



Electrophysiological measures reveal the role of anterior cingulate cortex in learning from unreliable feedback

Peng Li^{1,2} · Weiwei Peng¹ · Hong Li^{1,2,3} · Clay B. Holroyd⁴

Published online: 10 July 2018
© Psychonomic Society, Inc. 2018

Abstract

Although a growing number of studies have investigated the neural mechanisms of reinforcement learning, it remains unclear how the brain responds to feedback that is unreliable. A recent theory proposes that the reward positivity (RewP) component of the event-related brain potential (ERP) and frontal midline theta (FMT) power reflect separate feedback-related processing functions of anterior cingulate cortex (ACC). In the present study, the electroencephalogram (EEG) was recorded from participants as they engaged in a time estimation task in which feedback reliability was manipulated across conditions. After each response, they received a cue that indicated that the following feedback stimulus was 100%, 75%, or 50% reliable. The results showed that participants' time estimates adjusted linearly according to the feedback reliability. Moreover, presentation of the cue indicating 100% reliability elicited a larger RewP-like ERP component than the other cues did, and feedback presentation elicited a RewP of approximately equal amplitude for all of the three reliability conditions. By contrast, FMT power elicited by negative feedback decreased linearly from the 100% condition to 75% and 50% condition, and only FMT power predicted behavioral adjustments on the following trials. In addition, an analysis of Beta power and cross-frequency coupling (CFC) of Beta power with FMT phase suggested that Beta-FMT communication modulated motor areas for the purpose of adjusting behavior. We interpreted these findings in terms of the hierarchical reinforcement learning account of ACC, in which the RewP and FMT are proposed to reflect reward processing and control functions of ACC, respectively.

Keywords Frontal midline theta · Reward positivity · Anterior cingulate cortex · Feedback reliability

Introduction

Adaptive decision making depends on a network of neural systems for cognitive control that includes the anterior cingulate cortex (ACC), a brain area believed to contribute to multiple

cognitive functions, such as action selection (Holroyd & Coles, 2002), conflict monitoring (Carter & Van Veen, 2007), and task switching (Johnston, Levin, Koval, & Everling, 2007). In the reinforcement learning domain, the ACC is proposed to utilize reward-related feedback from the environment to regulate behavior (Botvinick, 2007; Holroyd & Coles, 2002), as suggested by converging evidence from human EEG studies (for reviews see Holroyd & Umemoto, 2016; Sambrook & Goslin, 2015; Walsh & Anderson, 2012), functional magnetic resonance imaging (fMRI) studies (Holroyd, Nieuwenhuis, Yeung, Nystrom, Mars, Coles, & Cohen, 2004; Nieuwenhuis, Slagter, Alting von Geusau, Heslenfeld, & Holroyd, 2005a), and nonhuman animal studies (Warren, Hyman, Seamans, & Holroyd, 2015). Recent developments of this proposal in the context of hierarchical reinforcement learning (HRL) hold the ACC responsible for motivating the selection and execution of extended goal-directed behaviors (HRL-ACC theory; Holroyd & McClure, 2015; Holroyd & Yeung, 2012). According to this proposal, the ACC utilizes the reward signals to learn the value of extended, goal-directed behaviors, as opposed to the values of more specific actions that comprise the extended behaviors. Nevertheless,

Electronic supplementary material The online version of this article (<https://doi.org/10.3758/s13415-018-0615-3>) contains supplementary material, which is available to authorized users.

✉ Hong Li
lihongszu@szu.edu.cn

¹ Brain Function and Psychological Science Research Center, Shenzhen University, No 3688, Nanhai Road, Nanshan District, Shenzhen 518060, China

² Shenzhen Key Laboratory of Affective and Social Cognitive Science, Shenzhen University, Shenzhen, China

³ Center for Language and Brain, Shenzhen Institute of Neuroscience, Shenzhen, China

⁴ Department of Psychology, University of Victoria, Victoria, Canada

exactly how the ACC utilizes these reward signals to regulate behavior remains unclear (Holroyd & Umemoto, 2016; Umemoto, HajiHosseini, Yates & Holroyd, 2017).

Although feedback processing by ACC has been studied extensively, relatively few studies have examined the effect of unreliable or uncertain feedback on ACC. In a task in which feedback probabilities varied dynamically, the fMRI blood-oxygen-level dependent signal in the ACC was correlated with a computational estimate of reward volatility, in which simulations were used to adjust subject learning rates (Behrens et al., 2007). Similarly, Khamassi et al. (2011) developed a computational model that simulates how ACC interacts with lateral prefrontal cortex to regulate behavior in uncertain and volatile environments. Rodent electrophysiological (Karlsson, Tervo, & Karpova, 2012) and human fMRI (O'Reilly, Schüffelgen, Cuell, Behrens, Mars, & Rushworth, 2013) studies also indicated ACC sensitivity to uncertainty when updating a model of the environment. Although as noted numerous human electrophysiological studies have examined feedback processing by ACC (for reviews, see Sambrook & Goslin, 2015; Walsh & Anderson, 2012), only a few of these studies investigated how reliability modulates feedback processing (Ernst & Steinhauser, 2015; Ernst & Steinhauser, 2017; Schiffer, Siletti, Waszak, & Yeung, 2017). Moreover, these studies examined the electrophysiological effects of feedback processing only in the time domain, i.e., using the event-related brain potential (ERP), and reported inconsistent results. Furthermore, most of these studies only tested two feedback conditions (reliable or not), leaving unresolved the question of whether the ACC control signal scales parametrically with feedback reliability.

To address these questions, we examined the effect of feedback reliability on several electrophysiological measures of feedback processing. In particular, our analysis focuses on frontal midline theta oscillations (FMT, Cavanagh & Frank, 2014; Cohen et al., 2011), the reward positivity (RewP; Baker & Holroyd, 2011; Holroyd, Pakzad-Vaezi, & Krigolson, 2008), and Beta power (HajiHosseini & Holroyd, 2015a; Marco-Pallarés, Münte, & Rodríguez-Fornells, 2015), all of which have been associated with feedback processing by ACC in humans.

Previous studies have shown that FMT, which consists of 4–8 Hz EEG oscillations distributed over frontal-central areas of the scalp, is modulated by feedback valence with greater power for negative than positive feedback stimuli (Cohen, Wilmes, & Vijver, 2011; Hajihosseini & Holroyd, 2013; Li, Baker, Warren, & Li, 2016). There is widespread agreement that FMT reflects a contribution of ACC to cognitive control (Cavanagh & Frank, 2014; Holroyd & Umemoto, 2016; Verguts, 2017). FMT is associated with various cognitive processes related to memory (Hsieh & Ranganath, 2013;

Rutishauser et al., 2010), conflict processing, and sensitivity to punishment (Cavanagh & Frank, 2014; Cavanagh & Shackman, 2015). Increases in FMT power also seem to improve behavioral performance by modulating neural synchrony in other brain regions via phase-amplitude coupling (Verguts, 2017; see also Holroyd, 2016). Furthermore, a meta-analysis by Cavanagh and Shackman (2015) revealed that FMT power is positively correlated with trial-by-trial behavioral adjustments across various tasks that demand cognitive control, although it remains unclear whether those signals are causally involved in the behavioral changes (Holroyd & Umemoto, 2016).

A large number of ERP studies have examined the RewP, a robust deflection in the ERP that is more negative-going for error or nonreward feedback stimuli than for correct or reward feedback stimuli in trial-and-error learning and guessing tasks (Miltner et al., 1997). An early theory of this ERP component emphasized the role of negative reward prediction error (RPE) signals carried to ACC by the midbrain dopamine system in enhancing the negative deflection, suggesting the more common name feedback error-related negativity or feedback-related negativity (FRN) (Holroyd & Coles, 2002). More recently, this ERP component has been renamed because of evident greater sensitivity to positive RPE signals than negative RPE signals (Holroyd, Pakzad-Vaezi, & Krigolson, 2008; Holroyd & Umemoto, 2016; Proudfit, 2015).

Importantly, both the FRN (to negative feedback) and RewP (to positive feedback) are typically calculated by taking the difference between the ERPs to positive and negative feedback stimuli, which removes overlap with other ERP components (Holroyd & Krigolson, 2007); thus, when measured by using the difference wave approach, the two components actually reflect the same phenomenon (Holroyd & Umemoto, 2016). In line with the dopamine hypothesis, substantial evidence confirms that the signal encodes a reward prediction error signal (Sambrook & Goslin, 2015). On the other hand, evidence that RewP amplitude relates to behavioral adjustments, as suggested by basic principles of reinforcement learning (Sutton & Barto, 1998), is less consistent (Holroyd & Umemoto, 2016; Walsh & Anderson, 2012). Instead, RewP amplitude may reflect motivation by ACC for performing the task at hand at a global level, rather than for instigating changes in task performance from one trial to the next (Holroyd & Umemoto, 2016; Umemoto et al., 2017).

Beta oscillation also are involved in feedback processing (HajiHosseini et al., 2012; Li et al., 2016; Marco-Pallarés et al., 2008). Enhanced Beta power has been elicited over different brain areas to various task events, such as over frontal cortex to feedback indicating monetary gain (HajiHosseini et al., 2012; Marco-Pallarés, Münte, & Rodríguez-Fornells, 2015), over sensorimotor areas to correct feedback (Luft et al., 2013), and over DLPFC in a working memory task (Altamura et al., 2010). By contrast, we recently found that

high-frequency Beta (and low-frequency Gamma) activity was reduced when feedback information was uninformative compared with informative (Li et al., 2016). These studies suggest that Beta oscillations recorded over different cortical regions may correspond to specific cognitive mechanisms that are activated in different task contexts. Meanwhile, several studies have suggested that the neuronal communication between large-scale cortical networks underlying cognitive processes could be manifested by cross-frequency coupling (CFC) between lower and higher frequencies (Siegel, Donner, & Engel, 2012; Canolty & Knight, 2010). In a reinforcement learning task in particular, Beta power was reported to couple with FMT (Cohen et al., 2009). Therefore, we also examined CFC between FMT and Beta to understand how the brain uses the feedback information to update behavior from trial to trial.

