

Goal-Directed Recruitment of Pavlovian Biases Through Selective Visual Attention

Johannes Algermissen and Hanneke E. M. den Ouden
Donders Institute for Brain, Cognition and Behaviour, Radboud University

Prospective outcomes bias behavior in a “Pavlovian” manner: Reward prospect invigorates action, while punishment prospect suppresses it. Theories have posited Pavlovian biases as global action “priors” in unfamiliar or uncontrollable environments. However, this account fails to explain the strength of these biases—causing frequent action slips—even in well-known environments. We propose that Pavlovian control is additionally useful if flexibly recruited by instrumental control. Specifically, instrumental action plans might shape selective attention to reward/punishment information and thus the input to Pavlovian control. In two eye-tracking samples ($N = 35/64$), we observed that Go/NoGo action plans influenced when and for how long participants attended to reward/punishment information, which in turn biased their responses in a Pavlovian manner. Participants with stronger attentional effects showed higher performance. Thus, humans appear to align Pavlovian control with their instrumental action plans, extending its role beyond action defaults to a powerful tool ensuring robust action execution.

Public Significance Statement

This study suggests that Pavlovian biases, a fast-and-frugal decision strategy that may trigger suboptimal choices in certain contexts, are not a permanent, immutable force upon behavior. Instead, they are flexibly recruited depending on the action a person is planning: Given the plan to make/withhold actions, people preferably attend to reward-/punishment-related information, which in turn triggers Pavlovian biases that facilitate the implementation of these plans. Stronger “outsourcing” of action implementation to such an attentional recruitment of Pavlovian biases leads to higher performance. These findings highlight how Pavlovian biases are more flexible than previously thought and how strong biases can be of advantage.

Keywords: decision-making, motivation, Pavlovian biases, attention, eye-tracking

Supplemental materials: <https://doi.org/10.1037/xge0001425.supp>

The valence of potential outcomes biases action selection: The prospect of rewards invigorates action (“Go”), while the prospect of punishment suppresses it (“NoGo”). These so-called motivational, or “Pavlovian,” biases have first been observed in animal studies in which the presence of a reward-associated cue invigorated cue-unrelated behaviors (Estes, 1943, 1948; LoLordo et al., 1974; Lovibond, 1983; Schwartz, 1976). While at first interpreted as

seemingly irrational, recent theorizing has suggested that these biases in fact constitute a decision-making strategy that is particularly “fast-and-frugal” (Boureau et al., 2015; Dayan et al., 2006). Past theorizing has assumed that, while inflexible, these biases are fast, computationally cheap, and likely attuned to global environmental statistics (Dayan et al., 2006). They can thus act as sensible “default” response strategies in situations in which instrumental, goal-directed

This article was published Online First May 18, 2023.

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

Results described in this article have previously been presented in the form of poster presentations at the “Computational Properties of Prefrontal Cortex (CPPC)” conference, March 23–25, 2022, Oxford, United Kingdom, and the “Nederlandse Vereniging voor Psychonomie (NVP)” conference, April 28–30, 2022, Egmond aan Zee, the Netherlands.

We thank Jesse van der Spek and Helena Olraun for assistance with data collection and Peter Dayan for helpful comments and suggestions on the article. This work was supported by a Netherlands Organization for Scientific Research (NWO) VIDI (Grant 452-17-016) to Hanneke E. M. den Ouden. The authors declare that they have no competing interests.

Johannes Algermissen served as lead for conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, software, validation, visualization, writing—original draft, and writing—review and editing. Hanneke E. M. den Ouden served as lead for funding acquisition and supervision and served in a supporting role for conceptualization, methodology, resources, validation, writing—original draft, and writing—review and editing.

Correspondence concerning this article should be addressed to Johannes Algermissen or Hanneke E. M. den Ouden, Donders Institute for Brain, Cognition and Behaviour, Radboud University, Thomas van Aquinostraat 4, 6526 GD Nijmegen, The Netherlands. Email: johannes.algermissen@donders.ru.nl or hanneke.denouden@donders.ru.nl

control fails to deliver rewards beyond chance levels, such as novel or uncontrollable environments (Daw et al., 2005; Dorfman & Gershman, 2019; O'Doherty et al., 2017). These accounts assume that Pavlovian and instrumental control coexist, largely segregated from another, and merely compete at the behavioral output level. In case of conflict, the former has to be actively suppressed—a requirement humans only imperfectly master (Brelan & Brelan, 1961; Cavanagh et al., 2013; Hershberger, 1986; Swart et al., 2018).

Several fields within psychology, including research on decision-making, motor control, and attention, have recognized that, in order to solve a given problem, an agent can use different strategies. To pick a situationally appropriate strategy, it does not only matter whether or how well the strategy solves the problem (e.g., how many rewards it returns), but also what the invested costs are (e.g., how long it takes, how many mental resources it takes; Bettman et al., 1990; Boureau et al., 2015). Hence, seemingly suboptimal or “irrational” behavior can turn out to be rational when seen in the light of costs or resource constraints—a term called “bounded rationality” (Simon, 1957) or more recently “resource rationality” (Griffiths et al., 2015; Lieder & Griffiths, 2020). Strategies that return higher-quality solutions in some situations might become unfeasible in other situations due to resource constraints, calling for simpler, less costly strategies. This viewpoint has the potential to not only explain the choice of seemingly inferior options violating the axioms of rational choice (Palminteri et al., 2015; Tversky, 1969), but also motor errors (Du et al., 2022; Hardwick et al., 2019; McDougale et al., 2016; Wolpert & Landy, 2012) and seemingly nonstrategic, imprecise, or inefficient (“lazy”) visual search (Araujo et al., 2001; Ballard et al., 1995; Draschkow et al., 2021; Horowitz & Wolfe, 1998; Steinman et al., 2003; Wolfe et al., 2000). In all these psychological domains, humans (and other animals) have seemingly multiple independent decision-making systems at their disposal.

One class of particularly simple decision strategies are so-called “heuristics” or “biases,” fast-and-frugal decision strategies which are rather inflexible, but perform well in a restricted set of situations (Gigerenzer & Gaissmaier, 2011; Hutchinson & Gigerenzer, 2005). These heuristics have likely been acquired as adaptations to specific environmental challenges through both biological and cultural evolution (Haselton et al., 2009; Todd & Brighton, 2016). The fact that many heuristics are present even in animals (Fawcett et al., 2014) speaks for their evolutionary ancientness and possibly genetic hardwiring. However, the “meta”-question arises how to determine which heuristic to use in a given situation (Lieder & Griffiths, 2017; Marewski & Link, 2014; Rieskamp & Otto, 2006). Apparently, humans and other animals frequently misapply heuristics (in such cases typically called “biases”; Beck et al., 2012; Fawcett et al., 2014; Rahnev & Denison, 2018), raising the question of why these biases are so seemingly strong and hard to suppress.

A relevant question in many fields of psychology is whether distinct strategies operate in isolation, conflict with each other, or even work in synergy. Specifically, more sophisticated strategies might “outsource” certain subroutines to simpler strategies, yielding a “division of labor.” Such a synergy is frequently assumed to evolve over time, with initial acquisition through more “explicit” rule-driven strategies, which are later outsourced to more “implicit,” incremental, habit-like strategies, a division prominent in goal-directed versus habitual decision-making (Balleine & Dickinson, 1998; Daw et al., 2005), response preparation (Du et al., 2022;

Hardwick et al., 2019), explicit versus implicit motor skill learning (Mazzoni & Krakauer, 2006; McDougale et al., 2016; McDougale & Taylor, 2019), and goal-directed versus history-guided attention (Anderson, 2016; Theeuwes, 2018). Beyond such a sequential labor division in which one system hands over control to another, there are even examples of systems that are active simultaneously, where one system trains the other, for example, in reward revaluation (Gershman et al., 2014; Robinson & Berridge, 2013), credit assignment (Moran et al., 2019), and memory replay (Matar & Daw, 2018). Crucially, such a simultaneous collaboration requires both systems to be permanently active. In this article, we propose that also instrumental and Pavlovian control can work in such a synergy.

In contrast to previous literature that has assumed a parallel, strictly segregated arrangement of instrumental and Pavlovian control, we suggest that the instrumental system can adaptively recruit and steer the Pavlovian system by selecting its input via visual attention. Humans are not just passively exposed to reward and punishment cues that drive these biases. Instead, they can actively seek out or ignore these cues and thereby modulate their influence via selective visual attention (“active sensing”; Friston et al., 2010; Gottlieb & Oudeyer, 2018; Yang et al., 2016). In a world full of distractions, where actions unfold over time and are prone to interference, instrumental control could harness the power of cue-driven, “automatic” behavioral tendencies by directing visual attention to cues that activate them and then automatically trigger the intended action. In this scenario, it might be warranted to keep the Pavlovian system permanently “online,” accepting a few infrequent errors for the benefit of overall more robust action implementation. This view contrasts with earlier assumptions that Pavlovian biases are mere “defaults” to fall back to in novel or uncontrollable environments. Instead, keeping Pavlovian control constantly online during instrumental goal pursuit might be advantageous. However, previous task designs measuring Pavlovian biases do not match such scenarios in which agents actively seek out information that helps them achieve their goals. We developed a new paradigm that temporally separates action selection, attention to reward and punishment information, and action execution. We then tested whether humans seek out reward and punishment information—and allow Pavlovian biases to shape responding—in a way that is aligned with their action goals. Note that, in the following, we will use the term “goal-directed” to denote such a synchronization between action goals and information search—remaining tacit about whether the underlying cognitive process involves prospective planning or devaluation sensitivity, features typically taken as indicators of “goal-directedness” of behavior (Balleine & Dickinson, 1998).

Research in the past decade supports the notion that overt attention (eye gaze) toward positive aspects of choice options predicts their eventual selection (Cavanagh et al., 2014; Fiedler & Glöckner, 2012; Krajbich et al., 2010), while attention to negative aspects predicts their rejection (Armell et al., 2008; Pachur et al., 2018; Westbrook et al., 2020). In these studies, positive and negative information is required for making the correct choice. Theoretical perspectives have speculated that longer attention to an option facilitates memory retrieval of its features, which could accentuate its value (Shadlen & Shohamy, 2016; Weilbacher et al., 2021). However, attention to task-irrelevant positive or negative cues—which have no apparent relationship to the choice options and thus

cannot serve as anchors for memory retrieval—might have similar effects. Indeed, in Pavlovian-to-instrumental transfer (PIT) paradigms, incidental background cues associated with positive/negative outcomes induce Go/NoGo actions (Estes, 1943, 1948; Geurts et al., 2013a, 2013b; Huys et al., 2011; Rescorla & Solomon, 1967). Linking those PIT effects to the role of attention in value-based choice implies that directing attention to (task-irrelevant) reward or punishment cues should activate the Pavlovian system and, in this way, automatically invigorate or suppress choice.

Beyond the effects of attention on action, there is also evidence that action plans themselves can direct attention (Heuer et al., 2020; Olivers & Roelfsema, 2020; van Ede, 2020). Task goals modulate which stimulus features we are sensitive to and distracted by (Eimer & Kiss, 2008; Folk et al., 1992; van der Stigchel & Hollingworth, 2018). “Active sensing” perspectives frame attention as a tool to actively interrogate the environment while implementing action plans (Cisek & Pastor-Bernier, 2014; Gottlieb & Oudeyer, 2018; Yang et al., 2016). The premotor theory of attention goes as far as proposing that the primary purpose of attention is to monitor target features relevant for preparing an action toward the target (Rizzolatti et al., 1987; Shelliga et al., 1997). Studies have indeed found perceptual sensitivity to be selectively sharpened for features relevant to an ongoing action, for example, object location for reaching movements or object size and orientation for grasping movements (Bekkering & Neggers, 2002; Craighero et al., 1999; Fagioli et al., 2007). However, in the domain of value-based decision-making, similar evidence for task goals shaping attention is scarce. One relevant finding might be that humans tend to seek out a choice option one final time before selecting it (“last fixation” or “late onset” bias) even if they already know this option to be superior to other options (Hunt et al., 2016; Kaanders et al., 2021). In this case, attention appears to be guided by choice rather than vice versa, extending the premotor theory of attention to value-based decision-making.

