Title: The interplay between experimental heat pain and non-invasive stimulation of the medial prefrontal cortex on reinforcement learning with manipulated outcome controllability

Running title: Effects of pain on reinforcement learning

Samy Babiker[1], Federica Luzzi[2,3], Matthias Mittner[1], Gábor Csifcsák[1]*

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

[2]Department of Neuroscience and Rehabilitation, University of Ferrara, Ferrara, Italy

[3]Center for Translational Neurophysiology of Speech and Communication, Italian Institute of Technology, Ferrara, Italy

*Corresponding author: Gábor Csifcsák (gabor.csifcsak@uit.no), Department of Psychology, Faculty of Health Sciences, UiT The Arctic University of Norway, Huginbakken 32, 9037 Tromsø, Norway

Acknowledgements

We wish to thank all our participants for their time and effort. Additionally, we would like to express our gratitude to Steffen Rygg Aasen and Ragnhild Nicolaisen Drevland for helping in the blinding procedure.

Abstract

Pain negatively affects several cognitive abilities, but knowledge about its effect on reinforcement learning (RL) is limited. During RL, instrumental choices can be influenced by heuristic tendencies to approach rewards or inhibit actions when facing potentially aversive events, introducing "Pavlovian bias" in behavior. Recent studies suggest that compromised outcome controllability enhances Pavlovian bias, a phenomenon that may be linked to suboptimal decision-making in learned helplessness (LH). Since LH is common in chronic pain syndromes, this study sought to establish a link between experimental heat pain (EHP), uncontrollable reward/loss and RL performance in healthy adults. In addition, we investigated if intermittent theta burst stimulation (iTBS) targeting the medial prefrontal/dorsal anterior cingulate cortex (mPFC/dACC) alleviates the deleterious effects of EHP on choice behavior. In a pre-registered, 2x2 between-group, double-blind study (N = 100), healthy adult participants underwent 3 blocks of an orthogonalized Go/NoGo task with two interleaved bouts of active or sham iTBS, and either EHP or warm skin stimulation combined with compromised outcome controllability during the task. Although EHP did not impact overall performance, it invigorated actions for rewards, reflecting enhanced Pavlovian bias. While two bouts of iTBS attenuated Pavlovian tendencies, this effect was counteracted by EHP, indicating antagonistic effects of pain and iTBS-modulated mPFC activity on Pavlovian-instrumental interactions. Surprisingly, EHP and iTBS exerted largely similar effects on other latent parameters of RL (go-bias, learning rate, exploration) in a manner that resembled LH. These findings shed light on the role of experimental pain and mPFC/dACC activity in LH-like choice behavior.

Key words: reinforcement learning, learned helplessness, Pavlovian bias, pain, repetitive transcranial magnetic stimulation, medial prefrontal cortex

Introduction

Several studies have provided evidence for the bidirectional relationship between pain and cognition. For example, when attention is given to a cognitive task, the sensation of the pain becomes less intense (Bushnell *et al.*, 1999; Peyron *et al.*, 1999). The analgesic effects of cognitively demanding tasks might be a result of limited executive resources resulting in an attentional distraction effect, which limits pain processing (Seminowicz & Davis, 2007; Phelps *et al.*, 2021). However, nociceptive stimuli are rarely completely unattended, and therefore, pain also draws on cognitive resources, leading to worse performance in tasks requiring executive control (Seminowicz & Davis, 2007; Buhle & Wager, 2010).

With respect to decision-making, experimental heat pain in healthy adults impacts choices, with more impulsive and higher risk-seeking behavior for gains, but not for losses (Koppel *et al.*, 2017). Conversely, patients with chronic pain show impaired value-based decisions, characterized by altered sensitivity to cues predicting reward and punishment (Apkarian *et al.*, 2004; Hess *et al.*, 2014; Attridge *et al.*, 2019). Along with these findings, exaggerated avoidance behavior to potentially nociceptive events (Harvie *et al.*, 2017), as well as less selective classical (Pavlovian) conditioning in chronic pain syndromes (Vlaeyen & Linton, 2000) point towards maladaptive reinforcement learning (RL), potentially driven by patients' low perceived control over painful events (Seville & Robinson, 2000; Seymour, 2019). In turn, a persisting feeling of compromised environmental controllability in chronic pain can generalize to areas of life that are not directly affected by pain itself, resulting in maladaptive coping associated with learned helplessness (Samwel *et al.*, 2006; Edwards *et al.*, 2011; Phelps *et al.*, 2021). Despite these clinical insights, it is less clear if the pairing of experimental pain with diminished environmental controllability influences RL performance in healthy adults.

Decisions in RL tasks largely depend on the interaction between instrumental and Pavlovian valuation (Rangel *et al.*, 2008). While instrumental learning encompasses the formation of stimulus-response and response-outcome associations established via trial-and-error, Pavlovian stimulus-outcome mappings are constrained by stimulus valence, and once established, they invigorate actions towards potential rewards, and inhibit behavior in the face of aversive events independent of action taken ("Pavlovian bias"; Guitart-Masip et al., 2014; Rangel et al., 2008). Such motivational drives are largely automatic (Rangel *et al.*, 2008) and therefore, they do not rely on cognitive effort (Boureau *et al.*, 2015).

However, in Pavlovian-instrumental conflict situations, cognitive control must be recruited to suppress Pavlovian bias over instrumental valuation to optimize performance (Guitart-Masip *et al.*, 2014). For instance, when a person is on a diet, the Pavlovian system would still advocate to approach the immediate reward of chocolate consumption ("reward-based invigoration"), and therefore, cognitive control would be needed to resist this motivational urge in order to maintain the goal of losing weight. Conversely, during vaccination the Pavlovian system prompts the avoidance of pain ("punishment-based suppression"), whereas the instrumental system facilitates approaching the syringe to achieve immunization, a process that also requires cognitive control. Importantly, recruitment of cognitive control during Pavlovian-instrumental conflict is associated with midline frontal theta-band activity recorded via electroencephalography, arising from the medial prefrontal/dorsal anterior cingulate cortex (mPFC/dACC) (Cavanagh *et al.*, 2013; Cavanagh & Frank, 2014; Guitart-Masip *et al.*, 2014; Swart *et al.*, 2018). Thus, this brain region might be directly involved in suppressing maladaptive Pavlovian bias in conflict situations, and advancing choices governed by the instrumental system.

In the past years, both theoretical and empirical work have implicated the role of outcome controllability in shaping decision strategies underlying RL (Ly *et al.*, 2019; Wang

et al., 2021). In healthy adults, compromised control over rewards or losses during RL has been linked to inefficient cognitive control and stronger Pavlovian bias, as well as alterations in latent parameters of RL such as learning rate, exploration tendencies and a predisposition for response initiation (Dorfman & Gershman, 2019; Csifcsák *et al.*, 2020; Csifcsák, Bjørkøy, *et al.*, 2021). Some of these effects are also present following the manipulation of controllability (i.e., when participants regained control), indicating carry-over effects (Csifcsák *et al.*, 2020; Csifcsák, Bjørkøy, *et al.*, 2021). Interestingly, behavioral passivity and reduced exploration of choice alternatives that arise due to compromised environmental control and can also generalize to novel situations, resemble behavior underlying learned helplessness (LH) in conditions such as depression, anxiety and chronic pain (Samwel *et al.*, 2006; Edwards *et al.*, 2011; Maier & Seligman, 2016; Phelps *et al.*, 2021). Chronic pain is particularly relevant for our current study, since the relationship between LH and clinical conditions with chronic and uncontrollable pain has been vastly documented (Seville & Robinson, 2000; Apkarian *et al.*, 2004; Samwel *et al.*, 2006; Harvie *et al.*, 2017), while the link between pain, outcome controllability and RL performance remains to be established in healthy adults.

A recent re-evaluative ("learned controllability") account of LH argued that passivity and the avoidance of threatening situations are not learned, but instead, they reflect innate and automatic response tendencies that dominate behavior when agents face reduced controllability over harmful events (Maier & Seligman, 2016). In contrast, experiencing adequate levels of outcome controllability increases activity in the ventral part of the mPFC, which via top-down suppression of LH-like behavior, fosters active exploration and adaptive coping (Maier & Seligman, 2016). This reconceptualization of LH bears similarities with the arbitration between Pavlovian vs. instrumental valuation during RL, whereby reduced reward/loss controllability is associated with stronger reliance on automatic (Pavlovian)

response patterns, probably driven by suboptimal cognitive control and mPFC/dACC activity (Csifcsák *et al.*, 2020; Csifcsák, Bjørkøy, *et al.*, 2021).

So far, two non-invasive brain stimulation (NIBS) studies attempted to counteract the unfavorable effects of low outcome controllability on RL performance via stimulating the mPFC in healthy adults (Csifcsák, Bjørkøy, *et al.*, 2021; Sedlinská *et al.*, 2023). Both studies utilized high-definition transcranial direct current stimulation (HD-tDCS), and expected that relative to a sham condition, real stimulation would improve task performance via reducing suboptimal Pavlovian bias during Pavlovian-instrumental conflict. While Csifcsák and colleagues (2021) found effects in the expected direction, with improved task performance following HD-tDCS, the study by Sedlinská et al. (2023) reported that mPFC stimulation increased (rather than reduced) Pavlovian bias, without changing overall response accuracy. These conflicting results can be attributed to substantial differences in task parameters and their effects on the participants' ratings of perceived controllability, but it is also noteworthy that the effects of HD-tDCS are less focal and relatively weak when compared to other NIBS techniques (Priori *et al.*, 2009; Turi *et al.*, 2021). Repetitive transcranial magnetic stimulation (rTMS) is a more potent tool to modulate neural excitability in circumscribed cortical areas (Ridding & Rothwell, 2007), with the intermittent theta burst stimulation (iTBS) pattern leading to relatively prolonged and robust excitability enhancements in the targeted area (Suppa *et al.*, 2016). Therefore iTBS is a promising method to investigate the causal relationship between mPFC/dACC activity and RL performance during and following painful stimulation and compromised outcome controllability.

In the current randomized, sham-controlled study, we adopted a 2 x 2 between-subject design to investigate whether experimental heat pain (EHP) and/or iTBS above the mPFC/dACC influenced response accuracy and latent parameters of RL in healthy adults. We postulated that EHP during periods of reduced outcome controllability would lead to impaired

task performance both during and following the intervention, driven by enhanced Pavlovian bias, behavioral passivity and reduced exploration tendencies. Such a result would provide empirical support for the link between pain, outcome controllability and stronger reliance on Pavlovian valuation during RL, and establish a behavioral model for LH-like behavior in healthy adults.

Furthermore, we tested whether stimulating the mPFC/dACC via iTBS would optimize RL performance by alleviating the detrimental effects of low controllability both when paired with EHP and, as a control condition, when participants receive non-painful warm stimulation. The mPFC/dACC has been associated with estimations of environmental controllability and volatility (Behrens *et al.*, 2007; Diener *et al.*, 2010; Gershman *et al.*, 2021; Ligneul *et al.*, 2022), and a meta-analysis also pointed at the dACC as a hub for integrating nociceptive information, negative affect and cognitive control (Shackman *et al.*, 2011). Therefore, we anticipated that targeting this cortical area with iTBS would exert favorable effects on choice behavior in our experimental setup.

Materials and Methods

Study design

The study protocol complied with the Declaration of Helsinki and was approved by the Regional Committees for Medical and Health Research Ethics of Northern Norway (REK Nord; Case number: 318704). The purpose, hypotheses, methods and analysis plan of the study were pre-registered at Open Science Framework (https://osf.io/s6e8r) and the data is publicly available (https://osf.io/r74uk/). The study adopted a between-subject, randomized, sham-controlled design where 100 participants were assigned to one of the 4 experimental groups: (1) Sham iTBS + warm (non-painful) stimulation ("Sham-Warm"), (2) sham iTBS + EHP ("Sham-Pain"), (3) active iTBS + warm stimulation ("Active-Warm"), (4) active iTBS + EHP ("Active-Pain") (Figure 1). Based on *a priori* power analysis (G*Power 3.1.9.7) (Faul *et*

al., 2007), a total sample size of $N = 100$ enables detecting a small-to-medium effect size (Cohen's f = 0.155) with 80% probability, and an effect of Cohen's f = 0.175 with 90% probability (repeated-measures analysis of variance, within-between factor interaction, alpha = .05). Therefore, we decided to collect data from 25 participants in each of the 4 groups.

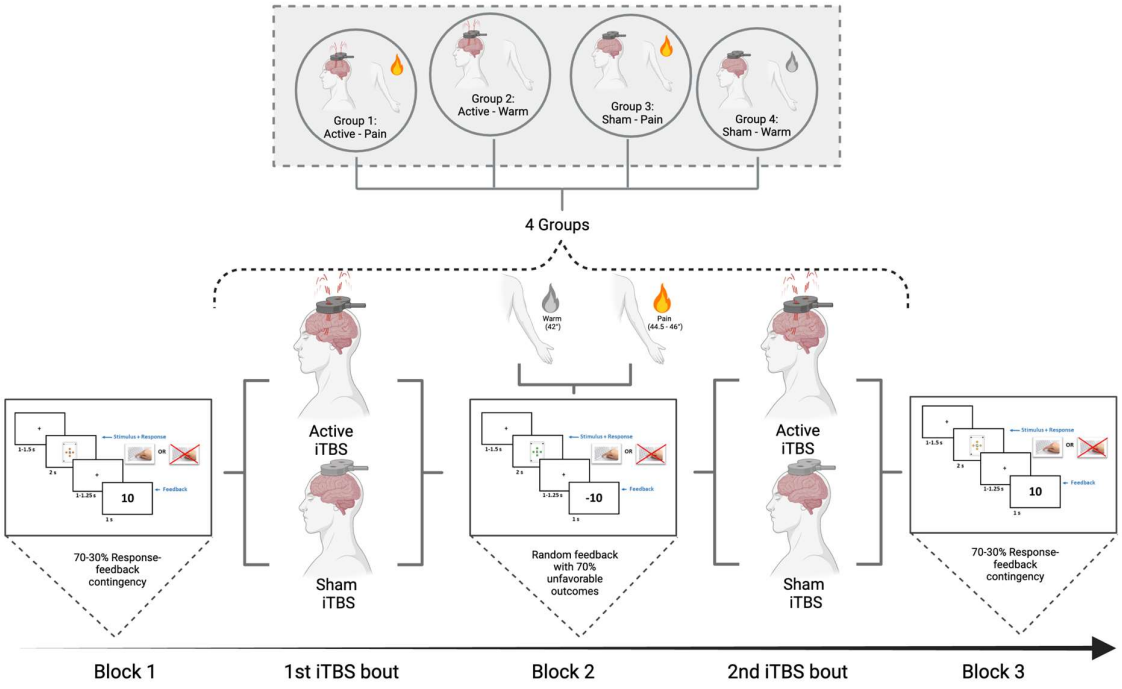


Figure 1. Overview of the experimental paradigm. Participants were either placed in the "Active-Pain", "Active-Warm", "Sham-Pain" or "Sham-Warm" groups, depending on whether they received active or sham iTBS before blocks 2 and 3, and whether they were exposed to warm or painful stimulation during block 2. Blocks 1 and 3 were identical across groups, with maintained outcome controllability (70-30% response-feedback contingency). In block 2, however, all participants underwent the "Control-Reward-Loss Manipulation" (CRLM) protocol with the aim to induce shifts in RL performance resembling LH, an effect that was expected to carry-over to the subsequent block.

