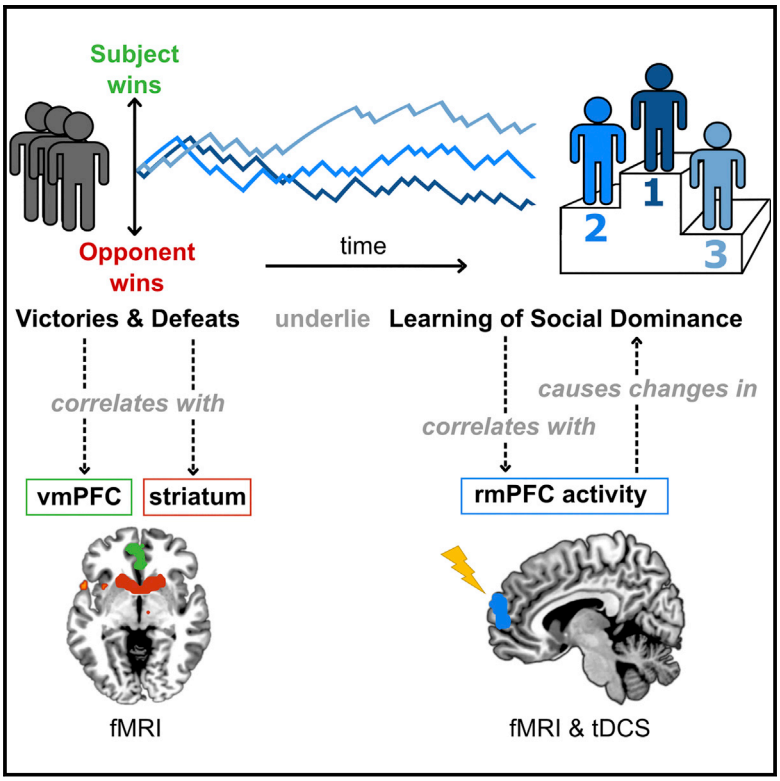
Article

Current Biology

Dynamical Representation of Dominance Relationships in the Human Rostromedial Prefrontal Cortex

Graphical Abstract



Authors

Romain Ligneul, Ignacio Obeso, Christian C. Ruff, Jean-Claude Dreher

Correspondence

r.ligneul@donders.ru.nl (R.L.), dreher@isc.cnrs.fr (J.-C.D.)

In Brief

Ligneul et al. demonstrate that BOLD activity in the rmPFC represents social dominance relationships as learned from competitive interactions, whereas the vmPFC and ventral striatum encode social victories and defeats, respectively. Electrical stimulation of the rmPFC modulates learning and updating of social dominance representations.

Highlights

- Social dominance is learned by a reinforcement-learning process
- Social dominance status and associated prediction errors are encoded in the rmPFC
- rmPFC stimulation shifts the weight of victories and defeats for dominance learning

Ligneul et al., 2016, Current Biology *26*, 1–9
December 5, 2016
http://dx.doi.org/10.1016/j.cub.2016.09.015

CellPress

Please cite this article in press as: Ligneul et al., Dynamical Representation of Dominance Relationships in the Human Rostromedial Prefrontal Cortex, Current Biology (2016), http://dx.doi.org/10.1016/j.cub.2016.09.015

Current Biology
Article

Dynamical Representation of Dominance Relationships in the Human Rostromedial Prefrontal Cortex

Romain Ligneul,[1,4,*] Ignacio Obeso,[1,5] Christian C. Ruff,[2,3] and Jean-Claude Dreher[1,3,*]
[1]Neuroeconomics, Reward and Decision Making Group, Center for Cognitive Neurosciences, CNRS, University of Lyon 1, 69675 Lyon, France
[2]Laboratory for Social and Neural Systems Research (SNS Lab), Department of Economics, University of Zurich, 8091 Zurich, Switzerland
[3]Co-senior author
[4]Present address: Motivational and Cognitive Control Group, Donders Center for Cognitive Neuroimaging, 6525 EN Nijmegen, the Netherlands
[5]Present address: Centro Integral en Neurociencias A.C. (CINAC), HM Hospitales-Puerta del Sur and CEU-San Pablo University, 28938 Madrid, Spain
*Correspondence: r.ligneul@donders.ru.nl (R.L.), dreher@isc.cnrs.fr (J.-C.D.)
http://dx.doi.org/10.1016/j.cub.2016.09.015

SUMMARY

Humans and other primates have evolved the ability to represent their status in the group's social hierarchy, which is essential for avoiding harm and accessing resources. Yet it remains unclear how the human brain learns dominance status and adjusts behavior accordingly during dynamic social interactions. Here we address this issue with a combination of fMRI and transcranial direct current stimulation (tDCS). In a first fMRI experiment, participants learned an implicit dominance hierarchy while playing a competitive game against three opponents of different skills. Neural activity in the rostromedial PFC (rmPFC) dynamically tracked and updated the dominance status of the opponents, whereas the ventromedial PFC and ventral striatum reacted specifically to competitive victories and defeats. In a second experiment, we applied anodal tDCS over the rmPFC to enhance neural excitability while subjects performed a similar competitive task. The stimulation enhanced the relative weight of victories over defeats in learning social dominance relationships and exacerbated the influence of one's own dominance over competitive strategies. Importantly, these tDCS effects were specific to trials in which subjects learned about dominance relationships, as they were not present for control choices associated with monetary incentives but no competitive feedback. Taken together, our findings elucidate the role of rmPFC computations in dominance learning and unravel a fundamental mechanism that governs the emergence and maintenance of social dominance relationships in humans.

INTRODUCTION

Social dominance hierarchies are ubiquitous in the animal kingdom and have a clear evolutionary significance, as they spontaneously emerge from competition for energetic resources and sexual partners [1, 2]. Some authors have thus argued for the existence of a human "dominance behavioral system" [3, 4] that may largely determine inter-individual differences in social behavior [5–7]. Social dominance refers to situations in which an "individual controls or dictates others' behavior, primarily in competitive situations" [8]. The concept is most frequently applied to learned relationships that are shaped by a history of victories and defeats within dyads of individuals [1], and a dominance relationship is generally considered as "established" when one individual reproducibly tries to avoid competitive encounters with another individual [9].

Epidemiological studies support the hypothesis [10] that many mental disorders such as addiction, anxiety, and depression [2, 11, 12] may result from the influence of dominance hierarchies on the brain circuitry. Understanding the brain processes underlying our ability to track dominance relationships may therefore be crucial for understanding inter-individual variability in social cognition [3] and to develop new therapeutic alternatives for several neuropsychiatric diseases [12].

In the past, human social dominance has been mostly studied using pre-established ranks depicted by perceptual cues [5, 13], thereby sidestepping the issue that social dominance is usually dynamically changing and needs to be learned. Numerous cortical and subcortical areas have been involved in social hierarchy processing [14], but it is unclear whether these brain structures indeed causally influence dominance-related behaviors and how their contributions in this respect differ. Moreover, most existing studies of social competition have used monetary incentives to energize competition, hence resulting in uncertainty about whether learning-related signals represent social dominance hierarchies or the processing of monetary reward (e.g., [15, 16]). Nonetheless, substantial evidence indicates that the rostromedial prefrontal cortex (rmPFC) is a key candidate for learning dominance relationships by experiencing victories and defeats during direct social competition. Indeed, this brain region has been found to participate in social learning [17], be engaged by various types of explicit judgments about others' resources, expertise, height, or intelligence [16, 18, 19], and be linked with cognitive functions influenced by dominance representations [20] such as imputing intentions and assessing the influence of others [17, 19, 21, 22]. In humans, the rmPFC may thus

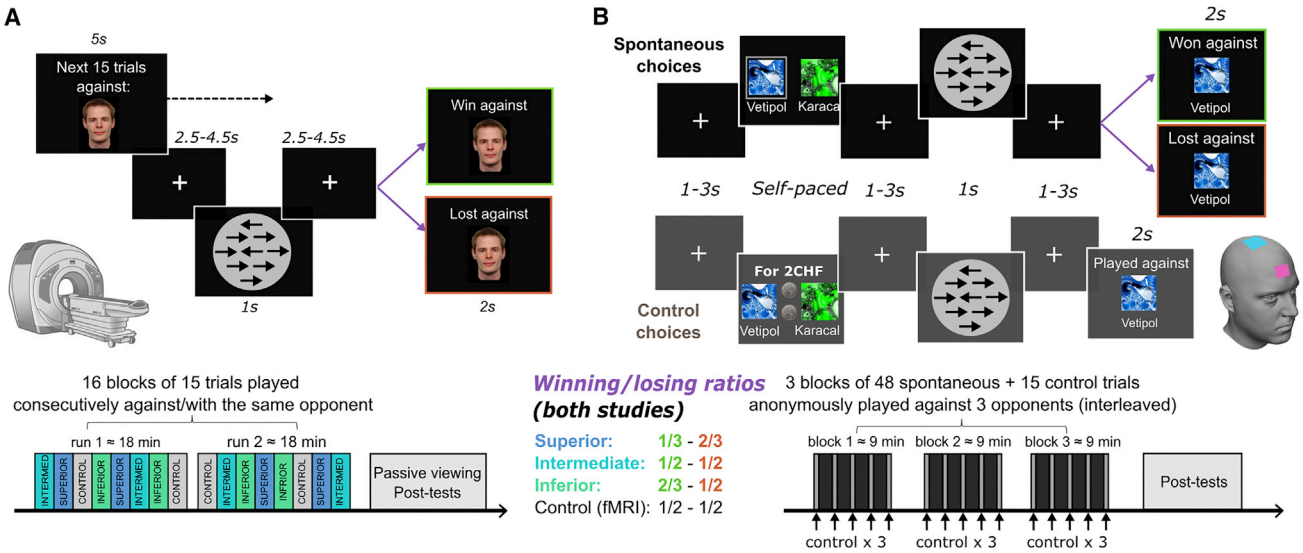


Figure 1. Experimental Procedures

(A) Time course of the fMRI experiment. During 15 trials of a "mini-block," subjects played against (or with) the same player in the competitive (or control) situation. The competitive task required subjects to evaluate a series of 46 stationary arrows, indicating in which direction the majority of these arrows pointed (left or right). The task was performed against one of three opponents who were implicitly associated with three frequencies of winning and losing. In order to succeed in the competition, subjects were instructed to answer accurately and faster than their opponents. The association between faces and winning/losing frequency was counterbalanced across subjects. In the control task, uninformative horizontal bars were displayed on the screen and joint successes/failures occurred when subjects responded with the same/opposite button as the other player.

(B) Time course of the tDCS experiment. Subjects performed a similar perceptual task but, in this experiment, opponents were marked by visual symbols and artificial names rather than a face photograph. Subjects had to choose from two alternatives of which opponent to play against (three opponents per block) in two types of trials designed to distinguish dominance-oriented (spontaneous) and reward-oriented (control) choices. In half of the subjects, we stimulated over the rmPFC for 30 min with the excitatory anodal electrode of the tDCS apparatus (magenta; the reference electrode on the vertex is in blue); in the sham group, stimulation was interrupted after only 30 s.

See also Figure S1.

mediate the interaction of high-level cognition with dynamical dominance representations, hence complementing the role of subcortical structures such as the amygdala and ventral striatum in long-term social conditioning [23, 24].

Here we combined reinforcement learning (RL), neuroimaging, and brain stimulation to demonstrate that rmPFC computations implement the learning and monitoring of social dominance relationships. We assumed that this process is instantiated by integrating competitive feedback (i.e., victories and defeats) into a flexible dominance-value representation (i.e., social dominance status; SDS) that is used to guide decision making in competitive contexts. More precisely, our RL algorithm updated an SDS_i for each specific opponent i against which the participant played, by comparing in each trial the prior SDS_i value with the observed outcome, resulting in a competitive prediction error (cPE). In cognitive terms, the SDS_i thus represents the chances of prevailing over a specific individual i. This neurocomputational architecture may thus account for the fast, context-dependent adjustment of social behaviors to the relevant dimension of the opponent's skills during social interactions. Besides assessing the role of the rmPFC in social dominance learning, our design also allowed us to identify the pattern of neural activity involved in the affective appraisal of competitive outcomes, a process most likely sculpted by a long-timescale experience with social competition [4].

In experiment 1, 28 participants underwent fMRI while they completed a perceptual decision-making task framed as measuring "cognitive efficiency" (Figure 1A; Experimental Procedures). A cover story described the experiment as an online competitive game played in real time against three other individuals. However, unbeknownst to the subjects, social victories and defeats were controlled by the computer in order to elicit an implicit skill-based hierarchy. To capture the dynamics of dominance representations from trial to trial, we regressed blood-oxygen-level-dependent (BOLD) activity on the time-resolved estimates of SDS_i and their associated cPEs. In a control condition, subjects simply had to press the same button as a fourth player. This condition allowed us to examine whether the observed neural activity was specific to situations where the subjects learned the social ranks rather than just tracking the probability of positive and negative outcomes.

As expected from our hypothesis, experiment 1 showed that the rmPFC robustly encoded SDS_i and cPEs whereas the ventral striatum and ventromedial prefrontal cortex (vmPFC) encoded competitive outcomes in a learning-independent manner. Having established this correlative evidence for the representation of SDS_i in the rmPFC, we applied anodal transcranial direct current stimulation (tDCS) or a sham stimulation over the coordinates that best encoded competitive prediction errors in our fMRI data (Figures 3A and 3B) while participants completed a similar competitive task (Figure 1B; Experimental Procedures).

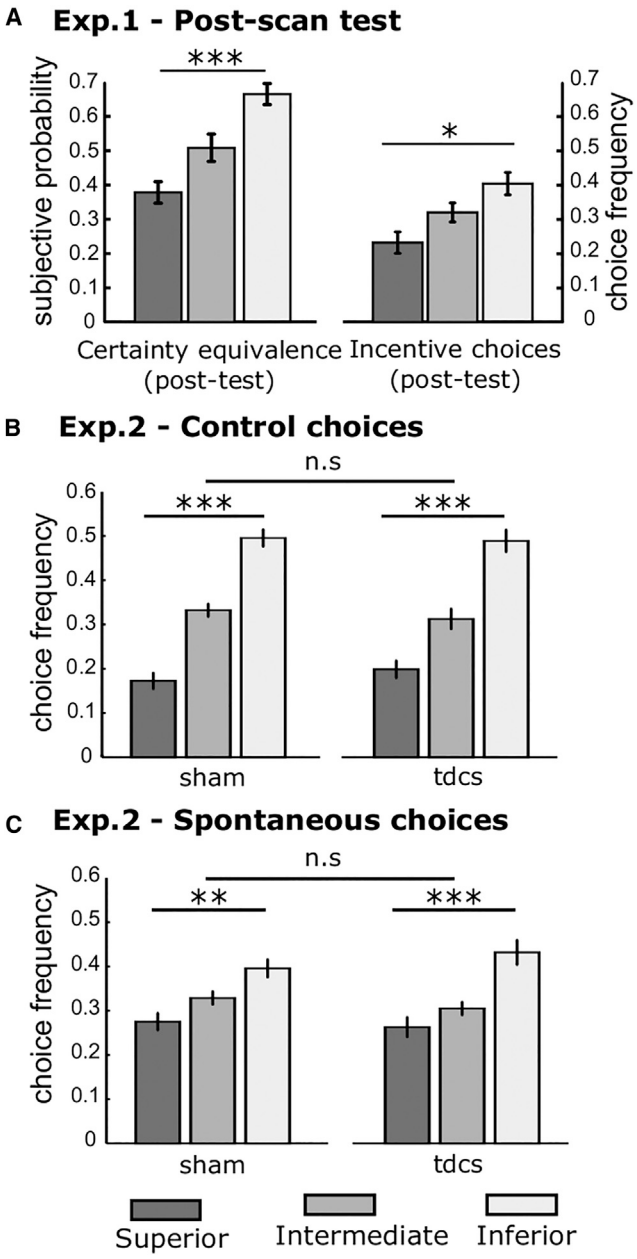


Figure 2. Learning of Social Hierarchy

(A) In experiment 1 (fMRI), learning was demonstrated by two post-scan tests in which subjects were incentivized to maximize their payoffs by using their acquired knowledge about the social hierarchy (see also Figures S1A and S1B).