We asked participants to conduct a time-estimation task in which the reliability of feedback varied from 100%, 75%, to 50% across three different stimulus conditions, despite holding constant the overall probabilities of positive versus negative feedback across the conditions. After participants made a response on each trial, a cue indicating the reliability of the following feedback stimulus was presented. We examined the RewP, FMT, and Beta power to elucidate how ACC processes reliable versus unreliable reward information to regulate behavior. We predicted that behavioral adjustments should correlate more strongly with FMT power than with RewP amplitude (Holroyd & Umemoto, 2016). Additionally, to investigate modulation by ACC of motor areas (Cohen & Ranganath, 2007; Luft et al., 2013), we explored CFC between FMT phase and Beta power over motor cortex. In this way, this experiment contributes to fMRI research in humans (Behrens et al., 2007; O'Reilly et al., 2013) and to invasive studies in nonhuman animals (Karlsson, Tervo, & Karpova, 2012) that have explored the role of ACC in regulating behavior in dynamic learning environments.

Methods

Subjects

Twenty (8 females) healthy, right-handed adults with mean age of 22 (± 2.2) years participated in the study. The study was approved by the local ethics committee, and subjects gave written informed consent before the experiments began. Subjects had normal or corrected-to-normal vision and reported no neurological disorders. The subjects received 50 Chinese Yuan (approximately 7.6 U.S. dollars) for their participation and extra money when they made a correct response during the task (25–35 Yuan total).

Task

Subjects performed a modified time-estimation task (Miltner et al., 1997). As shown in Fig. 1, at the beginning of each trial, a fixation “+” appeared on the screen for 500 ms to indicate participants to prepare to respond. Then, following delivery of a 50-ms pure tone (1,500 Hz, 65 dB) via an earphone, participants were required to press the space key when they believed that 1 s had elapsed. A blank screen was then presented for a random duration, selected from a uniform distribution between 800–1,200 ms at 1-ms intervals, and a color circle appeared on the screen for 1,000 ms. Participants were instructed that a red, white, and blue circle indicated that the feedback on that trial would be 100% reliable, 75% reliable, and 50% reliable (see below); the mappings between the different cue colors and probabilities were counterbalanced across subjects. Next, the cue stimulus was replaced with a blank screen that lasted for a randomly determined 800–1,200-ms interval. Finally, a 1,000-ms feedback stimulus indicated the trial outcome: subjects were instructed that “√” indicated a correct response, and “×” indicated an incorrect response. Each condition consisted of 65 trials, for 390 trials total. All of the visual stimuli subtended less than 5°.

In the 100% reliable condition, participants conducted the standard time estimation task in which positive and negative feedback stimuli were determined with equal frequency (50%/50%) according to a staircase procedure (Miltner et al., 1997). Correct feedback was delivered when the response time occurred within a time window that was initialized at 900–1,100 ms poststimulus; the size of this window decreased by 10 ms each time the participant received correct feedback and increased by 10 ms each time they received error feedback. In the 75% reliable condition, however, the feedback stimuli were delivered according to this procedure on only 75% of the trials; on the remaining 25% of the trials, which were selected at random, the feedback stimuli were delivered at random with equal probability. Finally, in the 50% reliable condition, the feedback stimuli on all of the trials were delivered at random (with equal probability). Note that even in the 100% condition, the outcomes were difficult to predict, because the subjects always received 50%/50% feedback, such that they could not determine the outcomes on their own accord unless their response was especially bad or especially good. The task was administered using E-prime software (Psychology Software Tools, Inc).

Post-task questionnaires

After the task was completed, participants were required to answer two questions based on their subjective experience. First, they were asked to rate their perception of the subjective reliability of feedback in each of the conditions by choosing from a list of five values (0, 25%, 50%, 75%, and 100%).

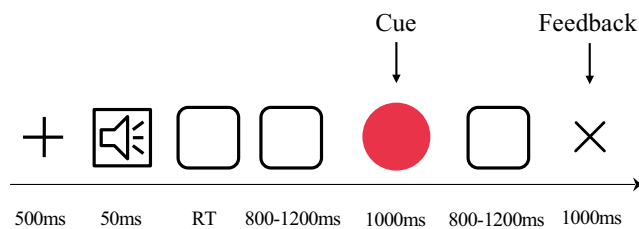


Fig. 1 Task design

Second, they were asked to rate how happy the different cues made them feel according to a 7-point Likert scale (1 = very unhappy, 7 = very happy).

Electrophysiological data recording and analysis

EEG data were recorded with a 64-channel amplifier system (BrainProduct, München, Germany) with reference electrodes placed on FCz. The impedances of all the electrodes were kept less than 10 k Ω . Vertical electrooculograms (EOGs) were recorded below the left eye and the horizontal EOG was recorded from the right orbital rim. The EEG and EOG were amplified and filtered using a 0.05–100-Hz bandpass and continuously digitized at 1,000-Hz/channel for offline analysis. The EEG data were re-referenced to both of the left and right mastoids. Trials contaminated by eye-blinks and movements were corrected using an infomax Independent Component Analysis algorithm (runica) (Delorme and Makeig, 2004). In addition, any trials containing a peak-to-peak deflection exceeding 80 μ V were excluded from the analysis. Less than 5% trials in each condition were excluded after artifact rejection.

For EEG offline analysis, EEG signals were bandpass filtered between 0.1–30 Hz and segmented from –200 ms before and 800 ms after the three cue events (100%, 75%, and 50% reliability) and six feedback events (three reliabilities \times two valences) separately. “RewP-like” ERP components to each of the three cues were constructed by subtracting the ERP to 100% cue from the ERPs to the other two cues and then assessed as the mean amplitude of these difference waves at channel Fz within 240–340 ms following feedback onset, as suggested by a meta-analysis (Sambrook & Goslin, 2015). Likewise, RewPs were constructed by subtracting the ERP to correct feedback from the ERP to incorrect feedback, separately for each reliability condition and assessed as the mean amplitude at channel Fz within the 240–340-ms time window. Fz was selected as the channel of interest, because RewP amplitude peaked at this site. Additionally, the P3b, a positive-going ERP component within 300–600-ms time window following stimulus onset, was analyzed, because this component has been frequently reported in the literature on feedback processing (Bernat et al., 2011; Li et al., 2010). Given that theories of P3b generation are not closely related to ACC function (Nieuwenhuis, Aston-Jones, & Cohen,

2005b), the method and results of the P3b analysis are reported in supplementary materials (SOM, Figure S1).

To extract both phase-locked (ERP) and nonphase-locked (ERD/ERS) brain responses, a time-frequency representation (TFR) of each single EEG epoch was calculated using the continuous wavelet transform (CWT) (Mouraux et al., 2003; Mouraux and Iannetti, 2008), from 1 to 40 Hz in steps of 0.5 Hz, and from –500 preceding to 1,000 ms after the event of interest (Cavanagh et al., 2010) in steps of 1 ms. For each frequency band, the magnitude of the event-related changes in oscillation amplitude was estimated as follows: $ER\% (t, f) = [P (t, f) - R (f)] / R (f) \times 100$, where $P (t, f) = |F(t, f)|^2$ is the power spectral density at each time-frequency point (t, f), and $R (f)$ is the average power in baseline interval (–400 to –100 ms) before cue onset and feedback onset. Hence, CWT results are expressed in ER% as a function of time and frequency. For statistical analysis, time-frequency magnitudes were averaged for each condition within specific time-frequency-spatial regions of interests: FMT (4–8 Hz, 200–400 ms) at channel Fz (Cavanagh et al., 2010; Li et al., 2016) and Beta (13–25 Hz, 400–800 ms) at channels F3&F4 (HajiHosseini & Holroyd, 2015a). For completeness, the analyses of cue-related beta power (Figure S2) and feedback-related delta power (Figure S3) are reported in the supplementary materials.

An envelope-to-signal (ESC) measure was used to assess cross-frequency coupling (CFC) between FMT oscillations and Beta oscillations (Bruns & Eckhom, 2004; Onslow, Bogacz, & Jones, 2011). Single-trial EEG data were separately bandpass filtered from 4–8 Hz (FMT) and 13–25 Hz (Beta). FMT phase and Beta power were then extracted using the Hilbert transform. For each subject and each experimental condition, correlations between the amplitude envelope of amplitude of the Beta oscillation (A_{beta}) and the phase angle of the FMT oscillation (Y_{FMT}) were calculated as $ESC = r (A_{\text{beta}}, Y_{\text{FMT}})$. The obtained ESC values (in the range between –1 and 1) were transformed to Z scores using the Fisher r-to-z transformation.

Correlation analysis

To evaluate the relationship between behavioral performance and electrophysiological responses to feedback, Pearson correlation analyses were performed between Δ RT values and RewP amplitude, FMT power, and Beta power. For each electrophysiological measure, we calculated a difference of differences, first by subtracting the EEG activities following correct feedback from that following incorrect feedback, separately for the 100% condition and 50% conditions, and then by subtracting these difference values in the 50% condition from the difference values in the 100% condition. For instance, the FMT difference value was calculated as $(100\%_{\text{FMT}}_{\text{Incorrect}} - 100\%_{\text{FMT}}_{\text{Correct}}) - (50\%_{\text{FMT}}_{\text{Incorrect}} - 50\%_{\text{FMT}}_{\text{Correct}})$. In addition, Δ RT values were calculated using a comparable

formula. These difference measures isolated the interaction of the effect feedback valence with the effect of feedback validity. Finally, three Pearson correlations were performed separately between Δ RT data and each of the EEG measures, namely, the difference RewP, the difference FMT, and the difference Beta values.

Results

Subjective rating scores

A one-way ANOVA on subjective reports of feedback reliability revealed that the reports differed across the reliability conditions, $F(2, 38) = 23.82, p < 0.001, \eta^2 = 0.56$ (Fig. 2a). Pairwise comparisons revealed that subjective reliability in the 100% condition (87%, standard error of the mean [SEM] = 3.3%) was significantly larger than subjective reliability in the 75% condition (63%, SEM = 4.6%, $t(19) = 4.9, p < 0.001$), which in turn was significantly larger than subjective reliability in the 50% condition (mean = 44%, SEM = 5.1%, $t(19) = 2.6, p < 0.02$).