Taken together, there appear to be mechanisms synchronizing agents’ attention with their action plans, and there is tentative evidence for attention to reward and punishment information triggering automatic responses in the fashion of Pavlovian biases. Hence, it seems indeed possible that an instrumental system could “recruit” the Pavlovian system to “aid” the execution of action plans by strategically steering attention toward relevant information. We tested this idea in two samples (the second one was a direct, preregistered replication) using eye-tracking. For this purpose, we designed a novel Go/NoGo learning task in which action planning and execution were separated by a phase in which participants could preview the positive or negative outcomes at stake. Notably, information about these outcomes was not informative for the selection of the correct action. We predicted that action plans would shape attention to reward and punishment stakes, that is, that participants’ first fixation (not confounded by bottom-up saliency effects due to a gaze-contingent design) would be more often on the reward information when participants planned a Go (compared to a NoGo) action. Vice versa, we predicted an effect of attention duration to rewards versus punishments on the final response, that is, that longer attention to reward compared to punishment information would lead to more Go responses and speed up reaction times (Figure 1A and 1B). Such a goal-directed recruitment of Pavlovian biases would extend their role beyond mere “default” strategies in novel environments toward a powerful aiding robust action execution.

Method

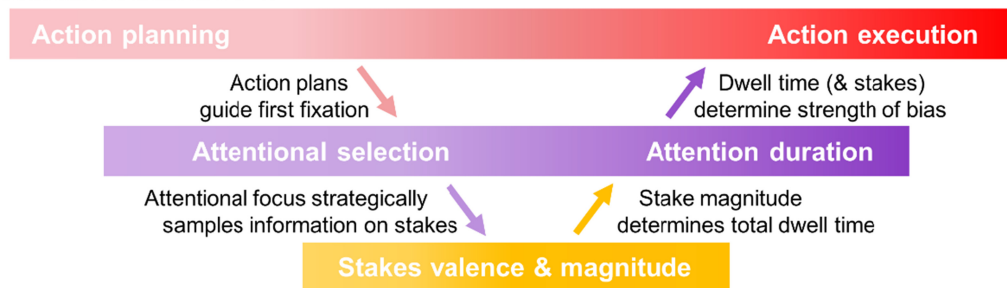
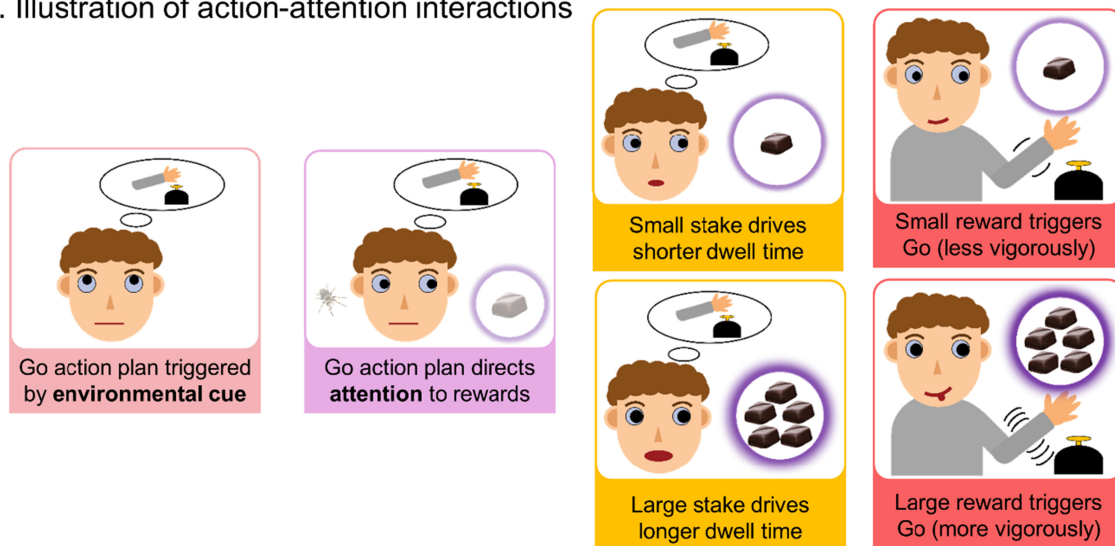
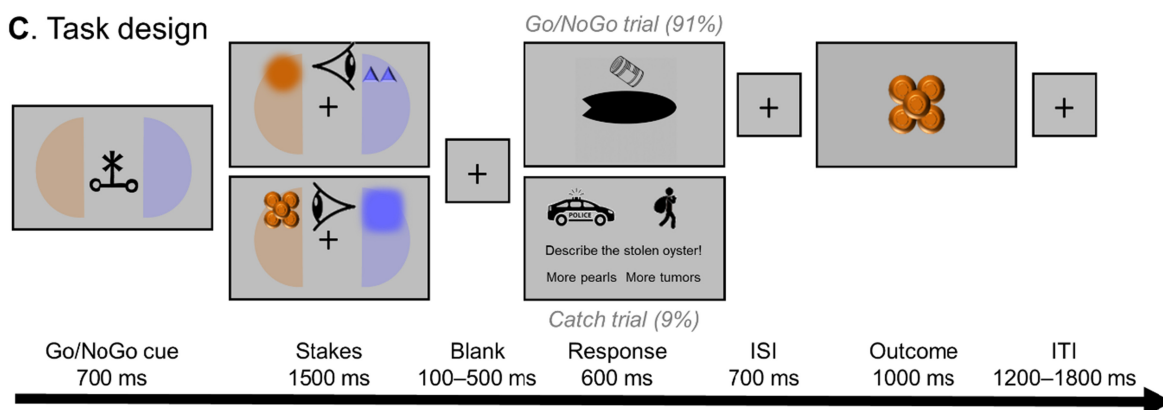
Participants and Exclusion Criteria

In Sample 1, we recorded eye-tracking data from 35 participants ($M_{\text{age}} = 23.7$, $SD_{\text{age}} = 4.1$, range = 18–35, one outlier at age 58; 27 women, eight men; 30 right-handed; 21 with the right eye as the dominant eye). In Sample 2 (replication sample), we recorded data from 64 participants ($M_{\text{age}} = 21.5$, $SD_{\text{age}} = 3.0$, range = 18–34; 50 women, 13 men, one other; 62 right-handed; 41 with the right eye as the dominant eye). In this replication sample, the study design, hypotheses, and analysis plan were preregistered (<https://osf.io/nsy5x>). The sample size for this sample was based on the effect size of the primary effect of interest in Sample 1, that is, action requirements affecting first fixations ($z = 2.89$, Cohen’s $d = 0.49$), which yielded a required sample of $N = 57$ to detect such an effect with 95% power (two-sided one-sample t test; Murayama et al., 2022). We initially collected data from 57 participants, but given that seven participants did not perform significantly above the chance level, we collected additional seven participants. Performance above 56% in 240 trials was significantly above chance (one-sided binomial test). Note that, in line with our preregistration, all results in the main text are based on all participants (see online Supplemental Material 1 for an overview of all results); results for only those participants that performed significantly above chance are reported in online Supplemental Material 2 and led to identical conclusions.

Participants were recruited via the SONA Radboud Research Participation System of Radboud University. Exclusion criteria comprised glasses, color blindness, and prior treatment for neurological or psychiatric disorders. The study protocol was identical for both samples. Participants took part in a 1 hr session that comprised informed consent, eye-tracker calibration, a 10-min practice phase including written instructions and practice trials, and finally the 30-min eye-tracking experiment. Upon completion of the task, participants filled in a structured debriefing about their presumed hypothesis of the experiment, and any strategies they applied. None of the participants guessed the study hypotheses. Participants received a participation fee of €10 or 1 hr of course credit plus a performance-dependent bonus of €0–2 (Sample 1: $M = €0.77$, $SD = €0.43$, range = €0.09–1.58; Sample 2: $M = €0.91$, $SD = €0.47$, range = €0.10–1.67). The research was approved by the local ethics committee of the Faculty of Social Sciences at Radboud University (proposal no. ECSW-2018-171).

Apparatus

Reporting follows recently suggested guidelines for eye-tracking studies (Fiedler et al., 2020). The experiment was performed in a dimly lit, sound-attenuated room, with participants’ heads stabilized with a chin rest. The experimental task was coded in PsychoPy 2020.2.7 on Python 3.7.0, presented on a 24" BenQ XL2420Z screen of resolution (1,920 × 1,080 pixels resolution, refresh rate 144 Hz). Manual button presses were applied via a custom-made button box with two buttons (index and middle finger of the dominant hand). Participants’ dominant eye was tracked with an EyeLink 1000 tracker (SR Research, Mississauga, Ontario, Canada; sampling rate of 1,000 Hz; spatial resolution of 0.01° of visual angle, monocular recording), controlled via Pylink for Python 3.7.0. The eye-tracker was placed 20 cm in front of the screen, and participants’ chin rest 90 cm in front of the screen. Before the task, participants performed

Figure 1*Theoretical Framework and Task Design***A. Theoretical framework****B. Illustration of action-attention interactions****C. Task design**

Note. (A) Theoretical framework of the interaction between action and attention. An environmental cue elicits an action plan, which directs top-down attention (first fixation) toward information about potential reward/punishment outcomes (stakes). The first fixation anchors attention and (partly) determines which stakes will receive more attention, which is additionally modulated by bottom-up signals such as the magnitude of the stakes. The relative attention on reward versus punishment stakes (dwell time) biases the final Go/NoGo action in a Pavlovian manner. (B) Cartoon illustration of the proposed interaction of action planning and attention. (C) Task design. Participants learned Go/NoGo responses to various cues (cover story: feed/not feed various oyster types to maximize pearls and minimize toxic tumors). Cue presentation (instructing the correct action) and action execution are separated by a phase in which rewards (pearls, here orange circles) and punishments (toxic tumors, here blue triangles) at stake for correct/incorrect responses are presented in a gaze-contingent manner. Afterward, the oyster (black oval) can be fed, and for Go responses, participants have to press the button on the side where it is “still open.” Outcomes are delivered in a probabilistic manner (75% feedback validity). On catch trials, participants have to indicate whether the oyster featured more pearls or tumors (cover story: The oyster is stolen by thieves and has to be retrieved back from the police, which requires identification). ISI = inter-stimulus interval; ITI = intertrial interval. See the online article for the color version of this figure.

a 9-point calibration and validation procedure (software provided by SR Research). Calibration was repeated until an error $<1^\circ$ was achieved for all points. The screen background gray tone (RGB 180, 180, 180) was constant across calibration and the experimental task.

Task

Participants performed a Go/NoGo learning task with delayed response execution, called the Oyster Farming Task (Figure 1C). On each trial, participants cultivated an oyster that could either grow 1–5 pearls or 1–5 hazardous tumors. Pearls gained money while tumors cost money for disposal. To maximize the probability that oysters grew pearls, participants needed to learn which oysters to “feed” (Go) and which ones not to feed (“NoGo”). Crucially, participants could choose to reveal the reward (number of pearls) and punishment (number of tumors) at stake prior to action execution in a gaze-contingent design. Participants’ score of accumulated money was turned into a bonus of €0–2 at the end of the task. Participants performed 264 trials split into three blocks of 88 trials (80 trials of the Go/NoGo task, eight catch trials), each with a new set of four oyster types. For detailed information on the instructions, see the original materials used in this study available in the data sharing collection under <https://doi.org/10.34973/05tj-3w64>.

Each trial started with one (of four) abstract action cues (letters from the Agathodaimon alphabet; size $5.2^\circ \times 5.2^\circ$) presented for 700 ms in the center of the screen, representing an oyster type. For each oyster type, there was an optimal action (feed or not feed) that participants needed to learn by trial-and-error. Feeding was only possible when the oysters “opened” later in the trial. The optimal action led to rewards (pearls) in 75% of (valid) trials, otherwise to punishments (tumors; on “invalid trials”). Vice versa, suboptimal actions led to punishments on valid trials, but to rewards on invalid trials. During the action cue presentation, participants were informed about the sides (left vs. right) on which upcoming stakes information (rewards vs. punishments) would appear via faintly colored semicircles in the respective colors (blue and orange, counterbalanced across participants).