(B) In experiment 2 (tDCS), the analysis of control choices showed that all subjects learned to discriminate the opponents and chose to play less frequently against the superior opponents.

(C) Results from the same analysis as in (B) for spontaneous choices. Direct comparisons between spontaneous and control choices indicated that subjects explored the dominance hierarchy more frequently when no incentives were at stake but competitive feedback was delivered.

Shaded areas and errors bars represent SEM. n.s, not significant. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. See also Figure S4.

This enabled us to assess the causal influence of rmPFC computations over decision making in competitive contexts. Indeed, experiment 2 was similar to experiment 1, except that it required the participants to choose in each trial with which opponents they wanted to compete. In the majority of trials, termed "spontaneous," no monetary rewards were associated with competitive feedback. However, a minority of "control" trials involving monetary incentives (but no competitive feedback) were also included. We used these two types of trials to demonstrate that tDCS over the rmPFC specifically alters dominance-oriented rather than general reward-oriented decision making.

RESULTS

Behavioral Evidence for Social Dominance Learning

In experiment 1 (fMRI), we assessed learning success outside the scanner after completion of the main competition task (see the Supplemental Experimental Procedures; Figures S1A and S1B). Subjects were well able to learn the ranks of the opponents from competitive feedback in the main fMRI task (Figure 2A).

In experiment 2 (tDCS), learning of social ranks could be directly estimated from the choices made in the control trials, for which subjects had to maximize monetary payoffs through rational selection of the weakest opponents [25] (Figure 2B). Every subject selected more often the weaker opponent available in each pair ($t(1,33) = 11.0$, $p < 0.001$), which validated our methodological approach and demonstrated that subjects were able to infer rapidly the skills of their competitors in the perceptual decision-making game. In the spontaneous trials of experiment 2 (Figure 2C), a similar pattern was observed, with subjects on average choosing more often to compete against the weaker opponents ($t(1,33) = 4.6$, $p < 0.001$).

Importantly, experiencing defeats during the competitive tasks led to clear-cut enhancements of perceptual performances in the next trial, which presumably reflects the cognitive engagement of our subjects and their willingness to win—or reluctance to lose—in the competitive game. Reaction times significantly decreased after a defeat, whereas accuracy remained constant, and this phenomenon was of similar amplitude in the two experiments (Figure S4). Extended analyses related to the aforementioned effects are reported in [25].

Neural Correlates of Social Dominance Learning: Experiment 1

To identify the brain regions engaged in the dynamical representation and the update of the SDS_i by cPEs, we applied an RL algorithm to the sequence of competitive choices and outcomes delivered by the computer program (see the Experimental Procedures). We assigned a value of 1 to victories and a value of 0 to defeats, so that the SDS_i variable tracked the anticipated probability of winning against each opponent i (which is high against inferior and low against superior opponents). We then used a model-based fMRI approach to regress the trial-by-trial variables derived from the algorithm on BOLD signals (general linear model 1; GLM1).

We found that the rmPFC (Montreal Neurological Institute [MNI] coordinates: [6, 59, 10]) was the only brain region that encoded both positive and negative cPEs at canonical statistical thresholds (Figures 3A and 3B; Table S1; $p < 0.05$ cluster-level

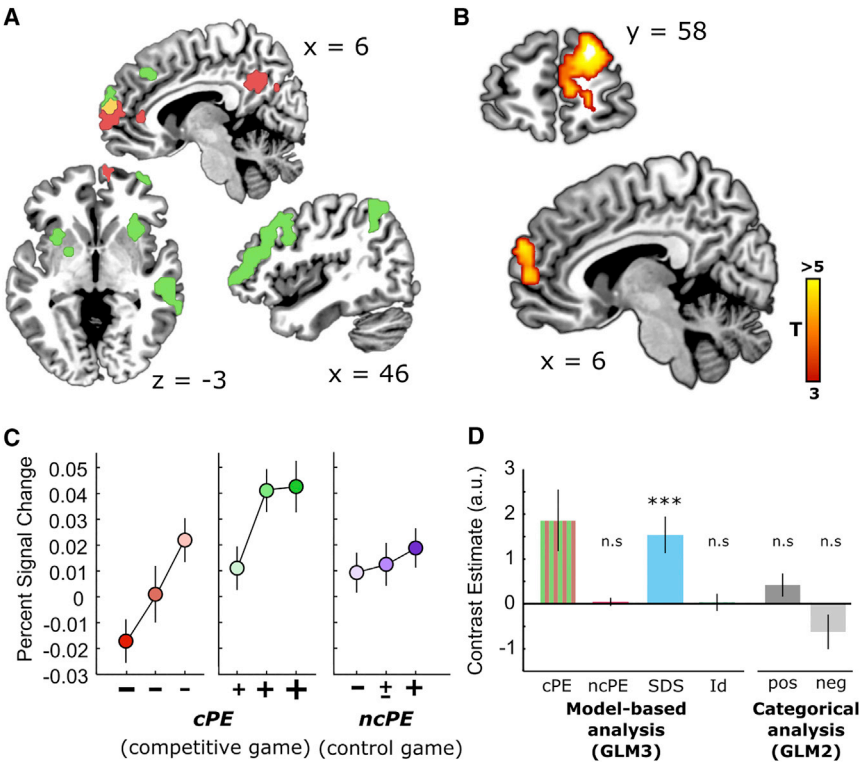


Figure 3. Encoding of Competitive Prediction Errors in the rmPFC

(A) Statistical maps ($p < 0.05$ FWE, cluster-corrected threshold $p < 0.001$) for positive competitive prediction error (cPE) (green) and negative cPE (red) overlapped in the rmPFC (yellow).

(B) The right bank of the medial rmPFC was the only region to encode cPEs of both affective valences.

(C) Percent signal changes from the conjunction cluster shown in (B) showed that rmPFC activity increased with the size of the cPE (green, positive; red, negative; purple, non-competitive).

(D) An ROI analysis showed that the activity changes observed in the rmPFC reflected both the cPEs and the anticipated opponents' social dominance status (SDS_i) (GLM3, light blue). It also demonstrated that the rmPFC did not reflect winning/losing per se, the identity of the opponent, or interactions of these two factors (i.e., no significant contrast [win against superior/inferior] > [lose against inferior/superior], dark gray; vice versa, light gray). ***$p < 0.001$.

Error bars and shaded areas represent SEM. See also Figure S5 and Table S1.

corrected for multiple comparisons, initial height threshold $p < 0.001$). Re-estimating the hemodynamic responses separately for three different intensities of cPE showed that rmPFC activity encoded a "signed" prediction error (Figure 3C), with more positive signals in response to more positive cPEs (i.e., victories against opponents currently predicted to be superior) and more negative signals in response to more negative cPEs (i.e., defeats against opponents predicted to be inferior).

In order to ascertain that the rmPFC monitored *dynamical* fluctuations in SDS_i rather than overall performance or identity of opponents, we ran two complementary analyses. First, a new GLM of simple interactions between outcome valence (i.e., win or loss) and opponent categories (i.e., superior, intermediate, or inferior) could not explain BOLD activity in the rmPFC (Figure 3D; GLM2). Second, a new GLM that included outcome valence, opponent category, and cPE still revealed a clear effect for competitive prediction errors, although all of the variance related to opponent category and outcome valence had already been explained by the other two parametric regressions (Figure 3D; GLM3). The separate analysis of victory- and defeat-related prediction errors showed that they were indeed both represented in the rmPFC (Figure 3A, yellow) but also engaged partly distinct neural networks. Negative cPEs were additionally encoded in the posterior cingulate cortex (PCC; Figure 3A, red), whereas positive cPEs also recruited a frontoparietal network, the bilateral anterior insula, and a right medial temporal lobe region (Figure 3A, green; Table S1).

Importantly, the rmPFC not only encoded cPEs during the outcome phase but also the SDS_i at the decision-making stage (MNI coordinates: [48, 8, 25]; Figure 3D, SDS blue graph bar; region-of-interest (ROI) analysis within the cluster of Figure 3B: $t(1,27) = 3.8$, $p < 0.001$). A similar SDS_i representation was also found in the right dorsolateral prefrontal cortex (Table S1). This observation is consistent with the notion that the rmPFC not only encodes competitive prediction errors but also anticipates the probabilities of victories and defeats during the choice itself, a process needed to drive adaptive decision making in socio-competitive contexts. Importantly, we found no evidence for a significant encoding of non-competitive prediction errors (ncPEs) ($t(1,27) = 0.32$, $p = 0.75$; Figures 3C and 3D), reaction times, or accuracy in the rmPFC (Figures S5A–S5C). These control analyses excluded that the rmPFC merely monitored the frequency of positive/negative outcomes and associated reward prediction errors, or that perceptual performance mediated the effects observed here. Moreover, the absence of correlation between cPE and ncPE signals ($r = 0.09$, $p = 0.63$) further suggested that the rmPFC (Figure 3B) did not encode a domain-general reward prediction error signal but instead a prediction error specific for competitive interactions.

Finally, we investigated the affective processing of victories and defeats using a non-parametric approach (GLM2) to compare the BOLD responses elicited by social feedback in the intermediate competitive and non-competitive conditions (maximally matched in terms of outcome frequencies and visual stimulation). A whole-brain analysis revealed a significant interaction between task type (compete/coordinate) and outcome valence (positive/negative) in the ventral striatum (Figure 4A), ventromedial prefrontal cortex (Figure 4B), and, to a lesser extent, amygdala (Figure 4C) and midbrain (Table S5). Compared to control failures, social defeats de-activated more the ventral striatum, amygdala, and midbrain. By contrast, the vmPFC was the only region to respond with increased BOLD signals when subjects experienced competitive victories, as compared to control successes. Subsequent analyses revealed

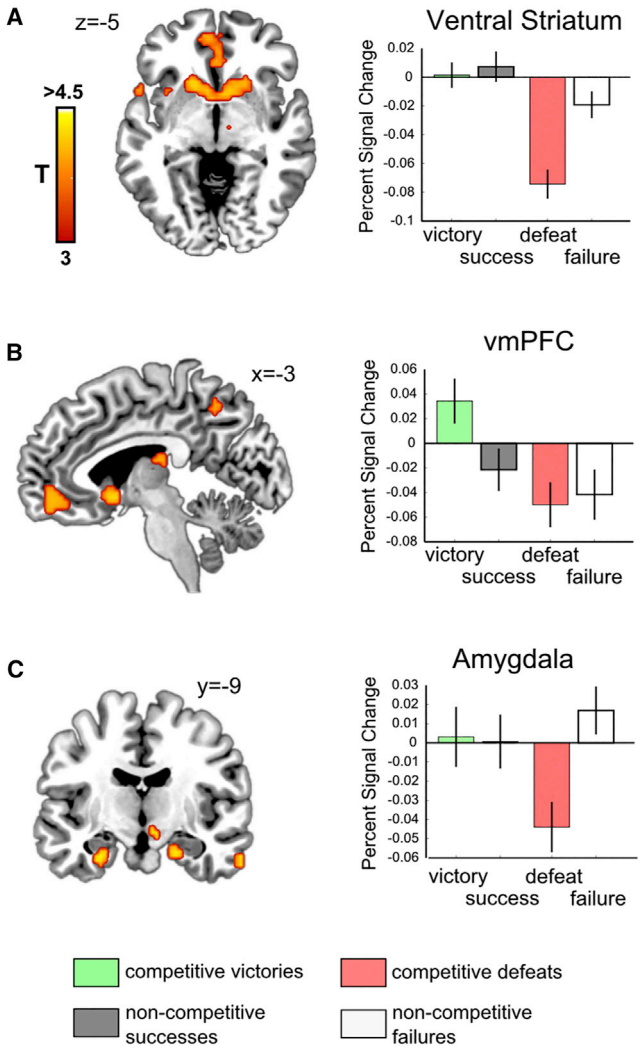


Figure 4. Interaction between Positive versus Negative Outcomes and Collaborative versus Competitive Games

Competition-specific outcome signals from the intermediate competitive condition (50% victories) as compared to successes/failures from the non-competitive condition (50% successes). The interaction contrast "competitive victory and control failure > competitive defeat and control success" yielded significant results ($p < 0.05$ FWE, cluster-corrected threshold $p < 0.001$) in the (A) bilateral ventral striatum (MNI coordinates: [−15, 17, −2]), (B) vmPFC (MNI coordinates: [−3, 47, −11]), and (C) bilateral amygdala (MNI coordinates: [−24, −19, −17] and [21, −7, −26]; small-volume correction [SVC], 8 mm sphere) and midbrain (MNI coordinates: [6, −13, −11]; SVC within the brainstem mask). Graphs represent the percentages of signal change for competitive victories (filled green bars) and defeats (filled red bars) in the competitive social task. Gray and empty bars denote successes and failures in the control condition, respectively. Error bars represent SEM. See also Table S2.

a significant relationship between the personality trait of behavioral inhibition [26] (BIS-BAS scales) and the striatal sensitivity to competitive outcomes, suggesting that those BOLD signals relate to the affective processing of social defeats (see Figure S5D). Moreover, the analysis of these general, prediction-independent affective responses to competitive outcomes yielded no significant difference in the rmPFC (Table S2).

