€ gift card vouchers. Research was approved by the local ethics committee of the Faculty of Social Sciences at Radboud University (proposal no. ECSW-2018-171).

Task

Participants performed an adapted version of the Motivational Go/ NoGo learning task termed “reverse-dot-probe version” (Fig. S06A). On each trial, they first saw how many points they could win for a correct response (printed in green font with a “+”) or lose for an incorrect response (printed in red font with a “-”, termed “stakes”). Stakes varied between 10 and 90 points drawn from a uniform distribution. Reward and punishment stake were presented on the left/ right side of the screen, with positions counterbalanced across blocks. Participants were instructed to attend to the stakes because these were relevant for a catch task implemented on some of the trials (see below). After 500 ms, in addition to the stakes, one out of four action cues (letter from the Agathodaimon alphabet) appeared on the screen, which required either a Go response (space bar press) or a NoGo response (no button press). Participants had to learn the correct response from trial-and-error and respond within 1,500 ms. The action cue was presented in close proximity to either the reward stake or the punishment stake, nudging participants to direct more attention to one of the two stakes. Cue position was counterbalanced across trials and orthogonal to action requirements. After a brief fixation cross screen (700 ms), participants received the outcome (either the reward or the punishment stake previously shown) displayed for 1,500 ms. Feedback was probabilistic in that 86% (12 out of 14) trials were "valid" with a correct response winning points and an incorrect response losing points, while the remaining 14% of trials were "invalid" with a correct response losing points and an incorrect response winning points. Trials ended with a variable inter-trial interval (uniform distribution from 1,100 ms till 1,900 ms in steps of 100 ms).

On 12 trials within the first two blocks, after the outcome phase, a catch task occurred. Reward and punishment stake magnitudes were presented together with a "decoy" number (all numbers printed in white font on black boxes without +/- signs, random assignment of numbers to positions). Participants had to indicate the "other" outcome they could have received (i.e., points-to-be-won in case they lost points, points-to-be-lost in case they won points) by clicking on it with the mouse within 20 seconds. The catch task required participants to memorize the exact stake magnitudes seen earlier in the trial, incentivizing attention to them. For the latter two blocks, we did not include any catch trials to not interfere with participants applying the additional instructions (see below).

After the second block, participants received additional instructions that explicitly encouraged them to look at the reward stake in case they planned to perform a Go response, and look at the punishment stake in case they planned to perform a NoGo response. In this way, we aimed to test whether participants could voluntarily align their attention with their action plans and in this way reduce the effect of the action cue's position on responses.

Participants completed 224 trials split into four blocks à 56 trials, each blocks featuring four novel cues with 14 repetitions. Trial features (action cue identity, action requirement, stake magnitudes and positions, ITI) were controlled by one of ten pseudo-randomly drawn "spreadsheets" (preventing cue to repeated on more than two consecutive trials) randomly allocated to participants.

Data Preprocessing

In line with the pre-registration, we excluded reaction times shorter than 300 ms from all analyses (as those are too fast to be induced by the presented cue). Using 200 ms as alternative cut-off (as used in our eye-tracking samples) did not change the conclusions.

Analyses

We analyzed participants' responses (Go/ NoGo) using mixed-effects logistic regression models and their reaction times using mixed-effects linear regression as implemented in the lme4 package in R (Bates, Mächler, Bolker, & Walker, 2015). For all categorical independent variables, sum-to-zero coding was used. Continuous dependent and independent variables were standardized such that regression weights can be interpreted as standardized regression coefficients. We included all possible random intercepts, slopes, and correlations to achieve a maximal random effects structure (Barr, Levy, Scheepers, & Tily, 2013). *P*-values were computed using likelihood ratio tests with the package afex (Singmann, Bolker, Westfall, & Aust, 2018). We considered *p*-values smaller than $\alpha = 0.05$ as statistically significant.

