



The Role of Frontostriatal Systems in Instructed Reinforcement Learning: Evidence From Genetic and Experimentally-Induced Variation

Nathan Tardiff*, Kathryn N. Graves and Sharon L. Thompson-Schill

Department of Psychology, University of Pennsylvania, Philadelphia, PA, United States

Instructions have a powerful effect on learning and decision-making, biasing choice even in the face of disconfirming feedback. Detrimental biasing effects have been reported in a number of studies in which instruction was given prior to trial-and-error learning. Previous work has attributed individual differences in instructional bias to variations in prefrontal and striatal dopaminergic genes, suggesting a role for prefrontally-mediated cognitive control processes in biasing learning. The current study replicates and extends these findings. Human subjects performed a probabilistic reinforcement learning task after receiving inaccurate instructions about the quality of one of the options. In order to establish a causal relationship between prefrontal cortical mechanisms and instructional bias, we applied transcranial direct current stimulation over dorsolateral prefrontal cortex (anodal, cathodal, or sham) while subjects performed the task. We additionally genotyped subjects for the COMT Val158Met genetic polymorphism, which influences the breakdown of prefrontal dopamine, and for the DAT1/SLC6A3 variable number tandem repeat, which affects expression of striatal dopamine transporter. We replicated the finding that the COMT Met allele is associated with increased instructional bias and further demonstrated that variation in DAT1 has similar effects to variation in COMT, with 9-repeat carriers demonstrating increased bias relative to 10-repeat homozygotes. Consistent with increased top-down regulation of reinforcement learning, anodal subjects demonstrated greater bias relative to sham, though this effect was present only early in training. In contrast, there was no effect of cathodal stimulation. Finally, we fit computational models to subjects' data to better characterize the mechanisms underlying instruction bias. A novel choice bias model, in which instructions influence decision-making rather than learning, was found to best account for subjects' behavior. Overall, these data provide further evidence for the role of frontostriatal interactions in biasing instructed reinforcement learning, which adds to the growing literature documenting both costs and benefits of cognitive control.

Keywords: reinforcement learning, prefrontal cortex, striatum, tDCS, dopamine, COMT, DAT

OPEN ACCESS

Edited by:

Xiaolin Zhou,
Peking University, China

Reviewed by:

Xiaochu Zhang,
University of Science and Technology
of China, China
Jie Hu,
University of Zurich, Switzerland

*Correspondence:

Nathan Tardiff
ntardiff@sas.upenn.edu

Received: 15 May 2018

Accepted: 12 November 2018

Published: 17 December 2018

Citation:

Tardiff N, Graves KN and
Thompson-Schill SL (2018) The Role
of Frontostriatal Systems in Instructed
Reinforcement Learning: Evidence
From Genetic and
Experimentally-Induced Variation.
Front. Hum. Neurosci. 12:472.
doi: 10.3389/fnhum.2018.00472

INTRODUCTION

Successful learning and decision-making require a balance between exploiting prior information and learning from new experiences that may contradict it. One pervasive source of prior information in humans is instruction from others. Such instruction has clear benefits on both ontogenetic and historical timescales, allowing children to rapidly learn about the world and allowing culture and technology to develop and evolve (Tomasello, 1999). On an individual level, advice and information received from friends, professionals, and the media shape our view of the world and our choices.

The alternative to learning from advice and instruction is learning from direct experience of the world. One well-characterized method of learning from experience is reinforcement learning (RL), in which actions are selected so as to maximize reward (see Niv, 2009; Dolan and Dayan, 2013 for reviews). Recent work exploring the effects of instruction on RL has found that accurate advice can significantly improve performance (Biele et al., 2009; Doll et al., 2011). Yet such instruction is often detrimental when it is inaccurate. A potential consequence of inaccurate instruction and, more generally, inaccurate prior information, is *confirmation bias*, whereby data that are consistent with a prior hypothesis are sought, attended to, or valued over disconfirming data, which are neglected, filtered, or devalued (Nickerson, 1998). Confirmation bias is thought to be pervasive in human reasoning, affecting children and adults' scientific reasoning as well as that of professional scientists (Mahoney, 1977; Kuhn, 1989; MacCoun, 1998; Hergovich et al., 2010).

Biases have been induced in both social and non-social RL tasks utilizing various methods of information delivery. Information indicative of the moral character of computerized partners in a repeated trust game biases share decisions to "good" and "bad" partners despite identical behavior by the computer (Delgado et al., 2005; Fareri et al., 2012). Poor advice provided by fellow subjects impairs performance on the Iowa Gambling Task (Biele et al., 2009, 2011). Finally, in an RL task in which subjects learn to discriminate among pairs of probabilistically rewarded symbols, subjects instructed that a particular symbol is desirable persist in choosing that symbol more than would be expected given negative feedback, selecting it more frequently than symbols rewarded at an equal rate (Doll et al., 2009, 2011, 2014; Staudinger and Büchel, 2013). In sum, instructional biases appear to be persistent, and they are only partially ameliorated by feedback.

The neural substrates of instructed learning are still emerging, though as in uninstructed RL, frontostriatal areas are commonly implicated (Doll et al., 2009; Wolfensteller and Ruge, 2012). Neuroimaging has supported a role for prefrontal cortex (PFC) in representing instructions or prior information (Li et al., 2011; Fouragnan et al., 2013), with activity in instructed conditions found in dorsolateral PFC (DLPFC) and medial PFC. Connectivity analyses further support a role for PFC, reporting increased functional connectivity between frontal and striatal regions during instructed/prior knowledge conditions, consistent

with top-down influence on striatal reward prediction errors (Li et al., 2011; Fouragnan et al., 2013).

Evidence of PFC altering striatal learning comports well with accounts of PFC-mediated cognitive control biasing or filtering information in other brain regions. Such top-down modulation focuses information processing on task-relevant information while suppressing irrelevant information (Shimamura, 2000; Miller and Cohen, 2001; Chrysikou et al., 2014). Performance should be optimal when the level of filtering is suitable to the demands of the task (Chrysikou et al., 2014). Consequently, increased top-down control can incur both costs and benefits. This is the case in instructed RL, where instruction-induced bias has been shown to vary according to individual differences in PFC dopaminergic tone caused by the catechol-O-methyltransferase (COMT) Val158Met genetic polymorphism (Doll et al., 2011). In particular, the Met allele, which has been shown to confer benefits in tests of working memory and cognitive control as compared to the Val allele (Durstewitz and Seamans, 2008; Witte and Flöel, 2012), is associated with a cost in the form of increased adherence to inaccurate instructions.

The goal of the present study was three-fold. First, we sought to replicate the effect of COMT on instructed reinforcement learning, providing further evidence for the role of PFC-mediated top-down control in biasing RL.

Second, we aimed to expand the understanding of the impact of striatal dopaminergic genes on instructed RL. While Doll et al. (2011) examined the effects of genetic polymorphisms specific to approach or avoidance learning in the striatum, we examined the effect of the DAT1/SLC6A3 variable number tandem repeat (VNTR), which affects striatal dopamine (DA) reuptake by altering expression of the dopamine transporter (DAT; Vandenberg et al., 1992; Faraone et al., 2014). Though there are conflicting reports on the exact effects of the DAT1/SLC6A3 VNTR, a recent meta-analysis suggests that in healthy individuals the 9-repeat allele is associated with increased DAT expression in human striatum, and thus potentially more efficient reuptake of DA as compared to the 10-repeat variant (Faraone et al., 2014; cf. Costa et al., 2011). Striatal DA levels have previously been linked to cognitive flexibility (Cools and D'Esposito, 2009; Beeler et al., 2010; Garcia-Garcia et al., 2010), making DAT1 a plausible modulator of instructed RL.

Finally, while genetic and neuroimaging evidence is compelling, it falls short of establishing a causal role for PFC in biasing RL. We therefore hoped to establish this causal link by directly modulating PFC via transcranial direct current stimulation (tDCS). In keeping with a costs/benefits framework, we predicted that anodal stimulation—which has been successfully applied to PFC in order to improve cognitive control (Fregni et al., 2005; Cattaneo et al., 2011; Zaehle et al., 2011; Nozari and Thompson-Schill, 2013; Karuza et al., 2016)—would lead to increased bias due to increased top-down regulation. Cathodal stimulation over PFC has produced inconsistent results in cognitive domains (Jacobson et al., 2012; Nozari et al., 2014). However, supporting the costs/benefits framework, it has been linked to decreased working memory (Zaehle et al., 2011) and selective attention (Nozari et al., 2014; Zmigrod et al., 2016), but improved dual

task performance (Filmer et al., 2013) and cognitive flexibility (Chrysikou et al., 2013). Therefore we tentatively predicted that cathodal stimulation would lead to decreased bias due to decreased top-down control of RL.

METHODS

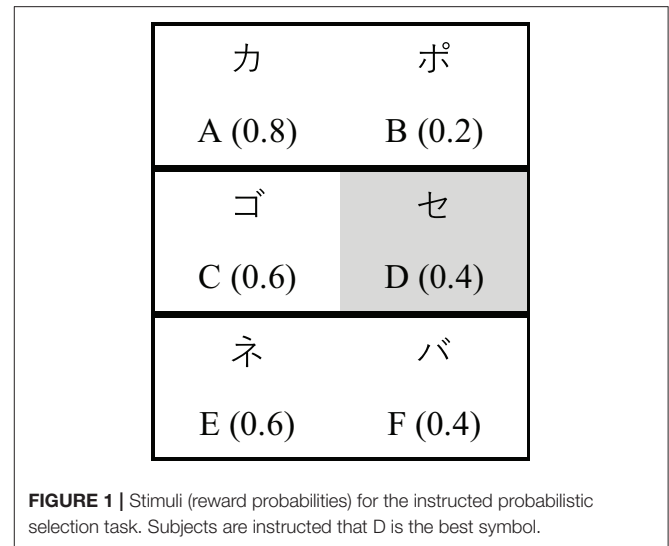
Subjects

One-hundred and twenty-six right-handed subjects (42 per condition, 80 female, $M_{age} = 22.20$ years) participated in the study, receiving \$20 in compensation, regardless of performance. Informed consent was obtained from each subject in accordance with the University of Pennsylvania IRB. Subjects were randomly assigned to stimulation condition. We excluded a total of 23 subjects from the analyses for failure to meet the performance criteria described in section Performance Criteria (9 anodal, 6 cathodal, 8 sham), for a final sample of 103 (65 female, $M_{age} = 21.84$ years). Of these subjects, genotyping failed for one subject. For the Val158Met single-nucleotide polymorphism (SNP) of the COMT gene (rs4680), frequencies per allele in the final sample were 34:53:15 (Val/Val:Val/Met:Met/Met). For the DAT1/SLC6A3 VNTR in the 3' untranslated region, frequencies per allele were 65:26:6:2:1:1:1 (10/10:9/10:9/9:10/11:8/9:8/8:6/10). Subjects were placed in a 10/10 group if they had two repeats of 10+; otherwise they were placed in a 9-repeat carrier group (67 10/10, 35 9c). Neither gene differed from Hardy-Weinberg equilibrium either across the whole sample (all $ps > 0.14$) or within racial/ethnic subgroups (all $ps > 0.15$; see **Supplementary Tables 3–6** for sample demographic breakdown). There was no association between COMT and DAT genotypes ($p > 0.35$, Fisher's Exact Test), nor were there any associations between the two genes and stimulation condition (all $ps > 0.3$). For the dopamine genotype composite, the distribution of subjects was: 25:43:27:7 (0:1:2:3). The composite was also not significantly associated with stimulation condition ($p = 0.09$).

Materials and Procedure

Subjects completed an instructed probabilistic selection task (iPST), presented on a 13" laptop computer via PsychoPy (Peirce, 2009). This task required subjects to learn the value of symbols initially presented in 3 pairs (AB, CD, EF; **Figure 1**). Within each pair, one symbol had a higher probability of reward. Symbols were rendered as Japanese Hiragana characters, and the assignment of Japanese character to underlying stimulus was randomized across subjects. During the instructions, each symbol was presented individually for 5 s to familiarize subjects with the stimuli. Crucially, when symbol D was presented the screen also displayed the following false advice: "This symbol has the best chance of being correct."

During the training phase, subjects had to learn the value of each symbol via probabilistic feedback, which was delivered according to the symbol's underlying P(reward). Importantly, subjects were expected to learn to select the more highly rewarded symbol within each pair. Subjects completed 4 training blocks. Each block contained 20 repetitions of each pair, for a total of 60 trials per block and 240 total training trials. Trial order and feedback were randomized within each



block. During the test phase, all possible symbol pairings were presented (e.g., AB, AC, AD, AE, AF, ...) without feedback. Each pair was presented 6 times, for a total of 90 trials. Order was randomized across subjects. See **section 1.1** of the Supplementary Material for further details regarding task design and presentation.

Performance Criteria

Subjects had to meet the following performance criteria for the uninstructed symbols in order to be included in the analyses: $\geq 60\%$ accuracy on the AB pair and $\geq 50\%$ accuracy on the EF pair in at least one training block after the first block, with both criteria met in the same block. These criteria are similar to training phase learning criteria used in previous reports (Frank et al., 2007; Doll et al., 2009, 2011, 2014), but were relaxed slightly for AB to allow for additional variability in learning performance, given a previous report of tDCS effects on this pair (Turi et al., 2015). Subjects were also excluded if they failed to respond on $> 10\%$ of training trials.

In addition to excluding subjects who failed to pay attention or learn, these criteria helped ensure that subjects with arbitrary biases for one of the uninstructed symbols were excluded from the analyses. However, to further protect against arbitrary affinities introducing bias into the between-group analyses, we further tested for the presence of genotype or stimulation differences in the first 10 training trials of the uninstructed training pairs. There were no significant effects (all $ps > 0.10$), indicating that none of our genotype or stimulation groups entered the training phase with arbitrary stimulus preferences.