Participants

One hundred healthy adults participated in the study (M_{age} = 23.9 years, SD_{age} = 5.7; 47 female; 88 right-handed). The four experimental groups did not differ significantly in terms of age ($F(3,96) = 0.527$, $p = .655$), gender ($\chi^2(3) = 6.86$, $p = .076$) or handedness ($\chi^2(3) = 3.78$, $p = .285$). Participants were mainly recruited from UiT The Arctic University of Norway (campus Tromsø), via personal contact, flyers and social media. According to our

inclusion criteria, participants had to be between 18 and 50 years of age, fluent in either English or Norwegian, with good or corrected vision, and without any current or previous psychiatric/neurological disorder (e.g., depression, bipolar disorder, epilepsy, ADHD, etc.) or chronic pain syndromes, no pregnancy, no metal implants in the head, and no electronic devices (e.g., pacemaker) in the body. They were also asked to arrive to the lab well-rested, with piercings or other metal jewelry removed from the body, and to avoid taking psychotropic drugs (with the exception of caffeine and nicotine) or painkillers on the day of data collection. In order to increase their motivation throughout the experiment, participants were told that they would receive a gift card worth 400 NOK (approximately 40 Euros) at the local shopping center for participating, and that high levels of performance in the task would be rewarded with a bonus of 100 NOK (approximately 10 Euros). However, all participants completing the experiment received gift cards worth of 500 NOK, irrespective of task performance. All participants complied with the experimental protocol, and they all fulfilled our pre-registered inclusion criterion of producing at least 1 Go and 1 NoGo response in all 3 blocks of the RL task.

Overview of the procedure

Every participant signed the informed consent. Participants completed five questionnaires online, before arriving to the lab: the Positive and Negative Affect Schedule (PANAS) (Watson *et al.*, 1988), Beck's Hopelessness Scale (BHS) (Beck *et al.*, 1974), the Behavioral Inhibition System/Behavioral Activation System (BIS/BAS) Scale (Carver & White, 1994), the Locus of Control Behavior (LCB) Scale (Craig *et al.*, 1984), and the Need for Cognition scale (Cacioppo *et al.*, 1984). These questionnaires assessed participants' mood states (PANAS) and personality traits related to either exhibiting Pavlovian-type response tendencies in everyday life (BIS/BAS), their tendency to develop hopelessness under

challenging situations (BHS), the attribution of controllability to internal or external sources (LCB) or the predisposition to implement cognitive effort in demanding tasks (NFC).

Next, we prepared for iTBS by calibrating the neuronavigation device to the participants' heads and determined individual motor thresholds for iTBS. After motor threshold had been estimated, participant and experimenter left the room, and the other experimenter blinded the iTBS stimulation. Following this, we measured pain tolerance levels to determine the intensity of painful stimulation during the RL task. Next, participants were given instructions about the card game, did a short practice block, and completed a mini-quiz to ensure that all main aspects of the task were understood.

After completing block 1 of the task, we delivered the first bout of iTBS. Before the stimulation, participants went through a familiarization process, consisting of a short bout (1 second) at low intensity (25% of the maximum stimulator output, MSO), followed by another short familiarization bout at their actual stimulation intensity. After confirming that participants tolerated TMS, we proceeded with the iTBS (active or sham). This was followed by block 2 of the RL task, which involved warm or painful stimulation. The second bout of iTBS was delivered after block 2, and was identical to the first one (i.e., every participant received either active or sham iTBS on both occasions). Finally, participants completed block 3 of the card game.

After each block of the task, participants were asked to rate their subjective level of outcome controllability ("Please rate the degree to which you felt you could control the outcomes (points), by either picking or not picking up the cards!") and success ("Please rate how successful you think you were in collecting points!"). In addition, two questions were used to measure mean and peak levels of perceived pain in each block ("Please rate the maximum pain level you felt in this block!", "Please rate the average pain level you felt in

this block!"). These questions were answered by dragging an arrow across a visual analogue scale (VAS) ranging from 0 (absent) to 100 (maximal).

At the end of the experimental session, participants filled out a TMS-checklist about possible side-effects (Brunoni *et al.*, 2011) and were asked to guess whether they received active or sham stimulation. The same forced-choice guess was made by the experimenter to collect data on blinding efficacy for both the participant and the experimenter.

The RL task and manipulation of outcome controllability

We used a modified version of the orthogonalized Go/NoGo task previously reported by Csifcsák et al. (2020, 2021) and Sedlinská et al. (2023), available in two languages (English or Norwegian). The task was originally designed to investigate the neural correlates of Pavlovian bias over instrumental choices during RL (Guitart-Masip *et al.*, 2011). Our version was framed as a computerized card game, consisting of a short practice block and 3 main experimental blocks, each associated with a set of 4 cards (16 cards in total). Every card contained a unique character (letter from the Latin alphabet), in addition to a colored geometric shape (practice block: black circle, main blocks: yellow diamond, green square, orange circle, randomly assigned to one of the 3 blocks) that was shared between cards belonging to the same card set, and hence was used to differentiate between the experimental blocks. Only one card was shown at each trial (Figure 1), with every card being presented 20 times in total for the main blocks, resulting in 80 trials in each (lasting for 430 seconds on average). Cards during the practice block were presented only 5 times each (20 trials in total).

The goal of the game was to maximize the total points earned by finding out via trial-and-error whether it was more favorable for each card to be "picked up" by pressing the space bar (Go response), or to be left untouched (NoGo response). Participants were asked to make decisions as fast as possible, and to emit Go responses while the card was visible on the screen (for 2 seconds). Each card was followed by a short delay, after which a feedback

screen appeared, informing the participants about the number of points earned or lost at the corresponding trial (Figure 2a).

Within each card set, cards were randomly assigned to one of the four experimental conditions: “Go-to-Win”, “NoGo-to-Avoid”, “Go-to-Avoid”, and “NoGo-to-Win” (Figure 2b). For Win cards, participants’ aim was to receive a reward (10 points) instead of a neutral outcome (no-win: 0 points). For Avoid cards, participants had to aim for avoid losing (-10 points) by receiving a neutral outcome (no-loss: 0 points). Given that the Pavlovian system advocates approach toward rewards and response inhibition in the face of aversive events, the “Go-to-Win” and “NoGo-to-Avoid” cards are congruent with the Pavlovian system, whereas “Go-to-Avoid” and “NoGo-to-Win” cards induce Pavlovian-instrumental conflict. This task design enables learning Pavlovian-congruent cards more easily, whereby optimal responding for Pavlovian-conflict cards requires the employment of cognitive control to suppress Pavlovian bias (i.e., inhibiting a Go response for certain Win cards and emitting a Go response for certain Avoid cards).

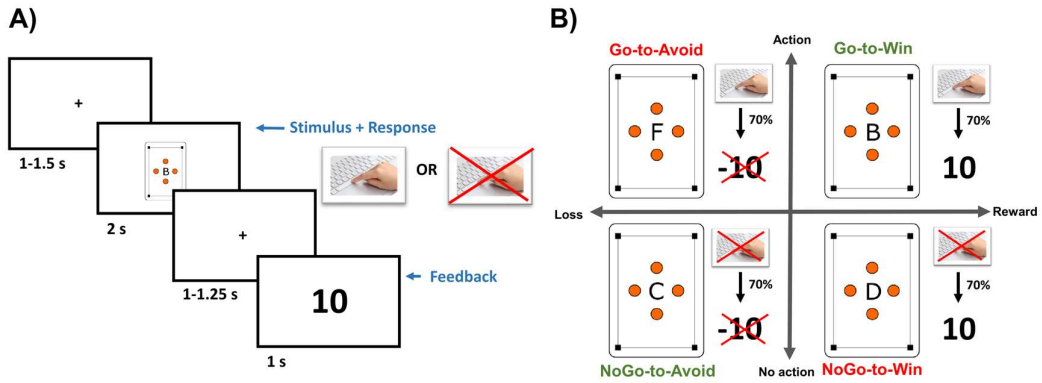


Figure 2. Overview of the RL task. A) In each trial, participants were asked to decide which action to take (Go vs. NoGo) during card presentation. Feedback (10 points, 0 point or -10 points) was shown after a short delay. **B)** The four card types of the task differ in terms of response requirement (action vs. no action; vertical axis) and outcome valence (reward vs. loss; horizontal axis). The Go-to-Win and NoGo-to-Avoid cards are congruent with predictions by the Pavlovian system (depicted in green), whereas the Go-to-Avoid and NoGo-to-Win cards induce Pavlovian conflict (depicted in red).

In blocks 1 and 3, outcomes were probabilistic with a response-feedback contingency of 70-30%. That is, correct (incorrect) responses were followed by favorable (unfavorable) outcomes in 70% of the trials, and feedback was misleading in 30% of trials. Crucially, we aimed to induce choice strategies resembling LH in block 2 by removing feedback controllability, and at the same time, fixing the frequency of favorable/unfavorable outcomes ("Control-Reward-Loss Manipulation"; CRLM). In this block, outcomes were uncontrollable by being presented randomly, irrespective of preceding instrumental (Go or NoGo) choices. While this manipulation was similar to those utilized in our previous studies (Csifcsák *et al.*, 2020; Csifcsák, Bjørkøy, *et al.*, 2021; Sedlinská *et al.*, 2023), we now also kept the rate of favorable/unfavorable outcomes for each card at 30-70%. Thus, Win cards in this block resulted in "no-win" (0 point) and "win" (10 points) outcomes in 70% and 30% of the trials, respectively. Conversely, outcomes for Avoid cards were fixed at 70% for "loss" (-10 points) and 30% for "no-loss" (0 point) trials. The design of block 2 was inspired by studies pointing at both low response-outcome contingency and low reward/high punishment prevalence as predictors of LH-like behavior in humans (Teodorescu & Erev, 2014; Ly *et al.*, 2019). Akin to animal models for LH (Maier & Seligman, 2016), our participants were not informed in advance about the CRLM in block 2, and therefore, they were expected to learn this via trial-and-error. We hoped that low/high rates of rewards/losses in this block would lead to diminished subjective controllability, which was argued to be more closely correlated with changes in choice behavior than objective controllability (Ly *et al.*, 2019). All participants were debriefed about the CRLM at the end of experiment.

Participants were not fully informed about the structure of the task. However, they were told that (1) some cards were "winning" and some were "losing", (2) winning cards were associated with favorable (10 points) or neutral (0 point) outcomes, while losing cards either had unfavorable (-10 points) or neutral (0 point) outcomes, (3) for some cards it could

be more worthwhile to be picked up (Go), while for other cards it would be beneficial to remain passive (NoGo), (4) the characteristics for each card (winning vs. losing; Go vs. NoGo) remained constant during the experimental block, (5) there was no relation between cards from different blocks so that the learning process had to start fresh at the beginning of each block, (6) the feedback was probabilistic (i.e., correct responses resulted in favorable outcomes with a high probability, but they were not guaranteed, and thus, there would also be misleading outcomes), and (7) that response times would also be measured and therefore, active Go-responses should be as quick as possible. Overall, participants were encouraged to try out both response options (Go vs. NoGo) for all cards on several trials to find out the optimal response strategy in each block.

Determining individual resting motor thresholds

For determining the resting motor threshold (RMT), participants were seated in a comfortable chair and equipped with three Ag/AgCl electrodes to record electromyography (EMG). The ground electrode was placed on the dorsum of the right hand, the reference electrode was placed on the lateral surface of the proximal interphalangeal joint of the 5th finger, while the active electrode was placed above the abductor digiti minimi muscle. Electrodes were connected to a Brain Products V-amp 8 digital DC amplifier (BrainProducts GmbH, Germany). Participants were asked to find a comfortable position in the chair, completely relax, and place their right arms on a pillow.

TMS pulses were delivered with a MAG & More PowerMAG lab 100 device, using a PMD70-pCool coil (MAG & More GmbH, München, Germany). The coil was registered with the neuronavigation system (TMS Navigator (Localite GmbH, Germany) using special markers prior to experiment start. We navigated the coil above the primary motor cortex (at MNI coordinates: x = -40, y = -20, z = 52), which we used as a starting location for hotspot hunting. The coil handle was oriented 45º from the midsagittal plane in posterior-lateral

orientation. The MSO was first set at 35%, and the coil was systematically moved within an approximately 2 cm radius to assess if there were consistent motor-evoked potentials (MEPs) with at least 50μV peak amplitude approximately 20-35ms post-stimulation. In cases we couldn't detect reliable MEPs, MSO was increased by 5% and the procedure was repeated until the scalp location corresponding to the cortical representation of the target muscle was identified. We marked the location in the neuronavigation software (Localite software, version 3.0.72) and proceeded with determining the RMT. For this purpose, we started at 35% MSO and assessed if at least 5 MEPs could be elicited out of 10 single-pulse stimulations. If not, we increased pulse intensity by 5% MSO, and evaluated MEPs again. This process was repeated until we observed at least 5 MEPs out of 10 attempts. Then, we reduced TMS intensity by 1% MSO, and repeated the MEP evaluation. We continued the gradual reduction of pulse intensity until we failed to record 5 MEPs out of 10 stimulations. The RMT value was defined as the last TMS pulse intensity we used + 1% MSO (Rossini et al., 2015).

Blinding for iTBS

We followed a double-blind TMS protocol using a randomization list that was prepared prior to the start of data collection. Sham iTBS was delivered with a PMD70-pCool-Sham coil (MAG & More GmbH, München, Germany). Out of the two researchers present in the lab, the one responsible for blinding opened an envelope corresponding to the participant's ID, which contained the predetermined stimulation condition (active or sham) and whether the participant would receive warm or painful stimulation during block 2 of the RL task. The coil was then either swapped for the sham coil or not, and registered with the neuronavigation system. The coil was also covered with a thin fabric tube manufactured from non-magnetizable material in order to hide tiny scratches and surface marks that could compromise the blinding of the experimenter. During stimulation, both the participant and the researcher wore earplugs to increase the chances of proper blinding.

Participants in the active iTBS groups guessed exactly at chance level about their stimulation protocol (guessed active: N = 25, guessed sham: N = 25), while more participants guessed correctly about receiving sham iTBS (guessed active: N = 18, guessed sham: N = 32). The difference was not significant ($\chi^2(1) = 2.00$, $p = 0.157$). Similar results were obtained for the experimenters' guesses, indicating successful double-blinding (active group: N = 31 for guessed active, N = 19 for guessed sham; sham group: N = 27 for guessed active, N = 23 for guessed sham; $\chi^2(1) = 0.66$, $p = 0.418$).

iTBS protocol

The stimulation protocol was controlled by a dedicated software that was integrated with the TMS device (Stimware, version 1.9). Stimulation of the mPFC/dACC was achieved with the iTBS protocol, which was demonstrated to increase neural excitability in the targeted area (Suppa *et al.*, 2016; Lowe *et al.*, 2018; Diederichs *et al.*, 2021; Xu *et al.*, 2024). The protocol consists of three bursts of 50Hz TMS pulses with 200ms breaks between each burst (hence, bursts are delivered at 5Hz, corresponding to the theta frequency band). Stimulation is intermittent, whereby each 2 second-long stimulation period (30 pulses) is followed by a rest period of 8 seconds, and this 10 second-long pattern is repeated for a total of 192 seconds (600 pulses). In our study, we implemented this iTBS protocol twice: Once preceding, and once immediately after block 2 of the card game (Figure 1), with a stimulation intensity corresponding to 80% of the participant's RMT. The mPFC/dACC was targeted using the neuronavigation system using MNI coordinates x = 0, y = 10, z = 46. These coordinates correspond to the portion of the dACC associated with both cognitive control and pain processing (Shackman *et al.*, 2011). The handle of the coil was oriented towards the occipital pole along the midsagittal line.

