Full title: Causal role of the dorsolateral prefrontal cortex in modulating the balance between Pavlovian and instrumental systems in the punishment domain

Short title: tDCS on dlPFC modulates Pavlovian bias in the punishment domain bias

Authors:

Hyeonjin Kim[1]*, Jihyun K. Hur[2], Mina Kwon[1], Soyeon Kim[1], Yoonseo Zoh[2], Woo-Young Ahn[1,3]

[1] Department of Psychology, Seoul National University, Seoul, Korea 08826

[2] Department of Psychology, Yale University, New Haven, CT 06511

[3] Department of Brain and Cognitive Sciences, Seoul National University, Seoul, Korea 08826

Corresponding author:

Woo-Young Ahn, Ph.D.

Department of Psychology

Seoul National University

Seoul, Korea 08826

Tel: +82-2-880-2538, Fax: +82-2-877-6428. E-mail: wahn55@snu.ac.kr

Conflict of Interest: The authors declare no competing financial interest.

Abstract

Previous literature suggests that a balance between Pavlovian and instrumental decision-making systems is critical for optimal decision-making. Pavlovian bias (i.e., approach toward reward-predictive stimuli and avoid punishment-predictive stimuli) often contrasts with the instrumental response. Although recent neuroimaging studies have identified brain regions that may be related to Pavlovian bias, including the dorsolateral prefrontal cortex (dlPFC), it is unclear whether a causal relationship exists. Therefore, we investigated whether upregulation of the dlPFC using transcranial current direct stimulation (tDCS) would reduce Pavlovian bias. In this double-blind study, participants were assigned to the anodal or the sham group; they received stimulation over the right dlPFC for 3 successive days. On the last day, participants performed a reinforcement learning task known as the orthogonalized go/no-go task; this was used to assess each participant's degree of Pavlovian bias in reward and punishment domains. We used computational modeling and hierarchical Bayesian analysis to estimate model parameters reflecting latent cognitive processes, including Pavlovian bias, go bias, and choice randomness. Several computational models were compared; the model with separate Pavlovian bias parameters for reward and punishment domains demonstrated the best model fit. When using a behavioral index of Pavlovian bias, the anodal group showed significantly lower Pavlovian bias in the punishment domain, but not in the reward domain, compared with the sham group. In addition, computational modeling showed that Pavlovian bias parameter in the punishment domain was lower in the anodal group than in the sham group, which is consistent with the behavioral findings. The anodal group also showed a lower go bias and choice randomness, compared with the sham group. These findings suggest that anodal tDCS may lead to behavioral suppression or change in Pavlovian bias in the punishment domain, which will help to improve comprehension of the causal neural mechanism.

Author summary

A decision-making bias guided by the Pavlovian system (i.e., approach reward and avoid punishment) is often useful and predominant across species but it is also related to several psychiatric conditions. The dorsolateral prefrontal cortex (dlPFC) is known to be related to such "Pavlovian bias" but it is unclear whether a causal relationship exists between them. Here, we evaluated whether decision-making biases including Pavlovian bias could be modulated by exogenous brain stimulation, transcranial current direct stimulation, over the right dlPFC for 3 successive days. A combination of behavioral analysis and computational modeling revealed that the anodal group had lower Pavlovian bias in the punishment domain compared with the sham group. In addition, the anodal group showed lower go bias and choice randomness than the sham group, which can also hamper instrumental learning. These findings suggest a causal role for the dlPFC in modulating the balance between the Pavlovian and instrumental decision-making systems.

Introduction

Decision-making is governed by multiple systems, including the fundamental Pavlovian and instrumental systems. The Pavlovian system involves a pre-preprogrammed behavioral tendency known as Pavlovian bias (i.e., approaching reward-predictive stimuli and avoiding punishment-predictive stimuli) [1]. In contrast, the instrumental system involves learning the optimal response to each stimulus by evaluating its outcomes without prior preparation. Although the Pavlovian bias has several benefits, it may hamper goal-directed behavior. For example, animals (e.g., pigeons) with strong Pavlovian bias fail to learn to withhold pecking in response to stimuli predictive of food, even when they can receive food only by withholding pecking [2,3]. Humans are also affected by Pavlovian bias in various

74 decision-making situations, such as dieting [4,5] or substance abuse [6]. Thus, there is a need
75 to investigate methods to effectively overcome such bias.

76 The neural mechanisms that underlie Pavlovian bias are not fully understood, but some
77 previous research has suggested that the prefrontal cortex plays a pivotal role in overcoming
78 Pavlovian bias [7–9]. A functional magnetic resonance imaging (fMRI) study of participants
79 who successfully employed the instrumental system during conflict with the Pavlovian system
80 found that such individuals showed hyperactivation of the bilateral inferior frontal gyri while
81 anticipating inhibition [9]. In addition, an electroencephalography study showed that the
82 activation of the anterior cingulate cortex, as measured by the midfrontal theta power of the
83 electroencephalogram signal, was associated with overcoming Pavlovian bias [7]. However,
84 these studies failed to provide conclusive evidence for a causal neural mechanism, and the brain
85 regions that control Pavlovian bias remained unknown.

86 We speculated that the dorsolateral prefrontal cortex (dlPFC) might be a key region
87 involved in controlling Pavlovian bias. The dlPFC has been implicated in higher-level
88 cognitive control and goal-directed actions [4,5,10–17]. For example, dieters showed
89 hyperactivation of the dlPFC when they successfully selected healthy food over tasty food [4].
90 In addition, the dlPFC was important in individuals who valued stimuli in a context-dependent
91 manner and performed goal-directed behavior to maximize reward [10]. Although a previous
92 fMRI studying the neural correlates of Pavlovian bias did not identify the dlPFC as a candidate
93 region [9], the negative results are related to the imaging strategy used in the study, rather than
94 the lack of a relationship. The imaging was focused on subcortical structures; the dlPFC regions
95 were not assessed.

96 In the present study, we evaluated the presence of a causal relationship between the
97 dlPFC and Pavlovian bias using non-invasive brain stimulation (i.e., transcranial direct current
98 stimulation [tDCS]). Using tDCS was based on several previous studies of modulating

decision-making biases. For example, the competition between the model-based and the model-free systems [18], as well as affective bias of instrumental action [19], were modulated by tDCS targeting the prefrontal cortex. .

Overall, we investigated whether anodal tDCS on dlPFC would suppress the Pavlovian bias (sham-controlled); we sought to identify the causal neural mechanism underlying such bias. We applied anodal tDCS over the right dlPFC [20–22] for 3 consecutive days [23–25] On the third day, we administered a reinforcement learning task known as the orthogonalized go/no-go task, which measured the degree of Pavlovian bias [9]. The task had four conditions; two were Pavlovian-congruent, where go was the action required to win the reward and no-go was the action required to avoid punishment; the two remaining conditions were Pavlovian-incongruent, where go was the action required to avoid punishment and no-go was the action required to win the reward. Participants were required to learn the correct action for each condition to maximize the reward and minimize punishment. We compared the degree of Pavlovian bias across tDCS groups using the difference in behavioral accuracy between Pavlovian-congruent and Pavlovian-incongruent conditions. We also used a model parameter (i.e., Pavlovian bias parameter) estimated by computational modeling and hierarchical Bayesian analysis (HBA) as another index of Pavlovian bias. Under the punishment domain, we found significantly lower Pavlovian bias in the anodal tDCS group than in the sham group.