Genotyping

DNA samples were collected via Oragene saliva kits (DNA Genotek) and genotyped at the Penn Molecular Profiling Facility using standard procedures (see **section 1.2** of the Supplementary Material).

TABLE 1 | Stimulation procedure and duration for verum stimulation (sham was identical except stimulation only lasted for 30 s, at the onset of the fixation period).

Phase	tDCS	Duration
Instructions	No	Variable
Fixation	Yes	3 min
Training	Yes	17 min
Test Instructions	No	Variable
Test	No	6 min

Transcranial Direct Current Stimulation

We delivered direct current via saline-soaked sponge electrodes with a 25 cm² surface area. Current was generated by a continuous current stimulator (Magstim Eldith 1 Channel DC Stimulator Plus, Magstim Company Ltd., Whitland, Wales). In all conditions, 1.0 mA direct current was applied after a 30 s ramp-up period and was followed by a 30 s ramp-down. In the verum conditions, current was applied for 20 min. Stimulation was applied for only 30 s during sham. In the anodal condition, the anode was placed over F7, in accordance with the 10–20 international system, and the cathode was placed over the right supraorbital. This placement was reversed in the cathodal condition.

The F7-RSO montage was chosen because current modeling (HDExplore Software, v2.3, SOTERIX) suggested it would maximize current through DLPFC sites found to be active during instructed reinforcement learning conditions (Li et al., 2011; Fouragnan et al., 2013). Stimulation at F7 has been shown to modulate prefrontally-mediated cognitive control across a range of tasks (Lupyan et al., 2012; Chrysikou et al., 2013; Nozari et al., 2014). The procedure for each subject is outlined in **Table 1**. Stimulation began 180 s prior to the start of the first trial while subjects were presented with a fixation cross. Stimulation has not been shown to produce after-effects at 1.0 mA unless applied for at least 3 min, and thus this period gives stimulation time to take full effect (Nitsche and Paulus, 2000). Additionally, though stimulation ended after the training phase, after-effects have been reported up to an hour after stimulation lasting 9–13 min, so it is possible tDCS could directly affect performance at test in addition to its indirect effect through modifying performance during training (Nitsche and Paulus, 2001; Nitsche et al., 2003).

Data Analysis

Statistical analyses were conducted in R (R Core Team, 2018) using logistic mixed models implemented in the lme4 package (Bates et al., 2015). By modeling both fixed and random effects, these models controlled for the non-independence inherent in within-subjects data. All models included random intercepts for subjects and random slopes for within-subjects variables and their interactions (Schieleth and Forstmeier, 2009; Barr et al., 2013). When making between group comparisons of factors with more than two levels without planned comparisons, the significance of main effects and interactions were computed using the car package (Fox and Weisberg, 2011). *Post-hoc* comparisons were computed using the lsmeans package

(Lenth, 2016). Significance levels for *post-hoc* comparisons were corrected using the Bonferroni-Holm method (Holm, 1979). Permutation tests were conducted via Monte Carlo sampling (1.0e6–1 permutations) using the perm package (Fay and Shaw, 2010).

Computational Modeling

Reinforcement learning models were fit to each subject's data in order to evaluate hypotheses regarding the mechanisms of instructional bias. Models were fit by maximizing the log likelihood of the data using MATLAB's fmincon (Mathworks, MA, USA). To avoid local minima, each model fit was repeated 25 times from different random starting points, using RMSEARCH. All models were fit to both training and test phase data. For the training phase, fits were optimized to account for subjects' trial-wise choices; for the test phase, they were optimized to result in learned Q-values after training that best account for choices during test (Frank et al., 2007).

Standard Model

This model implements a standard Q-learning model with separate learning rates for gains and losses (Frank et al., 2007). The value of each stimulus is updated according to the following learning rule:

$$Q_{t+1}(s) = Q_t(s) + [\alpha_g * \delta_t]_+ + [\alpha_l * \delta_t]_-$$

$$\delta_t = r_t - Q_t(s)$$

where $Q_t(s)$ is the action value of stimulus s at trial t , r_t is the reward (0 for losses, 1 for gains), and δ_t is the reward prediction error. The learning rate α_g applies only to gain trials, while the learning rate α_l applies only to loss trials.

Choice in the standard model and subsequently described models was implemented via a softmax function:

$$P_t(s_1) = \frac{\exp\left(\frac{Q_t(s_1)}{\beta}\right)}{\exp\left(\frac{Q_t(s_1)}{\beta}\right) + \exp\left(\frac{Q_t(s_2)}{\beta}\right)}$$

where $P_t(s_1)$ is the probability of choosing symbol s_1 over symbol s_2 , and β is a temperature parameter determining the extent to which choice is deterministic vs. random.

For this model and subsequent models, we placed the following bounds on the parameters: $\alpha \in [0.002, 1]$; $\beta \in [0.06, 20]$. The temperature parameter was additionally constrained by an empirical prior (Gershman, 2016): $\frac{1}{\beta} \sim \text{Gamma}(5.09, 0.83)$. Q-values for all stimuli were initialized at 0.5.

Learning Bias Model (Doll et al., 2011)

The learning bias model is identical to the standard model in all respects except that when symbol D is chosen, the baseline learning rate is distorted as follows:

$$Q_{t+1}(D) = Q_t(D) + [\alpha_g * \alpha_{bg} * \delta_t]_+ + \left[\frac{\alpha_l}{\alpha_{bl}} * \delta_t\right]_-$$

where α_{bg} increases the learning rate for instruction-consistent feedback (gains), α_{bl} diminishes the learning rate for instruction-inconsistent feedback (losses), and $\alpha_b \in [1, 10]^1$.

Bayesian Hypothesis Testing Model (Doll et al., 2011)

This model accounts for the possibility that the bias lies not in learning the value of the instructed stimulus D but in the decision to choose D. In this case, the choice bias requires that learners achieve a certain level of confidence that D is rewarded below chance before they abandon it. This model implements a Bayesian Q-learner with $Q_t(s) \sim \text{Beta}[\alpha_t(s), \beta_t(s)]$. After reward feedback, posterior Q-value distributions are updated as:

$$Q_{t+1}(s) \sim \text{Beta}[\alpha_t(s) + r_t, \beta_t(s) + (1 - r_t)]$$

which increments α by 1 after gains and β by 1 after losses. Additionally, after every trial the α and β counts decay toward uniform, controlled by free parameters γ_α and γ_β ; $\gamma \in [0, 1]$ (0 is full decay and 1 is no decay). Choice is implemented by submitting the mean of each symbol's beta distribution to the softmax function above. Crucially, when the instructed stimulus is encountered, a decision bias is implemented as follows:

$$P_t(s_{\text{alt}}) = \frac{\exp\left(\frac{0.5}{\beta}\right)}{\exp\left(\frac{0.5}{\beta}\right) + \exp\left(\frac{\mu_t(D) + \phi * \sigma_t(D)}{\beta}\right)}$$

with $\phi \in [0, 20]$ and $P_t(D) = 1 - P_t(s_{\text{alt}})$. This decision rule dictates that the mean value of D must be $> \phi$ standard deviations of D below chance before it is more probable that the alternative symbol, s_{alt} , is chosen. Thus, the more certain the learner is of the value of D, the lower the bias.

Decision Bias Model

Though the Bayesian hypothesis testing model has provided a reasonable fit to some subjects' training data and has been shown to be sensitive to individual differences, it has not overall outperformed the standard model in explaining training phase performance (Doll et al., 2009, 2011). It also compares the value of D to chance instead of to the value of the alternative stimulus, making it less effective as a possible model of test phase performance. Furthermore, interpretation of this model in comparison to the standard uninstructed model is complicated by the fact that they are not nested models. Therefore, we also implemented a novel decision bias model. This model uses the standard Q-learner described above, but the softmax decision rule is modified for choices involving the instructed stimulus in a manner similar to the Bayesian hypothesis testing model:

$$P_t(s_{\text{alt}}) = \frac{\exp\left(\frac{Q_t(s_{\text{alt}})}{\beta}\right)}{\exp\left(\frac{Q_t(s_{\text{alt}})}{\beta}\right) + \exp\left(\frac{Q_t(D) + \rho}{\beta}\right)}$$

with $\rho \in [0, 1]$. The free parameter ρ determines how much greater the value of the alternative symbol must be before it

¹In order to prevent learning rates exceeding 1.0, the learning bias parameters were also constrained such that $\alpha_b \leq \alpha^{-1}$ (Doll et al., 2009).

is more probable that it is chosen over D. Therefore, unlike the Bayesian model, this model: (a) assumes a fixed bias; (b) compares the value of D to the alternative symbol, making it more appropriate as a model of test phase choice; and (c) contains the standard model as a special case ($\rho = 0$), ensuring differences in fit will be attributable to the presence of the bias and not to differences in the learner.

Model Comparison

Goodness of fit was assessed using Akaike information criteria (AIC). We additionally submitted the AIC values to a Bayesian random effects analysis, which assumes there is a distribution of models in the population and attempts to identify which model is most prevalent. The quantity resulting from this analysis is a protected exceedance probability (PEP), which is the probability that a given model is the most frequent in the population, above and beyond chance (Rigoux et al., 2014). PEPs were computed using the VBA toolbox (Daunizeau et al., 2014). Model comparison was then made on the basis of both AIC and PEPs.

RESULTS

We begin by reviewing general performance across the sample. We then examine genotypic differences in instructional bias. To this end, we first attempt to replicate the effect of COMT genotype. We then extend the investigation of the influence of dopaminergic genes on instructional bias to the DAT1 gene. In brief, we partially replicated the effect of COMT and found effects of DAT1 on instructional bias as well. Motivated by these findings, we next ask whether a dopamine composite variable constructed from the COMT and DAT variables captures additional aspects of performance. These analyses demonstrated an overall graded effect of the dopamine composite on performance and also uncovered a small group of subjects who demonstrated more extreme bias. We then ask if we can causally manipulate instructional bias with tDCS, finding that anodal stimulation had a small but significant effect on performance during training. Finally, we fit computational models to test potential mechanisms underlying instructional bias, finding evidence in favor of a model incorporating a bias on the decision to choose the instructed stimulus, rather than a bias on the learned value of the instructed stimulus.

General Performance: Training Phase Instructed Learning

We first conducted analyses of choice behavior during training. In all analyses, accuracy was binary coded (incorrect: 0, correct: 1), where correct is defined as choosing the stimulus with the higher probability of reward, regardless of whether it was rewarded on that trial. Trial Type was treatment coded (CD: 0, EF: 1). This coding allows direct assessment of how much instruction biased learning. Block was reverse Helmert coded in order to capture learning-related changes in the mean level of responding across training (i.e., Block 2 was compared to Block 1, Block 3 was compared to the mean of Blocks 1 and 2, and Block 4 was compared to the mean of all prior blocks). We assessed the effects of genotype and stimulation

both by examining performance on the CD trials alone, and by contrasting performance on CD with the equally rewarded but uninstructed EF pair. Given our between-subjects design, this latter contrast serves to account for additional variance in learning unrelated to instructional control. Therefore, instructed training models included all two-way and three-way interactions of genotype or tDCS condition, Trial Type, and Block.

Subjects were below chance on the CD pair ($\beta = -0.27$, $z = -2.86$, $p = 0.004$). Performance was significantly better on the EF pair ($\beta = 0.93$, $z = 6.96$, $p < 0.0001$), validating the success of the instructional manipulation. Despite poor overall performance on the CD pair, subjects continued to learn away from the instructions throughout training, as demonstrated by the significance of all three Block regressors (Block 2 vs. 1: $\beta = 0.32$, $z = 2.81$, $p = 0.005$; Block 3 vs. (1, 2): $\beta = 0.24$, $z = 2.34$, $p = 0.02$; Block 4 vs. (1,2,3): $\beta = 0.31$, $z = 3.13$, $p = 0.002$).

Uninstructed Learning

Variable coding in uninstructed training models was the same as above, except Trial Type was effect coded (AB: 1, EF: -1). The three-way interactions were not included in these models as there were no hypotheses relevant to these contrasts.

Subjects performed significantly above chance on uninstructed trials ($\beta = 1.12$, $z = 14.67$, $p < 0.0001$). There was an effect of Trial Type ($\beta = 0.46$, $z = 10.55$, $p < 0.0001$), indicating that subjects performed significantly better on the AB pair over the EF pair, in line with the relative difficulty of the two discriminations. Subjects continued to learn throughout training, though the magnitude of this effect was numerically smaller in later blocks (Block 2 vs. 1: $\beta = 0.40$, $z = 4.79$, $p < 0.0001$; Block 3 vs. (1, 2): $\beta = 0.22$, $z = 2.74$, $p = 0.006$; Block 4 vs. (1, 2, 3): $\beta = 0.18$, $z = 2.17$, $p = 0.03$). There was additionally a Trial Type x Block 2 vs. 1 interaction ($\beta = 0.31$, $z = 3.97$, $p < 0.0001$), indicating a steeper learning trajectory for AB over EF during the initial blocks of the task, which again is unsurprising given the relative ease of the AB discrimination.

General Performance: Test Phase

The training and test phases are purported to represent different processes subserved by different neural systems (Frank et al., 2007). While the training phase is supposed to reflect hippocampally- and frontally-mediated memory and hypothesis-testing processes, the test phase is designed to give a “purer” measure of striatally-learned reinforcement values. The standard approach to assessing performance at test is to examine performance on trials in which a stimulus of interest is included in novel pairings, giving an estimate of how well underlying reward values were learned during training.