As pre-registered (https://osf.io/kzdhm), firstly, we tested whether the action cue position (i.e., the cue being closer to the reward stake or to the punishment stake) as a proxy for participants' induced attention affect their Go/ NoGo responses and reaction times, expecting a main effect of cue position. Secondly, we tested whether instructing people to attend to stake that matched their action plan reduced the effect of cue position, expecting an interaction between cue position and instructions. We tested both hypotheses in a single model (a logistic regression model for responses, a linear regression model for reaction times) featuring the regressors required response (Go/ NoGo), cue position (on the reward/ punishment side), and instructions (before /after) as well as all possible interactions. As mentioned in the pre-registration, we also report the interaction between required action and instructions as well as the three-way interaction between required action, cue position, and instructions.

Furthermore, we specified two exploratory analyses in our pre-registration. Firstly, we tested whether the difference in stakes (reward minus punishment stake) affected participants' responses and reaction times, expecting more positive differences to lead to more and faster Go responses. For this purpose, we fitted a model with stake difference as sole regressor. Secondly, we calculated participants' mean score on the self-control scale (SCS), BIS and BAS scales and regret judgements and tested whether these scores modulated participants' cue position effect. For this purpose, we fitted a new model for each score featuring cue position, the respective score, and their interaction.

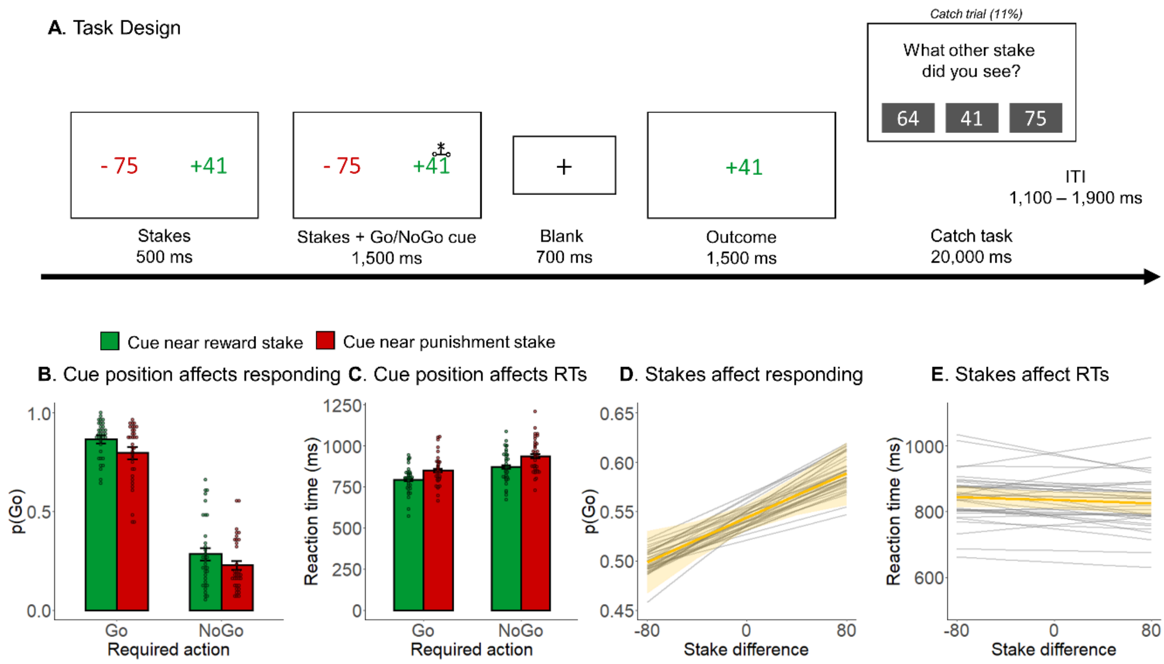

Figure SI06. Task design and results from the online study manipulation attention to reward and punishment information **A.** Task design. On each trial, participants saw many points they could win for correct responses or lose for incorrect responses ("stakes"). After 500 ms, a Go/ NoGo action cue was displayed either next to the reward or the punishment stake, nudging participants to direct more attention to the respective stake. Participants learned whether a cue required a Go or NoGo response from trial-and-error. Outcomes are delivered in a probabilistic manner (86% feedback validity). On catch trials, participants indicated which other stake (i.e., the one they did not receive as an outcome) they had seen before. **B.** Proportion of Go responses as a function of action requirement and cue position. Participants performed significantly more Go responses to Go cues than NoGo cues and when cues were presented next to the reward stake compared to the punishment stake. **C.** Reaction times as a function of action requirement and cue position. Participants showed significantly faster responses to Go cues than NoGo cues and when cues were presented next to the reward stake compared to the punishment stake. **D.** Proportion of Go responses as a function of stake difference (reward minus punishment stake). As net stakes became more positive, participants performed significantly more Go responses. **E.** Reaction times as a function of stake difference (reward minus punishment stake). As net stakes became more positive, participants became faster, but this effect was not significant.

Results