Results

Anodal and sham group characteristics

We analyzed data from 31 participants, including the basic demographic information (age and sex), psychiatric symptoms, and psychological characteristics (**Table 1**). There were no significant group-level differences in terms of demographic, psychiatric, and psychological

variables between the sham and the anodal groups. We also measured the perceived side effects of tDCS; we found no differences between groups in terms of itching, skin irritation, skin pain, fatigue, mood disturbance, and visual distortion ($p > 0.05$ for all; see **S1 Table** in supplementary material). However, the intensity of perceived tingling was significantly higher in the anodal group than in the sham group ($p < 0.05$). In addition, the degrees of headache and difficulty in concentration were significantly higher in the sham group than in the anodal group ($p < 0.05$ for both). The differences in perceived side effects did not affect the behavioral Pavlovian bias. However, there were significant differences in perceived duration and continuity of stimulation between the sham and anodal groups (**S2 Table** in supplementary material).

Table 1. Descriptive statistics.

	Sham (N = 14)	Anode (N = 17)	p value
Age	25.071 (3.731)	23.529 (3.484)	0.245
Sex; male	5 (35.7%)	8 (47.1%)	0.524
SCID[a]			
avoidant	2.357 (2.098)	2.294 (1.724)	0.927
dependent	1.286 (1.773)	1.412 (1.228)	0.817
obsessive-compulsive	2.714 (1.816)	3.353 (1.656)	0.315
passive-aggressive	1.000 (1.109)	1.353 (1.967)	0.555
depressive	1.714 (1.858)	2.059 (1.919)	0.618
paranoid	1.500 (1.871)	1.706 (1.929)	0.767
schizotypal	0.857 (1.027)	0.765 (1.480)	0.845
schizoid	0.929 (1.207)	1.412 (1.502)	0.339
histrionic	1.929 (1.269)	2.176 (1.590)	0.640
narcissistic	3.429 (2.503)	3.353 (2.783)	0.938
borderline	1.429 (1.742)	2.941 (3.750)	0.176
antisocial	0.857 (1.406)	0.412 (0.870)	0.289
Y-BOCS[b]			
obsessive	1.357 (2.530)	1.824 (3.206)	0.662
compulsive	4.143 (4.258)	3.688 (3.860)	0.761
BDI[c]			
	3.714 (2.920)	9.176 (12.259)	0.115
STAI-X[d]			

state	38.000 (9.397)	40.941 (9.523)	0.396
trait	37.071 (6.956)	40.176 (10.212)	0.342
BIS11[e]			
cognitive	16.429 (3.322)	16.412 (3.692)	0.990
motor	20.357 (3.934)	19.882 (5.171)	0.780
non-planning	24.571 (5.721)	24.176 (5.637)	0.848

Mean (standard deviation) for continuous variables and count (%) for categorical variables.

[a]SCID: Structured Clinical Interview for DSM-5

[b]Y-BOCS: Yale-Brown Obsessive Compulsive Scale

[c]BDI: Beck's Depression Inventory

[d]STAI-X: State-Trait Anxiety Inventory

[e]BIS 11: Barratt Impulsiveness Scale Version 11

Behavioral results

We used the difference in behavioral accuracy under Pavlovian-congruent and Pavlovian-incongruent conditions to compare the degree of Pavlovian bias across tDCS groups (see **Methods** for more information). In the punishment domain, the anodal group did not show any significant difference in behavior under the two punishment conditions (Fig 1A; $p > 0.05$). In contrast, the sham group exhibited significantly lower accuracy under the Pavlovian-incongruent punishment condition (e.g., go-to-avoid) than under the Pavlovian-congruent condition (e.g., no-go-to-avoid) (Fig 1B; $p < 0.05$). Neither group exhibited a significant difference in accuracy under the two reward conditions (e.g., go-to-win and no-go-to-win). Consistent with these findings, the behavioral Pavlovian bias index in the punishment domain was significantly lower in the anodal group than in the sham group ($p < 0.05$).

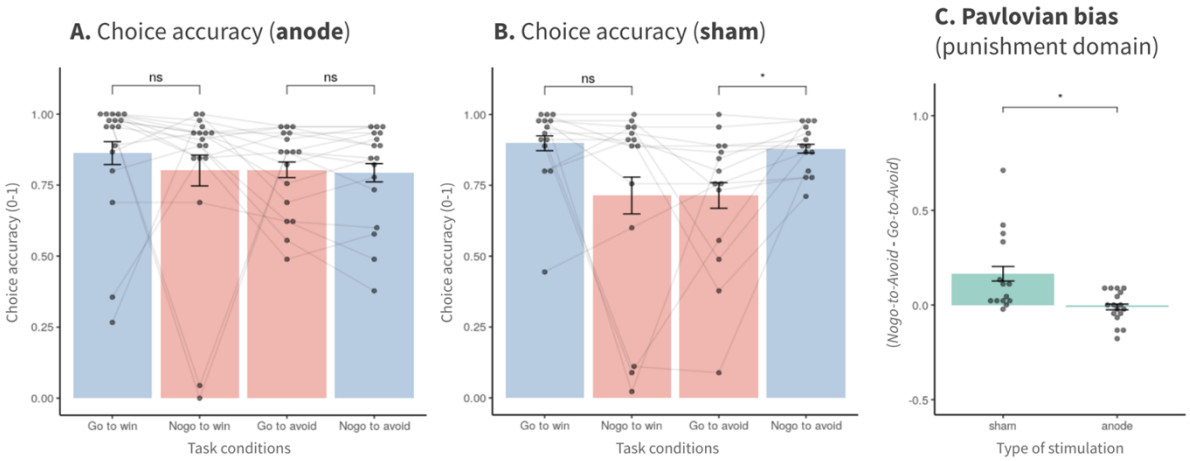


Fig 1. Pavlovian bias in the punishment domain decreased in the anodal session (accuracy). In the sham group, we found a significant difference in behavioral accuracy between the punishment conditions, which indicated the presence of Pavlovian bias, particularly in the punishment domain. This difference was not observed in the anodal group. The behavioral index of Pavlovian bias in the punishment domain also showed significantly lower bias in the anodal group than in the sham group.