Experimental heat pain stimulation

EHP was induced via a PC-controlled system (Pathway, Medoc, Ramat Yishay, Israel). Contact heat was delivered via a 3x3 cm Aluminum thermode that was attached to the volar surface of the lower arm. The system enabled delivering heat between -10 and +50°C via a dedicated software (error margin: 0.3°C), with built-in safety measures preventing long stimulation at temperatures capable of inducing skin damage. Individual pain tolerance levels were assessed via attaching the thermode to the dominant forearm, whereas for task-related stimulation, it was attached to the non-dominant forearm (to enable responding to the card game with the dominant arm). The thermode was also placed on the participants during blocks 1 and 3, but it was not active.

Prior to the task, we determined heat pain tolerance levels in all participants. For this, we used 8 runs of gradually increasing heat stimulation ("method of limits", from 32 to 50°C, at 0.5°C/second rate) (Defrin *et al.*, 2006), and asked participants to terminate the stimulation by a button press when they felt that "the pain becomes so intensive that they want it to stop". Pain tolerance levels were calculated by averaging temperature values from runs 3-8.

Nociceptive stimulation in the Sham-Pain and Active-Pain groups was delivered during block 2 of the RL task (Figure 1) for exactly 7.5 minutes. Stimulation intensity was set at 2°C below the participants' individual tolerance levels, but never below 44°C or above 46.5°C. This was to ensure that we stimulated at the dynamic range of thermal nociceptors in the skin, but never at intensities capable of inducing skin damage (Arendt-Nielsen & Chen, 2003; Nielsen et al., 2005) In the Sham-Warm and Active-Warm groups, the temperature was fixed at 42°C, irrespective of individual pain tolerance levels. Importantly, since EHP stimulation never exceeded 46.5°C, heat intensities were within the safety range that can be applied for 7.5 minutes to induce medium-intensity pain in healthy participants (Naert *et al.*, 2008), and also complied with the safety measures of the Pathway ATS System.

Statistical analysis

In line with our pre-registered analysis plan, we tested with a series of univariate analysis of variance (ANOVA) tests whether participants in the 4 groups differed in any of our questionnaires: the PANAS subscores for positive and negative affect, the BHS, the NFC, the BIS/BAS subscores, and the LCB. If any scores showed a significant ($p < 0.05$) main effect of Group (4 levels), they were to be entered as covariates for all subsequent analyses (see below).

Our main analysis consisted of 7 repeated-measures ANOVAs (rmANOVAs) to assess how EHP and/or iTBS interfered with subjective ratings (controllability, success, mean pain, peak pain), changes in task performance (response accuracy, reaction time) and the degree of Pavlovian bias (Pavlovian Performance Index) over the 3 blocks of the RL task. Significant ($p < 0.05$) interactions were followed-up with post hoc Bonferroni-corrected pairwise comparisons (p_{Bonf}). Where necessary, violations of the assumption of sphericity were adjusted with Greenhouse-Geisser's epsilon (ε).

For details about the expected significant main effects and interactions for all analyses, we refer the reader to our pre-registration (https://osf.io/s6e8r). With respect to response accuracy, mean accuracy calculated for each participant, block and card type was entered as dependent variable into rmANOVA, with within-subject factors Block (1-3), card Valence (Win, Avoid) and card Congruence (Congruent, Conflict), and between-subject factors iTBS (active, sham) and EHP (warm, pain).

Analysis of reaction times (RTs) was restricted to correct responses (Go-trials) only. Median RTs were calculated for each participant, block and card type separately, and assessed via rmANOVA with Block and card Congruence as within-, and iTBS and EHP as between-subject factors.

Finally, the analysis of the change in Pavlovian bias across task blocks and groups was done via calculating the Pavlovian Performance Index (PPI) as the mean of two measures, Reward-Based Invigoration (RBI; the number of Go responses on Win trials/total number of Go) and Punishment-Based Suppression (PBS; the number of NoGo responses on Avoid trials/total number of NoGo). These indices represent the likelihood to initiate actions towards rewards and inhibit responses when facing potential loss, respectively (Cavanagh et al., 2013). PPI was used as dependent variable in rmANOVA, whereas Block served as within-, and iTBS and EHP as between-subject factors. In a separate analysis, we separated PPI to RBI and PBS (captured by the within-subject variable Index). The rationale behind this was to test the expected stronger effects of EHP on Pavlovian bias in the loss domain (PBS), resulting in a significant Block x EHP x Index interaction.

Computational modeling

We applied the computational model used in previous studies (Cavanagh *et al.*, 2013; Csifcsák *et al.*, 2020; Csifcsák, Bjørkøy, *et al.*, 2021; Sedlinská *et al.*, 2023) to estimate how latent parameters of RL are affected by EHP and/or iTBS.

To evaluate changes RL parameters throughout the task, we estimated them for each individual *i* and block *j* separately. Based on earlier work involving a largely similar orthogonalized Go/NoGo RL task, we modeled participants (index i) decisions at each trial with the soft-max function:

$$p\left(Go\middle|s_{t,j,i}\right) = \frac{exp\left[W_t\left(Go\middle|s_{t,j,i}\right)\beta_{j,i}^{-1}\right]}{exp\left[W_t\left(Go\middle|s_{t,j,i}\right)\beta_{j,i}^{-1}\right] + exp\left[W_t\left(NoGo\middle|s_{t,j,i}\right)\beta_{j,i}^{-1}\right]}$$

whereby the probability to pick up *p(*Go) a given card ($s_{t,j,i}$) on trial *t* was largely based on the 'weight' (W_t) of each choice alternative, but at the same time, it was inversely influenced by the exploration parameter of each participant on each block ($\beta_{j,i}$). Parameter β controls the degree to which choices are driven by learning via trial-and-error (low values), or made rather at random (high values being indicative of exploratory behavior). At each trial, W_t was

modeled as a linear combination of the value of the corresponding stimulus-action combination (Q_t), the value (V_t) of the stimulus scaled with the participant's Pavlovian parameter in that block ($\pi_{j,i}$), and the participant's go-bias parameter in the same block ($b_{j,i}$).

$$W_t(a|s_{t,j,i}) = \begin{cases} Q_t(Go|s_{t,j,i}) + b_{j,i} + \pi_{j,i}V_t(s_{t,j,i}) & if\ a = Go \\ Q_t(NoGo|s_{t,j,i}) & if\ a = NoGo \end{cases}$$

Thus, participants with larger b values were more predisposed to emit Go responses irrespective of their learning history, whereas reduced values indicate a general behavioral shift towards passivity. On the other hand, larger π values increase the probability of picking up cards associated with rewards (i.e., cards with positive V_t values), while negatively valued cards (Avoid cards) reduce the chances of actions for individuals with strong Pavlovian bias. However, in the absence of Pavlovian bias ($\pi = 0$) action weights were not directly influenced by card valuation.

Stimulus (V_t) and stimulus-action values (Q_t) were both updated trial-by-trial in a Rescorla-Wagner manner by the prediction error (formulated as the difference between the received and the expected reward) scaled by the learning rate parameter (α) of that block and individual,

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

and

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

where $r_{t,j,i}$ denotes the current reward, and learning rate α quantifies the degree to which participants rely on the prediction error to update their values. Overall, individuals were assumed to continuously update four different stimulus values (one for each card) and eight Q-values (one for each card-action coupling) during every block of the RL task.

In line with earlier work (Csifcsák *et al.*, 2020; Csifcsák, Bjørkøy, *et al.*, 2021; Sedlinská *et al.*, 2023), hierarchical Bayesian modeling using Hamiltonian Monte-Carlo algorithms (Hoffman & Gelman, 2011) were implemented in Stan (Carpenter *et al.*, 2017). We used six parallel chains with warm-up period of 1000 samples each such that 6000 samples were drawn from the converged chains. The trace plots for all variables were visually inspected for convergence, and the Gelman-Rubin diagnostic (Gelman & Rubin, 1992) was $\hat{R} \leq 1.05$ for all parameters.

Within the context of LH-like behavior, we expected that the combination of EHP and CRLM would increase Pavlovian bias (parameter π), but at the same time, reduce exploration (parameter β), learning rate (parameter α) and the general tendency to initiate actions (go-bias parameter b). Further, we speculated that active iTBS would counteract some of these effects (primarily the change in parameter π), due to the strong association between cognitive control, nociception and the mPFC/dACC (Shackman *et al.*, 2011).

To test these assumptions, effects of Block, EHP, iTBS, and their interactions on each RL parameter were included directly at the group-level in the hierarchical model. We estimated posterior densities for every model coefficient and evaluated if their 95% highest density interval (95% HDI) excluded zero. Specifically, for each coefficient we estimated the posterior mean change (MC) relative to block 1, as well as the 95% HDI and the posterior probability of the MC being either positive, $p(MC>0)$ or negative, $p(MC<0)$. Finally, we used these probabilities to calculate evidence ratios (ERs) that either favor a positive (ER^+) or a negative effect (ER^-). The ER represents the ratio of two probabilities: For ER^+, the probability of the effect being positive, divided by the probability of the effect being zero or negative, or its inverse for ER^-.

Results

Pre-registered analysis

In line with our pre-registered protocol, we first ran seven separate one-way ANOVA's to test whether participants in the four experimental groups showed similar baseline scores in our questionnaires assessing mood or personality traits related to hopelessness, behavioral tendencies, control attribution and engagement in cognitively challenging situations. Neither model indicated a significant main effect of Group ($F < 2.27$, $p > 0.086$ for all; Supplementary Table 1) and thus, none of these questionnaire scores were added as covariates in further analyses.

With regards to VAS ratings on subjective levels of outcome controllability and success, rmANOVA indicated a significant main effect of Block for both measures (control: $F(2,192) = 40.32$, $p < 0.001$, $\eta_p^2 = 0.30$, Figure 3A; success: $F(2,192) = 68.22$, $p < 0.001$, $\eta_p^2 = 0.42$, Figure 3B). Despite the significant reduction in block 2 relative to block 1 (control: $p_{Bonf} < 0.001$; success: $p_{Bonf} < 0.001$), we did not detect carry-over effects from CRLM to the last block, since ratings in block 3 were significantly higher than in the baseline block (control: $p_{Bonf} = 0.012$; success: $p_{Bonf} < 0.001$). Moreover, our analysis did not reveal any interactions of Block with either iTBS or EHP ($F < 1.66$, $p > 0.195$ for all; Supplementary Table 2).

As for our participants' ratings of mean pain after each block, descriptive statistics indicated medium vs. low levels of subjective pain intensity in block 2 for the two EHP conditions, with substantial variability (Pain groups: $M = 55.0$; $SD = 21.2$; Warm groups: $M = 15.9$; $SD = 17.2$; Figure 3C). There was a significant main effect of Block ($F(1.4,134.5) = 258.60$, $\varepsilon = 0.7$, $p < 0.001$, $\eta_p^2 = 0.73$) and EHP ($F(1,96) = 62.98$, $p < 0.001$, $\eta_p^2 = 0.40$), qualified by a Block x EHP interaction ($F(1.4,134.5) = 89.58$, $\varepsilon = 0.7$, $p < 0.001$, $\eta_p^2 = 0.48$). Post hoc comparisons revealed a significant difference between the pain vs. warm groups in

block 2 only ($p_{Bonf} < 0.001$; $p_{Bonf} > 0.47$ for the other blocks), verifying the efficacy of our pain-stimulation protocol (Figure 3C). However, iTBS did not interfere with pain perception in our sample ($F < 2.55$, $p > 0.099$ for all interactions, Supplementary Table 2). Largely similar effects were found for peak pain ratings (Supplementary Table 2). We note, however, that there was a significant main effect of iTBS for peak pain ratings ($F(1,96) = 5.89$, $p = 0.017$, $\eta_p^2 = 0.06$), with higher values for sham vs. active stimulation (Sham iTBS: $M = 18.2$, $SD = 12.4$; Active iTBS: $M = 13.4$, $SD = 10.8$), but the interaction between iTBS and Block was not significant ($F(1.5,142.5) = 0.48$, $\varepsilon = 0.7$, $p = 0.563$, $\eta_p^2 < 0.01$).

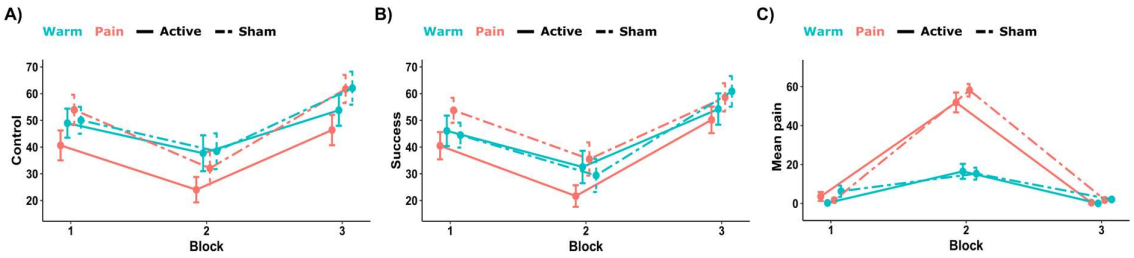


Figure 3. Differences in subjective ratings across all blocks between the four experimental groups for A) control, B) success and C) mean pain. Error bars represent standard errors.

With respect to response accuracy, we found a significant main effect of Block ($F(1.8,169.7) = 37.15$, $\varepsilon = 0.9$, $p < 0.001$, $\eta_p^2 = 0.28$), indicating reduced accuracy in block 2 relative to block 1 ($p_{Bonf} < 0.001$) and block 3 ($p_{Bonf} < 0.001$). However, we note that, due to CRLM, accuracy was expected to be around chance level in this block, since the complete random presentation of outcomes prevented learning. Accuracy was comparable between blocks 1 and 3 ($p_{Bonf} > 0.999$), being suggestive of the absence of a carry-over effect from CRLM to the last experimental block (Figure 4A). As expected, participants responded significantly better to Pavlovian-congruent vs. conflict cards (main effect of Congruence: $F(1,96) = 92.84$, $p < 0.001$, $\eta_p^2 = 0.49$), qualified by a significant Congruence x Valence interaction ($F(1,192) = 27.90$, $p < 0.001$, $\eta_p^2 = 0.23$). Follow-up post hoc comparisons indicated significant congruence (Win cards: $M_{diff} = 0.41$, $p_{Bonf} < 0.001$; Avoid cards: $M_{diff} = 0.18$, $p_{Bonf} < 0.001$) as well as valence effects (Congruent cards: $M_{diff} = 0.12$, $p_{Bonf} < 0.001$;

Conflict cards: M_{diff} = -0.12, p_{Bonf} < 0.001). That is, participants responded more accurately to congruent cards in both the gain and loss domains, and at the same time, showed improved performance for congruent-win (vs. congruent-avoid) and conflict-avoid (vs. conflict-win) cards (Supplementary Figure 1).

