Title: Intermittent absence of control during reinforcement learning interferes with Pavlovian bias in action selection

Gábor Csifcsák[1], Eirik Melsæter[1], Matthias Mittner[1]

[1]Department of Psychology, UiT The Arctic University of Norway, 9037 Tromsø, Norway

Corresponding author: Gábor Csifcsák, Department of Psychology, UiT The Arctic University of Norway, Huginbakken 32, 9037 Tromsø, Norway, Phone: +47 77 64 67 76, Fax: +47 77 64 52 91, Email: gabor.csifcsak@uit.no

Number of words in abstract: 230

Number of words in introduction, methods, results & discussion: 9337

Number of figures: 5

Number of tables: 1

Declaration of interest: None

Abstract

The ability to control the occurrence of rewarding and punishing events is crucial for our well-being. Two ways to optimize performance are to follow heuristics like Pavlovian biases to approach reward and avoid loss, or to rely more on slowly accumulated stimulus-action associations. Although reduced control over outcomes has been linked to suboptimal decision-making in clinical conditions associated with learned helplessness, it is unclear how uncontrollability of the environment is related to the arbitration between different response strategies.

This study directly tested whether a behavioral manipulation designed to induce learned helplessness in healthy adults (intermittent loss of control over feedback in a reinforcement learning task; "yoking") would modulate the magnitude of Pavlovian bias and the neurophysiological signature of cognitive control (frontal midline theta power) in healthy adults. Using statistical analysis and computational modeling of behavioral data and electroencephalographic signals, we found stronger Pavlovian influences and alterations in frontal theta activity in the yoked group. However, these effects were not accompanied by reduced performance in experimental blocks with regained control, indicating that our behavioral manipulation was not potent enough for inducing helplessness and impaired coping ability with task demands.

We conclude that the level of contingency between instrumental choices and rewards/punishments modulates Pavlovian bias during value-based decision-making, probably via interfering with the implementation of cognitive control. These findings might have implications for understanding the mechanisms underlying helplessness in various psychiatric conditions.

Introduction

Our value-based decisions are influenced by automatic preparatory behaviors, such as approach for potential rewards or response inhibition when facing threat (Clark, Hollon, & Phillips, 2012; Dayan & Berridge, 2014; Dolan & Dayan, 2013; Rangel, Camerer, & Montague, 2008). These (in)action tendencies are controlled by the Pavlovian valuation system, and can be extremely beneficial when agents are required to act or suppress their actions rapidly, without deliberate evaluation of the situation (e.g., halting before stepping on the road if a fast vehicle is approaching). However, Pavlovian influences over instrumental responses can also introduce conflict in decision-making (e.g., eating a delicious-looking cake when one is on a diet), hindering participants' performance under various circumstances (Cavanagh, Eisenberg, Guitart-Masip, Huys, & Frank, 2013; Guitart-Masip et al., 2012; Huys et al., 2012; Swart et al., 2018). In these situations, Pavlovian bias can either exert synergistic or antagonistic effects with other valuation systems that control habitual and goal-directed responding (Dayan & Berridge, 2014).

Decision-making in conflicting situations can be improved via recruiting cognitive control (CC; Ridderinkhof et al., 2004; Alexander and Brown, 2010). On the neural level, successful control over Pavlovian bias under conflict has been linked to theta-band (4-8 Hz) oscillations recorded above the medial prefrontal cortex with electroencephalography (EEG; Cavanagh et al., 2013; Cavanagh and Frank, 2014; Swart et al., 2018). In healthy participants, stronger frontal midline theta (FMθ) activity during decision-making was accompanied by weaker Pavlovian bias and more accurate responding both between and within subjects (Cavanagh et al., 2013; Swart et al., 2018). Thus, as a neural signature of CC, FMθ can be informative on the trial-by-trial variability of top-down inhibition of Pavlovian response

tendencies, and hence on the balance between Pavlovian vs. instrumental (habitual and goal-directed) controllers.

Maladaptive value-based decision-making is a key feature of clinical conditions such as major depression (Chen, Takahashi, Nakagawa, Inoue, & Kusumi, 2015; Eshel & Roiser, 2010). Several streams of evidence point towards suboptimal Pavlovian influences over instrumental choices in states with depressed mood and/or increased anxiety levels. For example, healthy adults with mild depressive symptoms (Huys et al., 2012), those undergoing experimentally induced acute stress (de Berker et al., 2016) or survivors of a severe traumatic event (Ousdal et al., 2018) showed more intensive Pavlovian bias across different experimental settings, whereas the action specificity of Pavlovian influences was found to be attenuated in patients diagnosed with major depression (Huys et al., 2016). Although alterations in Pavlovian bias in these conditions might be related to inefficient CC, to our knowledge, direct evidence for this notion is missing.

Learned helplessness (LH) is a characteristic behavioral response first described in rodents: After receiving uncontrollable painful shocks, animals showed motor passivity and compromised coping ability in aversive situations (Maier & Seligman, 2016; Maier & Seligman, 1976). The concept of LH has been linked to cognitive and affective symptoms of depression (Diener, 2013; Pryce et al., 2011). Originally, it was proposed that during LH induction ("yoking", an experimental protocol in which an animal receives identical shocks as a paired control, but without being able to control them), what animals learn is the non-contingency between their responses and stressful events. The resulting emotional-cognitive state is being transferred to new situations, causing anxiety, reduced motivation, cognitive deficits and ultimately, impaired coping behavior (Maier & Seligman, 1976; Pryce et al., 2011). However, Maier and Seligman (2016) recently proposed that helplessness is not really learned; rather,

the LH response might belong to the animals' behavioral repertoire for responding to aversive events. Based on this new theoretical framework, animals with control over environmental stimuli can override helplessness by recruiting CC mechanisms arising from medial prefrontal areas. Thus, "being in control" is actually the crucial information that animals learn, while helplessness can be regarded as an innate and automatic response. This reformulation of helplessness resembles a special case of Pavlovian bias called "punishment-based suppression" (PBS). Both LH and PBS are (1) elicited by aversive stimuli, (2) dominated by motor inhibition, and (3) associated with altered serotoninergic signaling (Clark et al., 2012; Dayan & Berridge, 2014; Guitart-Masip, Duzel, Dolan, & Dayan, 2014; Maier & Seligman, 2016).

Based on the similarities between LH and Pavlovian bias in action selection, we sought to determine if the intermittent absence of control over outcomes (yoking) in a reinforcement learning (RL) task would induce a state resembling helplessness in healthy adults. Specifically, we anticipated that yoking would enhance Pavlovian bias over performance and weaken CC (indexed by FMθ) during decision-making.

Methods

Participants

Forty-six healthy adults were randomized to 2 experimental groups (N = 23 in each; control group: 14 female, age (M ± SD): 24.1 ± 4.6 years, 20 right-handed; yoked group: 12 female, age: 23.7 ± 2.6 years, 22 right-handed). The two groups did not differ in age (BF_{01} = 3.33), sex (BF_{01} = 3.03), handedness (BF_{01} = 2.22) or other basic personality and affective-cognitive characteristics (Table 1). Randomization was performed in a double-blind manner, so that the experimenter (author E. M.) was unaware of group membership.

Table 1. Descriptives and statistical results for the comparison of data from questionnaires and cognitive tests between the two groups and the two assessments.

Baseline measures	PANAS-Past Positive	PANAS-Past Negative	BAS-Drive	BAS-Fun Seeking	BAS-Reward Resp.	BIS	BHS	OSPAN	Fluency
Control:									
Mean	32.95	17.91	9.91	11.39	16.17	18.78	2.85	11.65	54.26
SD	5.85	6.95	2.10	1.97	2.22	5.24	2.81	2.96	12.41
Yoked:									
Mean	35.69	16.87	9.91	12.00	16.78	19.52	2.69	12.04	49.22
SD	4.44	4.14	2.08	2.15	2.02	3.64	2.03	2.44	13.37
$BF_{01\ Group}$	0.96	2.94	3.44	2.27	2.32	3.03	3.33	3.12	1.69

Repeated measures	PANAS-Present Positive		PANAS-Present Negative		Success score		Control Score	
	Pre	Post	Pre	Post	Day 1	Day 2	Day 1	Day 2
Control:								
Mean	32.26	28.27	12.17	11.13	48.00	60.48	52.39	57.57
SD	5.46	8.26	2.22	1.76	19.59	11.86	18.38	15.39
Yoked:								
Mean	32.78	30.74	12.17	11.83	56.48	65.13	59.87	61.57
SD	6.08	7.95	2.74	3.07	21.00	12.48	14.79	11.65
$BF_{01\ Group}$	2.08		3.12		1.33		1.26	
$BF_{01\ Time}$	0.01		1.51		0.02		2.08	
$BF_{01\ Group\ x\ Time}$	1.75		2.38		2.85		2.77	

Abbreviations: BAS: Behavioral Approach System; BAS-Reward Resp.: BAS-Reward Responsiveness; BHS: Beck Hopelessness Scale; BIS: Behavioral Inhibition System; OSPAN: Operation Span task; PANAS: Positive and Negative Affect Schedule; SD: standard deviation.

All participants met our inclusion criteria (age ≥ 18 years, no history of psychiatric or neurological conditions based on self-report, not under the influence of psychotropic agents or drugs modulating activity of the central nervous system, good or corrected eyesight, sufficient sleep in the preceding night). Participants were informed to receive gift cards worth 100 NOK (approximately equal to 11.7 USD) upon successful completion of the experiment, with the possibility of receiving a bonus of the same amount if their task performance exceeds a pre-defined threshold (not specified in the information sheet). Eventually all participants received the bonus and were told that their performance was satisfactory. All participants provided written consent before the start of data collection. The detailed study protocol was approved by the Institutional Ethics Committee of the Department of Psychology, UiT The

Arctic University of Norway, and complied with the Declaration of Helsinki. Study materials and data are available at https://osf.io/89mdr.

Experimental design

Data collection was scheduled for two different days, separated by a minimum of 1 and a maximum of 12 days. On day 1, they first signed the informed consent and then completed the Norwegian versions of the Positive and Negative Affect Schedule (PANAS) asking about their mood in the past 30 days (PANAS-Past; Gullhaugen and Nøttestad, 2012; Watson et al., 1988), the Behavioral Inhibition System/Behavioral Approach System (BIS/BAS) scales (Brunborg, Johnsen, Mentzoni, Molde, & Pallesen, 2011; Carver & White, 1994) and the Beck Hopelessness Scale (BHS; Beck et al., 1974; Hjemdal et al., 2012). The PANAS consists of 20 statements describing affective states, organized into 2 subscales (PANAS-Positive and PANAS-Negative). The BIS/BAS measures personality attitudes towards approach vs. avoidance behavior in appetitive and aversive situations, respectively. While BIS is regarded to be a unitary concept, BAS scores are divided into three subscales, the BAS-Drive, BAS-Fun Seeking and BAS-Reward Responsiveness. The BHS assesses people's tendency to become or feel hopeless in certain real-life situations, with scores between 0-20. Hopelessness is a psychological construct closely related both to helplessness and the pathogenesis of major depressive disorder (Pryce et al., 2011). Data for the BHS was not collected for 2 control participants.