The Impact of rmPFC Electrical Stimulation on Social Dominance Learning: Experiment 2

In the homologous task performed under tDCS or sham stimulation (Figure 1B), we accounted for the dynamic and progressive emergence of dominance representations throughout each block by means of a reinforcement-learning scheme that included a decision rule (softmax policy; Figures 5A and 5B; Supplemental Experimental Procedures). Bayesian group comparison of alternative RL models (Figure S2; Table S3) provided strong evidence for a dual-learning rate model that differentiated the weights of victories and defeats in the update of the opponent-specific SDSi values (see also [25]). Interestingly, only the best model accounting for spontaneous choices ignored the (private) information about response correctness, so that learning was based only on the information jointly available to the participant and his opponent. Such intersubjective consistency (or symmetry) of SDSi values is crucial for a learned variable that should be common to both members of a dyad because it indexes the status of an inter-personal relationship. Moreover, the degrees of learning asymmetry (i.e., learning rates for victory, *minus* those for defeats) associated with each type of choice were not correlated ($r = -0.12$, $p = 0.51$), which further confirmed that SDS learning differed from reward-based learning in our task. As expected, the analysis of inverse temperature parameters also showed that choices were less stochastic when an incentive was at stake as compared to spontaneous choices (paired-sample t test: $t(1,33) = 3.97$, $p < 0.001$; Figure S3A), but tDCS had no influence on this parameter (main effect: $F(1,29) = 0$, $p = 0.98$; interaction with trial type: $F(1,29) = 1.67$, $p = 0.21$).

In order to test our main hypothesis that the rmPFC regulates social dominance learning, we investigated the effect of rmPFC stimulation on the parameters that best accounted for choices in spontaneous and control trials. More specifically, given that fMRI analyses revealed that cPEs of positive and negative valence were encoded in different directions (Figure 3C), we expected that the unidirectional (excitatory) tDCS would enhance positive cPEs but actually reduce negative cPEs. This is because anodal tDCS is thought to enhance resting potentials of the stimulated neurons, thereby facilitating the triggering of action potentials by excitatory post-synaptic potentials but reducing the impact of inhibitory post-synaptic potentials on neuronal firing (see also [25]).

We first tested these predictions with a GLM that included stimulation type (i.e., sham/tDCS) as a fixed, between-subject effect. Two within-subject factors were also included as repeated measures: positive versus negative learning rates and spontaneous versus control condition (GLM-A; Table S4). A three-way interaction of stimulation condition, trial type, and valence indicated that the effects of tDCS on spontaneous and control choices differed significantly ($F(1,29) = 6.96$, $p = 0.013$). Thus, we separated the analysis for each trial type. This analysis showed that, in the spontaneous choices, electrical stimulation strongly interacted with the learning weights associated with victories and defeats in the spontaneous trials (Figure 5C, left; Table S5; $F(1,29) = 13.5$, $p < 0.001$; GLM-B; Table S4). As predicted, this interaction showed that tDCS enhanced the weight of victories and diminished the weight of defeats for the dynamical representation of SDS. Gender did not affect learning rates

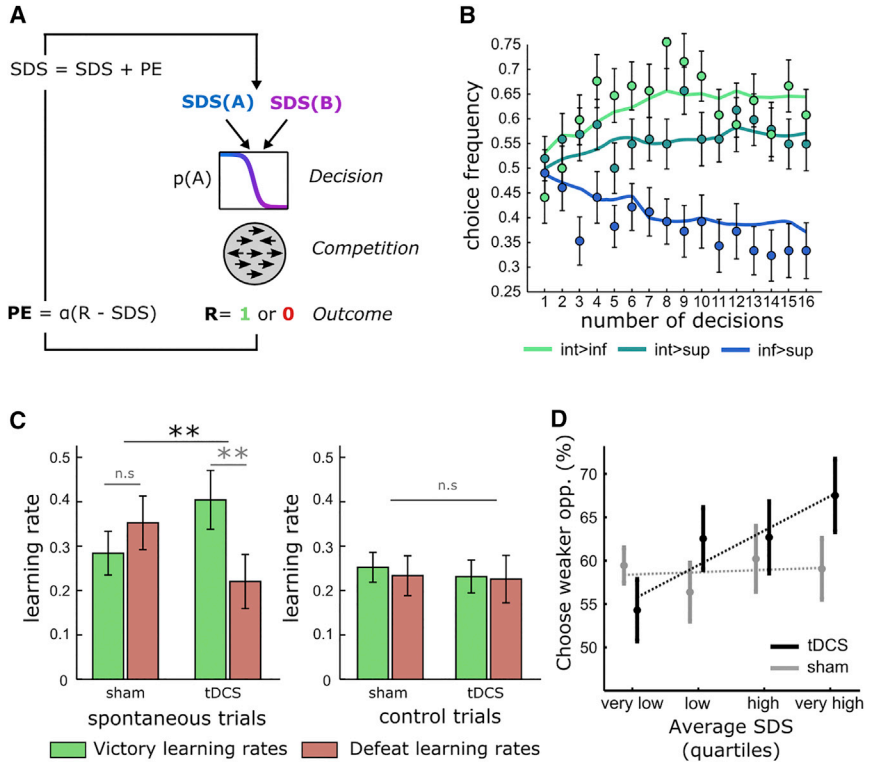


Figure 5. Reinforcement-Learning Modeling and Effects of tDCS on Social Dominance Learning

(A) Overview of the computational model. The reinforcement-learning (RL) algorithm assumes that decisions are taken probabilistically (softmax policy) according to the social dominance status attributed to each opponent i. Once the competition has occurred, the SDS_i value of the selected opponent is updated for the next trial, in proportion to the prediction error elicited by the outcome (cPE = outcome R − SDS_i), multiplied by the learning rate α.

(B) Observed choices (dots) and modeled choices (lines) in the spontaneous trials (merged across sham and tDCS groups for display purposes). Dark blue, blue-green, and green lines represent the probability of preferring the intermediate over the inferior opponent, the intermediate over the superior opponent, and the inferior over the superior opponent, respectively.

(C) In spontaneous trials (left), learning rates related to defeats and victories were balanced in the sham group. However, anodal tDCS over the rmPFC induced a significant imbalance in the learning rates, with more weight placed on victories and less weight placed on defeats. This effect did not occur in control trials (right). **$p < 0.01$.

(D) Significant interaction between the averaged SDS and stimulation group (black, tDCS; gray, sham) in the proneness to select the weaker opponent. Subjects under rmPFC tDCS alternated between periods in which they challenged superior opponents despite increased frequency of defeats (low averaged SDS) and periods in which they consistently challenged inferior opponents despite having already established their dominance (high averaged SDS).

Error bars represent SEM. See also Figures S2–S4 and Tables S3–S5.

in the spontaneous trials and did not interact with the corresponding tDCS effects ($F(1,29) = 0.01$, $p = 0.92$ and $F(1,29) = 0.3$, $p = 0.87$, respectively). Crucially, we observed no significant influence of electrical stimulation over control choices that were driven by the motivation to maximize monetary reward (Figure 5C, right; GLM-C; Table S4; stimulation condition by valence: $F(1,29) = 0.01$, $p = 0.9$; main effect: $F(1,29) = 0.01$, $p = 0.93$). This demonstrated that tDCS acted specifically on the choices that had an impact on the emergence of dominance relationships.

Complementing this key finding, we observed that the degree of learning asymmetry (corrected for tDCS effects) was correlated with the internal locus of control (iLOC) scale across subjects, so that subjects with a higher sense of control weighted victories more and defeats less on average ($r = 0.41$, $p = 0.016$; Figure S3C). Interestingly, in the sham group, people scoring low on the iLOC scale exhibited a marked over-weighting of defeats as compared to victories, which was absent in the tDCS group (Figure S3B). Yet it is important to stress that the effect of tDCS on learning rates did not interact with this personality scale.

Because tDCS induced an asymmetry in the learning rates, favoring victory- over defeat-based learning in the spontaneous trials, modeled SDS_i values were logically higher in the tDCS group ($t(1, 32) = 2.77$, $p = 0.009$). Therefore, we hypothesized that rmPFC-stimulated subjects would be less affected by the experience of subordination but also more prone to establish their dominance repeatedly, despite having already experienced their dominance. To test this prediction without introducing biases, we used the learning parameters estimated in control trials and averaged the SDS_i values (A-SDS) over all opponents i, resulting in a global estimate of subjects' own dominance on a trial-by-trial basis (higher when subjects are on average more dominant). Then, we investigated whether this averaged SDS variable influenced behavior differently in the sham/tDCS groups. In line with our expectations, the effect of A-SDS on the propensity to select the weaker opponent interacted with the stimulation condition ($F(3,87) = 3.2$, $p = 0.029$) (Figure 5D). An effect of A-SDS was clearly observed in the tDCS group ($F(3,42) = 3.6$, $p = 0.02$; linear trend: $F(1,14) = 11.5$, $p = 0.004$), but not in the sham group ($F(3,42) = 0.5$, $p = 0.66$). The choices of the subjects in the tDCS group were thus more influenced by the representation of their own (averaged) dominance status in the task, hence translating into a lower dependence of their choices on the victories and defeats in the immediately preceding trials (Figure S3D). Besides, the difference in selection of weak opponents between different A-SDS levels (fourth minus first quartile) correlated positively with the degree of asymmetry in learning rates ($r = 0.37$, $p = 0.03$), confirming that this phenomenon derived from the effect of tDCS on learning rates.

DISCUSSION

We investigated the neural circuitry underlying the learning of dominance asymmetries during social competition. Neural

activity in the rmPFC correlated with competitive prediction errors, a teaching signal necessary to learn dominance representations from social defeats and victories in the absence of external incentives. More importantly, enhancing neural excitability with anodal tDCS modulated the learning parameters that governed the update of dominance representations and enhanced the influence of participants' dominance on competitive choices, whereas perceptual performances and choices driven by monetary incentives were unaffected. Taken together, our findings demonstrate for the first time that the rmPFC is causally involved in learning and monitoring social dominance, a pivotal feature of our relationship with others.

The rmPFC activation reported here was located in Brodmann area 10, one of the latest portions of the prefrontal cortex to evolve in primates [27], which has been linked with cognitive processes required to navigate complex social environments such as mentalizing [28], forming first impressions about others [29], and updating beliefs about others [17, 19, 30]. Recently, a number of neuroimaging studies showed that the rmPFC is more specifically engaged by competitive contexts or comparative judgments than collaborative contexts or non-comparative judgments [15, 16, 31, 32]. Medial PFC lesions in humans can impair correct attributions of dominance from the verbal description of social relationships [33], and a recent study indicated that individuals scoring higher on a dominance personality scale performed better in a social learning task known to rely on rmPFC computations [34].

The peak coordinates of cPE encoding in our study lie exactly at the boundary of two regions known to encode representations relating oneself and others [28, 30]. This area is therefore ideally located to represent an inter-personal variable integrating both types of information. That tDCS increased the influence of an ego-centered dominance variable (A-SDS) over participants' choices is also consistent with studies showing that the sense of control or "agency"—a construct close to the dominance concept [8]—also depends on rmPFC computations [35–37]. Supporting the idea that rmPFC activity is sensitive to the sense of control in social context, the relative weights of victories and defeats correlated with the iLOC scale [38] across subjects, as though tDCS seemed to "re-balance" social dominance learning in low-iLOC subjects (Figure S3B). Such a phenomenon might be of clinical interest because low iLOC scores have also been linked to lower stress resilience [38] and to dysfunctional dominant-subordinate relationships at work [39]. More particularly, tDCS may help reduce the risk of aggressive disorders in some of these subjects [40, 41].

Although our results strongly suggest that the rmPFC plays a central and causal role in the emergence of social dominance based on competitive interactions, neuroimaging studies have also highlighted the importance of other brain regions in the appraisal of social rank [42]. For example, the ventral striatum is recruited in social tasks delivering feedback about one's own monetary payoff relative to that of another individual [5, 15, 16, 31]. The amygdala has also been highlighted as a key structure for learning social hierarchies by observation [24] rather than by direct competition (as in the current experiments). Although we found no learning-related signals in the ventral striatum or amygdala, these two regions differentiated social victories and defeats, and an ROI analysis indicated that the amplitude of defeat-related de-activations in the striatum correlated positively with a behavioral inhibition scale [26] often viewed as reflecting more subordinate personality profiles [4] (Figure S5D). This correlation suggests that social experience may shape brain responses to competitive feedback on a much longer timescale than a single fMRI session. Accordingly, two studies showed that the ventral striatal responses to social status cues depended on the subjective socio-economic status [7] and culture of subjects [6]. Neural plasticity within the dopaminergic system may underlie this phenomenon, as animal research showed that chronic dominance translates into higher expression of pre-synaptic dopamine receptor type 2 (D2) receptors [43] whereas chronic subordination involves enhanced dopamine receptor type 1 (D1) activity in this subcortical structure [23].

In conclusion, our fMRI-tDCS approach cross-validates the key involvement of rmPFC activity in social dominance learning and sheds light on the computational principles governing this process, which is distinct from the establishment of behavioral inhibition and avoidance in the course of long-term dominance relationships (for additional discussion of these findings, see [25]). Contributing to action selection in a flexible and context-dependent way [44], the rmPFC might be necessary to assess dominance relationships based on limited competitive feedback and to decide rapidly whether one should adopt a dominant or subordinate social role, depending on how social interactions unfold in specific contexts. A better understanding of the interplay of subcortical and rmPFC computations might eventually open the path to new therapeutic interventions tailored to reduce dysfunctional imbalances in dominance-related behaviors as seen, for example, in depression, psychopathy, social anxiety, and pathological aggression [3].

EXPERIMENTAL PROCEDURES

For a full description of cover stories, experimental designs, model fitting, model comparison, complementary tasks, and the exact procedures used for fMRI and tDCS, please see the Supplemental Experimental Procedures. All participants met the inclusion criteria (right-handedness, no history of psychiatric or neurologic disorders, Caucasian ethnicity) and gave written, informed consent.

Experiment 1: fMRI

Experimental Design: Overview

Twenty-eight men (mean age: 22 years; age range: 18–27 years) entered the scanner and were trained on the perceptual decision-making task during 100 trials, which included (non-social) feedback about their accuracy. The fMRI experiment itself was a social-competition task divided into two runs of 18 min each, followed by a passive viewing test and two post-tasks aiming at assessing learning in the main task (Figure 1A). Participants were confronted with three different opponents (plus one control player in the non-competitive condition). Trials were organized into mini-blocks (four per condition) during which participants played 15 trials in a row against/with the same player, whose face was presented before starting the 15 trials and at each outcome. The subjects were told that they interacted in real time with the other individuals, and received the following instructions: "If both responses are correct, the fastest player wins. If one player gives an incorrect response, the accurate player wins. If both responses are incorrect, the slowest player wins. If one player doesn't respond, then he loses automatically."

Unbeknownst to the subjects, outcomes were covertly manipulated to induce the three different competitive conditions (33%, 50%, or 66% of victories) or the control condition (50% of cooperative successes). This latter condition was a simplified coordination game in which subjects viewed horizontal

lines instead of arrows and had to respond by pressing the left or right button without any perceptual judgment. They were told that the trial would result in a shared win or loss if both players responded in the same direction or opposite directions, respectively.