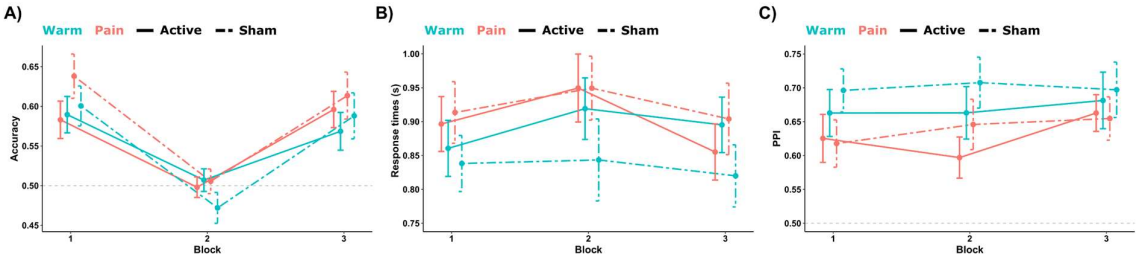


Figure 4. Differences in overall A) accuracy, B) reaction times, and C) Pavlovian bias across all blocks between the four experimental groups. Error bars represent standard errors. PPI: Pavlovian Performance Index.

Contrary to our hypotheses, we did not find significant Block x iTBS (F(1.8,169.7) = 1.53, ε = 0.88, p = 0.220, η_p^2 = 0.02), Block x EHP (F(1.8,167.7) = 0.15, ε = 0.88, p = 0.836, η_p^2 < 0.01) or Block x EHP x iTBS interactions (F(1.8,180.8) = 0.46, ε = 0.88 p = 0.609, η_p^2 < 0.01). With respect to the anticipated effects of CRLM, pain and/or iTBS on card-congruence across the three blocks, we also failed to find significant interactions between Block, Congruence and EHP or iTBS (F < 1.67, p > 0.191 for all; Supplementary Table 3).

Next, we analyzed RTs for Go-trials across experimental blocks and EHP or iTBS interventions. Here, we note that 15 participants did not produce Go responses to conflicting Go-to-Avoid cards in all 3 blocks, and therefore were excluded from the analysis. Hence, this analysis was restricted to data from 85 participants (Sham-Warm: N = 19, Active-Warm: N = 20, Sham-Pain: N = 22, Active-Pain: N = 24). We found a significant main effect of Congruence (F(1,81) = 24.09, p < 0.001, η_p^2 = 0.23) due to faster responses to Go-to-Win vs. Go-to-Avoid cards, which confirmed the presence of Pavlovian bias in the sample (Supplementary Figure 2). However, neither the main effect of Block (F(1.9,150.4) = 3.11, p = 0.051, η_p^2 = 0.04), nor any of the interactions between Block and EHP or iTBS were significant (F < 0.56, p > 0.567 for all; Figure 4B). Moreover, no effects of iTBS or EHP

were found on responding to Pavlovian-congruent vs. conflict cards over time ($F < 2.38$, $p > 0.096$ for all interactions; Supplementary Table 4).

Finally, we tested if Pavlovian bias (as reflected by the PPI measure) would be influenced by our experimental manipulations of CRLM, EHP and/or iTBS. We found no significant main effect of Block ($F(2,192) = 1.59$, $p = 0.206$, $\eta_p^2 = 0.02$, Figure 4C) or interactions between Block and EHP or iTBS ($F < 0.125$, $p > 0.288$ for all interactions). When repeating the analysis by entering the two subscores of PPI separately, a significant main effect of Index was observed ($F(1,192) = 20.67$, $p < 0.001$, $\eta_p^2 = 0.177$), pointing at stronger Pavlovian bias in the loss vs. gain domain (i.e., PBS > RBI). While the Index x iTBS interaction was also significant ($F(1,96) = 4.02$, $p = 0.048$, $\eta_p^2 = 0.04$), we refrained from interpreting it since the crucial 3-way interaction between Block, Index and iTBS did not reach significance ($F(1.8,175.8) = 1.04$, $p = 0.350$, $\eta_p^2 = 0.01$). No other main effects or interactions were detected ($F < 2.81$, $p > 0.096$ for all; Supplementary Table 5).

Exploratory analysis

In our pre-registered analysis, we noticed that some of the baseline measurements from block 1 showed substantial variations between the four groups (Figure 4). Therefore, we normalized data from blocks 2 and 3 by calculating differences relative to the initial block. The resulting "change variables" (block 2 – block 1; block 3 – block 1) for accuracy, RT and PPI were then entered into rmANOVAs to assess the potential effect of EHP and/or iTBS on RL performance, while controlling for random fluctuations across the 4 experimental groups in the baseline block. Even though this analysis did not yield significant results for response accuracy or PPI (Supplementary Tables 6-7), RT data revealed a significant main effect of Block ($F(1,81) = 7.83$, $p = .006$, $\eta_p^2 = 0.09$), as well as a significant Block x Congruence x EHP interaction ($F(1,81) = 4.33$, $p = .041$, $\eta_p^2 = 0.05$; Figure 5). This effect reflected a change in Pavlovian bias following EHP, with speeding-up of responses in the pain-groups for

Pavlovian-congruent cards in block 3 relative to block 2 (p_{Bonf} = .006), and an indication for faster RTs following pain relative to warm stimulation in block 3 for Pavlovian congruent cards, an effect that did not survive Bonferroni correction ($p_{uncorrected}$ = .040). Other pairwise comparisons were not significant ($p_{uncorrected}$ > .105). Details of the complete analysis is presented in Supplementary Table 8.

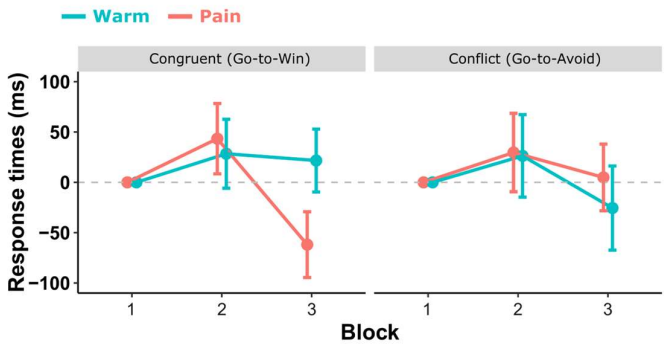


Figure 5. Effect of experimental heat pain on changes in reaction times relative to block 1, plotted separately for Pavlovian-congruent (left) and conflict trials (right). Error bars represent standard errors.

Computational modeling

We applied a computational model to our data to estimate if latent parameters of RL were influenced by CRLM, EHP and/or iTBS. This approach provided a more nuanced view on the effects of our interventions than that obtained via our pre-registered analysis of response accuracy, RT and PPI, since parameters of Pavlovian bias (π), go-bias (b), learning rate (α) and exploration/temperature (β) were estimated by taking into account the global learning process on a trial-by-trial level. However, we note that this analysis was not pre-registered, and hence results presented herein should be regarded as exploratory. Below we provide an overview of the key findings, whereas the coefficient estimates (MC values) for each effect are presented in Supplementary Table 9, along with the corresponding posterior probabilities ($p_{(MC>0)}$ or $p_{(MC<0)}$) and evidence ratios (ER^+ or ER^-).

With respect to the Pavlovian parameter π (Figure 6A), we did not observe changes from block 1 to block 2 or to block 3, indicating that CRLM was not interfering with the

computational estimate of Pavlovian bias in our task when combined with sham iTBS and warm stimulation ($p_{(MC>0)} < .926$, $ER^+ < 12.37$). In line with results from the PPI analysis, EHP was also ineffective in influencing Pavlovian response tendencies ($p_{(MC<0)} < .867$, $ER^- < 6.50$). While one bout of iTBS did not modulate this parameter during block 2 (MC = -0.064 [-0.310, 0.185], $p_{(MC<0)} = .695$, $ER^- = 2.28$), two iTBS bouts led to reduced Pavlovian bias in the last block (MC = -0.274 [-0.517, -0.017], $p_{(MC<0)} = .983$, $ER^- = 57.82$). Importantly, this effect was only present in participants previously receiving warm stimulation, but not in those being exposed to pain, as indicated by evidence for the 3-way Block 3 x EHP x iTBS interaction (MC = 0.365 [0.039, 0.698], $p_{(MC>0)} = .984$, $ER^+ = 59.61$).

The go-bias (parameter *b*) was increased in the Sham-Warm group during CRLM in block 2 (MC = 0.839 [0.479, 1.196], $p_{(MC>0)} > .999$, $ER^+ = \infty$), but it did not carry over to block 3 (MC = -0.041 [-0.362, 0.263], $p_{(MC<0)} = .591$, $ER^- = 1.44$; Figure 6B). Conversely, participants in the pain-groups showed a reduced tendency to emit an active response both in block 2 (MC = -1.102 [-1.588, -0.598], $p_{(MC<0)} > .999$, $ER^- = 2999.00$) and block 3 (MC = -0.540 [-0.985, -0.085], $p_{(MC<0)} = .989$, $ER^- = 88.66$). Surprisingly, iTBS exerted a similar effect, reducing go-bias both after one and two bouts (block 2: MC = -0.791 [-1.332, -0.229], $p_{(MC<0)} = .995$, $ER^- = 205.90$; block 3: MC = -0.495 [-0.950, -0.018], $p_{(MC<0)} = .981$, $ER^- = 51.17$). When combining EHP with iTBS, we observed a non-additive (ceiling) effect in block 2, as evidenced by a positive coefficient of a similar magnitude than those observed for either intervention alone (MC = 0.995 [0.177, 1.712], $p_{(MC>0)} = .995$, $ER^+ = 180.82$). In block 3, however, the positive coefficient for the 3-way interaction was estimated to be higher than those estimated for either EHP or iTBS only, pointing at antagonism between 2 bouts of mPFC/dACC stimulation and post-pain exposure (MC = 1.002 [0.276, 1.756], $p_{(MC>0)} = .997$, $ER^+ = 249.00$).

Relative to the baseline block, learning rate (parameter α) was increased both during and after CRLM in the Sham-Warm group (block 2: MC = 0.531 [0.358, 0.728], $p_{(MC>0)}$ > .999, $ER^+ = \infty$; block 3: MC = 0.652 [0.439, 0.849], $p_{(MC>0)}$ > .999, $ER^+ = \infty$; Figure 6C). These effects were counteracted by painful stimulation in both blocks (block 2: MC = -0.868 [-2.000, -0.547], $p_{(MC<0)}$ > .999, $ER^- = \infty$; block 3: MC = -0.359 [-0.659, -0.040], $p_{(MC<0)}$ = .985, $ER^- = 67.18$), and a similar modulation was detected for iTBS in block 2 (MC = -0.638 [-0.935, -0.316], $p_{(MC<0)}$ > .999, $ER^- = \infty$), but not in block 3 (MC = -0.141 [-0.579, 0.277], $p_{(MC<0)}$ = .728, $ER^- = 2.68$). Finally, when EHP and iTBS was applied in combination, learning rate was of similar magnitude to EHP alone in block 2, being indicative of a ceiling effect (MC = 0.578 [0.054, 1.138], $p_{(MC>0)}$ = .981, $ER^+ = 52.57$). In contrast, no interaction between EHP and iTBS was found in block 3 (MC = 0.002 [-0.532, 0.552], $p_{(MC>0)}$ = .511, $ER^+ = 1.05$).

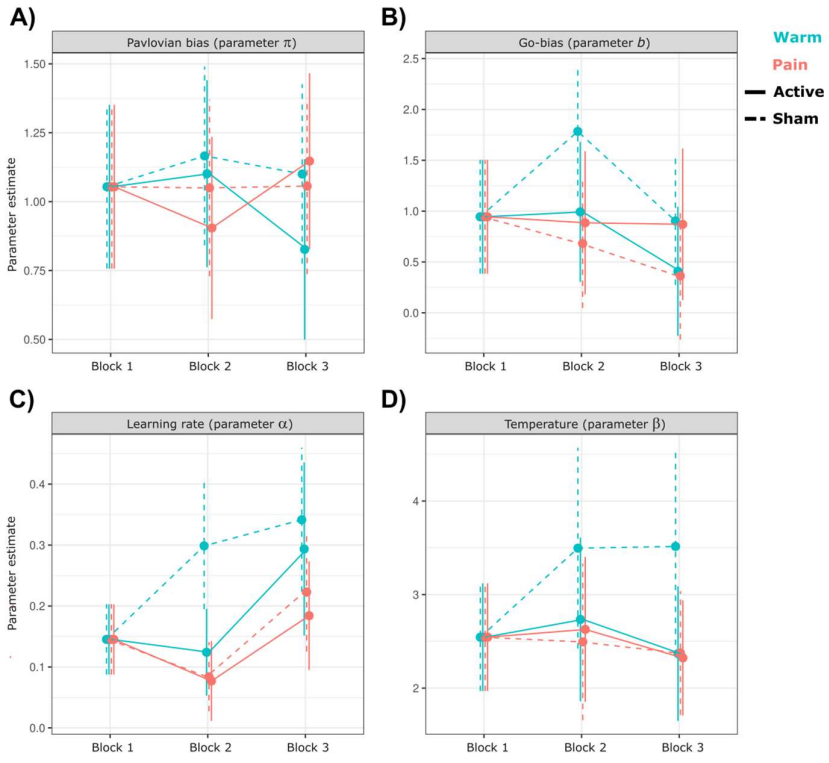


Figure 6. Changes in latent parameters of reinforcement learning across the 3 experimental blocks with respect to EHP, iTBS or both. (A) Pavlovian bias, (B) go-bias, (C) learning rate (D) temperature/exploration. Error bars represent the 95% highest density intervals.

Our participants' tendency to deviate from their reinforcement history and make exploratory choices (parameter β) increased both in block 2 (MC = 0.313 [0.077, 0.547], $p_{(MC>0)}$ = .995, ER^+ = 213.29) and block 3 (MC = 0.318 [0.092, 0.529], $p_{(MC>0)}$ = .998, ER^+ = 427.57; Figure 6D) in the Sham-Warm group. Similarly to its influence on learning rates, pain also suppressed exploratory behavior (block 2: MC = -0.340 [-0.685, -0.010], $p_{(MC<0)}$ = .974, ER^- = 31.71; block 3: MC = -0.390 [-0.682, -0.099], $p_{(MC<0)}$ = .996, ER^- = 239.00). Even though we observed effects in the same direction after two bouts of iTBS (block 3: MC = -0.395 [-0.684, -0.076], $p_{(MC<0)}$ = .994, ER^- = 170.43), one bout did not substantially influence this parameter (block 2: MC = -0.247 [-0.582, 0.081], $p_{(MC<0)}$ = .926, ER^- = 12.45). EHP and iTBS did not interact in block 2 (MC = 0.302 [-0.151, 0.720], $p_{(MC>0)}$ = .905, ER^+ = 9.47), but there was some indication for a ceiling effect in block 3 (MC = 0.373 [-0.015, 0.769], $p_{(MC>0)}$ = .968, ER^+ = 30.25), despite the 95% HDI for the coefficient estimate included zero.