Overall, participants learned the Go/ NoGo task (% correct: M = 79.0, SD = 12.0, range 52.7–94.2), performing significantly more Go responses to Go cues than NoGo cues (main effect of required action: b = 1.60, 95% CI [1.33 1.88], $\chi^2(1)$ = 54.53, p < .001). Three participants did not perform significantly above chance (per-participant logistic regression with response as dependent and required response as independent variable, which is significant for accuracy levels of at least 56%). In line with our pre-registration, we report results with and without these participants. Performance in the catch task was above chance (3 response options imply a chance level of 33.3%; a one-sided binomial test based on 12 trials is significant for 63% accuracy and higher) in only 25 out of 34 participants. Also, the group-level performance was hardly above chance (% correct: $M = 66.4$, $SD = 18.6$, range 25.0–81.7), likely reflecting that this task was very demanding.

Firstly, in line with our pre-registration, we tested whether the cue position (action cue on the reward/ punishment side) affected participants' Go/ NoGo responses. Participants performed more Go responses when the action cue was on the side of the reward stake compared to the side of the punishment stake (main effect of cue position: $b = 0.19$, 95% CI [0.09 0.29], $\chi^2(1) = 10.90$, $p < .001$; Fig. S06B), suggesting that increased attention to rewards (compared to punishments) induced more Go responses. Similarly, participants performed faster Go responses when the action cue was on the side of the reward stake compared to the side of the punishment stake (main effect of cue position: $b = -0.03$, 95% CI [-0.04 -0.02], $\chi^2(1) = 25.70$, $p < .001$; Fig. S06C). These results suggested that more attention directed to reward/ punishment stake causally affects participants' responses and reaction times in the fashion of Pavlovian biases.

Secondly, in line with our pre-registration, we tested whether the effect of cue position became smaller after participants were instructed to attend to the stake that matched their action plan. The interaction effect between cue position and instructions was not significant ($b = -0.03$, 95% CI [-0.10 0.04], $\chi^2(1) = 0.55$, $p = .458$), providing no evidence for responses becoming less affected by the cue position once participants tried to voluntarily deploy their attention. In fact, the sign of the effect suggested the effect of cue position to become stronger (instead of weaker) after additional instructions were administered. However, there was a significant interaction between required action and instructions ($b = -0.38$, 95% CI [-0.50 -0.25], $\chi^2(1) = 29.28$, $p < .001$), suggesting that participant overall performed better after receiving instructions. In absence of a control group, this effect cannot be disentangled from an increase in performance over time, providing inconclusive evidence for whether instructions affected participants' responses or not. The three-way interaction effect between required action, cue position, and instruction was not significant (b = 0.01, 95% CI [-0.06 0.07], $\chi^2(1)$ = 1.78, p = .182). Apart from responses, also the effect of cue position on reaction times was not significantly changed by instructions (b = 0.01, 95% CI [-0.003 0.02], $\chi^2(1)$ = 1.65, p = .199), and neither was the interaction between required action and instructions (b = 0.001, 95% CI [-0.01 0.01], $\chi^2(1)$ = 0.04, p = .840) nor the three-way interaction effect between required action, cue position, and instruction (b = -0.0003, 95% CI [-0.01 0.01], $\chi^2(1)$ = 0.004, p = .948) significant.