Following the three questionnaires, participants were asked to read carefully the instructions about our RL task, perform a short practice session (consisting of 20 trials in total), and to complete one "baseline" block of the task (80 trials in total) with the standard response-feedback contingency of 70-30% for correct/incorrect responses (task difficulty was set to be identical as in Cavanagh et al., 2013). At the end of the task, participants were shown two

visual analogue scales to rate between 0 and 100 the degree to which they felt they were successful during the task in obtaining as many points as possible (success score) and how much they think they could control the outcomes by choosing the appropriate response at each trial (control score).

Finally, to ensure similar levels of working memory capacity and executive functioning between groups, participants performed the Operation Span (OSPAN) task (Turner & Engle, 1989), and the phonemic fluency task with letters F, A and S (Ruff, Light, Parker, & Levin, 1997). The fluency score was not collected for 1 yoked participant. Data collection on day 1 lasted for 1-1.5 hours with short breaks between the tests.

On day 2, participants were asked to complete the PANAS, but now focusing on their affect states at the moment (PANAS-Present, Pre-task). After setting up the EEG for recording, participants completed a short practice block (20 trials) and started the long version of the task, consisting of 9 experimental blocks (720 trials in total: 4 cards x 20 repetitions x 9 blocks). This was followed by registering success and control ratings, and the session ended with the repeated collection of PANAS scores (PANAS-Present, Post-task). The whole procedure on day 2 lasted for 2-2.5 hours.

Reinforcement learning task and yoking procedure

We used a modified version of the orthogonalized Go/NoGo task that was specifically designed to investigate the neural correlates of Pavlovian bias (and the control thereof) during instrumental learning and value-based decision-making (Cavanagh et al., 2013; Guitart-Masip et al., 2012). Custom-made cards containing characters (letters) and colored symbols served as cue stimuli (Figure 1A). Participants were told that they were about to play a card game in which 4 cards would be presented in a random order during each of the experimental blocks.
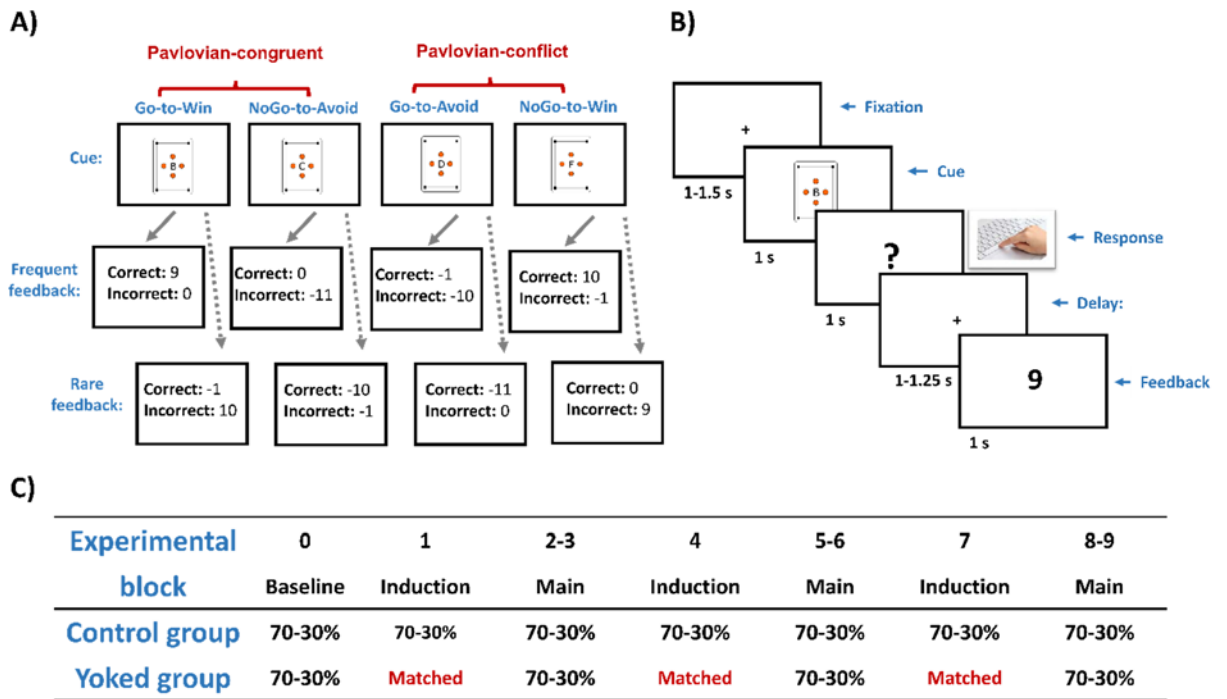
Figure 1. Overview of the behavioral task. (A) In each experimental block, participants were presented with four cards, differing in their action requirement (Go vs. NoGo) and in their associated outcomes (reward vs. loss). Feedback was probabilistic (70-30%); rewards and losses were defined as 10 and -10 points, respectively, but outcomes following an active response were penalized by a Go-cost (-1 points). Action requirements were congruent with the Pavlovian system for two card types (Go-to-Win, NoGo-to-Avoid), whereas two cards (NoGo-to-Win, Go-to-Avoid) induced Pavlovian conflict. (B) In each trial, participants were asked to make decisions about their actions during card presentation, but to make responses only for the response screen (question mark). Feedback was shown after a short delay. (C) On day 1 (block 0: baseline), all participants completed one "normal" block of the task, with probabilistic feedback contingent on their responses. On induction blocks (1, 4, 7) of day 2, yoked participants received random feedback matched to their control pair, i.e., outcomes recorded previously for the same block and card type in a control participant. Other (main) blocks (2, 3, 5, 6, 8, 9) were identical between the two groups, with behavioral control over feedback upon successful learning.

Their task was to maximize their earned points by the end of the game by finding out via trial-and-error, which card should be "picked up" by pressing the space bar with their dominant hand, and which card should be left untouched by remaining passive. Participants were told that (1) some cards were "winning" and some were "losing", (2) there were cards with favorable outcomes following a response and cards associated with no response, (3) these characteristics for each card remained constant for the duration of the experimental block, (4) there was no relationship between cards belonging to different blocks, (5) the task was difficult because feedback was probabilistic and thus there were infrequently presented

misleading outcomes as well (probability levels not told explicitly), (6) so participants were encouraged to explore both response options for all cards on multiple trials, and (7) all outcomes were numerically smaller if they were preceded by an active response relative to remaining passive (i.e., there was a small "Go-cost" of responding, which would reduce wins, modify neutral outcomes and increase losses by -1 points). We informed the participants that the Go-cost resembled the effort of exploring by action, mimicking real-life situations (Teodorescu & Erev, 2014). By introducing the Go-cost, our aim was to promote tendencies of remaining inactive in the yoked group, since behavioral passivity and reduced exploratory behavior are central features of helplessness (Maier & Seligman, 2016; Maier & Seligman, 1976; Teodorescu & Erev, 2014).

Each block contained a new set of cards and there was no relationship between the card sets (i.e., participants had to start learning again at the start of each block). Cards sets were assigned randomly to the 9 experimental blocks. To emphasize the distinction between card sets and blocks, cards within each card set contained the same unique combination of a symbol (circle, diamond, square or star) and color (yellow, rose, blue, green and orange). Importantly, each card contained a unique character that was not used in other experimental blocks (44 characters in total, all being letters from the alphabet, including special characters from Norwegian, German, Hungarian and Serbian: 4 characters for the practice, 4 characters for the "baseline" block on day 1 and 9 blocks x 4 characters for the EEG session on day 2). Except for the practice session, cards in each block were shown 20 times, yielding a total of 80 trials per block.

In each block, the 4 cards within each card set were randomly assigned to one of our four experimental conditions: "Go-to-Win", "NoGo-to-Avoid", "Go-to-Avoid", "NoGo-to-Win" (Figure 1A). For winning cards, our participants' aim was to collect points and avoid the

absence of winning, while for avoid cards, they had to avoid losing points by obtaining neutral outcomes (0 or -1 points, depending on Go-cost). Given that the Pavlovian system promotes approach towards potential rewards and inhibits response tendencies for stimuli associated with losses, the "Go-to-Win" and "NoGo-to-Avoid" conditions are congruent with the Pavlovian system, whereas "Go-to-Avoid" and "NoGo-to-Win" cards induce Pavlovian conflict.

Participants sat in a darkened, sound-shielded room, in front of a 19" Sony Trinitron CRT monitor with a viewing distance of approximately 57 cm and a refresh rate of 100 Hz. Stimuli were presented and responses were recorded with PsychoPy 1.83.04 (Peirce, 2007). Each trial started with a fixation sign (1° x 1° of visual angle) presented at the center of the screen for a duration of 1-1.5 second, followed by the cue (card) presentation (7.8° x 12.8°) for 1 second (Figure 1B). Participants were told that the script would register responses only during the 1-second long presentation of the response screen containing a question mark (3.5° x 5.5°) that would appear after cue offset, so they should refrain from responding directly to the cues. The response screen was followed by a delay period (1-1.25 seconds) with the fixation sign, and the trial ended with the presentation of a feedback screen (1 second) depicting the numerical value of the outcome (3.5-10° x 5.5°). Outcomes could be rewards (10 or 9 points), neutral values (0 or -1 points) or losses (-10 or -11 points), depending on the presence/absence of the Go-cost. Outcomes were probabilistic with a response-feedback contingency of 70-30%, so that correct (incorrect) responses were followed by the favorable (unfavorable) outcome in 70% of the trials, while feedback was misleading in the remaining 30%.

We aimed to induce helplessness in the yoked group by manipulating action-outcome contingency in induction blocks 1, 4 and 7 of day 2 (Figure 1C). Unbeknownst to the participants, each yoked participant was paired with one member of the control group making

sure that data of the control pair was collected first. During induction blocks, yoked participants received random feedback (reward, neutral, loss) recorded earlier for their control pair, saved separately for each specific card in these blocks. However, the outcomes were not necessarily identical between pair members, as the Go-cost was applied for each participant and trial individually. In other words, a feedback for a control participant for a given Go-to-Win trial could be 9 points following a Go response (10 points of reward minus 1 point of Go-cost), whereas the corresponding yoked pair could receive 10 points if no button press was made for the same trial. This experimental design enabled that the net outcome for control and yoked participants during induction blocks was numerically almost identical, and that it was only the uncontrollability of feedback (rather than the prevalence of favorable vs. unfavorable outcomes) that distinguished between the two groups. This aspect of our setup closely resembled the one used in the seminal animal studies of LH, enabling to assess the impact of non-contingency between actions and outcomes on behavior and neural activity (Maier & Seligman, 2016; Maier & Seligman, 1976; Pryce et al., 2011).

