Memory-reliant Post-error Slowing Is Associated with Successful Learning and Fronto-occipital Activity

Björn C. Schiffler, Rita Almeida, Mathias Granqvist, and Sara L. Bengtsson

Abstract

■ Negative feedback after an action in a cognitive task can lead to devaluing that action on future trials as well as to more cautious responding when encountering that same choice again. These phenomena have been explored in the past by reinforcement learning theories and cognitive control accounts, respectively. Yet, how cognitive control interacts with value updating to give rise to adequate adaptations under uncertainty is less clear. In this fMRI study, we investigated cognitive control-based behavioral adjustments during a probabilistic reinforcement learning task and studied their influence on performance in a later test phase in which the learned value of items is tested. We provide support for the idea that functionally relevant and memory-reliant behavioral adjustments in the form of post-error slowing during reinforcement learning are associated with test performance. Adjusting response speed after negative feedback was correlated with BOLD activity in right inferior frontal gyrus and bilateral middle occipital cortex during the event of receiving the feedback. Bilateral middle occipital cortex activity overlapped partly with activity reflecting feedback deviance from expectations as measured by unsigned prediction error. These results suggest that cognitive control and feature processing cortical regions interact to implement feedback-congruent adaptations beneficial to learning. ■

INTRODUCTION

Reinforcement learning and theories of cognitive control have both been successful in accounting for aspects of human behavior concerning learning from feedback and adjusting future responses. Reinforcement learning theory (Sutton & Barto, 1998) has been reliably used to explain human and animal behavior in tasks where one learns by trial-and-error associations that lead to maximizing reward (Schultz, 2015). These learning models rely on the measure of a prediction error, that is, how a given reward differs from what is expected. Prediction errors estimated from behavioral data and their neuronal correlates have been used to study learning from positive and negative feedback. In particular, they have been used to study how responses are adjusted according to the size of the prediction error (Steinberg et al., 2013; Cavanagh, Frank, Klein, & Allen, 2010; Cohen & Ranganath, 2007). One brain area that has consistently been implicated in the coding of positive prediction error is the striatum (see reviews by Chase, Kumar, Eickhoff, & Dombrovski, 2015; Garrison, Erdeniz, & Done, 2013). How the brain codes for negative prediction errors is less apparent. However, several studies point to a similar role of the striatum in encoding aversive outcomes and prediction errors (Asaad & Eskandar, 2011; Seymour, Daw, Dayan, Singer, & Dolan, 2007; for a review, see Delgado, Li, Schiller, & Phelps, 2008).

Research on cognitive control focuses on how selection of perceptual input, working memory (WM), and response regulation is adjusted for successful performance (Botvinick, Braver, Barch, Carter, & Cohen, 2001). One example of cognitive control is the act of slowing or outright stopping a response after the commission of an error. Post-error slowing refers to relatively higher RTs on the next trial after previous negative than after positive feedback and most likely reflects an increase in response caution, which is in accord with traditional accounts of cognitive control (Dutilh et al., 2012; Botvinick et al., 2001). ACC and medial as well as lateral PFC are thought to be associated with cognitive control, specifically in processing negative feedback and in resolving conflicts (Kerns et al., 2004; Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004; Aron, Behrens, Smith, Frank, & Poldrack, 2007). Although post-error slowing has been linked to greater error awareness (Nieuwenhuis, Ridderinkhof, Blom, Band, & Kok, 2001), specific learning benefits of slowing response speed in accordance with errors or negative feedback are still under discussion (Danielmeier & Ullsperger, 2011; Hester, Barre, Mattingley, Foxe, & Garavan, 2007; Hajcak, McDonald, & Simons, 2003).

Behavioral adjustments as postulated by cognitive control have been observed in reinforcement learning paradigms (Cavanagh et al., 2010; Frank, Moustafa, Haughey, Curran, & Hutchison, 2007). These studies showed that participants are able to adjust behavior according to stimulus-specific feedback received several seconds and trials before.

Karolinska Institutet, Stockholm, Sweden

However, it is still unclear how trial-to-trial behavioral adjustments interact with reinforcement learning processes (Ullsperger, Danielmeier, & Jocham, 2014) and if behavioral adjustments in reinforcement learning are associated with learning outcomes. In this study, we set out to explore these questions. In concrete, we aimed at investigating the association between reinforcement learning constructs and behavioral adjustments in speed of reaction and action selected, probing whether feedback-congruent behavioral adjustments that rely on memory systems lead to a better learning outcome, and finally, studying possible neuronal correlates of these processes using fMRI.

METHODS

Participants

Forty-eight healthy right-handed participants (24 women) were recruited from advertisement on a Web page (student-kaninen.se) and the Karolinska Institutet campus with surrounding areas. All participants had Swedish as their native tongue. They gave written informed consent before taking part in the study. The study was approved by the ethics committee in Stockholm, Sweden (Dnr No. 2012/1980-32).

In total, 11 participants were excluded before data analysis. Three participants were excluded because of incomplete learning or testing phase sessions. One participant had to be excluded because of excessive head motion (displacement of more than one voxel for multiple volumes). Three participants did not reach sufficient accuracy scores in the testing phase of the probabilistic learning task (scores less than 10 of a possible 16 in choosing the best symbol and, at the same time, less than 10 of 16 in avoiding the worst symbol), indicating that they did not succeed in learning the task. Finally, four participants had to be excluded because they were aware of a connection of the probabilistic learning task and the semantic priming during the scrambled sentence task (see below) as evaluated via a debriefing form subsequent to the scanning phase. The final sample thus consisted of 37 participants (17 women; age range = 18–30 years, mean age = 23.19 years, SD = 3.35 years).

This study is part of an ongoing project on the influence of self-associations on learning. The participants in this study have been primed by written sentences containing associations to "stupid" or "clever" (20 participants primed "clever" and 17 participants primed "stupid"). The priming was achieved by using the scrambled sentence task (Bargh & Chartrand, 2000). Before every 30 trials of the probabilistic learning task, during the learning phase, 12 trials of scrambled sentences were given. Thus, in total, three sets of sentences were given. The priming setup was adapted from a previous study, and further information can be found in Bengtsson, Dolan, and Passingham (2011). In this study, we address the overall effects in the whole group and control for the manipulation of self-associations in all analyses.

Probabilistic Learning Task

We used an adapted version of the probabilistic selection task (Frank, Seeberger, & O'Reilly, 2004), consisting of two phases: a learning phase and a testing phase (Figure 1).

During the learning phase, participants were presented with three different sets of symbol pairs (90 pairs overall) on the computer screen. In a forced-choice paradigm, participants were asked to choose either of the symbols by indicating left or right on a keypad. Left was indicated by the button corresponding to the participant's index finger, and right was indicated by the button corresponding to the participant's middle finger. After response, written feedback was presented that read either "correct" or "wrong" (in Swedish). Through trial and error, over the course of learning, participants acquired information about the intrinsic probabilities of feedback for each symbol. For the symbol pair AB, symbol A had an 80% chance of returning positive feedback, whereas symbol B only had a 20% chance of returning positive feedback. For the symbol pair CD, the distribution was 70%/30%, and for EF, 60%/40%, respectively (Figure 1, left). The visual character of the symbols in the AB and CD pairs was switched for 19 of the 37 participants to account for differences in visual recognizability of particular symbols. Furthermore, pair presentation order and stimulus position (left/right) were pseudorandomized across participants. Each symbol pair was presented for 4000 msec. Afterward, feedback was presented for a duration of 1000 msec, which in turn was followed by a delay of 2500 msec before the next pair was presented. Because of the pseudorandomized presentation of each symbol pair, the interval in which the same pair was presented again ranged between 7.5 (the direct next trial) and 98 sec, with an average of 21.8 sec. The learning phase was divided into three units of 30 symbol pairs each.

In the testing phase of the experiment, symbols A and B from the learning phase were paired up with symbols
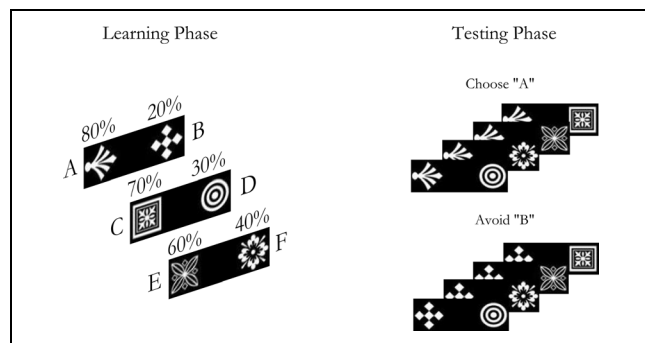

Figure 1. Probability learning task. During the learning phase, participants learned stimulus–value correspondences through trial-and-error. Symbol pairs differed in the relative amount of positive feedback: Pair AB had a distribution of 80% (A) and 20% (B) positive feedback; CD, 70%/30%; and EF, 60%/40%. After the learning phase, symbols A and B were each paired up with symbols C, D, E, and F in a testing phase in which no feedback was given.