*p < 0.05

Computational modeling

We tested three computational models to explain the data (see **Methods** for more information). Model 1 was a reinforcement learning model suggested by Guitart-Masip et al. (2012), which included five parameters (ξ: irreducible noise; ε: learning rate; ρ: outcome sensitivity; b: go bias; π: Pavlovian bias). Model 2 was a model with six parameters, including separate feedback sensitivity parameters for reward and punishment cues (ρ_{rew} and ρ_{pun}), compared to Model 1. Model 3 further separated Pavlovian bias parameters for reward and punishment cues (π_{rew} and π_{pun}) compared with Model 2. We compared the models using the leave-one-out information criterion (LOOIC) values, which were calculated using leave-one-out cross-validation [26] (**Table 2**). Data from the sham and anodal groups were fitted

separately. Model 3, which had separate Pavlovian bias parameters for reward and punishment domains the best fit. Models 2 and 1 were the second and third best-fitting models, respectively. Thus, we used estimated parameter values from Model 3 for the subsequent analyses.

Table 2. Model comparison (LOOIC).

	Parameters	sham	anode
Model 1	ξ, ε, b, π, ρ	1852.5	2168.7
Model 2	ξ, ε, b, π, ρ_{rew}, ρ_{pun}	1813.6	2129.7
Model 3	ξ, ε, b, π_{rew}, π_{pun}, ρ_{rew}, ρ_{pun}	**1769.7**	**2093.0**

Lower LOOIC values indicated better model performance.

Model parameters

We calculated the posterior distributions of all group-level parameters from Model 3; we compared the results between the anodal and sham groups (**Table 3, Fig 2**). The anodal group displayed credibly lower irreducible noise (ξ), compared with the sham group. Go bias (b) was also credibly lower in the anodal group than in the sham group. Finally, the anodal group had credibly lower Pavlovian bias in the punishment domain (π_{pun}), compared with the sham group; this is consistent with the behavioral analysis findings that behavioral Pavlovian bias in the punishment domain was significantly lower in the anodal group than in the sham group. Increased involvement of the frontal-striatal network after anodal stimulation might suppress the biases (e.g. Pavlovian bias in the punishment domain, go bias, and choice randomness), thereby interrupting goal-directed behavior of the instrumental system.

Table 3. Posterior mean (95% HDI highest density interval) of group mean parameters in anodal and sham groups

	Anode	Sham	Difference
ξ irreducible noise	0.027 [0.009, 0.060]	0.092 [0.056, 0.141]	-0.064 [-0.117, -0.017]

ε learning rate	0.430 [0.316, 0.554]	0.321 [0.224, 0.442]	0.107 [-0.052, 0.264]
b Go-bias	-0.173 [-0.689, 0.334]	1.27 [0.437, 2.46]	-1.45 [-2.72, -0.461]
π_{rew} Pavlovian bias (reward)	0.045 [-0.223, 0.322]	-0.141 [-0.428, 0.351]	0.183 [-0.364, 0.587]
π_{pun} Pavlovian bias (punishment)	-0.063 [-0.212, 0.091]	0.411 [-0.033, 0.855]	-0.473 [-0.939, -0.006]
ρ_{rew} reward sensitivity	12.4 [7.72, 21.8]	12.3 [7.63, 21.4]	0.146 [-9.91, 10.7]
ρ_{pun} punishment sensitivity	6.24 [4.69, 8.31]	8.83 [5.92, 14.0]	-2.58 [-8.02, 1.02]

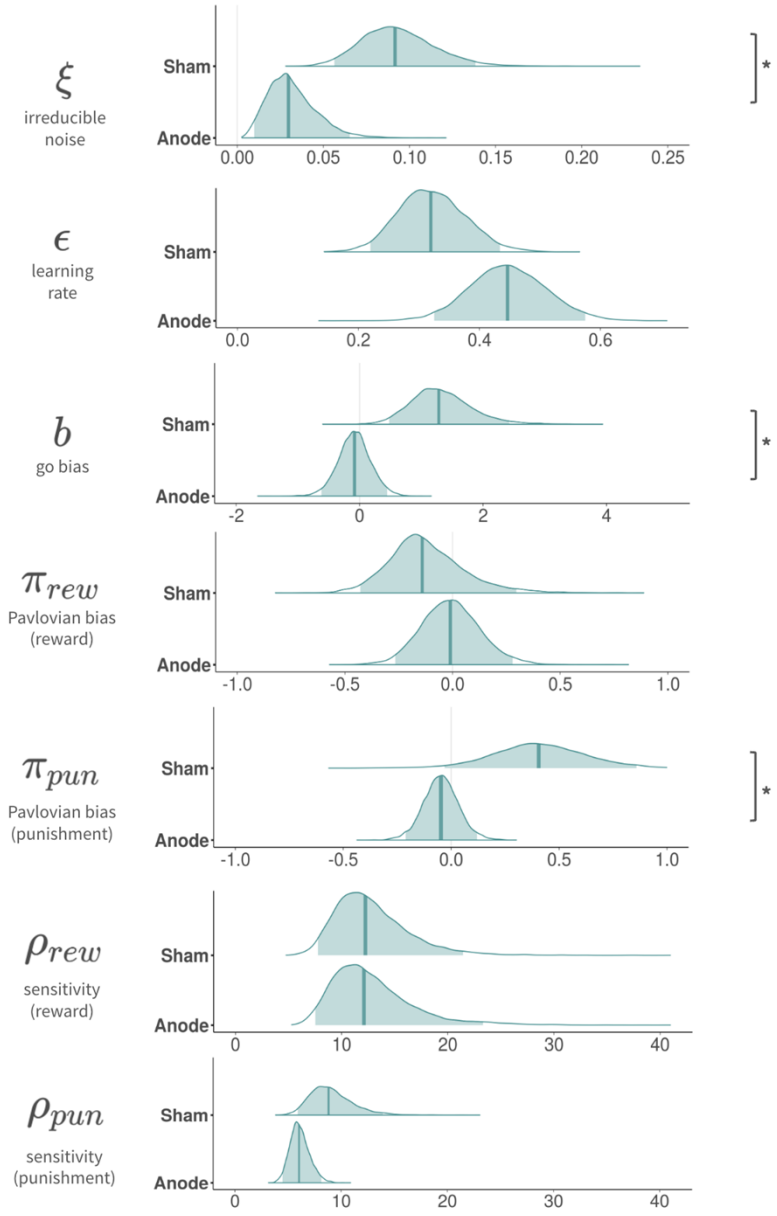


Fig 2. Pavlovian bias parameter in the punishment domain and other parameters decreased in the anodal session (modeling parameter).

We found lower parameter values in irreducible noise, go bias, and Pavlovian bias in the punishment domain in the anodal session, compared with the sham session. The decrease in Pavlovian bias in the punishment domain is consistent with the behavioral analysis.

* 95% highest density interval of the posterior difference did not include zero.

However, there were no credible differences between groups in terms of other parameters, such as learning rate (ε), Pavlovian bias in reward domain (π_{rew}), reward sensitivity (ρ_{rew}), and punishment sensitivity (ρ_{pun}).