Analysis of behavioral data

Response accuracy was calculated for each participant, experimental block and card type separately. In induction blocks, accuracy was defined as the proportion of "correct" responses, taking into account the arbitrary assignment of cards to one of the 4 card types. Based on previous work, we also determined two measures of Pavlovian performance bias (PPB): Reward-based invigoration (RBI) was quantified as the number of Go responses on Win trials/Total number of Go, whereas Punishment-based suppression (PBS) was calculated as the number of NoGo responses on Avoid trials/Total number of NoGo. These indices quantify the participants' likelihood to produce Go (RBI) or NoGo (PBS) responses only for cues associated with reward and punishment, respectively (Cavanagh et al., 2013). Although

usually averaged to estimate the overall magnitude of PPB during the task, we analyzed the RBI and the PBS separately, since we assumed a closer correspondence between PBS and LH. We also calculated the number of Go responses (NumGo) for each experimental block, separately for Pavlovian-congruent (Go-to-Win, NoGo-to-Avoid) and conflict (Go-to-Avoid, NoGo-to-Win) cards to assess if our yoking procedure induced behavioral passivity in yoked participants. Finally, we extracted the total amount of earned points for every block, to compare task performance between the two groups more directly. Accuracy, PPB values, NumGo and total scores were entered into the Bayesian repeated-measures ANOVA function of JASP 0.9.2 (JASP Team, 2018), which implements the Bayesian linear mixed effects model of the BayesFactor package in R (Morey, Rouder, & Jamil, 2015). For all behavioral measures, Block (1-9) was used as within- and Group (control vs. yoked) as between-subject factors, with additional within-subject factors of Valence (win vs. avoid cards; for accuracy), Congruency (Pavlovian-congruent vs. conflict; for accuracy and NumGo) and Index (RBI vs. PBS; for PPB). In contrast to conventional null hypothesis significance testing, Bayesian statistics enabled the estimation of evidence favoring either the alternative ($BF_{10} > 3$) or the null hypothesis ($BF_{01} > 3$). Interactions were assessed using a Bayes Factor ($BF_{inclusion}$) that compares models containing the interaction vs. equivalent models without the interaction term. All Bayesian analyses were performed using default prior scales (r scale fixed effects: 0.5; r scale random effects: 1).

EEG recording and analysis

EEG was recorded at 1000 Hz from 33 channels using a QuickAmp system and actiCAP Ag/AgCl electrodes (Brain Products GmbH, Gilching, Germany), without online frequency filters. Thirty-two electrodes were placed on the scalp in an equidistant arrangement, 2 electrodes under and above the left eye to record vertical electrooculogram (in a bipolar

montage), while the ground and reference electrodes were positioned at locations AFz and FCz, respectively. Data was preprocessed with BrainVision Analyzer 2.1.2 (Brain Products GmbH, Gilching, Germany). First, EEG was high-pass filtered at 0.5 Hz (zero phase shift Butterworth filter, order = 4), followed by ICA-based removal of artifacts related to vertical eye movements. Cue-locked epochs were extracted starting at 1000 ms before and 2000 ms after stimulus onset. Data was baseline corrected (from -100 ms to 0 ms) and epochs containing artifacts related to eye-movements, muscle activity or other non-cerebral sources were removed in a semi-automatic manner combining automatic artifact detection (gradient threshold: 50 μV/ms; amplitude criteria: below -100 μV and above 100 μV; low activity criterion: 0.5 μV/100 ms) and visual inspection of the data. Subsequently, data was re-referenced to the average of mastoids, a frontocentral pooled channel was created from data of four channels (Fz, Cz, FC1, FC2), and epochs containing a blink marker between 0-500 ms were also removed to ensure that stimulus processing in this time interval was not influenced by disruptions in visual input.

We applied three strategies to analyze modulations in cue-locked FMθ between groups and experimental conditions. First, we measured condition-averaged FMθ in an *a priori* defined scalp region, time interval and frequency range (Cavanagh et al., 2013). Following time-frequency transformation of segmented data using continuous complex Morlet wavelets (from 1 Hz to 30 Hz in 30 linear-spaced frequency steps, Morlet parameter c = 3, baseline correction between -300 ms and -200 ms), data was averaged separately for the four trial types (Go-to-Win, NoGo-to-Avoid, Go-to-Avoid, NoGo-to-Win) and two block types (main & induction). FMθ power was extracted from our frontocentral pooled channel for each participant between 175-350 ms and 4-8 Hz (the mean of three wavelet layers with central frequencies at 5.17 Hz, 5.81 Hz and 6.53 Hz, respectively). These FMθ power values were

analyzed using Bayesian repeated-measures ANOVA: Block type (main vs. induction), Valence of the trial (win vs. avoid) and Congruency (congruent vs. conflict) were entered as within-subject and Group (control vs. yoked) as between-subject factors.

To assess if conflict-related modulations in oscillatory brain activity were restricted to the theta band (4-8 Hz) above frontocentral scalp sites in the 175-350 ms time interval, we analyzed our data across all scalp electrodes and in a wider time-frequency range. For this data-driven approach, we down-sampled the data to 250 Hz and entered it into the STUDY structure of EEGLab v14.1.2 (Delorme & Makeig, 2004) in Matlab R2018b (The Mathworks, Natick, MA, United States). Between-group (control vs. yoked) and between-condition (Pavlovian-congruent vs. conflict) modulations in event-related spectral perturbation (ERSP) were analyzed separately for data corresponding to main and induction blocks. ERSP decomposition was performed with Hanning-tapered sinusoidal wavelets starting with 3 cycles at 3 Hz and increasing in 48 log-spaced frequency steps to 10 cycles at 50 Hz, with baseline correction between -300 ms and -200 ms. We used permutation-based 2x2 ANOVA (1000 permutations, false discovery rate (FDR) method to control for multiple comparisons, p < 0.05) to assess (1) frontocentral effects in the 3-12 Hz frequency range up until 450 ms, and (2) scalp-wide effects in our time-frequency cluster of interest.

Finally, we extracted single-trial FMθ power for all artifact-free trials, and used these values in our model-based approach (see: Computational modeling) to investigate how variations in FMθ are related to our Pavlovian bias parameter (π) between experimental groups (control vs. yoked) and block types (main vs. induction).

Computational modeling

To evaluate if yoking influenced latent parameters of RL and decision-making, we implemented three computational models of increasing complexity (M1-M3) to our

behavioral data. Our primary interest was to look for potential group differences in the temporal evolution of the Pavlovian bias parameter π, and to assess if single-trial modulations in FMθ would be related to the magnitude of Pavlovian bias over action selection on the same trial (theta-scaling parameter; ω). In addition, we extracted parameters representing block-wise modulations in the randomness of choice (temperature; β), learning rate (α) and the general tendency to initiate actions (Go-bias; b). We chose to fit these parameters to each block separately (in a hierarchically constrained manner) because we speculated that our yoking manipulation would interfere with other latent processes of learning and decision-making. Specifically, it has been shown that healthy participants can dynamically adjust their learning rates to the volatility of the environment (Behrens, Woolrich, Walton, & Rushworth, 2007), and changes of similar nature were anticipated for choice randomness and Go-bias in the yoked group (Teodorescu & Erev, 2014).

The decision about which action (either "Go" or "NoGo") subject i picks in trial t of block j when stimulus s_t (one of the four possible cards) is presented, is made by a Bernoulli-experiment with probabilities $p(\text{Go})$ and $p(\text{NoGo}) = 1 - p(\text{Go})$. These probabilities are calculated as

$$p(\text{Go}|s_{t,j,i}) = \frac{\exp[W_t(\text{Go}|s_{t,j,i})/\beta_{j,i}]}{\exp[W_t(\text{Go}|s_{t,j,i})/\beta_{j,i}] + \exp[W_t(\text{NoGo}|s_{t,j,i})/\beta_{j,i}]} \qquad [1]$$

i.e., a softmax-function based on the "weight" W_t of each action. The probability is therefore dependent on the relative weight of the possible actions (Go vs. NoGo) as well as the temperature parameter $\beta_{j,i}$ of subject i in block j. This parameter, controling the exploration-exploitation tradeoff, determines how rigid the decisions are biased in favor of the higher-weighted option.

The response-weights W_t are functions of the accumulated Q-values based on the past reinforcement history of the stimulus. For a given, learned stimulus value Q_t,

$$W_t(a|s_{t,j,i}) = \begin{cases} Q_t(\text{Go}|s_{t,j,i}) + b_{j,i} + \pi_{j,i}V(s_{t,j,i}) & \text{if } a = \text{Go} \\ Q_t(\text{NoGo}|s_{t,j,i}) & \text{if } a = \text{NoGo.} \end{cases} \quad [2]$$

Here, a is either Go or NoGo, parameter $b_{j,i}$ codes for a general bias for or against Go responses and $\pi_{j,i}$ is the Pavlovian bias parameter that captures the effect of previous reinforcement history of a stimulus $V(s_{t,j,i})$, independent of the actions that were taken. This parameter has previously been shown to be important for the paradigm employed in our study (Cavanagh et al., 2013; Swart et al., 2018). The reinforcement history of stimulus $s_{t,j,i}$ is cumulated (learned) in a value-representation $V(s_{t,j,i})$ for that stimulus

$$V(s_{t,j,i}) = V_{t-1}(s_{t,j,i}) + \alpha_{j,i}\big[r_{t,i,j} - V_{t-1}(s_{t,j,i}))\big] \quad [3]$$

where $\alpha_{j,i}$ is the learning rate and $r_{t,j,i}$ is the reward (feedback) obtained in trial t. According to the experimental protocol, the feedback is a combination of a reward or punishment of 10 or -10 points plus a small cost of -1 for Go choices. These action outcomes are presented probabilistically, depending on choice accuracy (active response for Go cards and passivity in NoGo trials). Thus, $r_{t,j,i}$ can be one of the following values: $\{-11, -10, -1, 0, 9, 10\}$. This implies that for stimuli that are appetitive (associated with reward) the decision made on the basis of the weights $W_t(a|s_{t,j,i})$ will be biased towards Go responses, while it will bias behavior to passivity (NoGo) when the stimulus is aversive (associated with punishment or loss of reward), given the Pavlovian bias parameter π is positive. The final bit of the model is a standard Q-learning mechanism where stimulus/action pairs receive a value $Q_t(a|s_{t,j,i})$ that are updated to the standard rule

$$Q_t(a|s_{t,j,i}) = Q_{t-1}(a|s_{t,j,i}) + \alpha_{j,i}\big[r_{t,j,i} - Q_{t-1}(a|s_{t,j,i}))\big]. \quad [4]$$

At the start of each block j, the Q-values are initialized to 0 for all action/card pairs which implies a random choice (Go/NoGo) for each card at the first trial.