C, D, E, and F, respectively. If participants have correctly learned the relative values of symbols during the learning phase, they should choose A against all other symbols and should not choose B against any of the other symbols (Figure 1, right). No feedback was given during the testing phase. Here, 32 trials were presented to every participant (four trials for every pairing).

Participants practiced the learning phase of the experiment twice, once outside and once inside the scanner. For the practice sessions, 25 trials were presented. During practice, different symbols were used that were unrelated to the symbols used for data collection.

Reinforcement Learning Model

A standard reinforcement learning model (van den Bos, Cohen, Kahnt, & Crone, 2012; Sutton & Barto, 1998) was set up to analyze data from the learning phase. We assumed that participants learn differently from positive and negative feedback (Klein et al., 2007; Frank, Woroch, & Curran, 2005; Frank et al., 2004) and investigated whether this differential learning would be reflected in a correspondence between a feedback-specific learning rate and behavioral parameters (e.g., post-error slowing, staying/shifting after positive/negative feedback). Therefore, we estimated two learning rates (α_{pos}, α_{neg}) to capture the influence from positive and negative feedback on behavior, respectively (van den Bos et al., 2012; Kahnt et al., 2009; Frank et al., 2007).

Decision weights of individual symbols were initialized at zero (Niv, Edlund, Dayan, & O'Doherty, 2012; van den Bos et al., 2012; Jocham, Klein, & Ullsperger, 2011; Frank, Doll, Oas-Terpstra, & Moreno, 2009; Frank et al., 2007) and updated for the chosen symbol according to the positive or negative learning rate (depending on feedback) and prediction error on a trial-by-trial basis:

$$Q_{t+1} = Q_t + \alpha_{(pos/neg)} \times \delta_t$$

Prediction errors were calculated as the difference between received feedback ($r = 1$ for positive and $r = 0$ for negative feedback) and the current decision weight of the chosen stimulus:

$$\delta_t = r_t - Q_t$$

Choice behavior was modeled by entering symbol weights on each trial into a softmax function together with the inverse temperature parameter β, which captures individuals' exploitation/exploration tendency, for example, for the symbol A in the stimulus pair AB, the probability of choosing A at time t is given by:

$$P(A)_t = \frac{1}{1 + e^{-\beta\left(Q(A)_t - Q(B)_t\right)}}$$

The two learning rates and the inverse temperature were estimated for each participant by fitting the model predictions to participants' decisions. We used the con-

strained nonlinear optimization function *fmincon* in the optimization toolbox of MATLAB to implement maximum a posteriori estimation (Daw, 2011) with the constraints $0 \leq \alpha_{pos,neg} \leq 1$ and $\beta \geq 0$. For $\alpha_{pos,neg}$, we used beta distributed priors with both shape parameters equal to 1.2, and for β, a normal distributed prior with mean = 0 and variance = 10, as used in previous work (den Ouden et al., 2013). We initialized the learning rates at 0.5 and β at 1. After model fitting, we also calculated trial-by-trial decision confidence as the absolute divergence of the probability of one of the two symbols from 50%, here shown again for stimulus pair AB:

$$\text{Conf(AB)}_t = \left|0.5 - P(A)_t\right|$$

Model Validation

To verify that our model indeed captures important aspects of participants' behavior, we simulated behavior using the obtained model parameters ($\alpha_{pos,neg}$, β) from each respective participant to make softmax-based probabilistic decisions on a trial-by-trial basis for the trial sequences of the experimental study. We repeated this procedure 10,000 times for every participant and calculated average decision accuracies over all repetitions, separately for each of the three learning pairs and the three units of the learning phase.

Behavioral Analysis

All behavioral data were analyzed within R (R version 3.0.3; R Core Team, 2014) and MATLAB (The MathWorks, Natick, MA). Mixed-level model analyses with participants as random effects were conducted using the linear mixed-effects model R package *lme4* (Bates, Maechler, Bolker, & Walker, 2014) and restricted maximum likelihood estimation.

Learning Phase Accuracy and ΔRT

Accuracy during the learning phase was calculated for every symbol pair as the fraction of choices for the symbol that was rewarded more on average (i.e., higher accuracy for choosing symbols ACE over symbols BDF). We determined accuracy on every symbol pair for each of the three units of the learning phase.

We calculated RT differences in the learning phase on a trial-to-trial basis by subtracting the subsequent trial RT from the current trial RT (ΔRT). In accordance with previous research (Cavanagh et al., 2010), we differentiated between ΔRT on the next same symbol pair (ΔRTpair, e.g., RT on subsequent pair AB trial minus RT on current pair AB trial; Figure 2) and ΔRT on the direct next trial only when the subsequent pair was different (ΔRTdirect).

We studied the effect of learning phase (Unit 1, 2, or 3) and symbol pair (AB, CD, or EF) on accuracy and probabilities

Figure 2. Post-error slowing in the probability learning task. Three different pairs (AB, CD, and EF) are presented in pseudorandomized order and symbol position (left/right). Displayed is the choice of symbol "B," followed by negative feedback and corresponding hypothesized RT slowing on the subsequent same pair trial. Symbol denominations are not shown to participants during actual experiment. Inset shows an example trial as displayed during the experiment.
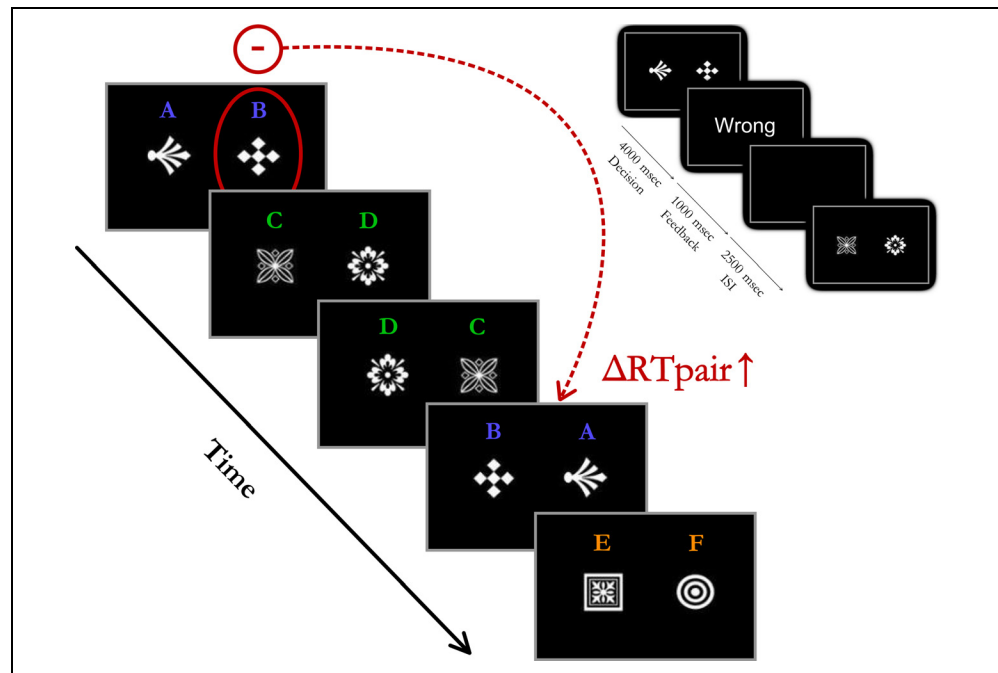
computed from the reinforcement learning model using linear mixed models with participant as random factor and prime, learning phase, and symbol pair as fixed factors. We used *F* tests and Satterthwaite's approximation to the degrees of freedom to test the main effects of learning phase and symbol pair, as implemented in the R package *lmerTest* (Kuznetsova, Brockhoff, & Christensen, 2015). Post hoc pairwise tests (single-step method) were performed using the R package *multcomp* (Hothorn et al., 2015).

Learning Phase Measures and Their Relation to Testing Phase

To determine whether relevant post-error behavioral adjustments (i.e., on the next same pair) have an effect on the testing phase of the probabilistic learning task, we used a multiple regression analysis with $\Delta RTpair$ after negative feedback ($\Delta RTpair_{neg}$), percentage of negative feedback-congruent shifting behavior, and accuracy of the three symbol pairs during the learning phase (i.e., percentage of choosing the better over the worse symbols; Zaghloul et al., 2012) as independent variables and total test score as the dependent variable. Using mixed models, we investigated whether individual ΔRT after positive and negative feedback could be predicted on a trial-by-trial basis by prediction errors and subjective confidence estimated by our reinforcement model as well as whether that trial was followed by staying with the same or shifting response toward the other symbol. Furthermore, we explored whether $\Delta RTpair_{neg}$ was associated with trial onset, that is, whether it changes over time during the course of the experiment or with one

of the three symbol pairs in particular using onset time and symbol pair as independent variables in a mixed model.