Likewise, a one-way ANOVA on the participants' ratings of subjective happiness to the cue stimuli revealed a main effect of reliability, $F(2, 38) = 24.21, p < 0.001, \eta^2 = 0.56$ (Fig. 2b). Follow-up t tests demonstrated that participants felt happier when they saw the 100% reliability cue (5.75, SEM = 0.28) than when they saw the 75% reliability cue (4.2, SEM = 0.26, $t(19) = 5.82, p < 0.001$) or when they saw the 50% reliability cue (3.65, SEM = 0.2, $t(19) = 6.19, p < 0.001$). However, happiness scores were not statistically different between the 75% cue condition and 50% cue condition ($p = 0.11$).

RT results

To measure participants' behavioral adjustments, the change in RT was calculated as the absolute value of the change between trial n and trial $n+1$, separately for each feedback condition. A two-way ANOVA on these RT differences with reliability condition (100%, 75%, and 50%) and feedback valence (positive and negative) as factors revealed a main effect of valence, $F(1, 19) = 41.12, p < 0.001, \eta^2 = 0.68$, and an interaction between reliability and valence, $F(2, 38) = 24.42, p < 0.001, \eta^2 = 0.56$ (Fig. 2c). Follow-up t tests indicated effects of valence in the 100% reliability condition ($p < 0.001, \eta^2 = 0.74$), the 75% reliability condition ($p < 0.001, \eta^2 = 0.56$), and the 50% reliability condition ($p < 0.05, \eta^2 = 0.2$). The main effect of reliability was not significant, $F(1.2, 23.63) < 1, p = 0.53, \eta^2 = 0.03$ (Fig. 2c).

These absolute changes in RT suggest that participants' behavioral adjustments were more sensitive to reliable than unreliable feedback. Evidently, they adjusted their behavior

more in the 100% reliability condition, because these trials provided the most useful information for doing so. This observation aligns with the subjective rating of happiness data mentioned above, which indicated that participants preferred the 100% reliability condition.

To investigate whether subjects realized that they were too fast or too slow based on an internal representation of the response, we also examined the direction of the changes as a function of RT. For this purpose, we first divided the error trials (in which subjects received negative feedback) into fast error (mean - 1 SD) and slow error trials (mean + 1 SD) for each participant. Then, we calculated the RT adjustment by subtracting the RT on each error trial (fast or slow) from the RT on the subsequent trial and averaged these difference values separately in the three reliability conditions. These data were submitted to a two (trial type: slow or fast) by three (reliability: 100%, 75%, and 50%) repeated measures ANOVA. There was a significant main effect of trial type, $F(1, 19) = 95.12, p < 0.001, \eta^2 = 0.83$ (Figure S4). Post-hoc tests revealed that the difference value for fast error trials (232 ms) was significantly larger than that for the slow error trials (-327 ms), $p < 0.001$. The main effect of reliability did not reach significance, $F(2, 38) = 2.19, p = 0.13, \eta^2 = 0.10$, and there was no significant interaction effect between reliability and trial type, $F(2, 38) < 1, p = 0.46, \eta^2 = 0.04$. Post-hoc one-sample t tests against zero revealed that the difference value after fast error trials was positive, $t(19) = 11.82, p < 0.001$, indicating response slowing, and the difference value after slow error trials was negative, $t(19) = 6.97, p < 0.001$, indicating response speeding. In other words, participants sped up when they were too slow and slowed down when they were too fast following trials with error feedback, irrespective of the reliability of the feedback. When considered with the previous analysis (Fig. 2c), these RT results indicate that the participants adjusted their task performance based both on external feedback and on an internal representation of their response.

Surprisingly, the preceding analysis indicates that the participants' RT adjustments were sensitive to the degree of error even in the 50% reliability condition when the feedback was random. To examine this more closely, we averaged the RT data in the 50% reliability condition separately according to four subconditions: actual correct and feedback correct; actual incorrect and feedback correct; actual correct and feedback incorrect; actual incorrect and feedback incorrect. Actual correct trials were defined as trials where the participant's actual RT occurred within a 900–1,100-ms window, and actual incorrect trials were defined otherwise. We then considered the absolute change in RT following each trial type and submitted these values to a two (actual correctness: correct and incorrect) by two (feedback valence: correct and incorrect) ANOVA. This test revealed statistically significant main effects of feedback valence, $F(1, 19) = 4.48, p < 0.05, \eta^2 = 0.19$, and actual correctness, $F(1, 19) = 15.63, p = 0.001, \eta^2 = 0.45$.

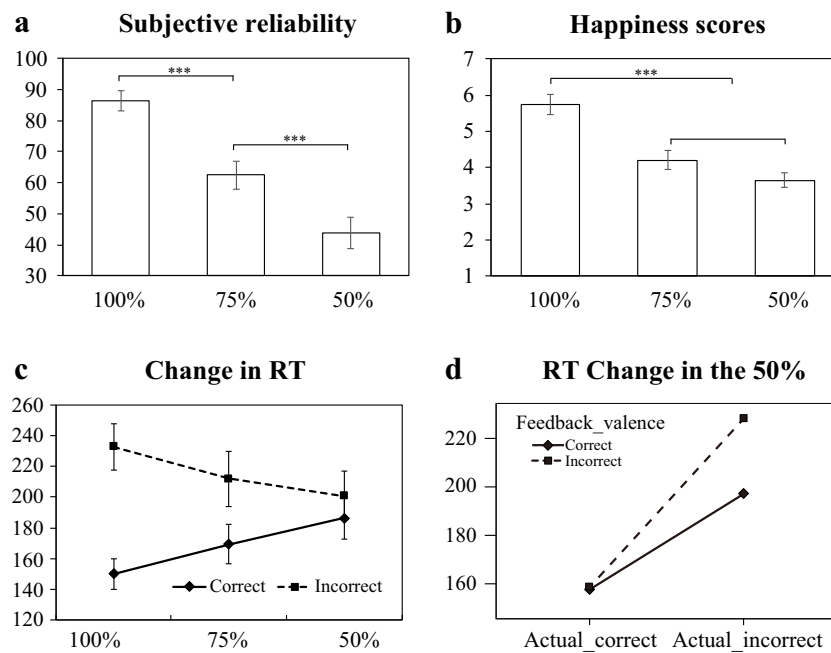


Fig. 2 Behavioral data. **a** Subjective ratings of reliability in each condition. **b** Subjective ratings of happiness in each condition. **c** Change in RT between trials n and $n+1$ as a function of perceived reliability, separately following correct (solid) and incorrect (dashed)

Furthermore, there was a significant interaction effect, $F(1, 19) = 9.63$, $p < 0.01$, $\eta^2 = 0.34$. As illustrated in Fig. 2d, participants made larger adjustments after incorrect feedback (197 ms) than after correct feedback (158 ms) but only when they were actually incorrect ($p < 0.001$), not when they were actually correct ($p = 0.93$). These findings indicate that participants adjusted their behavior more on the following trial when, based on an internal evaluation of their performance on the present trial, they determined that their response was probably incorrect rather than correct. Nevertheless, these adjustments also were largest on error trials when the feedback confirmed, rather than disconfirmed, their internal evaluation—indicating that they also considered external information about their performance, even when this feedback was entirely unreliable.

ERP results

Cue-related activities

As seen by inspection, the cue stimuli elicited a RewP-like ERP component (Fig. 3a). Paired sample t tests conducted on the amplitude of the difference waves between the 75% and 100% conditions, and between the 50% and 100% conditions, revealed that this RewP-like component in the 75% condition (mean = $-2.84 \mu\text{V}$, SEM = 0.89) was not significantly different in amplitude from that in the 50% condition (mean = $-2.71 \mu\text{V}$, SEM = 0.89), $t(19) = 0.35$, $p = 0.73$. Notably, both the RewP-like component in the 75% condition and the 50%

condition were significantly different from $0 \mu\text{V}$, $t(19) = 3.21$, $p = 0.005$, and $t(19) = 3.06$, $p < 0.01$ (Fig. 3b). The scalp distributions showed that the observed RewP-like component peaked over a frontocentral area of the scalp (Fig. 3c).

Feedback-related activities

Feedback-related ERPs are presented in Fig. 4a. A one-way repeated measures ANOVA on RewP amplitude elicited by the feedback stimuli across the 100% (mean = $-4 \mu\text{V}$, SEM = 0.78), 75% (mean = $-3.95 \mu\text{V}$, SEM = 0.77) and 50% (mean = $-3.41 \mu\text{V}$, SEM = 1.01) conditions did not reveal a statistically significant main effect, $F(2, 38) < 1$, $p = 0.86$, $\eta^2 = 0.01$ (Fig. 4b).¹ For all of three reliability conditions, the RewP was distributed over a frontocentral area of the scalp (Fig. 4c).

Given that we found an interaction between feedback valence and actual correctness on RT adjustments in the 50% reliability condition (Fig. 2d)—indicating that participants evaluated their performance not only on the basis of external feedback but also on an internal representation of their performance—we also examined whether RewP amplitude was modulated by the consistency of the feedback with their internal representations in the 50% reliability condition. However, there was no statistically significant difference in the size of

¹ Note that the RewP effect was robust to different time window utilized to calculate its amplitude (i.e., 270–300ms and 250–300ms) and to alternative measurement approaches (i.e., peak detection).

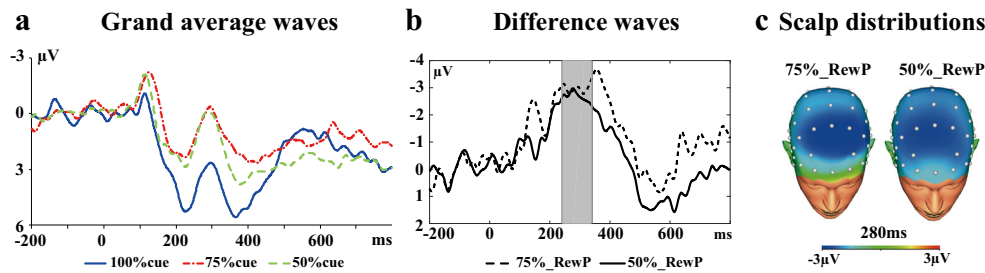


Fig. 3 Cue-related ERP results. **a** Grand average ERPs at channel Fz elicited by the cue stimuli for the 100% (blue solid line), 75% (red dotted line), and 50% (green dash line) conditions. **b** Difference waves between the ERPs at channel Fz in 100% cue condition and 75% cue condition (dash line) and between the ERPs in 100% cue condition and

50% cue condition (solid line). The time window (240–340 ms) for the RewP-like ERP component is shaded in grey. **c** The scalp distributions of two difference waves at the time of peak amplitude: 75% minus 100% (left) and 50% minus 100% (right).

these signals between expectancy conditions (please see the SOM).