Discussion

Our results suggest a causal role of the dlPFC in modulating Pavlovian bias in the punishment domain. Moreover, we found that other decision-making tendencies (i.e., go bias and irreducible noise) were also modulated.

We found that anodal stimulation of the dlPFC reduced Pavlovian bias, which might be related to goal-directed control in the frontal-striatal circuit. The frontal-striatal circuit connects the prefrontal cortex and striatum (including following key areas: ventral striatum [nucleus accumbens], dorsal striatum [caudate and putamen], ventromedial prefrontal cortex, dlPFC, and dorsal anterior cingulate cortex [dACC]) [17,27]. The dlPFC neurons that project to the striatum may modulate the action-outcome contingency that is encoded and updated in the striatum and ventromedial prefrontal cortex. Therefore, anodal stimulation over the right dlPFC may facilitate high-level cognitive control [4,5,12–17] and enhance goal-directed behavior by suppressing Pavlovian bias. Another possible mechanism is that anodal stimulation of the dlPFC increases dopamine release in the striatum [28,29]; when the dlPFC is stimulated, information about instrumental control is transmitted to the striatum, which responds by increasing dopamine release and overcoming the Pavlovian bias. Furthermore, the tDCS might lead to increased connectivity between the dlPFC and dACC. The dACC, along with the dlPFC, plays a critical role in updating action values and modulating the integration of subjective value and action-outcome contingency. Previous studies showed that Pavlovian bias was suppressed by frontal midline theta power, an electroencephalography correlate of dACC [7,8]. Therefore,

it is plausible that the tDCS over the dlPFC facilitated the dACC activation, which would reduce Pavlovian bias.

However, the current study only found suppression of Pavlovian bias in the punishment domain, not the reward domain. Thus, the present findings contribute to knowledge about aversion-related decision-making in the Pavlovian system [30–32]. The underlying neural mechanisms of appetitive-related decision-making have been widely investigated, but the mechanisms that underlie aversive-related decision-making have received less attention [33]. A recent study found that aversive stimuli were associated with active escape response or passive avoidance response [31]. The authors suggested that serotonin might be involved in passive inhibitory responses [34–36], while dopamine might be involved in active escape responses; this is similar to the active approach response toward appetitive stimuli [30,37]. Therefore, the current results concerning suppression of Pavlovian bias in the aversive domain, obtained by connecting behavioral activation and avoidance, might reflect a similar neural process for active escape response. These results are consistent with previous evidence that increased dopamine release after anodal tDCS of the dlPFC might suppress Pavlovian bias [32,38]. Future tDCS studies should separate avoidance and escape trials to further explore the mechanism that underlies suppression of Pavlovian bias in the punishment domain.

We also observed decreases in go bias and choice randomness in the anodal group. Because go bias and choice randomness interrupt the goal of maximizing benefit, a similar mechanism for interrupting goal-directed behavior may exist, as previously discussed. Increased involvement of the frontal-striatal network (dlPFC, striatum, and dACC) after electrical stimulation of the dlPFC might lead to the suppression of go bias and choice randomness. We presume that the instrumental system may gain preference under conflicting conditions between the instrumental and Pavlovian systems.

Our result showed that decision-making biases were modulated by external intervention, which may have clinical relevance, particularly for substance misuse and other addictive behaviors. For example, increased Pavlovian bias has been linked to substance use and gambling disorders [39,40], while increased go bias, which may reflect impaired response inhibition, has also been associated with various addictive disorders [41,42]. Choice randomness (e.g. decision-making noise or inverse temperature) was greater in patients with cocaine abuse and gambling disorders than in healthy individuals [43,44]. Thus, the current findings may aid in the development of treatments that can reduce the decision-making biases implicated in the various psychiatric conditions. Because the current study only included healthy participants, future studies should include individuals with psychiatric disorders.

A potential limitation of the current study is that the second session data were affected by the task practice effect. Therefore, overall accuracy of task performance was significantly higher in the second session than in the first session. To control for the practice effect, we only analyzed data from the first session and performed between-subject analyses (see Experimental protocol). Future studies should attempt to eliminate the practice effect from the experimental protocol.

In conclusion, our results suggest a causal relationship between non-invasive dlPFC stimulation and corresponding decision-making behavior. We found the reduced Pavlovian bias in the punishment domain, go bias, and choice randomness after dlPFC facilitation using anodal stimulation. However, further clarification using neuroimaging techniques is needed to identify the neural mechanism that underlies the effects of tDCS; efforts are also needed to determine how biases are modulated by neural changes in the dlPFC and connected brain networks. In addition, because decision-making biases have been implicated in addictive disorders, our results have practical implications for the treatment of individuals with such disorders. Furthermore, only Pavlovian bias in the punishment, but not the reward domain, was

modulated; thus, there is a need for further studies concerning aversive-related decision-making to explain why behavior related to avoiding an aversive state was only modulated by tDCS.

Materials and Methods

Participants

We recruited 39 participants from Seoul National University in Seoul, Korea, using online and offline advertisements. The experimental protocol was approved by the Seoul National University Research Ethics Committee and all participants provided informed consent before participation. Participants were excluded if they were unwilling to participate in the study, or were not fluent in Korean; they were also excluded if they reported impaired color discrimination, psychiatric medication use, neurological or psychiatric illness, or any health conditions that would make them unsuitable for the experiments. In addition, participants were excluded if they had low-quality data such as sleep during the experiment or results that indicated an inability to understand the task. Finally, we eliminated participants with a go-to-win accuracy of < 0.1 because learning failure in the easiest go-to-win condition indicated a lack of understanding or concentration. In total, data from 17 and 14 participants in the anodal and sham sessions, respectively, were analyzed (see below for more information).

Experimental protocol

First, we collected data regarding the participants' basic demographic information (age and sex) and psychological characteristics. We administered the Structured Clinical Interview for DSM-5 to detect mental illnesses (**Table 1**). In addition, we evaluated the psychological characteristics of obsession-compulsion (Yale-Brown Obsessive Compulsive Scale),

depression (Beck's Depression Inventory), anxiety (State-Trait Anxiety Inventory), and impulsivity (Barratt Impulsiveness Scale version 11) (**Table 1**). The participants visited the laboratory for 3 consecutive days and repeated the visits to counterbalance the tDCS polarity (six total sessions). For the first 2 days, participants received tDCS for 20 min; on the third day, participants performed an orthogonalized go/no-go task after they had received tDCS stimulation for 20 min. The daily visiting time was matched on a within-participant basis to remove the confounding effect of circadian rhythm [38,45]. Participants were randomly assigned to receive anodal or sham stimulation on the first or second 3 days of visits. The first and second sets of visits were separated by a mean of 24 days. We found significantly better performance in the second task (see **S1 Fig** in the supplementary material for more information), suggesting a practice effect. Therefore, we analyzed behavior data only from the first task to avoid any potential confounding effects.

tDCS stimulation

During each session, tDCS was applied for 20 min using circular sponge electrodes (size = 25 cm^2) and the Starstim system (Neuroelectrics, Barcelona, Spain). The target electrode was positioned on the right dlPFC (i.e., F4 according to the 10-10 International 10-20 electroencephalogram electrode system); the return electrode was positioned on the left cheek (**Fig 3**). The stimulation protocol was based on previous studies that used tDCS targeting dlPFC [20]. The left cheek was selected as the return position to avoid confounding cortical activation [46–49]. The stimulation included 30 s of ramp-up and ramp-down at the beginning and end of the stimulation, respectively. During the anodal session, anodal stimulation to F4 was performed for 19 min between ramp-up and ramp-down stimulations; however, in the sham session, participants were not stimulated between the ramp-up and ramp-down stimulations. During the stimulation, participants were instructed to sit with their gaze fixed on

a crosshair on the computer monitor. The double-blind mode in the Starstim software was used to ensure that all experimenters and participants remained unaware of the order of polarity. The software blinds the type of current stimulation using a 4-digit password lock set by a third-party administrator.