We model the data from all subject and sessions in the framework of hierarchical Bayesian modeling. We refer the reader to Gelman et al. (2013) for in-depth coverage of the advantages of this approach. In this setup, we model each subjects' parameters in the baseline block β_i, α_i, b_i and π_i as coming from a group-level distribution with means $\mu_\beta, \mu_\alpha, \mu_b$ and μ_π and standard deviations $\sigma_\beta, \sigma_\alpha, \sigma_b$ and σ_π. All subjects took part in a baseline session (block $j = 1$) of our task on day 1, under identical conditions (i.e., no yoking for the yoked group). We therefore model all subject-specific parameters in the first block as coming from the same group-level distribution. We follow the approach taken by Ahn et al. (2016), who model all group-level parameters as coming from a normal distribution with unit information prior and transform the samples to the respective range of the subject-level parameters. Therefore,

$$\Phi^{-1}(\alpha_{j,i}) \sim \mathcal{N}(\mu_\alpha, \sigma_\alpha) \tag{5}$$

where Φ^{-1} is the inverse standard cumulative normal density function,

$$\log \beta_{j,i} \sim \mathcal{N}(\mu_\beta, \sigma_\beta), \tag{6}$$

$$\pi_{j,i} \sim \mathcal{N}(\mu_\pi, \sigma_\pi) \tag{7}$$

and

$$b_{j,i} \sim \mathcal{N}(\mu_b, \sigma_b). \tag{8}$$

To account for changes across experimental blocks j, we allow all group-level parameters to depend on the block parameter and group-membership g such that

$$\begin{aligned} \mu_\alpha &= \mu_{\alpha_0} + \alpha_1^{g,j} \\ \mu_\beta &= \mu_{\beta_0} + \beta_1^{g,j} \\ \mu_\pi &= \mu_{\pi_0} + \pi_1^{g,j} \\ \mu_b &= \mu_{b_0} + b_1^{g,j} \end{aligned} \tag{9-12}$$

where g indicates group (control vs. yoked) and j is the block. The block-level parameters for the first block are set to zero because they constitute the baseline.

Finally, we investigate trial-by-trial variations of the Pavlovian bias with the single-trial estimate of theta power $\theta_{i,j,t}$ extracted from each individual's EEG-data by adding it as a covariate to the Pavlovian parameter π in conflict trials only

$$W_t(a|s_{t,j,i}) = \begin{cases} Q_t(Go|s_{t,j,i}) + b_{j,i} + (\pi_{j,i} + \omega_i\theta_{t,j,i})V(s_{t,j,i}) & \text{if } a = \text{Go and conflict} \\ Q_t(NoGo|s_{t,j,i}) & \text{if } a = \text{NoGo.} \end{cases} \quad [13]$$

This parameter ω_i is modeled similarly as the other group-level parameters such that

$$\omega_i \sim \mathcal{N}(\mu_{\omega_0} + \omega_{\text{induction}} \times \text{induction}, \ \sigma_{\omega_0}) \quad [14]$$

where we added a separate term $\omega_{\text{induction}}$ that is intended to capture deviations of the parameter during induction (blocks 1, 4 and 7). Here, the dummy-variable induction is 1 whenever an induction block is presented (in the yoked group, only) and 0 otherwise. Similarly to Cavanagh et al., 2013, we sigmoid-transformed the standardized theta-estimates, such that the used $\theta_{i,j,t}$ values were

$$\theta_{t,j,i} = \frac{2}{1+\exp\left(\frac{\theta^{raw}{}_{t,j,i} - \langle\theta^{raw}{}_{t,j,i}\rangle_{tj}}{\text{SD}\left(\theta^{raw}{}_{t,j,i}\right)_{tj}}\right)} - 1 \quad [15]$$

where mean and standard deviation (SD) of the raw theta-values were calculated for each subject (across blocks t and sessions j). Missing values were imputed as the median of the remaining theta-values for this subject and block.

The intercepts of the group-level model received the following prior distributions:

$$\begin{aligned}
\mu_{\alpha_0} &\sim \mathcal{N}(-0.4, 0.7), & \sigma_\alpha &\sim \text{HalfCauchy}(0, 0.02) \\
\mu_{\beta_0} &\sim \mathcal{N}(-1.5, 0.8), & \sigma_\beta &\sim \text{HalfCauchy}(0, 0.1) \\
\mu_{\pi_0} &\sim \mathcal{N}(0, 1), & \sigma_\pi &\sim \text{HalfNormal}(0, 1) \quad [16\text{-}20] \\
\mu_{b_0} &\sim \mathcal{N}(0, 1), & \sigma_b &\sim \text{HalfNormal}(0, 1) \\
\mu_{\omega_0} &\sim \mathcal{N}(0, 1), & \sigma_\omega &\sim \text{HalfNormal}(0, 1).
\end{aligned}$$

All the block-level variables $\xi_1^{g,j}$ where $\xi \in \{\alpha, \beta, \pi, b, \omega\}$ and $\omega_{\text{induction}}$, received unit-information priors $\xi_1^{g,j} \sim \mathcal{N}(0,1)$. These prior distributions were picked such that the implied prior on the subject-level parameters was in a reasonable range informed by previous studies.

All models where implemented using Hamiltonian Monte-Carlo algorithms (Hoffman & Gelman, 2014) implemented in Stan (Carpenter et al., 2017). We used six parallel chains with warm-up period of a 1000 samples each such that 6000 samples were drawn from the converged chains. Traceplots for all variables were manually screened for convergence. In addition, we calculated the Gelman-Rubin diagnostic (Gelman & Rubin, 1992) to ensure that all $\widehat{R} \leq 1.05$. We used the Watanabe–Akaike information criterion (WAIC) for model-selection (Watanabe, 2013). WAIC differences larger than 10 can be considered strong (Pratte & Rouder, 2012).

Based on previous work (Cavanagh et al., 2013; Swart et al., 2018), we implemented three models of increasing complexity. The first model (M1) did not differentiate between the groups, such that $\xi_1^{\text{control},j} = \xi_1^{\text{yoked},j}$ for all parameters ξ and blocks j and did not implement any theta-specific effects on the Pavlovian bias, $\omega_i = 0$. The second model (M2) differentiated between the two groups but still left $\omega_i = 0$ and the third model (M3) implemented the full architecture described above. A model selection procedure showed that the second model (WAIC= 38496.7), implementing group differences, was superior to M1 (WAIC= 38526.3), ΔWAIC=14.8 (SE=29.7). Furthermore, M3 (WAIC= 38486.5) implementing the neural covariate was superior to the other two, ΔWAIC= 5.1 (SE= 5.2). We report parameter estimates from the final model implementing all components.

Posterior predictive distributions were calculated from the final model by randomly drawing $n_{rep} = 1000$ samples for the parameters from the subject-level variables from the posterior distribution and generating artificial datasets according to these parameter values.

Each of these 1000 datasets was then summarized in the same way as the real data, i.e., P(Go) was calculated by summing all Go-responses and dividing by the number of subjects. The resulting distribution of P(Go) values was then summarized by their mean and 5%-quantiles.

Results

Yoking enhances the impact of Pavlovian bias on choice behavior

On day 1 (baseline session), we found strong evidence indicating more correct responses in Pavlovian-congruent trials (main effect of Congruency: BF_{10} = 10459.10). There was also an interaction with Valence (BF_{10} = 64.19) suggesting that participants showed stronger Pavlovian bias for Win cards (more correct responses in Go-to-Win relative to NoGo-to-Win trials; Figure 2A). Importantly, the main effect of Group (BF_{01} = 3.03) and all interactions containing this factor (BF_{01} > 1.69) pointed towards similar baseline performance in the two experimental groups. With respect to our two Pavlovian performance bias (PPB) parameters (RBI and PBS), we found strong evidence for higher PBS vs. RBI scores (BF_{10} = 175.63; PBS (M ± SD): 0.61 ± 0.16; RBI: 0.56 ± 0.10). Again, we did not find group differences ($BF_{01 \, Group}$ = 1.75; $BF_{01 \, Group \, x \, Index}$ = 3.44), indicating uniform motivational bias in our cohort. The two groups produced comparable amount of Go responses (BF_{01} = 4.54) and although controls earned more points (control group: 15.52 ± 65.11; yoked group: -5.91 ± 62.41), statistical assessment was inconclusive for this measure and pointing towards an absence of an effect (BF_{01} = 2.04).

On day 2 (EEG session), participants completed the long version of our task (9 blocks consisting of 720 trials). Crucially, we aimed at inducing a state resembling helplessness in the yoked group by withdrawing control over action outcomes in three induction blocks (1, 4 and 7). As for the subjective ratings about the perceived level of success that were collected on both experimental days, we found higher scores on day 2, indicating that participants were

more satisfied with their performance after 9 blocks of the card game relative to their responses after the single baseline block (Table 1).
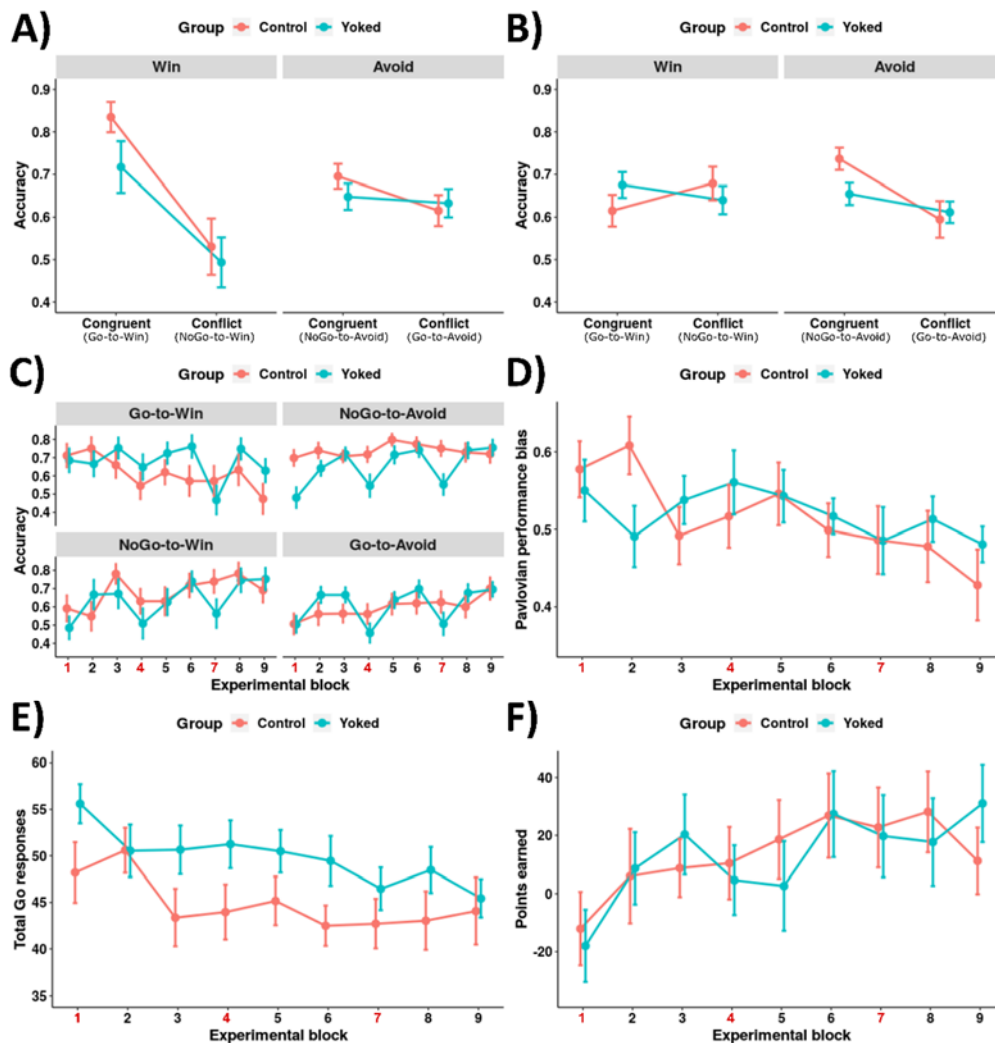
Figure 2. Conventional analysis of behavioral data. Response accuracy obtained from both groups, calculated separately for the four card types in the baseline (A) and EEG sessions (B), and throughout the whole task on Day 2 (C). Changes in Pavlovian performance bias (D), the total number of Go responses (E) and the total amount of points earned (F) across the 9 experimental blocks for both groups. Error bars represent standard errors, block numbers depicted with red are induction blocks in the yoked group.