Time frequency results

Cue-related FMT

A one-way ANOVA on FMT power with reliability condition (100%, 75%, and 50%) as the factor failed to reveal a main effect of reliability, $F(2, 38) = 1.65, p = 0.21, \eta^2 = 0.08$ (Fig. 5).

Feedback-related FMT Power

A two-way repeated measure ANOVA on FMT power with reliability (100%, 75%, and 50%) and valence

(correct and incorrect) as independent variables revealed a main effect of reliability, $F(2, 38) = 10.67, p < 0.001, \eta^2 = 0.36$, a main effect of valence, $F(1, 19) = 19.98, p < 0.001, \eta^2 = 0.51$, and an interaction between reliability and valence, $F(2, 38) = 6.5, p < 0.005, \eta^2 = 0.26$ (Figs. 6 and 7a). Post-hoc tests revealed the reliability effect only for incorrect feedback, $F(2, 18) = 10.62, p = 0.001, \eta^2 = 0.54$, but not for correct feedback, $F(2, 18) < 1, p = 0.44, \eta^2 = 0.09$. Post-hoc *t* tests indicated that incorrect feedback elicited greater FMT power in the 100% reliability condition than that in the 75% reliability condition ($p < 0.03$), which was significantly larger than FMT power in the 50% reliability condition ($p < 0.01$).

Because we hypothesized that RewP amplitude and FMT power reflect different ACC functions, we predicted

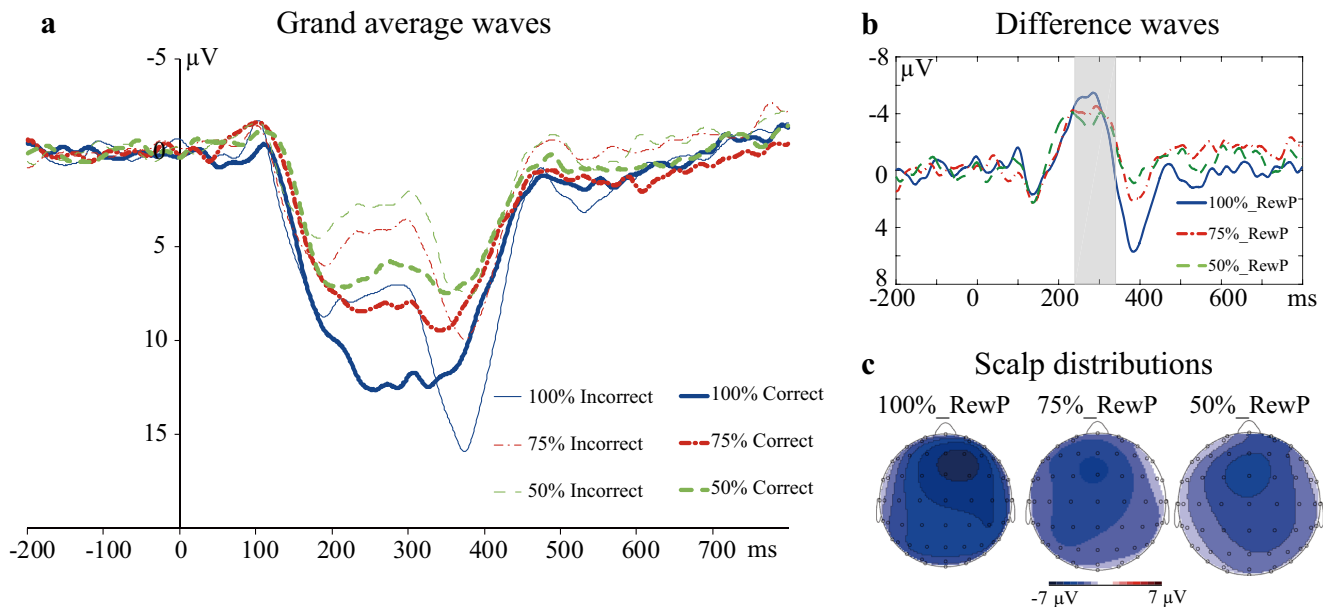


Fig. 4 Feedback-related ERP results. **a** Grand average ERPs at channel Fz for the six conditions: 100% incorrect (thin, green solid line), 100% correct (thick, green solid line), 75% incorrect (thin, red dotted line), 75% correct (thick, red dotted line), 50% incorrect (thin, blue dash line), and 50% correct (thick, blue dash line). **b** Difference waves between incorrect

and correct feedback, separately for the 100% (blue solid line), 75% (red dotted line), and 50% (green dash line) conditions. The time window (240–340 ms) for the RewP is shaded in grey. **c** The scalp distributions of the RewP for the 100%, 75%, and 50% conditions.

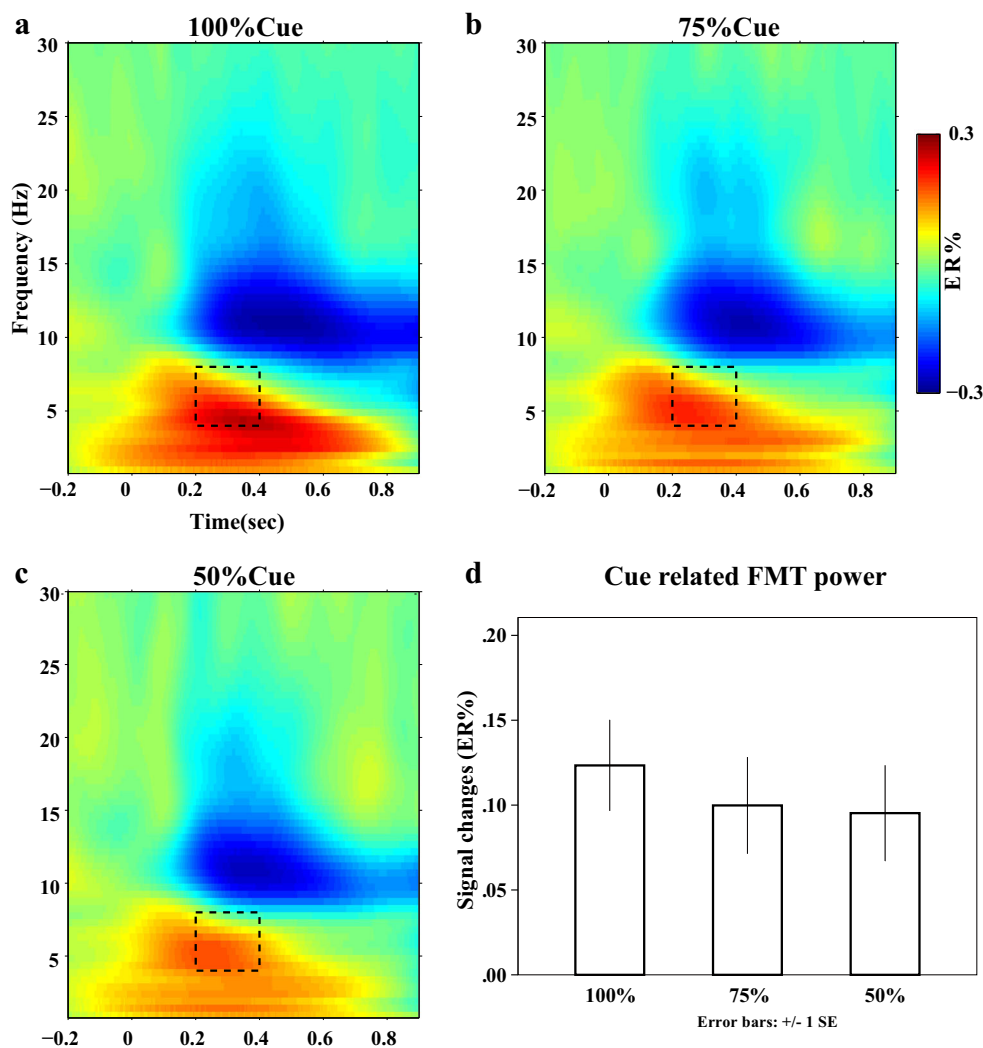


Fig. 5 Cue-related time-frequency results. Time-frequency plots of EEG data recorded at channel Fz elicited by the cues for the 100% (a), 75% (b), and 50% (c) reliability conditions. a–c The black boxes with dashed lines

indicate the time window and frequencies of interest for FMT. d FMT power within 200–400 ms for the three reliability conditions. The error bars indicate standard error mean.

that they would show different patterns of activity as a function of feedback reliability and valence. To compare these two signals directly, we first calculated the difference values between the incorrect feedback condition and the correct feedback condition separately for the three reliability conditions, and then calculated the Z-scores of FMT power and RewP amplitude for the three reliability conditions for each participant. A two-way ANOVA with reliability (100%, 75%, and 50%) and measurement type (FMT and RewP) as within-subjects factors revealed no significant main effect of measurement type, $F(1, 19) < 1$, $p = 1.0$, $\eta^2 = 0.0$, and no significant main effect of feedback reliability, $F(2, 38) = 2.75$, $p = 0.10$, $\eta^2 = 0.13$. However, the interaction between reliability and measurement type was marginally significant, $F(1.5, 28.2) = 3.57$, $p = 0.054$, $\eta^2 = 0.16$ (Fig. 7b). To explore this interaction further, a post-hoc analysis suggested that

the normalized difference in FMT power between incorrect and correct trials in the 100% reliability condition was significantly larger than that in the 50% reliability condition ($p < 0.001$). No statistically significant differences were found with respect to the normalized RewP amplitudes between the 100% and 50% reliability condition ($p = 0.88$). These results suggest qualitatively different behaviors between the RewP and FMT in this task.