We employed some strategies to reduce the potential limitations of tDCS in the current stimulation protocol. First, the electrode placement and size can affect the spatial distribution of stimulation [50]. Therefore, we placed a return electrode in an extracephalic area (i.e., left chick) to minimize the stimulation of other cortical areas and the shunting effect caused by a short inter-electrode distance [51]. In addition, the effects of tDCS can be confounded by biological and lifestyle factors [51]. We mitigated such factors by stimulating participants over 3 consecutive days before the task to produce cumulative and larger effects. The participants visited the laboratory at the same time (variation of < 3 h) to reduce the effects of circadian rhythm. Finally, to reduce confounding factors related to the experimental design, we used a sham-controlled double-blind protocol.

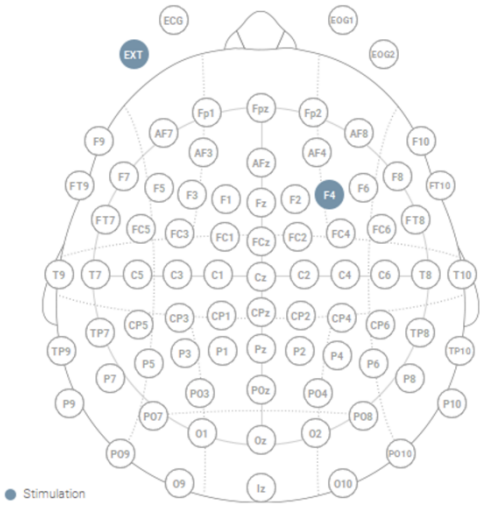
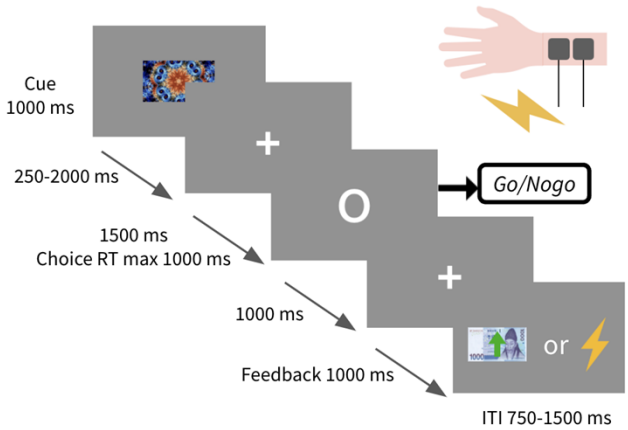


Fig 3. Montage of tDCS

342 A sponge was placed over the right dlPFC (F4) to stimulate the brain using weak electric
343 current (2 mA) for 20 min. Another sponge was placed on the left cheek. This figure was
344 adapted from the protocol summary panel in Starstim software NIC (copyright notice ©
345 Neuroelectrics SLU).
346

347 # Experimental task

348 We used the orthogonalized go/no-go task reported by Guitart-Masip et al. (2013) (**Fig**
349 **4**). At the beginning of each trial, a 1000-ms cue (fractal image) was presented to indicate one
350 of four conditions; go-to-win reward, go-to-avoid punishment, no-go-to-win reward, and no-
351 go-to-avoid punishment. After a variable interval of 250–2000 ms, a target circle appeared for
352 a maximum of 1500 ms, after which the participants responded with go or no-go within 1000
353 ms. After 1000 ms, participants received feedback according to their response and cue
354 condition. The feedback included virtual monetary gain as a reward, a yellow bar as a neutral
355 outcome, and an electric shock as punishment. The optimal response led to a beneficial
356 outcome for each condition, with a probability of 0.7. Therefore, participants learned the
357 optimal response to each cue from trial and error. The task included 180 trials, with 45 trials
358 for each condition.
359



	Rewards	*Punishments*
Go	**Go to win** Pavlovian-congruent	**Go to avoid** Pavlovian-incongruent
Nogo	**Nogo to win** Pavlovian-incongruent	**Nogo to avoid** Pavlovian-congruent

Fig 4. Orthogonalized go/no-go task.

Four types of stimuli were presented. Two stimuli were Pavlovian-congruent: go action to win reward and no-go action to avoid punishment. The remaining two stimuli were Pavlovian-incongruent: go action to avoid punishment and no-go action to win reward. Participants were instructed to maximize reward and minimize punishment by learning the correct action for each stimulus. Participants were asked to select an action when a target was presented. The reward was a picture of money, 1000 won (approximately US$ 1), whereas the punishment was an electric shock to the wrist.

RT: response time, ITI: intertrial interval

Although money is secondary feedback and shock is primary feedback, we decided to use monetary gain as the reward and electric shock as the punishment based on previous studies [52–55]. The electric shock was applied to each participant's left wrist. The intensity of the electric shock (2–12.4mA) was adjusted to cause a "moderately unpleasant" sensation (5 points

on an 11-point Likert scale [i.e., 0 = not at all unpleasant, 10 = very un- pleasant]). (see **supplementary material** for more information).