Discussion

This study investigated how uncontrollable heat pain and non-invasive stimulation of the mPFC/dACC influenced RL performance during and following manipulated outcome controllability and reward/loss frequency. We anticipated that pain combined with CRLM would trigger maladaptive learning and choice strategies resembling LH, whilst iTBS was expected to counteract this effect and optimize performance. Such finding would pinpoint top-down cognitive control arising from the mPFC/dACC as the mechanism for shaping Pavlovian-instrumental interactions during decision-making in the presence of uncontrollable stressors, and pave the way towards establishing a laboratory-based model for inducing LH-like behavior in healthy humans.

With CRLM, a novel protocol for shifting choice strategies towards LH was introduced, consisting of the withdrawal of outcome controllability coupled with low reward/high loss frequency in block 2 of the RL task. The effect of this intervention in itself

was not the focus of the study, since all four groups were exposed to CRLM, with no control condition. Therefore, we refrain from the detailed discussion of the related results (i.e., the main effect of Block in the statistical models), but our insights can be found in the Supplementary Discussion.

As for the main aims of the study, we found that EHP triggered subtle LH-like changes in RL, without influencing overall task performance. In addition, EHP counteracted the effect of mPFC/dACC stimulation on Pavlovian bias, indicating an antagonism between nociception and cognitive control in modulating Pavlovian bias during RL. Finally, we present evidence that effects of EHP on RL were associated with activity in the mPFC/dACC, pointing at the role of this cortical area in developing LH-like behavior in humans.

Experimental pain triggers subtle LH-like changes in RL

In the framework of RL, pain is commonly viewed as a signal that guides behavior towards minimizing current and future harm (Seymour, 2019). Nociceptive stimuli trigger the cessation of ongoing actions, typically followed by active escape/avoidance behavior, whereby the initial freezing phase is governed by the Pavlovian system, and subsequent actions are under instrumental control (LeDoux & Daw, 2018; Seymour, 2019; Vlaeyen & Crombez, 2020). If pain persists regardless of one's action, it is believed to induce LH-like behavior in the form of passivity, anxiety and compromised coping (Samwel *et al.*, 2006; Maier & Seligman, 2016).

Perceived low controllability over the environment is a key element of states associated with LH (Maier & Seligman, 2016; Ly *et al.*, 2019). Given that EHP was uncontrollable in our study, we assumed that it would sensitize participants' to the effect of CRLM in block 2, and possibly, manifest in lower subjective ratings of controllability in block 3 as a carry-over effect. However, ratings were comparable across experimental groups throughout the task, which is in contrast with a finding that prior exposure to uncontrollable

painful stimuli is associated with suboptimal controllability estimations in an "explore-and-predict" RL task (Ligneul *et al.*, 2022). It is possible that neither our (rather crude) self-report measure, nor our RL task may have been sufficiently sensitive to detect subtle shifts in evaluations of task controllability. Moreover, participants in the current task were informed about EHP being both task-irrelevant and uncontrollable throughout block 2, while in the "stress induction phase" by Ligneul et al. (2022), painful stimuli were delivered as response outcomes, rendering them task-relevant, and therefore, probably also more potent in interfering with subjective estimations of control. In our study, we chose to apply EHP as a task-irrelevant, uncontrollable contextual factor of constant intensity to mimic tonic pain in chronic pain syndromes and hence to investigate its possible effects on choice behavior.

Surprisingly, EHP did not influence response accuracy or RTs in our pre-registered analyses. Here, we speculate that our protocol may not have induced sufficiently strong and/or long-lasting pain to interfere with attentional/executive processes required for optimal task-focus, feedback learning and choice deliberation in the final task-block. Notably, a recent study found that cognitive distraction at any load level reduced pain scores at lower pain intensities in healthy participants, while moderate to severe pain impaired task performance at high cognitive load (Lier *et al.*, 2022; but see also: Seminowicz & Davis, 2007). Given that pain ratings for block 2 indicated medium pain intensity, and our RL task did not rely heavily on goal-directed/executive resources (Rangel *et al.*, 2008; Guitart-Masip *et al.*, 2014), it is feasible that the two-way interaction between nociception and cognition did not reach the intensity/load threshold at which EHP could be detrimental for response accuracy with our experimental paradigm.

Despite the null-effect of EHP on our pre-registered measure of Pavlovian bias (PPI), exploratory analysis revealed action invigoration following painful vs. warm stimulation for cards predicting rewards. This result provides some support for our *a priori* assumption that

uncontrollable pain enhances Pavlovian bias over instrumental choices even in cases when pain is task-irrelevant, being suggestive of the generalization of pain-associated Pavlovian values to conceptually similar situations (Seymour, 2019). However, it remains puzzling why EHP did not influence PPI or parameter π from the computational modeling, both being derived from participants' choice preferences instead of response speed. Accuracy- and RT-based measures of Pavlovian bias were shown to be differentially sensitive to pharmacological manipulation (Swart *et al.*, 2017), and hence they probably reflect distinct (albeit conceptually overlapping) phenomena. In our data, RT proved to be a more sensitive measure in this regard, manifesting in more pronounced action vigor for reward- vs. loss-associated cues following exposure to EHP. This finding also aligns well with pain-related enhancement of sensitivity to gains vs. losses both in healthy participants and patients with chronic pain (Gandhi *et al.*, 2013; Hess *et al.*, 2014), which may be related to motivational aspects of pain towards seeking pain-relief (Navratilova *et al.*, 2013).

While the main aim of this study was to assess whether EHP enhanced Pavlovian bias over instrumental valuation, and therefore, could be linked to LH-like shifts in RL strategies, the relationship between Pavlovian valuation and LH is still debated (Sedlinská *et al.*, 2023). However, perturbations of other latent RL parameters have been more directly associated with LH, such as reductions in go-bias (reflecting freezing/behavioral passivity; Maier & Seligman, 2016) and weaker exploration tendency (Teodorescu & Erev, 2014; Chen *et al.*, 2015). Crucially, estimates for both measures were substantially reduced in our study, not only during EHP exposure, but also in block 3. This points at a carry-over effect, which is another key feature of LH. Finally, we also found lower learning rates during and following EHP, indicating weaker updating of stimulus/action values during RL in these participants. Previously, attenuated learning rates were linked to compromised behavioral flexibility in healthy adults exposed to uncontrollable thermal pain (Raio *et al.*, 2017), as well as to the

severity of anhedonia symptoms (Chase *et al.*, 2010), a state commonly associated with LH (MacAulay *et al.*, 2014).

Based on the above, we conclude that EHP triggered subtle shifts in RL strategies in a manner that resembles choice behavior in LH, and that LH may be linked to stronger Pavlovian bias over instrumental valuation, as reflected by valence-dependent action invigoration. Overall, these results underline the utility of our experimental protocol for inducing pain-associated LH-like states in healthy adults. Crucially, these effects did not affect overall performance, and thus, they cannot be regarded as maladaptive. Future work could build upon our findings and aim at improving the current protocol, and to investigate whether it leads to more robust changes in RL performance in patients with chronic pain.

Experimental pain antagonizes the effect of mPFC/dACC stimulation on Pavlovian bias

Given the role of the mPFC/dACC in cognitive control, pain perception and estimations of environmental controllability (Shackman *et al.*, 2011; Cavanagh *et al.*, 2013; Ligneul *et al.*, 2022), we anticipated that active stimulation targeting this area would reduce Pavlovian bias in our task, with accumulating effects from block 2 (one bout of iTBS) to block 3 (two bouts). Moreover, we predicted an antagonism between EHP and iTBS in modulating Pavlovian bias, a finding that would lend support to the learned controllability account of LH (Maier & Seligman, 2016), which posits that exaggerated automatic (Pavlovian) response tendencies in LH can be controlled via top-down inhibition arising from the mPFC.

Effects of different rTMS protocols on RL performance are mixed (Csifcsák, Forstmann, *et al.*, 2021), and even within its rather well-documented impact on executive functioning and working memory (Lowe *et al.*, 2018; Wu *et al.*, 2021; Xu *et al.*, 2024), not all rTMS studies found beneficial changes in task performance among healthy adults (Chung *et al.*, 2018; Lowe *et al.*, 2018). Our results confirm that iTBS may induce only subtle shifts in

behavior without altering overall accuracy or response times, resulting in attenuated Pavlovian bias (parameter π) following two bouts of stimulation. Interestingly, no effect was observed after one bout of iTBS, being indicative of the build-up of after-effects over repeated stimulations within the same session (Tse *et al.*, 2018; Aasen *et al.*, 2024).

There is converging evidence that midfrontal theta oscillations arising from the dACC are linked to cognitive control under various conflicting situations (Cavanagh & Frank, 2014), including Pavlovian bias in RL tasks (Cavanagh *et al.*, 2013; Swart *et al.*, 2018; Csifcsák *et al.*, 2020; but see: Sedlinská *et al.*, 2023). Although theta-band activity was not measured in this study, our result is in accordance with the pre-registered hypothesis and support the utility of the current iTBS protocol in enhancing cognitive control under motivational conflict during RL. However, we note that not all results converge on this conclusion, since the PPI or conflict-associated RTs remained unaffected by active stimulation. This is in line with earlier conclusions about the increased sensitivity of model-based estimates to experimental manipulations in the context of Pavlovian-instrumental interactions (Csifcsák *et al.*, 2020; Sedlinská *et al.*, 2023), and also supports the distinction between accuracy- and RT-based indices of Pavlovian bias, with the former being more sensitive to iTBS, whereas the latter being influenced by EHP only (see above).

An important aspect of the effect of iTBS on parameter π was that it developed only in the Sham-Warm group, but not in participants receiving painful skin stimulation between the two bouts of active iTBS. In other words, delivering EHP in block 2 prevented the build-up of iTBS-associated enhancement of cognitive control over Pavlovian bias. After-effects of rTMS are known to show state-dependent properties, whereby the magnitude of neuromodulation is influenced by the excitability state of the underlying neural population (Bergmann, 2018). Given that EHP was delivered in-between the two bouts of iTBS, we postulate a similar

mechanism here, with uncontrollable pain shifting the sensitivity of mPFC/dACC neurons away from the threshold of iTBS to enhance cognitive control.

The dACC is known for its role in pain perception as well as cognitive control, and even though the these functional properties are spatially overlapping, they involve distinct subregions within this area (Shackman *et al.*, 2011; Kragel *et al.*, 2018). While there is empirical data supporting the antinociceptive effect of rTMS above the mPFC/dACC (Kanda *et al.*, 2003), other studies point at higher pain ratings following excitatory stimulation (Yoo *et al.*, 2006), rendering the evidence about the effect inconclusive (Lefaucheur *et al.*, 2008). Our participants' pain ratings remained unchanged following active iTBS, further questioning if stimulation of the mPFC/dACC modulates nociception. In contrast, EHP was sufficient to interfere with processes linked to the recruitment of cognitive control. Whether the interacting effect of EHP and iTBS on Pavlovian bias is restricted to the mPFC/dACC, and if it is mediated by shared or distinct neural assemblies within this region, remains to be determined. Nevertheless, our finding provides insights into the multifaceted role of the mPFC/dACC in the context of pain and RL, which could be linked to suboptimal cognitive control and decision-making in patients with chronic pain (Apkarian *et al.*, 2004; Moriarty *et al.*, 2011; Berryman *et al.*, 2014; Attridge *et al.*, 2019). Finally, our data provide some evidence for the hypothesized link between exaggerated Pavlovian bias and cognitive mechanisms underlying LH as predicted by the learned controllability account (Maier & Seligman, 2016), since stimulation of the mPFC/dACC attenuated Pavlovian bias (possibly via promoting the recruitment of cognitive control), and this effect was counteracted by uncontrollable pain, a factor commonly associated with LH.

Effects of experimental pain on RL are associated with activity in the mPFC/dACC

Despite our prediction that iTBS would exert opposite effects on RL to painful skin stimulation, when applied in isolation, the two interventions induced largely similar changes

in model-based estimates of go-bias, learning rate and exploration tendency. Given this substantial overlap between the consequences of iTBS and EHP on RL, we speculate that the potential of EHP to induce LH-like effects in learning and choice behavior may have been related to activity in the mPFC/dACC.

In addition to its association with nociception and cognitive control, the mPFC/dACC has also been implicated in latent processes of RL. More specifically, both neuroimaging and single-cell recordings support the role of distinct subregions withing this area in response selection and exploratory behavior (Rushworth, 2008; Algermissen *et al.*, 2021; Holroyd & Verguts, 2021; Clairis & Pessiglione, 2024), driven by computations regarding the value of actions and cognitive control in the face of uncertainty (Behrens *et al.*, 2007; Rushworth & Behrens, 2008; Shenhav *et al.*, 2013; Silvetti *et al.*, 2014). Furthermore, our finding regarding the direction of effects following iTBS aligns well with prior results pointing at the mPFC/dACC in contributing to LH-like behavior both in healthy adults and in patients with anxiety/mood disorders or chronic pain (Bauer *et al.*, 2003; Diener *et al.*, 2010; Salomons *et al.*, 2012).

Non-invasive stimulation of the mPFC/dACC via electric currents yielded similar results using the same Pavlovian-instrumental Go/NoGo task, albeit combined with a different manipulation of task controllability (Csifcsák, Bjørkøy, *et al.*, 2021; Sedlinská *et al.*, 2023). In particular, Csifcsák et al. (2021) reported that HD-tDCS coupled with reduced control over outcomes (while reward/loss rates were matched to a control group) resulted in weaker exploration tendency, without substantially influencing go-bias and learning rates. Conversely, Sedlinská et al. (2023) introduced simultaneous manipulation of controllability and reward/loss rates, which shifted the effect of active mPFC/dACC stimulation towards suppressed go-bias and reduced learning rates, but interestingly, also to more pronounced exploration. Even though it is challenging to reconcile findings from HD-tDCS with those

reported herein due to differences in brain stimulation techniques (electric currents vs. magnetic fileds) both in terms of their intensity and spatial focality (Priori *et al.*, 2009; Turi *et al.*, 2021), as well the way outcome controllability was manipulated, our results provide converging evidence for the involvement of the mPFC/dACC in adopting a RL profile that resembles LH. Given that EHP exerted largely similar effects on go-bias, learning rate and exploration tendency to that observed for iTBS above the mPFC/dACC, and that affective/cognitive aspects of pain are also closely linked to the mPFC/dACC (Vogt, 2005; Shackman *et al.*, 2011), we propose that pain-induced shifts in choice behavior resembling LH were driven by activity in this cortical region.