Directly after the action cue offset, participants were cued with the exact locations of the stakes and given 1,500 ms to unveil the tumors and pearls at stake on the respective trial. Stakes were revealed in a gaze-contingent fashion: fuzzy circular color patches appeared on the semicircles, which changed into the number of pearls/tumors at stake when participants fixated on them. This eliminated any bottom-up saliency effects (e.g., of stake magnitude) on peripheral vision that could affect participants’ first fixations. To prevent exact preprogramming of saccades, the exact locations of stakes varied across trials. Stakes were located on an invisible circle with a radius of 5.2° visual angle around the screen center (i.e., distance of stakes from the center was kept constant), with a potential vertical displacement of -45 to $+45^\circ$ from the horizontal midline. Vertical displacement was always identical for both pearls and tumors. Stakes were represented by circular areas of interest of 150 pixels (2.7°), with a minimal distance between stakes (at maximal vertical displacement) of 514 pixels (9.4°) and a maximal distance (positioned on the horizontal midline) of 852 pixels (15.6°). Stakes were presented in orange (RGB 200, 100, 7) and blue (RGB 104, 104, 255) of equal luma. Stakes varied in magnitude (1–5 items; total display size $2.6^\circ \times 2.6^\circ$) and magnitude was balanced within action cues (i.e., each of the 20 possible combinations used once per

cue, excluding the five possible combinations in which both magnitudes were identical). The mapping of pearls and tumors to the left/right side varied across trials and was balanced within action cues (each side 10 times per cue) to control for possible participant-specific side biases in gaze.

Stakes offset was followed by a variable interval of 100–500 ms (uniform distribution in steps of 100 ms), after which a release cue (black oyster shape and a food can, $5.2^\circ \times 5.2^\circ$) appeared for 600 ms, indicating that the oyster was about to close and could be fed if necessary. The oyster remained open on either the left or right side, indicating the side where the oyster could be fed. If participants chose to feed the oyster, they had to press the respective button on the open side. The uncertainty about the response side (left/right) at the time of the action cue, which was only resolved with the release cue, prevented premature responding. In-time responses were confirmed by the food can ($1.7^\circ \times 1.7^\circ$) tipping over to the respective side. Seven hundred milliseconds after the release cue offset, the outcome ($3.5^\circ \times 3.5^\circ$) was presented for 1,000 ms. Late responses during the release cue–outcome interval were recorded, but did not affect the outcome. Pressing the incorrect button (i.e., the oyster was open on the left/right, but participants pressed the right/left button) counted as incorrect (i.e., yielded tumors on valid trials) and was confirmed by the can tipping over in the respective direction. Participants received a number of either pearls or tumors, depending on the stakes, their response, and trial validity. Trials finished with a variable intertrial interval between 1,200 and 1,800 ms (uniform distribution in steps of 100 ms).

On eight out of 88 trials per block, participants performed a catching task which incentivized attention to the stakes: instead of the release cue, participants had to report whether the reward or punishments stakes were of greater magnitude (Figure 1C). These catch trials encouraged participants to monitor both stakes and process their magnitude.

Data Preprocessing

Behavior

Catch trials were excluded from all analyses of responses and reaction times (RTs). We further excluded trials with RTs below 200 ms (% trials with button presses per participant: Sample 1: $M = 0.1$, $SD = 0.3$, range = 0–1.5; Sample 2: $M = 0.2$, $SD = 0.3$, range = 0–1.1) because such fast responses could not be expected to incorporate processing of the cue. Likewise, we excluded trials RTs above 800 ms (% trials with button presses per participant: Sample 1: $M = 0.9$, $SD = 1.6$, range = 0–6.8; Sample 2: $M = 0.5$, $SD = 1.8$, range = 0–14.0). This deadline was 200 ms after the release cue offset (i.e., the closing of the response window) as we reasoned that any later responses could have been triggered by the release cue offset. Go responses with the incorrect hand were very rare (% trials with incorrect hand response per participant: Sample 1: $M = 1.7$, $SD = 3.1$, range = 0–14.6; Sample 2: $M = 1.3$, $SD = 2.4$, range = 0–13.3) and not significantly influenced by stakes or dwell times.

Eye-Tracking Preprocessing

Gaze data were processed in R with custom-code. Continuous data were epoched into trials of 1,500 ms relative to stakes onset. Gaps of missing samples up to a duration of 75 ms (due to blinks or saccades) were interpolated using linear interpolation. Trials with more than 50% of missing samples were discarded altogether

(% trials per participant: Sample 1: $M = 4.5$, $SD = 8.0$, range = 0–34.1; Sample 2: $M = 3.5$, $SD = 7.9$, range = 0–52.7). The gaze position was marked as being on the reward/punishment stakes when the gaze position was less than 150 pixels away from the center of the respective stakes image, which was also the criterion in our gaze-contingent design for rendering stakes visible. For each trial, we computed the first fixation on any stakes object (reward or punishment) as well as the total duration (in ms) with which rewards and punishments were fixated over the entire trial (“dwell time”). Absolute dwell times were converted into dwell time difference (reward time minus punishment time; Westbrook et al., 2020).

In some trials, participants only fixated one stake (% trials with at least one fixation per participant: Sample 1: $M = 11.0$, $SD = 14.6$, range = 0–61.4; Sample 2: $M = 10.0$, $SD = 14.4$, range = 0–50.4), leading to ratios of 0 or 1. We thus deviated from our preregistration and reported results for dwell time difference (reward minus punishment dwell time) in the main text, which avoids such an accumulation of values at the edges; analyses of dwell time ratio are reported in [online Supplemental Material 1](#) and led to identical conclusions. Analyses using only the trials on which participants fixated both stakes led to largely identical conclusions.

Data Analysis

General Strategy

We tested hypotheses using mixed-effects linear regression (function `lmer`) and logistic regression (function `glmer`) as implemented in the package `lme4` in R (Bates et al., 2015). We used generalized linear models with a binomial link function (i.e., logistic regression) for binary dependent variables such as responses (Go vs. NoGo) and first fixation, and linear models for continuous variables such as RTs or dwell time. We used zero-sum coding for categorical independent variables. All continuous dependent and independent variables were standardized such that regression weights can be interpreted as standardized regression coefficients. All regression models contained a fixed intercept. We added all possible random intercepts, slopes, and correlations to achieve a maximal random effects structure (Barr et al., 2013). p Values were computed using likelihood ratio tests with the package `afex` (Singmann et al., 2018). We considered p values smaller than $\alpha = 0.05$ as statistically significant.

The main analyses were preregistered for Sample 2 (replication sample; preregistration available under <https://osf.io/nsy5x>). We deviated from our preregistration by reporting results based on dwell time differences (reward minus punishment dwell time) instead of dwell time ratios (reward dwell time divided by reward plus punishment dwell time) in the main text. When participants fixated only one stake, the dwell time ratios were either 0 or 1, regardless of the absolute dwell time on every single fixated option, leading to a loss of information and an accumulation of values at the edges, yielding a distribution with three modes. In contrast, dwell time differences were approximately normally distributed and statistically more comparable to stake differences. Nonetheless, analyses of dwell time ratio and dwell time differences led to identical conclusions as reported in [online Supplemental Material 1](#).

Baseline Learning and Pavlovian Biases

First, following previously established motivational Go–NoGo learning tasks (Guitart-Masip et al., 2011; Swart et al., 2017), we

tested (a) the degree to which participants learned the task, that is, performed more Go responses to Go cues than NoGo cues, and (b) whether responses were influenced by the magnitude of the reward and punishment stakes, reflecting the presence of a Pavlovian bias. For this purpose, we fitted mixed-effects regressions with responses (Go/NoGo) and (as a secondary variable) reaction times as dependent variables and (a) the required action (Go/NoGo) as well as (b) the difference in reward and punishment stake magnitude (ranging from -4 to $+4$) as independent variables. A significant effect of stake difference was followed up with post hoc analyses separating the effects of reward and punishment stake magnitudes, reported in [online Supplemental Material 3](#).

Analysis of Gaze Patterns

Our first key prediction was that action plans, elicited by the oyster cues, directed attention toward action-congruent stakes (reward stake for Go requirement, punishment stake for NoGo requirement). The crucial test of this prediction was whether the action requirement elicited by the cue affected the location of the first fixation (on the reward vs. the punishment stake). This first fixation was not confounded by any bottom-up saliency effects since, in our gaze-contingent design, the magnitudes of the stakes were not visible yet. We used both required action (Go or NoGo) and the difference in the modeled Q values for Go relative to NoGo responses as independent variables to predict the first fixation. These analyses also included catch trials since, during the stakes phase, participants were unaware of whether the trial would be a Go/NoGo or a catch trial. All eye-tracking analyses contained a regressor capturing any participant-specific side biases (overall preference to fixate on the left or right).

Computational Modeling of Action Values

We tested the impact of participants' action intentions on their attention toward the reward and punishment stakes using two operationalizations: Firstly, we approximated participants' intentions by the action required by the presented cue (oyster type). However, this operationalization assumes that participants (have learned and) know the required action. This assumption is violated (a) at the beginning of blocks when participants cannot know the required action yet and still have to acquire it through trial-and-error and (b) even more so in participants who fail to learn the correct response for (some of) the cues. Thus, secondly, as a more proximate measure of participants' beliefs of what they should do, we fitted a simple Rescorla–Wagner Q -learning model to the Go/NoGo response data of each participant. This model uses outcomes r ($+1$ for rewards, -1 for punishments; given that the exact outcome magnitude is irrelevant for learning) to update the action value Q for the chosen action a toward cue s :

$$Q_t(a_t, s_t) = Q_{t-1}(a_t, s_t) + \alpha \times (r - Q_{t-1}(a_t, s_t)) \quad (1)$$

Action values were then translated to action probabilities using a Softmax choice rule:

$$p(\text{Go}, s_t) = \frac{\beta \times e^{Q_t(\text{Go}, s_t)}}{\beta \times e^{Q_t(\text{Go}, s_t)} + \beta \times e^{Q_t(\text{NoGo}, s_t)}} \quad (2)$$

The model featured the free parameters α and β . The learning rate α determines the impact of prediction errors (i.e., higher α leads to stronger incorporation of recent outcomes and discounting of past

outcomes). The inverse temperature β determines the stochasticity of choices (i.e., higher β leads to more deterministic choices in line with action values and lower β to more noisy, stochastic choices). Both parameters were estimated to each participants' data using a grid search, with α in the range [0, 1] in steps of 0.01 (Sample 1: $M = 0.07$, $SD = 0.08$, range = 0.01–0.35; Sample 2: $M = 0.14$, $SD = 0.18$, range = 0.001–0.84) and β in the range of [1, 40] in steps of 0.1 (Sample 1: $M = 8.27$, $SD = 8.21$, range = 1.0–32.7; Sample 2: $M = 8.64$, $SD = 9.57$, range = 1.0–34.8). Starting values for Q_{Go} and Q_{NoGo} were set to 0. Using each participant's best-fitting parameters as well as their actions and outcomes on each trial, we then simulated the action values for Go and NoGo responses on each trial using one-step-ahead predictions (Steingroever et al., 2014). We used the difference term $Q_{Go} - Q_{NoGo}$ as a more proximate measure of participants' action intentions on each trial based on their past experience with each cue. On catch trials (on which participants did not make a Go/NoGo response and did not receive feedback), Q values were not updated but were carried over from the last cue encounter. Similarly, Q values were not updated on trials on which participants responded in the incorrect direction (i.e., pressed left when the oyster was open on the right or vice versa) since participants were instructed that such "directional" errors were always counted as incorrect. Feedback was thus not informative as to whether a Go or NoGo response would have been correct for this cue.

Analysis of Effects of Attention on Responses and Reaction Times

Our second key prediction was that attention to the reward and punishment stakes would shape action execution. To test this prediction, we tested whether the dwell time difference (milliseconds spent on reward stakes minus milliseconds spent attending to punishment stakes) predicted responses (Go vs. NoGo) and response speed (RT, for Go responses only). These analyses excluded catch-task trials (where responses did not relate to learning but to comparing stake magnitudes). All analyses involving responses or reaction times as dependent variables controlled for the required response as well as participant-specific side biases (overall preference to first fixate the left or right). Results did not change when controlling for the Q value difference instead of the required response.