Two measures from the literature were used to assess the effect of instruction on test phase performance (Doll et al., 2014). The first analysis compared performance on Avoid-D (AD, DE) vs. performance on Avoid-F (AF, CF). For both measures, the target stimulus should not be chosen, as it has been paired with stimuli that had a higher probability of reward during training. Given that D and F had identical reward probabilities during training, subjects should perform equally well on both measures. However, if instruction biased the ultimate reward values subjects learned,

or if subjects’ choices continue to be biased at test, they should avoid D at a lower rate than they avoid F.

Choice Type was entered as an effect-coded factor (Avoid-D: 1, Avoid-F: -1) in a logistic mixed effects model of choice performance. The intercept was significant ($\beta = 0.73$, $z = 5.53$, $p < 0.0001$), indicating that subjects’ overall avoidance of D and F was above chance. There was also a main effect of Choice Type ($\beta = -0.61$, $z = -5.43$, $p < 0.0001$). As expected, subjects showed a confirmation bias effect, avoiding D significantly less than they avoided the equally rewarded symbol F.

The second analysis of instructed learning examined performance on DF trials in order to directly compare the relative subjective value of the two stimuli. A greater effect of instruction on learning, and thus a greater bias, should be associated with an increased tendency to choose D over F.

In this model, choice on DF trials was the dependent variable (D: 1, F: 0). The intercept was significant ($\beta = 1.58$, $z = 5.86$, $p < 0.0001$). Subjects demonstrated a strong bias—they were almost five times more likely to choose D, as indicated by an odds ratio (OR) of 4.86. In sum, our training and test results replicate previous investigations (Doll et al., 2009, 2011, 2014) and confirm that the instructional manipulation was successful.

COMT: Training Phase Instructed Learning

We next sought to replicate the effect of the COMT Met allele on adherence to the instructions during training (Doll et al., 2011). COMT genotype was effect coded. All other variables were coded as above.

There was a significant COMT x Trial Type interaction ($\chi^2_{(2)} = 13.94$, $p = 0.0009$). Met homozygotes were significantly worse overall on the instructed pair (**Figure 2A, Supplementary Table 7**), as compared to both heterozygotes ($\beta = -0.98$, $z = -3.70$, $p_{corrected} = 0.001$) and Val homozygotes ($\beta = -0.92$, $z = -3.26$, $p_{corrected} = 0.006$). Met homozygotes also demonstrated better performance at a trend level on the uninstructed EF pair compared to Val/Met subjects, but this did not survive correction for multiple comparisons ($\beta = 0.41$, $z = 1.87$, $p = 0.06$, $p_{corrected} = 0.25$). Notably, no other comparisons reached significance, including the comparison of instructed performance between Val/Met and Val/Val subjects (all $ps > 0.2$), indicating impaired performance was specific to Met homozygotes.

Because our Met/Met group was somewhat small ($N = 15$) due to the low frequency of this genotype in the general population (Auton et al., 2015), we took a number of additional steps to ensure these results were not spurious. First, we reran our analyses comparing Val homozygotes to Met carriers (Metc), which was also the analysis performed by Doll et al. (2011). In this case, we failed to replicate the effect of Met-carrier status on instructed learning. The Metc x Trial Type interaction was not significant ($\chi^2_{(1)} = 0.16$, $p = 0.69$), nor were there any other significant effects of Met carrier status (all $ps > 0.42$). We then asked whether the full COMT model or the Metc model provided a better fit to the data, finding that the COMT model was a modestly better fit, despite including additional parameters ($AIC_{COMT} = 19966$, $AIC_{Metc} = 19969$). Finally, we conducted permutation tests on CD trials, averaged across all blocks, to

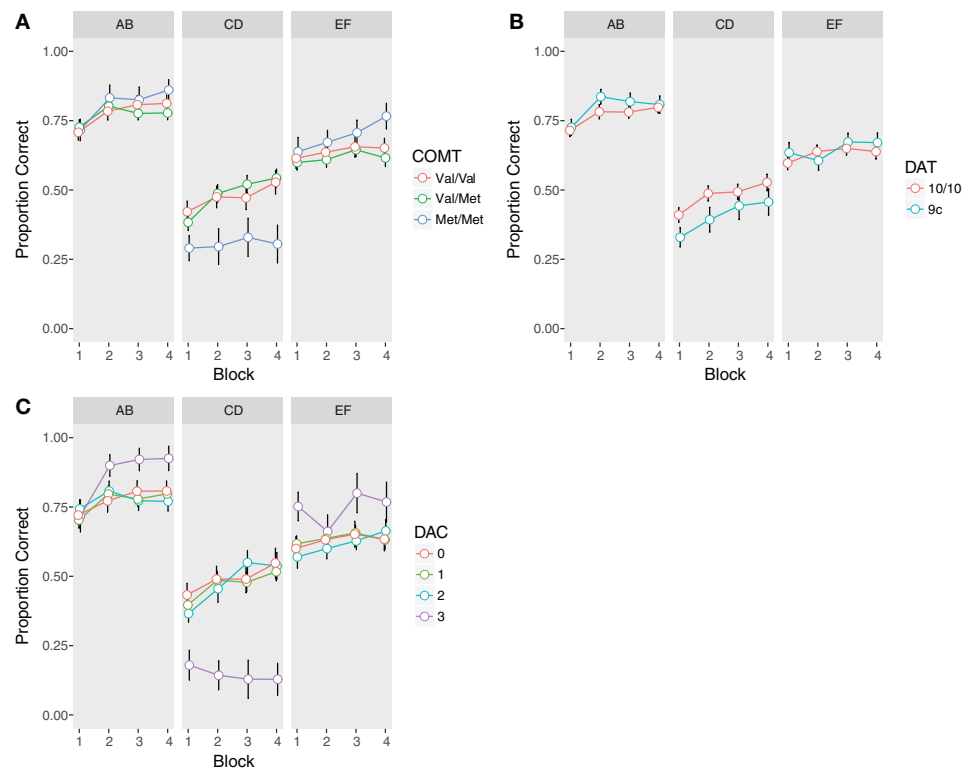


FIGURE 2 | Training phase performance by trial type (AB, CD, EF) and genotype. Accuracy is defined as the proportion of time the symbol with the higher reward probability was chosen, regardless of whether it was rewarded or not. Error bars are standard errors of the mean. **(A)** COMT, **(B)** DAT, **(C)** Dopamine composite (DAC).

further guard against the possibility that our Met homozygote results could have arisen under the null. Confirming our results, Met homozygotes' performance was reliably below the mean on CD trials ($p = 0.004$), and this group performed worse than both Val homozygotes ($p_{corrected} = 0.006$) and heterozygotes ($p_{corrected} = 0.001$). We therefore utilize the full breakdown of COMT genotype for the remainder of the results.

Uninstructed Learning

In contrast to instructed learning, we found no effects of COMT genotype on uninstructed learning (all $ps > 0.2$; **Figure 2A**, **Supplementary Table 8**).

COMT: Test Phase

Instructed test phase performance demonstrated evidence of a gene-dose effect (**Figures 3A,D**). COMT status significantly predicted performance on DF trials ($\chi^2_{(2)} = 9.06$, $p = 0.01$). Val homozygotes were less likely to choose D on DF trials compared to heterozygotes ($\beta = -1.11$, $z = -2.09$, $p_{corrected} = 0.07$) and to Met homozygotes ($\beta = -2.28$, $z = -2.82$, $p_{corrected} = 0.01$). There was no significant difference between Val/Met and Met/Met groups, but Met/Met subjects were numerically more likely to choose D ($\beta = 1.18$, $z = 1.53$, $p_{corrected} = 0.13$). Supporting this pattern, an exploratory gene-dose analysis demonstrated a significant linear effect of the number of Met alleles on choosing D over F ($\beta = 1.60$, $z = 3.01$, $p = 0.003$).

There were no significant effects of COMT genotype on the Avoid-D/Avoid-F measure (all $ps > 0.17$), but quantitatively, differences were indicative of a similar gene-dose relationship on Avoid-D. An exploratory gene-dose analysis demonstrated a trend-level Met x Trial Type interaction ($\beta = -0.44$, $z = -1.86$, $p = 0.06$). While increasing Met alleles negatively predicted performance on Avoid-D ($\beta = -0.90$, $z = -2.09$, $p_{corrected} = 0.07$), there was no relationship with uninstructed Avoid-F ($\beta = -0.02$, $z = -0.05$, $p_{corrected} = 0.96$).

The above results refine, but only partially replicate, the effect of COMT genotype on instructed RL. While Doll et al. (2011) found that Met carriers demonstrate greater instructional bias relative to Val homozygotes during training, we found increased bias exclusively for Met homozygotes. Our COMT test phase results provide novel evidence for a gene-dose effect, though differences on the Avoid-instructed measure were not as robust as reported previously. The prior report included a somewhat greater percentage of Met homozygotes out of all Met carriers (28.3%) than the present study (22.1%), which could have impacted the results given that instructional bias appears to be strongest in the former group. Additionally, a number of methodological differences could have contributed to these discrepancies. These differences aside, as COMT is thought to be particularly and differentially important to the regulation of prefrontal dopamine levels (Durstewitz and Seamans, 2008; Tunbridge, 2010), the present findings further implicate prefrontal cortex

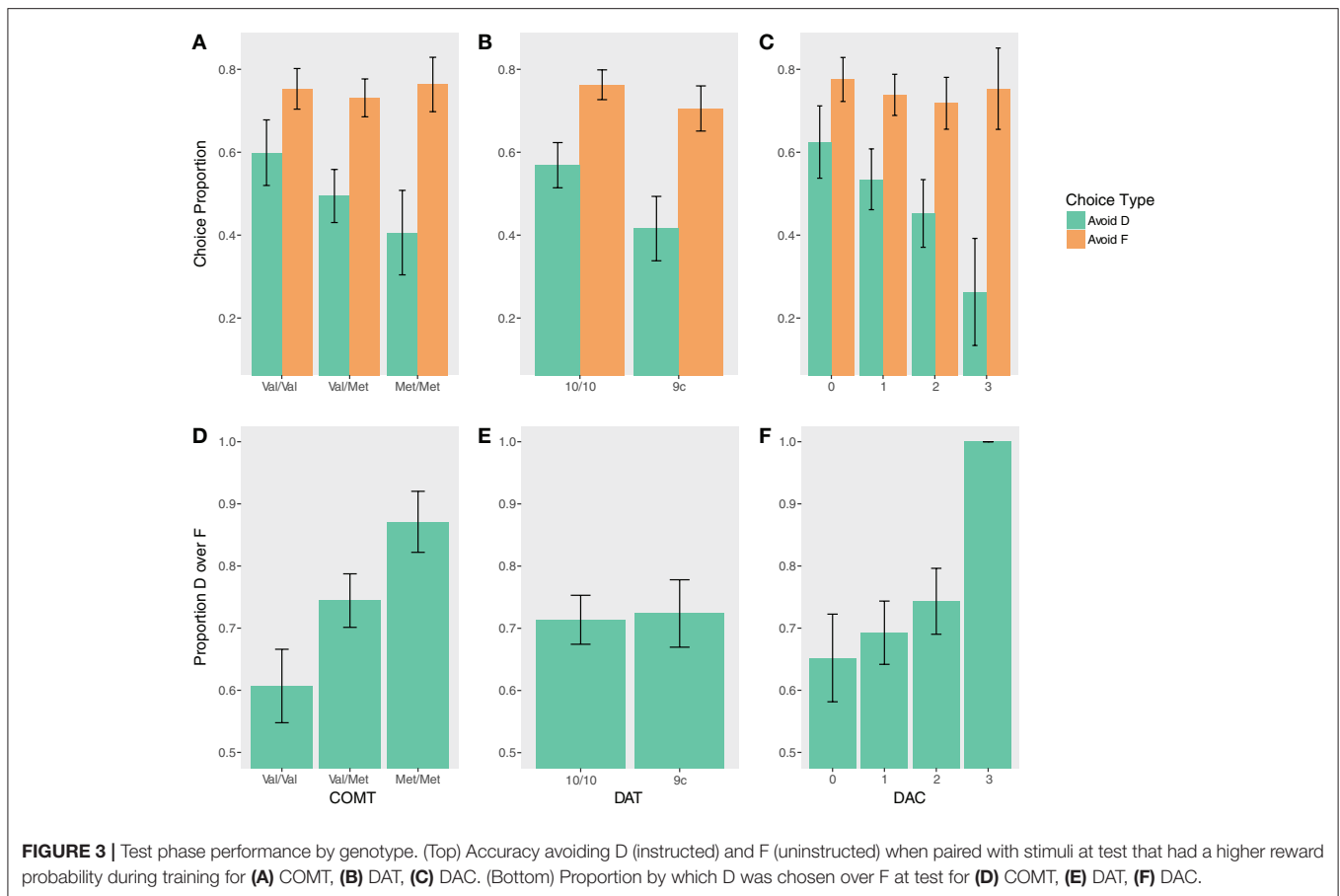


FIGURE 3 | Test phase performance by genotype. (Top) Accuracy avoiding D (instructed) and F (uninstructed) when paired with stimuli at test that had a higher reward probability during training for (A) COMT, (B) DAT, (C) DAC. (Bottom) Proportion by which D was chosen over F at test for (D) COMT, (E) DAT, (F) DAC.

in biasing responding to instructed stimuli at both training and test.

DAT: Training Phase

Expanding the investigation of the effect of dopaminergic genes on instructional bias, we next examined the effect of DAT1 genotype. In our regression models, DAT was simple coded with 10/10 homozygotes as the reference (9c: 0.5, 10/10: -0.5).