On the basis of previous work (Frank et al., 2007), we also investigated whether a general working memory (gWM) component as reflected in feedback-congruent staying/shifting on the first five trials of every pair would show a relation to the executive process measured by post-error slowing:

$$\text{gWM} = \sum_{i=1}^{5} x_i + \sum_{j=1}^{5} y_j$$

where i and j index positive and negative feedback trials, respectively, and

$$x_i = \begin{cases} 1 \text{ if stay} \,|\, \text{pos feedback} \\ 0 \text{ if shift} \,|\, \text{pos feedback} \end{cases} \quad y_j = \begin{cases} 1 \text{ if shift} \,|\, \text{neg feedback} \\ 0 \text{ if stay} \,|\, \text{neg feedback} \end{cases}$$

Trials with a missing response and corresponding direct next trials or next pair trials were removed from all analyses. Missed response trials corresponded to less than 1% of all trials in all analyses.

For the testing phase, we calculated an overall test score for the new symbol combinations corresponding to how often participants chose A and avoided B against all other symbols.

Image Acquisition: fMRI and Anatomical Data

Imaging data were acquired on a 3-T GE scanner (Discovery MR750; GE, Fairfield, CT) using an eight-channel head coil. We acquired 40 axial slices in interleaved order using a gradient-echo EPI sequence (flip angle = 90°, echo

time = 30 msec, repetition time = 2600 msec) with a field of view of 28.8 cm, matrix size of 96 × 96 (zero-filled before inverse fast Fourier transform to 128 × 128), slice thickness of 3 mm, and slice spacing of 0.5 mm. Three hundred eighty-six volumes were acquired for the learning session, whereas the amount of acquired volumes for the testing session varied depending on participants' RTs.

In addition, we acquired high-resolution T1-weighted anatomical images of every participant after the fMRI sessions, using a fast inversion-recovery-prepared 3-D gradient-echo sequence (BRAVO) with a flip angle of 12°, inversion time of 450 msec, repetition time of 7.9 msec, echo time of 3.1 msec, field of view of 24 cm, matrix size of 240 × 240 (zero-filled before inverse fast Fourier transform to 256 × 256), slice thickness of 1 mm (no gap), and interleaved acquisition.

Image Preprocessing

Imaging data were preprocessed in SPM (SPM 12b, version 6033; Wellcome Trust Centre for Neuroimaging, UCL, London, United Kingdom) as run in MATLAB, following standard preprocessing analysis protocols. Briefly, the first six volumes of every session were discarded to account for early saturation effects, and the remaining volumes were realigned to the mean image of every session. For both anatomical and functional images, the origin was set to the anterior commissure. Anatomical images were bias corrected via Segment, and functional images were then coregistered to the bias-corrected T1. Using the forward deformation field estimated from the segmentation step, the bias-corrected T1 as well as the functional images were normalized into Montreal Neurological Institute (MNI) 152 space with a spatial resolution of 2 × 2 × 2 mm. Finally, smoothing in form of a Gaussian kernel (FWHM = 8 mm) was applied to the functional images.

fMRI Data Analysis

To further our understanding of the link between typical cognitive control parameters and parameters extracted from reinforcement learning modeling with brain imaging, we first investigated parametric neural processes of post-error slowing during the feedback presentation in the learning phase. We then investigated BOLD that changed with absolute and signed prediction errors at the event of seeing the feedback. Subsequently, we studied if there was any anatomical overlap between the neural processes involved in the response to absolute prediction error and post-error slowing.

We used an informed basis set (canonical hemodynamic response function plus time and dispersion derivatives; Friston et al., 1998) to model the hemodynamic response at the single participant level. Separate general linear models were fit to the high-pass filtered data (cutoff frequency = 1/128 Hz).

For each participant, in a first level analysis, we modeled 12 regressors and their derivatives (42 regressors overall including six movement regressors). We modeled onsets of positive and negative feedback trials parametrically modulated by ΔRTpair as well as decision phase trials divided according to whether people were slowing (ΔRTpair > 0) or speeding (ΔRTpair < 0) during this phase and whether previous feedback was positive or negative (i.e., four regressors concerning the decision phase). Further regressors included the unmodulated onsets of positive and negative feedback trials. The phase between the key response and the feedback (expectation phase) was modeled from onsets of the participant's keypress until the feedback was received and divided into trials with average positive expectation (i.e., decisions for symbols A, C, or E) and average negative expectation (i.e., decisions for symbols B, D, or F). Furthermore, we modeled the sentences from the scrambled sentence task and corresponding keypresses. All regressors were modeled as events (delta stick functions) apart from the expectation phase, which was modeled as epochs with variable duration corresponding to the time from keypress until feedback and the sentences that were modeled as epochs of 8 sec.

The first level design differed slightly for the analysis of signed prediction error for which positive and negative feedback onsets were parametrically modulated by respective prediction errors instead of ΔRT and for the analysis of unsigned prediction error, in which we did not differentiate between positive and negative feedback and parametrically modulated all feedback trials by the unsigned prediction error (36 regressors in total).

First, we addressed the question whether brain activity during the reception of negative feedback was associated with slowing down on the next same pair trial. We used negative feedback trials as onsets and ΔRTpair_neg as a parametric modulator for these events.

Second, we investigated the effect of deviance from expectations (i.e., surprise) on brain activity when receiving feedback. This was implemented by using positive and negative feedback trials as onsets and modulating them by the unsigned prediction error on that trial, reflecting brain activity that increases with amount of surprise.

Parameter estimates from these single-participant analyses were taken up to the respective second level and included in one-sample ANOVAs, testing for the effects across the group. Here, we included the canonical hemodynamic response function plus time and dispersion derivatives as well as prime as regressors of no interest. Cluster results are reported at FWE whole-brain corrected level ($p_{\text{FWE-cluster}} < .05$) with an initial cluster defining threshold of $p < .001$, uncorrected.

Finally, we analyzed whether signed prediction errors showed a specific relation to brain activity. On the basis of previous studies (Chase et al., 2015; Garrison et al., 2013; Asaad & Eskandar, 2011), we used an a priori ROI mask of the entire striatum, which was created in WFU PickAtlas

(Maldjian, Laurienti, Kraft, & Burdette, 2003), and report small volume corrected (SVC) statistics for this contrast.

We report anatomical locations of significant second level activations using the SPM Anatomy toolbox (Eickhoff et al., 2005) in MNI space and visualize group level activations on a group-averaged normalized structural template and rendered on the MNI 152 template brain using MRIcroGL (www.mccauslandcenter.sc.edu/mricrogl/).

RESULTS

Behavioral Results

Learning Phase Accuracy, Computed Probabilities, and ΔRT

As expected, accuracy depended both on learning phase ($F(2, 292) = 17.119, p < .001$) and symbol pair ($F(2, 292) = 11.445, p < .001$), and this effect was similar for accuracies computed from the reinforcement learning model simula-

tions for learning phase ($F(2, 292) = 81.287, p < .001$) and for symbol pair ($F(2, 292) = 69.693, p < .001$). As revealed in the post hoc pairwise tests, significant differences in accuracy were found between all units ($2 > 1: z = 2.619, p = .024; 3 > 1: z = 5.841, p < .001; 3 > 2: z = 3.222, p = .004$; Figure 3A) and between symbol pairs AB and CD (AB > CD: $z = 4.486, p < .001$) as well as AB and EF (AB > EF: $z = 3.683, p < .001$), but not between CD and EF ($p = .7$). We found similar differences in the accuracies of the reinforcement learning model simulations between the learning units ($2 > 1: z = 10.59, p < .001; 3 > 1: z = 11.44, p < .001$; no significant difference between 3 and 2, $p = .672$; Figure 3B) and between symbol pairs (AB > CD: $z = 9.088, p < .001$; AB > EF: $z = 11.07, p < .001$; no significant difference between EF and CD, $p = .117$).