Note that, in our preregistration for Sample 2, we mentioned the plan to fit reinforcement-learning drift-diffusion models to the combined choice and RT data. See online Supplemental Material 3 for a discussion of why these models were unable to reproduce important qualitative patterns present in the empirical data, which was likely due to the tight response deadline and the NoGo response option.

Between-Subjects Correlations of Accuracy

If humans synchronized their attention with their action plans such that Pavlovian biases would align with instrumental action requirements, one would expect this process to facilitate task performance and lead to higher accuracy. To test whether participants with stronger effects of attention on the final response indeed showed higher accuracy, we performed exploratory analyses by computing between-subjects correlations between overall task accuracy and (a) the degree to which stake differences (reward minus punishment stake magnitude) affected responses as well as (b) the degree to which relative dwell time (reward minus punishment dwell time)

affected responses. For this purpose, we refit the respective models on all participants, collapsing across both samples (total $N = 99$), and computed between-subjects correlations between participants' percent correct responses and their respective regression coefficient (fixed + random effect extracted).

Transparency and Openness

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. All data, analysis code, and research materials are available at <https://doi.org/10.34973/05tj-3w64>. Code will be maintained at <https://github.com/johalgermissen/Algermissen2023JEPG>, with a permanent copy at the time of publication at <https://github.com/denoudenlab/Algermissen2023JEPG>. The study design, hypotheses, and analysis plan for Sample 2 were preregistered at <https://osf.io/nsy5x>. Data were analyzed using R, Version 4.1.3 (R Core Team, 2022). Models were fitted with the package lme4, Version 1.1.31 (Bates et al., 2015). Plots were generated with ggplot, Version 3.4.2 (Wickham, 2016).

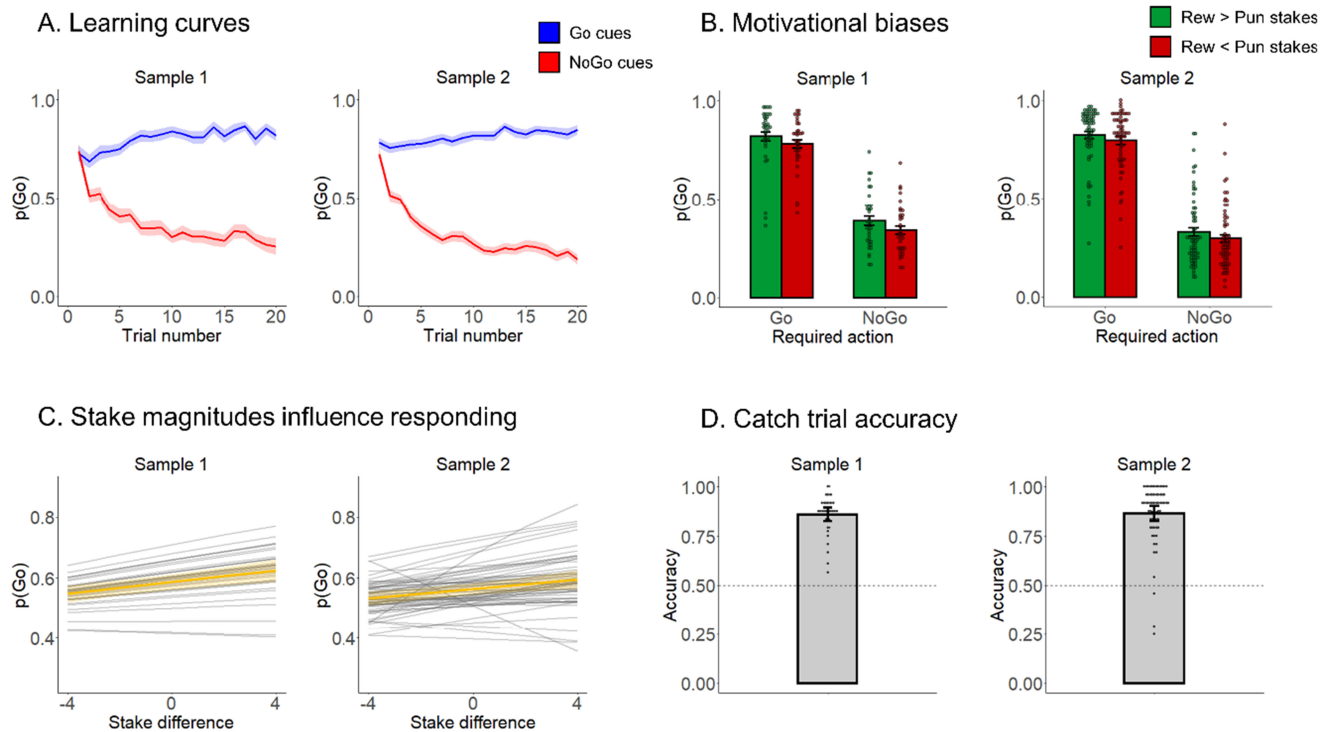
Results

Learning and Pavlovian Biases

Overall, participants learned the Go/NoGo task (% correct, Sample 1: $M = 70.0$, $SD = 10.4$, range = 50.0–87.1; Sample 2: $M = 73.4$, $SD = 13.2$, range = 36.3–91.7), performing significantly more Go responses to Go cues than NoGo cues (Sample 1: $b = 1.08$, 95% CI [0.88, 1.27], $\chi^2(1) = 53.19$, $p < .001$; Sample 2: $b = 1.27$, [1.09, 1.44], $\chi^2(1) = 89.19$, $p < .001$; Figure 2A). Participants also performed well in the catch trials (% correct: Sample 1: $M = 85.8$, $SD = 10.1$, range = 56.5–100; Sample 2: $M = 86.2$, $SD = 15.5$, range = 25.0–100; Figure 2D). Five (seven) people in Sample 1 (2) did not perform significantly above chance (56% correct based on a one-sided binomial test with 240 trials) in the Go/NoGo task. In line with our preregistration, we still included these subjects in all our analyses (for results without these participants, see online Supplemental Material 2). To account for variability in learning, we estimated action (Q) values for each trial based on a Rescorla–Wagner learning model.

Beyond outcome-based learning, responding was affected by the stake magnitudes in a way similar to previously observed Pavlovian biases. A more positive stake difference (reward minus punishment stake magnitude) increased the proportion of Go responses (Sample 1: $b = 0.12$, 95% CI [0.06, 0.17], $\chi^2(1) = 15.32$, $p < .001$; Sample 2: $b = 0.09$, [0.03, 0.15], $\chi^2(1) = 7.92$, $p = .005$; Figure 2B and 2C) and increased response speed (Sample 1: $b = -0.04$, [-0.07, -0.01], $\chi^2(1) = 7.32$, $p = .007$; Sample 2: $b = -0.03$, [-0.05, -0.004], $\chi^2(1) = 6.31$, $p = .012$). The effect of stakes differences did not become weaker over trials or blocks (see online Supplemental Material 3). Separating these effects for the reward and punishment stakes showed that effects were driven by both valences: higher (relative to lower) reward stake magnitude increased responding and speeded up responses, while higher (relative to lower) punishment stake magnitude decreased responding and slowed responses (see online Supplemental Material 3).

In sum, we found evidence that participants learned the task and that the reward and punishment stake magnitudes biased responding in opposite directions, reflecting Pavlovian biases. For reaction times, we found larger reward stake magnitudes to speed up

Figure 2*Task Performance and Pavlovian Biases*

Note. (A) Performance in the Pavlovian Go/NoGo task. Trial-by-trial proportion of Go actions (\pm SEM) for Go cues (blue lines) and NoGo cues (red lines). Shadows indicate standard errors for per-condition-per-participant means. Participants clearly learn whether to make Go actions or not (blue lines go up; red lines go down). (B) Pavlovian biases. Participants perform more Go responses on trials where the reward stake was higher than the punishment stake (green bars; left side of each pair of bars) than vice versa (red bars; right side of each pair of bars). Individual data points reflect response proportion per participant. (C) Stake magnitudes biased responding in a continuous fashion. A higher stake difference (i.e., a reward stake minus punishment stake) resulted in a higher proportion of Go responses. Faint gray lines represent regression lines per participant as predicted by the mixed-effects regression model; the bronze solid line represents the group-level regression line; bronze shading represents mean and 95% confidence intervals. Note the two strong outliers in Sample 2; excluding these outliers did not change conclusions. (D) Performance in the catch trials. Individual data points reflect accuracy per participant. See the online article for the color version of this figure.

responding and larger punishment stake magnitudes to slow down responding, again in line with Pavlovian biases as reported in previous literature (Guitart-Masip et al., 2011; Swart et al., 2017).

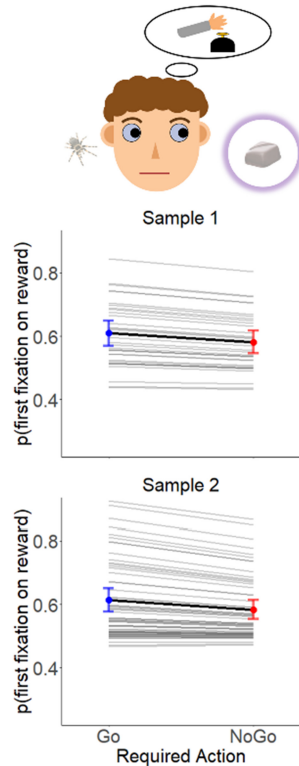
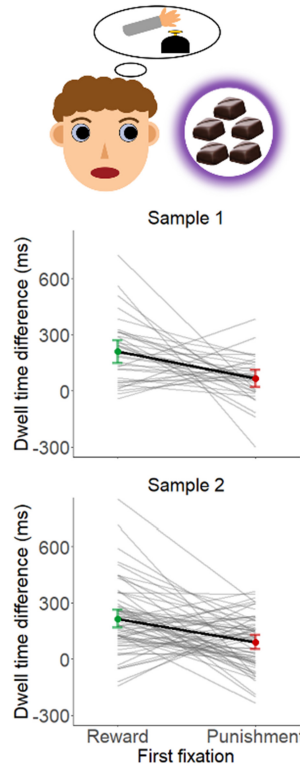
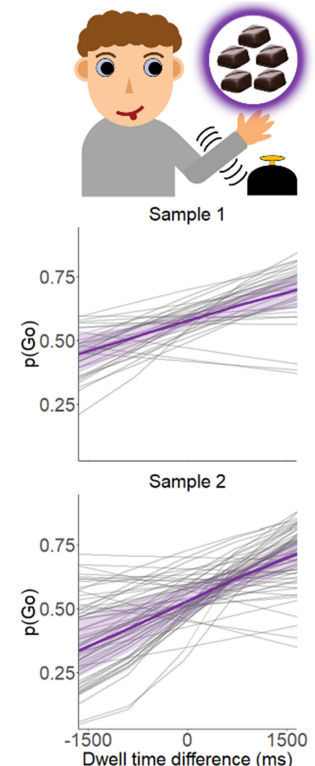
Action Plans Direct Eye Gaze

Next, we tested whether participants' attention was synchronized to their action plans. Such a link would allow Pavlovian biases to be elicited specifically by reward/punishment cues that trigger an action in line with participants' intentions. As a measure of goal-directed attention, we used the first fixation on each trial (Konovalov & Krajbich, 2016), which was unaffected by any bottom-up saliency effects of the (yet to be uncovered) stakes in our gaze-contingent design. On trials that required a Go response, participants were significantly more likely to first fixate on rewards than on trials that required a NoGo response (Sample 1: $b = 0.11$, 95% CI [0.04, 0.19], $\chi^2(1) = 13.92$, $p < .001$; Sample 2: $b = 0.09$, [0.03, 0.15], $\chi^2(1) = 7.88$, $p = .005$; Figure 3A).