There was no difference in the levels of perceived control between ratings obtained on the two days. Importantly, neither success nor control scores differed between groups, nor was there an interaction between group and session in this respect. This result was surprising as we expected that the yoked group would rate their levels of success and control lower after the second session. However, considering that during induction, yoked participants received

feedback collected from their control pair that – depending on the control pair's performance – could result in a high number of favorable ("reward" and "no loss") outcomes, the comparable success and control ratings of the two groups is perhaps more understandable. In line with this result, other studies also reported that behavioral signatures of helplessness are not correlated with healthy individuals' perceived control over the environment (Teodorescu & Erev, 2014) and that relatively high reward probabilities can create the "illusion of control" despite the absence of a causal relationship between actions and outcomes (Ly, Wang, Bhanji, & Delgado, 2019). The two groups did not differ in their positive and negative mood scores either, with both groups reporting lower positive affect at the end of the second session, probably due to the long and demanding nature of the task.

Despite our hypothesis regarding reduced response accuracy for the yoked group, their performance deteriorated only during the yoked blocks. Given the non-contingency between responses and feedback, accuracy in these blocks was around chance level, resulting in a very strong main effect of Block (BF_{10} = 45344.46) and moderate evidence for a Block x Group interaction (BF_{10} = 3.11). In addition, we observed a three-way Congruency x Valence x Group interaction (BF_{10} = 38.48), which was due to the surprising low overall accuracy of control participants to rewarding-congruent Go-to-Win cards relative to the conflicting NoGo-to-Win stimuli (Figure 2B). Thus, when averaged across the 9 experimental blocks, response accuracy in the control group indicated no Pavlovian bias for rewarding cards whatsoever. For avoid cards, both groups responded more accurately to congruent NoGo-to-Avoid cards, and this effect was more prominent in controls. More detailed examination of task performance for all four card types across experimental block revealed that with the exception of the Go-to-Win condition, response accuracy was gradually increasing with task progress in both groups, with the expected intermittent drops in performance in the critical induction blocks in the yoked

group (Figure 2C). On Go-to-Win trials, however, the control group showed a surprising gradual reduction in accuracy, from 71.1 ± 30.6 % to 47.4 ± 39.4 % (M ± SD; first vs. final blocks).

This pattern suggests that while successfully overcoming their Pavlovian (Go) response tendencies for rewarding cards in the conflicting NoGo-to-Win condition (accuracy increase from 59.1 ± 34.0 % in block 1 to 69.1 ± 32.4 % in block 9), control participants overcompensated and adopted the same response strategy in Go-to-Win trials, despite being maladaptive for this card type. The yoked group distinguished better between congruent and conflict trials in main blocks, as their improving performance for conflict cards was not accompanied by reductions in task performance in the congruent trials (Figure 2C). These results suggest that the two groups might have also differed in how they implemented CC during the task: Stronger inhibition of Pavlovian bias via CC can be expected exclusively for conflicting cards in the yoked group, whereas the reduction in Pavlovian bias for congruent Go-to-Win cards might be related to strong CC in control participants (see "Reduced control over outcomes disrupts the relationship between FMθ and Pavlovian bias").

With respect to the two pre-defined PPB parameters (RBI and PBS), we found strong evidence for reducing Pavlovian bias by the end of the task (main effect of Block: $BF_{10} = 38.29$). As for our baseline measures, participants showed stronger response inhibition upon aversive stimuli (main effect of Index: $BF_{10} = 14.51$). Although the Bayes Factor for the Block x Group interaction term suggested only weak evidence in favor of between-group differences in the temporal evolution of Pavlovian bias throughout the task ($BF_{10} = 2.14$; Figure 2D), separate analysis for data obtained in the two groups revealed robustly reducing Pavlovian bias in the control group only (Control: $BF_{10} = 8843.61$; Yoked: $BF_{01} = 100.00$). Changes in the magnitude

of PPB across the 9 experimental blocks were found to be similar for RBI and PBS in both groups (Control: BF_{01} = 166.66; Yoked: BF_{01} = 142.85).

With respect to numGo, we found strong evidence for more Go responses in the yoked group (BF_{10} = 156.19; Figure 2E), an effect that was not restricted to main or induction blocks ($BF_{01\ Group\ x\ Block}$ = 250.00). Finally, the amount of points earned throughout the EEG session did not differ between groups ($BF_{01\ Group}$ = 8.13; $BF_{01\ Group\ x\ Block}$ = 40.00; Figure 2F).

In order to investigate the Pavlovian bias and its influence on reinforcement learning, we implemented a computational RL model for our task (Cavanagh et al., 2013; Swart et al., 2018). Model selection revealed that the computational model M3 incorporating single-trial FMθ power and five groups of free parameters (α: learning rate; β: temperature; π: Pavlovian parameter; b: Go-bias; ω: theta-scaling; out of which α, β, π and b were allowed to vary on a block-by-block basis) was superior to the two simpler models M1 and M2 in accounting for variations in behavioral responses. Posterior predictive simulations using the fitted parameter distributions were performed to generate new choices and outcomes for both groups and all main blocks (baseline; blocks 2, 3, 5, 6, 8 and 9 on day 2) according to the winning model (Figure 3A). We found satisfactory correspondence between simulated and real data, as response accuracy was gradually increasing for all four card types (increasing/reducing Go probabilities for Go/NoGo trials, respectively), and we observed decreasing Pavlovian bias with task progress (reducing advantage in accuracy for Pavlovian-congruent Go-to-Win and NoGo-to-Avoid cards relative to conflicting Go-to-Avoid and NoGo-to-Win cards). Model-derived posterior densities for the group-level parameters, indicative of estimates for the first block (Figure 3B), revealed comparable values with earlier studies (Cavanagh et al., 2013; Swart et al., 2018). Notably, the Go-bias parameter b was in the positive value range excluding

zero ($\mu_{b_0} = 0.87 \, [0.45, 1.28]$), successfully capturing participants' overall tendencies to

respond with Go rather than NoGo irrespective of card valence.
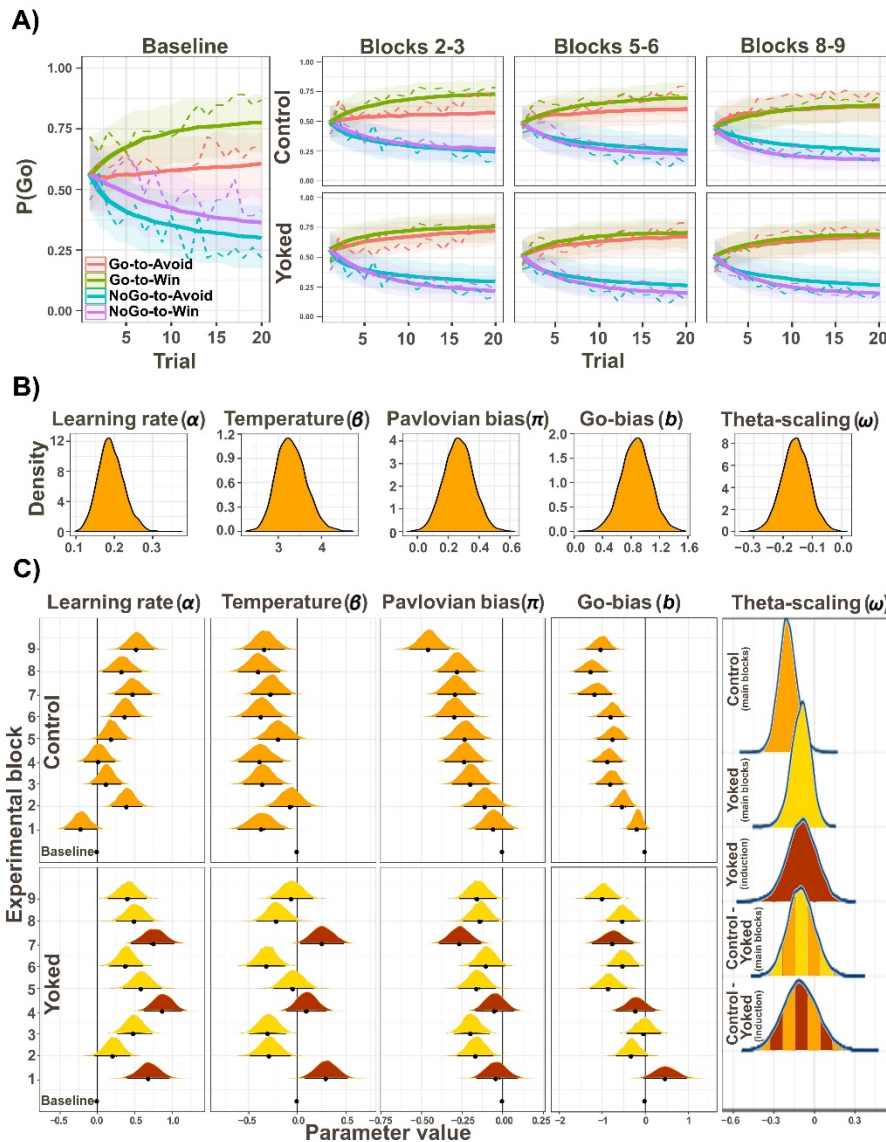
Figure 3. Computational modeling results. (A) Our winning model (M3) captured the main aspects of the trial-by-trial variation of choice behavior for the four card types. Original data is depicted with dashed lines, simulated data using M3 is shown with solid lines along with the 95% highest density interval. (B) Posterior estimates for the five types of free parameters of M3. (C) Changes of model-parameters in the two groups and across all experimental blocks. Estimates from induction blocks (1, 4, 7) are depicted in red. All densities are shown as deviations from values calculated for the baseline block, represented by point estimates.