Feedback-related Beta power

A comparable, three-way ANOVA on Beta power that also included channel as a factor (F3 and F4) demonstrated a main effect of reliability, $F(1.2, 23.2) = 18.82$, $p < 0.001$, $\eta^2 = 0.50$, a main effect of valence, $F(1, 19) = 6.57$, $p < 0.02$, $\eta^2 = 0.26$, and an interaction between valence and reliability, $F(1.5, 29) = 4.04$, $p < 0.02$, $\eta^2 = 0.18$ (Fig. 7c). Post-hoc t tests revealed

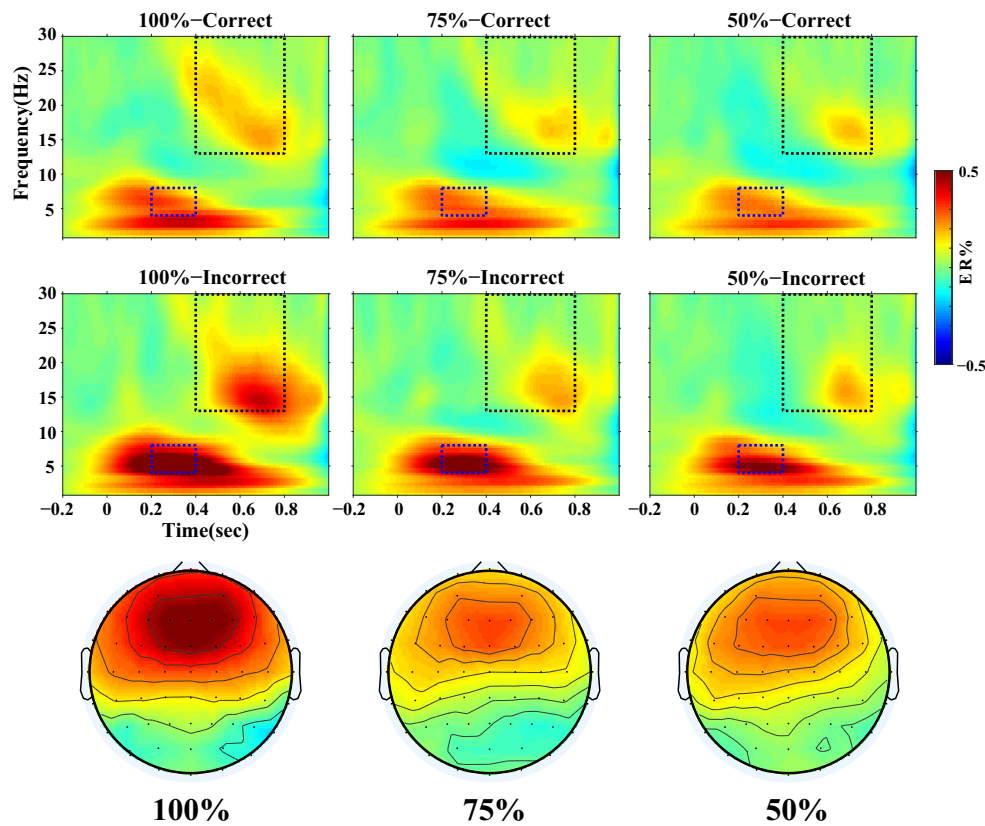


Fig. 6 Feedback-related time-frequency results. The time-frequency plots at channel Fz associated with feedback presentation for six conditions of reliability and valence (upper and middle panel) and the scalp distributions of the difference in FMT power between incorrect feedback and

correct feedback for the three reliability conditions (lower panel). The black and blue boxes in the upper and middle panels represent the time window and frequencies of interest for Beta and FMT, respectively.

an effect of valence in the 100% reliable condition ($p < 0.005$) but not in the 75% reliable condition ($p = 0.29$) nor the 50% reliable condition ($p = 0.70$). There was no main effect of electrode and no interaction between electrode and other factors: all $F < 1$. However, the three-way interaction effect

among electrode, valence, and reliability was significant, $F(2, 38) = 3.68, p < 0.05, \eta^2 = 0.16$ (Fig. 8a). Post-hoc t tests revealed an effect of valence in the 100% reliable condition at channel F3 ($p < 0.005$) and at channel F4 ($p < 0.05$); no other significant results were observed.

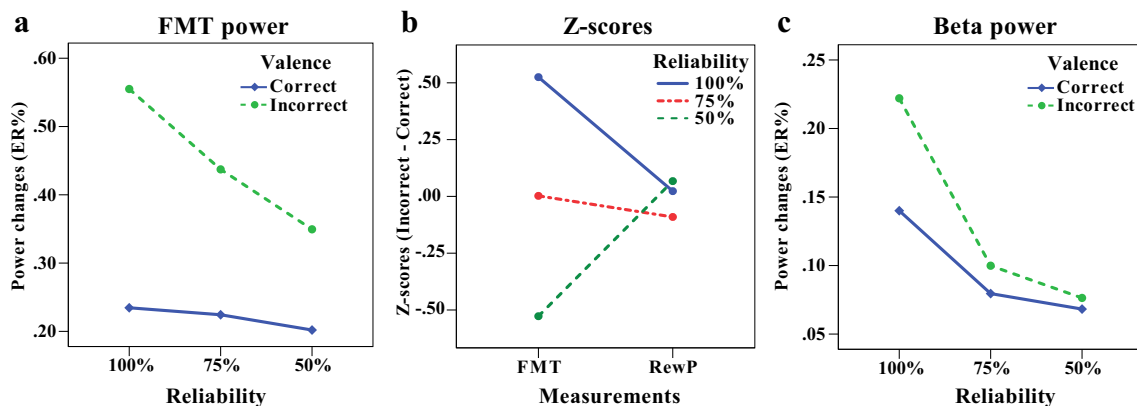


Fig. 7 Mean power in the FMT and Beta bands and Z-scores for FMT power and RewP amplitude. **a** FMT power as a function of feedback valence and reliability. **b** The interaction effect between measurement (FMT power and RewP amplitude) and reliability condition (100%, 75%, and 50%) on normalized difference values. Z-scores were calculated within-subjects across the three reliability conditions on 1) the

difference in FMT power elicited by correct feedback and incorrect feedback and 2) RewP amplitude. **c** Beta power as a function of feedback valence and reliability. **a** and **c** X-axis indicates reliability condition. Y-axis indicates power relative to baseline (ER%). Green dash lines indicate incorrect trials. Blue solid lines indicate correct trials.

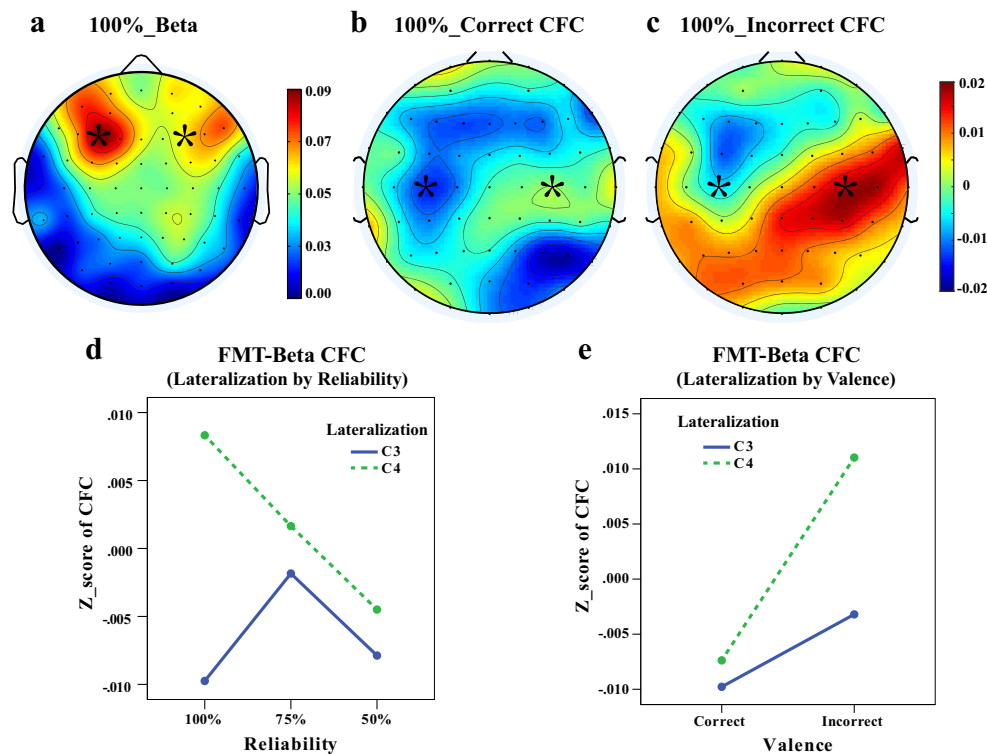


Fig. 8 Topography of Beta power and CFC results between FMT phase and Beta power. Scalp distributions of Beta power between incorrect feedback and correct feedback in the 100% reliability condition (**a**); black asterisks indicate channel F3 (left) and F4 (right). Topographies of CFC results between FMT and Beta following correct feedback (**b**)

and incorrect feedback (**c**); black asterisks indicate channel C3 (left) and C4 (right). (**d**) FMT-Beta CFC as a function of laterality and reliability condition (left panel) and (**e**) laterality and feedback valence (right panel). Green dash lines indicate CFC results at channel C4 and blue solid lines indicate CFC results at channel C3.