Behavioral data analysis

Behavioral data were analyzed using R [56]. Response accuracy was calculated as the proportion of correct choices. The difference in accuracy between Pavlovian-congruent and Pavlovian-incongruent conditions was evaluated using Student's t-test. The behavioral Pavlovian bias index was evaluated as the difference between the accuracy of Pavlovian-congruent and Pavlovian-incongruent conditions; it was calculated individually for each domain. For example, Pavlovian bias index in the punishment domain was calculated by subtracting the accuracy of the go-to-avoid condition from the accuracy of the no-go-to-avoid condition.

$$Pavlovian\ bias = (go\_to\_win + no\_go\_to\_avoid) - (no\_go\_to\_win + go\_to\_avoid)$$

$$Pavlovian\ bias\ (reward) = (go\_to\_win) - (no\_go\_to\_win)$$

$$Pavlovian\ bias\ (punishment) = (no\_go\_to\_avoid) - (go\_to\_avoid)$$

Computational modeling

We tested three models. A previous study suggested that Model 1 had the best fit and consisted of five parameters (Model RW + noise + bias + Pav; Guitart-Masip et al., 2012). Model 1 calculates the probability of performing (or withholding it) an action in response to the stimulus in each trial, based on action weights. If a participant successfully learned the action-reward contingency, the probability of performing a correct action was higher. It is calculated as follows:

$$p(a_1 \mid s) = \left\{\frac{exp(w(a_1,s))}{exp(W_t(a_1,s)) + exp(W_t(a_2,s))}\right\}(1-\xi) + \frac{\xi}{2} \quad \ldots (1)$$

In particular, the go action probability was larger if the W value for go (a_1) was greater using squashed softmax and for no-go (a_2), vice versa. Here, t is the trial number ($1 \leq t \leq 180$) and s is the stimulus ($s \in \{1, 2, 3, 4\}$). The four stimuli indicate four conditions, respectively: go-to-win reward, go-to-avoid punishment, no-go-to-win reward, and no-go-to-avoid punishment. In addition, a is the action ($a \in \{0, 1\}$), where 1 is go and 0 is no-go. ξ is the irreducible noise ($0 \leq \xi \leq 1$), where a value closer to 1 indicates random choice less considering the W value. W(a, s) is the action weight, which is defined as follows:

$$W_t(a,s) = \begin{cases} Q_t(a,s) + b + \pi V_t(s) & if\ a = go \\ Q_t(a,s) & if\ else \end{cases} \quad \ldots (2)$$

Q(a, s) and V(s) are updated by each trial according to the equations below:

$$Q_t(a_t,s_t) = Q_{t-1}(a_t,s_t) + \varepsilon(\rho r_t - Q_{t-1}(a_t,s_t)) \quad \ldots (3)$$

$$V_t(s_t) = V_{t-1}(s_t) + \varepsilon(\rho r_t - V_{t-1}(s_t)) \quad \ldots (4)$$

In equation (3), r is the feedback ($r \in \{-1, 0, 1\}$), where 1 is the reward, 0 is neutral, and −1 is punishment. ε is the learning rate ($0 \leq \varepsilon \leq 1$); If ε is closer to 1, it is more likely to reflect the previous feedbacks to update Q values. Furthermore, ρ is outcome sensitivity ($0 \leq \rho$). A larger ρ indicates the participant subjectively exaggerates the outcome value. Using this process, the Q value converges to the high-probability outcome for each stimulus when the correct action for the stimulus is accumulated.

In equation (4), the V value is updated in a manner similar to the Q value, but it converges to the high-probability feedback for each stimulus, regardless of the performed actions. In equation (2), for the updated Q values when the action was go, the go bias parameter b and V value multiplied by the Pavlovian bias parameter π ($0 \leq \pi$) were added to the Q values; they consisted of the W values. A large go bias was correlated with large W(go, s). When the V value converged to reward, and the action was go, the large Pavlovian bias parameter was correlated with generally large W(go, reward) and generally small W(no-go, reward). When the V value converged to punishment and the action was go, the large Pavlovian bias parameter was correlated with generally small W(go, punishment) and generally large W(no-go, punishment). This suggests that a large Pavlovian bias parameter was correlated with greater predisposition to Pavlovian-congruent choices.

Model 2 shares almost all equations and updating rules with Model 1, although it has distinct feedback sensitivity parameters for reward and punishment cues: ρ_{rew}, and ρ_{pun}, respectively. Therefore, Model 2 contains six parameters. Model 3 shares almost all equations and updating rules with Model, although it has different Pavlovian bias parameters for reward and punishment cues: π_{rew} and π_{pun}, respectively. Model 3 contains seven parameters and was used to test the distinct effect found in behavioral data, where Pavlovian bias was only significant in the punishment domain.

Model parameter estimation using HBA

The model parameters were estimated using HBA [57–59]. HBA has some advantages over the traditional maximum likelihood estimation (MLE) method. First, HBA provides estimated parameters as posterior distributions, rather than the point estimates provided by MLE. The distributions provide additional information, particularly regarding the uncertainty of estimated values. Second, the hierarchical structure of HBA allows stable and reliable

estimation of individual parameters. Individual-level MLE estimates are often noisy and unreliable; group-level MLE estimates do not include information concerning individual differences. In HBA, each individual estimate informs the group estimate (hyperparameter), and the individual commonalities reflected in the hyperparameter inform individual estimates. Therefore, individual estimates are more stable and reliable, even when data are insufficient. Previous studies have found that parameters estimated by HBA are more accurate than parameters estimated by MLE [60].

We separately fitted the models for anodal and sham groups to make stable and reliable individual estimates that reflected similarities within each group. HBA was conducted by hBayesDM (v. 1.1.1) [61] and R Stan (v. 2.21.0) [62]. Stan is a probabilistic program used for Bayesian modeling; it provides inferences based on Markov chain Monte Carlo (MCMC) algorithms, such as the Hamiltonian Monte Carlo, for sampling from high-dimensional parameter spaces. Weakly informative priors were used to reduce their influence on the posterior distributions [61]. In addition, non-centered parameterization (Matt trick) was used to optimize the sampling process [63]. We used four independent chains and a sample size of 4000, including 2000 burn-in samples per chain. The use of four independent chains ensured that the estimated parameters were stable, despite variations in the starting points [64]. We also confirmed the accuracy of parameter estimation by inspecting well-mixed trace plots and the Rhat values (Rhat < 1.1).

Model comparison

We used LOOIC to compare the models [26]. The LOOIC value for each model was calculated by estimating the out-of-sample prediction accuracy of the fitted models. This method uses the log-likelihood from posterior simulations of the estimated parameters. We

used R package loo to identify the model with the lowest LOOIC value, which had the best fit [26].

Group comparison of model parameters

For each group-level parameter, we subtracted the posterior distribution of the sham group from the posterior distribution of the anodal group for analysis of group-level differences. Group differences were considered credible when the 95% highest density intervals of posterior difference distributions did not include the value 0 [65].

Code accessibility

The codes are publicly available in the GitHub repository (behavioral data and R codes for behavioral and modeling analyses will be made available in a Github repository upon publication).

Acknowledgments

This work was supported by the Basic Science Research Program through the National Research Foundation (NRF) of Korea funded by the Ministry of Science, ICT, & Future Planning (NRF-2018R1C1B3007313 and NRF-2018R1A4A1025891); the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2021M3E5D2A0102249311); and the Creative Pioneering Researchers Program through Seoul National University to W.-Y.A.

Reference

1. Bach DR, Dayan P. Algorithms for survival: a comparative perspective on emotions. Nat Rev Neurosci. 2017;18: 311–319.

2. Williams DR, Williams H. Auto-maintenance in the pigeon: sustained pecking despite contingent non-reinforcement. J Exp Anal Behav. 1969;12: 511–520.