During the last unit of the learning phase, accuracy on pair AB was high (70% accuracy or higher) for all participants but one (50% accuracy). Exclusion of this participant did not have an impact on any of the conclusions
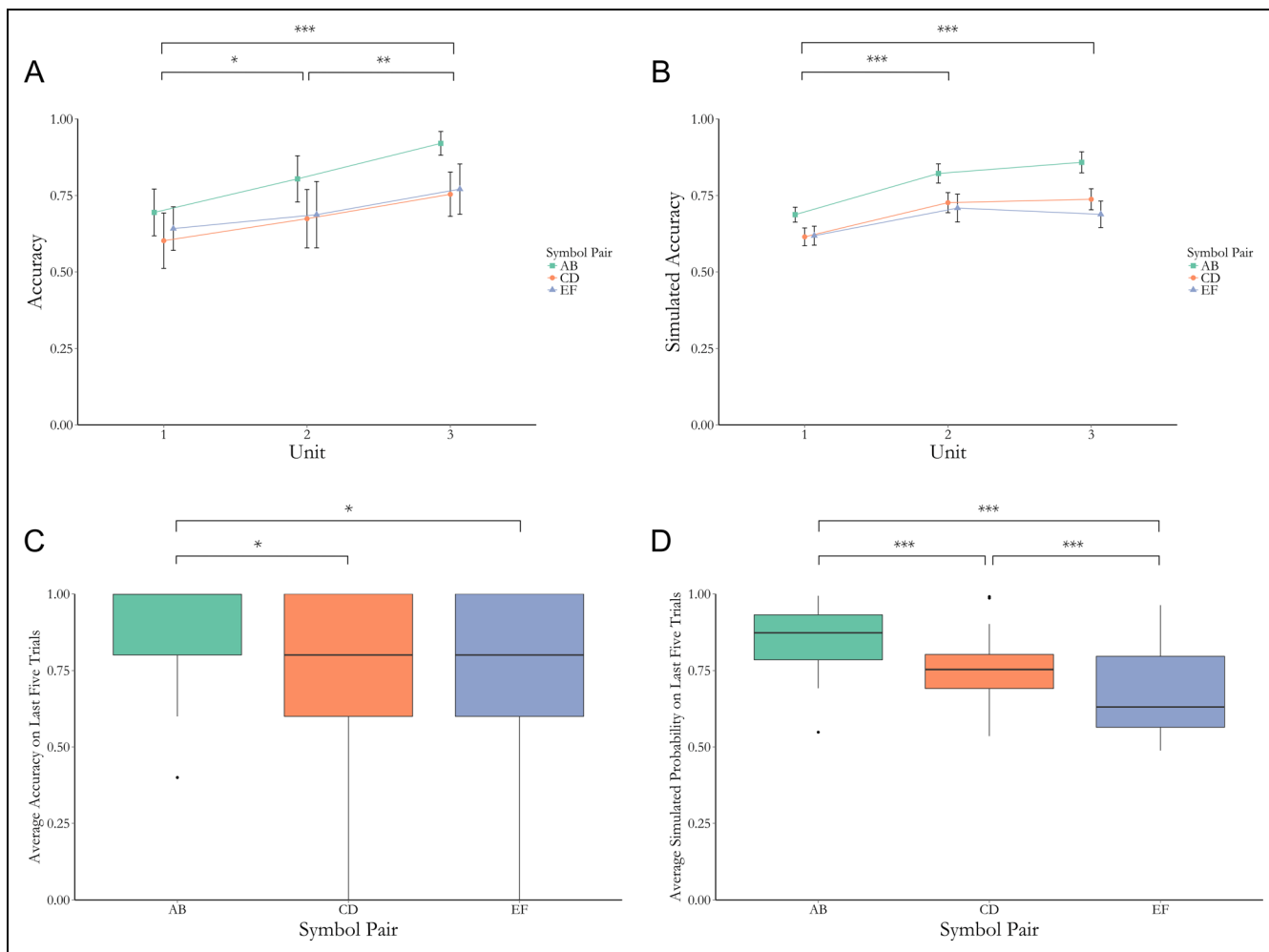

Figure 3. Learning phase accuracy and model simulations. (A) Average accuracy for all three symbol pairs over the three units of the learning phase. One unit consisted of 30 trials. Error bars indicate 95% CI. (B) Average simulated accuracy for all three symbol pairs over the three units of the learning phase. (C) Boxplots of average accuracy of the last five learning trials for symbol pairs AB, CD, and EF, respectively. (D) Boxplots of average simulated reinforcement learning model probabilities of the last five learning phase trials for symbol pairs AB, CD, and EF, respectively. *$p < .05$, **$p < .01$, ***$p < .001$.
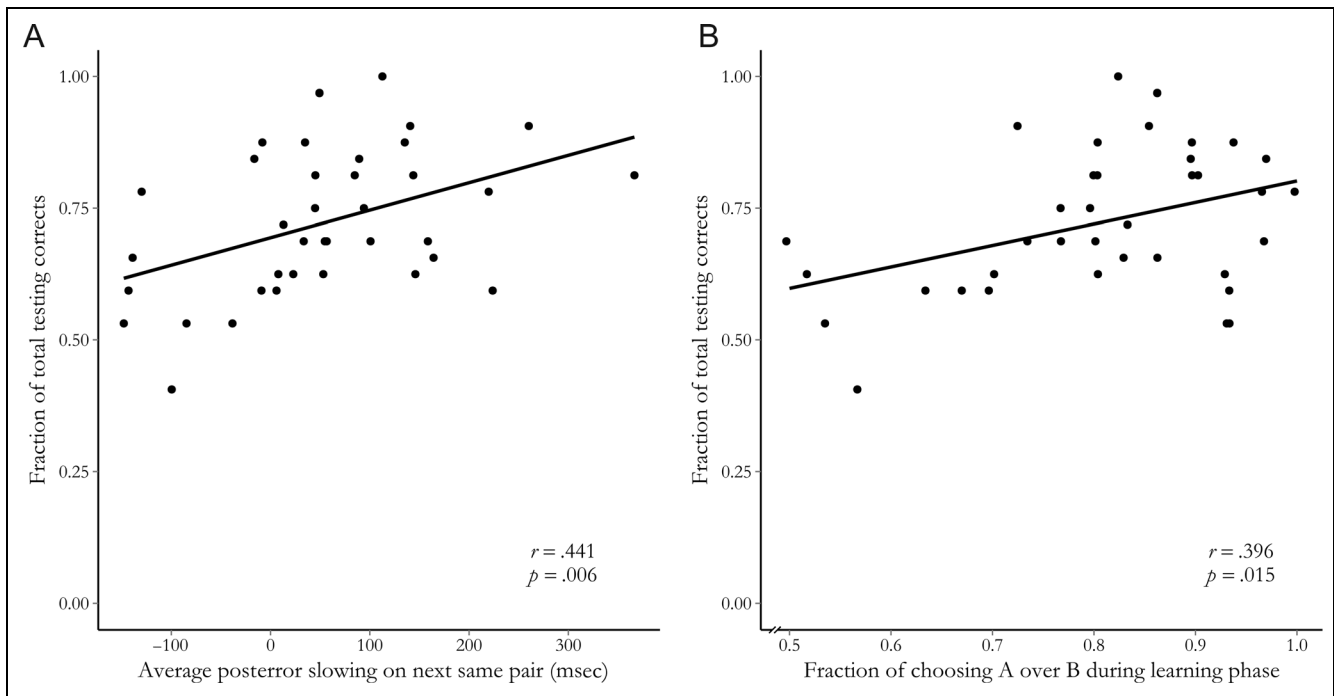
Figure 4. Visualization of relationship between learning phase and testing phase measures. (A) Scatterplot illustrating the correlation between post-error slowing on the next same pair during the learning phase and total score during the testing phase over participants. (B) Scatterplot illustrating the correlation between accuracy on pair AB during the learning phase and total score during the testing phase over participants. A random jitter of 0.005 has been applied to x axis values to display overlapping data points.

drawn from the results, so we are presenting all results with this participant included.

Accuracy on the last five trials of the learning task for every pair demonstrated that participants performed better on symbol pair AB compared with the other two symbol pairs (overall effect of symbol pair: $F(2, 72) = 4.734$, $p = .012$; $accuracy_{AB} > accuracy_{CD}$: $z = 2.845, p = .012$; $accuracy_{AB} > accuracy_{EF}$: $z = 2.438, p = .039$; no significant difference between $accuracy_{CD}$ and $accuracy_{EF}$, $p = .913$; Figure 3C). The averages of the last five probabilities computed from the reinforcement learning model simulations for every symbol pair during the learning phase reflect this learned differentiation between AB and the other two symbol pairs (overall effect of symbol pair: $F(2, 72) = 53.753, p < .001$; p(A) > p(C): $z = 5.386, p < .001$; p(A) > p(E): $z = 10.366, p < .001$; p(C) > p(E): $z = 4.98, p < .001$; Figure 3D). These results indicate that simulations using fitted parameters from our reinforcement learning model fit well with participants' behavior in this paradigm.

Trial-by-trial $\Delta RTpair$ differed depending on previously received feedback. Negative feedback led to slowing down of RT during the next same pair trial compared with positive feedback ($b = 123.6, t(3180) = 4.59, p < .001$). This effect persisted even when normalizing $\Delta RTpair$ by the current scale of RT, that is, dividing $\Delta RTpair$ by current RT + previous RT ($b = 78.11, t(3180) = 5.04, p < .001$). On the other hand, negative feedback led to speeding during the next irrelevant direct trial compared with positive feedback ($b = -70.07, t(2653) = -2.16, p = .03$).