FMT-Beta cross-frequency coupling (CFC) following feedback

To explore functional connectivity between FMT and Beta power, CFC was computed between the phase angle of FMT at channel Fz and Beta power across the entire scalp during 0–800-ms period following feedback onset. Inspection of Fig. 8b and c suggested different CFC patterns over left and right motor areas, as predicted. To test this result, a three-way ANOVA on these CFC values with lateralization (C3, C4), reliability (100%, 75%, 50%), and valence (correct, incorrect) as factors revealed a main effect of valence, $F(1, 19) = 10.92$, $p < 0.005$, $\eta^2 = 0.37$, indicating that the CFC value for the incorrect feedback (0.004) was significantly larger than that for the correct feedback (−0.009). Furthermore, the main effect of lateralization, $F(1, 19) = 5.22$, $p < 0.04$, $\eta^2 = 0.22$, and the interaction between reliability and lateralization (Fig. 8d), $F(2, 38) = 3.64$, $p < 0.04$, $\eta^2 = 0.16$, were statistically significant; paired t tests revealed that the CFC value at C4 was significantly larger than that at C3 in the 100% condition ($p < 0.005$) but not in the 75% ($p = 0.44$) and 50% conditions ($p = 0.56$). Additionally, there was a marginally significant interaction between valence and lateralization (Fig. 8e), $F(1, 19) = 4.11$, $p < 0.06$, $\eta^2 = 0.18$; further tests showed that the CFC

value at C4 was significantly larger than that at C3 only following incorrect ($p < 0.005$) but not following correct feedback ($p = 0.64$).

Correlation analysis

The difference in difference values (see methods) for FMT power was significantly correlated with the change in response times following these trials, $r = 0.47$, $p < 0.04$, indicating that greater differences in FMT power between incorrect feedback and correct feedback in the 100% reliability condition compared with the 50% condition were associated, across participants, with larger behavioral adjustments following these trials (Fig. 9, left panel). In contrast, the difference in difference values in RewP amplitude was not related to the change in response times, $r = -0.17$, $p = 0.49$ (Fig. 9, right panel); a comparison of these slopes using the R package function `cocor` (Diedenhofen & Musch, 2015), which implements Hittner et al.'s Z procedure, confirmed that these two correlations were significantly different from each other ($p < 0.05$, Hittner, May, & Silver, 2003). The difference in difference values in Beta power was not related to the change in response times, $r = 0.2$, $p = 0.39$.

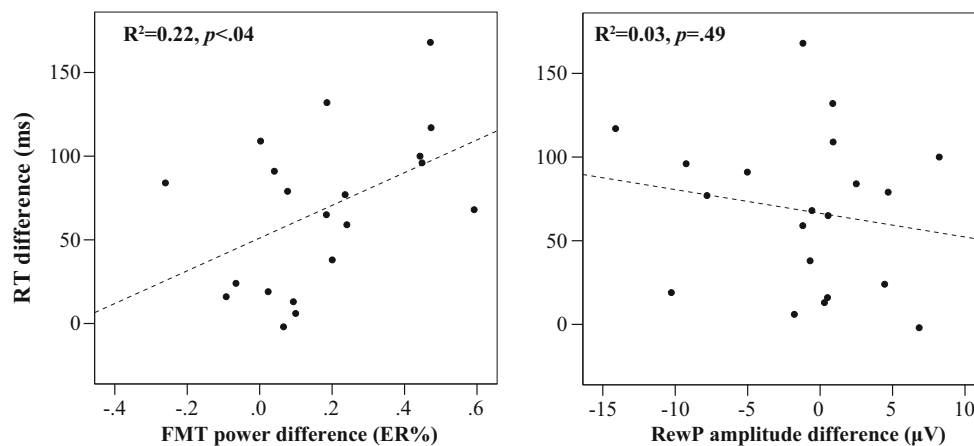


Fig. 9 Scatter plots of the correlation between electrophysiological measures and trial-to-trial change in response times Δ RT. Left panel: FMT power differences. Right panel: RewP amplitude differences (see

Methods). Differences in FMT power were significantly correlated with Δ RT data, $r = 0.47$, $p < 0.04$, whereas differences in RewP amplitude were not significantly correlated with Δ RT data, $r = -0.17$, $p = 0.49$.

Finally, given arguments that FMT power reflects the reward positivity in the frequency domain (Bernat et al., 2008; Bernat et al., 2011; Cohen et al., 2011), we also correlated the difference in differences for FMT power against the difference in differences in RewP amplitude across subjects. As expected, the two measures were uncorrelated ($p > 0.05$).

Discussion

The current EEG study investigated the involvement of ACC in regulating behavior in response to unreliable feedback. We utilized three EEG signals related to feedback processing by ACC for this purpose: FMT as a putative measure of cognitive control, the RewP as an index reward processing, and Beta and FMT-Beta CFC as indices of binding between the first two processes. We adopted a modified time-estimation task in which the reliability of the feedback stimuli was systematically varied across conditions from 100%, 75%, to 50%, as indicated by a cue presented to the participants after their response on each trial. Participants' subjective ratings of reliability decreased with reliability across the conditions, indicating that the manipulation was successful. The RT data showed that participants made larger behavioral adjustments when feedback was more reliable than when feedback was less reliable, suggesting that feedback reliability influenced their reinforcement learning process. Nevertheless, they continued to rely somewhat on the feedback even when it was completely random, especially when error feedback agreed with their own internal representation of response correctness. Accordingly, we found that the 100% reliable cue elicited larger RewP than the 75% and 50% reliable cues did, which is in line with participants' subjective rating of happiness. Moreover, an interaction effect between reliability and feedback valence was found for both feedback-elicited FMT and Beta power, suggesting that reliability modulated the psychological processes

indexed by these two EEG signals. Furthermore, we found that coupling of FMT phase with Beta power was associated with behavioral adjustments on the following trial.

The relationship between the RewP and FMT, and whether it reflects a signed or unsigned RPE signal, has been the subject of some controversy (Cavanagh, Zambrano-Vazquez, & Allen, 2012; Holroyd, HajiHosseini, & Baker, 2012). Because FMT power has been previously said to be related to RPEs (Cohen, Elger, & Ranganath, 2007; Marco-Pallarés et al., 2008), one might expect to observe similar patterns of FMT and RewP in this experiment. However, recent evidence suggests that these phenomena are dissociable. On this account, unexpected, task-related events elicit an increase in FMT power, with the portion of FMT oscillations that is coherent in phase across trials appearing in the ERP as a series of positive and negative deflections, including a negative-going ERP component at approximately 250 ms post-feedback called the N200 (Holroyd et al., 2008). Like FMT, the N200 is elicited by unexpected, task-relevant stimuli irrespective of whether the task involves reinforcement learning (Holroyd, 2004). In line with the dopamine theory of this component, negative RPE signals conveyed to ACC enhance N200 amplitude following unexpected error feedback (Warren & Holroyd, 2012), whereas positive RPE signals suppress N200 amplitude (Baker & Holroyd, 2011; Holroyd et al., 2008; Holroyd, Krigolson, & Lee, 2011; Holroyd & Umemoto, 2016).

Consistent with this inference, we observed clear dissociations between FMT power and RewP amplitude. In particular, FMT power was not correlated with RewP amplitude. Consistent with this, increased FMT power following feedback presentation was associated with greater feedback reliability, but the same was not true of RewP amplitude, which was about the same amplitude across conditions (Figs. 4b and 7a). In addition, FMT power, but not RewP amplitude, was associated with feedback-related adjustments to behavior. Similar discrepant patterns between

RewP amplitude and FMT power have been reported in our prior studies (Hajihosseini & Holroyd, 2013; Li et al., 2016). For example, we previously found that neutral feedback elicited larger RewP amplitudes than did negative feedback, whereas no significant difference in FMT power was observed between these two conditions (Li et al., 2016). This discrepancy may stem from the fact that the RewP is a phasic signal, not an oscillation. Because phasic EEG peaks contain energy in a wide range of frequency bands (Yeung, Bogacz, Holroyd, Nieuwenhuis, & Cohen, 2007), the ERP technique may be better able than wavelet- and Fourier-based methods to discriminate the RewP from noise.

These observations support the argument that the RewP and FMT power index distinct cognitive phenomena (Holroyd et al., 2012; Holroyd & Umemoto, 2016; Hajihosseini & Holroyd, 2013), namely, that the RewP reflects RPE signals communicated to ACC, whereas FMT power reflects the role of ACC in sustaining effortful control over the task at hand (Holroyd & Umemoto, 2016). Multiple lines of evidence indicate that FMT reflects a signal produced by ACC for cognitive control (Cavanagh & Frank, 2014; Cavanagh & Shackman, 2015; Holroyd, 2016; Holroyd & Umemoto, 2016). For example, converging findings suggest that the FMT oscillations play an essential role in feedback learning and behavioral adaptation (Cavanagh & Shackman, 2015; Hajihosseini & Holroyd, 2013). More specifically, FMT oscillations have been associated with action adjustments across various tasks, including probabilistic learning (Cavanagh, 2010), time-estimation (Irene, Ridderinkhof, & Cohen, 2011), and learning efficiency (Luft et al., 2013). In the present study, we found that FMT power linearly decreased from the most reliable feedback condition to the least reliable feedback condition. The interpretation of this finding is straightforward: the control level was maximized on the trials when it was the most useful and minimized otherwise.

By contrast, RewP amplitude was not correlated with behavioral adjustments in the present study. This lack of correspondence between RewP amplitude and trial-to-trial changes in behavior echoes many similar findings from previous studies. Whereas several studies have reported that the RewP amplitude predicts future behavior (Cohen & Ranganath, 2007; Holroyd & Krigolson, 2007), several other RewP studies have not (Chase, Swainson, Durham, Benham, Cools, 2011; Luft et al., 2013; see Walsh & Anderson, 2012; Holroyd & Umemoto, 2016 for reviews). For example, Walsh and Anderson (2011) observed that instructions about reward probability in a learning task dramatically changed participant behavior but did not affect RewP amplitude, revealing a clear dissociation between RewP amplitude and behavior. Although the reinforcement learning theory of the RewP holds that this ERP component indexes the use by ACC of a dopamine-dependent RPE signal for modulating future behavior (Holroyd & Coles, 2002), this modulation was conceived

as regulating a high-level decision system over action policies, not over the actions themselves (Holroyd & Coles, 2002). More recently, the HRL-ACC theory developed this account by proposing that the RewP reflects motivational factors related to task performance (Holroyd & Yeung, 2012; Holroyd & Umemoto, 2016), whereas FMT power indexes the cognitive control signal applied by ACC over other action-production systems (Holroyd & Umemoto, 2016; see also Holroyd, 2016). On this view, behavioral changes should be better predicted by FMT power than by RewP amplitude, as we observed.