Thirdly, as part of the exploratory analyses mentioned in the pre-registration, we tested whether the difference in stakes (reward minus punishment stake) affected participants' responses or reaction times. As expected, as the difference in stakes increased (relatively more points to win than to lose), participants performed significantly more Go responses (b = 0.08, 95% CI [0.03 0.12], $\chi^2(1)$ = 8.15, p = .004; Fig. S06D), suggesting that the difference in available rewards/ punishments biased their responses in the fashion of Pavlovian biases. Reaction times were not significantly affected by the stake difference (b = -0.004, 95% CI [-0.01 0.004], $\chi^2(1)$ = 1.01, p = .316; Fig. S06E).

Fourthly, as part of the exploratory analyses mentioned in the pre-registration, we tested whether the effect of cue position on responses was predicted by participants' score on the self-control scale (SCS), the BIS and BAS scales, or the regret and responsibility ratings in the omission bias vignettes. We did not find any significant modulation of the cue position effect by SCS scores (b = -0.03, 95% CI [-0.09 0.06], $\chi^2(1)$ = 0.70, p = .403), BAS Drive scores (b = -0.04, 95% CI [-0.11 0.03], $\chi^2(1)$ = 1.03, p = .310), BAS Reward Responsiveness scores (b = -0.01, 95% CI [-0.08 0.05], $\chi^2(1)$ = 0.10, p = .756), rated regret for changing the match plan after a previous football win (b = -0.02, 95% CI [-0.10 0.07], $\chi^2(1)$ = 0.14, p = .710), rated responsibility asymmetry when changing/ keeping the match plan after a previous football win (b = 0.02, 95% CI [-0.04 0.08], $\chi^2(1)$ = 0.39, p = .532), rated regret for changing the match plan after a previous football defeat (b = -0.01, 95% CI [-0.07 0.05], $\chi^2(1)$ = 0.10, p = .750), or rated responsibility asymmetry when changing/ keeping the match plan after a previous football defeat (b = -0.004, 95% CI [-0.09 0.08], $\chi^2(1)$ = 0.01, p = .933). However, the cue position effect was significantly modulated by BIS scores (b = -0.07, 95% CI [-0.13 -0.01], $\chi^2(1)$ = 4.32, p = .038) with participants with higher BIS scores showing weaker cue position effects, and by BAS Fun Seeking scores (b = -0.07, 95% CI [-0.14 -0.01], $\chi^2(1)$ = 4.64, p = .031) with participants with higher BAS scores showing again weaker cue position effects. Given the sample only comprised 34 participants and several between-participants analyses were run, these results should be interpreted with caution.