Following the functional imaging session, we invited subjects to complete two complementary tasks designed to evaluate the avoidance of dominant individuals and the ability to recall the implicit hierarchy learned during the fMRI experiment (Figures S1A and S1B). The protocol was approved by the local ethics committee (Comité de Protection des Personnes – Sud-Est 2, Lyon, France).

fMRI Analyses: Overview

To capture the dynamics of the social learning process elicited by our implicit hierarchy manipulation, we applied a Rescorla-Wagner rule (learning rate 0.1; see also Figure 5A and Equation 1 below) to model the sequence of victories and defeats associated with each competitive condition. This model-based fMRI approach was used to probe the neural substrates of cPEs and SDS_i by regressing brain activity on the trial-by-trial variables derived from our RL algorithm (GLM1). We focused this parametric analysis on positive and negative cPEs (at the outcome stage), their non-competitive counterparts (ncPE; outcome stage), and the momentary dominance status against the current opponent (SDS_i; perceptual decision stage).

A second GLM was set up to uncover brain activity that was categorically associated with different types of outcomes (GLM2). This GLM included no parametric regressors, and fractionated competitive outcomes into six regressors (i.e., victories and defeats against superior, intermediate, and inferior opponents). A third GLM similar to GLM1 was used to ascertain that the encoding of prediction errors was not confounded by baseline effects or opponent identities (GLM3). In this model, competitive outcomes were modeled using a single categorical regressor to which three parametric regressors were added in the following order: opponent category (superior, intermediate, or inferior), outcome valence (victory or defeat), and cPE. The serial orthogonalization performed by SPM8 (http://www.fil.ion.ucl.ac.uk/spm) ensured that cPE effects could not be explained by the first two variables. So that confounds in the interpretation of cPE signals could be avoided, reaction times were always added as a first parametric regressor when modeling the onset of the perceptual stage.

Statistical analyses were performed using a conventional two-level random-effects approach with SPM8 (http://www.fil.ion.ucl.ac.uk/spm). All GLMs included regressors for motor responses, pauses between mini-blocks, as well as for the six motion parameters estimated from the realignment step. Statistical inference was performed at a standard threshold of $p < 0.05$, family-wise error (FWE) cluster-level corrected for multiple comparisons, with an initial cluster-forming threshold of $p < 0.001$.

Experiment 2: tDCS of the rmPFC while Learning Dominance Relationships

Experimental Design

Thirty-four healthy subjects (mean age: 22.1 years; 17 males) were told they would compete against each other anonymously. The protocol was approved by the ethics committee of the Canton of Zurich. As in experiment 1, the experiment began with 100 training trials with accuracy feedback. During the competitive task, outcomes against all triplets of opponents were again controlled covertly in order to reach 33%, 50%, and 66% of victories against the subject, thereby creating an implicit social hierarchy composed of an inferior, an intermediate, and a superior opponent.

In each block, subjects were told they would play against three new opponents (depicted by fractals and nicknames) randomly and anonymously chosen among the other players in the room. Each block was composed of 48 "spontaneous choices" in which there was no money at stake and in which feedback (win, even, or lose) was given to the subjects. Additionally, three "control choices" without feedback had to be performed at the beginning of each block and then after each series of 12 spontaneous choices. For those trials only, we instructed subjects that 2 Swiss Francs (CHF) would be added to the initial endowment of the winning player (40 CHF) and that the same amount of money would be removed from the endowment of the losing player. Time courses were the same as in the spontaneous trials, except that non-informative feedback was delivered in order to avoid additional learning effects. Finally, for all trial types, only two opponents were proposed at the choice stage; subjects were thus told that the opponent they would not select would automatically compete against the third (undisplayed) player engaged in the block.

Reinforcement-Learning Modeling

In order to capture the dynamic influence of victories and defeats on the competitive choices of the participants and to estimate the learning rates associated with the two stimulation regimes, we used an RL algorithm. This assumed that the probability of choosing to defy one opponent i over another opponent j depended on the relative difference in their social dominance statuses SDS_i and SDS_j.

$$SDSi(t+1) = SDSi(t) + \alpha * (R - SDSi(t)) \quad \text{(Equation 1)}$$

$$p(i) = \frac{\exp(\beta * SDS_i)}{\exp(\beta * SDS_i) + \exp(\beta * SDS_j)} \quad \text{(Equation 2)}$$

Equation 1 determines how the social dominance status of the chosen opponent i (SDS_i) is updated according to the feedback received. R is the "reward" collected in the ongoing trial, which was arbitrarily set to 0 for defeats, 0.5 for evens, and 1 for victories. This means that higher competitive values in any given trial correspond to weaker opponents (against whom the subject recently "harvested" victories), and vice versa. Constrained between 0 and 1, the free parameter α represents the learning rate of the model (high α implies high volatility in the representation of values) and could differ for victories and defeats, hence offering the possibility to learn asymmetrically. Equation 2 defines a stochastic decision rule (softmax) that calculates the probability of choosing the opponent i given the other available opponent j. The free parameter β is the inverse temperature parameter, and dictates to what extent the decision is deterministic relative to the values of the available options ($\beta \in \mathbb{R}$). We tested six variants of our reinforcement-learning scheme, testing for three degrees of self-performance monitoring (no monitoring, accuracy, and accuracy and response speed) as well as single- versus dual-learning rate applied to the update of opponent's values after victories and defeats. Alternative models were compared using Bayesian group comparison (Figure S2). For each trial type, only the model presenting the highest exceedance probability was analyzed.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and five tables and can be found with this article online at http://dx.doi.org/10.1016/j.cub.2016.09.015.

AUTHOR CONTRIBUTIONS

R.L. designed, programmed, and performed the experiments, analyzed the data, and wrote the paper. I.O. designed and performed the tDCS experiment and wrote the paper. C.C.R. designed the tDCS experiment and wrote the paper. J.-C.D. designed the experiments and wrote the paper.

ACKNOWLEDGMENTS

This work was performed within the framework of the Laboratory of Excellence (LABEX; grant ANR-11-LABEX-0042 to J.-C.D.) of the Université de Lyon within the program "Investissement d'Avenir" (ANR-11-IDEX-0007) and a grant from the European Institute for Advanced Study Fellowship Program at the Hanse-Wissenschaftskolleg to J.-C.D. This work was also supported by grants from the Agence Nationale pour la Recherche (ANR "Brain Choice" 14-CE13-0006; to J.-C.D.) and grants of the SNSF (105314_152891, CRSII3_141965, and 51NF40_144609; to C.C.R.). We thank R. Girard, the staff of CERMEP-Imagerie du Vivant, and the staff of the SNS Lab at the University of Zurich for helpful assistance with data collection.

Received: March 31, 2016
Revised: August 16, 2016
Accepted: September 8, 2016
Published: October 27, 2016

REFERENCES

1. Bernstein, I.S. (1981). Dominance: the baby and the bathwater. Behav. Brain Sci. *4*, 419–429.
2. Sapolsky, R.M. (2005). The influence of social hierarchy on primate health. Science *308*, 648–652.
3. Johnson, S.L., Leedom, L.J., and Muhtadie, L. (2012). The dominance behavioral system and psychopathology: evidence from self-report, observational, and biological studies. Psychol. Bull. *138*, 692–743.
4. Keltner, D., Gruenfeld, D.H., and Anderson, C. (2003). Power, approach, and inhibition. Psychol. Rev. *110*, 265–284.
5. Zink, C.F., Tong, Y., Chen, Q., Bassett, D.S., Stein, J.L., and Meyer-Lindenberg, A. (2008). Know your place: neural processing of social hierarchy in humans. Neuron *58*, 273–283.
6. Freeman, J.B., Rule, N.O., Adams, R.B., Jr., and Ambady, N. (2009). Culture shapes a mesolimbic response to signals of dominance and subordination that associates with behavior. Neuroimage *47*, 353–359.
7. Ly, M., Haynes, M.R., Barter, J.W., Weinberger, D.R., and Zink, C.F. (2011). Subjective socioeconomic status predicts human ventral striatal responses to social status information. Curr. Biol. *21*, 794–797.
8. Wilson, R.A., and Keil, F.C. (1999). The MIT Encyclopedia of the Cognitive Sciences (MIT Press).
9. Rowell, T.E. (1974). The concept of social dominance. Behav. Biol. *11*, 131–154.
10. Price, J. (1967). The dominance hierarchy and the evolution of mental illness. Lancet *290*, 243–246.
11. Sandi, C., and Haller, J. (2015). Stress and the social brain: behavioural effects and neurobiological mechanisms. Nat. Rev. Neurosci. *16*, 290–304.
12. Sloman, L., and Gilbert, P., eds. (2000). Subordination and Defeat: An Evolutionary Approach to Mood Disorders and Their Therapy (Lawrence Erlbaum Associates).
13. Santamaría-García, H., Burgaleta, M., and Sebastián-Gallés, N. (2015). Neuroanatomical markers of social hierarchy recognition in humans: a combined ERP/MRI study. J. Neurosci. *35*, 10843–10850.
14. Amat, J., Aleksejev, R.M., Paul, E., Watkins, L.R., and Maier, S.F. (2010). Behavioral control over shock blocks behavioral and neurochemical effects of later social defeat. Neuroscience *165*, 1031–1038.
15. Kätsyri, J., Hari, R., Ravaja, N., and Nummenmaa, L. (2013). The opponent matters: elevated fMRI reward responses to winning against a human versus a computer opponent during interactive video game playing. Cereb. Cortex *23*, 2829–2839.
16. Bault, N., Joffily, M., Rustichini, A., and Coricelli, G. (2011). Medial prefrontal cortex and striatum mediate the influence of social comparison on the decision process. Proc. Natl. Acad. Sci. USA *108*, 16044–16049.
17. Behrens, T.E.J., Hunt, L.T., Woolrich, M.W., and Rushworth, M.F.S. (2008). Associative learning of social value. Nature *456*, 245–249.
18. Lindner, M., Hundhammer, T., Ciaramidaro, A., Linden, D.E.J., and Mussweiler, T. (2008). The neural substrates of person comparison—an fMRI study. Neuroimage *40*, 963–971.
19. Boorman, E.D., O'Doherty, J.P., Adolphs, R., and Rangel, A. (2013). The behavioral and neural mechanisms underlying the tracking of expertise. Neuron *80*, 1558–1571.
20. Galinsky, A.D., Magee, J.C., Inesi, M.E., and Gruenfeld, D.H. (2006). Power and perspectives not taken. Psychol. Sci. *17*, 1068–1074.
21. Hampton, A.N., Bossaerts, P., and O'Doherty, J.P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. Proc. Natl. Acad. Sci. USA *105*, 6741–6746.
22. Yoshida, W., Seymour, B., Friston, K.J., and Dolan, R.J. (2010). Neural mechanisms of belief inference during cooperative games. J. Neurosci. *30*, 10744–10751.
23. Hollis, F., van der Kooij, M.A., Zanoletti, O., Lozano, L., Cantó, C., and Sandi, C. (2015). Mitochondrial function in the brain links anxiety with social subordination. Proc. Natl. Acad. Sci. USA *112*, 15486–15491.
24. Kumaran, D., Melo, H.L., and Duzel, E. (2012). The emergence and representation of knowledge about social and nonsocial hierarchies. Neuron *76*, 653–666.
25. Ligneul, R., Obeso, I., Ruff, C.C., and Dreher, J.-C. (2016). Dynamical representation of dominance relationships in the human medial prefrontal cortex: additional information. http://dx.doi.org/10.6084/m9.figshare.3811365.v2.
26. Carver, C.S., and 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.
27. Allman, J., Hakeem, A., and Watson, K. (2002). Two phylogenetic specializations in the human brain. Neuroscientist *8*, 335–346.
28. Denny, B.T., Kober, H., Wager, T.D., and Ochsner, K.N. (2012). A meta-analysis of functional neuroimaging studies of self- and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. J. Cogn. Neurosci. *24*, 1742–1752.
29. Mende-Siedlecki, P., Cai, Y., and Todorov, A. (2013). The neural dynamics of updating person impressions. Soc. Cogn. Affect. Neurosci. *8*, 623–631.
30. Nicolle, A., Klein-Flügge, M.C., Hunt, L.T., Vlaev, I., Dolan, R.J., and Behrens, T.E.J. (2012). An agent independent axis for executed and modeled choice in medial prefrontal cortex. Neuron *75*, 1114–1121.
31. Le Bouc, R., and Pessiglione, M. (2013). Imaging social motivation: distinct brain mechanisms drive effort production during collaboration versus competition. J. Neurosci. *33*, 15894–15902.
32. Fareri, D.S., and Delgado, M.R. (2014). Differential reward responses during competition against in- and out-of-network others. Soc. Cogn. Affect. Neurosci. *9*, 412–420.
33. Karafin, M.S., Tranel, D., and Adolphs, R. (2004). Dominance attributions following damage to the ventromedial prefrontal cortex. J. Cogn. Neurosci. *16*, 1796–1804.
34. Cook, J.L.L., den Ouden, H.E.M., Heyes, C.M.M., and Cools, R. (2014). The social dominance paradox. Curr. Biol. *24*, 2812–2816.
35. Spengler, S., von Cramon, D.Y., and Brass, M. (2009). Was it me or was it you? How the sense of agency originates from ideomotor learning revealed by fMRI. Neuroimage *46*, 290–298.
36. Renes, R.A., van Haren, N.E.M., Aarts, H., and Vink, M. (2015). An exploratory fMRI study into inferences of self-agency. Soc. Cogn. Affect. Neurosci. *10*, 708–712.
37. Naylor, J.C., Borckardt, J.J., Marx, C.E., Hamer, R.M., Fredrich, S., Reeves, S.T., and George, M.S. (2014). Cathodal and anodal left prefrontal tDCS and the perception of control over pain. Clin. J. Pain *30*, 693–700.
38. Judge, T.A., Erez, A., Bono, J.E., and Thoresen, C.J. (2002). Are measures of self-esteem, neuroticism, locus of control, and generalized self-efficacy indicators of a common core construct? J. Pers. Soc. Psychol. *83*, 693–710.
39. Kinicki, A.J., and Vecchio, R.P. (1994). Influences on the quality of supervisor–subordinate relations: the role of time-pressure, organizational commitment, and locus of control. J. Organ. Behav. *15*, 75–82.
40. Dambacher, F., Schuhmann, T., Lobbestael, J., Arntz, A., Brugman, S., and Sack, A.T. (2015). Reducing proactive aggression through non-invasive brain stimulation. Soc. Cogn. Affect. Neurosci. *10*, 1303–1309.
41. Deming, A.M., and Lochman, J.E. (2008). The relation of locus of control, anger, and impulsivity to boys' aggressive behavior. Behav. Disord. *33*, 108–119.
42. Watanabe, N., and Yamamoto, M. (2015). Neural mechanisms of social dominance. Front. Neurosci. *9*, 154.
43. Morgan, D., Grant, K.A., Gage, H.D., Mach, R.H., Kaplan, J.R., Prioleau, O., Nader, S.H., Buchheimer, N., Ehrenkaufer, R.L., and Nader, M.A. (2002). Social dominance in monkeys: dopamine D2 receptors and cocaine self-administration. Nat. Neurosci. *5*, 169–174.
44. Donoso, M., Collins, A.G.E., and Koechlin, E. (2014). Foundations of human reasoning in the prefrontal cortex. Science *344*, 1481–1486.