Feedback-congruent Staying/Shifting

Previous feedback also influenced the relative propensity to stay with the same symbol or shift to the other one. As expected, participants were more likely to stay after previous positive feedback relative to previous negative feedback ($b = 1.605, z = 17.449, p < .001$).

Furthermore, the overall tendency to shift after negative feedback over participants correlated with the negative learning rate (α_{neg}) estimated by our reinforcement learning model ($r = .3298, p = .042$) in accordance with previous studies (van den Bos et al., 2012; Kahnt et al., 2009).

We did not find a significant relation between a general WM component calculated for the first five trials and average post-error slowing over participants ($r = .2014, p = .231$).

Learning Phase Measures and Relation to Testing Phase

Average $\Delta RTpair$ after negative feedback, accuracy during the learning phase on the three pairs ($accuracy_{AB}$, $accuracy_{CD}$, $accuracy_{EF}$), and shifting versus staying in response to negative feedback predicted learning transfer as reflected in the testing phase results ($F(6, 30) = 3.47$, $p = .01$, adjusted $R^2 = .29$). Test scores were significantly associated with $\Delta RTpair_{neg}$ ($b = 15.35, t(30) = 2.653, p = .013$; correlation between $\Delta RTpair_{neg}$ and total test score in Figure 4A) and $accuracy_{AB}$ ($b = 14.33, t(30) = 2.918, p = .007$; correlation between $accuracy_{AB}$ and total test score in

Figure 4B), whereas stayshift$_{neg}$, $t(30) = 1.165, p = .253$, accuracy$_{CD}$ $(t(30) = -0.277, p = .784$, and accuracy$_{EF}$, $t(30) = -0.216, p = .83$, were not. Slowing down more on the direct next trial after negative feedback was not significantly associated with better performance during the testing phase ($r = .086, p = .613$; Figure 5). ΔRTpair$_{neg}$ was not significantly associated with accuracy$_{AB}$ ($t(32) = -0.671, p = .507$), accuracy$_{CD}$ ($t(32) = 1.764, p = .087$), or accuracy$_{EF}$ ($t(32) = -1.057, p = .299$) as assessed by a multiple regression.

Relation of ΔRT to Computational Model and Behavioral Measures

Prediction errors on negative feedback were associated with ΔRTpair$_{neg}$ ($b = -345.2, t(1254) = -4.962, p < .001$), indicating that a stronger deviation of the received negative feedback from expectation led to an increase in behavioral slowing. In addition, lower confidence was also associated with ΔRTpair$_{neg}$ ($b = -488.45, t(1254) = -2.966, p = .003$), whereas the decision to stay or switch was not significantly associated with ΔRTpair$_{neg}$ ($b = -90.38, p = .066$). On the other hand, prediction errors on positive feedback predicted speeding on the next same pair trial ($b = -115.02, t(1919) = -2.458, p = .014$).

We did not find a significant relation between average ΔRTpair$_{neg}$ and the computed negative learning rate across participants ($r = .1066, p = .532$), indicating that post-error slowing might not be directly coupled with trial-by-trial value updating from negative feedback.

ΔRTpair$_{neg}$ was neither significantly related to trial onset (ΔRTpair$_{neg}$ did not change over time in the experiment,
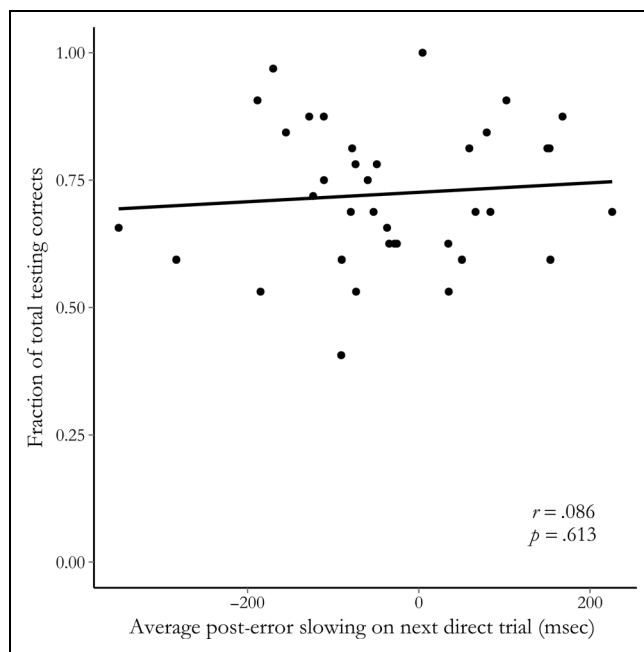

Figure 5. Scatterplot illustrating the correlation between post-error slowing on the next irrelevant direct trial during the learning phase and total score during the testing phase over participants.

$t(1255) = 1.338, p = .181$) nor to a particular symbol pair ($t(1255) = 0.992, p = .321$).

ΔRTs on the next same pair trial after positive feedback were faster when staying with the same symbol ($b = 273.45, t(1919) = 5.324, p < .001$) but were not affected by computed confidence ($b = 139.93, p = .272$).

A similar association was found between ΔRTdirect and prediction errors as well as confidence on direct next irrelevant trials. Prediction errors on negative feedback ($b = -265.23, t(1088) = -3.724, p < .001$) and confidence on the direct next trial ($b = -291.54, t(1088) = -1.996, p = .046$) were associated with slowing on the direct next trial. After positive feedback, confidence on the direct next trial ($b = -726.20, t(1560) = -6.178, p < .001$) and prediction error ($b = -362.45, t(1560) = -6.225, p < .001$) were both associated with RT speeding.

Absolute prediction errors as a measure of deviations from expectations after both positive and negative feedback were not significantly associated with speed adjustments on the next same pair trial ($b = 63.46, p = .078$) or direct next trial ($b = -65.88, p = .133$).

When using weights initialized at 0.5 instead of 0, we found a similar relation between ΔRTpair$_{neg}$ and negative prediction error ($b = -290.146, t(1254) = -4.21, p < .001$) as well as ΔRTpair$_{neg}$ and confidence ($b = -400.21, t(1254) = -2.469, p = .014$). Correlation between estimated prediction errors and confidence values with weights initialized at 0 and 0.5 was high ($r = .8925$ for prediction errors and $r = .8564$ for confidence on positive feedback trials, $r = .8923$ for prediction errors, and $r = .8422$ for confidence on negative feedback trials). This indicates that the results are robust with regard to weight initialization.

fMRI Results

Parametric Modulation of BOLD by ΔRTpair$_{neg}$

While viewing negative feedback, evoked brain activity in right inferior frontal gyrus (rIFG; peak x, y, z at 50, 20, 30: $t(1, 107) = 4.35, p_{\text{FWE-cluster}} = .013, k_E = 369$ voxels), bilateral middle occipital gyri (right peak at 32, −64, 32: $t(1, 107) = 4.55, p_{\text{FWE-cluster}} = .020, k_E = 331$ voxels; left peak at −32, −74, 26: $t(1, 107) = 3.99, p_{\text{FWE-cluster}} = .044, k_E = 263$ voxels), and right inferior occipital gyrus (peak at 26, −96, −6: $t(1, 107) = 4.34, p_{\text{FWE-cluster}} = .050, k_E = 253$ voxels) was associated with RT slowing on the next same pair trial (Figure 6 and Table 1).

Parametric Modulation of BOLD by Absolute Prediction Error

Absolute prediction error on all feedback trials was associated with activity in right middle frontal gyrus (peak at 44, 38, 18: $t(1, 107) = 4.44, p_{\text{FWE-cluster}} = .014, k_E = 353$ voxels) and bilateral parietal/occipital cortex (right peak at 30, −64, 40: $t(1, 107) = 4.36, p_{\text{FWE-cluster}} = .011, k_E = 373$ voxels;

Figure 6. fMRI results. BOLD activity on negative feedback associated with $\Delta RTpair_{neg}$, overlaid on group-averaged anatomical template in MNI space. Color bar indicates t values. Results are shown at $p < .001$ uncorrected for display purposes.
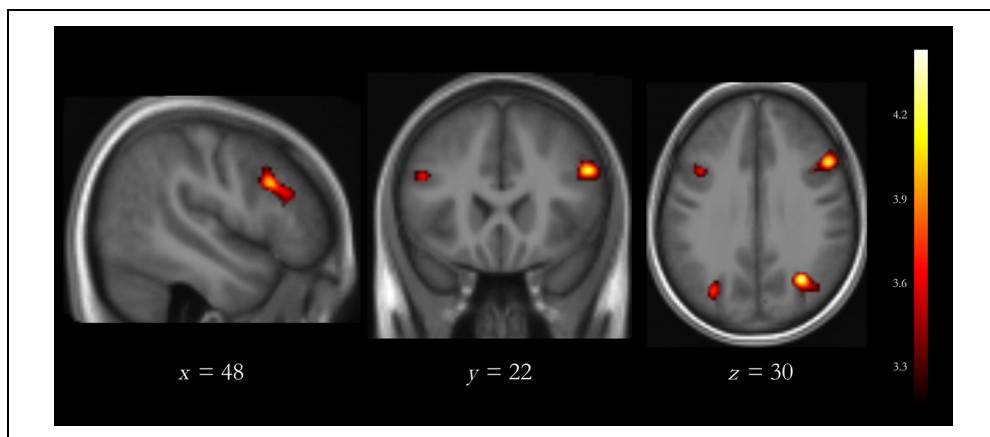
left cluster extending rostrally over left inferior parietal lobule and caudally over left middle occipital gyrus, peak at −46, −42, 46: $t(1, 107) = 5.05$, $p_{FWE\text{-}cluster} < .001$, $k_E = 1210$ voxels) as well as activity in left precentral gyrus (peak at −30, 0, 48: $t(1, 107) = 4.51$, $p_{FWE\text{-}cluster} = .024$, $k_E = 305$ voxels) and right superior frontal gyrus (peak at 26, 22, 58: $t(1, 107) = 4.45$, $p_{FWE\text{-}cluster} = .045$, $k_E = 254$ voxels; Figure 7 and Table 1).