This analysis used the required response as a predictor on every trial, which is globally appropriate given that participants learned the task. However, at the beginning of blocks, participants could not know the required response yet. Furthermore, some

participants failed to learn the correct response for (some of) the cues. Thus, as a more proximate measure of participants' beliefs of what they should do, we fitted a simple Rescorla–Wagner model (Rescorla & Wagner, 1972) to the Go/NoGo response data of each participant, simulated the action (Q) values for Go and NoGo responses on each trial, and used the difference $Q_{Go} - Q_{NoGo}$ as a regressor to quantify the trial-by-trial relative value of making a Go relative to NoGo response. At the beginning of a block, this regressor will be zero, and it will stay (close to) zero in case participants fail to learn the correct response. We found that the more Q values favored a Go compared to a NoGo response, the more likely were participants to first fixate the reward (Sample 1: $b = 0.09$, 95% CI [0.03, 0.19], $\chi^2(1) = 8.04$, $p = .005$; Sample 2: $b = 0.13$, [0.05, 0.22], $\chi^2(1) = 9.17$, $p = .002$; online Supplemental Material 4).

We furthermore performed exploratory analyses to test whether action plans affect attention beyond the first fixation, that is, also the overall difference in dwell time to the stakes (dwell time on the reward stake minus dwell time on the punishment stake). This difference was higher when the reward stake was fixated first (Sample 1: $b = 0.18$, 95% CI [0.07, 0.30], $\chi^2(1) = 8.81$, $p < .001$; Sample 2: $b = 0.16$, [0.08, 0.24], $\chi^2(1) = 13.23$, $p < .001$; not

Figure 3*Mutual Influences Between Action and Attention***A. Action plans direct first fixations****B. First fixations anchor dwell times****C. Dwell times influence responses**

Note. (A) Action plans direct first fixations. When required to make a Go action, participants are more likely to first fixate reward information than when a NoGo action was required. (B) First fixation anchors attention. Dwell times are longer on reward stakes compared to punishment stakes when the first fixation was already on reward stakes. Dwell times are additionally shaped by other factors such as the stake magnitudes. (C) Dwell time differences affect final responses. Longer attention to reward compared to punishment stakes resulted in a higher proportion of Go responses. Gray lines = regression lines per participant as predicted by the mixed-effects regression model; black line = group-level regression line; shading = 95% confidence interval. See the online article for the color version of this figure.

significant in either sample when only analyzing trials with both stakes fixated), showing that the first fixation anchored which stakes would receive overall more attention. Over and above this effect, action value kept shaping dwell times, such that people dwelt longer on the reward (compared with the punishment) stake for Go relative to NoGo cues (Sample 1: $b = 0.03$, [0.01, 0.05], $\chi^2(1) = 4.71$, $p = .030$; Sample 2: $b = 0.03$, [0.02, 0.05], $\chi^2(1) = 13.79$, $p < .001$; [online Supplemental Material 4](#)), corroborated when approximating action plans alternatively via Q values (Sample 1: $b = 0.03$, [0.01, 0.05], $\chi^2(1) = 4.36$, $p = .037$; Sample 2: $b = 0.04$, [0.02, 0.06], $\chi^2(1) = 24.82$, $p < .001$; [online Supplemental Material 4](#)). Furthermore, dwell time was influenced by the stake magnitudes, with significantly longer dwell time on the reward stake compared to the punishment stake for more positive stakes differences (Sample 1: $b = 0.09$, [0.05, 0.13], $\chi^2(1) = 16.49$, $p < .001$; Sample 2: $b = 0.12$, [0.09, 0.15], $\chi^2(1) = 41.59$, $p < .001$; see [Figure 3B](#)). This latter effect shows that total dwell time was not completely determined by the first fixation, which was shaped by “top-down” action values, but was additionally sensitive to bottom-up saliency effects of the stake magnitudes.

In sum, we found evidence that participants’ attention to valenced stakes information, in terms of both initial fixation and total dwell time, was synchronized to their initial action plans.

Eye Gaze Predicts Responses

We next assessed whether and how attention shaped the ultimate response. We used the difference in dwell times (reward minus punishment stakes) as an integral measure of total attention ([Kononov & Krajbich, 2016](#)). We controlled for the required action to show that attention predicted the eventual response even beyond participants’ likely intentions.

The longer participants attended to rewards compared to punishments, the more likely they were to make a Go response (Sample 1: $b = 0.13$, 95% CI [0.07, 0.20], $\chi^2(1) = 12.20$, $p < .001$; Sample 2: $b = 0.19$, [0.13, 0.26], $\chi^2(1) = 28.44$, $p < .001$; [Figure 3C](#)). Furthermore, in Sample 2 (but not Sample 1), longer attention to rewards compared to punishments led to faster reaction times (Sample 1: $b = -0.04$, [-0.09, 0.02], $\chi^2(1) = 1.90$, $p = .168$; Sample 2: $b = -0.03$, [-0.05, -0.01], $\chi^2(1) = 4.53$, $p = .033$). When considered in isolation, higher dwell time on rewards

increased responding, but did not significantly affect reaction times, while higher dwell time on punishment decreased responding and slowed responses (see [online Supplemental Material 5](#)). We did not observe any interaction effects between stakes and dwell time effects (see [online Supplemental Material 5](#)).

As action plans both affected attention as well as the ultimate response, one might wonder if the link between attention and the ultimate response was induced by action plans as a “common cause.” To exclude this possibility, all analyses using dwell times to predict responses included the required action as a regressor. Furthermore, we obtained causal evidence for the effect of attention on the ultimate response in a separate online study, in which we manipulated attention. In this study, action cues were presented simultaneously with stakes, but located in close spatial proximity to either the reward or the punishment stakes. We reasoned that the stakes closer to the action cue would receive more attention. Indeed, we observed that action cues were located closer to reward (instead of punishment) stakes resulted in more and faster Go responses. This additional dataset corroborates a causal effect of attention on the ultimate choice. For details, see [online Supplemental Material 6](#).

In sum, we found evidence in both samples that dwell time on rewards/punishments drove responses toward Go/NoGo and speeded/slowed responses, respectively, such that attention determined the eventual strength of Pavlovian biases. Tentative evidence suggested that the effects of stake magnitudes and dwell times were highly similar.

Stake Magnitude and Attentional Effects Differently Relate to Performance

Lastly, given that both stake magnitudes and dwell times affected responses and RTs in a highly similar way, we asked whether these effects also had similar consequences for participants’ overall performance. Crucially, stakes were controlled by the experimental protocol and were therefore unrelated to the required response on each trial. In contrast, attention was under the control of the participant. If participants fixated reward or punishment cues in line with their action goals and then let attention guide their eventual response, strong attention effects could putatively improve their performance. We performed exploratory analyses testing whether the effects of stake magnitudes and dwell times on responding were related to accuracy across participants.

The effect of stake difference on responses correlated significantly negatively with accuracy, $r(97) = -0.24$, 95% CI $[-0.42, -0.04]$, $p = .017$ (see [online Supplemental Material 7](#); after removing two outliers visible: $r(95) = -0.26$, $[-0.44, -0.06]$, $p = .010$; [Figure 4A](#)), while the effect of dwell time difference correlated significantly positively with accuracy, $r(97) = 0.45$, $[0.27, 0.60]$, $p < .001$ ([Figure 4B](#)). Effects were not exclusively driven by reward or punishment stakes/dwell times, but both (in opposite directions, respectively; see [online Supplemental Material 4](#)). We excluded two simpler explanations of the association between the attentional effect and task accuracy: First, this association was not driven by more accurate participants providing higher-quality eye-tracking data (see [online Supplemental Material 7](#)). Second, accuracy was not linked to a stronger focus on reward information (i.e., more first fixation on rewards or longer attention to rewards); if anything, more accurate participants showed a more variable gaze pattern, which supports the idea that these participants could rely in their responses on

their context-appropriate gaze patterns (see [online Supplemental Material 4](#)).

In sum, although correlational, these results suggest that strong attentional effects might facilitate performance, while strong stake magnitudes effects impair it. Based on these analyses, stake magnitude and attentional effects appear to be dissociable.

Discussion

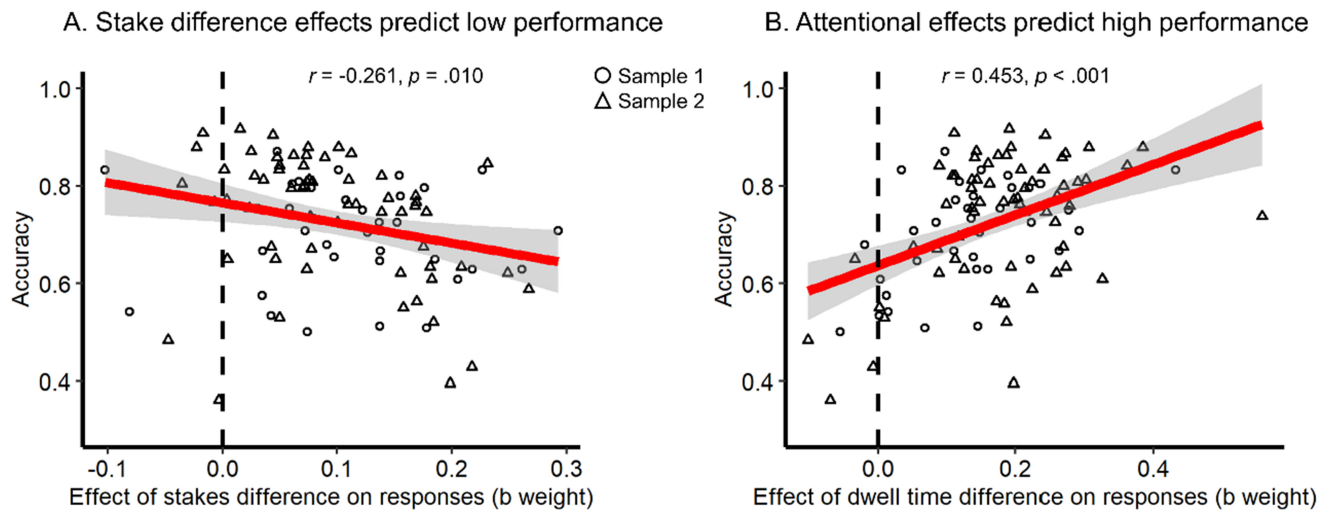
We report evidence from two independent samples showing that instrumental action plans steer attention toward rewards and punishments and in this way shape the input to the Pavlovian control system, triggering responses in line with those action plans. These results shed new light on the possible function of Pavlovian control. In contrast to current theories, we suggest that these biases have an important role beyond providing reasonable response defaults in novel or seemingly uncontrollable environments. Crucially, in addition, Pavlovian control can support instrumental control for efficient and robust action execution. In a novel task, participants successfully learned to perform Go and NoGo actions to various cues. Their responses and reaction times were biased by task-irrelevant information about potential reward/punishment outcomes (stakes), similar to previously reported Pavlovian biases. Most crucially, we found that participants aligned their attention to these stakes with their action plans: they paid more attention to reward stakes when they had to perform a Go action, and relatively more attention to punishment stakes when they had to perform a NoGo action. In turn, attention to these stakes biased ultimate responses, such that more attention to rewards increased the frequency and speed of Go responding. Exploratory between-subjects analyses showed that stronger attentional effects on choice were associated with higher performance, hinting at the adaptive nature of using attention to elicit an automatic response. In sum, these results support the notion that humans can adaptively direct attention to reward and punishment information to selectively elicit Pavlovian biases in line with their action plans.

Current theories often emphasize the “hardwired” nature of Pavlovian biases (Boureau et al., 2015; Dayan et al., 2006) that allow for fast, but inflexible responding. Under the assumption that these biases embody environmental statistics on an evolutionary time scale, they should lead to the correct response in most situations. Normative models assign a dominant role to these biases in contexts that cannot be controlled (yet) by instrumental knowledge about action–outcome relationships (Dorfman & Gershman, 2019). However, once an environment is controllable, biases should disappear. Frequent action slips reveal that Pavlovian biases continue to interfere with goal-directed behavior and require active suppression (Cavanagh et al., 2013; Swart et al., 2018). These cases of interference seem to question their putatively adaptive nature, warranting an update on previous theories.