3. Hershberger WA. An approach through the looking-glass. Anim Learn Behav. 1986;14: 443–451.

4. Hare TA, Camerer CF, Rangel A. Self-control in decision-making involves modulation of the vmPFC valuation system. Science. 2009;324: 646–648.

5. Hare TA, Malmaud J, Rangel A. Focusing attention on the health aspects of foods changes value signals in vmPFC and improves dietary choice. J Neurosci. 2011;31: 11077–11087.

6. Childress AR, Hole AV, Ehrman RN, Robbins SJ, McLellan AT, O'Brien CP. Cue reactivity and cue reactivity interventions in drug dependence. NIDA Res Monogr. 1993;137: 73–95.

7. Cavanagh JF, Eisenberg I, Guitart-Masip M, Huys Q, Frank MJ. Frontal theta overrides pavlovian learning biases. J Neurosci. 2013;33: 8541–8548.

8. Gershman SJ, Guitart-Masip M, Cavanagh JF. Neural signatures of arbitration between Pavlovian and instrumental action selection. PLoS Comput Biol. 2021;17: e1008553.

9. Guitart-Masip M, Huys QJM, Fuentemilla L, Dayan P, Duzel E, Dolan RJ. Go and no-go learning in reward and punishment: Interactions between affect and effect. Neuroimage. 2012;62: 154–166.

10. Rudorf S, Hare TA. Interactions between dorsolateral and ventromedial prefrontal cortex underlie context-dependent stimulus valuation in goal-directed choice. J Neurosci. 2014;34: 15988–15996.

11. Ridderinkhof KR, Ullsperger M, Crone EA, Nieuwenhuis S. The role of the medial frontal cortex in cognitive control. Science. 2004;306: 443–447.

12. Samejima K, Doya K. Multiple representations of belief states and action values in corticobasal ganglia loops. Ann N Y Acad Sci. 2007;1104: 213–228.

13. Gläscher J, Rudrauf D, Colom R, Paul LK, Tranel D, Damasio H, et al. Distributed neural system for general intelligence revealed by lesion mapping. Proc Natl Acad Sci U S A. 2010;107: 4705–4709.

14. Baumgartner T, Knoch D, Hotz P, Eisenegger C, Fehr E. Dorsolateral and ventromedial prefrontal cortex orchestrate normative choice. Nat Neurosci. 2011;14: 1468–1474.

15. MacDonald AW 3rd, Cohen JD, Stenger VA, Carter CS. Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. Science. 2000;288: 1835–1838.

16. Miller EK, Cohen JD. An integrative theory of prefrontal cortex function. Annu Rev Neurosci. 2001;24: 167–202.

17. Averbeck B, O'Doherty JP. Reinforcement-learning in fronto-striatal circuits. Neuropsychopharmacology. 2022;47: 147–162.

18. Weissengruber S, Lee SW, O'Doherty JP, Ruff CC. Neurostimulation Reveals Context-Dependent Arbitration Between Model-Based and Model-Free Reinforcement Learning. Cereb Cortex. 2019;29: 4850–4862.

19. Ly V, Bergmann TO, Gladwin TE, Volman I, Usberti N, Cools R, et al. Reduced Affective Biasing of Instrumental Action With tDCS Over the Prefrontal Cortex. Brain Stimul. 2016;9: 380–387.

20. Smittenaar P, Prichard G, FitzGerald THB, Diedrichsen J, Dolan RJ. Transcranial direct current stimulation of right dorsolateral prefrontal cortex does not affect model-based or model-free reinforcement learning in humans. PLoS One. 2014;9: e86850.

21. Ye H, Chen S, Huang D, Wang S, Luo J. Modulating activity in the prefrontal cortex changes decision-making for risky gains and losses: a transcranial direct current stimulation study. Behav Brain Res. 2015;286: 17–21.

22. Fecteau S, Knoch D, Fregni F, Sultani N, Boggio P, Pascual-Leone A. Diminishing risk-taking behavior by modulating activity in the prefrontal cortex: a direct current stimulation study. J Neurosci. 2007;27: 12500–12505.

23. Cohen Kadosh R, Soskic S, Iuculano T, Kanai R, Walsh V. Modulating Neuronal Activity Produces Specific and Long-Lasting Changes in Numerical Competence. Curr Biol. 2010;20: 2016–2020.

24. Alonzo A, Brassil J, Taylor JL, Martin D, Loo CK. Daily transcranial direct current stimulation (tDCS) leads to greater increases in cortical excitability than second daily transcranial direct current stimulation. Brain Stimul. 2012;5: 208–213.

25. Gálvez V, Alonzo A, Martin D, Loo CK. Transcranial direct current stimulation treatment protocols: should stimulus intensity be constant or incremental over multiple sessions? Int J Neuropsychopharmacol. 2013;16: 13–21.

558 26. Vehtari A, Gelman A, Gabry J. Practical Bayesian model evaluation using leave-one-out
559 cross-validation and WAIC. Stat Comput. 2017;27: 1413–1432.

560 27. Griffiths KR, Morris RW, Balleine BW. Translational studies of goal-directed action as a
561 framework for classifying deficits across psychiatric disorders. Front Syst Neurosci.
562 2014;8: 101.

563 28. Fonteneau C, Redoute J, Haesebaert F, Le Bars D, Costes N, Suaud-Chagny M-F, et al.
564 Frontal Transcranial Direct Current Stimulation Induces Dopamine Release in the Ventral
565 Striatum in Human. Cereb Cortex. 2018;28: 2636–2646.

566 29. Fukai M, Bunai T, Hirosawa T, Kikuchi M, Ito S, Minabe Y, et al. Endogenous dopamine
567 release under transcranial direct-current stimulation governs enhanced attention: a study
568 with positron emission tomography. Transl Psychiatry. 2019;9: 115.

569 30. Lloyd K, Dayan P. Safety out of control: dopamine and defence. Behav Brain Funct.
570 2016;12: 15.

571 31. Millner AJ, Gershman SJ, Nock MK, den Ouden HEM. Pavlovian Control of Escape and
572 Avoidance. J Cogn Neurosci. 2018;30: 1379–1390.

573 32. Paulus MP. Driven by Pain, Not Gain: Computational Approaches to Aversion-Related
574 Decision Making in Psychiatry. Biol Psychiatry. 2020;87: 359–367.

575 33. Dayan, Huys. Neurophysiology: Serotonin's many meanings elude simple theories. Elife.
576 2015. Available: https://elifesciences.org/articles/7390

577 34. Graeff FG, Silveira Filho NG. Behavioral inhibition induced by electrical stimulation of
578 the median raphe nucleus of the rat. Physiol Behav. 1978;21: 477–484.

579 35. Crockett MJ, Clark L, Robbins TW. Reconciling the role of serotonin in behavioral
580 inhibition and aversion: acute tryptophan depletion abolishes punishment-induced
581 inhibition in humans. J Neurosci. 2009;29: 11993–11999.