We also found that the three reliability conditions elicited a robust RewP but that, in contrast to FMT power, RewP amplitude did not differ in size across the three conditions. It is particularly surprising that the 50% condition generated a RewP, given that the participants were aware that the feedback types in the 50% reliable condition occurred totally at random. Although further analyses on RT in the 50% condition suggested that participants adjusted their behaviors based on both internal and external performance information, RewP amplitude was not detectably different for the feedback that agreed with the internal representations than for the feedback that did not. We see two possibilities underlying this result. First, the RPE may have been taken up by the cue, preventing modulations of RewP amplitude to the feedback. Consistent with this possibility, the scalp distributions and timing of the difference waves between the 100% condition and other two conditions suggest that the cue might have elicited a RewP. Furthermore, the subjective rating scores indicated that participants preferred the 100% reliable feedback more than the other two conditions, suggesting that the 100% reliable cue may have been rewarding to them. This observation replicates a finding in our previous study, in which we observed that a predictive cue that indicated whether feedback on that trial would be available or also not carried reward information (Wang et al., 2016). The finding is in line with a monkey study in which midbrain dopamine neurons signaled both the animal's preference for an advance cue, indicating the size of a forthcoming reward, as well as the standard RPE signal to the reward itself (Bromberg-Martin & Hikosaka, 2009). That said, a strict computational account of temporal difference learning would not predict that RPE signals should be elicited by cues that do not predict the subsequent outcome (Sutton & Barto, 1998).

Second, the RewP might reflect an obligatory response to the feedback valence that is relatively insensitive to the veridicality of the outcomes. This possibility is supported by the previous study that revealed a dissociation between the effects of instruction on RewP amplitude and behavior (Walsh & Anderson, 2011). Combined EEG and fMRI studies have demonstrated that the RewP amplitude is strongly related to in ventral striatal activation (Carlson, Foti, Mujica-Parodi, Harmon-Jones, & Hajcak, 2011; Nieuwenhuis, Slagter, Alting von Geusau, Heslenfeld, & Holroyd, 2005a),

suggesting that RewP amplitude might index “lower level” reward (Holroyd & Umemoto, 2016; Shahnazian & Holroyd, 2016). These speculations call for further examination in future studies.

Additionally, Beta power has been associated with feedback learning (Marco-Pallarés, Münte, & Rodríguez-Fornells, 2015). An influential hypothesis holds that Beta oscillations desynchronize when behavioral change is needed, encoding a so called status quo signal (Engel & Fries, 2010; Luft et al., 2013). Accordingly, in one of our previous studies, larger Beta power was observed following reward feedback compared with nonreward feedback, the source of which was localized to DLPFC (HajiHosseini & Holroyd, 2015a). In a following study, it was shown that Beta power is enhanced following incorrect feedback relative to correct feedback when task instructions emphasize learning from errors (HajiHosseini & Holroyd, 2015b). The current study replicated the scalp distribution of Beta power, which peaked at the same channels near DLPFC as in the previous studies (F3 & F4; Fig. 8a). Moreover, Beta power was enhanced following negative feedback relative to positive feedback, but only when the feedback was reliable. These findings suggest that Beta power might reflect an aspect of RL that is unrelated to feedback valence, because Beta was elicited both by positive and negative feedback (Engel & Fries, 2010; Luft et al., 2013; Li et al., 2016; HajiHosseini & Holroyd, 2015b). We speculated that Beta activity may reflect increasing levels of cognitive effort associated with attention, as suggested by previous studies (Buschman, Denovellis, Diogo, Bullock, & Miller, 2012; Buschman & Miller, 2007; Pesaran et al., 2008).

Lastly, neural synchrony across distributed neural networks plays a key role in communicating information between different neural groups (Canolty & Knight, 2010; Siegel, Donner, & Engel, 2012). In particular, FMT generated in ACC appears to regulate binding of information between different brain regions, a process that requires effortful control (Holroyd, 2016). This position is supported by recent computational modelling work that illustrates how CFC across wide brain regions between FMT and gamma oscillations contributes to task performance (Verguts, 2017). In the current study, an exploratory CFC analysis revealed coupling between ACC FMT phase and Beta power over central motor areas (channels C3 & C4) in two feedback conditions, but only when these feedback stimuli were 100% reliable. This observation is in line with previous studies that indicated that FMT couples with Beta oscillations, as observed both with intracranial (Axmacher, Henseler, Jensen, Weinreich, Elger, & Fell, 2010) and scalp (Cohen, Elger, & Fell, 2009) recordings, suggesting that FMT-Beta coupling may reflect prefrontal control over sensorimotor processing (Siegel, Donner, & Engel, 2012). The HRL-ACC theory proposes that a low-level actor module (including DLPFC and the dorsal striatum) is responsible for implementing action policies, e.g., guiding the motor

area to carry out movement (Holroyd & Yeung, 2012; Holroyd & Umemoto, 2016). Our present findings demonstrated that CFC between FMT and Beta may reflect this guidance function from actor module to motor areas. Although the precise function of Beta and Beta-FMT CFC remains unknown, these results suggest how different EEG signals can be utilized to investigate the different modules in the HRL-ACC framework.

Conclusions

Manipulating feedback reliability had distinct effects on different measures of ACC function. In particular, unreliable feedback reduced the impact of FMT power on behavioral adaptations and diminished an effect of FMT phase-amplitude coupling with Beta oscillations over motor cortex. By contrast, feedback reliability had no impact on RewP amplitude. These findings support a role for ACC in the hierarchical control of behavior.

Acknowledgements This study was supported by the National Natural Science Foundation of China (31671158 & 31671150), the (Key) Project of DEGP (2015WTSCX094), Shenzhen Peacock Plan (grant no. KQTD2015033016104926), and the youth Project of Humanities and Social Sciences of Shenzhen University (16QNFC51).

References

- Altamura, M., Goldberg, T. E., Elvevåg, B., Holroyd, T., Carver, F. W., & Weinberger, D. R., et al. (2010). Prefrontal cortex modulation during anticipation of working memory demands as revealed by magnetoencephalography. *Journal of Biomedical Imaging*, 2010(10), 12.
- Axmacher, N., Henseler, M. M., Jensen, O., Weinreich, I., Elger, C. E., & Fell, J. (2010). Cross-frequency coupling supports multi-item working memory in the human hippocampus. *Proceedings of the National Academy of Sciences*, 107(7), 3228–3233.
- Baker, T. E. and C. B. Holroyd (2011). "Dissociated roles of the anterior cingulate cortex in reward and conflict processing as revealed by the feedback error-related negativity and N200." *Biological Psychology* 87(1): 25–34.
- Behrens, T. E., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. (2007). Learning the value of information in an uncertain world. *Nature neuroscience*, 10(9), 1214–1221.
- Bernat, E. M., Nelson, L. D., Steele, V. R., Gehring, W. J., & Patrick, C. J. (2011). Externalizing psychopathology and gain–loss feedback in a simulated gambling task: Dissociable components of brain response revealed by time-frequency analysis. *Journal of Abnormal Psychology*, 120(2), 352–364.
- Bernat, E.M., Nelson, L.D., Holroyd, C.B., Gehring, W.J., and Patrick, C.J. (2008). Separating cognitive processes with principal components analysis of EEG time-frequency distributions. *Proceedings of the Society of Photo-Optical Instrumentation Engineers*, Vol. 7074, 70740S.
- Botvinick, M. M. (2007). Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function, *Cognitive Affective, & Behavioral Neuroscience*, 7(4), 356–366.