In support of this notion, effects of the combined delivery of EHP and iTBS were non-additive in terms of go-bias (block 2), learning rate (block 2) and exploration (block 3), so that their joint effects did not exceed their individual ones, when applied in isolation. Ceiling effects are well-known in the field of non-invasive brain stimulation (Bonaiuto & Bestmann, 2015), and may indicate that EHP and iTBS were modulating activity in overlapping neural populations. The only exception showing an antagonism between EHP and iTBS was the go-bias estimate in block 3. Here, the Active-Pain group did not demonstrate the tendency for behavioral passivity that was present in the Active-Warm and the Sham-Pain groups. This phenomenon points at a non-linear interaction between pain and mPFC/dACC stimulation with respect to general action tendency during RL, and warrants further investigations for the potential of iTBS to counteract pain-induced behavioral passivity in health and disease.

Conclusion

The current study provides insights into how uncontrollable pain interferes with value-based learning and decision-making, and adds to our understanding about the causal role of mPFC/dACC activity in choice behavior under circumstances that may lead to LH. Whereas EHP did not impair response accuracy, and was ineffective in modulating two of our

Pavlovian measures (PPI or parameter π), exploratory RT analysis provided some support for our pre-registered hypothesis about pain-induced enhancement of Pavlovian bias over instrumental choices. Furthermore, two bouts of iTBS reduced Pavlovian bias in participants receiving warm stimulation, underlining the role of mPFC/dACC-mediated cognitive control over Pavlovian valuation in RL tasks. Crucially, this effect was counteracted by EHP, being suggestive of antagonistic influences between pain and non-invasive mPFC/dACC stimulation in modulating Pavlovian-instrumental interactions.

In the context of LH, EHP reduced go-bias, learning rate and exploration tendencies not only during pain exposure, but also as a carry-over effect in block 3, pointing at a link between uncontrollable pain and LH-like behavior in healthy adults. Surprisingly, largely similar effects were observed after iTBS, implicating neural excitability in the mPFC/dACC in shifting choice behavior in a manner that is compatible with the concept of LH.

Altogether, our results highlight the intricacies of RL under manipulations of outcome controllability and reward/loss frequency, when coupled with uncontrollable painful stimulation. Moreover, the findings further our understanding about the causal role of mPFC/dACC activity in choice behavior under circumstances that may lead to LH.

Data Availability

Publicly available at: https://osf.io/r74uk/

Funding

This work was supported by The Research Council of Norway through the Student Research Programme in Psychology.

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethics Approval

The study protocol complied with the Declaration of Helsinki and was approved by the Regional Committees for Medical and Health Research Ethics of Northern Norway (REK Nord; Case number: 318704).

Author contributions

SB: conceptualization, formal analysis, investigation, methodology, writing – original draft, writing – review & editing. FL: conceptualization, formal analysis, investigation, methodology, visualization, writing – review & editing. MM: conceptualization, formal analysis, methodology, software, supervision, validation, visualization, writing – review & editing. GC: conceptualization, formal analysis, funding acquisition, methodology, project administration, supervision, validation, visualization, writing – original draft, writing – review & editing. All authors contributed to the article and approved the submitted version.

References

Aasen, S.R., Drevland, R.N., Csifcsák, G., & Mittner, M. (2024) Increasing Mind Wandering With Accelerated Intermittent Theta Burst Stimulation Over the Left Dorsolateral Prefrontal Cortex. *PsyArXiv*, https://doi.org/10.31234/osf.io/fkx3w.

Algermissen, J., Swart, J.C., Scheeringa, R., Cools, R., & den Ouden, H.E.M. (2021) Striatal BOLD and Midfrontal Theta Power Express Motivation for Action. *Cereb. Cortex*, bhab391.

Apkarian, A.V., Sosa, Y., Krauss, B.R., Thomas, P.S., Fredrickson, B.E., Levy, R.E., Harden, R.N., & Chialvo, D.R. (2004) Chronic pain patients are impaired on an emotional decision-making task. *Pain*, **108**, 129–136.

Arendt-Nielsen, L. & Chen, A.C.N. (2003) Lasers and other thermal stimulators for activation of skin nociceptors in humans. *Neurophysiol. Clin. Neurophysiol.*, **33**, 259–268.

Attridge, N., Pickering, J., Inglis, M., Keogh, E., & Eccleston, C. (2019) People in pain make poorer decisions. *PAIN*, **160**, 1662.

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

Beck, A.T., Weissman, A., Lester, D., & Trexler, L. (1974) The measurement of pessimism: The Hopelessness Scale. *J. Consult. Clin. Psychol.*, **42**, 861–865.

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

Bergmann, T.O. (2018) Brain State-Dependent Brain Stimulation. *Front. Psychol.*, **9**.

Berryman, C., Stanton, T.R., Bowering, K.J., Tabor, A., McFarlane, A., & Moseley, G.L. (2014) Do people with chronic pain have impaired executive function? A meta-analytical review. *Clin. Psychol. Rev.*, **34**, 563–579.

Bonaiuto, J.J. & Bestmann, S. (2015) Understanding the nonlinear physiological and behavioral effects of tDCS through computational neurostimulation. *Prog. Brain Res.*, **222**, 75–103.

Boureau, Y.-L., Sokol-Hessner, P., & Daw, N.D. (2015) Deciding How To Decide: Self-Control and Meta-Decision Making. *Trends Cogn. Sci.*, **19**, 700–710.

Brunoni, A.R., Amadera, J., Berbel, B., Volz, M.S., Rizzerio, B.G., & Fregni, F. (2011) A systematic review on reporting and assessment of adverse effects associated with transcranial direct current stimulation. *Int. J. Neuropsychopharmacol.*, **14**, 1133–1145.

Buhle, J. & Wager, T.D. (2010) Performance-dependent inhibition of pain by an executive working memory task. *PAIN*, **149**, 19.

Bushnell, M.C., Duncan, G.H., Hofbauer, R.K., Ha, B., Chen, J.-I., & Carrier, B. (1999) Pain perception: Is there a role for primary somatosensory cortex? *Proc. Natl. Acad. Sci.*, **96**, 7705–7709.

Cacioppo, J.T., Petty, R.E., & Feng Kao, C. (1984) The Efficient Assessment of Need for Cognition. *J. Pers. Assess.*, **48**, 306–307.

Carpenter, B., Gelman, A., Hoffman, M.D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M.A., Guo, J., Li, P., & Riddell, A. (2017) Stan: A Probabilistic Programming Language. *J. Stat. Softw.*, **76**, 1.

Carver, C.S. & White, T.L. (1994) Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS Scales. *J. Pers. Soc. Psychol.*, **67**, 319–333.

Cavanagh, J.F., Eisenberg, I., Guitart-Masip, M., Huys, Q., & Frank, M.J. (2013) Frontal Theta Overrides Pavlovian Learning Biases. *J. Neurosci.*, **33**, 8541–8548.

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

Chase, H.W., Frank, M.J., Michael, A., Bullmore, E.T., Sahakian, B.J., & Robbins, T.W. (2010) Approach and avoidance learning in patients with major depression and healthy controls: relation to anhedonia. *Psychol. Med.*, **40**, 433–440.
Chen, C., Takahashi, T., Nakagawa, S., Inoue, T., & Kusumi, I. (2015) Reinforcement learning in depression: A review of computational research. *Neurosci. Biobehav. Rev.*, **55**, 247–267.
Chung, S.W., Rogasch, N.C., Hoy, K.E., & Fitzgerald, P.B. (2018) The effect of single and repeated prefrontal intermittent theta burst stimulation on cortical reactivity and working memory. *Brain Stimulat.*, **11**, 566–574.
Clairis, N. & Pessiglione, M. (2024) Value estimation versus effort mobilization: a general dissociation between ventromedial and dorsomedial prefrontal cortex. *J. Neurosci.*, **44**.
Craig, A.R., Franklin, J.A., & Andrews, G. (1984) A scale to measure locus of control of behaviour. *Br. J. Med. Psychol.*, **57**, 173–180.
Csifcsák, G., Bjørkøy, J., Kuyateh, S., Reithe, H., & Mittner, M. (2021) Transcranial Direct Current Stimulation above the Medial Prefrontal Cortex Facilitates Decision-Making following Periods of Low Outcome Controllability. *eneuro*, **8**, ENEURO.0041-21.2021.
Csifcsák, G., Forstmann, B.U., & Mittner, M. (2021) Transcranial stimulation and decision-making.
Csifcsák, G., Melsæter, E., & Mittner, M. (2020) Intermittent Absence of Control during Reinforcement Learning Interferes with Pavlovian Bias in Action Selection. *J. Cogn. Neurosci.*, **32**, 646–663.
Defrin, R., Shachal-Shiffer, M., Hadgadg, M., & Peretz, C. (2006) Quantitative Somatosensory Testing of Warm and Heat-Pain Thresholds: The Effect of Body Region and Testing Method. *Clin. J. Pain*, **22**, 130.
Diederichs, C., DeMayo, M.M., Cole, J., Yatham, L.N., Harris, A.D., & McGirr, A. (2021) Intermittent Theta-Burst Stimulation Transcranial Magnetic Stimulation Increases GABA in the Medial Prefrontal Cortex: A Preliminary Sham-Controlled Magnetic Resonance Spectroscopy Study in Acute Bipolar Depression. *Front. Psychiatry*, **12**.
Diener, C., Kuehner, C., & Flor, H. (2010) Loss of control during instrumental learning: A source localization study. *NeuroImage*, **50**, 717–726.
Dorfman, H.M. & Gershman, S.J. (2019) Controllability governs the balance between Pavlovian and instrumental action selection. *Nat. Commun.*, **10**, 5826.
Edwards, R.R., Cahalan, C., Mensing, G., Smith, M., & Haythornthwaite, J.A. (2011) Pain, catastrophizing, and depression in the rheumatic diseases. *Nat. Rev. Rheumatol.*, **7**, 216–224.
Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007) G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods*, **39**, 175–191.
Gandhi, W., Becker, S., & Schweinhardt, P. (2013) Pain increases motivational drive to obtain reward, but does not affect associated hedonic responses: A behavioural study in healthy volunteers. *Eur. J. Pain*, **17**, 1093–1103.
Gelman, A. & Rubin, D.B. (1992) Inference from Iterative Simulation Using Multiple Sequences. *Stat. Sci.*, **7**, 457–472.
Gershman, S.J., Guitart-Masip, M., & Cavanagh, J.F. (2021) Neural signatures of arbitration between Pavlovian and instrumental action selection. *PLOS Comput. Biol.*, **17**, e1008553.
Guitart-Masip, M., Duzel, E., Dolan, R., & Dayan, P. (2014) Action versus valence in decision making. *Trends Cogn. Sci.*, **18**, 194–202.

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. *J. Neurosci.*, **31**, 7867–7875.

Harvie, D.S., Moseley, G.L., Hillier, S.L., & Meulders, A. (2017) Classical Conditioning Differences Associated With Chronic Pain: A Systematic Review. *J. Pain*, **18**, 889–898.

Hess, L.E., Haimovici, A., Muñoz, M.A., & Montoya, P. (2014) Beyond pain: modeling decision-making deficits in chronic pain. *Front. Behav. Neurosci.*, **8**.

Hoffman, M.D. & Gelman, A. (2011) The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *ArXiv E-Prints*,.

Holroyd, C.B. & Verguts, T. (2021) The best laid plans: computational principles of anterior cingulate cortex. *Trends Cogn. Sci.*, **25**, 316–329.

Kanda, M., Mima, T., Oga, T., Matsuhashi, M., Toma, K., Hara, H., Satow, T., Nagamine, T., Rothwell, J.C., & Shibasaki, H. (2003) Transcranial magnetic stimulation (TMS) of the sensorimotor cortex and medial frontal cortex modifies human pain perception. *Clin. Neurophysiol.*, **114**, 860–866.

Koppel, L., Andersson, D., Morrison, I., Posadzy, K., Västfjäll, D., & Tinghög, G. (2017) The effect of acute pain on risky and intertemporal choice. *Exp. Econ.*, **20**, 878–893.

Kragel, P.A., Kano, M., Van Oudenhove, L., Ly, H.G., Dupont, P., Rubio, A., Delon-Martin, C., Bonaz, B.L., Manuck, S.B., & Gianaros, P.J. (2018) Generalizable representations of pain, cognitive control, and negative emotion in medial frontal cortex. *Nat. Neurosci.*, **21**, 283–289.

LeDoux, J. & Daw, N.D. (2018) Surviving threats: neural circuit and computational implications of a new taxonomy of defensive behaviour. *Nat. Rev. Neurosci.*, **19**, 269–282.

Lefaucheur, J.-P., Antal, A., Ahdab, R., Ciampi de Andrade, D., Fregni, F., Khedr, E.M., Nitsche, M., & Paulus, W. (2008) The use of repetitive transcranial magnetic stimulation (rTMS) and transcranial direct current stimulation (tDCS) to relieve pain. *Brain Stimulat.*, **1**, 337–344.

Lier, E.J., Van Rijn, C.M., De Vries, M., Van Goor, H., & Oosterman, J.M. (2022) The interaction between pain and cognition: on the roles of task complexity and pain intensity. *Scand. J. Pain*, **22**, 385–395.

Ligneul, R., Mainen, Z.F., Ly, V., & Cools, R. (2022) Stress-sensitive inference of task controllability. *Nat. Hum. Behav.*, **6**, 812–822.

Lowe, C.J., Manocchio, F., Safati, A.B., & Hall, P.A. (2018) The effects of theta burst stimulation (TBS) targeting the prefrontal cortex on executive functioning: A systematic review and meta-analysis. *Neuropsychologia*, **111**, 344–359.

Ly, V., Wang, K.S., Bhanji, J., & Delgado, M.R. (2019) A Reward-Based Framework of Perceived Control. *Front. Neurosci.*, **13**, 65.

MacAulay, R.K., McGovern, J.E., & Cohen, A.S. (2014) Understanding Anhedonia: The Role of Perceived Control. In Ritsner, M.S. (ed), *Anhedonia: A Comprehensive Handbook Volume I*. Springer Netherlands, Dordrecht, pp. 23–49.

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

Moriarty, O., McGuire, B.E., & Finn, D.P. (2011) The effect of pain on cognitive function: a review of clinical and preclinical research. *Prog. Neurobiol.*, **93**, 385–404.

Naert, A.L.G., Kehlet, H., & Kupers, R. (2008) Characterization of a novel model of tonic heat pain stimulation in healthy volunteers. *PAIN*, **138**, 163.

Navratilova, E., Xie, J.Y., King, T., & Porreca, F. (2013) Evaluation of reward from pain relief. *Ann. N. Y. Acad. Sci.*, **1282**, 1–11.