Here, we suggest that a strong Pavlovian system can be adaptive, even in well-known environments, when it is actively brought into alignment with the goals of other (instrumental) systems. Pavlovian and instrumental control do not need to operate in a strictly parallel fashion and merely interact at the output stage. Instead, we show that instrumental control can determine the input to Pavlovian control by selectively steering attention to (potentially unrelated) reward or punishment information. In this way, it sets the Pavlovian system on a “ballistic” track that will eventually lead to the intended response. Having such an auxiliary mechanism that will

Figure 4

Between-Subjects Correlations Between Global Go/NoGo Task Performance and Stake Magnitude (A) and Attentional (B) Effects



Note. (A) Participants with stronger effects of the stake difference on responding (i.e., steeper slopes in Figure 2C) showed lower performance. (B) Participants with stronger effects of the dwell time difference on responding (i.e., steeper slopes in Figure 3C) showed higher performance. Individual data points reflect per-participant scores, and the red solid line reflects the regression of accuracy on stake magnitude/attentional effects (shades for ± 1 SE). Points = individual participant effects, purple line = regression line, shading = ± 1 SE. See the online article for the color version of this figure.

trigger the intended response might be particularly adaptive in real-life contexts in which the implementation of actions unfolds over time and is prone to interruption by distractors. By “aligning” Pavlovian with instrumental control, action selection becomes more robust against interference. Such a facilitatory effect of Pavlovian control is in line with our finding of better performance in participants with stronger attentional shaping of responses.

Our findings shed new light on the potential use of simple, “fast-and-frugal” systems in decision-making, motor control, and attention. These fields distinguish slow, more computationally demanding, but at the same time more flexible and “accurate” strategies against faster, less demanding, but inflexible and frequently incorrect strategies (Balleine & Dickinson, 1998; Du et al., 2022; McDougle et al., 2016; Theeuwes, 2018). The latter may yield adequate behavior only in a subset of situations, but are frequently misapplied (Beck et al., 2012; Fawcett et al., 2014; Rahnev & Denison, 2018), raising the question of why they are not permanently suppressed beyond contexts of high novelty or uncertainty. In the case of Pavlovian biases, we suggest that these biases can facilitate the implementation of instrumental action plans by making them more robust against distraction. The price of infrequent motor errors caused by Pavlovian biases might be worth paying if, at the same time, the robustness of instrumental action implementation is significantly enhanced. Future research needs to address under which exact circumstances an architecture with a more flexible, sophisticated strategy and more inflexible, simple strategy warrants the infrequent errors produced by the latter.

Beyond the context of Pavlovian biases, our results extend previous literature on the upstream determinants (rather than downstream consequences) of attention allocation. Previous studies have found that, at least early in the choice process, attention appears to be randomly allocated to choice options in a way that is independent of their value (Manohar & Husain, 2013; Westbrook et al., 2020). In

contrast, recent Bayesian accounts of “active sensing” have proposed that attention should be actively driven by the value and uncertainty of choice options in order to gather the maximal amount of information (Callaway et al., 2021; Jang et al., 2021; Sepulveda et al., 2020). We highlight yet another role of attention allocation: to stabilize (or even speed up) action implementation in face of delays and distraction. This role stipulates that (visual) attention is at least partly under the control of ongoing motor processes—as proposed by the premotor theory of attention (Olivers & Roelfsema, 2020; Rizzolatti et al., 1987; Sheliga et al., 1997)—as well as recent accounts highlighting that vision and visual working memory primarily serve action (Heuer et al., 2020; van Ede, 2020).

The idea of Pavlovian biases being recruited by instrumental action plans extends such accounts into the domain of value-based decision-making. It provides a potential explanation for why humans seek out a choice option right before selecting it, even when this will not reveal new information on what is the optimal choice (Hunt et al., 2016; Kaanders et al., 2021). Fixating an (appetitive) option might trigger Pavlovian biases that ensure its selection in face of distractors. Even more so, after participants have made the decision to select an option, its collection and consumption (potentially in face of competitors) might require further motor actions that can benefit from invigoration via these biases. Hence, the role of Pavlovian biases in invigorating motor programs might potentially explain phenomena of human (and animal) curiosity and information seeking (Cervera et al., 2020; Vasconcelos et al., 2015) even after the decision process is finished.

Our results also shed new light on the potential mechanisms by which attention to different choice options affects their eventual choices. Past research has not yet provided evidence on how fixating on a choice option (e.g., a well-known food item like a Snickers) helps its evaluation or affords more information about it. Some accounts have proposed that value-based decisions are made by

retrieving goal-relevant information or “preferences” from memory (Shadlen & Shohamy, 2016). Attention to an option could potentially facilitate the retrieval of value-related information about this option (Callaway et al., 2021). Other studies have observed the effects of attention also on perceptual choices that might not require memory retrieval, suggesting that attention can also affect visual stimulus processing directly (Smith & Krajbich, 2021; Tavares et al., 2017). In contrast to all of these studies, our results suggest that attentional effects might be uncoupled from any features of the choice option and instead be “Pavlovian” in nature: attending to (any) positive information disinhibits motor cortex and facilitates selection, while attending to (any) negative information inhibits motor cortex and leads to rejection—regardless of whether this information is related to the choice option or not.

Crucially, in our paradigm, positive and negative information was unrelated (and orthogonal) to the action that needed to be selected, and thus should not be incorporated into the choice process. However, even this unrelated information did bias choice. To dissociate whether attentional effects are truly driven by increased knowledge about an option’s features rather than a simple (dis-) inhibitory effect of its valence, future research should systematically manipulate the relevance of positive and negative option features to the eventual choice.

There are a few important considerations when generalizing our findings to real-world situations. First, possible outcomes of a choice are often not explicitly presented to an agent. Rather, agents must make a selection among many potentially relevant pieces of information on what they deem important. Our task tried to mimic such situations by allowing agents to freely choose how much to attend to information about rewards and punishments at stake. Still, attention allocation differed from “naturalistic” free-viewing settings in two important ways. Participants were not completely free to attend to the stakes, but were incentivized to do so by the secondary catch task. Furthermore, only two pieces of potential information—exemplary of positive and negative aspects of the situation—were presented, which is a drastic simplification of our information-dense environment. Future extensions of this research should provide participants with a larger set of information to select from, allowing them complete freedom to seek out any information during action preparation.

Second, in real-life situations as well as in this task, people might initiate an action plan, but then change their mind. We only had access to the participants’ ultimate response, which does not allow us to disentangle situations in which they maintained a determined action plan throughout the trial from situations in which action plans were changed based on reward/punishment information. Neuroimaging techniques with high temporal resolution such as electroencephalography and magnetoencephalography could shed light on the dynamic interactions between motor processes and how these change as a function of attentional focus.

Third and finally, exploratory analyses suggested that participants whose ultimate response relied more strongly on attentional inputs showed higher performance. This result corroborates the postulated adaptive nature of a strong Pavlovian system that can be harnessed by instrumental systems. In contrast, the degree to which responses were shaped by the stakes magnitudes (i.e., larger magnitudes resulting in stronger Pavlovian biases) was associated with lower performance. This—at first perhaps surprising—dissociation likely arose from our task design in which stakes magnitudes were orthogonal

to action requirements. When participants performed substantially above chance, stakes magnitudes had a greater potential to disturb action selection on “incongruent” trials (where the required action and the action triggered by the net stakes difference were mismatched) than to facilitate it on “congruent” trials. In contrast, in many real-world contexts, it is adaptive to take into account the size of available rewards or punishments when choosing whether and how vigorously to respond.

Still, even if stakes magnitudes and attention to stakes are both meaningful contributors to choices in real-world settings, it is noteworthy that both had different consequences for performance in our task, suggestive of dissociable behavioral phenotypes. While relying on stake magnitudes might be linked to “sign-tracking” behavior previously observed in animals and humans (Flagel et al., 2009, 2010; Schad et al., 2020) and suggested to constitute a risk factor for addiction (Chen et al., 2023; Garbusow et al., 2016; Robinson & Berridge, 1993), relying on attention might be a “novel” phenotype reflecting strategic recruitment of Pavlovian biases. To conclusively demonstrate the adaptive nature of using attention to invigorate Pavlovian biases, future studies would need to causally manipulate participants’ strategies. Such studies could, for example, train participants to strategically seek out reward or punishment information under a certain action plan. The ability to strategically up- or downregulate Pavlovian biases could then be relevant for future interventions in psychopathologies characterized by aberrant biases, such as depression (Huys et al., 2016) or alcohol addiction (Chen et al., 2023; Garbusow et al., 2016; Schad et al., 2020; Sommer et al., 2017).

In sum, our results suggest a broadening of the current view of Pavlovian control: in addition to providing sensible “default” actions in novel or uncontrollable environments, a strong Pavlovian system can be adaptive even in well-known environments when its robust, almost “ballistic” nature is recruited to ensure that an action plan is implemented even in face of distraction.

Context of This Research

Much literature on Pavlovian biases has focused on situations in which these biases are maladaptive, investigating how they can be suppressed via top-down control (Cavanagh et al., 2013; Swart et al., 2018). However, stronger biases have been found predictive of better recovery from depression (Huys et al., 2016). Furthermore, initial theoretical considerations have proposed that biases could be evaded by mentally reframing a given situation (Boureau & Dayan, 2011) rather than recruiting top-down control. We pursued this line of reasoning experimentally, testing whether humans use attention to reward/punishment cues to create a “Win”/“Avoid” situation that helps them pursue their action goals. This perspective highlights that instrumental and Pavlovian control might more often work in concert rather than oppose each other.

Constraints of Generality

Pavlovian biases might be a universal phenomenon shared by humans and many animal species. They have been described across the animal realm, suggesting a genetic basis shared among humans and other animals and/or a “mandatory” acquisition very early in life in a set of diverse environments. While there are considerable individual differences in the strength of these biases (as described in this article as well as previous work), the direction of their effects both

within and across species is highly consistent, with reward prospect invigorating responding and threat of punishment inhibiting it. Systematically inverted biases have never been observed. In contrast, for the strategic attentional recruitment of these biases, there might be no similar “hard-wired” basis and such a strategy might be acquired by different individuals to different degrees. We speculate that, similar to the biases themselves, the direction of this strategic effect is consistent across individuals. The existence of an “inverted” strategy is highly implausible. The authors would like to highlight that the studies reported in this article were conducted in English and that a significant portion of the participants were not Dutch natives (although this was not systematically assessed), suggesting that the strategic recruitment of biases occurs independently of the local culture where this research has taken place.