As compared to 10-repeat homozygotes, 9-repeat carriers were significantly worse on the instructed pair ($\beta = -0.43$, $z = -2.17$, $p = 0.03$; **Figure 2B, Supplementary Table 9**). There was also a trend-level DAT x Trial Type interaction ($\beta = 0.53$, $z = 1.91$, $p = 0.056$). While 9-repeat carriers were worse on the CD pair, there was no difference between genotypes on the EF pair ($p > 0.5$). There were no interactions between DAT and Block, indicating similar learning trajectories in both groups (all $ps > 0.4$). Nor were there any effects of DAT on uninstructed learning (all $ps > 0.4$; **Figure 2B, Supplementary Table 10**).

DAT: Test Phase

There was no effect of DAT on DF trials ($p = 0.74$; **Figure 3E**). There was a main effect of DAT on Avoid-D/Avoid-F ($\beta = -0.66$, $z = -2.40$, $p = 0.02$; **Figure 3B**) in the absence of a significant interaction ($p > 0.17$), suggesting that 9-repeat carriers were significantly worse overall on these measures. However, the effect

seems to be driven primarily by worse performance on Avoid-D (Avoid-D: $\beta = -0.49$, $z = -2.30$, $p_{corrected} = 0.04$; Avoid-F: $\beta = -0.17$, $z = -1.18$, $p = 0.24$).

Remarkably, though DAT plays little role in cortical DA clearance (Sulzer et al., 2016), it appears to be equally if not more predictive of training than test phase performance, the former of which is putatively more reliant on prefrontal function (Frank et al., 2007). This result is surprising, given that investigations assessing other striatal genes have found that striatal genotypic effects in both instructed and uninstructed learning are confined to the test phase only (Frank et al., 2007; Doll et al., 2011). Previous work has indicated that there is a reciprocal relationship between prefrontal and striatal DA, with more prefrontal DA leading to more cognitive stability, while more striatal DA leads to more cognitive flexibility (Cools and D'Esposito, 2009). Motivated by this and by prior studies in which composites of multiple DA genes have shown better predictive power than single genes (Nikolova et al., 2011; Kohno et al., 2016), we next asked whether a composite DA variable would better predict instructional bias.

DA Composite: Training Phase

To produce the DA composite (DAC), we recoded the COMT and DAT variables according to putative prefrontal-striatal DA balance (COMT: Val/Val = 0, Val/Met = 1, Met/Met = 2; DAT:

10/10 = 0, 9c = 1), and then summed the two variables. The resulting composite ranged between 0 (low frontal DA, high striatal DA) and 3 (high frontal DA, low striatal DA).

Reexamining training phase performance (**Figure 2C**, **Supplementary Table 11**), we found a significant effect of DAC ($\chi^2_{(3)} = 11.02$, $p = 0.01$), superseded by a significant DAC x Trial Type interaction ($\chi^2_{(3)} = 29.56$, $p < 0.0001$). *Post-hoc* comparisons revealed that the DAC3 group was significantly and uniquely impaired in learning away from the instructions compared to the other three groups (DAC3 vs. DAC 0: $\beta = -2.03$, $z = -5.52$, $p_{corrected} < 0.0001$; DAC3 vs. DAC 1: $\beta = -1.96$, $z = -5.56$, $p_{corrected} < 0.0001$; DAC3 vs. DAC 2: $\beta = -1.99$, $z = -5.46$, $p_{corrected} < 0.0001$). In contrast, DAC3 subjects demonstrated better performance on EF, though this did not survive correction for multiple comparisons: (DAC3 vs. DAC 0: $\beta = 0.62$, $z = 1.94$, $p = 0.053$, $p_{corrected} = 0.42$; DAC3 vs. DAC 1: $\beta = 0.58$, $z = 1.92$, $p = 0.056$, $p_{corrected} = 0.42$; DAC3 vs. DAC 2: $\beta = 0.68$, $z = 2.14$, $p = 0.03$, $p_{corrected} = 0.29$). No other comparisons between DAC groups were significant (all $ps > 0.6$). The DAC x Trial Type interaction was already present in the first block of training, suggesting it was not the result of extensive learning ($\chi^2_{(3)} = 15.89$, $p = 0.001$). Nor was it ameliorated by additional training, as the DAC3 group was the only group to show no evidence of learning away on CD from the first block to the last (DAC 0: $\beta = 0.53$, $z = 2.00$, $p_{corrected} = 0.09$; DAC 1: $\beta = 0.57$, $z = 2.83$, $p_{corrected} = 0.01$; DAC 2: $\beta = 0.80$, $z = 3.11$, $p_{corrected} = 0.008$; DAC 3: $\beta = -0.59$, $z = -1.03$, $p_{corrected} = 0.30$).

There were no significant differences between DAC groups in the analysis of uninstructed learning (all $ps > 0.3$), though as with EF, the DAC3 group's performance was quantitatively better on AB (**Figure 2C**, **Supplementary Table 12**). These differences in uninstructed learning are intriguing given that they are in the opposite direction of the instructed effect, but given the small sample size of the DAC3 group ($N = 7$) due to the lower prevalence of both the COMT Met allele (Auton et al., 2015) and the DAT 9-repeat variant (Vandenberg et al., 1992; Doucette-Stamm et al., 1995) in the general population, this study may not have had the statistical power to determine whether such small effects are reliable.

As with the COMT Met/Met results, because of the small sample size of the DAC3 group, we again took efforts to ensure these results did not arise by chance. First, we repeated the analysis with a modified DA composite created by summing the Metc and DAT variables (Metc: Val/Val = 0, Met carrier = 1; DAT: 10/10 = 0, 9c = 1), producing three DACmetc groups $Ns = 25:51:26$ (0:1:2). Repeating our analysis of instructed learning, we failed to find any effects of DACmetc (all $ps > 0.21$). However, the full DAC model provided a much better fit to the data, despite including additional parameters ($AIC_{DAC} = 19958$, $AIC_{DACmetc} = 19974$), and also provided a better fit than both the COMT and DAT instructed learning models ($AIC_{COMT} = 19966$, $AIC_{DAT} = 19961$). Permutation tests on the average performance on CD trials across training also support the results of the regression analysis. DAC3 subjects were reliably below the mean on CD trials ($p < 0.0001$), and this group performed worse than all other DAC groups (DAC3 vs. DAC0: $p_{corrected} = 0.0001$, DAC3 vs. DAC1: $p_{corrected} < 0.0001$, DAC3 vs. DAC2: $p_{corrected} = 0.0001$).

Given that it is highly unlikely that seven randomly chosen subjects would have performance at the level of the DAC3 group, we utilize the full DAC composite for the remainder of the results.

DA Composite: Test Phase

While there was only a marginal main effect of DAC on Avoid-D/Avoid-F ($\chi^2_{(3)} = 6.62$, $p = 0.085$), a gene-dose analysis revealed a significant linear effect of DAC ($\beta = -0.86$, $z = -2.56$, $p = 0.01$) qualified by a DAC x Choice Type interaction ($\beta = -0.62$, $z = -2.16$, $p = 0.03$). DAC status was negatively associated with avoiding D; it showed no relationship to avoiding F (Avoid-D: $\beta = -1.47$, $z = -2.88$, $p_{corrected} = 0.008$; Avoid-F: $\beta = -0.24$, $z = -0.68$, $p_{corrected} = 0.50$; **Figure 3C**). DF trials revealed a similar pattern; though there was no main effect of DAC ($\chi^2_{(3)} = 1.12$, $p = 0.77$), there was a significant gene-dose effect, with increasing choice of the instructed stimulus with increasing DAC status ($\beta = 1.72$, $z = 2.53$, $p = 0.01$). This effect appears to be driven primarily by the DAC3 group, all seven of whom remarkably chose D over F 100% of the time (**Figure 3F**).

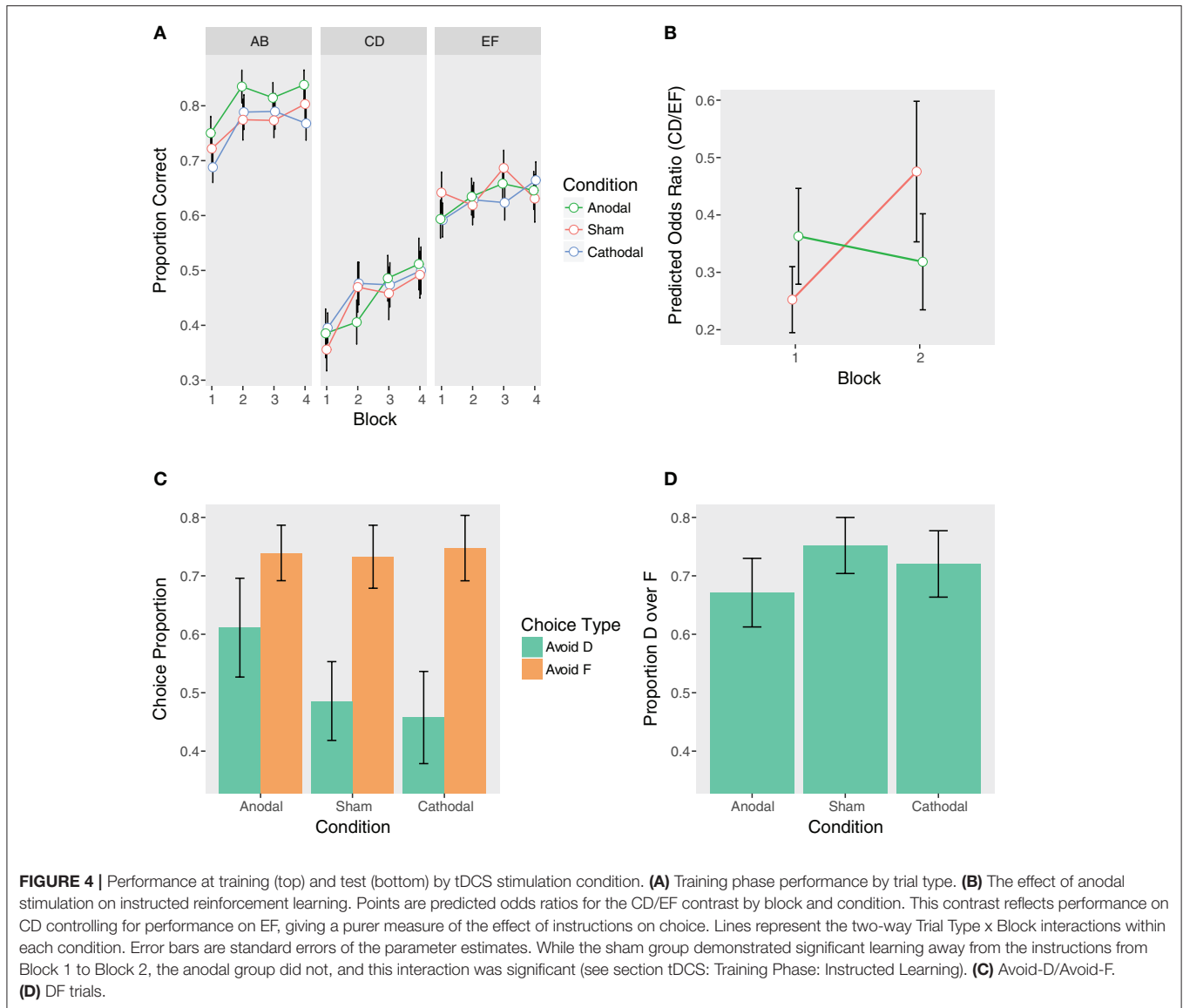
In sum, there was graded effect of DAC on test phase performance, with increasing frontal (decreasing striatal) DA predicting greater adherence to the instructions. This graded relationship was punctuated by the performance of the DAC3 group, who, as during training, demonstrated substantially greater instructional bias.

Taken together, the genotyping results implicate prefrontal cortex, and in particular the balance between prefrontal and striatal dopamine, in modulating instructed RL. This pattern motivates asking our next question: Does experimentally manipulating prefrontal function via tDCS alter the magnitude of instructional bias?

tDCS: Training Phase Instructed Learning

To examine the main hypotheses of the study—that anodal stimulation will increase confirmation bias, while cathodal stimulation may decrease it—our focal analyses concerned the contrasts of Anodal vs. Sham stimulation and Cathodal vs. Sham stimulation. These contrasts include Condition, or the overall effect of stimulation compared to Sham on instructed choice, and Condition x Trial Type, which allows for the same assessment while controlling for performance on EF. For a more fine-grained investigation of the time course of learning, we additionally examined the Condition x Block interactions, which indicate whether stimulation altered the extent to which subjects learned away from the instructions across training blocks, and the Condition x Trial Type x Block interactions, which allow for the same assessment while controlling for performance on EF. Condition was simple coded with sham as the reference (Anodal: 2/3 -1/3, Cathodal: -1/3 2/3, Sham: -1/3 -1/3).

We first examined the contrasts between anodal and sham stimulation. Supporting our hypothesis, there was a significant Anodal vs. Sham x Trial Type x Block 2 vs. 1 interaction ($\beta = 0.76$, $z = 2.22$, $p = 0.03$). When controlling for performance on EF, the sham group demonstrated significant learning away from the instructions from Block 1 to Block 2 on CD, while the anodal group did not (Sham: $\beta = 0.63$, $z = 2.60$, $p_{corrected} =$



0.046; Anodal: $\beta = -0.13$, $z = -0.53$, $p_{corrected} = 1.00$). The sham group nearly doubled their performance (OR = 1.88), but the anodal group demonstrated essentially no learning (OR = 0.88; **Figures 4A,B** and **Table 2**). Examining performance on CD without adjusting for EF, the Anodal vs. Sham x Block 2 vs. 1 interaction was at trend ($\beta = -0.45$, $z = -1.59$, $p = 0.11$). As above, the sham group showed significant learning from Block 1 to Block 2, while the anodal group did not (Sham: $\beta = 0.53$, $z = 2.68$, $p_{corrected} = 0.04$; Anodal: $\beta = 0.08$, $z = 0.42$, $p_{corrected} = 1.00$). In contrast, neither group demonstrated significant learning from Block 1 to Block 2 on EF (Sham: $\beta = -0.10$, $z = -0.64$, $p_{corrected} = 1.00$; Anodal: $\beta = 0.21$, $z = 1.33$, $p_{corrected} = 0.74$).