582 36. Crockett MJ, Clark L, Apergis-Schoute AM, Morein-Zamir S, Robbins TW. Serotonin
583 modulates the effects of Pavlovian aversive predictions on response vigor.
584 Neuropsychopharmacology. 2012;37: 2244–2252.

585 37. Navratilova E, Porreca F. Reward and motivation in pain and pain relief. Nat Neurosci.
586 2014. Available:
587 https://idp.nature.com/authorize/casa?redirect_uri=https://www.nature.com/articles/nn.381
588 1&casa_token=mV6mzYAzdI4AAAAA:iN9Jw421XS2oV0KY-
589 RPhvi0qotzoCn3K4pqIDqqeHXHrA8RBCySZLavOtFITL-GCF5jI0FBRkyEjWfo_9Gw

590 38. Krause B, Cohen Kadosh R. Not all brains are created equal: the relevance of individual
591 differences in responsiveness to transcranial electrical stimulation. Front Syst Neurosci.
592 2014;8: 25.

593 39. Everitt BJ, Belin D, Economidou D, Pelloux Y, Dalley JW, Robbins TW. Neural
594 mechanisms underlying the vulnerability to develop compulsive drug-seeking habits and
595 addiction. Philos Trans R Soc Lond B Biol Sci. 2008;363: 3125–3135.

596 40. Flagel SB, Robinson TE, Clark JJ, Clinton SM, Watson SJ, Seeman P, et al. An animal
597 model of genetic vulnerability to behavioral disinhibition and responsiveness to reward-
598 related cues: implications for addiction. Neuropsychopharmacology. 2010;35: 388–400.

599 41. Zilverstand A, Huang AS, Alia-Klein N, Goldstein RZ. Neuroimaging Impaired Response
600 Inhibition and Salience Attribution in Human Drug Addiction: A Systematic Review.
601 Neuron. 2018;98: 886–903.

602 42. Spechler PA, Chaarani B, Hudson KE, Potter A, Foxe JJ, Garavan H. Response inhibition
603 and addiction medicine: from use to abstinence. Prog Brain Res. 2016;223: 143–164.

604 43. Zhukovsky P, Puaud M, Jupp B, Sala-Bayo J, Alsiö J, Xia J, et al. Withdrawal from
605 escalated cocaine self-administration impairs reversal learning by disrupting the effects of
606 negative feedback on reward exploitation: a behavioral and computational analysis.
607 Neuropsychopharmacology. 2019;44: 2163–2173.

608 44. Peters J, Vega T, Weinstein D, Mitchell J, Kayser A. Dopamine and Risky Decision-
609 Making in Gambling Disorder. eNeuro. 2020;7. doi:10.1523/ENEURO.0461-19.2020

610 45. Li LM, Uehara K, Hanakawa T. The contribution of interindividual factors to variability of
611 response in transcranial direct current stimulation studies. Front Cell Neurosci. 2015;9:
612 181.

613 46. Nitsche MA, Paulus W. Sustained excitability elevations induced by transcranial DC
614 motor cortex stimulation in humans. Neurology. 2001;57: 1899–1901.

615 47. Nitsche MA, Liebetanz D, Lang N, Antal A, Tergau F, Paulus W. Safety criteria for
616 transcranial direct current stimulation (tDCS) in humans. Clin Neurophysiol. 2003;114:
617 2220–2; author reply 2222–3.

618 48. Nitsche MA, Fricke K, Henschke U, Schlitterlau A, Liebetanz D, Lang N, et al.
619 Pharmacological modulation of cortical excitability shifts induced by transcranial direct
620 current stimulation in humans. J Physiol. 2003;553: 293–301.

621 49. Im C-H, Park J-H, Shim M, Chang WH, Kim Y-H. Evaluation of local electric fields
622 generated by transcranial direct current stimulation with an extracephalic reference
623 electrode based on realistic 3D body modeling. Phys Med Biol. 2012;57: 2137–2150.

624 50. Wagner T, Fregni F, Fecteau S, Grodzinsky A, Zahn M, Pascual-Leone A. Transcranial
625 direct current stimulation: a computer-based human model study. Neuroimage. 2007;35:
626 1113–1124.

627 51. Thair H, Holloway AL, Newport R, Smith AD. Transcranial Direct Current Stimulation
628 (tDCS): A Beginner's Guide for Design and Implementation. Front Neurosci. 2017;11:
629 641.

630 52. Choi JM, Padmala S, Spechler P, Pessoa L. Pervasive competition between threat and
631 reward in the brain. Soc Cogn Affect Neurosci. 2014;9: 737–750.

632 53. Talmi D, Dayan P, Kiebel SJ, Frith CD, Dolan RJ. How humans integrate the prospects of
633 pain and reward during choice. J Neurosci. 2009;29: 14617–14626.

634 54. Lawson RP, Seymour B, Loh E, Lutti A, Dolan RJ, Dayan P, et al. The habenula encodes
635 negative motivational value associated with primary punishment in humans. Proc Natl
636 Acad Sci U S A. 2014;111: 11858–11863.

637 55. Vlaev I, Seymour B, Dolan RJ, Chater N. The price of pain and the value of suffering.
638 Psychol Sci. 2009;20: 309–317.

639 56. R Core Team. R: A Language and Environment for Statistical Computing. Vienna,
640 Austria: R Foundation for Statistical Computing; 2020. Available: https://www.R-
641 project.org/

642 57. Berger JO. Statistical Decision Theory and Bayesian Analysis. Springer Science &
643 Business Media; 2013.

644 58. Gelman A, Carlin JB, Stern HS, Rubin DB. Bayesian data analysis. Chapman and
645 Hall/CRC; 1995.

646 59. Lee MD. How cognitive modeling can benefit from hierarchical Bayesian models. J Math
647 Psychol. 2011;55: 1–7.

648 60. Ahn W-Y, Krawitz A, Kim W, Busmeyer JR, Brown JW. A Model-Based fMRI Analysis
649 with Hierarchical Bayesian Parameter Estimation. J Neurosci Psychol Econ. 2011;4: 95–
650 110.

651 61. Ahn W-Y, Haines N, Zhang L. Revealing Neurocomputational Mechanisms of
652 Reinforcement Learning and Decision-Making With the hBayesDM Package. Comput
653 Psychiatr. 2017;1: 24–57.

654 62. Stan Development Team. RStan: the R interface to Stan. 2022. Available: https://mc-
655 stan.org/

656 63. Papaspiliopoulos O, Roberts GO, Sköld M. A General Framework for the Parametrization
657 of Hierarchical Models. Stat Sci. 2007;22: 59–73.

658 64. Vehtari A, Gelman A, Simpson D, Carpenter B, Bürkner P-C. Rank-normalization,
659 folding, and localization: An improved Rˆ for assessing convergence of MCMC (with
660 discussion). Bayesian Anal. 2021;16. doi:10.1214/20-ba1221

661 65. Kruschke J. Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan. Academic
662 Press; 2014.