We repeated all analyses while excluding three participants who did not perform significantly above chance in the Go/ NoGo task. Firstly, still, participants performed more (b = 0.18, 95% CI [0.08 0.29], $\chi^2(1)$ = 10.13, p = .001) and faster (b = -0.03, 95% CI [-0.05 -0.02], $\chi^2(1)$ = 26.84, p < .001) Go responses when the action cue was on the side of the reward stake compared to side of the punishment stake. Secondly, the effect of cue position on responses was again not significantly different after compared to before additional instructions were administered (b = -0.02, 95% CI [-0.10 0.06], $\chi^2(1)$ = 0.24, p = .623), but the effect of required action was again stronger after compared to before responses (b = -0.41, 95% CI [-0.55 -0.27], $\chi^2(1)$ = 23.39, p < .001), with again no significant three-way interaction (b = 0.01, 95% CI [-0.07 0.09], $\chi^2(1)$ = 0.06, p = .800). Regarding reaction times, again, neither the effect of cue position (b = 0.01, 95% CI [-0.003 0.02], $\chi^2(1)$ = 1.98, p = .159) nor the effect of required action (b = 0.003, 95% CI [-0.01 0.02], $\chi^2(1)$ = 0.28, p = .597) was significantly modulated by instructions, and neither was the three-way interaction significant (b = -0.001, 95% CI [-0.01 0.01], $\chi^2(1)$ = 0.08, p = .779). Thirdly, as the stake difference increased, participants again performed significantly more Go responses (b = 0.06, 95% CI [0.01 0.11], $\chi^2(1)$ = 5.72, p = .017), but not significantly faster responses (b = -0.006, 95% CI [-0.01 0.002], $\chi^2(1)$ = 2.33, p = .127). Fourthly, we again did not find any significant modulation of the cue position effect by SCS scores (b = -0.04, 95% CI [-0.11 0.03], $\chi^2(1)$ = 1.25, p = .264), BAS Drive scores (b = -0.02, 95% CI [-0.09 0.05], $\chi^2(1)$ = 0.30, p = .582), BAS Reward Responsiveness scores (b = -0.01, 95% CI [-0.08 0.05], $\chi^2(1)$ = 0.10, p = .751), rated regret for changing the match plan after a previous football win (b = 0.02, 95% CI [-0.07 0.09], $\chi^2(1)$ = 0.27, p = .603), rated responsibility asymmetry when changing/ keeping the match plan after a previous football win (b = 0.02, 95% CI [-0.04 0.09], $\chi^2(1)$ = 0.23, p = .632), rated regret for changing the match plan after a previous football defeat (b = -0.004, 95% CI [-0.07 0.06], $\chi^2(1)$ = 0.01, p = .909), or rated responsibility asymmetry when changing/ keeping the match plan after a previous football defeat (b = 0.007, 95% CI [-0.08 0.10], $\chi^2(1)$ = 0.02, p = .877). The modulation by BIS scores was not significant any more (b = -0.06, 95% CI [-0.13 0.004], $\chi^2(1)$ = 2.33, p = .127), while the modulation by

BAS Fun Seeking scores was still significant (*b* = -0.06, 95% CI [-0.13 -0.003], $\chi^2(1) = 4.19$, *p* = .041). Overall, analyses excluding the three participants who did not perform the Go/ NoGo task significantly above chance led to identical conclusions as analyses including all participants.

Discussion

In this study, we manipulated attention by displaying Go/ NoGo action cues next to either the reward or punishment stake, nudging participants to pay relatively more attention to the stake that we next to the action cue. We obtained causal evidence that attention to reward information (compared to punishment information) leads to more Go (compared to NoGo) responses as well as to faster responses. We did not find evidence for instructions to voluntarily deploy attention in line action plans reducing the attentional effect. Potentially, the task was too demanding and the trial time course too fast for participants to voluntarily steer attention in a way that supported their action plans. Future studies might use different instructions or an altered task design that gives participants more time to deploy attention before they perform an action.

Furthermore, we found evidence for overall stake differences (reward minus punishment stake) biasing responses (but not reaction times) in the fashion of Pavlovian biases. These results support the effect of stake differences on responses reported in the main text. Finally, we did not find any strong modulation of the attentional effect by self-reported measures such as the Self-Control Scale, the BIS/ BAS scales, or regret and responsibility ratings in two vignettes measuring omission biases. Although there was some evidence for stronger BIS and BAS Fun Seeking scores predicting weaker attention effects, these results should be treated with caution given the limited sample size and the higher number of tests. Future studies should test for such links in larger samples. In sum, the core conclusion is that the results of this study support a causal effect of attention on Go/ NoGo responses.

Supplementary Material 7:

Effects of stake magnitudes and dwell times on responses predict interindividual differences in task performance

Both stakes and dwell times affected Go/ NoGo responses (and reaction times) in a similar way, i.e., a higher reward stake as well as more attention to it increased Go responding and speeded responses, while a higher punishment stake as well as more attention to it decreased Go responding and slowed responses. Given such highly similar effects, one might expect them to operate through the same underlying mechanism. First, one consequence following from such a shared architecture is that effects should influence each other, i.e., the presence of a higher stake could alter the impact of dwell times on responses, or vice versa, which predicts an interaction effect. However, we observed no evidence for such an interaction effect (see S05), tentatively suggesting that effects operate independently of each other (though curiously with highly similar consequences).