We also sought to ensure that the effect of anodal stimulation early in learning was not driven by the presence of DAC3 subjects. Controlling for DAC, the Anodal vs. Sham x Trial Type x Block 2 vs. 1 interaction remained significant ($\beta = 0.86$, $z = 2.45$, $p =$

0.01) and the Anodal vs. Sham x Block 2 vs. 1 interaction for CD remained at trend ($\beta = -0.49$, $z = -1.74$, $p = 0.08$), confirming that the effect was not driven by genotypic differences between stimulation conditions.

Taken together, these results indicate that anodal stimulation significantly impeded learning away from the instructions during the initial blocks. No other Anodal vs. Sham contrasts were significant (**Table 2**), including the overall effect of anodal stimulation ($p = 0.97$) and the Anodal vs. Sham x Trial Type interaction ($p = 0.88$), suggesting that anodal stimulation only weakly and transiently affected performance. In contrast to the anodal condition, there were no significant effects of cathodal stimulation (all $ps > 0.14$).

Uninstructed Learning

We also explored the effect of stimulation on accuracy during training for the uninstructed symbol pairs (AB, EF). Quantifying

TABLE 2 | Mixed effects logistic regression model of the effect of instruction (CD vs. EF) and tDCS on training phase performance.

Predictor	β	OR ^a	z	p
Intercept	-0.27	0.76	-2.86	0.004
Anodal vs. Sham	-0.01	0.99	-0.04	0.97
Cathodal vs. Sham	0.09	1.10	0.40	0.69
Trial Type	0.93	2.53	6.96	<0.0001
Block 2 vs. 1	0.32	1.38	2.81	0.005
Block 3 vs. (1, 2)	0.24	1.27	2.35	0.02
Block 4 vs. (1, 2, 3)	0.31	1.36	3.13	0.002
Anodal vs. Sham x Trial Type	-0.05	0.95	-0.15	0.88
Cathodal vs. Sham x Trial Type	-0.17	0.84	-0.54	0.59
Anodal vs. Sham x Block 2 vs. 1	-0.45	0.64	-1.59	0.11
Anodal vs. Sham x Block 3 vs. (1, 2)	0.18	1.20	0.73	0.47
Anodal vs. Sham x Block 4 vs. (1, 2, 3)	0.08	1.07	0.31	0.75
Cathodal vs. Sham x Block 2 vs. 1	-0.19	0.83	-0.69	0.49
Cathodal vs. Sham x Block 3 vs. (1, 2)	-0.10	0.91	-0.41	0.68
Cathodal vs. Sham x Block 4 vs. (1, 2, 3)	-0.17	0.85	-0.70	0.48
Trial Type x Block 2 vs. 1	-0.22	0.80	-1.58	0.11
Trial Type x Block 3 vs. (1, 2)	-0.03	0.97	-0.22	0.82
Trial Type x Block 4 vs. (1, 2, 3)	-0.16	0.85	-1.11	0.27
Anodal vs. Sham x Trial Type x Block 2 vs. 1	0.76	2.15	2.22	0.03
Anodal vs. Sham x Trial Type x Block 3 vs. (1, 2)	-0.25	0.78	-0.75	0.46
Anodal vs. Sham x Trial Type x Block 4 vs. (1, 2, 3)	0.06	1.06	0.17	0.86
Cathodal vs. Sham x Trial Type x Block 2 vs. 1	0.47	1.60	1.41	0.16
Cathodal vs. Sham x Trial Type x Block 3 vs. (1, 2)	-0.12	0.89	-0.37	0.71
Cathodal vs. Sham x Trial Type x Block 4 vs. (1, 2, 3)	0.50	1.65	1.46	0.14

Boldfaced text indicates $p < 0.05$. ^aOR: Odds Ratio.

the effect of stimulation on uninstructed learning is important in order to show that effects on instruction are not in some way due to generally altered learning, especially given a prior report of altered performance on the AB pair under anodal stimulation (Turi et al., 2015).

Though there were no significant effects of stimulation condition at the $p < 0.05$ level, there was a trend-level Anodal vs. Sham x Trial Type interaction ($\beta = 0.16$, $z = 1.66$, $p = 0.097$; **Figure 4A, Supplementary Table 13**), reflecting somewhat better average performance on the AB pair by the anodal group. This difference is intriguing given increasing evidence that working memory processes contribute to RL performance (Collins and Frank, 2012; Collins et al., 2017), and anodal stimulation has been shown to improve working memory (Fregni et al., 2005; Zaehle et al., 2011; Nozari and Thompson-Schill, 2013). However, in light of the marginal nature of this unhypothesized effect, we do not interpret it further. As with instructed learning, there were no significant effects of cathodal stimulation (all $ps > 0.12$).

tDCS: Test Phase

In contrast to the training phase, there were no significant effects of stimulation on either Avoid-D/Avoid-F or DF trials at test (all $ps > 0.19$; **Figures 4C,D**). This suggests that unlike COMT genotype, to the extent that tDCS modulated instructed learning, it biased choice during training without impacting the learned value of the instructed stimulus.

Computational Modeling

While the behavioral analyses above confirm the existence of instructional bias, they are only weakly informative with respect to the underlying mechanisms. Two classes of models have been suggested to account for instructional bias on the PST: models in which instructions bias striatal reward learning (learning bias models), and those in which instructions affect choice rather than learning (choice bias models; Doll et al., 2009). Prior work has provided weak evidence for a choice bias operating during training, while test phase performance has been best explained by a learning bias mechanism (Doll et al., 2009, 2011). Two results from the present study bear on this question. First, the early-developing, persistent bias of the DAC3 group during training, coupled with their exclusive choice of D over F at test, would seem to be more consistent with a choice bias during both phases. However, these effects could also plausibly arise from a very strong learning bias, making this interpretation far from definitive. Second, the unaltered performance by the anodal group at test also appears more consistent with tDCS influencing a choice bias early in training, though caution is warranted in interpreting a null result.

We therefore fit computational models to subjects' data—one learning bias model and two choice bias models—each of which encapsulates a different hypothesis about the nature of instructional control (see section Methods: Computational Modeling). Briefly, the *learning bias model* (Doll et al., 2009) assumes instructional bias arises from an increase in learning rate for gains and a decrease in learning rate for losses when the instructed symbol D is selected. The *Bayesian hypothesis testing model* (Doll et al., 2009) assumes that subjects veridically learn the reward value of D in a Bayesian fashion, but must have a certain level of confidence that the value of D is below chance before they reliably stop choosing it. We additionally implemented a novel choice bias model, the *decision bias model*, which assumes a standard RL learner with a fixed bias added to the value of D during choice. Finally, we fit a standard RL model, which tests the null hypothesis of no bias.

Contrary to prior work, both the training and test phase were best explained by the decision bias model (**Table 3**). However, while AIC strongly supported this model at both training and test, the protected exceedance probabilities and estimated model frequencies did not provide strong evidence that this model was more frequent in the population for the training phase than the Bayesian hypothesis testing model. We therefore examined the correlation between each model's bias parameter and performance on CD trials across training, in order to ascertain whether one or the other model better accounted for behavior on instructed learning trials. The ϕ parameter of the Bayesian hypothesis testing model was significantly correlated

TABLE 3 | Model comparison of reinforcement learning model fits to subject data.

Model	FP	–LL	AIC	PEP	EF
TRAINING					
Standard	3	13543.8	27705.7	0.007	0.239
Learning bias	5	13298.2	27626.4	5.081e-05	0.003
Bayes HT	4	13402.9	27629.8	0.339	0.362
Decision bias	4	13284.2	27392.4	0.654	0.397
TEST					
Standard	3	4755.4	10128.8	3.816e-16	0.164
Learning bias	5	4362.8	9755.7	3.816e-16	0.002
Bayes HT	4	4748.4	10320.8	3.816e-16	0.015
Decision bias	4	4304.6	9433.3	1.000	0.818

FP: number of free parameters; –LL: negative log-likelihood; AIC: Akaike information criteria; PEP: protected exceedance probability, the probability that a given model is the most frequent in the population, above and beyond chance; EF: estimated model frequency, the frequency of the model in the population as estimated by the Bayesian random-effects analysis.

with performance on CD trials ($r_{(101)} = -0.23$, $p = 0.02$). However, the correlation between the ρ parameter of the decision bias model and CD performance was much stronger ($r_{(101)} = -0.66$, $p < 0.0001$), and the difference between the correlations was significant (Steiger's $Z = -3.82$, $p = 0.0001$). In accordance with our tentative hypothesis based on the behavioral results, we conclude that both training and test phase performance can be parsimoniously accounted for by a single choice bias mechanism.

We also reexamined genotypic and stimulation group differences with respect to the ρ parameter of the decision bias model. These results are reported fully in **section 2** of the Supplementary Material and average parameter estimates are reported in **Supplementary Tables 1 and 2**. Briefly, we found effects of COMT and DAC on ρ at both training and test, in the same direction as the behavioral results. For DAT, 9-repeat carriers were fit with a higher ρ parameter during training, but test phase differences were best explained by the 9-repeat carrier group being fit with a lower learning rate for losses as compared to 10/10 group. We were, however, unable to confirm the anodal tDCS behavioral effect in the parameters of the decision bias model. While this does not invalidate the effect, it does warrant additional caution in interpreting the result.

DISCUSSION

There is mounting evidence that reward learning is far more complex and dynamic than can be accounted for by simple model-free theories of reinforcement. This complexity has been explored with respect to goal-directed planning processes (i.e., model-based RL; Dolan and Dayan, 2013) and instructional control (Wolfensteller and Ruge, 2012), among others. Both model-based RL and instructional control have been associated with cognitive control and frontostriatal function (Daw et al., 2005; Doll et al., 2009, 2011, 2014; Li et al., 2011; Wolfensteller and Ruge, 2012; Fouragnan et al., 2013; Smittenaar et al., 2013; Otto et al., 2015). While the importance of cognitive control to healthy cognitive functioning is indisputable, top-down

control can be detrimental to learning and cognitive flexibility (Chrysikou et al., 2014; Gopnik et al., 2015).

In the case of instructed reinforcement learning, increased top-down control can be detrimental in that it leads to greater instructional bias toward inaccurate instructions. This study expands on the finding that instructional bias is associated with dopaminergic genes affecting PFC and striatal function (Doll et al., 2011), suggesting that the balance between PFC DA (COMT) and striatal DA (DAT1) modulates instructed learning. We further establish a causal link between PFC and biases found in instructed RL. In accord with our hypothesis, anodal subjects demonstrated more protracted learning away from the instructions during the early blocks of training, complementing the genetic evidence that individual differences associated with PFC function are linked to individual differences in instructional control of RL.

A Dopamine Genetic Composite Is Associated With Instructed Learning

While both COMT Met/Met genotype and DAT1 9-repeat carrier genotype were individually significant predictors of greater instructional control during training, the DA composite revealed that this effect was selective to Met/Met:9-repeat carriers (DAC3). This greater bias emerged early in training and persisted throughout the training phase, unaffected by feedback. During test, a gene-dose effect, confirmed both within each gene and with the composite, demonstrated greater bias with increasing Met alleles and decreasing DAT1 repeats. These results are consistent with the known reciprocal relationship between PFC and striatal DA (Kolachana et al., 1995; King et al., 1997; Meyer-Lindenberg et al., 2005). It has been hypothesized that the balance between cognitive stability and cognitive flexibility is mediated via corticostriatal interactions and the differential modulation of prefrontal and striatal circuits by DA. While increases in prefrontal relative to striatal DA have been linked to cognitive stability, increases in striatal relative to prefrontal DA have been linked to cognitive flexibility (Cools and D'Esposito, 2009). We propose that increasing PFC DA, indexed by increasing Met alleles, coupled with decreasing tonic striatal DA, indexed by decreasing DAT1 repeats, shifts the balance away from bottom-up striatal learning based on reward prediction errors and toward PFC-mediated top-down control of RL.

While extracellular DA is primarily recycled via reuptake by DAT in striatal regions, there is little DAT expression in PFC, where levels of DA are controlled by reuptake via the norepinephrine transporter (NET) and enzymatic breakdown via COMT (Seamans and Yang, 2004; Sulzer et al., 2016). With regard to COMT, PFC DA plays a critical role in stabilizing working memory representations (Durstewitz and Seamans, 2008), which are thought to facilitate top-down control (Miller and Cohen, 2001). Notably, carriers of the Met allele of the Val158Met genetic polymorphism have diminished COMT enzyme activity and concomitantly higher levels of prefrontal dopamine (see Tunbridge, 2010 for review). Elevated DA in PFC may then cause increased D1 receptor stimulation, which further drives activity in PFC afferents such as the striatum (Bilder et al.,

2004). Indeed, frontostriatal functional connectivity varies with COMT genotype (Tan et al., 2007; Krugel et al., 2009; Tunbridge et al., 2013). Behaviorally, the Met allele has been associated with enhanced working memory and cognitive control (see Witte and Flöel, 2012 for review). Carriers of the Val allele have more rapid breakdown of prefrontal dopamine and thus somewhat weaker working memory, but potentially greater cognitive flexibility (Krugel et al., 2009; Witte and Flöel, 2012). Replicating previous findings (Doll et al., 2011), the Met allele in our study was associated with greater instructional bias and therefore indicative of greater top-down control.