- Bromberg-Martin, E. S., and Hikosaka, O. (2009). Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron*, 63, 119–126.
- Bruns, A., & Eckhom, R. (2004). Task-related coupling from high- to low-frequency signals among visual cortical areas in human subdural recordings. *International Journal of Psychophysiology*, 51(2), 97.
- Buschman, T. J., Denovellis, E. L., Diogo, C., Bullock, D., & Miller, E. K. (2012). Synchronous oscillatory neural ensembles for rules in the prefrontal cortex. *Neuron*, 76(4), 838–846.
- Buschman, T.J., and Miller, E.K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science*, 315, 1860–1862.
- Canolty R. T. & Knight, R. T. (2010). The functional role of cross-frequency coupling. *Trends in Cognitive Sciences*, 14(11), 506–515.
- Carlson, J. M., Foti, D., Mujica-Parodi, L. R., Harmon-Jones, E., & Hajcak, G. (2011). Ventral striatal and medial prefrontal BOLD activation is correlated with reward-related electrocortical activity: a combined ERP and fMRI study. *Neuroimage*, 57(4), 1608–1616.
- Carter, C. S., & Van Veen, V. (2007). Anterior cingulate cortex and conflict detection: an update of theory and data. *Cognitive Affective, & Behavioral Neuroscience*, 7(4), 367–379.
- Cavanagh, J. F., & Shackman, A. J. (2015). Frontal midline theta reflects anxiety and cognitive control: meta-analytic evidence. *Journal of Physiology-Paris*, 109(1–3), 3–15.
- Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, 18(8), 414–421.
- Cavanagh, J. F., Frank, M. J., Klein, T. J., & Allen, J. J. (2010). Frontal theta links prediction errors to behavioral adaptation in reinforcement learning. *Neuroimage*, 49(4), 3198–3209.
- Cavanagh, J. F., Zambrano-Vazquez, L., & Allen, J. J. (2012). Theta lingua franca: A common mid-frontal substrate for action monitoring processes. *Psychophysiology*, 49(2), 220–238.
- Chase, H. W., Swainson, R., Durham, L., Benham, L., & Cools, R. (2011). Feedback-related negativity codes prediction error but not behavioral adjustment during probabilistic reversal learning. *Journal of Cognitive Neuroscience*, 23(4), 936–946.
- Cohen, M. X., Elger, C. E., & Fell, J. (2009). Oscillatory activity and phase–amplitude coupling in the human medial frontal cortex during decision making. *Journal of cognitive neuroscience*, 21(2), 390–402.
- Cohen, M. X., Elger, C. E., & Ranganath, C. (2007). Reward expectation modulates feedback-related negativity and eeg spectra. *Neuroimage*, 35(2), 968–978.
- Cohen, M. X., & Ranganath, C. (2007). Reinforcement learning signals predict future decisions. *Journal of Neuroscience*, 27(2), 371.
- Cohen, M. X., Wilmes, K. A., & Vijver, I. V. D. (2011). Cortical electrophysiological network dynamics of feedback learning. *Trends in Cognitive Sciences*, 15(12), 558–566.
- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*, 134(1), 9–21.
- Diedenhofen, B. & Musch, J. (2015). cocor: A comprehensive solution for the statistical comparison of correlations. *PLoS ONE*, 10(4): e0121945.
- Engel, A. K., & Fries, P. (2010). Beta-band oscillations—signalling the status quo?. *Current opinion in neurobiology*, 20(2), 156–165.
- Ernst, B., & Steinhauser, M. (2015). Effects of invalid feedback on learning and feedback-related brain activity in decision-making. *Brain and cognition*, 99, 78–86.
- Ernst, B., & Steinhauser, M. (2017). Top-down control over feedback processing: the probability of valid feedback affects feedback-related brain activity. *Brain & Cognition*, 115, 33.
- Hsieh, L. T., & Ranganath, C. (2013). Frontal midline theta oscillations during working memory maintenance and episodic encoding and retrieval. *Neuroimage*, 85(2), 721–729.
- HajiHosseini, A., & Holroyd, C. B. (2013). Frontal midline theta and N200 amplitude reflect complementary information about expectancy and outcome evaluation. *Psychophysiology*, 50(6), 550–562.
- HajiHosseini, A., & Holroyd, C. B. (2015a). Reward feedback stimuli elicit high-beta eeg oscillations in human dorsolateral prefrontal cortex. *Scientific Reports*, 5, 13021.
- HajiHosseini, A., & Holroyd, C. B. (2015b). Frontal beta oscillations reflect encoding of information related to desired task performance irrespective of feedback valence. Program No. 352.24. 2015 Neuroscience Meeting Planner. Washington, DC: Society for Neuroscience, 2015. Online
- HajiHosseini, A., Rodríguez-Fornells, A., Marco-Pallarés, J., 2012. The role of beta-gamma oscillations in unexpected rewards processing. *Neuroimage*, 60, 1678–1685.
- Hittner, J. B., May, K., & Silver, N. C. (2003). A Monte Carlo evaluation of tests for comparing dependent correlations. *The Journal of general psychology*, 130(2), 149–168.
- Holroyd, C. B. (2004). A note on the oddball N200 and the feedback ERN. *Neurophysiology*, 78, 447–455.
- Holroyd, C. B., (2016). The waste disposal problem of effortful control. In: Braver, T. (Ed.), Motivation and cognitive control. Psychology Press, New York, NY, pp. 235–260.
- Holroyd, C. B., HajiHosseini, A., & Baker, T. E. (2012). ERPs and EEG oscillations, Best friends forever: comment on Cohen et al. *Trends in Cognitive Sciences*, 16, 192.
- Holroyd, C. B. & Krigolson O. E. (2007). Reward prediction error signals associated with a modified time estimation task. *Psychophysiology*, 44(6): 913–917.
- Holroyd, C. B. and M. G. H. Coles (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, 109, 679–709.
- Holroyd, C. B., Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-related positivity: sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology*, 45(5), 688–697.
- Holroyd, C. B., & McClure, S. M. (2015). Hierarchical control over effortful behavior by rodent medial frontal cortex: A computational model. *Psychological review*, 122(1), 54.
- Holroyd, C. B., Nieuwenhuis, S., Yeung, N., Nystrom, L., Mars, R. B., Coles, M. G., & Cohen, J. D. (2004). Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nature Neuroscience*, 7(5), 497.
- Holroyd, C. B., & Umemoto, A. (2016). The research domain criteria framework: The case for anterior cingulate cortex. *Neuroscience & Biobehavioral Reviews*, 71, 418–443.
- Holroyd, C. B., & Yeung, N. (2012). Motivation of extended behaviors by anterior cingulate cortex. *Trends in Cognitive Sciences*, 16(2), 122–128.
- Irene, V. D. V., Ridderinkhof, K. R., & Cohen, M. X. (2011). Frontal oscillatory dynamics predict feedback learning and action adjustment. *Journal of Cognitive Neuroscience*, 23(12), 4106–4121.
- Johnston, K., Levin, H. M., Koval, M. J., & Everling, S. (2007). Top-down control-signal dynamics in anterior cingulate and prefrontal cortex neurons following task switching. *Neuron*, 53(3), 453–462.
- Karlsson, M. P., Tervo, D. G., Karpova, A. Y. (2012). Network resets in medial prefrontal cortex mark the onset of behavioral uncertainty. *Science*, 338(6103):135–139.
- Khamassi, M., Lallée, S., Enel, P., Procyk, E., & Dominey, P. F. (2011). Robot cognitive control with a neurophysiologically inspired reinforcement learning model. *Frontiers in neurobotics*, 5.
- Li, P., Jia, S., Feng, T., Liu, Q., Suo, T., & Li, H. (2010). The influence of the diffusion of responsibility effect on outcome evaluations: Electrophysiological evidence from an ERP study. *Neuroimage*, 52(4), 1727–1733.
- Li, P., Baker, T. E., Warren, C., & Li, H. (2016). Oscillatory profiles of positive, negative and neutral feedback stimuli during adaptive

- decision making. *International Journal of Psychophysiology*, 107, 37–43.
- Luft, C. D. B., et al. (2013). High-Learners Present Larger Mid-Frontal Theta Power and Connectivity in Response to Incorrect Performance Feedback. *Journal of Neuroscience* 33(5): 2029–2038.
- Marco-Pallarés, J., Cucurell, D., Cunillera, T., García, R., Andrés-Pueyo, A., Münte, T. F., & Rodríguez-Fornells, A. (2008). Human oscillatory activity associated to reward processing in a gambling task. *Neuropsychologia*, 46, 241–248.
- Marco-Pallarés, J., Münte, T. F., & Rodríguez-Fornells, A. (2015). The role of high-frequency oscillatory activity in reward processing and learning. *Neuroscience & Biobehavioral Reviews*, 49, 1–7.
- Miltner, W. H. R., Braun, C. H., & Coles, M. G. H. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a “generic” neural system for error detection. *Journal of Cognitive Neuroscience*, 9, 788–798.
- Mouraux, A., Guerit, J. M., & Plaghki, L. (2003). Non-phase locked electroencephalogram (EEG) responses to CO₂ laser skin stimulations may reflect central interactions between A δ - and C-fibre afferent volleys. *Clinical neurophysiology*, 114(4), 710–722.
- Mouraux, A., & Iannetti, G. D. (2008). Across-trial averaging of event-related eeg responses and beyond. *Magnetic Resonance Imaging*, 26(7), 1041–54.
- Nieuwenhuis, S., Astonjones, G., & Cohen, J. D. (2005a). Decision making, the p3, and the locus coeruleus-norepinephrine system. *Psychological Bulletin*, 131(4), 510–32.
- Nieuwenhuis, S., Slagter, H., Alting von Geusau, N., Heslenfeld, D.J., & Holroyd, C.B. (2005b). Knowing good from bad: Differential activation of human cortical areas by positive and negative outcomes. *European Journal of Neuroscience*, 21, 3161–3168.
- Onslow, A. C., Bogacz, R., & Jones, M. W. (2011). Quantifying phase-amplitude coupling in neuronal network oscillations. *Progress in Biophysics & Molecular Biology*, 105(1–2), 49–57.
- O’Reilly, J. X., Schüffelgen, U., Cuell, S. F., Behrens, T. E., Mars, R. B., & Rushworth, M. F. (2013). Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proceedings of the National Academy of Sciences*, 110(38), E3660–E3669.
- Pesaran, B., Nelson, M.J., and Andersen, R.A. (2008). Free choice activates a decision circuit between frontal and parietal cortex. *Nature*, 453, 406–409.
- Proudfit, G. H. (2015). The reward positivity: From basic research on reward to a biomarker for depression. *Psychophysiology*, 52(4), 449–459.
- Rutishauser, U., Ross, I. B., Mamelak, A. N., & Schuman, E. M. (2010). Human memory strength is predicted by theta-frequency phase-locking of single neurons. *Nature*, 464(7290), 903–7.
- Sambrook, T. D., & Goslin, J. (2015). A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychological Bulletin*, 141(1), 213–235.
- Schiffer, A. M., Siletti, K., Waszak, F., & Yeung, N. (2017). Adaptive behaviour and feedback processing integrate experience and instruction in reinforcement learning. *NeuroImage*, 146, 626–641.
- Siegel, M., Donner, T. H., & Engel, A. K. (2012). Spectral fingerprints of large-scale neuronal interactions. *Nature Reviews Neuroscience*, 13(2), 121–134.
- Sutton, R. S., & Barto, A. G. (1998). Reinforcement learning: An introduction (Vol. 1, No. 1). Cambridge: MIT press.
- Umemoto, A., Hajihosseini, A., Yates, M. E., & Holroyd, C. B. (2017). Reward-based contextual learning supported by anterior cingulate cortex. *Cognitive Affective & Behavioral Neuroscience*, 17(3), 642.
- Verguts, T. (2017). Binding by random bursts: a computational model of cognitive control. *Journal of Cognitive Neuroscience*, 29(6), 1103–1118.
- Walsh, M. M., & Anderson, J. R. (2011). Modulation of the feedback-related negativity by instruction and experience. *Proceedings of the National Academy of Sciences of the United States of America*, 108(47), 19048–53.
- Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience & Biobehavioral Reviews*, 36(8), 1870–1884.
- Warren, C. M., & Holroyd, C. B. (2012). The impact of deliberative strategy dissociates ERP components related to conflict processing vs. reinforcement learning. *Frontiers in neuroscience*, 6.
- Warren, C. M., Hyman, J. M., Seamans, J. K., & Holroyd, C. B. (2015). Reward processing in the rodent anterior cingulate cortex. *Journal of Physiology, Paris*, 109 (1), 87–94.
- Wang, J., Chen, Z., Peng, X., Yang, T., Li, P., Cong, F., & Li, H. (2016). To know or not to know? theta and delta reflect complementary information about an advanced cue before feedback in decision-making. *Frontiers in psychology*, 7.