Current Biology, Volume *26*

Supplemental Information

Dynamical Representation of Dominance Relationships in the Human Rostromedial Prefrontal Cortex

Romain Ligneul, Ignacio Obeso, Christian C. Ruff, and Jean-Claude Dreher

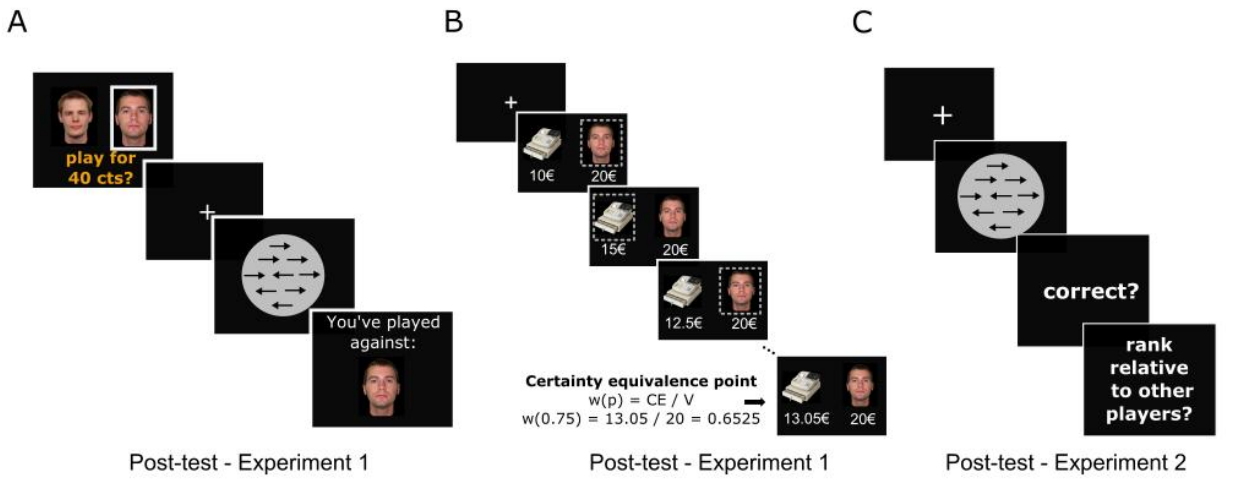


Figure S1 (related to Figure 1). Complementary behavioral tasks. (A) Choice opponent task (measured after Experiment 1)**.** Immediately after the fMRI session, subjects performed a modified version of the competition task in which they could select the opponent they wished to play against. After selection, they played the perceptual-decision making game which included a monetary incentive of 40 cents in each of the 36 trials (no feedback was given in order to avoid further learning). **(B)** Social 'certainty equivalence' task (measured after Experiment 1)**.** The recursive design of the certainty equivalence task (administered 1 hour after the main competitive task) allowed subjects to converge toward a 'certainty equivalence' point at which they had no preference between the two options displayed [S1]. The cash register on the left represented a sure amount of money which varied across trials. The face on the right depicts one of the opponents met in the competition task, coupled with a fixed amount of money. This money could be won with the probability equal to the frequency of winning against this opponent during the competition task, as learned during the main experiment. At the end of the task, one trial was randomly selected and played out, and the resulting money was added to subject's endowment. Each of the different opponents was evaluated in a separate block of 7-12 trials, until indifference was reached (for details, see[1]). Once the indifference point was reached, the subjective probability of winning against the presented opponent was computed as: $w(p) = \frac{CE}{V}$, where CE was the certainty equivalent and V the value associated with the risky amount (i.e., the amount associated with the opponent, here 20€). **(C)** In Experiment 2, immediately after 30 minutes of stimulation, a task was administered to verify that performance monitoring ability was not altered by tDCS stimulation. Subjects performed 80 perceptual decisions using the same stimuli as in the main competitive task and were asked to indicate whether they thought they had made an error after each response. In case of a correct response, they were also asked to guess their social rank on the basis of their reaction time (results of this experiment are reported in Table S8). Additional analyses of the results associated with Figure 2A and S1A/BN are available at https://dx.doi.org/10.6084/m9.figshare.3811365.v2

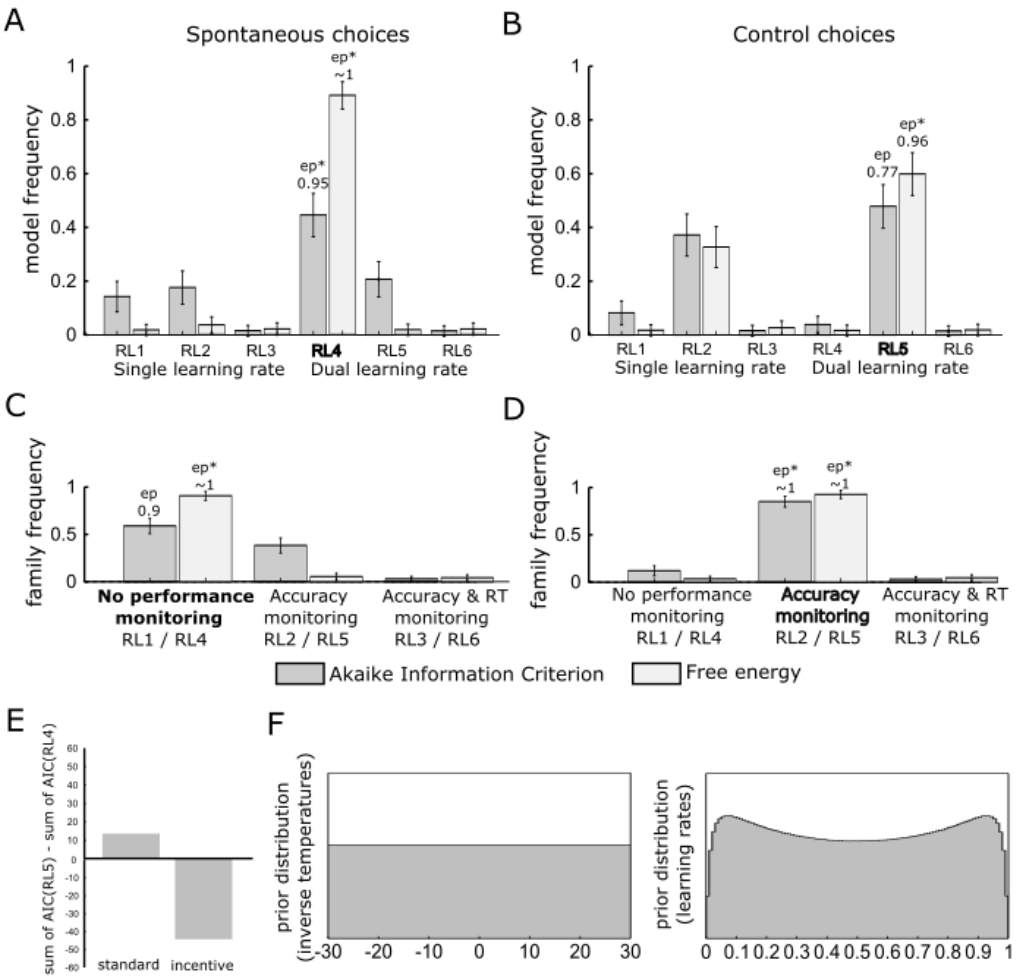


Figure S2 (related to Figure 5). Group Bayesian model selection in the two types of trials (Experiment 2) (see also Table S6-7 and Supplemental Experimental Procedures). **(A)** Model comparison for spontaneous choices driven by social dominance representations. These choices indicated that a model with two learning rates for victories and defeats (top) was more likely, which updated dominance statuses after errors. Light and dark bars represent the estimated frequency of each model in the population, using either Free Energy or the Akaike Information Criterion as comparison metrics, respectively. **(B)** Model comparison for control choices driven by rational maximization of monetary pay-off indicated the most likely model that was the one with two learning rates for victories and defeats (top), which did *not* update dominance statuses after errors. **(C)** Same analysis performed by pooling models with one and two learning rates into 3 families characterized by variation in the dependence between performance monitoring and update. In the spontaneous condition, the models that updated dominance representation were the most adequate. **(D)** Oppositely, in the control condition, models that conditioned update on accurate responding in the perceptual task were the most adequate. **(E)** Differences in Akaike values between the competing models 4 and 5, showing the relative advantage of RL4 (lower AIC) on RL5 to explain spontaneous choices, and vice versa. **(F)** Prior distributions used for the Variational Bayesian inversion of the learning models (left, inverse temperatures; right, learning parameters), designed to approach as much as possible non-informative, flat priors over the full range of plausible values. Additional discussion of the results associated with this figure is available at https://dx.doi.org/10.6084/m9.figshare.3811365.v2

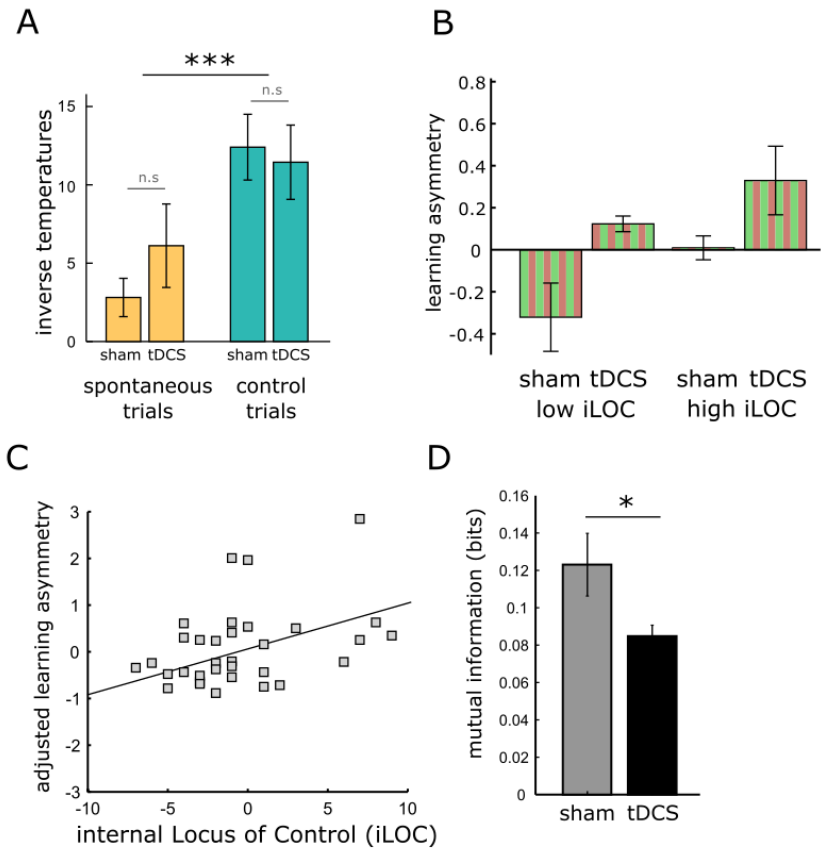


Figure S3 (related to Figure 5). (A) Inverse temperature parameters driving choices in each trial type showed that incentives made subjects more deterministic in their decision in order to maximize monetary pay-offs. tDCS did not affect this parameter. **(B)** Decomposition of the tDCS effect in participants with respect to the internal Locus of Control personality trait (high iLOC, low iLOC, median split) showing that tDCS stimulation and iLOC influenced learning asymmetry in a similar and additive way. **(C)** Learning asymmetries adjusted for tDCS effect (i.e. z-scored in each group independently) were significantly correlated with iLOC scores across subjects (r=0.41, p=0.016). **(D)** The mutual information between current choice (3 categories: superior, intermediate, inferior) and immediately preceding outcomes (6 categories: victory or defeat against each opponent type) was lower in the tDCS group as compared to the sham group (the first trial following each triplet of control choices was not included in this analysis). This is because averaged SDS values fluctuated more slowly than those pertaining to opponent-specific dominance relationships did. Confirming our interpretation, a lower mutual information between choices and the preceding competitive outcome was observed in the tDCS as compared to the sham group. Mutual information was computed using the MIToolbox (https://github.com/Craigacp/MIToolbox) according to the formula:

$$I(C;O) = \sum_{C \in [sup,mid,inf]} \sum_{O \in [win_{sup},win_{mid},win_{inf},lose_{sup},lose_{mid},lose_{inf}} p(C,O)\log\left(\frac{p(C,O)}{p(C) * p(O)}\right)$$

where C represents the 3 possible choices, O represents the 6 possible opponent-specific outcomes, and $p(C,O)$ represents the joint probability of a given combination of C and O. Error bars indicate s.e.m. Additional discussion of the results associated with this figure is available at https://dx.doi.org/10.6084/m9.figshare.3811365.v2

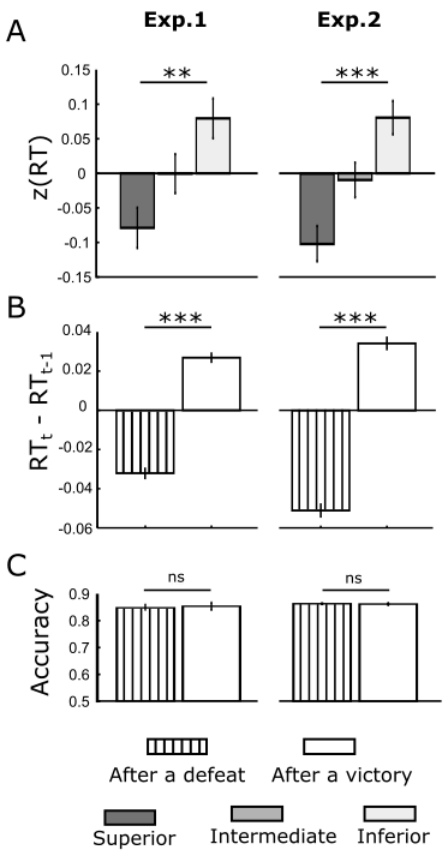


Figure S4 (related to Figure 2). Perceptual decision-making performances in both experiments. **(A)** Z-scored reaction times indicated a clear modulation of reaction times by the opponents' rank, with faster reaction times against better opponents in both studies (top panel). **(B-C)** Trial-to-trial fluctuations in reaction times and accuracies indicated that experiencing a defeat increased speed **(B)** but not accuracies **(C)** of perceptual decisions in the following trial (middle and bottom panels, respectively). Error bars and shaded areas represent s.e.m. Additional analyses related to this figure are available at https://dx.doi.org/10.6084/m9.figshare.3811365.v2