An alternative way of assessing how comparable these effects are is to probe their consequences for task performance across participants: Does letting responses be strongly guided by stake differences (reward minus punishment stake magnitudes) vs. strongly guided by dwell time differences (reward minus punishment dwell times) have similar or different consequences for overall performance in the Go/ NoGo task? For this purpose, we re-fitted regression models across both samples, extracted per-participant regression coefficients (fixed-effect plus participant-specific random effect), and correlated these coefficients with participant overall performance (% correct responses).

Performance was significantly lower in those participants in which stake difference more strongly shaped their responses (Figure S08A, B). This finding was in stark contrast to significantly higher performance in those participants in which dwell time differences (reward minus punishment dwell time) more strongly affected response. It is noteworthy that the stake differences are experimentally controlled, and thus purely "bottom-up", while in contrast, dwell time differences were under participants' control and synchronized to action plans, both directly (effect on dwell time difference) and indirectly (effect on first fixations).

We performed control analyses to exclude the possibility that the association between attentional effects on responses and task performance was driven by better performing participants showing higher eye-tracking data quality. First, we computed the number of trials with any (opposed to no) fixation on any of the two stakes. This number was significantly positively correlated with performance, $r(97) = 0.23$, 95% CI [0.03, 0.41], $p = .025$, but not with the attentional effect on responses, $r(97) = 0.13$, 95% CI [-0.07, 0.32], $p = .208$. When using both task performance and number of trials with any fixation to predict attention effects in a multiple linear regression, the effect of task performance was still strongly significant, $t(96) = 4.79$, $p < .001$. Second, we calculated the total time (in ms) that people attended to any of the two stakes objects. This number was neither significantly correlated with performance, $r(97) = 0.09$, 95% CI [-0.11, 0.28], $p = .389$, nor with the attentional effect on responses, $r(97) = 0.13$, 95% CI [-0.07, 0.32], $p = .183$, and when using both task performance and total fixation time to predict attention effects in a multiple linear regression, the effect of task performance was still strongly significant, $t(96) = 4.90$, $p < .001$. In sum, it is unlikely that the correlation between performance and attentional effects on responses is driven by more accurate participants providing higher-quality eye-tracking data.

Furthermore, we performed control analyses checking whether performance, being associated with how many rewards (rather than punishments) participants received, was associated with differential fixation patterns (more first fixations or longer fixations) to reward vs. punishment stakes. It is possible that performance affects information search: high performing participants can reasonably expect to receive rewards most of the time, so they might be more interested in and attend more to reward stakes. Vice versa, lower performing participants might expect occasional punishments and thus also attend to punishment stakes. There was no significant correlation between task performance and the number of first fixations on rewards vs. punishments, $r(97) = -0.11$, 95% CI [-0.30, 0.09], $p = .298$ and the association between task performance and the attentional effect on responses remained significant when controlling for the number of first fixations, $t(96) = 4.97$, $p < .001$. There was however though a significantly negative correlation between task performance and overall dwell time difference (dwell time on reward stakes minus dwell time on punishment stakes), $r(97) = -0.27$, 95% CI [-0.44, -0.08], $p = .007$: participants with higher performance showed a more variable (i.e., less biased towards reward stakes) gaze pattern and attended relatively more to punishments compared to participants with low performance. The association between task performance and the attentional effect on responses remained significant when controlling for the this overall dwell time difference, $t(96) = 5.20$, $p < .001$. In sum, we found no evidence for high performing participants exclusively focusing on reward stakes and low performing participants also attending to punishment stakes. If anything, we found the opposite pattern of high performing participants showing a more variable gaze pattern (also attending to punishment stakes), which chimes with the idea that these participants could rely their response on their (more adaptive/ flexible) gaze pattern.

Note that all these performance-dependent results are exploratory and should be interpreted with caution.