Common Activations of $\Delta RTpair_{neg}$ and Absolute Prediction Error

As activity associated with $\Delta RTpair_{neg}$ modulation on negative feedback trials and absolute PE modulation on all feedback trials recruited in part similar areas, we used a logical AND procedure to assess common significant effects across both contrasts. We find overlap in bilateral occipital areas active in both contrasts (left side: 104 voxels, right side: 102 voxels) as well as overlap in one voxel of rIFG (MNI coordinates: 48, 34, 24), which are shown in Figure 7.

Striatum ROI Analysis of Signed Prediction Error

In an ROI analysis focusing on the striatum, prediction errors on negative feedback scaled negatively with activity in left caudate (−14, 14, 2: $t(1, 107) = 3.87$, $p_{FWE} = .042$ [SVC]), that is, higher activity corresponded to a

Table 1. fMRI Results

Contrast/Cerebral Region	MNI Peak Coordinates				Cluster Extent
	x	y	z	t	
Negative feedback parametrically modulated by $\Delta RTpair_{neg}$					
Right inferior frontal gyrus*	50	20	30	4.35	369
Right middle occipital gyrus*	32	−64	32	4.55	331
Left middle occipital gyrus*	−32	−74	26	3.99	263
Right inferior occipital gyrus*	26	−96	−6	4.34	253
Positive and negative feedback parametrically modulated by absolute prediction error					
Right middle frontal gyrus*	44	38	18	4.44	353
Right middle/superior occipital gyrus*	30	−64	40	4.36	373
Left inferior parietal lobule and left middle occipital gyrus*	−46	−42	46	5.05	1210
Right superior frontal gyrus*	26	22	58	4.45	254
Left precentral gyrus*	−30	0	48	4.51	305
Negative feedback parametrically modulated by negative prediction error					
Left caudate**	−14	14	2	3.87	62

*Significant at $p < .05$ (whole-brain FWE cluster level-corrected).

**Significant at $p < .05$ (SVC).

Figure 7. fMRI results. fMRI analysis showing overlap (blue) of significant clusters for ΔRTpair$_{neg}$ analysis (red–yellow) and absolute prediction error analysis (green). Color bar indicates *t* values. Only clusters surviving a threshold of $p_{FWE-cluster} < .05$ with an initial cluster defining threshold of $p < .001$ uncorrected are shown.
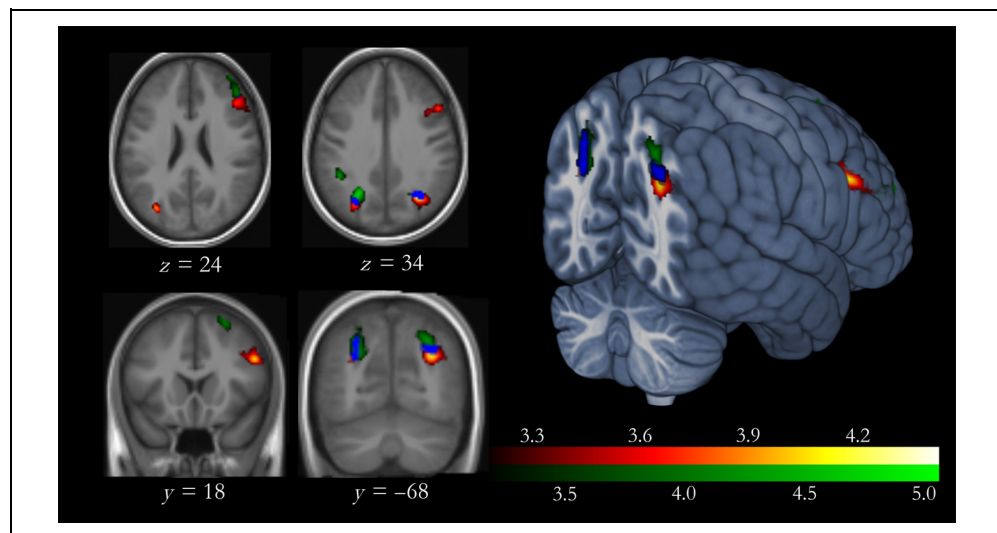
more negative prediction error. On the other hand, we did not find brain activity increasing or decreasing with positive prediction error in the striatum at this threshold.

DISCUSSION

This study demonstrates that behavioral slowing after negative feedback is associated with improved test phase performance, in which symbols are being compared to evaluate the unique probabilities learned. Importantly, we found that the post-error slowing concerns pair-specific slowing in particular that included memory-reliant processes of, on average, 22 sec. This memory-based adjustment was associated with activity in rIFG and bilateral middle occipital gyrus at the time when the participants were given feedback. Activity in these regions varied according to the size of the post-error slowing on the next same pair trial. The BOLD signal in similar occipital cortex regions correlated with absolute prediction errors when the participants were presented with feedback, and we further discovered a relation between negative prediction error and post-error slowing. These findings highlight an intriguing interplay between reinforcement learning and cognitive control processes.

Pair-specific Post-error Slowing

Previously, delayed adjustments of RTs and neural correlates to behavioral adjustments have been demonstrated in this type of reinforcement learning paradigm (Cavanagh et al., 2010; Frank et al., 2007). For example, Cavanagh and colleagues found both behavioral slowing after negative feedback on the next same pair as well as behavioral speeding after positive feedback. Behavioral speeding on the next same pair was predicted by heightened theta power in lateral PFC.

This study supports these behavioral findings and points toward a function of lateral PFC in regulating re-

sponse speed adaptations, even for a response that will occur several trials later, in accordance with current feedback. Our results suggest that rIFG implements memory-reliant inhibitory processes after negative feedback in particular as indicated by our fMRI results and thus plays a role in generating the post-error slowing. Unlike tasks that require immediate error correction, such as the classical Eriksen Flanker task (Siegert et al., 2014) or a Simon task (e.g., Danielmeier, Eichele, Forstmann, Tittgemeyer, & Ullsperger, 2011; King, Korb, von Cramon, & Ullsperger, 2010), feedback has little conceptual meaning for the immediate subsequent pair-unspecific trial in the current reinforcement learning task. Previously, no speed adjustments have been found on the direct next trial (Cavanagh et al., 2010; Frank et al., 2007), and we even find post-error speeding on the direct next trial. It is possible that this difference stems from the fact that we utilized a slowed down version of the same task, suitable for fMRI.

Role of Inferior Frontal Cortex in Processing Feedback in Accordance with Subsequent Behavioral Adjustments

During the reception of negative feedback, we observed increased activity in rIFG related to slowing down on the next same pair trial. Behaviorally, we observed that this slowing was positively related to success in the testing phase. Bilateral IFG has previously been found to play a role in instrumental learning as IFG activity differentiated learners from nonlearners when inhibiting a response that was required to obtain a reward (Guitart-Masip et al., 2012). It is reasonable to assume that maintenance of negative feedback in memory over several trials as present in our task relies on similar mechanisms instigated by rIFG and that feedback-congruent responses are beneficial for later learning outcome.

In fact, inferior frontal cortex, particularly on the right side, has previously been implicated to form part of a

cognitive control network that is involved in braking or stopping responses (Aron, Robbins, & Poldrack, 2014; Aron, Behrens, Smith, Frank, & Poldrack, 2007). A recent effective connectivity study showed that rIFG modulates the excitatory influence between the pre-SMA and subthalamic nucleus (Rae, Hughes, Anderson, & Rowe, 2015). Consistent with the idea that response inhibition requires online maintenance of relevant information, rIFG has also been implicated in storing behaviorally relevant information in memory (Spitzer, Goltz, Wacker, Auksztulewicz, & Blankenburg, 2014; Marklund & Persson, 2012; Clark et al., 2007). More generally, tasks in which storage or maintenance of value or knowledge is required over several trials evoke associated activity in lateral PFC (Curtis & D'Esposito, 2003). This stands in contrast to paradigms in which the immediate next trial calls for adjustments, such as in Simon or Stroop tasks, where brain activity predicting response threshold adaptation is often seen predominantly in medial error-related processing regions, for instance, in ACC (King et al., 2010; Kerns et al., 2004; however, see also Hester, Barre, Murphy, Silk, & Mattingley, 2008, for an involvement of medial PFC in maintenance of feedback information over several trials).