In the case of the DAT1/SLC6A3 VNTR, our behavioral results are consistent with increased DAT expression with the 9-repeat allele (Faraone et al., 2014) leading to reductions in tonic DA concentrations in the striatum. Reduced tonic DA in striatum has been shown to facilitate PFC input (Goto and Grace, 2005), which would in turn allow for greater biasing of RL. Furthermore, human imaging studies have demonstrated that DAT1 and COMT affect activity in prefrontal and striatal regions during reward anticipation. While the results of these studies are not entirely consistent, anticipatory activity in striatum is generally greater for DAT1 9-repeat carriers and is modulated by COMT genotype (Dreher et al., 2009; Aarts et al., 2010; cf. Yacubian et al., 2007), with one study finding the highest activity in both lateral PFC and ventral striatum for Met/Met:9-repeat carriers (Dreher et al., 2009).

However, this interpretation must be qualified by the considerable uncertainty surrounding the effect of the DAT1/SLC6A3 VNTR on dopaminergic function. Both *in vivo* and *in vitro* studies have produced conflicting results, with some supporting greater DAT expression for the 9-repeat allele compared to the 10-repeat allele, while others report the opposite, or no relationship (Costa et al., 2011; Faraone et al., 2014). A recent meta-analysis of human imaging studies supports the first possibility when restricting the analysis to normal controls (Faraone et al., 2014). Disease status, development, and ancestry may all play a role in the functional consequences of DAT1 (Franke et al., 2010; Shumay et al., 2011; Faraone et al., 2014). Even in the absence of changes in overall DAT expression, heterogeneities in DAT density and variations in neuronal morphology can substantially affect dopamine reuptake, which could contribute to the diversity of findings (Kaya et al., 2018).

It is also unclear the extent to which variation in DAT expression should be expected to influence tonic vs. phasic DA. Phasic DA bursts are associated with salient stimuli and have been shown to be associated with learning via reward prediction errors (Schultz et al., 1997; Berridge, 2012). Various roles have been ascribed to tonic DA, including modulation of response vigor (Niv et al., 2007), exploration (Beeler et al., 2010), and the relative weighting of effort costs (Salamone et al., 2007). DAT has a clear role in maintaining tonic DA concentrations (Efimova et al., 2016; Sulzer et al., 2016). Accordingly, DAT has been attributed a major role in synaptic DA clearance after phasic release (Bilder et al., 2004), and pharmacological blockade of DAT alters DA transients and leads to long lasting increases in tonic

DA (Floresco et al., 2003; Ford et al., 2010). However, detailed biophysical modeling suggests that diffusion is responsible for synaptic clearance of DA, with DAT having a (potentially limited) role in shaping the radius and duration at which DA bursts could activate receptors via volume transmission (Cragg and Rice, 2004; Arbuthnott and Wickens, 2007; Rice and Cragg, 2008). Notably, increasing burst firing of DA neurons in the ventral tegmental area does not cause tonic increases in extracellular DA in the nucleus accumbens without DAT blockade (Floresco et al., 2003). Tonic DA may also indirectly influence phasic activity, though the direction of this influence is complicated to determine; elevated tonic DA due to increased tonic DA neuron firing may augment the peak and duration of DA bursts (Dreyer et al., 2010), but tonic concentrations may also inhibit phasic DA via autoreceptor feedback mechanisms (Bilder et al., 2004).

The performance of patients with schizophrenia provides an interesting counterpoint to the combined effect of COMT and DAT. Opposite to the Met/Met:9-repeat carrier genotype, the pathology of schizophrenia includes hyperdopaminergic tone in striatum and hypodopaminergic tone in PFC (Weinberger et al., 1992; da Silva Alves et al., 2008; Brisch et al., 2014; Slifstein et al., 2015; Grace and Gomes, 2018). Notably, patients with schizophrenia demonstrate reduced instructional bias on the PST (Doll et al., 2014). They also seem to rely less on putatively PFC-mediated processes in uninstructed learning, including reduced use of win-stay, lose-shift strategies and poorer performance on the easiest AB pair, potentially indicative of a reduced tendency to maximize or otherwise use rule-based strategies (Waltz et al., 2007, 2011; Doll et al., 2014). Though the elevated performance on AB in the Met/Met:9-repeat carrier group in the present study was not significant, it provides further evidence of opposite behavioral effects of opposite dopaminergic profiles.

Our findings of reduced flexibility with increasing ratio of PFC to striatal DA are also in accord with the effects of COMT and DAT1 on reversal learning. Compared to Met homozygotes, Val homozygotes show greater learning-rate adaptation around reversals, leading to improved performance (Krugel et al., 2009). Notably, Val homozygotes have more differentiated prediction error signals in striatal regions and greater learning-rate-dependent modulation of frontostriatal connectivity, suggestive of more adaptive prefrontal modulation of striatal RL (Krugel et al., 2009). On the other hand, the DAT1 9-repeat allele is associated with greater perseveration after a reversal (den Ouden et al., 2013). It is interesting to note that this perseveration effect was explained by the 9-repeat allele conferring a more rapidly decreasing learning rate with increasing experience, which may be related to the decreased learning rate modulation of COMT Met homozygotes. Direct comparison is difficult, however, as different computational models were used in the two studies. Importantly, while den Ouden and colleagues attributed their findings to more robust striatal learning of the previous reward contingencies, in the case of Met/Met:9-repeat carriers in the present study, their performance in the training phase cannot be due to greater ingraining of previous experience; the bias in the present case was due to instruction, not experience, was robustly evident in the first training block, and persisted throughout training.

Stimulation Weakly Increased Instructional Bias

In contrast to the genetic effects, the effect of tDCS on performance was far more limited. In accord with our hypothesis, anodal subjects demonstrated modestly more protracted learning away from the instructions during the early blocks of training. However, there was no effect of cathodal stimulation, and no effect of either stimulation condition during the test phase.

While the isolation of the effect to the training phase makes sense in light of the postulated division between frontal and striatal systems during training and test (Frank et al., 2007), it is at odds with the finding of increased bias at test associated with the COMT Met allele. It may be the case that genotypic effects on frontostriatal DA balance or frontostriatal connectivity (discussed above) allow for greater biasing of striatum by PFC than is possible with single-session tDCS.

Mechanisms of Instructional Bias

The mechanisms underlying instructional bias are under debate. Proposals include models in which instructions bias striatal reward learning (learning bias models; Biele et al., 2009; Doll et al., 2009) or those in which instructions affect choice rather than learning (choice bias models; Doll et al., 2009). Evidence in favor of each of these classes of models has been mixed. Past computational modeling has tended to support learning bias models (Biele et al., 2009, 2011; Doll et al., 2009, 2011) but does not unequivocally rule out choice bias models (Doll et al., 2009, 2011). A number of neuroimaging studies have favored neither class of models, finding blunted activation in basal ganglia structures during instructed/prior knowledge conditions, suggesting a suppression of RL (Delgado et al., 2005; Biele et al., 2011; Li et al., 2011; Fouragnan et al., 2013). However, one study found overall decreased activity in reward structures but activity consistent with a learning bias in the form of an “outcome bonus” for choosing the instructed stimulus (Biele et al., 2011).

Adding to this debate, we find that our training phase results can be explained by a novel choice bias model—the decision bias model—containing a fixed bias for choosing the instructed symbol. This is in contrast to past work, which has found that a standard RL model without instructional bias best fits training phase performance, despite clear behavioral effects of instruction during training (Doll et al., 2009, 2011). Our model also better predicted behavioral performance on CD trials compared to the Bayesian hypothesis testing model, a choice bias model previously shown to provide a reasonable fit to some subjects’ training data and to be sensitive to effects of COMT (Doll et al., 2009, 2011). These results thus provide stronger evidence for the existence of a choice bias mechanism during training.

The decision bias and Bayesian hypothesis testing models differ in a number of regards (see section Methods: Computational Modeling), with the most prominent differences being in the type of learner (standard Q-learning vs. Bayesian Q-learning) and in the nature of the bias (fixed vs. variable). We cannot say with certainty which of these factors most contributes to the superior performance of the decision bias model, though comparing our pattern of results to past work suggests that the Bayesian learner detracted from the performance of the model; all else equal, a variable bias should presumably better capture

the behavior of a putative fixed bias agent than no bias. That said, an important direction for future work is to introduce a variable bias into the standard Q-learning framework and compare this to a fixed bias. This poses a challenge, since the uncertainty information used to implement adaptivity in the Bayesian framework is not present in the standard framework.

Again contrary to prior results, the decision bias model also best explained performance at test. While model comparison and striatal dopaminergic genetic effects have been previously taken as evidence of a learning bias mechanism at test (Doll et al., 2009, 2011), the supposition that the test phase primarily measures learning free of choice effects has recently come into question (Shiner et al., 2012; Smittenaar et al., 2012), in keeping with a broader role of DA in modulating motivation and learned value representations (Cagniard et al., 2006; Berridge, 2012; Medic et al., 2014). Further supporting our finding, a recent reevaluation of test phase performance using an alternative model redescribed the learning bias for one striatal genotype as a choice bias (Collins and Frank, 2014). These discrepancies highlight the fact that model comparison results are dependent on the models tested. Additionally, in light of the evidence from other studies, there is no reason to think choice bias and learning bias mechanisms are mutually exclusive. However, the complexity of a model implementing both forms of bias would likely pose identifiability issues. We suggest that along with continued refinements to computational models, novel experimental designs capable of teasing apart these different possibilities will be necessary to advance our understanding of the mechanisms of instructional control.

Specificity of the Effects and Limitations

While there is good evidence that the expression of COMT and DAT1 are regionally specific, caution must be taken in interpreting the results of stimulation, as the lack of focality of tDCS prevents strong claims about effects on specific brain regions. Stimulation could have altered the function of other brain areas involved in RL, including orbitofrontal cortex (O’Doherty, 2004). Neuroimaging and current modeling have even shown tDCS effects in subcortical structures, including the basal ganglia (Sadleir et al., 2010; Weber et al., 2014). However, the lack of stimulation effects on uninstructed learning and test phase performance somewhat militates against these possibilities.

Importantly, while our sample size was large for a tDCS study (Minarik et al., 2016) and was larger than the original report of the effects of COMT on instructed RL (Doll et al., 2011), these results should be replicated, particularly in light of the weakness of the tDCS effects and the small sample size of some genotypes. In the latter case, the low frequencies of the COMT Met and DAT1 9-repeat alleles in the population make collecting adequate samples of these groups challenging (Vandenbergh et al., 1992; Doucette-Stamm et al., 1995; Auton et al., 2015). Because access to such samples is difficult outside of large cohort studies, we took statistical steps within our sample to ensure the robustness of our genetic results. Given the known interaction of COMT and task on the effects of prefrontal stimulation (Plewania et al., 2013; Nieratschker et al., 2015), larger samples would also permit an examination of genotype x stimulation interactions. Though a between-subjects design was necessary

in this study due to the use of deception, future examinations of this topic could also be improved by the development of within-subjects designs. Finally, it is conceivable that there is more opportunity to decrease bias than increase it, given the overwhelming feedback subjects receive in contradiction to the instructions. Unfortunately, cathodal tDCS, which could in principle be used to test this hypothesis, failed to elicit an effect in the present case and is demonstrably unreliable (Jacobson et al., 2012; Nozari et al., 2014). Future studies using theta-burst transcranial magnetic stimulation may be an appropriate alternative.

Conclusion

In sum, the present study provides further evidence for the role of PFC in biasing instructed RL, and additionally highlights the importance of frontostriatal DA balance in modulating top-down inputs. Such top-down regulation of learning by PFC is consistent with increased cognitive control leading to both costs and benefits (Chrysikou et al., 2014). Understanding the interplay of cognitive control and learning is thus key to establishing what level of control is most adaptive in a given situation. This endeavor will ultimately require delineating the relationship between computational and neurocognitive factors in learning and choice.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this manuscript are available from the corresponding author upon reasonable request.

REFERENCES

- Aarts, E., Roelofs, A., Franke, B., Rijpkema, M., Fernández, G., Helmich, R. C., et al. (2010). Striatal dopamine mediates the interface between motivational and cognitive control in humans: evidence from genetic imaging. *Neuropsychopharmacology* 35, 1943–1951. doi: 10.1038/npp.2010.68
- Arbuthnott, G. W., and Wickens, J. (2007). Space, time and dopamine. *Trends Neurosci.* 30, 62–69. doi: 10.1016/j.tins.2006.12.003
- Auton, A., Abecasis, G. R., Altshuler, D. M., Durbin, R. M., Bentley, D. R., Chakravarti, A., et al. (2015). A global reference for human genetic variation. *Nature* 526, 68–74. doi: 10.1038/nature15393
- Barr, D. J., Levy, R., Scheepers, C., and Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *J. Mem. Lang.* 68, 255–278. doi: 10.1016/j.jml.2012.11.001
- Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67, 1–48. doi: 10.18637/jss.v067.i01
- Beeler, J. A., Daw, N., Frazier, C. R. M., and Zhuang, X. (2010). Tonic dopamine modulates exploitation of reward learning. *Front. Behav. Neurosci.* 4:170. doi: 10.3389/fnbeh.2010.00170
- Berridge, K. C. (2012). From prediction error to incentive salience: mesolimbic computation of reward motivation. *Eur. J. Neurosci.* 35, 1124–1143. doi: 10.1111/j.1460-9568.2012.07990.x
- Biele, G., Rieskamp, J., and Gonzalez, R. (2009). Computational models for the combination of advice and individual learning. *Cogn. Sci.* 33, 206–242. doi: 10.1111/j.1551-6709.2009.01010.x
- Biele, G., Rieskamp, J., Krugel, L. K., and Heekeren, H. R. (2011). The neural basis of following advice. *PLoS Biol.* 9:e1001089. doi: 10.1371/journal.pbio.1001089
- Bilder, R. M., Volavka, J., Lachman, H. M., and Grace, A. A. (2004). The catechol-O-methyltransferase polymorphism: relations to the tonic-phasic dopamine

AUTHOR CONTRIBUTIONS

NT and ST-S formulated the experiment. NT implemented the experiment and performed data analysis. NT and KG collected the data. NT wrote the first draft of the manuscript, and all authors contributed to manuscript revision.