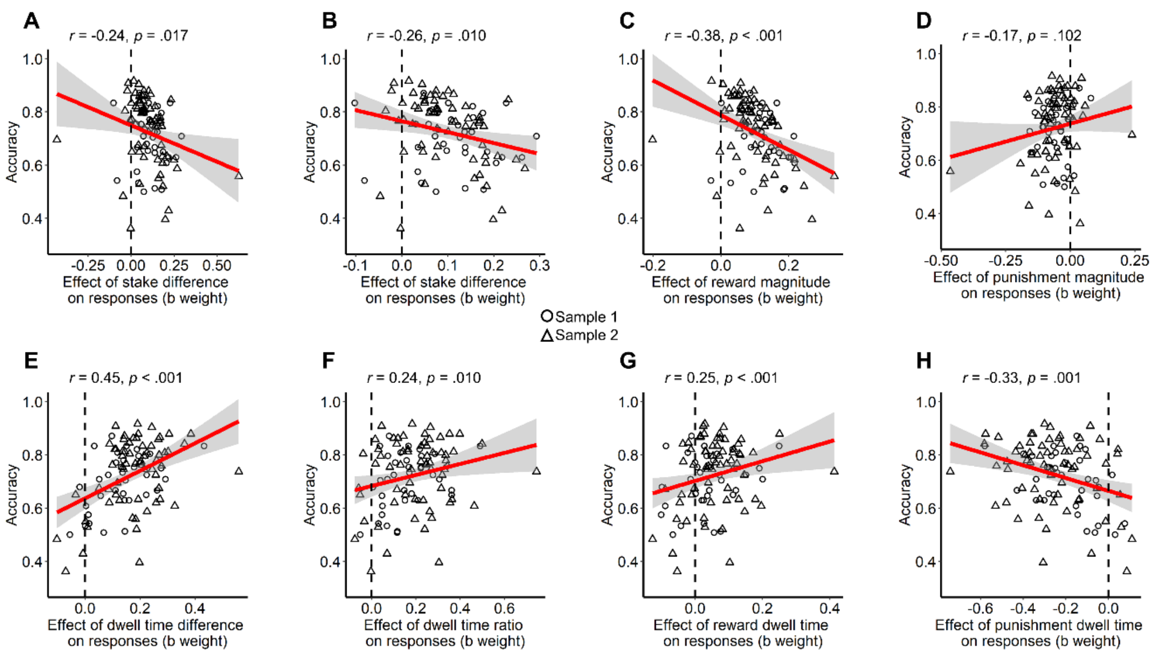

Figure SI07. Association between interindividual variability of accuracy and in the effects of stake magnitudes and dwell times on responses. Participants' mean accuracy correlated significantly negatively with their respective effect of stake differences on responses (**A**), also when two outliers removed (**B**), which was driven both by a negative correlation with the effect of the reward stake (**C**; note that these effects tend to be positive) as well as a positive correlation with the effect of the punishment stake (**D**; note that these effects tend to be negative, i.e., participants with stronger negative effects showed worse performance). These correlations suggest that participants with strong stake difference effects showed poor performance. The opposite pattern occurred for the effect of dwell time on responses: This effect correlated significantly positively with accuracy, both for the difference between reward and punishment dwell times (**E**) as well as the relative dwell time (ratio) on rewards (**F**). Again, this effect was driven by reward dwell times (**G**) rather than punishment dwell times (**H**). These correlations suggest that participant with strong attention effects showed high performance.

Supplemental References

Anwyl-Irvine, A., Dalmaijer, E. S., Hodges, N., & Evershed, J. K. (2020). Realistic precision and accuracy of online experiment platforms, web browsers, and devices. *Behavior Research Methods*. doi: 10.3758/s13428-020-01501-5

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

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

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *Journal of Personality and Social Psychology*, *67*(2), 319–333. doi: 10.1037/0022-3514.67.2.319

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

Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality*, *72*(2), 271–324. doi: 10.1111/j.0022-3506.2004.00263.x

Zeelenberg, M., van den Bos, K., van Dijk, E., & Pieters, R. (2002). The inaction effect in the psychology of regret. *Journal of Personality and Social Psychology*, *82*(3), 314–327. doi: 10.1037/0022-3514.82.3.314