Sensory Input to Inferior Frontal Cortex and Its Relation to Unsigned Prediction Error

The abovementioned findings suggest that activity in lateral PFC goes beyond merely implementing error signals received from medial PFC. Here, we demonstrate that brain activity in rIFG correlates with relevant adjustments even when several intervening trials are presented. We also find that, in addition to rIFG activation, bilateral middle occipital cortex activity scales with the amount of post-error slowing on the next same pair trial. The parametric modulation of activity in these higher-order visual regions may reflect that they are involved in processing stimulus features when the features are particularly relevant for memory storage (Danielmeier et al., 2011; King et al., 2010; Ishai, Ungerleider, Martin, & Haxby, 2000). This is supported by our finding that bilateral middle occipital cortex activity overlapped with activity that varied parametrically with the unsigned prediction error. Unsigned prediction error can be viewed as a measure of absolute deviation from the expected outcome (feedback) and therefore as a proxy for surprise, which is a typical learning signal (Hayden, Heilbronner, Pearson, & Platt, 2011; Pearce & Hall, 1980).

Ventrolateral PFC receives converging input from the ventral visual stream, for example, information about the shape and color of stimuli (Takahashi, Ohki, & Kim, 2012; Sakagami & Pan, 2007), which it can then convert into templates for motor commands (Sakagami & Pan, 2007). Previous studies have shown that different regions in medial and lateral frontal cortex selectively interact with task-relevant and task-irrelevant brain areas to maintain sensory information needed for future decisions in mem-

ory (Spitzer et al., 2014; Danielmeier et al., 2011; King et al., 2010; see Gazzaley & Nobre, 2012, for a review). The current finding opens up the question as to whether there is a dynamic interplay between frontal areas and occipital regions that relies on information from lower level processes of detailed visual stimuli features that are in themselves calling for response speed adaptations dependent on feedback or whether the interaction between reinforcement processes and cognitive control takes place higher up in the processing hierarchy. We suggest that future studies address this particular question for a more comprehensive understanding of learning.

Test Phase Performance

Interestingly, we find that both being accurate on the symbol pair AB (i.e., choosing A more often than B) and slowing down after negative feedback on the next same pair trial contribute to instrumental learning as instantiated by the scores of a later testing phase. To our knowledge, effects of feedback-congruent adjustments on test performance have not been reported. $\Delta RTpair_{neg}$ and $accuracy_{AB}$ were associated with outcome in the testing phase to a similar degree, but with independent contributions. Previous studies have reported that greater accuracy during the learning phase benefits overall test score (Zaghloul et al., 2012), although a dissociation between learning and testing accuracy has also been described. For instance, Shohamy and colleagues found that dopamine medication selectively impaired Parkinson's patients' ability to learn action-value correspondences during a learning phase but did not have an effect on later generalization of that knowledge (Shohamy, Myers, Geghman, Sage, & Gluck, 2006). Furthermore, Klein and colleagues reported that participants with relatively increased dopamine D2 receptor density displayed more brain activity in the rostral cingulate zone during negative feedback in the learning phase, which in turn predicted positive and negative learning scores during the testing phase (Klein et al., 2007). Thus, the link between performance during the learning phase and performance during a later testing phase is under debate, and we believe that our finding that the relative duration of the memory-reliant post-error slowing impacts on learning as manifested in the test score is a valuable focus for future studies. It can be speculated that the additionally recruited time during the decision phase may promote a better integration of previous negative feedback with the stored value of a particular symbol and right inferior frontal cortex is contributing to this process (see Dixon & Christoff, 2014, for a review on the role of lateral PFC in value-based learning). In that sense, response slowing after unexpected/contrary feedback may aid the value updating process, although not relating to immediate stay or shift decisions.

We did not find evidence for a significant difference in post-error slowing when choosing to stay with the same symbol or shift to the respective other symbol. In addition,

decisions to stay with the same symbol after positive feedback or shift to the other symbol after negative feedback were not directly associated with test performance. Note that, after negative feedback, participants only shifted their response to the other symbol on about 42% of all trials (compared with about 87% staying after positive feedback), indicating that negative feedback did not always lead to an immediate choice correction. Importantly, the probabilistic structure of the task would allow for trials with negative feedback, which do not actually require a shifting decision—for example, if participants encounter negative feedback on symbol A, which had previously accumulated a large positive weight in the reinforcement history. The appropriate response in this case would be not to switch toward symbol B, yet a slowing in response speed as we observed in this study could still indicate the conflicting feedback.

As the computed symbol–action values are dependent on the participant's reinforcement history, we cannot tease apart the interaction between these two variables in this study. Yet, including post-error slowing a priori into the computational model could provide an interesting ground for exploration in the future (e.g., Niv, Daw, Joel, & Dayan, 2007).

Negative Prediction Error

We observed that post-error slowing correlated negatively with previous prediction error after negative feedback as well as confidence in the current decision. That is, a more negative prediction error on previous feedback and less confidence on the subsequent same pair trial both led to increased post-error slowing. This is in line with earlier research showing that action preparation can be influenced by uncertainty on a trial and the stimulus-related surprise (Bestmann et al., 2008). Yet, we find that negative prediction error was both associated with RT slowing on the direct next trial and the next same pair trial, indicating a general impact of negative prediction error on response speed and not a pair-specific one.

Our findings on the coding of negative prediction errors in the caudate nucleus replicate a recent study in non-human primates (Asaad & Eskandar, 2011) and provide support for the idea that negative reward prediction error is related to increasing rather than decreasing activity in error-detecting brain regions.

Conclusion

In summary, our results suggest that feedback-congruent response speed adaptations benefit learning in a reinforcement learning context. We illustrate that the slowing in RT could be predicted by using trial-by-trial estimated prediction errors via a reinforcement learning model. The brain imaging data showed that rIFG plays a role in integrating feedback to adjust memory-dependent responses relevant to the task at hand and that dorsal occipital cortex

contributes to these speed adjustments, conceivably signaling deviations from reward expectations.

REFERENCES

Aron, A. R., Behrens, T. E., Smith, S., Frank, M. J., & Poldrack, R. A. (2007). Triangulating a cognitive control network using diffusion-weighted magnetic resonance imaging (MRI) and functional MRI. *Journal of Neuroscience, 27,* 3743–3752.

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2014). Inhibition and the right inferior frontal cortex: One decade on. *Trends in Cognitive Sciences, 18,* 177–185.

Asaad, W. F., & Eskandar, E. N. (2011). Encoding of both positive and negative reward prediction errors by neurons of the primate lateral prefrontal cortex and caudate nucleus. *Journal of Neuroscience, 31,* 17772–17787.

Bargh, J. A., & Chartrand, T. L. (2000). The mind in the middle: A practical guide to priming and automaticity research. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (pp. 253–285). Cambridge: Cambridge University Press.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4 (R Package Version 1.1-6). Retrieved from ftp://ftp.ctex.org/mirrors/CRAN/web/packages/lme4/lme4.pdf.

Bengtsson, S. L., Dolan, R. J., & Passingham, R. E. (2011). Priming for self-esteem influences the monitoring of one's own performance. *Social Cognitive and Affective Neuroscience, 6,* 417–425.

Bestmann, S., Harrison, L. M., Blankenburg, F., Mars, R. B., Haggard, P., Friston, K. J., et al. (2008). Influence of uncertainty and surprise on human corticospinal excitability during preparation for action. *Current Biology, 18,* 775–780.

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

Cavanagh, J. F., Frank, M. J., Klein, T. J., & Allen, J. J. B. (2010). Frontal theta links prediction errors to behavioral adaptation in reinforcement learning. *Neuroimage, 49,* 3198–3209.

Chase, H. W., Kumar, P., Eickhoff, S. B., & Dombrovski, A. Y. (2015). Reinforcement learning models and their neural correlates: An activation likelihood estimation meta-analysis. *Cognitive, Affective & Behavioral Neuroscience, 15,* 435–459.

Clark, L., Blackwell, A. D., Aron, A. R., Turner, D. C., Dowson, J., Robbins, T. W., et al. (2007). Association between response inhibition and working memory in adult ADHD: A link to right frontal cortex pathology? *Biological Psychiatry, 61,* 1395–1401.

Cohen, M. X., & Ranganath, C. (2007). Reinforcement learning signals predict future decisions. *Journal of Neuroscience, 27,* 371–378.