Conversely, the group-level Pavlovian-bias parameter π also showed a fully positive

distribution ($\mu_{\pi_0} = 0.27 \, [0.08, 0.46]$), confirming that in addition to the Go-bias, participants

were more likely to initiate/suppress actions once they learned that a given card was

potentially rewarding/punishing. Finally, the negative theta-scaling ω parameter ($\mu_{\omega_0} =$ $-0.16\,[-0.25, -0.06]$) was in accordance with earlier observations regarding the negative relationship between single-trial FMθ power and the magnitude of the Pavlovian parameter (π), supporting views that FMθ represents the recruitment of top-down cognitive control over prepotent (in)action tendencies (Cavanagh et al., 2013; Cavanagh & Frank, 2014; Swart et al., 2018).

Crucial to our behavioral manipulation of yoking, our groups differed in how parameter estimates changed during the EEG session (day 2) relative to the baseline block (day 1): Controls showed gradual reductions in their Go-bias (*b*) and Pavlovian-bias (π) parameters, whereas the learning rate and the temperature parameters remained relatively stable throughout the task (Figure 3C). This indicates, that subjects slowly learned to overcome their biases as experience with the task increased. This effect appeared to be weaker in the yoked group. To quantitatively assess if block-by-block variations in α, β, *b* and π parameters differed between the two groups, we estimated the magnitude of parameter changes throughout the task by fitting models where changes in parameter values were linearly constrained across experimental blocks (i.e., linearly in- or decreasing with block number rather than being free to vary across blocks), and calculated 95% highest density intervals (HDIs) for control-yoked group differences for these regression coefficients. This analysis yielded a mean difference of -0.031 [-0.042, -0.018] for Pavlovian bias, -0.018 [-0.065, 0.031] for the Go-bias, 0.023 [-0.005, 0.050] for the learning rate and -0.020 [-0.043, 0.002] for the temperature parameter. Given that 95% HDI estimates were entirely below 0 for π, we conclude that reduction in Pavlovian bias was more pronounced in the control group, indicating that intermittent loss of control in yoked participants interfered with dynamic adjustments of prepotent response tendencies as

the task progressed. Since the 95% HDIs for the slope differences calculated for the other three parameters included zero, similar conclusions could not be drawn for these estimates.

In addition, we observed higher α and β values in induction vs. main blocks in the yoked group (Figure 3C), suggesting that participants assigned more weights to prediction errors during Q-learning updates (α), and produced more stochastic choices characteristic of exploratory behavior (β) when control over feedback was compromised. To test if this effect was robust, we calculated posterior distributions for main-induction block differences for all parameters in the yoked group, merged across blocks 2,3,5,6,8,9 and 1,4,7 respectively. This analysis revealed negative estimates not only for the learning rate (M = -0.346 [-0.402, -0.292]) and temperature parameters (M = -0.419 [-0.459, -0.377]), but also for Pavlovian bias (M = -0.034 [-0.056, -0.011]) and Go-bias (M = -0.371 [-0.477, -0.266]), suggesting that participants increased their learning rates (α), exploration tendencies (β) and predispositions for emitting an active response (b) when the environment became more volatile (Behrens et al., 2007; Pulcu & Browning, 2017). Crucially, the negative estimate for parameter π confirmed our hypothesis regarding the stronger reliance on Pavlovian influences when instrumental learning was ineffective due to yoking. Altogether, the finding of higher π estimates in the yoked group both during and following induction blocks point towards weaker CC in these participants.

Yoking reduces FMϑ in Pavlovian-congruent trials and disrupts the relationship between FMϑ and Pavlovian bias

Since previous work identified frontal 4-8 Hz theta oscillation (FMθ) as a neural signature for CC (Cavanagh & Frank, 2014), we hypothesized that enhanced Pavlovian bias in yoked participants would be accompanied by weaker FMθ during decision-making. First, we measured trial-averaged FMθ power in our *a priori* defined time interval (175-350 ms),

frequency range (4-8 Hz) and scalp region (Fz, FC1, FC2, Cz) and compared it between conditions (Pavlovian congruent vs. conflict), block types (main vs. induction) and groups (control vs. yoked). Bayesian repeated-measures ANOVA revealed the absence of overall conflict-related or group-associated differences ($BF_{01\ Congruency}$ = 6.66 and $BF_{01\ Group}$ = 2.38). However, we found moderate evidence for an interaction between Congruency and Group (BF_{10} = 5.28), an effect that was not influenced by block type ($BF_{01\ Congruency\ x\ Group\ x\ Block\ type}$ = 3.33) or the valence of the card ($BF_{01\ Congruency\ x\ Group\ x\ Valence}$ = 4.36; $BF_{01\ Congruency\ x\ Group\ x\ Block\ type\ x\ Valence}$ = 1.76). We measured stronger FMθ for congruent trials in the control group, and the anticipated effect of congruency (conflict > congruent) was present only in yoked participants (Figure 4).
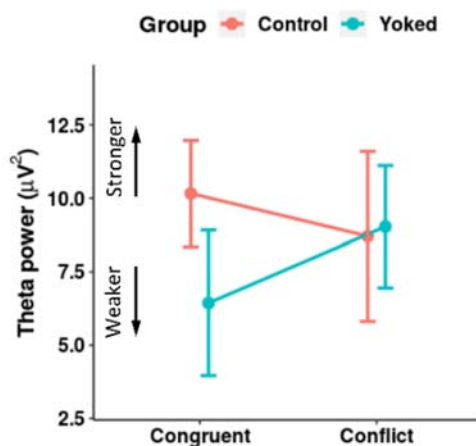
Figure 4. Pavlovian conflict-induced modulations in FMθ power in the two groups. Error bars represent standard errors.

This finding was unexpected since two previous studies have reported enhanced FMθ in very similar tasks for stimuli that induced conflict between prepotent Pavlovian response tendencies and instrumental task requirements (Cavanagh et al., 2013; Swart et al., 2018). However, the current EEG result aligns well with our behavioral observations, namely, that control participants showed deteriorating performance in the Pavlovian-congruent Go-to-Win condition, presumably because they recruited FMθ-related CC mechanisms in a rather

maladaptive way for these cards, resulting in stronger Pavlovian bias by the end of the task. To exclude the possibility that increased FMθ power for Pavlovian-congruent cues in the control group was exclusively related to general motor inhibition in Go-to-Win trials rather than to CC over Pavlovian bias *per se*, we ran Bayesian paired-samples *t*-tests to pairwise compare FMθ power obtained for the four card types and found evidence for comparable FMθ for Go-to-Win and NoGo-to-Avoid cards (main blocks: BF_{01} = 3.95; induction blocks: BF_{01} = 3.99).

To further investigate whether brain responses related to control over conflicting stimuli would occur at different latencies and/or above other brain regions, we conducted a data-driven analysis of event-related spectral perturbations up until 450 ms post-cue onset and between 3-12 Hz, involving all scalp channels. Permutation-based 2x2 ANOVA confirmed the enhanced FMθ power for congruent cards in the control group, causing significant (p < 0.05, FDR-corrected) congruency effects in the yoked data only (Figure 5). However, this effect emerged in a somewhat later time window (200-400 ms) and was shifted posteriorly towards central electrodes (C3, Cz, C4, CP1, CP2) relative to our pre-defined time-frequency cluster.

We also extracted single-trial FMθ power to test if variations in this measure would modulate the impact of Pavlovian bias in subsequent choices on a trial-by-trial basis. We followed earlier reports (Cavanagh et al., 2013; Swart et al., 2018) and incorporated FMθ into the computational model M3 with a scaling parameter ω that captures the strength and the direction of the relationship between parameter π and FMθ (see Methods). We hypothesized that yoking would weaken control over Pavlovian bias, which would manifest in the disrupted relationship between FMθ and π (i.e., in ω values around 0). For this purpose, we also estimated $\omega_{induction}$, an additional parameter that was extracted during induction blocks only,

exclusively in the yoked group. Relative to the value estimations for the control group (-0.198 [-0.327 -0.053]), yoked participants had more positive ω values for both block types, with the corresponding 95% HDIs including zero (Figure 3C; main blocks: -0.100 [-0.243 0.041]; induction blocks: -0.095 [-0.310 0.115]). While these effects indicate that the negative relationship between single-trial FMθ and Pavlovian bias is weaker in the yoked group, the evidence for a group difference in this regard was not compelling as 95% HDI estimates for between-group effects also included zero (Figure 3C; group difference for main blocks: -0.097 [-0.287 0.108]; group difference for induction blocks: -0.103 [-0.361 0.145]).
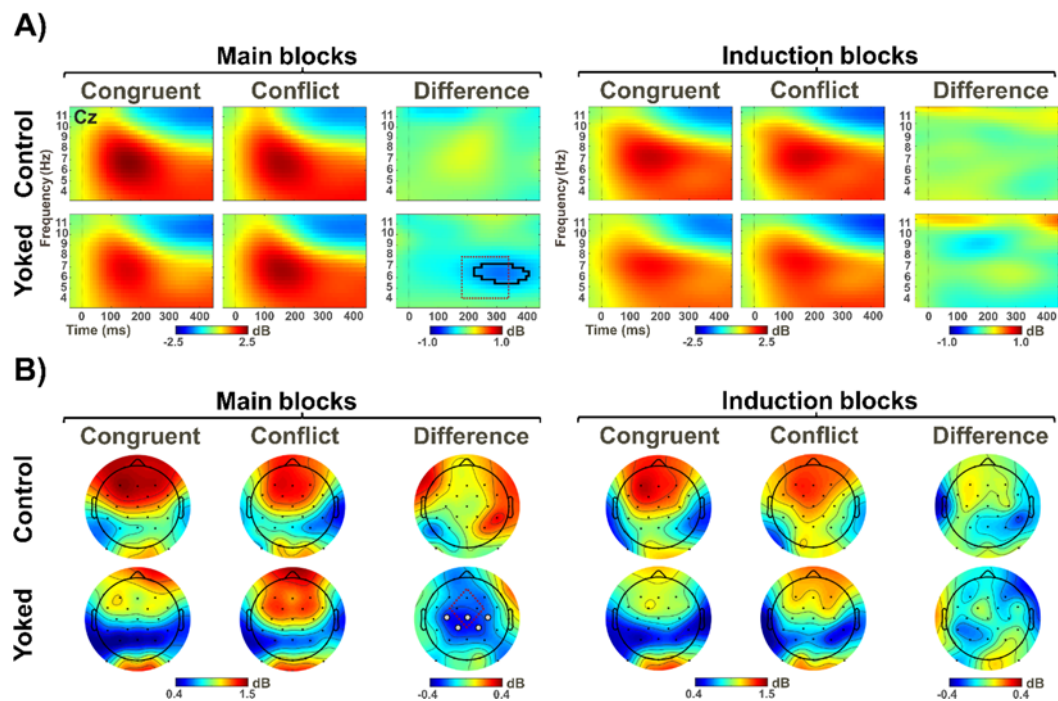
Figure 5. Data-driven analysis of changes in FMθ power. (A) Permutation-based ANOVA revealed significant (p < 0.05, FDR-corrected) conflict > congruent FMθ difference in the yoked group only (region highlighted with solid black line). Although this effect was present in our pre-defined frequency range (4-8 Hz; region highlighted with red dotted line), it was shifted to a later time interval (200-400 ms post-cue onset). No effects were found in induction blocks, indicating that FMθ was altered in yoked participants when feedback signals were unreliable. (B) Scalp distribution of FMθ power (4-8 Hz) in the 200-400 ms time interval showed significant effects in the yoked group only (electrodes highlighted with gray discs), exclusively for main block data. Relative to our pre-defined scalp distribution (electrodes within the red dotted line), the effect shifted to more posterior central electrodes. Again, no effects were found in the induction blocks.