Nielsen, C.S., Price, D.D., Vassend, O., Stubhaug, A., & Harris, J.R. (2005) Characterizing individual differences in heat-pain sensitivity. *PAIN*, **119**, 65.
Peyron, R., García-Larrea, L., Grégoire, M.-C., Costes, N., Convers, P., Lavenne, F., Mauguière, F., Michel, D., & Laurent, B. (1999) Haemodynamic brain responses to acute pain in humans: sensory and attentional networks. *Brain*, **122**, 1765–1780.
Phelps, C.E., Navratilova, E., & Porreca, F. (2021) Cognition in the Chronic Pain Experience: Preclinical Insights. *Trends Cogn. Sci.*, **25**, 365–376.
Priori, A., Hallett, M., & Rothwell, J.C. (2009) Repetitive transcranial magnetic stimulation or transcranial direct current stimulation? *Brain Stimulat.*, **2**, 241–245.
Raio, C.M., Hartley, C.A., Orederu, T.A., Li, J., & Phelps, E.A. (2017) Stress attenuates the flexible updating of aversive value. *Proc. Natl. Acad. Sci.*, **114**, 11241–11246.
Rangel, A., Camerer, C., & Montague, P.R. (2008) A Framework for Studying the Neurobiology of Value-Based Decision Making. *Nat. Rev. Neurosci.*, **9**, 545–556.
Ridding, M.C. & Rothwell, J.C. (2007) Is there a future for therapeutic use of transcranial magnetic stimulation? *Nat. Rev. Neurosci.*, **8**, 559–567.
Rushworth, M.F. & Behrens, T.E. (2008) Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat. Neurosci.*, **11**, 389–397.
Rushworth, M.F.S. (2008) Intention, Choice, and the Medial Frontal Cortex. *Ann. N. Y. Acad. Sci.*, **1124**, 181–207.
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. *Eur. J. Neurosci.*, **35**, 1481–1487.
Samwel, H.J.A., Evers, A.W.M., Crul, B.J.P., & Kraaimaat, F.W. (2006) The Role of Helplessness, Fear of Pain, and Passive Pain-Coping in Chronic Pain Patients. *Clin. J. Pain*, **22**, 245.
Sedlinská, T., Bolte, L., Melsæter, E., Mittner, M., & Csifcsák, G. (2023) Transcranial direct-current stimulation enhances Pavlovian tendencies during intermittent loss of control. *Front. Psychiatry*, **14**, 1164208.
Seminowicz, D.A. & Davis, K.D. (2007) A re-examination of pain–cognition interactions: Implications for neuroimaging. *PAIN*, **130**, 8.
Seville, J.L. & Robinson, A.B. (2000) Locus of control in the patient with chronic pain. In *Personality Characteristics of Patients with Pain*. American Psychological Association, Washington, DC, US, pp. 165–179.
Seymour, B. (2019) Pain: A Precision Signal for Reinforcement Learning and Control. *Neuron*, **101**, 1029–1041.
Shackman, A.J., Salomons, T.V., Slagter, H.A., Fox, A.S., Winter, J.J., & Davidson, R.J. (2011) The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nat. Rev. Neurosci.*, **12**, 154–167.
Shenhav, A., Botvinick, M.M., & Cohen, J.D. (2013) The Expected Value of Control: An Integrative Theory of Anterior Cingulate Cortex Function. *Neuron*, **79**, 217–240.
Silvetti, M., Alexander, W., Verguts, T., & Brown, J.W. (2014) From conflict management to reward-based decision making: actors and critics in primate medial frontal cortex. *Neurosci. Biobehav. Rev.*, **46**, 44–57.
Suppa, A., Huang, Y.-Z., Funke, K., Ridding, M.C., Cheeran, B., Di Lazzaro, V., Ziemann, U., & Rothwell, J.C. (2016) Ten Years of Theta Burst Stimulation in Humans: Established Knowledge, Unknowns and Prospects. *Brain Stimulat.*, **9**, 323–335.
Swart, J.C., Frank, M.J., Määttä, J.I., Jensen, O., Cools, R., & Ouden, H.E.M. den (2018) Frontal network dynamics reflect neurocomputational mechanisms for reducing maladaptive biases in motivated action. *PLOS Biol.*, **16**, e2005979.

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**, e22169.

Teodorescu, K. & Erev, I. (2014) Learned Helplessness and Learned Prevalence: Exploring the Causal Relations Among Perceived Controllability, Reward Prevalence, and Exploration. *Psychol. Sci.*, **25**, 1861–1869.

Tse, N.Y., Goldsworthy, M.R., Ridding, M.C., Coxon, J.P., Fitzgerald, P.B., Fornito, A., & Rogasch, N.C. (2018) The effect of stimulation interval on plasticity following repeated blocks of intermittent theta burst stimulation. *Sci. Rep.*, **8**, 8526.

Turi, Z., Normann, C., Domschke, K., & Vlachos, A. (2021) Transcranial magnetic stimulation in psychiatry: is there a need for electric field standardization? *Front. Hum. Neurosci.*, **15**, 639640.

Vlaeyen, J.W.S. & Crombez, G. (2020) Behavioral Conceptualization and Treatment of Chronic Pain. *Annu. Rev. Clin. Psychol.*, **16**, 187–212.

Vlaeyen, J.W.S. & Linton, S.J. (2000) Fear-avoidance and its consequences in chronic musculoskeletal pain: a state of the art. *Pain*, **85**, 317–332.

Vogt, B.A. (2005) Pain and emotion interactions in subregions of the cingulate gyrus. *Nat. Rev. Neurosci.*, **6**, 533–544.

Wang, K.S., Yang, Y.-Y., & Delgado, M.R. (2021) How perception of control shapes decision making. *Curr. Opin. Behav. Sci.*, Value based decision-making, **41**, 85–91.

Watson, D., Clark, L.A., & Tellegen, A. (1988) Development and validation of brief measures of positive and negative affect: The PANAS scales. *J. Pers. Soc. Psychol.*, **54**, 1063–1070.

Wu, X., Wang, L., Geng, Z., Wei, L., Yan, Y., Xie, C., Chen, X., Ji, G.-J., Tian, Y., & Wang, K. (2021) Improved Cognitive Promotion through Accelerated Magnetic Stimulation. *eNeuro*, **8**.

Xu, M., Nikolin, S., Samaratunga, N., Chow, E.J.H., Loo, C.K., & Martin, D.M. (2024) Cognitive Effects Following Offline High-Frequency Repetitive Transcranial Magnetic Stimulation (HF-rTMS) in Healthy Populations: A Systematic Review and Meta-Analysis. *Neuropsychol. Rev.*, **34**, 250–276.

Yoo, W.-K., Kim, Y.-H., Doh, W.-S., Lee, J.-H., Jung, K.-I., Park, D.-S., & Park, E.-S. (2006) Dissociable modulating effect of repetitive transcranial magnetic stimulation on sensory and pain perception. *Neuroreport*, **17**, 141–144.

Supplementary Material

Supplementary Discussion

Effects of CRLM on reinforcement learning

There is a vast amount of research regarding LH-induction in animals (for review see: Maier & Seligman, 2016), to our knowledge, experimental protocols for triggering LH-like behavior in healthy humans are still lacking (Henkel *et al.*, 2002; Pryce *et al.*, 2011), and therefore, cognitive/affective mechanisms underlying LH in remain underexplored.

One of our key assumptions was that CRLM would enhance Pavlovian bias, an effect that would support views on the role of outcome controllability in shaping Pavlovian-instrumental interactions (Dorfman & Gershman, 2019; Csifcsák *et al.*, 2020). Instead, both the PPI measure and the model-derived Pavlovian parameter remained unchanged throughout the task in the Sham-Warm group, being in contrast with previous findings (Dorfman & Gershman, 2019; Csifcsák *et al.*, 2020, 2021). Here we speculate that the balance between instrumental and Pavlovian valuation is not only sensitive to response-outcome contingency and reward/loss frequency as previously assumed, but also to subjective evaluations of control over the environment. In earlier studies, outcome controllability during RL was withdrawn while keeping reward/loss rates matched to non-manipulated participants, leading to "illusion of control" (i.e., high self-reported controllability), coupled with higher degree of Pavlovian bias (Csifcsák et al. 2020, 2021). However, the present study kept reward rates low and loss rates high in block 2, which prevented developing a false sense of being in control, as evidenced by the reduced controllability ratings in all groups. Several authors have argued for the need to distinguish between objective and subjective controllability with respect to their influence on cognitive performance and behavior (Ly *et al.*, 2019; Wang *et al.*, 2021), which we believe could have contributed to the absence of the predicted enhancement of Pavlovian (in)action tendencies by CRLM (for similar results and conclusions, see Sedlinská *et al.*, 2023).

As anticipated, response accuracy was around chance level in block 2 during CRLM, however, this effect was not carried over to block 3, in which control over outcomes was restored. With respect to other latent parameters of RL, CRLM exerted opposite effects than those predicted for LH-like behavior, enhancing go-bias, learning rate and temperature/exploration tendency. Controllability in can be viewed as an appetitive signal prompting agents to seek out stable environments (Ly *et al.*, 2019), and therefore, the immediate effects of CRLM in our study may reflect an attempt to regain control over outcomes in the card game. Thus, we argue that CRLM-induced shifts in RL may be associated with the high resilience of our sample (consisting of healthy adults), demonstrating proactive behaviour and more vigorous updating of internal stimulus/action values once the increased volatility of the environment was recognized (Behrens *et al.*, 2007; Pulcu & Browning, 2017). In conclusion, our CRLM protocol was not successful to trigger LH-like decision-making patterns, but nevertheless, we cannot rule out the possibility that it augmented the effects of EHP or iTBS when applied in combination.

References

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

Csifcsák, G., Bjørkøy, J., Kuyateh, S., Reithe, H., & Mittner, M. (2021) Transcranial Direct Current Stimulation above the Medial Prefrontal Cortex Facilitates Decision-Making following Periods of Low Outcome Controllability. *eneuro*, **8**, ENEURO.0041-21.2021.

Csifcsák, G., Melsæter, E., & Mittner, M. (2020) Intermittent Absence of Control during Reinforcement Learning Interferes with Pavlovian Bias in Action Selection. *J. Cogn. Neurosci.*, **32**, 646–663.

Dorfman, H.M. & Gershman, S.J. (2019) Controllability governs the balance between Pavlovian and instrumental action selection. *Nat. Commun.*, **10**, 5826.

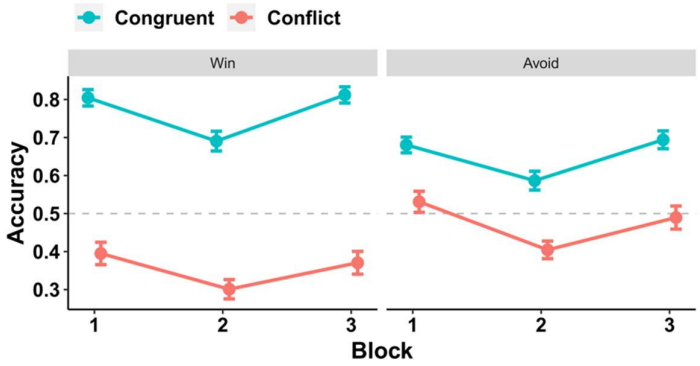
Henkel, V., Bussfeld, P., Möller, H.-J., & Hegerl, U. (2002) Cognitive-behavioural theories of helplessness/hopelessness: Valid models of depression? *Eur. Arch. Psychiatry Clin. Neurosci.*, **252**, 240–249.

Ly, V., Wang, K.S., Bhanji, J., & Delgado, M.R. (2019) A Reward-Based Framework of Perceived Control. *Front. Neurosci.*, **13**, 65.

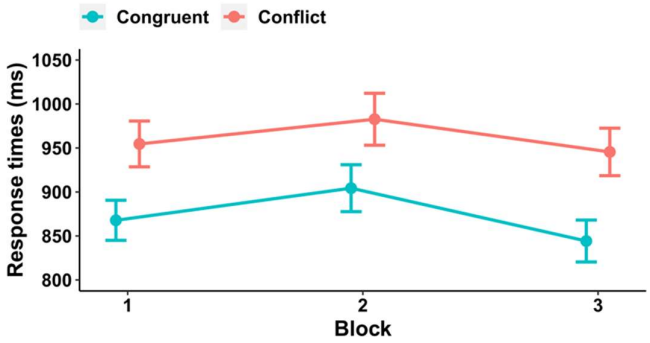
Maier, S.F. & Seligman, M.E.P. (2016) Learned helplessness at fifty: Insights from neuroscience. *Psychol. Rev.*, **123**, 349–367.

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. *Pharmacol. Ther.*, **132**, 242–267.
Pulcu, E. & Browning, M. (2017) Affective bias as a rational response to the statistics of rewards and punishments. *Elife*, **6**, e27879–e27879.
Sedlinská, T., Bolte, L., Melsæter, E., Mittner, M., & Csifcsák, G. (2023) Transcranial direct-current stimulation enhances Pavlovian tendencies during intermittent loss of control. *Front. Psychiatry*, **14**, 1164208.
Wang, K.S., Yang, Y.-Y., & Delgado, M.R. (2021) How perception of control shapes decision making. *Curr. Opin. Behav. Sci.*, Value based decision-making, **41**, 85–91.

Supplementary Figures



Supplementary Figure 1. Effect of Pavlovian congruence across task blocks on response accuracy for all participants (irrespective of experimental group), plotted separately for win (left) and avoid (right) trials. Error bars represent standard errors.



Supplementary Figure 2. Effect of Pavlovian congruence across task blocks on reaction times for all participants (irrespective of experimental group). Error bars represent standard errors.

Supplementary Tables

Supplementary Table 1.

One-way ANOVA for the main effect of Group for all questionnaire subscores.

Measure	***df***	***F***	***p***	η^2_p
PANAS positive	3, 96	2.26	.087	0.07
PANAS negative	3, 96	0.61	.607	0.02
BAS Drive	3, 96	1.93	.130	0.06
BAS Fun Seeking	3, 96	0.71	.550	0.02
BAS Reward Responsiveness	3, 96	0.54	.654	0.02
BHS	3, 96	0.07	.976	< 0.01
BIS	3, 96	2.02	.116	0.06
NFC	3, 96	0.57	.634	0.02
LCB	3, 96	0.30	.827	< 0.01

Note. PANAS: Positive and Negative Affect Schedule, BAS: Behavioral Activation System, BHS: Beck Hopelessness Scale, BIS: Behavioral Inhibition System, NFC: Need For Cognition, LCB: Locus of Control of Behavior.

Supplementary Table 2.
Pre-registered analysis of subjective ratings of outcome controllability, success, mean and peak pain.

Measure	**Effects**	***df***	***F***	***p***	η^2_p
Controllability	Block	2, 192	40.32	**< .001***	0.30
	iTBS	1, 96	2.51	.117	0.03
	EHP	1, 96	1.18	.279	0.01
	Block ✻ iTBS	2, 192	1.04	.355	0.01
	Block ✻ EHP	2, 192	1.25	.288	0.01
	iTBS ✻ EHP	1, 96	0.81	.370	<0.01
	Block ✻ iTBS ✻ EHP	2, 192	0.16	.855	< 0.01
Success	Block	2, 192	68.22	**< .001***	0.42
	iTBS	1, 96	1.73	.192	0.02
	EHP	1, 96	0.07	.798	< 0.01
	Block ✻ iTBS	2, 192	0.13	.878	< 0.01
	Block ✻ EHP	2, 192	0.71	.495	< 0.01
	iTBS ✻ EHP	1, 96	1.40	.240	0.01
	Block ✻ iTBS ✻ EHP	2, 192	1.65	.196	0.02
Mean pain	Block	1.4, 134.5	258.60	**< .001***	0.73
	iTBS	1, 96	1.68	.198	0.02
	EHP	1, 96	62.98	**< .001***	0.40
	Block ✻ iTBS	1.4, 134.5	0.03	.928	<0.01
	Block ✻ EHP	1.4, 134.5	89.58	**< .001***	0.48
	iTBS ✻ EHP	1, 96	0.01	.905	<0.01
	Block ✻ iTBS ✻ EHP	1.4, 135.5	2.54	.100	0.03
Peak pain	Block	1.5, 142.5	272.73	**< .001***	0.74
	iTBS	1, 96	5.89	**.017***	0.06
	EHP	1, 96	47.41	**< .001***	0.33
	Block ✻ iTBS	1.5, 142.5	0.48	.563	<0.01
	Block ✻ EHP	1.5,142.5	83.49	**< .001***	0.47
	iTBS ✻ EHP	1, 96	<0.01	.994	<0.01
	Block ✻ iTBS ✻ EHP	1.5, 142.5	2.58	.095	0.03

Note. The stars denote significant ($p < .05$) effects.
EHP: experimental heat pain; iTBS: intermittent theta burst stimulation.