FUNDING

This work was supported by a National Institutes of Health Grant (R01DC009209) to ST-S and by a Department of Defense/Army Research Office National Defense Science and Engineering Graduate (NDSEG) Fellowship (32 CFR 168a) to NT.

ACKNOWLEDGMENTS

Thanks to Roy Hamilton, Olu Faseyitan, and the LCNS lab for assistance with tDCS current modeling and stimulation parameters and to Hetty Rodriguez, Jennifer-Marie Rosado, and the Penn Molecular Profiling Facility for genotyping services. We also thank Brianna Owairu for help with data collection.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnhum.2018.00472/full#supplementary-material>

- hypothesis and neuropsychiatric phenotypes. *Neuropsychopharmacology* 29, 1943–1961. doi: 10.1038/sj.npp.1300542
- Brisch, R., Saniotis, A., Wolf, R., Bielau, H., Bernstein, H.-G., Steiner, J., et al. (2014). The role of dopamine in schizophrenia from a neurobiological and evolutionary perspective: old fashioned, but still in vogue. *Front. Psychiatry* 5:47. doi: 10.3389/fpsyt.2014.00047
- Cagniard, B., Beeler, J. A., Britt, J. P., McGehee, D. S., Marinelli, M., and Zhuang, X. (2006). Dopamine scales performance in the absence of new learning. *Neuron* 51, 541–547. doi: 10.1016/j.neuron.2006.07.026
- Cattaneo, Z., Pisoni, A., and Papagno, C. (2011). Transcranial direct current stimulation over Broca's region improves phonemic and semantic fluency in healthy individuals. *Neuroscience* 183, 64–70. doi: 10.1016/j.neuroscience.2011.03.058
- Chrysikou, E. G., Hamilton, R. H., Coslett, H. B., Datta, A., Bikson, M., and Thompson-Schill, S. L. (2013). Noninvasive transcranial direct current stimulation over the left prefrontal cortex facilitates cognitive flexibility in tool use. *Cogn. Neurosci.* 4, 81–89. doi: 10.1080/17588928.2013.768221
- Chrysikou, E. G., Weber, M. J., and Thompson-Schill, S. L. (2014). A matched filter hypothesis for cognitive control. *Neuropsychologia* 62, 341–355. doi: 10.1016/j.neuropsychologia.2013.10.021
- Collins, A. G. E., Ciullo, B., Frank, M. J., and Badre, D. (2017). Working memory load strengthens reward prediction errors. *J. Neurosci.* 37, 4332–4342. doi: 10.1523/JNEUROSCI.2700-16.2017
- Collins, A. G. E., and Frank, M. J. (2012). How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis. *Eur. J. Neurosci.* 35, 1024–1035. doi: 10.1111/j.1460-9568.2011.07980.x

- Collins, A. G. E., and Frank, M. J. (2014). Opponent actor learning (OpAL): modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive. *Psychol. Rev.* 121, 337–366. doi: 10.1037/a0037015
- Cools, R., and D'Esposito, M. (2009). "Dopaminergic modulation of flexible cognitive control in humans," in *Dopamine handbook*, eds. L. L. Iversen, S. D. Iversen, S. B. Dunnett, and A. Bjorklund (New York, NY: Oxford University Press), 249–260.
- Costa, A., Riedel, M., Müller, U., Möller, H.-J., and Ettinger, U. (2011). Relationship between SLC6A3 genotype and striatal dopamine transporter availability: a meta-analysis of human single photon emission computed tomography studies. *Synapse* 65, 998–1005. doi: 10.1002/syn.20927
- Cragg, S. J., and Rice, M. E. (2004). DAncing past the DAT at a DA synapse. *Trends Neurosci.* 27, 270–277. doi: 10.1016/j.tins.2004.03.011
- da Silva Alves, F., Figue, M., van Amelsvoort, T., Veltman, D., and de Haan, L. (2008). The revised dopamine hypothesis of schizophrenia: evidence from pharmacological MRI studies with atypical antipsychotic medication. *Psychopharmacol. Bull.* 41, 121–132. doi: 10.1016/S0920-9964(08)70291-3
- Dauvizeau, J., Adam, V., and Rigoux, L. (2014). VBA: A probabilistic treatment of nonlinear models for neurobiological and behavioural data. *PLoS Comput. Biol.* 10:e1003441. doi: 10.1371/journal.pcbi.1003441
- Daw, N. D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 8, 1704–1711. doi: 10.1038/nn1560
- Delgado, M. R., Frank, R. H., and Phelps, E. A. (2005). Perceptions of moral character modulate the neural systems of reward during the trust game. *Nat. Neurosci.* 8, 1611–1618. doi: 10.1038/nn1575
- den Ouden, H. E. M., Daw, N. D., Fernandez, G., Elshout, J. A., Rijpkema, M., Hoogman, M., et al. (2013). Dissociable effects of dopamine and serotonin on reversal learning. *Neuron* 80, 1090–1100. doi: 10.1016/j.neuron.2013.08.030
- Dolan, R. J., and Dayan, P. (2013). Goals and habits in the brain. *Neuron* 80, 312–325. doi: 10.1016/j.neuron.2013.09.007
- Doll, B. B., Hutchison, K. E., and Frank, M. J. (2011). Dopaminergic genes predict individual differences in susceptibility to confirmation bias. *J. Neurosci.* 31, 6188–6198. doi: 10.1523/JNEUROSCI.6486-10.2011
- Doll, B. B., Jacobs, W. J., Sanfey, A. G., and Frank, M. J. (2009). Instructional control of reinforcement learning: a behavioral and neurocomputational investigation. *Brain Res.* 1299, 74–94. doi: 10.1016/j.brainres.2009.07.007
- Doll, B. B., Waltz, J. A., Cockburn, J., Brown, J. K., Frank, M. J., and Gold, J. M. (2014). Reduced susceptibility to confirmation bias in schizophrenia. *Cogn. Affect. Behav. Neurosci.* 14, 715–728. doi: 10.3758/s13415-014-0250-6
- Doucette-Stamm, L. A., Blakely, D. J., Tian, J., Mockus, S., and Mao, J. (1995). Population genetic study of the human dopamine transporter gene (DAT1). *Genet. Epidemiol.* 12, 303–308. doi: 10.1002/gepi.1370120307
- Dreher, J.-C., Kohn, P., Kolachana, B., Weinberger, D. R., and Berman, K. F. (2009). Variation in dopamine gene influences responsivity of the human reward system. *Proc. Natl. Acad. Sci.* 106, 617–622. doi: 10.1073/pnas.0805517106
- Dreyer, J. K., Herrik, K. F., Berg, R. W., and Hounsgaard, J. D. (2010). Influence of phasic and tonic dopamine release on receptor activation. *J. Neurosci.* 30, 14273–14283. doi: 10.1523/JNEUROSCI.1894-10.2010
- Durstewitz, D., and Seamans, J. K. (2008). The dual-state theory of prefrontal cortex dopamine function with relevance to catechol-O-methyltransferase genotypes and schizophrenia. *Biol. Psychiatry* 64, 739–749. doi: 10.1016/j.biopsych.2008.05.015
- Efimova, E. V., Gainetdinov, R. R., Budygin, E. A., and Sotnikova, T. D. (2016). Dopamine transporter mutant animals: a translational perspective. *J. Neurogenet.* 30, 5–15. doi: 10.3109/01677063.2016.1144751
- Faraone, S. V., Spencer, T. J., Madras, B. K., Zhang-James, Y., and Biederman, J. (2014). Functional effects of dopamine transporter gene genotypes on *in vivo* dopamine transporter functioning: a meta-analysis. *Mol. Psychiatry* 19, 880–889. doi: 10.1038/mp.2013.126
- Fareri, D. S., Chang, L. J., and Delgado, M. R. (2012). Effects of direct social experience on trust decisions and neural reward circuitry. *Front. Neurosci.* 6:148. doi: 10.3389/fnins.2012.00148
- Fay, M. P., and Shaw, P. A. (2010). Exact and asymptotic weighted logrank tests for interval censored data: the interval R package. *J. Stat. Softw.* 36:2. doi: 10.18637/jss.v036.i02
- Filmer, H. L., Mattingley, J. B., and Dux, P. E. (2013). Improved multitasking following prefrontal tDCS. *Cortex* 49, 2845–2852. doi: 10.1016/J.CORTEX.2013.08.015
- Floresco, S. B., West, A. R., Ash, B., Moore, H., and Grace, A. A. (2003). Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. *Nat. Neurosci.* 6, 968–973. doi: 10.1038/nn1103
- Ford, C. P., Gantz, S. C., Phillips, P. E. M., and Williams, J. T. (2010). Control of extracellular dopamine at dendrite and axon terminals. *J. Neurosci.* 30, 6975–6983. doi: 10.1523/JNEUROSCI.1020-10.2010
- Fouragnan, E., Chierchia, G., Greiner, S., Neveu, R., Avesani, P., and Coricelli, G. (2013). Reputational priors magnify striatal responses to violations of trust. *J. Neurosci.* 33, 3602–3611. doi: 10.1523/JNEUROSCI.3086-12.2013
- Fox, J., and Weisberg, S. (2011). *An R Companion to Applied Regression. 2nd ed.* Thousand Oaks, CA: Sage.
- Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., and Hutchison, K. E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proc. Natl. Acad. Sci.* 104, 16311–16316. doi: 10.1073/pnas.0706111104
- Franke, B., Vasquez, A. A., Johansson, S., Hoogman, M., Romanos, J., Boreatti-Hümmer, A., et al. (2010). Multicenter analysis of the SLC6A3/DAT1 VNTR haplotype in persistent ADHD suggests differential involvement of the gene in childhood and persistent ADHD. *Neuropsychopharmacology* 35, 656–664. doi: 10.1038/npp.2009.170
- Fregni, F., Boggio, P. S., Nitsche, M., Berman, F., Antal, A., Feredoes, E., et al. (2005). Anodal transcranial direct current stimulation of prefrontal cortex enhances working memory. *Exp. Brain Res.* 166, 23–30. doi: 10.1007/s00221-005-2334-6
- Garcia-Garcia, M., Barceló, F., Clemente, I. C., and Escera, C. (2010). The role of the dopamine transporter DAT1 genotype on the neural correlates of cognitive flexibility. *Eur. J. Neurosci.* 31, 754–760. doi: 10.1111/j.1460-9568.2010.07102.x
- Gershman, S. J. (2016). Empirical priors for reinforcement learning models. *J. Math. Psychol.* 71, 1–6. doi: 10.1016/j.jmp.2016.01.006
- Gopnik, A., Griffiths, T. L., and Lucas, C. G. (2015). When younger learners can be better (or at least more open-minded) than older ones. *Curr. Dir. Psychol. Sci.* 24, 87–92. doi: 10.1177/0963721414556653
- Goto, Y., and Grace, A. A. (2005). Dopaminergic modulation of limbic and cortical drive of nucleus accumbens in goal-directed behavior. *Nat. Neurosci.* 8, 805–812. doi: 10.1038/nn1471
- Grace, A. A., and Gomes, F. V. (2018). The circuitry of dopamine system regulation and its disruption in schizophrenia: insights into treatment and prevention. *Schizophr. Bull.* 2018:sbx199. doi: 10.1093/schbul/sbx199
- Hergovich, A., Schott, R., and Burger, C. (2010). Biased evaluation of abstracts depending on topic and conclusion: further evidence of a confirmation bias within scientific psychology. *Curr. Psychol.* 29, 188–209. doi: 10.1007/s12144-010-9087-5
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* 6, 65–70. doi: 10.2307/4615733
- Jacobson, L., Koslowsky, M., and Lavidor, M. (2012). tDCS polarity effects in motor and cognitive domains: a meta-analytical review. *Exp. Brain Res.* 216, 1–10. doi: 10.1007/s00221-011-2891-9
- Karuza, E. A., Balewski, Z. Z., Hamilton, R. H., Medaglia, J. D., Tardiff, N., and Thompson-Schill, S. L. (2016). Mapping the parameter space of tDCS and cognitive control via manipulation of current polarity and intensity. *Front. Hum. Neurosci.* 10:665. doi: 10.3389/fnhum.2016.00665
- Kaya, C., Cheng, M. H., Block, E. R., Bartol, T. M., Sejnowski, T. J., Sorkin, A., et al. (2018). Heterogeneities in axonal structure and transporter distribution lower dopamine reuptake efficiency. *eNeuro* 5:ENEURO.0298-17.2017. doi: 10.1523/ENEURO.0298-17.2017
- King, D., Zigmond, M. J., and Finlay, J. M. (1997). Effects of dopamine depletion in the medial prefrontal cortex on the stress-induced increase in extracellular dopamine in the nucleus accumbens core and shell. *Neuroscience* 77, 141–153. doi: 10.1016/S0306-4522(96)00421-6
- Kohn, M., Nurmi, E. L., Laughlin, C. P., Morales, A. M., Gail, E. H., Hellemann, G. S., et al. (2016). Functional genetic variation in dopamine signaling moderates prefrontal cortical activity during risky decision making. *Neuropsychopharmacology* 41, 695–703. doi: 10.1038/npp.2015.192
- Kolachana, B. S., Saunders, R. C., and Weinberger, D. R. (1995). Augmentation of prefrontal cortical monoaminergic activity inhibits dopamine release in the