Discussion

Yoking prevents adopting maladaptive decision-making strategies to Pavlovian-congruent cues

In this study, we report that in an environment with stable action-outcome contingencies, healthy adults in our control group gradually adopted maladaptive decision-making strategies when being exposed to the same task multiple times. Specifically, these subjects showed overly reduced Pavlovian response tendencies by avoiding Go responses to rewarding Go-to-Win cards. This effect was accompanied by comparable FMθ power in Pavlovian-congruent vs. conflict trials, as well as a clear negative relationship (indexed by parameter ω) between single-trial FMθ and Pavlovian bias. Given that FMθ is widely recognized as the neural correlate of CC, these results indicate that the control group implemented top-down inhibitory mechanisms over Pavlovian response tendencies also in trials when they were actually beneficial (i.e., Go-to-Win trials).

An intriguing finding of the current study is that intermittent absence of control over rewards and losses during RL prevented our yoked participants from suppressing their Pavlovian bias in Pavlovian-congruent trials, and thus, to rely on Pavlovian influences in a more successful way. Even though our conventional analysis focusing on two indices of Pavlovian performance bias (i.e., RBI and PBS) did not provide compelling evidence for a yoking effect, more sophisticated computational modeling showed (1) that the gradual reduction in parameter π was weaker in yoked participants, and (2) that estimates for π were increased during yoking blocks, confirming our hypothesis about the enhanced influence of Pavlovian response tendencies on choice behavior when the contingency between actions and outcomes was removed. In addition, FMθ was substantially weaker in the yoked group for Pavlovian-congruent vs. conflict cards, indicating that these participants were more successful

in determining whether CC was necessary or not. Finally, our data provided some evidence for the compromised efficacy of CC in reducing Pavlovian bias both during and following yoking, as reflected by the more positive value estimates for parameter ω. In summary, we found advantageous effects of intermittent absence of control over rewards and losses over choice behavior, an outcome that we did not expect when the study was designed.

Although our prediction of observing stronger Pavlovian bias in the yoked vs. control group was supported by the data, the underlying mechanisms turned out to be quite different from what we anticipated. Firstly, our yoked participants' overall response accuracy and the amount of earned points were not impaired in main blocks, pointing towards the absence of yoking-induced deficits in decision-making when feedback was reliable. Since impaired coping with task demands *following* yoking in a key feature of LH, this finding suggests that our yoking procedure was not strong enough to induce helplessness in our participants. Secondly, behavioral and neural data for Pavlovian-conflict trials in main blocks were largely similar in the two groups, indicating adequate conflict-related performance irrespective of group membership. The finding that the difference in FMθ between the two groups stemmed from Pavlovian-congruent instead of conflict trials is surprising, does not align with our hypotheses and is contradicting previous reports (Cavanagh et al., 2013; Swart et al., 2018). Interestingly, however, another study also failed to replicate the conflict-induced FMθ-enhancement in healthy adults (Albrecht, Waltz, Cavanagh, Frank, & Gold, 2016), providing some evidence against the robustness of this phenomenon. Overall, these results suggest that control participants overcompensated their conflict-associated strategies of implementing CC over Pavlovian bias in NoGo-to-Win trials (i.e., by inhibiting responding to win cards), and adopted the same strategy for all rewarding cards, despite being disadvantageous in Go-to-Win trials. Thus, it seems that by having more control over outcomes (i.e., not being yoked in three

blocks), the control group was less successful in evaluating whether CC was necessary or not, which was manifested in comparable FMθ power for congruent vs. conflict cards. What can be the explanation for the stronger CC in the control group and for more appropriate behavioral adjustments in yoked participants? The answer to this question might lie in how uncertainty, caused by impaired control over the environment, influences the valuation of mental effort and the implementation of top-down inhibition during learning and value-based decision-making.

Yoking might interfere with estimations of expected value of control

We propose that stronger uncertainty about the consequences of actions during yoking interfered with the calculation of the "expected value of control (EVC)" (Shenhav, Botvinick, & Cohen, 2013), resulting in more precise conflict-related implementation of CC over Pavlovian bias in the yoked group. It has been proposed that motivation to engage in CC depends on the trade-off between the estimated mental effort and the expected benefit of consuming executive resources (Boureau, Sokol-Hessner, & Daw, 2015; Pessiglione, Vinckier, Bouret, Daunizeau, & Le Bouc, 2017; Shenhav et al., 2013). In these models, the degree of control over environmental events is of key importance, which can be either captured by the advantage of any chosen action relative to a random response (Boureau et al., 2015) or by the probability of achievable outcomes via exerting control (Huys & Dayan, 2009; Shenhav et al., 2013). Since our yoking protocol was deliberately designed to invalidate any effort for influencing action outcomes, this behavioral manipulation could have easily led to the overestimation of CC costs in this group.

On the other hand, we also found evidence for yoking-induced alterations in how outcomes were used to update action values. Modeling results revealed increased learning rates in induction blocks, a sign of overreliance on recent feedback signals that can interfere

with the gradual accumulation of reinforcement history (Behrens et al., 2007; Pulcu & Browning, 2017). This finding indicates that with more uncertainty around choice outcomes, participants adopted a different strategy to utilize information about the value of exerting CC. From the argument above if follows that more specific conflict-associated enhancement of FMθ in the yoked group was likely due to more cautious implementation of CC, as the payoff of mental effort was most probably underestimated.

Uncertainty is generally regarded as the main driving force determining the selection between valuation systems (Boureau et al., 2015; Daw, Niv, & Dayan, 2005). In this study, we aimed to induce helplessness by introducing *unexpected* uncertainty to the environment, because response-feedback contingency was disrupted without any prior warning, and choices could not be reliably shaped via trial-and-error. This type of uncertainty has been linked to norepinephrinergic activity and higher propensity to make more random, exploratory decisions (Cohen, McClure, & Angela, 2007; Pulcu & Browning, 2017). Consistent with this account, we also detected more intensive exploration during yoking (reflected by the increased temperature parameter), but not in main blocks that were dominated by *expected* uncertainty (i.e., stable response-feedback contingencies).

A puzzling finding of the current study is the unexpected increase in cognitive control and the consequential gradual reduction in Go-to-Win performance and Pavlovian bias in our control group. None of the previous studies utilizing similar orthogonalized Go/NoGo task designs reported a similar behavioral phenomenon (Albrecht et al., 2016; Cavanagh et al., 2013; Guitart-Masip et al., 2012; Swart et al., 2018). We note, however, that in those studies, the task was much shorter (40 vs. 180 presentations per stimulus) and often also easier (80-20% vs. 70-30% feedback validity). Thus, it seems that by prolonging the task, maladaptive patterns of behavior can emerge under stable response-feedback contingencies, whereas

increased environmental volatility might protect participants from adopting these suboptimal decision strategies. This finding agrees with a recent study arguing that, with reduced instrumental control over the environment, the instrumental valuation system will necessarily provide more imprecise predictions about outcomes, which in turn, will shift the arbitration between instrumental vs. Pavlovian controllers towards the latter one, leading to enhanced Pavlovian bias in decision-making (Dorfman & Gershman, 2019).

Another potential task parameter underlying the observed effects is the introduction of a small Go-cost. We designed the Go-cost to have minimal impact on the total amount of points earned during an experimental session: With the size of the Go-cost, it is suboptimal to withhold Go responses to all cards relative to emitting at least one correct Go response (+3 points in net income with each correct Go). Still, it is possible that yoking reduced the participants' sensitivity to the Go-cost, resulting in higher proportions of Go responses in general (Figure 2E). Despite this, control participants' response accuracy for Pavlovian-congruent NoGo-to-Avoid cards were higher than for conflicting NoGo-to-Win cards, which would not be the case if their performance was exclusively driven by group differences in Go-cost sensitivity. In a similar vein, increased sensitivity to the Go-cost (or reduced motivation leading to inaction tendency) would not explain why the control group showed gradually increasing Go-to-Avoid accuracy (Figure 3C). Finally, control participants' enhanced FMθ power in congruent trials cannot be exclusively explained by CC related to motor inhibition (rather than to Pavlovian bias), since we found no evidence for difference for this neural measure between congruent Go-to-Win and NoGo-to-Avoid trials.

Yoking modulates activity in the medial prefrontal cortex

Several lines of research support the notion that the effects of yoking in our study are linked to activity in the dorsal anterior cingulate cortex (dACC). The dACC has extensive

connections with striatal and lateral prefrontal regions (Holroyd & Yeung, 2012), and has been associated with conflict detection (Botvinick, Braver, Barch, Carter, & Cohen, 2001), the processing of prediction errors (Holroyd & Yeung, 2012; Walsh & Anderson, 2012), monitoring environmental volatility (Behrens et al., 2007), uncertainty-based competition between valuation systems (Daw et al., 2005), cost/benefit analysis of control-demanding behavior (Holroyd & Yeung, 2012; Pessiglione et al., 2017; Shenhav et al., 2013), signaling the need for cognitive control (Cavanagh & Frank, 2014) and overriding Pavlovian bias (Cavanagh et al., 2013; Swart et al., 2018). Importantly, dACC activity is also modulated by experimental manipulations aiming at inducing LH in healthy volunteers (Bauer, Pripfl, Lamm, Prainsack, & Taylor, 2003; Diener, Kuehner, & Flor, 2010; Salomons et al., 2012). The present results align well with these findings because (1) our computational modeling revealed increased learning rates during LH induction pointing towards dACC involvement (Behrens et al., 2007), and (2) we found that in Pavlovian-congruent trials, yoking influenced FMθ, an electrophysiological signature of dACC activity (Cavanagh & Frank, 2014; Narayanan, Cavanagh, Frank, & Laubach, 2013; Wang, Ulbert, Schomer, Marinkovic, & Halgren, 2005; Womelsdorf, Johnston, Vinck, & Everling, 2010).