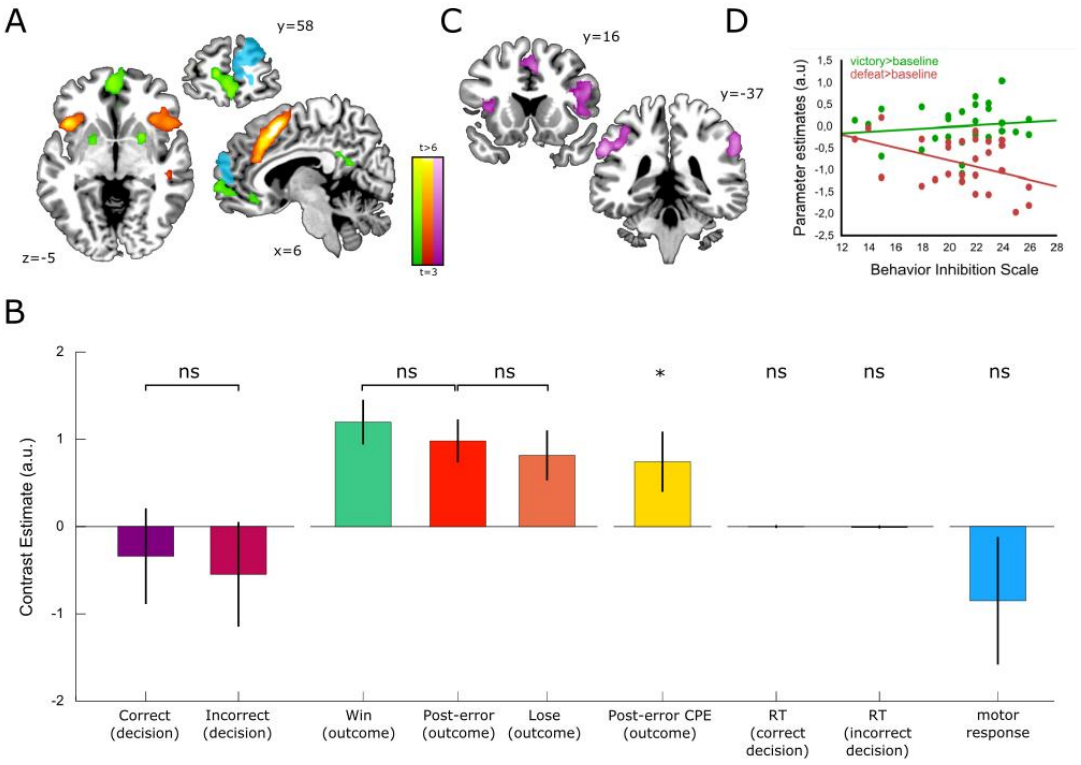


Figure S5 (related to Figure 3). Additional fMRI analyses (GLM1). **(A)** Statistical maps (cluster-forming threshold: p=0.001 unc.) for the correct>incorrect (in green) and incorrect>correct contrasts (in red) show the effect of response accuracy at the perceptual decision-making stage. This analysis revealed a classical error-monitoring network encompassing the anterior insula and the posterior dmPFC (incorrect>correct) as well as a classical valuation network encompassing the vmPFC and the lateral parts of the ventral striatum (correct>incorrect). Importantly, these accuracy-related activities did not overlap with learning-related rmPFC activities reported in Figure 3B (shown here in transparent blue). **(B)** ROI analyses in the rmPFC (transparent blue cluster; see also Figure 3B) indicate that the rmPFC activity was insensitive to perceptual errors at the decision-making stage (p>0.25; paired t-test, dark and light purple) as well as at the outcome stage (red *vs* green and red *vs* orange: both p>0.25; paired t-tests). rmPFC activity was also uncorrelated with reaction times at the decision-making stage (independently of accuracy) and was unaffected by motor responses (blue). However, the analysis of cPE encoding following erroneous perceptual decisions (yellow) revealed a significant, albeit reduced, encoding of prediction errors (one sample t-test: $t(1,27)=2.2$; $p<0.05$). **(C)** A whole-brain analysis revealed positive correlations between reaction times and BOLD activities within the bilateral parietal cortex, the right posterior dlPFC and the anterior cingulate cortex (in purple). These effects are likely the result of increased perceptual efforts in trials showing longer reaction times (due, for example, to the central positioning of several incongruent arrows). Again, effects of reaction times did not overlap with the rmPFC cluster of interest. No significant correlations were observed for the reverse contrast (i.e. more activity associated with decreasing RT). **(D)** A regression analysis between BIS and BAS personality traits[S2,S3] with striatal BOLD responses to competitive outcomes (as defined from Figure 4A) reveal that defeat-induced deactivations correlated with the Behavioral Inhibition Scale ($t=-2.68$, $p<0.05$) but not with the Behavioral Activation Scale ($t=0.76$; $p>0.45$). Overall, the ventral striatum of more inhibited subjects is more sensitive to the victory>defeat contrast ($t=3.76$, $p<0.01$), contrary to the BAS scale ($t=-1.41$; $p=0.17$). Bars and shaded areas represent s.e.m. Additional discussion associated with the results reported Figure S5 is available at https://dx.doi.org/10.6084/m9.figshare.3811365.v2

Region	BA	side	cluster FWE	n voxels	T	Z	X	Y	Z
Competitive prediction errors (positive and negative; conjunction analysis)			Outcome stage						
Superior Frontal Gyrus (rmPFC)	10	R	p<0.001	280	5.29	4.35	6	59	10
Competitive prediction errors (positive only)			Outcome stage						
Insula / Putamen	47	R	0.014	73	5.59	4.52	30	20	-5
Temporal Lobe posterior	21/22	R	p<0.001	128	5.15	4.26	66	-43	-2
Superior Frontal Gyrus	8/6	R	0.013	74	5.14	4.26	18	29	55
Frontal Gyrus middle	9/8/46	R	p<0.001	443	5.09	4.22	42	11	34
Insula / Putamen*	13	L	0.089	46	4.74	4.01	-30	14	-2
Superior Frontal Gyrus (rmPFC)	9/10	R	p<0.001	137	4.53	3.87	6	65	19
Parietal Lobe inferior/lateral	40/4	R	p<0.001	159	4.53	3.87	45	-58	49
Competitive prediction errors (negative only)			Outcome stage						
Superior Frontal Gyrus (rmPFC)	9/10	L/R	0,001	121	4.41	3.80	9	62	22
Cingulate Gyrus posterior	31	L/R	0.002	108	4.25	3.69	-6	-64	22
Social dominance status			Perceptual decision-making stage						
Precentral Gyrus	9	R	0.033	56	5.72	4.59	48	8	25
Superior Frontal Gyrus* (rmPFC)	10	R	0.12	39	5.35	4.38	21	50	31

Table S1 (related to Figure 3). Brain regions encoding continuous learning-related variables derived from the reinforcement-learning algorithm. Cluster-forming threshold: p=0.001. Cluster-wise multiple comparisons threshold: p^{FWE}=0.05. Asterisks (*) denote regions which did not pass the statistical threshold for multiple comparisons but which are of scientific interest in the context of the present study.

Region	BA	side	cluster FWE	n voxels	T	Z	X	Y	Z
Insula | Orbital Gyrus Posterior	38/47	R	0.005	99	5.60	5.24	51	11	-14
Amygdala*	#	R	0.21	37	5.16	4.86	21	-7	-26
Precentral Gyrus	6	R	0.01	86	5.08	4.80	9	-22	58
Caudate | Accumbens | Insula	#	L/R	p<0.001	350	4.81	4.56	-15	17	-2
Hippocampus /Amygdala*	#	L	0.11	48	4.52	4.32	-24	-19	-17
SNc / VTA*	#	R	0.55	22	4.31	4.13	6	-13	-11
Superior Frontal Gyrus vmPFC	11/12	L	p<0.001	166	3.99	3.85	-3	47	-11

Table S2 (related to Figure 4). Brain regions were activity showed the categorical interaction [intermediate opponent win & control failure]>[intermediate lose & control success], as revealed by a 2-way flexible factorial model at the second level. Cluster-forming threshold p = 0.001. Cluster-wise multiple comparisons threshold: p^{FWE}=0.05. Asterisks (*) denote regions which did not pass the statistical threshold for multiple comparisons but which are of scientific interest in the context of the present study.

Spontaneous choices

Model	Description	n param.	F - sum	EP (F)	AIC - sum	EP (AIC)	α vic	α def	β	ω
RL1	1α - error: update	2	-3734		6110		0.33±0.25		6.9 ±12.3	
RL2	1α - error: no update	3	-3717		6079		0.3±0.23		6.6±11.2	
RL3	1α - 1ω error: no update	3	-3710		6168		0.32±0.22		4.5±7.3	0.32±0.21
RL4	2α - error: update	3	**-3643**	>0.999	**5996**	>0.95	0.34±0.24	0.29±0.226	4.46±8.6	
RL5	2α - error: no update	3	-3693		6009		0.33±0.22	0.28±0.25	4.2±7.8	
RL6	2α - 1ω - error: no update	4	-3687		6122		0.33±0.20	0.32±0.25	3.8±6.8	0.29±0.22

Control choices

Model	Description	n param.	F - sum	EP (F)	AIC - sum	EP (AIC)	α vic	α def	β	ω
RL1	1α - error: update	2	-993		1727		0.21±0.14		9.8±0.7.9	
RL2	1α - error: no update	3	-928		1676		0.19±0.14		10.8±8.8	
RL3	1α - 1ω error: no update	3	-953		1757		0.23±0.15		8.2±7	0.31±0.2
RL4	2α - error: update	3	-974		1702		0.27±0.14	0.25±0.21	12.6±11	
RL5	2α - error: no update	3	**-915**	>0.95	**1658**	0,77	0.24±0.14	0.23±0.2	11.9±9.1	
RL6	2α - 1ω - error: no update	4	-949		1775		0.29±0.15	0.27±0.19	8.9±7.8	0.31±0.22

Table S3 (related to Figure 5). Model selection and parameters explaining spontaneous choices (top) and control choices (down). F = Free energy. AIC = Akaike information criterion. EP = exceedance probabilities. α = learning rate. β = inverse temperature. ω = sensitivity to rapidity of responding. See also the "Computational model: estimation and comparison procedure" section for an exhaustive description of the models.

	sham		tDCS		
	mean	SD	mean	SD	p(stim)
Age	21.76	2.22	22.53	2.58	0.40
Subjective Social Status	6.24	1.60	6.65	1.37	0.39
Self-Esteem	23.41	4.95	22.06	2.68	0.36
Competitive Personality	31.12	5.30	29.29	4.51	0.31
Cooperative Personality	22.82	4.61	24.06	2.79	0.31
Perceived Stress	14.06	6.04	13.12	4.12	0.63
BAS Drive	8.41	3.16	9.18	1.98	0.41
BAS Fun	8.06	2.68	7.35	2.60	0.46
BAS Reward	8.41	3.43	8.94	3.34	0.67
BIS	15.65	2.21	15.76	2.11	0.95
LOC Internal	1.06	4.16	-2.29	3.26	**0.02**
LOC external	2.94	5.88	4.94	5.23	0.23
LOC chance	4.18	5.77	1.12	4.34	0.10
Mood-before	5.12	1.05	5.24	1.03	0.77
Mood-after	5.18	1.07	5.29	0.99	0.75
Positive emotions	5.18	0.95	5.29	1.31	0.79
Negative emotions	3.82	1.55	3.29	1.31	0.29
Perceptual accuracy - spontaneous	0.83	0.07	0.85	0.06	0.42
Perceptual accuracy - control	0.86	0.08	0.86	0.06	0.74
Perceptual RT – spontaneous trials	0.47	0.05	0.48	0.11	0.66
Perceptual RT – control trials	0.46	0.05	0.46	0.10	0.90
Error detection	0.92	0.05	0.93	0.05	0.91
Self-performance sensitivity	0.28	0.18	0.30	0.19	0.84
Self-ranking	3.85	1.09	3.76	0.85	0.90

Table S4 (related to Figure 1 & 5). Demographic information and performance in the perceptual task compared between the tDCS and sham groups. The last three rows correspond to measures derived from the performance monitoring post-test (see also Figure S1) administered immediately after the stimulation protocol: “error detection” refers to the percentage of correctly reported errors, “self-performance sensitivity” refers to the coefficient of correlation (Spearman) between reaction times and guessed social ranks, “self-ranking” refers to the average social ranks guessed by subjects (a lower value indicating more a favorable self-ranking).

	SS (type III)	ddl	Mean squares	F	Sig.
GLM-A (all trials)					
TrialType * Valence * ILOC	.106	1	.106	2.590	.118
TrialType * Valence * Gender	.004	1	.004	.098	.756
TrialType * Valence * Stimulation	.285	1	.285	6.959	**.013**
TrialType * Valence * Gender * Stimulation	.001	1	.001	.017	.898
GLM-B (spontaneous trials)					
Intercept[#]	7.723	1	7.723	63.679	.000
Valence	.104	1	.104	2.630	.116
Valence * iLOC	.231	1	.231	5.859	**.022**
Valence * Gender	.000	1	.000	.009	.924
Valence * Stimulation	.535	1	.535	13.5	**<.001**
Valence * Gender * Stimulation	.005	1	.005	.126	.725
iLOC	.034	1	.034	.278	.602
Gender	.003	1	.003	.029	.867
Stimulation	.006	1	.006	.048	.828
Gender * Stimulation	.045	1	.045	.373	.546
GLM-C (control trials)					
Intercept	4.035	1	4.035	93.942	.000
Valence	.001	1	.001	.020	.887
Valence * iLOC	.000	1	.000	.011	.917
Valence * Gender	.005	1	.005	.134	.717
Valence * Stimulation	.001	1	.001	.015	.903
Valence * Gender * Stimulation	.012	1	.012	.311	.581
iLOC	.027	1	.027	.640	.430
Gender	.009	1	.009	.204	.655
Stimulation	.000	1	.000	.009	.927
Gender * Stimulation	.005	1	.005	.108	.745

Table S5 (related to Figure 5). GLM results with differential effects of tDCS stimulation on learning asymmetry as a function of trial type. GLM-A corresponds to the full model, which showed a significant 3-way interaction between trial type (spontaneous/control), learning rate valence (positive/negative) and stimulation condition (sham/tDCS). For GLM-A, effects are only reported down to the level where the 3-way interaction was split for further analysis. GLM-B corresponds to GLM-A restricted to the spontaneous trial types and shows a strongly significant interaction of stimulation condition with stimulation condition (see also Figure 5C). GLM-C is the same as GLM-B for control trial type, showing no effect of stimulation condition in this situation. Internal Locus of Control was included as a covariate in all GLM due to a significantly different score in the two groups.

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Experiment 1: fMRI

Participants

Twenty-eight young men were included in the analyses (mean age: 22; age range: 18-27) using the main mailing list addressed to the students of the University of Lyon. All participants met the inclusion criteria (right-handedness, no history of psychiatric or neurologic disorders, Caucasian ethnicity) and gave written, informed consent for a protocol approved by the local ethics committee (Comité de Protection des Personnes – Sud-Est 2, Lyon, France). 5 additional participants were scanned but not included in the study due to excessive motion in the scanner (3 subjects) or because they spontaneously expressed doubts regarding the cover story (2 subjects).