References

- Anderson, B. A. (2016). The attention habit: How reward learning shapes attentional selection. *Annals of the New York Academy of Sciences*, 1369(1), 24–39. <https://doi.org/10.1111/nyas.12957>
- Araujo, C., Kowler, E., & Pavel, M. (2001). Eye movements during visual search: The costs of choosing the optimal path. *Vision Research*, 41(25–26), 3613–3625. [https://doi.org/10.1016/S0042-6989\(01\)00196-1](https://doi.org/10.1016/S0042-6989(01)00196-1)
- Armell, K. C., Beaumel, A., & Rangel, A. (2008). Biasing simple choices by manipulating relative visual attention. *Judgment and Decision Making*, 3(5), 396–403. <https://doi.org/10.1017/S1930297500000413>
- Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience*, 7(1), 66–80. <https://doi.org/10.1162/jocn.1995.7.1.66>
- Balleine, B. W., & Dickinson, A. (1998). Goal-directed instrumental action: Contingency and incentive learning and their cortical substrates. *Neuropharmacology*, 37(4–5), 407–419. [https://doi.org/10.1016/S0028-3908\(98\)00033-1](https://doi.org/10.1016/S0028-3908(98)00033-1)
- 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. <https://doi.org/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. <https://doi.org/10.18637/jss.v067.i01>
- Beck, J. M., Ma, W. J., Pitkow, X., Latham, P. E., & Pouget, A. (2012). Not noisy, just wrong: The role of suboptimal inference in behavioral variability. *Neuron*, 74(1), 30–39. <https://doi.org/10.1016/j.neuron.2012.03.016>
- Bekker, H., & Neggers, S. F. W. (2002). Visual search is modulated by action intentions. *Psychological Science*, 13(4), 370–374. <https://doi.org/10.1111/j.0956-7976.2002.00466.x>
- Bettman, J. R., Johnson, E. J., & Payne, J. W. (1990). A componential analysis of cognitive effort in choice. *Organizational Behavior and Human Decision Processes*, 45(1), 111–139. [https://doi.org/10.1016/0749-5978\(90\)90007-V](https://doi.org/10.1016/0749-5978(90)90007-V)
- Boureau, Y.-L., & Dayan, P. (2011). Opponency revisited: Competition and cooperation between dopamine and serotonin. *Neuropsychopharmacology*, 36(1), 74–97. <https://doi.org/10.1038/npp.2010.151>
- Boureau, Y.-L., Sokol-Hessner, P., & Daw, N. D. (2015). Deciding how to decide: Self-control and meta-decision making. *Trends in Cognitive Sciences*, 19(11), 700–710. <https://doi.org/10.1016/j.tics.2015.08.013>
- Breland, K., & Breland, M. (1961). The misbehavior of organisms. *American Psychologist*, 16(11), 681–684. <https://doi.org/10.1037/h0040090>
- Callaway, F., Rangel, A., & Griffiths, T. L. (2021). Fixation patterns in simple choice reflect optimal information sampling. *PLoS Computational Biology*, 17(3), Article e1008863. <https://doi.org/10.1371/journal.pcbi.1008863>
- Cavanagh, J. F., Eisenberg, I., Guitart-Masip, M., Huys, Q. J. M., & Frank, M. J. (2013). Frontal theta overrides Pavlovian learning biases. *Journal of Neuroscience*, 33(19), 8541–8548. <https://doi.org/10.1523/JNEUROSCI.5754-12.2013>
- Cavanagh, J. F., Wiecki, T. V., Kochar, A., & Frank, M. J. (2014). Eye tracking and pupillometry are indicators of dissociable latent decision processes. *Journal of Experimental Psychology: General*, 143(4), 1476–1488. <https://doi.org/10.1037/a0035813>
- Cervera, R. L., Wang, M. Z., & Hayden, B. Y. (2020). Systems neuroscience of curiosity. *Current Opinion in Behavioral Sciences*, 35(1), 48–55. <https://doi.org/10.1016/j.cobeha.2020.06.011>
- Chen, H., Belanger, M. J., Garbusow, M., Kuitunen-Paul, S., Huys, Q. J. M., Heinz, A., Rapp, M. A., & Smolka, M. N. (2023). Susceptibility to interference between Pavlovian and instrumental control predisposes risky alcohol use developmental trajectory from ages 18 to 24. *Addiction Biology*, 28(2), Article e13263. <https://doi.org/10.1111/adb.13263>
- Cisek, P., & Pastor-Bernier, A. (2014). On the challenges and mechanisms of embodied decisions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1655), Article 20130479. <https://doi.org/10.1098/rstb.2013.0479>
- Craighero, L., Fadiga, L., Rizzolatti, G., & Umiltà, C. (1999). Action for perception: A motor-visual attentional effect. *Journal of Experimental Psychology: Human Perception and Performance*, 25(6), 1673–1692. <https://doi.org/10.1037/0096-1523.25.6.1673>
- 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. <https://doi.org/10.1038/nn1560>
- 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. <https://doi.org/10.1016/j.neunet.2006.03.002>
- Dorfman, H. M., & Gershman, S. J. (2019). Controllability governs the balance between Pavlovian and instrumental action selection. *Nature Communications*, 10(1), Article 5826. <https://doi.org/10.1038/s41467-019-13737-7>
- Draschkow, D., Kallmayer, M., & Nobre, A. C. (2021). When natural behavior engages working memory. *Current Biology*, 31(4), 869–874.e5. <https://doi.org/10.1016/j.cub.2020.11.013>
- Du, Y., Krakauer, J. W., & Haith, A. M. (2022). The relationship between habits and motor skills in humans. *Trends in Cognitive Sciences*, 26(5), 371–387. <https://doi.org/10.1016/j.tics.2022.02.002>
- Eimer, M., & Kiss, M. (2008). Involuntary attentional capture is determined by task set: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience*, 20(8), 1423–1433. <https://doi.org/10.1162/jocn.2008.20099>
- Estes, W. K. (1943). Discriminative conditioning. I. A discriminative property of conditioned anticipation. *Journal of Experimental Psychology*, 32(2), 150–155. <https://doi.org/10.1037/h0058316>
- Estes, W. K. (1948). Discriminative conditioning. II. Effects of a Pavlovian conditioned stimulus upon a subsequently established operant response. *Journal of Experimental Psychology*, 38(2), 173–177. <https://doi.org/10.1037/h0057525>
- Fagioli, S., Hommel, B., & Schubotz, R. I. (2007). Intentional control of attention: Action planning primes action-related stimulus dimensions. *Psychological Research*, 71(1), 22–29. <https://doi.org/10.1007/s00426-005-0033-3>
- Fawcett, T. W., Fallenstein, B., Higginson, A. D., Houston, A. I., Mallpress, D. E. W., Trimmer, P. C., & McNamara, J. M. (2014). The evolution of decision rules in complex environments. *Trends in Cognitive Sciences*, 18(3), 153–161. <https://doi.org/10.1016/j.tics.2013.12.012>
- Fiedler, S., & Glöckner, A. (2012). The dynamics of decision making in risky choice: An eye-tracking analysis. *Frontiers in Psychology*, 3, Article 335. <https://doi.org/10.3389/fpsyg.2012.00335>
- Fiedler, S., Schulte-Mecklenbeck, M., Renkewitz, F., & Orquin, J. L. (2020). *Guideline for reporting standards of eye-tracking research in decision sciences*. PsyArXiv. <https://doi.org/10.31234/osf.io/f6qcy>

- Flagel, S. B., Akil, H., & Robinson, T. E. (2009). Individual differences in the attribution of incentive salience to reward-related cues: Implications for addiction. *Neuropharmacology*, 56(Suppl. 1), 139–148. <https://doi.org/10.1016/j.neuropharm.2008.06.027>
- Flagel, S. B., Robinson, T. E., Clark, J. J., Clinton, S. M., Watson, S. J., Seeman, P., Phillips, P. E. M. M., & Akil, H. (2010). An animal model of genetic vulnerability to behavioral disinhibition and responsiveness to reward-related cues: Implications for addiction. *Neuropsychopharmacology*, 35(2), 388–400. <https://doi.org/10.1038/npp.2009.142>
- Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, 18(4), 1030–1044. <https://doi.org/10.1037/0096-1523.18.4.1030>
- Friston, K. J., Daunizeau, J., Kilner, J., & Kiebel, S. J. (2010). Action and behavior: A free-energy formulation. *Biological Cybernetics*, 102(3), 227–260. <https://doi.org/10.1007/s00422-010-0364-z>
- Garbusow, M., Schad, D. J., Sebold, M., Friedel, E., Bernhardt, N., Koch, S. P., Steinacher, B., Kathmann, N., Geurts, D. E. M., Sommer, C., Müller, D. K., Nebe, S., Paul, S., Wittchen, H.-U., Zimmermann, U. S., Walter, H., Smolka, M. N., Sterzer, P., Rapp, M. A., ... Heinz, A. (2016). Pavlovian-to-instrumental transfer effects in the nucleus accumbens relate to relapse in alcohol dependence. *Addiction Biology*, 21(3), 719–731. <https://doi.org/10.1111/adab.12243>
- Gershman, S. J., Markman, A. B., & Otto, A. R. (2014). Retrospective revaluation in sequential decision making: A tale of two systems. *Journal of Experimental Psychology: General*, 143(1), 182–194. <https://doi.org/10.1037/a0030844>
- Geurts, D. E. M., Huys, Q. J. M., den Ouden, H. E. M., & Cools, R. (2013a). Aversive Pavlovian control of instrumental behavior in humans. *Journal of Cognitive Neuroscience*, 25(9), 1428–1441. https://doi.org/10.1162/jocn_a_00425
- Geurts, D. E. M., Huys, Q. J. M., den Ouden, H. E. M., & Cools, R. (2013b). Serotonin and aversive Pavlovian control of instrumental behavior in humans. *Journal of Neuroscience*, 33(48), 18932–18939. <https://doi.org/10.1523/JNEUROSCI.2749-13.2013>
- Gigerenzer, G., & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology*, 62(1), 451–482. <https://doi.org/10.1146/annurev-psych-120709-145346>
- Gottlieb, J., & Oudeyer, P.-Y. (2018). Towards a neuroscience of active sampling and curiosity. *Nature Reviews Neuroscience*, 22(4), 541–548. <https://doi.org/10.1038/s41583-018-0078-0>
- Griffiths, T. L., Lieder, F., & Goodman, N. D. (2015). Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Topics in Cognitive Science*, 7(2), 217–229. <https://doi.org/10.1111/tops.12142>
- Guitart-Masip, M., Fuentemilla, L., Bach, D. R., Huys, Q. J. M., Dayan, P., Dolan, R. J., & Duzel, E. (2011). Action dominates valence in anticipatory representations in the human striatum and dopaminergic midbrain. *Journal of Neuroscience*, 31(21), 7867–7875. <https://doi.org/10.1523/JNEUROSCI.6376-10.2011>
- Hardwick, R. M., Forrence, A. D., Krakauer, J. W., & Haith, A. M. (2019). Time-dependent competition between goal-directed and habitual response preparation. *Nature Human Behaviour*, 3(12), 1252–1262. <https://doi.org/10.1038/s41562-019-0725-0>
- Haselton, M. G., Bryant, G. A., Wilke, A., Frederick, D. A., Galperin, A., Frankenhuys, W. E., & Moore, T. (2009). Adaptive rationality: An evolutionary perspective on cognitive bias. *Social Cognition*, 27(5), 733–763. <https://doi.org/10.1521/soco.2009.27.5.733>
- Hershberger, W. A. (1986). An approach through the looking-glass. *Animal Learning & Behavior*, 14(4), 443–451. <https://doi.org/10.3758/BF03200092>
- Heuer, A., Ohl, S., & Rolf, M. (2020). Memory for action: A functional view of selection in visual working memory. *Visual Cognition*, 28(5–8), 388–400. <https://doi.org/10.1080/13506285.2020.1764156>
- Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature*, 394(6693), 575–577. <https://doi.org/10.1038/29068>
- Hunt, L. T., Rutledge, R. B., Malalasekera, W. M. N., Kennerley, S. W., & Dolan, R. J. (2016). Approach-induced biases in human information sampling. *PLoS Biology*, 14(11), Article e2000638. <https://doi.org/10.1371/journal.pbio.2000638>
- Hutchinson, J. M. C., & Gigerenzer, G. (2005). Simple heuristics and rules of thumb: Where psychologists and behavioural biologists might meet. *Behavioural Processes*, 69(2), 97–124. <https://doi.org/10.1016/j.beproc.2005.02.019>
- Huys, Q. J. M., Cools, R., Gölzer, M., Friedel, E., Heinz, A., Dolan, R. J., & Dayan, P. (2011). Disentangling the roles of approach, activation and valence in instrumental and Pavlovian responding. *PLoS Computational Biology*, 7(4), Article e1002028. <https://doi.org/10.1371/journal.pcbi.1002028>
- Huys, Q. J. M., Gölzer, M., Friedel, E., Heinz, A., Cools, R., Dayan, P., & Dolan, R. J. (2016). The specificity of Pavlovian regulation is associated with recovery from depression. *Psychological Medicine*, 46(5), 1027–1035. <https://doi.org/10.1017/S0033291715002597>
- Jang, A. I., Sharma, R., & Drugowitsch, J. (2021). Optimal policy for attention-modulated decisions explains human fixation behavior. *eLife*, 10, Article e63436. <https://doi.org/10.7554/eLife.63436>
- Kaanders, P., Nili, H., O'Reilly, J. X., & Hunt, L. (2021). Medial frontal cortex activity predicts information sampling in economic choice. *Journal of Neuroscience*, 41(40), 8403–8413. <https://doi.org/10.1523/JNEUROSCI.0392-21.2021>
- Kononov, A., & Krajbich, I. (2016). Gaze data reveal distinct choice processes underlying model-based and model-free reinforcement learning. *Nature Communications*, 7(1), Article 12438. <https://doi.org/10.1038/ncomms12438>
- Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, 13(10), 1292–1298. <https://doi.org/10.1038/nn.2635>
- Lieder, F., & Griffiths, T. L. (2017). Strategy selection as rational metareasoning. *Psychological Review*, 124(6), 762–794. <https://doi.org/10.1037/rev0000075>
- Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, 43, Article e1. <https://doi.org/10.1017/S0140525X1900061X>
- LoLordo, V. M., McMillan, J. C., & Riley, A. L. (1974). The effects upon food-reinforced pecking and treadle-pressing of auditory and visual signals for response-independent food. *Learning and Motivation*, 5(1), 24–41. [https://doi.org/10.1016/0023-9690\(74\)90035-6](https://doi.org/10.1016/0023-9690(74)90035-6)
- Lovibond, P. F. (1983). Facilitation of instrumental behavior by a Pavlovian appetitive conditioned stimulus. *Journal of Experimental Psychology: Animal Behavior Processes*, 9(3), 225–247. <https://doi.org/10.1037/0097-7403.9.3.225>
- Manohar, S., & Husain, M. (2013). Attention as foraging for information and value. *Frontiers in Human Neuroscience*, 7, Article 711. <https://doi.org/10.3389/fnhum.2013.00711>
- Marewski, J. N., & Link, D. (2014). Strategy selection: An introduction to the modeling challenge. *Wiley Interdisciplinary Reviews: Cognitive Science*, 5(1), 39–59. <https://doi.org/10.1002/wcs.1265>
- Mattar, M. G., & Daw, N. D. (2018). Prioritized memory access explains planning and hippocampal replay. *Nature Neuroscience*, 21(11), 1609–1617. <https://doi.org/10.1038/s41593-018-0232-z>
- Mazzoni, P., & Krakauer, J. W. (2006). An implicit plan overrides an explicit strategy during visuomotor adaptation. *Journal of Neuroscience*, 26(14), 3642–3645. <https://doi.org/10.1523/JNEUROSCI.5317-05.2006>
- McDougall, S. D., Ivry, R. B., & Taylor, J. A. (2016). Taking aim at the cognitive side of learning in sensorimotor adaptation tasks. *Trends in Cognitive Sciences*, 20(7), 535–544. <https://doi.org/10.1016/j.tics.2016.05.002>