- caudate nucleus: an *in vivo* neurochemical assessment in the rhesus monkey. *Neuroscience* 69, 859–868. doi: 10.1016/0306-4522(95)00246-F
- Krugel, L. K., Biele, G., Mohr, P. N. C., Li, S.-C., and Heekeren, H. R. (2009). Genetic variation in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt decisions. *Proc. Natl. Acad. Sci.* 106, 17951–17956. doi: 10.1073/pnas.0905191106
- Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychol. Rev.* 96, 674–689. doi: 10.1037/0033-295X.96.4.674
- Lenth, R. V. (2016). Least-squares means: the R package lsmeans. *J. Stat. Softw.* 69, 1–33. doi: 10.18637/jss.v069.i01
- Li, J., Delgado, M. R., and Phelps, E. A. (2011). How instructed knowledge modulates the neural systems of reward learning. *Proc. Natl. Acad. Sci.* 108, 55–60. doi: 10.1073/pnas.1014938108
- Lupyan, G., Mirman, D., Hamilton, R., and Thompson-Schill, S. L. (2012). Categorization is modulated by transcranial direct current stimulation over left prefrontal cortex. *Cognition* 124, 36–49. doi: 10.1016/j.cognition.2012.04.002
- MacCoun, R. J. (1998). Biases in the interpretation and use of research results. *Annu. Rev. Psychol.* 49, 259–287. doi: 10.1146/annurev.psych.49.1.259
- Mahoney, M. J. (1977). Publication prejudices: an experimental study of confirmatory bias in the peer review system. *Cognit. Ther. Res.* 1, 161–175. doi: 10.1007/BF01173636
- Medic, N., Ziauddeen, H., Vestergaard, M. D., Henning, E., Schultz, W., Farooqi, I. S., et al. (2014). Dopamine modulates the neural representation of subjective value of food in hungry subjects. *J. Neurosci.* 34, 16856–16864. doi: 10.1523/JNEUROSCI.2051-14.2014
- Meyer-Lindenberg, A., Kohn, P. D., Kolachana, B., Kippenhan, S., McNerney-Leo, A., Nussbaum, R., et al. (2005). Midbrain dopamine and prefrontal function in humans: interaction and modulation by COMT genotype. *Nat. Neurosci.* 8, 594–596. doi: 10.1038/nn1438
- Miller, E. K., and Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202. doi: 10.1146/annurev.neuro.24.1.167
- Minarik, T., Berger, B., Althaus, L., Bader, V., Biebl, B., Brotzeller, F., et al. (2016). The importance of sample size for reproducibility of tDCS effects. *Front. Hum. Neurosci.* 10:453. doi: 10.3389/fnhum.2016.00453
- Nickerson, R. S. (1998). Confirmation bias: a ubiquitous phenomenon in many guises. *Rev. Gen. Psychol.* 2, 175–220. doi: 10.1037//1089-2680.2.2.175
- Nieratschker, V., Kiefer, C., Giel, K., Krüger, R., and Plewnia, C. (2015). The COMT Val/Met polymorphism modulates effects of tDCS on response inhibition. *Brain Stimul.* 8, 283–288. doi: 10.1016/j.brs.2014.11.009
- Nikolova, Y. S., Ferrell, R. E., Manuck, S. B., and Hariri, A. R. (2011). Multilocus genetic profile for dopamine signaling predicts ventral striatum reactivity. *Neuropsychopharmacology* 36, 1940–1947. doi: 10.1038/npp.2011.82
- Nitsche, M. A., Nitsche, M. S., Klein, C. C., Tergau, F., Rothwell, J. C., and Paulus, W. (2003). Level of action of cathodal DC polarisation induced inhibition of the human motor cortex. *Clin. Neurophysiol.* 114, 600–604. doi: 10.1016/S1388-2457(02)00412-1
- Nitsche, M. A., and Paulus, W. (2000). Excitability changes induced in the human motor cortex by weak transcranial direct current stimulation. *J. Physiol.* 527, 633–639. doi: 10.1111/j.1469-7793.2000.t01-1-00633.x
- Nitsche, M. A., and Paulus, W. (2001). Sustained excitability elevations induced by transcranial DC motor cortex stimulation in humans. *Neurology* 57, 1899–1901. doi: 10.1212/WNL.57.10.1899
- Niv, Y. (2009). Reinforcement learning in the brain. *J. Math. Psychol.* 53, 139–154. doi: 10.1016/j.jmp.2008.12.005
- Niv, Y., Daw, N. D., Joel, D., and Dayan, P. (2007). Tonic dopamine: Opportunity costs and the control of response vigor. *Psychopharmacology* 191, 507–520. doi: 10.1007/s00213-006-0502-4
- Nozari, N., and Thompson-Schill, S. L. (2013). More attention when speaking: does it help or does it hurt? *Neuropsychologia* 51, 2770–2780. doi: 10.1016/j.neuropsychologia.2013.08.019
- Nozari, N., Woodard, K., and Thompson-Schill, S. L. (2014). Consequences of cathodal stimulation for behavior: when does it help and when does it hurt performance? *PLoS ONE* 9:e84338. doi: 10.1371/journal.pone.0084338
- O'Doherty, J. P. (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr. Opin. Neurobiol.* 14, 769–776. doi: 10.1016/j.conb.2004.10.016
- Otto, A. R., Skatova, A., Madlon-Kay, S., and Daw, N. D. (2015). Cognitive control predicts use of model-based reinforcement learning. *J. Cogn. Neurosci.* 27, 319–333. doi: 10.1162/jocn_a_00709
- Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Front. Neuroinform.* 2:10. doi: 10.3389/neuro.11.010.2008
- Plewnia, C., Zwissler, B., Längst, I., Maurer, B., Giel, K., and Krüger, R. (2013). Effects of transcranial direct current stimulation (tDCS) on executive functions: influence of COMT Val/Met polymorphism. *Cortex* 49, 1801–1807. doi: 10.1016/j.cortex.2012.11.002
- R Core Team (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available online at: <https://www.r-project.org>
- Rice, M. E., and Cragg, S. J. (2008). Dopamine spillover after quantal release: rethinking dopamine transmission in the nigrostriatal pathway. *Brain Res. Rev.* 58, 303–313. doi: 10.1016/j.brainresrev.2008.02.004
- Rigoux, L., Stephan, K. E., Friston, K. J., and Daunizeau, J. (2014). Bayesian model selection for group studies—revisited. *Neuroimage* 84, 971–985. doi: 10.1016/j.neuroimage.2013.08.065
- Sadleir, R. J., Vannorsdall, T. D., Schretlen, D. J., and Gordon, B. (2010). Transcranial direct current stimulation (tDCS) in a realistic head model. *Neuroimage* 51, 1310–1318. doi: 10.1016/j.neuroimage.2010.03.052
- Salamone, J. D., Correa, M., Farrar, A., and Mingote, S. M. (2007). Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology* 191, 461–482. doi: 10.1007/s00213-006-0668-9
- Schielzeth, H., and Forstmeier, W. (2009). Conclusions beyond support: overconfident estimates in mixed models. *Behav. Ecol.* 20, 416–420. doi: 10.1093/beheco/arn145
- Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599. doi: 10.1126/science.275.5306.1593
- Seamans, J. K., and Yang, C. R. (2004). The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Prog. Neurobiol.* 74, 1–58. doi: 10.1016/j.pneurobio.2004.05.006
- Shimamura, A. P. (2000). The role of the prefrontal cortex in dynamic filtering. *Psychobiology* 28, 207–218. doi: 10.3758/BF03331979
- Shiner, T., Seymour, B., Wunderlich, K., Hill, C., Bhatia, K. P., Dayan, P., et al. (2012). Dopamine and performance in a reinforcement learning task: evidence from Parkinson's disease. *Brain* 135, 1871–1883. doi: 10.1093/brain/aw083
- Shumay, E., Chen, J., Fowler, J. S., and Volkow, N. D. (2011). Genotype and ancestry modulate brain's DAT availability in healthy humans. *PLoS ONE* 6:e22754. doi: 10.1371/journal.pone.0022754
- Slifstein, M., van de Giessen, E., Van Snellenberg, J., Thompson, J. L., Narendran, R., Gil, R., et al. (2015). Deficits in prefrontal cortical and extrastriatal dopamine release in schizophrenia: a positron emission tomographic functional magnetic resonance imaging study. *JAMA Psychiatry* 72:316. doi: 10.1001/jamapsychiatry.2014.2414
- Smittenaar, P., Chase, H. W., Aarts, E., Nusslein, B., Bloem, B. R., and Cools, R. (2012). Decomposing effects of dopaminergic medication in Parkinson's disease on probabilistic action selection—learning or performance? *Eur. J. Neurosci.* 35, 1144–1151. doi: 10.1111/j.1460-9568.2012.08043.x
- Smittenaar, P., FitzGerald, T. H. B., Romei, V., Wright, N. D., and Dolan, R. J. (2013). Disruption of dorsolateral prefrontal cortex decreases model-based in favor of model-free control in humans. *Neuron* 80, 914–919. doi: 10.1016/j.neuron.2013.08.009
- Staudinger, M. R., and Büchel, C. (2013). How initial confirmatory experience potentiates the detrimental influence of bad advice. *Neuroimage* 76, 125–133. doi: 10.1016/j.neuroimage.2013.02.074
- Sulzer, D., Cragg, S. J., and Rice, M. E. (2016). Striatal dopamine neurotransmission: regulation of release and uptake. *Basal Ganglia* 6, 123–148. doi: 10.1016/j.baga.2016.02.001
- Tan, H.-Y., Chen, Q., Sust, S., Buckholtz, J. W., Meyers, J. D., Egan, M. F., et al. (2007). Epistasis between catechol-O-methyltransferase and type II metabotropic glutamate receptor 3 genes on working memory brain function. *Proc. Natl. Acad. Sci.* 104, 12536–12541. doi: 10.1073/pnas.0610125104
- Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press.

- Tunbridge, E. M. (2010). "The catechol-O-methyltransferase gene: its regulation and polymorphisms," in *International Review of Neurobiology*, ed. E. Nissinen (Academic Press), 7–27.
- Tunbridge, E. M., Farrell, S. M., Harrison, P. J., and Mackay, C. E. (2013). Catechol-O-methyltransferase (COMT) influences the connectivity of the prefrontal cortex at rest. *Neuroimage* 68, 49–54. doi: 10.1016/j.neuroimage.2012.11.059
- Turi, Z., Mittner, M., Opitz, A., Popkes, M., Paulus, W., and Antal, A. (2015). Transcranial direct current stimulation over the left prefrontal cortex increases randomness of choice in instrumental learning. *Cortex* 63, 145–154. doi: 10.1016/j.cortex.2014.08.026
- Vandenberg, D. J., Persico, A. M., Hawkins, A. L., Griffin, C. A., Li, X., Jabs, E. W., et al. (1992). Human dopamine transporter gene (DAT1) maps to chromosome 5p15.3 and displays a VNTR. *Genomics* 14, 1104–1106. doi: 10.1016/S0888-7543(05)80138-7
- Waltz, J. A., Frank, M. J., Robinson, B. M., and Gold, J. M. (2007). Selective reinforcement learning deficits in schizophrenia support predictions from computational models of striatal-cortical dysfunction. *Biol. Psychiatry* 62, 756–764. doi: 10.1016/j.biopsych.2006.09.042
- Waltz, J. A., Frank, M. J., Wiecki, T. V., and Gold, J. M. (2011). Altered probabilistic learning and response biases in schizophrenia: behavioral evidence and neurocomputational modeling. *Neuropsychology* 25, 86–97. doi: 10.1037/a0020882
- Weber, M. J., Messing, S. B., Rao, H., Detre, J. A., and Thompson-Schill, S. L. (2014). Prefrontal transcranial direct current stimulation alters activation and connectivity in cortical and subcortical reward systems: a tDCS-fMRI study. *Hum. Brain Mapp.* 35, 3673–3686. doi: 10.1002/hbm.22429
- Weinberger, D. R., Berman, K. F., and Daniel, D. G. (1992). Mesoprefrontal cortical dopaminergic activity and prefrontal hypofunction in schizophrenia. *Clin. Neuropharmacol.* 15, 568A–569A. doi: 10.1097/00002826-199201001-00296
- Witte, A. V., and Flöel, A. (2012). Effects of COMT polymorphisms on brain function and behavior in health and disease. *Brain Res. Bull.* 88, 418–428. doi: 10.1016/j.brainresbull.2011.11.012
- Wolfensteller, U., and Ruge, H. (2012). Frontostriatal mechanisms in instruction-based learning as a hallmark of flexible goal-directed behavior. *Front. Psychol.* 3:192. doi: 10.3389/fpsyg.2012.00192
- Yacubian, J., Sommer, T., Schroeder, K., Glascher, J., Kalisch, R., Leuenberger, B., et al. (2007). Gene-gene interaction associated with neural reward sensitivity. *Proc. Natl. Acad. Sci.* 104, 8125–8130. doi: 10.1073/pnas.0702029104
- Zaehle, T., Sandmann, P., Thorne, J. D., Jäncke, L., and Herrmann, C. S. (2011). Transcranial direct current stimulation of the prefrontal cortex modulates working memory performance: combined behavioural and electrophysiological evidence. *BMC Neurosci.* 12:2. doi: 10.1186/1471-2202-12-2
- Zmigrod, S., Zmigrod, L., and Hommel, B. (2016). Transcranial direct current stimulation (tDCS) over the right dorsolateral prefrontal cortex affects stimulus conflict but not response conflict. *Neuroscience* 322, 320–325. doi: 10.1016/j.neuroscience.2016.02.046

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Tardiff, Graves and Thompson-Schill. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.