Experimental design: full description

All sessions started between 2 and 3 pm. The first hour was dedicated to instructions, initial training and contextualization. Then, subjects entered the scanner and performed a second training on the perceptual task of 100 trials (see below for detailed description) which included an accuracy feedback but no social manipulation. This allowed performance to reach a stable and controlled level before the beginning of the competition phase. Moreover, we used an implicit staircase procedure to ensure that the perceptual difficulty of the task would be homogenous across subjects (3:1 ratio corresponding to a desired rate of 80% of correct responses; overall performance during the competitive task: 78.2% +/- 5%).

The fMRI experiment itself involved four conditions that were matched for visual stimulation and timing parameters. Three conditions were associated with the competitive context. In these types of trials, subjects had to perform better than their opponents in the perceptual task. This perceptual

task required estimating as fast and accurately as possible the main direction indicated by a set of *non-moving* 46 arrows pointing either to the left or to the right. Perceptual difficulty was constant and corresponded to the proportion of arrows pointing in the majority direction. The rules of the competitive game were explicitly defined as follow: “The fastest player wins if both responses are correct. The accurate player wins if the other gives an incorrect response. If one player doesn’t respond, then he loses automatically. If both responses are incorrect, the slowest player wins”. This latter rule was set to discourage fast-guessing responses. The fourth condition was dedicated to the social but non-competitive control situation. In this type of trials, similar to a simplified coordination game, the subject viewed bars instead of arrows (no perceptual judgment) and had to respond by pressing the left or right button. They were told that the trial would result in shared win or loss if both players responded in the direction or opposition directions, respectively. Each condition comprised 60 trials (control, superior competitive, intermediate competitive, inferior competitive), resulting in a total of 240 trials for the whole experiment. Importantly, the association between the four faces and the four conditions was counterbalanced across subjects in order to avoid any bias due to visual stimulation when analyzing group results.

The experiment was divided in two fMRI runs of 18 minutes. These two runs were identical and comprised 8 blocks of 15 trials played sequentially against or with the same player in the competitive or control situations, respectively. Each block began with a 5 s yellow fixation cross followed by a 10 s screen presenting the picture of the other player as well as short instructions relating to the situation. Each trial proceeded as follow. A jittered white fixation cross (2.5 - 4.5 s, uniform distribution) initiated the trial. Then, the stimulus associated with the competitive or control situation (arrows or bars, respectively) was displayed for 350 ms within a grey disk, which subsequently remained alone on the screen for 650 ms; subjects were instructed to respond before this disk disappeared (all *miss* trials were removed from the analyses, on average, we observed 1.93 *misses* per subject). A second jittered fixation

cross (2.5 - 4.5 s, uniform distribution) separated the competition phase and the outcome presentation. The latter lasted 2 s and presented one of the following messages “Won against:” or “Lost against:” (competitive situation), “Won with:” or “Lost with:” (control situation) on top of the other player’s face.

Note that although participation in the study was compensated (70€ per subject), there was no monetary incentive associated with any events in this experiment, in order to avoid confounds when interpreting the meaning of BOLD signals elicited by social interactions and social outcomes in both situations.

Cover story

A strong contextualization procedure was used to ensure genuine cognitive and emotional engagement in the task. This procedure was largely inspired by a previous study on a related topic [S3]. Although they were fully simulated by the computer, the four other players were presented as real participants playing simultaneously over the internet. During the preliminary inclusion visit, we took a picture of each subject in order to “allow other participants to view his face during the experiment” (target faces, counterbalanced across subjects, were in the same age range than our participants, that is, between 18 and 30). We also insisted on the extreme importance of punctuality for this kind of multi-subject experiment. On the day of the experiment, participants had to certify that they did not recognize any of the four other participants’ faces. During scanner configuration, a fake webcam broadcast (showing the other participants being instructed by a second experimenter) was displayed next to some suggestive LINUX commands (e.g. “$ register participant 1 –namecode”). Finally, the subject had to wait approximately one minute in front of a ‘connection panel’ showing three other participants’ faces and the experimenter informed him that this was because the last participant connected late. Among the 28 participants effectively included in the study, none spontaneously expressed doubts concerning the procedure. Although six subjects declared that they were not really surprised when the experimenter

made the manipulation explicit (after the experiment was completed), those subjects reported that they had not been overly unsure about the cover story. Note that removing these subjects from the analysis did not change the results at the group level.

Optimization of the experimental design

Sequences of outcomes were precomputed to match the desired frequencies of winning and losing against each opponent. In the competitive conditions, winning frequencies for the subject were set to 33% when playing against the superior opponent, 50% against the intermediate opponent and 66% against the inferior opponent. In the control condition, winning frequencies were set to 50%, allowing unbiased, direct comparisons with the intermediate opponent from the competitive condition. Moreover, for each condition and subject, the sequence of winning and losing events was optimized in order to maximize statistical sensitivity to the win-lose contrast using homemade scripts derived from a validated toolbox [S4]. Moreover, this procedure normalized detection power across subjects, thus improving reliability of inter-individual variability analyses. Nonetheless, based on preliminary tests, we realized that it was necessary to attenuate performance-outcome independence in order to maintain believability. Indeed, enduring a defeat following an extremely fast correct response was perceived as very implausible by our pilot subjects. Therefore, we decided to replace on-line a precomputed defeat by a victory, if the subject produced a correct response with a RT below the 20^{th} percentile of his own accurate RT distribution. In this case, the suppressed defeat was later reinstated when an inaccurate response occurred against the same opponent in a trial precomputed to be victorious (this affected less than 10% of the trials). This strategy was successful to achieve believability while keeping target frequencies of winning in the desired range. Effective frequencies of winning were: 37% +/- 6% for the superior, 51% +/- 3% for the intermediate, and 63% +/- 3% for the inferior opponent.

fMRI acquisition

Imaging was conducted on a 1.5T Siemens Sonata scanner, using an eight-channel head coil. Twenty six interleaved slices tilted relative to the anterior commissure – posterior commissure line (20-30°) were acquired per volume (field of view, 220 mm; matrix 64 x 64; voxel size, 3.4 x 3.4 x 4mm; interslice gap, 0.4 mm). Functional images were obtained using a gradient-echo echoplanar imaging (EPI) T2*-weighted sequence (TR, 2.5 s; TE, 60 ms; flip angle, 90°). To improve the local field homogeneity and hence minimize susceptibility artifacts in the orbitofrontal cortex, a manual shimming was performed and a map of the magnetic field was acquired to correct for residual inhomogeneity-related distortion in the functional scans. Following the three fMRI runs, a high-resolution T1-weighted anatomical scan was acquired.

fMRI data preprocessing

All preprocessing steps were conducted using SPM8. The first four volumes of each run were removed to allow for T1 equilibrium effects. For each subject, functional images were time-corrected, realigned, unwarped using the magnitude and phase images, and coregistered to the anatomical scan. The anatomical scan was then normalized to the MNI space using the ICBM152 template brain and the resulting transformation matrix was applied to the functional images. Finally, the normalized functional images were spatially smoothed with an 8 mm Gaussian kernel.

fMRI analyses: full description

Statistical analyses of fMRI signals were performed using a conventional two-levels random-effects approach with SPM8 (http:/www.fil.ion.ucl.ac.uk/spm). All general linear models (GLM) described below included the 6 motion parameters estimated from the realignment step, in order to covary out potential movement-related artifacts in the BOLD signal. All regressors of interest were convolved with the canonical hemodynamic response function (HRF) using a Dirac function for motor events and a boxcar lasting the duration of the visual stimulus associated with each regressors. All GLM

models included a high-pass filter to remove low-frequency artifacts from the data (cut-off = 128 s) as well as a run-specific intercept. Temporal autocorrelation was modeled using an AR(1) process. For the computation of evoked responses (i.e. profiles of BOLD activity), we chose the well-validated and widely used *rfxplot* toolbox[S5]. BOLD signals were extracted from the cluster of interest (i.e., rmPFC activity as defined at Figure 3B) and were adjusted in each subject for block- and movement-related variances.

In order to capture the dynamics of the social learning process elicited by our implicit hierarchy manipulation, we applied a Rescorla-Wagner rule (learning rate = 0.1; see Equation 1, below) to model the sequence of victories and defeats associated with each competitive condition, in each subject separately. This model-based fMRI approach was thus used to probe the neural substrates of competitive prediction errors (cPE) and anticipated opponents' values (i.e. social dominance status, SDS) (GLM1). At the first-level, 3 categorical regressors modeled the decision-making phase (onset of the target arrows): control decisions, incorrect competitive decisions, correct competitive decisions. Four categorical regressors were dedicated to the outcome phase: competitive victories, competitive defeats, control outcomes, plus one regressor for outcomes which followed erroneous perceptual decisions in the competitive condition. Then, six parametric regressors were added to their corresponding categorical regressors in order to estimate brain activity correlating with the trial-by-trial variables derived from our RL algorithm. Positive, negative, and 'post-error' competitive prediction error values (cPE) were added as parametric modulators of victories, defeats and outcomes following an erroneous decision, respectively. In order to avoid confounds in the interpretation of cPE signals, RTs and trial by trial SDS values (i.e., expected chances of winning) derived from the RL algorithm were added as separate parametric modulators of the correct and incorrect perceptual-decisions regressors. Finally, control outcomes were parametrically modulated by a non-competitive prediction error term derived from the same algorithm used to describe learning in the competitive trials (ncPE).

A second GLM was run to uncover competition-specific brain activities associated with outcome processing (GLM2) and to compare model-based results with a GLM not accounting for the ongoing learning process. This GLM fractionated competitive victories and defeats in 3 regressors each (i.e., against superior, intermediate and inferior opponents), while control successes and failures were modeled as 2 separate regressors, resulting in 8 categorical regressors for the outcome period. The decision phase was also modeled using three categorical regressors: one for correct perceptual decisions, one for incorrect perceptual decisions, and one for control decisions.

A third GLM was used to confirm that the encoding of cPE was not confounded with block or opponent effect. This GLM3 was the similar to GLM1 except that victories and defeats were modeled as single outcome events, with 3 parametric regressors serially added to explain activities related to opponent category (superior, intermediate, inferior), outcome valence (positive, negative) and finally cPE. Due to the orthogonalization procedure of SPM, the cPE regressor could not explain any activity related to opponent category or outcome valence. Non-competitive outcomes were modeled as a single event modulated by prediction errors.

All statistical maps were obtained with at a statistical threshold of $p<0.05$ FWE-corrected for multiple comparisons at the cluster level, with a cluster-forming threshold of $p<0.001$. For the display of effects within regions of interest (ROI), we extracted contrast estimates from the corresponding cluster (resulting from a threshold of $p<0.001$ unc.) and computed the mean estimates. The fMRI results from the competitive task were analyzed using 2 GLMs described below. Each GLM included the onset of motor responses, pauses between mini-blocks, and six un-convolved motions regressors as covariates of no interest.

Model-based fMRI analysis

The following Rescorla-Wagner rule was used :

$$SDSi(t+1) = SDSi(t) + \alpha * (R(t) - SDSi(t)) \quad \text{(Equation 1)}$$

Where α is the learning rate of the algorithm, $SDSi(t)$ is the momentary dominance of the subject against a given opponent *i* at the beginning of trial *t*, and $R(t)$ is the outcome of the competitive interaction, coded as 1 in case of victory and -1 in case of defeat. Because we collected no choice data in Experiment 1, we set the learning rate at the standard level 0.1 for all the reported results (learning rates of 0.05, 0.2 and 0.3 were also investigated and revealed very similar results). This procedure has recently been proved valid, as the exact value of learning rates used in model-based fMRI has very limited impact on final results [S6]. Note that on each trial, $\mathrm{SDSi(t)}$ represents the anticipated probability of winning against a given opponent i. Since SDS_i values were updated for each opponent separately, they are higher against weaker opponents and lower against stronger opponents. In other words, they reflect the expected dominance of the subject over his/her opponents (the regressor was inverted when entered in the design matrix, so that the positive correlation, as displayed Figure 3D, indicates more BOLD activity when facing a momentarily dominant opponent).

The competitive prediction error variable regressed against BOLD signals correspond to the following term: $cPE = R(t) - SDSi(t)$. For each trial, actual and anticipated outcomes are thus compared by the algorithm to generate a teaching signal (the cPE) used to update the social dominance status (SDS) momentarily associated with the opponent under consideration. In this computational scheme, with victories coded as R=1 and defeats coded as R=0, SDS values thus fluctuate around the true underlying chances of winning associated with each opponent.

Experiment 2: tDCS

Participants and variables controlled across experimental groups

Thirty-nine subjects were recruited to participate in this experiment through the student mailing-list of the University of Zurich. The protocol was approved by the ethics committee of the Canton of Zurich. Testing was performed in groups of 9-11 participants in the behavioral room of the Laboratory for Social and Neural System research (SNS-Lab). In each of the 4 experimental sessions, subjects were randomly and evenly assigned to the tDCS or sham condition (4-6 tDCS subjects per condition and session). We used a double blind design where the subjects and the experimenters in charge of the instructions and electrode placement were unaware of the condition assignment. Five participants were excluded because they did not exhibit rational choices during the trials which included monetary incentives (i.e. their choices were not transitive).

Among the 39 participants recruited in the study, five were excluded because they were not transitive in their competitive preferences during triplets of control trials interleaved with the trials of interest (for example, if they preferred A over B and B over C, they didn't choose A over C more than 50% of the time) and they did not select inferior opponent more often than superior opponent in this condition, suggesting that they were not engaged in the task and/or that they did not detect the presence of an implicit hierarchy.

To control for potential confounds related to demographic and personality variables (see Table S8), all included participants (17 males, 17 females; mean age 22,1 +/- 2.4 STD) completed questionnaires measuring self-esteem[S7], perceived stress [S8], competition-cooperation strategy [S9], Behavioral Inhibition and Activation Scales [S2] (BIS-BAS), subjective social status and Locus of Control [S10]. A complementary task performed at the end of the experiment was also used to exclude that the effects of rmPFC stimulation on learning would be mediated by alteration of self-performance

monitoring. We performed two-way ANOVAs with stimulation condition and gender as fixed effects to detect eventual differences between the four subgroups of participants for these cognitive, demographic and personality factors (see Table S4). This showed that the two stimulation groups were reasonably matched ($p>0.15$) with respect to any variables that may affect social learning and strategic reasoning under competitive settings, except for the chance and internal Locus of Control subscales (chance LOC, $p=0.10$; internal LOC, $p=0.02$; sham group higher on both subscale). Including these subscales as covariates did not affect statistical inferences concerning our main conclusions, but we chose to systematically include iLOC as a covariate due to its a *priori* relevance for social dominance and rmPFC computations (see Results and Discussion sections).