- McDugle, S. D., & Taylor, J. A. (2019). Dissociable cognitive strategies for sensorimotor learning. *Nature Communications*, 10(1), Article 40. <https://doi.org/10.1038/s41467-018-07941-0>
- Moran, R., Keramati, M., Dayan, P., & Dolan, R. J. (2019). Retrospective model-based inference guides model-free credit assignment. *Nature Communications*, 10(1), Article 750. <https://doi.org/10.1038/s41467-019-08662-8>
- Murayama, K., Usami, S., & Sakaki, M. (2022). Summary-statistics-based power analysis: A new and practical method to determine sample size for mixed-effects modeling. *Psychological Methods*, 27(6), 1014–1038. <https://doi.org/10.1037/met0000330>
- O'Doherty, J. P., Cockburn, J., & Pauli, W. M. (2017). Learning, reward, and decision making. *Annual Review of Psychology*, 68(1), 73–100. <https://doi.org/10.1146/annurev-psych-010416-044216>
- Olivers, C. N. L., & Roelfsema, P. R. (2020). Attention for action in visual working memory. *Cortex*, 131(12), 179–194. <https://doi.org/10.1016/j.cortex.2020.07.011>
- Pachur, T., Schulte-Mecklenbeck, M., Murphy, R. O., & Hertwig, R. (2018). Prospect theory reflects selective allocation of attention. *Journal of Experimental Psychology: General*, 147(2), 147–169. <https://doi.org/10.1037/xge0000406>
- Palminteri, S., Khamassi, M., Joffily, M., & Coricelli, G. (2015). Contextual modulation of value signals in reward and punishment learning. *Nature Communications*, 6(1), Article 8096. <https://doi.org/10.1038/ncomms9096>
- Rahnev, D., & Denison, R. N. (2018). Suboptimality in perceptual decision making. *Behavioral and Brain Sciences*, 41, Article e223. <https://doi.org/10.1017/S0140525X18000936>
- R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.r-project.org/>
- Rescorla, R. A., & Solomon, R. L. (1967). Two-process learning theory: Relationships between Pavlovian conditioning and instrumental learning. *Psychological Review*, 74(3), 151–182. <https://doi.org/10.1037/h0024475>
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). Appleton-Century-Crofts.
- Rieskamp, J., & Otto, P. E. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, 135(2), 207–236. <https://doi.org/10.1037/0096-3445.135.2.207>
- Rizzolatti, G., Riggio, L., Dascola, I., & Umiltà, C. (1987). Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention. *Neuropsychologia*, 25(1), 31–40. [https://doi.org/10.1016/0028-3932\(87\)90041-8](https://doi.org/10.1016/0028-3932(87)90041-8)
- Robinson, M. J. F., & Berridge, K. C. (2013). Instant transformation of learned repulsion into motivational “wanting”. *Current Biology*, 23(4), 282–289. <https://doi.org/10.1016/j.cub.2013.01.016>
- Robinson, T. E., & Berridge, K. C. (1993). The neural basis of drug craving: An incentive-sensitization theory of addiction. *Brain Research Reviews*, 18(3), 247–291. [https://doi.org/10.1016/0165-0173\(93\)90013-P](https://doi.org/10.1016/0165-0173(93)90013-P)
- Schad, D. J., Rapp, M. A., Garbusow, M., Nebe, S., Sebold, M., Obst, E., Sommer, C., Desemo, L., Rabovsky, M., Friedel, E., Romanczuk-Seiferth, N., Wittchen, H.-U., Zimmermann, U. S., Walter, H., Sterzer, P., Smolka, M. N., Schlagenhauf, F., Heinz, A., Dayan, P., & Huys, Q. J. M. (2020). Dissociating neural learning signals in human sign- and goal-trackers. *Nature Human Behaviour*, 4(2), 201–214. <https://doi.org/10.1038/s41562-019-0765-5>
- Schwartz, B. (1976). Positive and negative conditioned suppression in the pigeon: Effects of the locus and modality of the CS. *Learning and Motivation*, 7(1), 86–100. [https://doi.org/10.1016/0023-9690\(76\)90019-9](https://doi.org/10.1016/0023-9690(76)90019-9)
- Sepulveda, P., Usher, M., Davies, N., Benson, A. A., Ortoleva, P., & De Martino, B. (2020). Visual attention modulates the integration of goal-relevant evidence and not value. *eLife*, 9, Article e60705. <https://doi.org/10.7554/eLife.60705>
- Shadlen, M. N., & Shohamy, D. (2016). Decision making and sequential sampling from memory. *Neuron*, 90(5), 927–939. <https://doi.org/10.1016/j.neuron.2016.04.036>
- Sheliga, B. M., Craighero, L., Riggio, L., & Rizzolatti, G. (1997). Effects of spatial attention on directional manual and ocular responses. *Experimental Brain Research*, 114(2), 339–351. <https://doi.org/10.1007/PL00005642>
- Simon, H. A. (1957). *Models of man; social and rational*. Wiley.
- Singmann, H., Bolker, B., Westfall, J., & Aust, F. (2018). *afex: Analysis of factorial experiments*. <https://cran.r-project.org/package=afex>
- Smith, S. M., & Krajbich, I. (2021). Mental representations distinguish value-based decisions from perceptual decisions. *Psychonomic Bulletin & Review*, 28(4), 1413–1422. <https://doi.org/10.3758/s13423-021-01911-2>
- Sommer, C., Garbusow, M., Jünger, E., Poosch, S., Bernhardt, N., Birkenstock, J., Schad, D. J., Jabs, B., Glöckler, T., Huys, Q. M., Heinz, A., Smolka, M. N., & Zimmermann, U. S. (2017). Strong seduction: Impulsivity and the impact of contextual cues on instrumental behavior in alcohol dependence. *Translational Psychiatry*, 7(8), Article e1183. <https://doi.org/10.1038/tp.2017.158>
- Steingrover, H., Wetzels, R., & Wagenmakers, E.-J. (2014). Absolute performance of reinforcement-learning models for the Iowa Gambling Task. *Decision*, 1(3), 161–183. <https://doi.org/10.1037/dec0000005>
- Steinman, R. M., Pizlo, Z., Forofonova, T. I., & Epelboim, J. (2003). One fixates accurately in order to see clearly not because one sees clearly. *Spatial Vision*, 16(3), 225–241. <https://doi.org/10.1163/156856803322467509>
- Swart, J. C., Frank, M. J., Määttä, J. I., Jensen, O., Cools, R., & den Ouden, H. E. M. (2018). Frontal network dynamics reflect neurocomputational mechanisms for reducing maladaptive biases in motivated action. *PLoS Biology*, 16(10), Article e2005979. <https://doi.org/10.1371/journal.pbio.2005979>
- Swart, J. C., Froböse, M. I., Cook, J. L., Geurts, D. E., Frank, M. J., Cools, R., & den Ouden, H. E. (2017). Catecholaminergic challenge uncovers distinct Pavlovian and instrumental mechanisms of motivated (in)action. *eLife*, 6, Article e22169. <https://doi.org/10.7554/eLife.22169>
- Tavares, G., Perona, P., & Rangel, A. (2017). The attentional drift diffusion model of simple perceptual decision-making. *Frontiers in Neuroscience*, 11, Article 468. <https://doi.org/10.3389/fnins.2017.00468>
- Theeuwes, J. (2018). Visual selection: Usually fast and automatic; seldom slow and volitional. *Journal of Cognition*, 1(1), Article 29. <https://doi.org/10.5334/joc.13>
- Todd, P. M., & Brighton, H. (2016). Building the theory of ecological rationality. *Minds and Machines*, 26(1–2), 9–30. <https://doi.org/10.1007/s11023-015-9371-0>
- Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, 76(1), 31–48. <https://doi.org/10.1037/h0026750>
- van der Stigchel, S., & Hollingworth, A. (2018). Visuospatial working memory as a fundamental component of the eye movement system. *Current Directions in Psychological Science*, 27(2), 136–143. <https://doi.org/10.1177/0963721417741710>
- van Ede, F. (2020). Visual working memory and action: Functional links and bi-directional influences. *Visual Cognition*, 28(5–8), 401–413. <https://doi.org/10.1080/13506285.2020.1759744>
- Vasconcelos, M., Monteiro, T., & Kacelnik, A. (2015). Irrational choice and the value of information. *Scientific Reports*, 5(1), Article 13874. <https://doi.org/10.1038/srep13874>
- Weilbacher, R. A., Krajbich, I., Rieskamp, J., & Gluth, S. (2021). The influence of visual attention on memory-based preferential choice. *Cognition*, 215, Article 104804. <https://doi.org/10.1016/j.cognition.2021.104804>

- Westbrook, A., van den Bosch, R., Määttä, J. I., Hofmans, L., Papadopetraki, D., Cools, R., & Frank, M. J. (2020). Dopamine promotes cognitive effort by biasing the benefits versus costs of cognitive work. *Science*, 367(6484), 1362–1366. <https://doi.org/10.1126/science.aaz5891>
- Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. <https://ggplot2.tidyverse.org>
- Wolfe, J. M., Alvarez, G. A., & Horowitz, T. S. (2000). Attention is fast but volition is slow. *Nature*, 406(6797), Article 691. <https://doi.org/10.1038/35021132>
- Wolpert, D. M., & Landy, M. S. (2012). Motor control is decision-making. *Current Opinion in Neurobiology*, 22(6), 996–1003. <https://doi.org/10.1016/j.conb.2012.05.003>
- Yang, S. C.-H., Wolpert, D. M., & Lengyel, M. (2016). Theoretical perspectives on active sensing. *Current Opinion in Behavioral Sciences*, 11, 100–108. <https://doi.org/10.1016/j.cobeha.2016.06.009>

Received June 7, 2022

Revision received March 13, 2023

Accepted March 19, 2023 ■