Supplementary Table 3.
Pre-registered analysis of response accuracy.

Effects	***df***	***F***	***p***	η^2_p
Block	1.8, 169.7	37.15	**< .001***	0.28
iTBS	1, 96	0.58	.450	< 0.01
EHP	1, 96	1.19	.278	0.01
iTBS x EHP	1, 96	0.72	.398	< 0.01
Block ∗ iTBS	1.8, 169.7	1.53	.220	0.02
Block ∗ EHP	1.8, 169.7	0.15	.836	< 0.01
Block ∗ iTBS ∗ EHP	1.8, 169.7	0.46	.609	< 0.01
Congruence	1, 96	92.84	**< .001***	0.49
Congruence ∗ iTBS	1, 96	0.65	.423	< 0.01
Congruence ∗ EHP	1, 96	2.98	.087	0.03
Congruence ∗ iTBS ∗ EHP	1, 96	0.06	.801	< 0.01
Valence	1, 96	0.52	.821	< 0.01
Valence ∗ iTBS	1, 96	0.10	.751	< 0.01
Valence ∗ EHP	1, 96	0.17	.683	< 0.01
Valence ∗ iTBS ∗ EHP	1, 96	0.78	.380	< 0.01
Block ∗ Congruence	2, 162	1.66	.192	0.02
Block ∗ Congruence ∗ iTBS	2, 162	1.09	.338	0.01
Block ∗ Congruence ∗ EHP	2, 162	0.86	.426	< 0.01
Block ∗ Congruence ∗ iTBS ∗ EHP	2, 162	0.38	.685	< 0.01
Block ∗ Valence	2, 192	0.05	.953	< 0.01
Block ∗ Valence ∗ iTBS	2, 192	1,07	.344	0.01
Block ∗ Valence ∗ EHP	2, 192	0.24	.790	< 0.01
Block ∗ Valence ∗ iTBS ∗ EHP	2, 192	0.75	.472	< 0.01
Congruence ∗ Valence	1, 96	27.90	**< .001***	0.23
Congruence ∗ Valence ∗ iTBS	1, 96	0.94	.334	0.01
Congruence ∗ Valence ∗ EHP	1, 96	0.51	.476	< 0.01
Congruence ∗ Valence ∗ iTBS ∗ EHP	1, 96	0.45	.504	< 0.01
Block ∗ Congruence ∗ Valence	1.9, 180.8	0.64	.521	< 0.01
Block ∗ Congruence ∗ Valence ∗ iTBS	1.9, 180.8	0.84	.426	< 0.01
Block ∗ Congruence ∗ Valence ∗ EHP	1.9, 180.8	0.18	.820	< 0.01
Block ∗ Congruence ∗ Valence ∗ iTBS ∗ EHP	1.9, 180.8	0.78	.455	< 0.01

Note. The stars denote significant ($p < .05$) effects.
EHP: experimental heat pain; iTBS: intermittent theta burst stimulation.

Supplementary Table 4.
Pre-registered analysis of reaction times.

Effects	***df***	***F***	***p***	η^2_p
Block	1.9, 150.4	3.11	.051	0.04
iTBS	1, 81	0.16	.692	< 0.01
EHP	1, 81	1.68	.199	0.02
iTBS x EHP	1, 81	0.79	.377	0.01
Block ∗ iTBS	1.9, 150.4	0.26	.757	< 0.01
Block ∗ EHP	1.9, 150.4	0.44	.629	0.01
Block ∗ iTBS ∗ EHP	1.9, 150.4	0.55	.568	0.01
Congruence	1, 81	24.09	**< .001***	0.23
Congruence ∗ iTBS	1, 81	0.70	.404	0.01
Congruence ∗ EHP	1, 81	0.20	.659	< 0.01
Congruence ∗ iTBS ∗ EHP	1, 81	0.12	.728	< 0.01
Block ∗ Congruence	2, 162	0.14	.868	< 0.01
Block ∗ Congruence ∗ iTBS	2, 162	0.30	.738	< 0.01
Block ∗ Congruence ∗ EHP	2, 162	2.37	.097	0.03
Block ∗ Congruence ∗ iTBS ∗ EHP	2, 162	0.65	.524	0.01

Note. The star denotes a significant ($p < .05$) effect.
EHP: experimental heat pain; iTBS: intermittent theta burst stimulation.

Supplementary Table 5.
Pre-registered analysis of PPI including Index (RBI vs. PBS) as factor.

Effects	***df***	***F***	***p***	η^2_p
Block	2, 192	1.59	.206	0.02
iTBS	1, 96	0.46	.501	< 0.01
EHP	1, 96	2.62	.109	0.03
iTBS x EHP	1, 96	0.10	.748	< 0.01
Block ∗ iTBS	2, 192	1.25	.289	0.01
Block ∗ EHP	2, 192	0.78	.462	< 0.01
Block ∗ iTBS ∗ EHP	2, 192	0.32	.728	< 0.01
Index	1, 96	20.68	**< .001***	0.18
Index ∗ iTBS	1, 96	4.02	**.048***	0.04
Index ∗ EHP	1, 96	2.80	.097	0.03
Index ∗ iTBS ∗ EHP	1, 96	0.04	.835	< 0.01
Block ∗ Index	1.8, 175.8	0.35	.683	< 0.01
Block ∗ Index ∗ iTBS	1.8, 175.8	1.04	.350	0.01
Block ∗ Index ∗ EHP	1.8, 175.8	0.78	.452	< 0.01
Block ∗ Index ∗ iTBS ∗ EHP	1.8, 175.8	0.15	.845	< 0.01

Note. The stars denote significant ($p < .05$) effects.
EHP: experimental heat pain; iTBS: intermittent theta burst stimulation; PBS: punishment-based suppression; RBI: reward-based invigoration.

Supplementary Table 6.
Exploratory analysis of response accuracy with change-scores from the baseline block.

Effects	***df***	***F***	***p***	η^2_p
Block	1, 96	37.77	**< .001***	0.28
iTBS	1, 96	1.60	.209	0.02
EHP	1, 96	0.27	.605	< 0.01
iTBS x EHP	1, 96	0.64	.427	< 0.01
Block ∗ iTBS	1, 96	1,35	.248	0.01
Block ∗ EHP	1, 96	0.72	.397	< 0.01
Block ∗ iTBS ∗ EHP	1, 96	0.94	.335	0.01
Congruence	1, 96	1.01	.316	0.01
Congruence ∗ iTBS	1, 96	0.57	.453	< 0.01
Congruence ∗ EHP	1, 96	0.24	.629	< 0.01
Congruence ∗ iTBS ∗ EHP	1, 96	1.30	.256	0.01
Valence	1, 96	0.51	.478	< 0.01
Valence ∗ iTBS	1, 96	1.04	.310	0.01
Valence ∗ EHP	1, 96	1.10	.298	0.01
Valence ∗ iTBS ∗ EHP	1, 96	0.87	.354	< 0.01
Block ∗ Congruence	1, 96	0.74	.392	< 0.01
Block ∗ Congruence ∗ iTBS	1, 96	2.65	.107	0.03
Block ∗ Congruence ∗ EHP	1, 96	0.61	.438	< 0.01
Block ∗ Congruence ∗ iTBS ∗ EHP	1, 96	0.05	.832	< 0.01
Block ∗ Valence	1, 96	0.21	.646	< 0.01
Block ∗ Valence ∗ iTBS	1, 96	0.56	.456	< 0.01
Block ∗ Valence ∗ EHP	1, 96	0.19	.667	< 0.01
Block ∗ Valence ∗ iTBS ∗ EHP	1, 96	0.12	.734	< 0.01
Congruence ∗ Valence	1, 96	2.84	.095	0.03
Congruence ∗ Valence ∗ iTBS	1, 96	1.40	.239	0.01
Congruence ∗ Valence ∗ EHP	1, 96	< 0.01	.994	< 0.01
Congruence ∗ Valence ∗ iTBS ∗ EHP	1, 96	1.52	.221	0.02
Block ∗ Congruence ∗ Valence	1, 96	< 0.01	.965	< 0.01
Block ∗ Congruence ∗ Valence ∗ iTBS	1, 96	0.63	.430	< 0.01
Block ∗ Congruence ∗ Valence ∗ EHP	1, 96	0.01	.910	< 0.01
Block ∗ Congruence ∗ Valence ∗ iTBS ∗ EHP	1, 96	0.05	.828	< 0.01

Note. The star denotes a significant ($p < .05$) effect.
EHP: experimental heat pain; iTBS: intermittent theta burst stimulation.

Supplementary Table 7.
Exploratory analysis of PPI with change-scores from the baseline block.

Effects	***df***	***F***	***p***	η^2_p
Block	1, 96	2.37	.127	0.02
iTBS	1, 96	0.22	.638	< 0.01
EHP	1, 96	0.17	.684	< 0.01
iTBS x EHP	1, 96	0.35	.556	< 0.01
Block ∗ iTBS	1, 96	2.58	.111	0.03
Block ∗ EHP	1, 96	1.57	.214	0.02
Block ∗ iTBS ∗ EHP	1, 96	0.28	.600	< 0.01
Index	1, 96	0.29	.594	< 0.01
Index ∗ iTBS	1, 96	0.64	.427	< 0.01
Index ∗ EHP	1, 96	0.08	.776	< 0.01
Index ∗ iTBS ∗ EHP	1, 96	0.05	.824	< 0.01
Block ∗ Index	1, 96	0.45	.503	< 0.01
Block ∗ Index ∗ iTBS	1, 96	1.63	.205	0.02
Block ∗ Index ∗ EHP	1, 96	1.77	.187	0.02
Block ∗ Index ∗ iTBS ∗ EHP	1, 96	0.29	.594	< 0.01

Note. EHP: experimental heat pain; iTBS: intermittent theta burst stimulation.

Supplementary Table 8.
Exploratory analysis of change-scores in reaction times from the baseline block.

Effects	***df***	***F***	***p***	η^2_p
Block	1, 81	7.83	**.006***	0.09
iTBS	1, 81	0.28	.597	< 0.01
EHP	1, 81	0.05	.830	< 0.01
iTBS x EHP	1, 81	0.58	.447	0.01
Block ∗ iTBS	1, 81	0.22	.643	< 0.01
Block ∗ EHP	1, 81	1.08	.301	0.01
Block ∗ iTBS ∗ EHP	1, 81	0.48	.489	0.01
Congruence	1, 81	< 0.01	.967	< 0.01
Congruence ∗ iTBS	1, 81	0.04	.836	< 0.01
Congruence ∗ EHP	1, 81	0.81	.370	0.01
Congruence ∗ iTBS ∗ EHP	1, 81	0.04	.838	< 0.01
Block ∗ Congruence	1, 81	0.32	.575	< 0.01
Block ∗ Congruence ∗ iTBS	1, 81	0.63	.429	0.01
Block ∗ Congruence ∗ EHP	1, 81	4.33	**.041***	0.05
Block ∗ Congruence ∗ iTBS ∗ EHP	1, 81	1.41	.238	0.02

Note. The stars denote significant ($p < .05$) effects.
EHP: experimental heat pain; iTBS: intermittent theta burst stimulation.

Supplementary Table 9.
Parameter estimates from the computational model reflecting the change from the baseline block to blocks 2 and 3 as a function of pain stimulation and/or active iTBS.

Parameter	Effect	MC	95% HDI	$p_{(MC>0)}$ or $p_{(MC<0)}$	ER^+ or ER^-
Pavlovian bias (π)	Block 2	0.112	[-0.041, 0.262]	.925	12.36
	Block 2 x EHP	-0.116	[-0.318, 0.084]	.866	6.49
	Block 2 x iTBS	-0.064	[-0.310, 0.185]	.695	2.28
	Block 2 x iTBS x EHP	-0.081	[-0.397, 0.261]	.686	2.19
	Block 3	0.046	[-0.123, 0.200]	.712	2.47
	Block 3 x EHP	-0.044	[-0.260, 0.165]	.657	1.92
	Block 3 x iTBS	**-0.274***	[-0.517, -0.017]	.983	57.82
	Block 3 x iTBS x EHP	**0.365***	[0.039, 0.698]	.984	59.61
Go-bias (b)	Block 2	**0.839***	[0.479, 1.196]	> .999	∞
	Block 2 x EHP	**-1.102***	[-1.588, -0.598]	> .999	2999.00
	Block 2 x iTBS	**-0.791***	[-1.332, -0.229]	.995	205.90
	Block 2 x iTBS x EHP	**0.995***	[0.177, 1.712]	.995	180.82
	Block 3	-0.041	[-0.362, 0.263]	.591	1.44
	Block 3 x EHP	**-0.540***	[-0.985, -0.085]	.989	88.66
	Block 3 x iTBS	**-0.495***	[-0.950, -0.018]	.981	51.17
	Block 3 x iTBS x EHP	**1.002***	[0.276, 1.756]	.997	249.00
Learning rate (α)	Block 2	**0.531***	[0.358, 0.728]	> .999	∞
	Block 2 x EHP	**-0.868***	[-2.000, -0.547]	> .999	∞
	Block 2 x iTBS	**-0.638***	[-0.935, -0.316]	> .999	∞
	Block 2 x iTBS x EHP	**0.578***	[0.054, 1.138]	.981	52.57
	Block 3	**0.652***	[0.439, 0.849]	> .999	∞
	Block 3 x EHP	**-0.359***	[-0.659, -0.040]	.985	67.18
	Block 3 x iTBS	-0.141	[-0.579, 0.277]	.728	2.68
	Block 3 x iTBS x EHP	0.002	[-0.532, 0.552]	.511	1.05
Exploration tendency (β)	Block 2	**0.313***	[0.077, 0.547]	.995	213.29
	Block 2 x EHP	**-0.340***	[-0.685, -0.010]	.974	37.71
	Block 2 x iTBS	-0.247	[-0.582, 0.081]	.926	12.45
	Block 2 x iTBS x EHP	0.302	[-0.151, 0.720]	.905	9.47
	Block 3	**0.318***	[0.092, 0.529]	.998	427.57
	Block 3 x EHP	**-0.390***	[-0.682, -0.099]	.996	239.00
	Block 3 x iTBS	**-0.395***	[-0.684, -0.076]	.994	170.43
	Block 3 x iTBS x EHP	0.373	[-0.015, 0.769]	.968	30.25

Note. The stars denote effects with the 95% HDI excluding zero.
EHP: experimental heat pain; iTBS: intermittent theta burst stimulation.