Experimental design

In the task used for the tDCS experiment, subjects were told that they would compete against each other anonymously during three similar blocks involving four players (the subject plus three opponents). Anonymity of social competition was primed by displaying a fractal image and a three-syllable word (not a name) attached to each of the three opponents. The perceptual game used for social competition and hierarchy induction was the same as in the fMRI task (see above), except that arrows were not hidden after 350ms and that the total number of arrows was reduced to 25 (instead of 46). As for fMRI, the experiment began with 100 training trials with accuracy feedback but no social manipulation or brain stimulation. Again, a staircase procedure enabled us to match perceptual difficulty across subjects (4:1 ratio corresponding to a desired rate of 85% of correct responses; overall performance during the competitive task: 84.2% +/- 6%).

Unbeknownst to the subjects, all triplets of opponents were programmed in order to reach 33%, 50% and 66% of victories against the subject, thereby creating an implicit social hierarchy composed of an inferior, an intermediate and a superior opponent (see next section). The three blocks of the task

were identical except that the nicknames and the fractals depicting players were changed and subjects were told that they would play against three new opponents, randomly and anonymously picked among the other players in the room. Each block was composed of 48 spontaneous choices in which there was no money at stake and in which feedbacks (win, even, lose) were given to the subjects. Each trial proceeded as follow (Figure 1): fixation cross (1-3 s), display of the two opponents available for choice (self-paced), fixation cross (1-3 s), perceptual decision-making (1s), fixation cross (1-3 s), display of the outcome (winning/losing/even). Additionally, 3 'control choices' without feedback had to be performed at the beginning of each block and after each series of 12 'spontaneous choices' with feedback. Subjects were then explicitly warned that 2 CHF would be added to the initial endowment of the winning player and that the same amount of money would be removed from the endowment of the losing player. Because we did not want learning to be affected by these incentivized trials, no immediate feedback was given but a (fake) summary of the extra-money won or lost was displayed at the end of each block. The time-course of these trials was strictly identical to the 'spontaneous choices' trials, except that no outcome was displayed. Finally, note that for all trial types, only 2 opponents were proposed at the choice stage. The subjects were told that the opponent they would not select would automatically compete against the third (undisplayed) player engaged in the block.

This design was used to disentangle social learning itself from competition- and status-related preferences. In the 'spontaneous choices' trials, subjects' choices could thus be driven by the desire to challenge the strongest opponent or by any form of information-seeking behavior. In the 'control choices' trials, there was a strong incentive for the subjects to choose the weakest opponent according to what they actually learned, whereas competitive and hierarchical motivations over choices were minimized by the absence of feedback.

Algorithm used to control competitive outcomes

The algorithm generating these frequencies of winning computed, after each correct response, to which percentile the current RT belonged, relative to the cumulative distribution of reaction times until the current trial. The distribution of reaction times was initialized using the last 30 trials of the training period. The threshold percentile under which a victory was delivered was determined by the target percentage of victories against the considered opponent multiplied by two *minus* the current percentage of victories against this opponent, so that a lack of victorious outcomes translated into more lenient threshold while an excess translated into a more severe threshold:

$$\textbf{victory}: \quad if \begin{cases} accurate\ perceptual\ decision \\ \textbf{and} \\ Percentile_{RT(t)} < (Pvic^{end}_{opponent} * 2) / Pvic^{t}_{opponent} \end{cases}$$

$$\textbf{defeat}: \quad if \begin{cases} inaccurate\ perceptual\ decision \\ \textbf{or} \\ Percentile_{RT(t)} \geq (Pvic^{end}_{opponent} * 2) / Pvic^{t}_{opponent} \end{cases}$$

$$\textbf{even}: \quad if \begin{cases} inaccurate\ perceptual\ decision \\ \textbf{and} \\ Pvic^{end}_{opponent} > Pvic^{t}_{opponent} \end{cases}$$

Where $Pvic^{end}_{opponent}$ is the target frequency of winning against the opponent, $Pvic^{t}_{opponent}$ is the current frequency of winning against the opponent, and $Percentile_{RT(t)}$ is the percentile of the RT relative to the cumulative probability distribution of the previous 30 trials (to accommodate fluctuation in performance across the experiment). This adaptive procedure was used to achieve good control over

outcomes while maintaining high believability and ecological validity. When this procedure was not sufficient to maintain hierarchical differences within reasonable limits, victories or defeats could be 'forced' irrespective of RT, provided that the subject gave a correct response or that the current RT percentile was superior to 10, respectively. Only 16,5% +/- 4% STD of the trials were adjusted in this latter way.

Contextualization procedure

For the tDCS experiment, the social competition was *a priori* highly credible because the experiment was performed in the same behavioral room of the Laboratory for Social and Neural Systems by ten to twelve subjects simultaneously. The computer screens of the room were shielded in view from one another and the experimenter stayed continuously in the room, preventing subjects from communicating and watching the screen of others. The beginning and termination of each block were synchronized to enhance the feeling of real-time gaming. At the beginning of each experimental session, the experimenter collectively instructed subjects that they would compete against each other anonymously – during three similar blocks involving four players (the subject plus three opponents). In addition, he indicated that the competition would have to be anonymous in order to exclude confounds related to social perceptions. At the beginning of each block, the subjects were attributed a nickname (3-syllabs pseudo-word) and a fractal image, and were presented with the nicknames and fractals associated with other opponents. Finally, the experimenter explained that some subjects would have the possibility to choose against whom they would play while others would just wait to be selected, and that these 'chooser' or 'non-chooser' roles would be randomly attributed. Covertly, all subjects were attributed the chooser role.

Transcranial direct current stimulation (tDCS) procedure

During the experiment, we applied tDCS over the rmPFC of participants using a commercially available multi-channel stimulator that allows stimulation of up to 16 participants with individually tailored stimulation protocols. Here, we applied anodal tDCS over the MNI coordinates [xyz=6, 59, 10] defined based on the fMRI experiment, using the peak coordinates of the results from the conjunction analysis for positive and negative competition errors, as defined in Figure 3B and Table S1. The electrode were placed on the scalp using the frameless stereotactic software Brainsight 2.0 and the T1-weighted anatomical scan of each participant, which enabled us to transform the standard MNI coordinates into individual scalp coordinates. The rmPFC was stimulated using a 5x5 cm (anodal) electrode allowing current to flow towards the 9x9 cm reference (cathodal) electrode placed on the vertex. The positioning and the large size of the reference electrode [S11] was used to circumvent influences on other cortical areas potentially relevant for executive control. In line with established procedure and safety recommendations, we stimulated with 1 mA current strength and used a 30 s ramping current increase to avoid tingling sensations caused by abrupt onset and offset of tDCS. In the tDCS group, stimulation lasted as long as the experimental session (never more than 30 minutes). In the sham group, we applied 30 s of ramping stimulation so that subjects could feel the onset of the current, but the stimulation was then immediately decreased and the tDCS was turned off after 1 minute. This procedure was adopted in order to avoid metacognitive biases associated with the detection by the participant of his/her stimulation condition. Although participants in the tDCS condition rated significantly higher on the post-test question “Did you felt the stimulation during the task?” (p=0.049, Mann-Whitney), it must be noted that answers were on average negative in both group (sham group: 1.81±0.55, tDCS group: 2.35±0.7; with ‘0’ corresponding to “Not at all”, 3 corresponding to ”Not sure”, and ‘7’ correspond to “Completely sure”). In addition, tDCS had no effect on self-reported mood state (from 0/”very negative” to 7/”very positive”) following the experiment (p=0.89 [t-test]) nor on the change between pre- and post-experiment mood state (p=0.99 [repeated-measures ANOVA], means reported Table S4). In summary,

the effect of tDCS on social learning could not be attributed to modulations of general emotional or metacognitive states during the experiment.

Computational modeling: estimation and comparison procedures

In order to capture the dynamic influence of victories and defeats on the competitive choices of the participants (Figure 2A) and to estimate the learning rates associated with the two stimulation regimes, we used a Rescorla-Wagner algorithm[S12] combined with a softmax decision policy (Equation 2) assuming that the probability of choosing to defy one opponent i over another opponent j depends on the relative difference in the social dominacen status associated with this opponent *versus* with the other opponent (both being presented simultaneously on the screen).

$$SDSi\,(t+1) = SDSi\,(t) + \alpha * (R(t) - SDSi\,(t)) \quad \text{(Equation 2)}$$

$$p(i) = \frac{\exp(\beta * SDS_i)}{\exp(\beta * SDS_i) + \exp(\beta * SDS_j)} \quad \text{(Equation 3)}$$

Equation (2) is identical to Equation (1) defined above. It determines how the social dominance status of the chosen opponent (SDS_i or SDS_j) is updated according to the feedback received. R(t) is the 'reward' collected in the ongoing trial: it was arbitrarily set to 0 for defeats, 0.5 for evens, and 1 for victories, so that higher competitive values in any given trial corresponds to weaker opponents (against who the subject recently "harvested" victories) and vice-versa. The free parameter α represents the learning rate of the model (high α imply high volatility in the representation of values) and was constrained between 0 and 1.

Equation (3) defines a softmax function which is a stochastic decision rule that calculates the probability of choosing an option given the available alternative. The free parameter β is the inverse temperature parameter and dictates to which extent the decision is deterministic relative to the SDS values of available opponents *i* and *j* (a high absolute β values mean that choices are strongly driven by

SDS values). Note that contrary to economic decision-making experiments, here β could be negative to account for a general preference towards stronger opponents.

Because subjects' performance could affect competitive outcomes, we elaborated a refined version of the Rescorla-Wagner rule which allowed prediction errors to be weighted by the performance of the subjects.

$$SDSi\,(t+1) = SDSi(t) + \omega * (1-P) * (R - SDSi(t)\,) + (1-\omega) * \alpha * (R - SDSi(t)\,)$$

(Equation 4,victory)

$$SDSi(t+1) = SDSi(t) + \omega * P * (R - SDSi(t)\,) + (1-\omega) * \alpha * (R - SDSi(t)\,)$$

(Equation 5, defeat)

Where ω is the performance weighting parameter (higher ω means higher determination of prediction errors by performances) and P is the normalized performance of the subject on the current trial, computed as:

$$P_t = 1 - \frac{\log(RT_t) - \min(\log(RT))}{\max(\log(RT)) - \min(\log(RT))} \qquad \text{(Equation 6)}$$

As previously described in the subsection dedicated to the fMRI analysis, the competitive prediction error term (referred to as "cPE") corresponds to the term $R - SDSi$.

In a preliminary step, we tested six variants of this reinforcement-learning scheme, to explained either spontaneous or control choices (Figure S2). Note that in both cases, the learning functions were only fed with the feedbacks delivered in the spontaneous condition, since no feedback was given in the control condition. These six models varied on two dimensions. First, three models used the same learning rate for victories and defeats (RL1, RL2, RL3), whereas the remaining models used separate learning rates for these different outcomes (RL4, RL5, RL6). Second, two models updated the values of

the opponent even after incorrect answers in the perceptual game (RL1, RL4), two models did not update opponents' value after incorrect answers (RL2, RL5) and the remaining models did not update after incorrect answer *and* included a performance-weighting parameter. The first dimension was used to determine whether social hierarchy learning is more likely driven by a single, symmetrical or a dual, asymmetrical learning process. The second dimension was used to address sensitivity to one's own accuracy in the update of opponents' value.

Estimation of optimal parameters and goodness of fit were performed in each subject independently using the variational Bayes (VB) approach proposed by [S13] and implemented in a validated Matlab toolbox (VBA toolbox, available at: https://code.google.com/p/mbb-vb-toolbox/). One advantage of the VB approach is that goodness of fit can be measured by the free energy, which depends not only on the maximum likelihood of observed choices and number of parameters but also on the posterior probability of the estimated parameters. In our model selection analyses, we used both the Free Energy and Akaike Information Criterion as metrics for goodness of fit. For all learning rates (α) and the performance-weighting parameters (ω), pseudo-uniform distribution of the priors over the [0,1] interval were used. For the priors of inverse temperature, we used Gaussian distributions centered on 0 and 1.4 for spontaneous and control choices, respectively; a value of 1.4 corresponds to the *a priori* expectation that the selection of the weakest opponents would reach 80% (for a RL problem where values range from 0 to 1; note however that given the extremely high variance used to generate prior distributions, this positively centering could only have very negligible effect on parameter estimation). Conversely, a value of 0 means that inverse temperature (β) are *a priori* equally likely to be positive or negative, which is justified by the absence of objective utility maximizing norm ruling spontaneous choices (subjects could perfectly prefer to defy systematically the strongest opponents). Because we had no specific hypothesis regarding the exact distribution of inverse temperature parameters in the population, we used a very large prior variance ($\sigma^2 = 10000$), which approximated a flat non-informative

distribution over the interval of interest. Note that using shrinkage non-informing priors (i.e. distributed around 0 with variance 1) instead of flat priors for the estimation of learning rate did not change conclusions based statistical inference.

SUPPLEMENTAL REFERENCES

S1. Ligneul, R., Sescousse, G., Barbalat, G., Domenech, P., and Dreher, J.-C. (2013). Shifted risk preferences in pathological gambling. Psychol. Med. *43*, 1059–1068.

S2. Carver, C. S., and 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. Available at: http://doi.apa.org/getdoi.cfm?doi=10.1037/0022-3514.67.2.319.

S3. Zink, C. F., Tong, Y., Chen, Q., Bassett, D. S., Stein, J. L., and Meyer-Lindenberg, A. (2008). Know Your Place: Neural Processing of Social Hierarchy in Humans. Neuron *58*, 273–283.

S4. Wager, T. D., and Nichols, T. E. (2003). Optimization of experimental design in fMRI: a general framework using a genetic algorithm. Neuroimage *18*, 293–309.

S5. Glascher, J. (2009). Visualization of group inference data in functional neuroimaging. Neuroinformatics *7*, 73–82.

S6. Wilson, R. C., and Niv, Y. (2015). Is Model Fitting Necessary for Model-Based fMRI? PLOS Comput. Biol. *11*, e1004237.

S7. Hatcher, J., and Hall, L. A. (2009). Psychometric Properties of the Rosenberg Self-Esteem Scale in African American Single Mothers. Issues Ment. Health Nurs. *30*, 70–77.

S8. Cohen, S., Kamarck, T., and Mermelstein, R. (1983). A global measure of perceived stress. J. Health Soc. Behav., 385–396.

S9. Simmons, C. H. C. H., Wehner, E. a. E. A., Tucker, S. S., and King, C. S. (1988). The Cooperative/Competitive Strategy Scale: A Measure of Motivation to Use Cooperative or Competitive Strategies for Success. J. Soc. Psychol. *128*, 199–205.

S10. Levenson, H. (1981). Research with the locus of control construct. In, L. HM, ed. (Academic Press, New York), pp. 15–63.

S11. Nitsche, M. A., and Paulus, W. (2001). Sustained excitability elevations induced by transcranial DC motor cortex stimulation in humans. Neurology *57*, 1899–1901.

S12. Sutton, R. S., and Barto, A. G. (1998). Reinforcement Learning: An Introduction (MIT Press. Cambrigde, Massachussets.).

S13. Daunizeau, J., Adam, V., and Rigoux, L. (2014). VBA: A Probabilistic Treatment of Nonlinear Models for Neurobiological and Behavioural Data. PLoS Comput. Biol. *10*.