Implications for learned helplessness

Although our paradigm was designed to induce a state resembling LH in healthy adults, we acknowledge that we did not accomplish this aim. First, explicit success and control ratings after the EEG session were similar between groups, and second, intermittent yoking was not accompanied by worse task performance in main blocks. Nevertheless, we did observe carry-over effects from induction blocks to non-yoked parts of the experiment, manifesting in better accuracy in Go-to-Win trials (Figure 2C), higher prevalence of Go responses (Figure 2E), larger Pavlovian bias parameter estimates (Figure 3C) and weaker FMθ power (Figure 5) in main

blocks. Moreover, we showed that while not having control over action outcomes, our subjects demonstrated higher learning rates, more random responding, stronger Go-bias and somewhat weaker relationship between Pavlovian bias and single-trial FMθ (Figure 3C). Therefore, we conclude that our behavioral manipulation was successful for uncovering neural and behavioral aspects of reduced control over reinforcers. Despite the above concerns, our findings lend some support to the reformulation of LH put forward by Maier and Seligman (2016) that behavior in yoked animals might not be learned, but rather be a part of their innate (Pavlovian) behavioral repertoire when top-down inhibition arising from the medial prefrontal cortex is weak. It is up to future studies to develop more potent experimental setups for investigating how more prominent versions of LH interfere with value-based decision-making in health and disease.

References

Ahn, W. Y., Haines, N., & Zhang, L. (2017). Revealing neurocomputational mechanisms of reinforcement learning and decision-making with the hBayesDM package. *Computational Psychiatry*, *1*, 24–57.

Albrecht, M. A., Waltz, J. A., Cavanagh, J. F., Frank, M. J., & Gold, J. M. (2016). Reduction of Pavlovian bias in schizophrenia: Enhanced effects in clozapine-administered patients. *PloS One*, *11*(4), e0152781.

Alexander, W. H., & Brown, J. W. (2010). Computational models of performance monitoring and cognitive control. *Topics in Cognitive Science*, *2*(4), 658–677.

Bauer, H., Pripfl, J., Lamm, C., Prainsack, C., & Taylor, N. (2003). Functional neuroanatomy of learned helplessness. *Neuroimage*, *20*(2), 927–939.

Beck, A. T., Weissman, A., Lester, D., & Trexler, L. (1974). The measurement of pessimism: The hopelessness scale. *Journal of Consulting and Clinical Psychology*, *42*(6), 861–865.

Behrens, T. E., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*(9), 1214–1221.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*(3), 624–652.

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.

Brunborg, G. S., Johnsen, B. H., Mentzoni, R. A., Molde, H., & Pallesen, S. ale. (2011). Individual differences in evaluative conditioning and reinforcement sensitivity affect bet-sizes during gambling. *Personality and Individual Differences*, *50*(5), 729–734.

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., … Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*(1), 1–32.

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

Cavanagh, J. F., Eisenberg, I., Guitart-Masip, M., Huys, Q., & Frank, M. J. (2013). Frontal theta overrides pavlovian learning biases. *The Journal of Neuroscience*, *33*(19), 8541–8548.

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, *18*(8), 414–421.

Chen, C., Takahashi, T., Nakagawa, S., Inoue, T., & Kusumi, I. (2015). Reinforcement learning in depression: A review of computational research. *Neuroscience & Biobehavioral Reviews*, *55*, 247–267.

Clark, J. J., Hollon, N. G., & Phillips, P. E. (2012). Pavlovian valuation systems in learning and decision making. *Current Opinion in Neurobiology*, *22*(6), 1054–1061.

Cohen, J. D., McClure, S. M., & Angela, J. Y. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *362*(1481), 933–942.

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.

Dayan, P., & Berridge, K. C. (2014). Model-based and model-free Pavlovian reward learning: Revaluation, revision, and revelation. *Cognitive, Affective, & Behavioral Neuroscience*, *2*(14), 473–492.

de Berker, A. O., Tirole, M., Rutledge, R. B., Cross, G. F., Dolan, R. J., & Bestmann, S. (2016). Acute stress selectively impairs learning to act. *Scientific Reports*, *6*, 29816.

Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Diener, C. (2013). Altered associative learning and learned helplessness in major depression. In *Psychiatric Disorders—New Frontiers in Affective Disorders (Schoepf D, ed)* (pp. 57–78). London: InTech.

Diener, C., Kuehner, C., & Flor, H. (2010). Loss of control during instrumental learning: A source localization study. *Neuroimage*, *50*(2), 717–726.

Dolan, R. J., & Dayan, P. (2013). Goals and habits in the brain. *Neuron*, *80*(2), 312–325.

Dorfman, H. M., & Gershman, S. J. (2019). Controllability Governs the Balance Between Pavlovian and Instrumental Action Selection. *BioRxiv*, 596577. https://doi.org/10.1101/596577

Eshel, N., & Roiser, J. P. (2010). Reward and punishment processing in depression. *Biological Psychiatry*, *68*(2), 118–124.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2013). *Bayesian Data Analysis, (Chapman & Hall/CRC Texts in Statistical Science)*.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 457–472.

Guitart-Masip, M., Duzel, E., Dolan, R., & Dayan, P. (2014). Action versus valence in decision making. *Trends in Cognitive Sciences*, *18*(4), 194–202.

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

Gullhaugen, A. S., & Nøttestad, J. A. (2012). Under the surface: The dynamic interpersonal and affective world of psychopathic high-security and detention prisoners. *International Journal of Offender Therapy and Comparative Criminology*, *56*(6), 917–936.

Hjemdal, O., Friborg, O., & Stiles, T. C. (2012). Resilience is a good predictor of hopelessness even after accounting for stressful life events, mood and personality (NEO-PI-R). *Scandinavian Journal of Psychology*, *53*(2), 174–180.

Hoffman, M. D., & Gelman, A. (2014). The No-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, *15*(1), 1593–1623.

Holroyd, C. B., & Yeung, N. (2012). Motivation of extended behaviors by anterior cingulate cortex. *Trends in Cognitive Sciences*, *16*(2), 122–128.

Huys, Q. J., 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.

Huys, Quentin JM, & Dayan, P. (2009). A Bayesian formulation of behavioral control. *Cognition*, *113*(3), 314–328.

Huys, Quentin JM, Eshel, N., O'Nions, E., Sheridan, L., Dayan, P., & Roiser, J. P. (2012). Bonsai trees in your head: How the Pavlovian system sculpts goal-directed choices by pruning decision trees. *PLoS Computational Biology*, *8*(3), e1002410.

JASP Team. (2018). *JASP (Version 0.9)*. Retrieved from https://jasp-stats.org/

Ly, V., Wang, K. S., Bhanji, J., & Delgado, M. R. (2019). A reward-based framework of perceived control. *Frontiers in Neuroscience*, *13*, 65. https://doi.org/10.3389/fnins.2019.00065

Maier, S. F., & Seligman, M. E. (2016). Learned helplessness at fifty: Insights from neuroscience. *Psychological Review*, *123*(4), 349–367.

Maier, Steven F., & Seligman, M. E. (1976). Learned helplessness: Theory and evidence. *Journal of Experimental Psychology: General*, *105*(1), 3–46.

Morey, R. D., Rouder, J. N., & Jamil, T. (2015). BayesFactor: Computation of Bayes factors for common designs. *R Package Version 0.9*.

Narayanan, N. S., Cavanagh, J. F., Frank, M. J., & Laubach, M. (2013). Common medial frontal mechanisms of adaptive control in humans and rodents. *Nature Neuroscience*, *16*(12), 1888–1895.

Ousdal, O. T., Huys, Q. J., Milde, A. M., Craven, A. R., Ersland, L., Endestad, T., … Dolan, R. J. (2018). The impact of traumatic stress on Pavlovian biases. *Psychological Medicine*, *48*(2), 327–336.

Peirce, J. W. (2007). PsychoPy—Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13.

Pessiglione, M., Vinckier, F., Bouret, S., Daunizeau, J., & Le Bouc, R. (2017). Why not try harder? Computational approach to motivation deficits in neuro-psychiatric diseases. *Brain*, *141*, 629–650.

Pratte, M. S., & Rouder, J. N. (2012). Assessing the dissociability of recollection and familiarity in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(6), 1591–1607.

Pryce, C. R., Azzinnari, D., Spinelli, S., Seifritz, E., Tegethoff, M., & Meinlschmidt, G. (2011). Helplessness: A systematic translational review of theory and evidence for its relevance to understanding and treating depression. *Pharmacology & Therapeutics*, *132*(3), 242–267.

Pulcu, E., & Browning, M. (2017). Affective bias as a rational response to the statistics of rewards and punishments. *Elife*, *6*, e27879.

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience*, *9*(7), 545.

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science*, *306*(5695), 443–447.

Ruff, R. M., Light, R. H., Parker, S. B., & Levin, H. S. (1997). The psychological construct of word fluency. *Brain and Language*, *57*(3), 394–405.

Salomons, T. V., Moayedi, M., Weissman-Fogel, I., Goldberg, M. B., Freeman, B. V., Tenenbaum, H. C., & Davis, K. D. (2012). Perceived helplessness is associated with individual differences in the central motor output system. *European Journal of Neuroscience*, *35*(9), 1481–1487.

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron*, *79*(2), 217–240.

Swart, J. C., Frank, M. J., Määttä, J. I., Jensen, O., Cools, R., & den Ouden, H. E. (2018). Frontal network dynamics reflect neurocomputational mechanisms for reducing maladaptive biases in motivated action. *PLoS Biology*, *16*(10), e2005979.

Teodorescu, K., & Erev, I. (2014). Learned helplessness and learned prevalence: Exploring the causal relations among perceived controllability, reward prevalence, and exploration. *Psychological Science*, *25*(10), 1861–1869.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language*, *28*(2), 127–154.

Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: Event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience and Biobehavioral Reviews*, *36*(8), 1870–1884.

Wang, C., Ulbert, I., Schomer, D. L., Marinkovic, K., & Halgren, E. (2005). Responses of human anterior cingulate cortex microdomains to error detection, conflict monitoring, stimulus-response mapping, familiarity, and orienting. *The Journal of Neuroscience*, *25*(3), 604–613.

Watanabe, S. (2013). A widely applicable Bayesian information criterion. *Journal of Machine Learning Research*, *14*(Mar), 867–897.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063–1070.

Womelsdorf, T., Johnston, K., Vinck, M., & Everling, S. (2010). Theta-activity in anterior cingulate cortex predicts task rules and their adjustments following errors. *Proceedings of the National Academy of Sciences*, *107*(11), 5248–5253.