Curtis, C. E., & D'Esposito, M. (2003). Persistent activity in the prefrontal cortex during working memory. *Trends in Cognitive Sciences, 7,* 415–423.

Danielmeier, C., Eichele, T., Forstmann, B. U., Tittgemeyer, M., & Ullsperger, M. (2011). Posterior medial frontal cortex activity predicts post-error adaptations in task-related visual and motor areas. *Journal of Neuroscience, 31,* 1780–1789.

Danielmeier, C., & Ullsperger, M. (2011). Post-error adjustments. *Frontiers in Psychology, 2,* 233.

Daw, N. (2011). Trial-by-trial data analysis using computational models. *Decision Making, Affect, and Learning: Attention and Performance XXIII, 23,* 3–38.

Delgado, M. R., Li, J., Schiller, D., & Phelps, E. A. (2008). The role of the striatum in aversive learning and aversive prediction errors. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences, 363,* 3787–3800.

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.

Dixon, M. L., & Christoff, K. (2014). The lateral prefrontal cortex and complex value-based learning and decision making. *Neuroscience & Biobehavioral Reviews, 45,* 9–18.

Dutilh, G., Vandekerckhove, J., Forstmann, B. U., Keuleers, E., Brysbaert, M., & Wagenmakers, E.-J. (2012). Testing theories of post-error slowing. *Attention, Perception, & Psychophysics, 74,* 454–465.

Eickhoff, S. B., Stephan, K. E., Mohlberg, H., Grefkes, C., Fink, G. R., Amunts, K., et al. (2005). A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage, 25,* 1325–1335.

Frank, M. J., Doll, B. B., Oas-Terpstra, J., & Moreno, F. (2009). Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nature Neuroscience, 12,* 1062–1068.

Frank, M. J., Moustafa, A. A., Haughey, H., Curran, T., & Hutchison, K. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proceedings of the National Academy of Sciences, U.S.A., 104,* 16311–16316.

Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in parkinsonism. *Science, 306,* 1940–1943.

Frank, M. J., Woroch, B. S., & Curran, T. (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron, 47,* 495–501.

Friston, K. J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M. D., & Turner, R. (1998). Event-related fMRI: Characterizing differential responses. *Neuroimage, 7,* 30–40.

Garrison, J., Erdeniz, B., & Done, J. (2013). Prediction error in reinforcement learning: A meta-analysis of neuroimaging studies. *Neuroscience & Biobehavioral Reviews, 37,* 1297–1310.

Gazzaley, A., & Nobre, A. C. (2012). Top–down modulation: Bridging selective attention and working memory. *Trends in Cognitive Sciences, 16,* 129–135.

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

Hajcak, G., McDonald, N., & Simons, R. F. (2003). To err is autonomic: Error-related brain potentials, ANS activity, and post-error compensatory behavior. *Psychophysiology, 40,* 895–903.

Hayden, B. Y., Heilbronner, S. R., Pearson, J. M., & Platt, M. L. (2011). Surprise signals in anterior cingulate cortex: Neuronal encoding of unsigned reward prediction errors driving adjustment in behavior. *Journal of Neuroscience, 31,* 4178–4187.

Hester, R., Barre, N., Mattingley, J. B., Foxe, J. J., & Garavan, H. (2007). Avoiding another mistake: Error and post-error neural activity associated with adaptive post-error behavior change. *Cognitive, Affective & Behavioral Neuroscience, 7,* 317–326.

Hester, R., Barre, N., Murphy, K., Silk, T. J., & Mattingley, J. B. (2008). Human medial frontal cortex activity predicts learning from errors. *Cerebral Cortex, 18,* 1933–1940.

Hothorn, T., Bretz, F., Westfall, P., Heiberger, R. M., Schuetzenmeister, A., & Scheibe, S. (2015). Package "multcomp" (R Package Version 1.3-3). Retrieved from https://cran.r-project.org/web/packages/multcomp/multcomp.pdf.

Ishai, A., Ungerleider, L. G., Martin, A., & Haxby, J. V. (2000). The representation of objects in the human occipital and temporal cortex. *Journal of Cognitive Neuroscience, 12(Suppl. 2),* 35–51.

Jocham, G., Klein, T. A., & Ullsperger, M. (2011). Dopamine-mediated reinforcement learning signals in the striatum and ventromedial prefrontal cortex underlie value-based choices. *Journal of Neuroscience, 31,* 1606–1613.

Kahnt, T., Park, S. Q., Cohen, M. X., Beck, A., Heinz, A., & Wrase, J. (2009). Dorsal striatal-midbrain connectivity in humans predicts how reinforcements are used to guide decisions. *Journal of Cognitive Neuroscience, 21,* 1332–1345.

Kerns, J. G., Cohen, J. D., MacDonald, A. W., Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science, 303,* 1023–1026.

King, J. A., Korb, F. M., von Cramon, D. Y., & Ullsperger, M. (2010). Post-error behavioral adjustments are facilitated by activation and suppression of task-relevant and task-irrelevant information processing. *Journal of Neuroscience, 30,* 12759–12769.

Klein, T. A., Neumann, J., Reuter, M., Hennig, J., von Cramon, D. Y., & Ullsperger, M. (2007). Genetically determined differences in learning from errors. *Science, 318,* 1642–1645.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2015). Package "lmerTest" (R Package Version 2.0-6). Retrieved from https://cran.r-project.org/web/packages/lmerTest/lmerTest.pdf.

Maldjian, J. A., Laurienti, P. J., Kraft, R. A., & Burdette, J. H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage, 19,* 1233–1239.

Marklund, P., & Persson, J. (2012). Context-dependent switching between proactive and reactive working memory control mechanisms in the right inferior frontal gyrus. *Neuroimage, 63,* 1552–1560.

Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P., & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: Evidence from an antisaccade task. *Psychophysiology, 38,* 752–760.

Niv, Y., Daw, N. D., Joel, D., & Dayan, P. (2007). Tonic dopamine: Opportunity costs and the control of response vigor. *Psychopharmacology, 191,* 507–520.

Niv, Y., Edlund, J., Dayan, P., & O'Doherty, J. (2012). Neural prediction errors reveal a risk-sensitive reinforcement-learning process in the human brain. *Journal of Neuroscience, 32,* 551–562.

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review, 87,* 532–552.

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

Rae, C. L., Hughes, L. E., Anderson, M. C., & Rowe, J. B. (2015). The prefrontal cortex achieves inhibitory control by facilitating subcortical motor pathway connectivity. *Journal of Neuroscience, 35,* 786–794.

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

Sakagami, M., & Pan, X. (2007). Functional role of the ventrolateral prefrontal cortex in decision making. *Current Opinion in Neurobiology, 17,* 228–233.

Schultz, W. (2015). Neuronal reward and decision signals: From theories to data. *Physiological Reviews, 95,* 853–951.

Seymour, B., Daw, N., Dayan, P., Singer, T., & Dolan, R. (2007). Differential encoding of losses and gains in the human striatum. *Journal of Neuroscience, 27,* 4826–4831.

Shohamy, D., Myers, C. E., Geghman, K. D., Sage, J., & Gluck, M. A. (2006). L-dopa impairs learning, but spares generalization, in Parkinson's disease. *Neuropsychologia, 44,* 774–784.

Siegert, S., Herrojo Ruiz, M., Brücke, C., Huebl, J., Schneider, G.-H., Ullsperger, M., et al. (2014). Error signals in the subthalamic nucleus are related to post-error slowing in patients with Parkinson's disease. *Cortex, 60,* 103–120.

Spitzer, B., Goltz, D., Wacker, E., Auksztulewicz, R., & Blankenburg, F. (2014). Maintenance and manipulation of somatosensory information in ventrolateral prefrontal cortex. *Human Brain Mapping, 35,* 2412–2423.

Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., & Janak, P. H. (2013). A causal link between prediction errors, dopamine neurons and learning. *Nature Neuroscience, 16,* 966–973.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.

Takahashi, E., Ohki, K., & Kim, D.-S. (2012). Dissociation and convergence of the dorsal and ventral visual streams in the human prefrontal cortex. *Neuroimage, 65,* 488–498.

Ullsperger, M., Danielmeier, C., & Jocham, G. (2014). Neurophysiology of performance monitoring and adaptive behavior. *Physiological Reviews, 94,* 35–79.

van den Bos, W., Cohen, M. X., Kahnt, T., & Crone, E. A. (2012). Striatum-medial prefrontal cortex connectivity predicts developmental changes in reinforcement learning. *Cerebral Cortex, 22,* 1247–1255.

Zaghloul, K. A., Weidemann, C. T., Lega, B. C., Jaggi, J. L., Baltuch, G. H., & Kahana, M. J. (2012). Neuronal activity in the human subthalamic nucleus encodes decision conflict during action selection. *Journal of Neuroscience, 32,* 2453–2460.