NeuroImage 331 (2026) 121894

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Learning from feedback is independent from feedback visibility, but supported by aperiodic neural activity

DanDan Liu [a,b], Shiwei Jia [a,b,*], Yanliang Sun [a,b], Lorenza Colzato [a,b,*], Bernhard Hommel [a,b,*]

[a] Shandong Provincial Key Laboratory of Brain Science and Mental Health, Faculty of Psychology, Shandong Normal University, Jinan 250014, China
[b] Cognitive Neurophysiology, Department of Child and Adolescent Psychiatry, Faculty of Medicine, TU Dresden, Dresden, Germany

ARTICLE INFO

Keywords:
Reinforcement learning
Invisible feedback
Continuous flash suppression paradigm (CFS)
Reward positivity (REWP)
P3a
FOOOF
Aperiodic exponent

ABSTRACT

Humans and other animals learn from feedback, by tending to repeat rewarded behavior and change behavior that receives negative feedback. Previous findings suggest that feedback does not need to be consciously perceived in order to be effective. Using continuous flash suppression, we presented participants with visible and invisible positive and negative feedback during a time estimation task while recording EEG. Behavioral results showed that both visible and invisible feedback significantly influenced time estimation error and adjustment in trial $N+1$, suggesting that subliminal reward information can be effectively utilized. Electrophysiological indicators (reward positivity, P3a, theta activity, aperiodic exponent) showed feedback-valence effects, but only when the feedback was visible. Performance in trial $N+1$ was successfully predicted by the aperiodic exponent only. These findings suggest that (1) behavioral control is independent from conscious perception of feedback signals; (2) the valence sensitivity of electrophysiological indicators is not informative for effective learning or the impact of feedback on behavioral control; and (3) the general neural state, as characterized by the aperiodic exponent, is predictive of the quality of learning from feedback, with steeper exponents providing the most supportive conditions for learning.

1. Introduction

Learning from feedback constitutes a core mechanism of adaptive behavior. Since Thorndike's seminal formulation of the law of effect (1898), it has been robustly demonstrated that actions followed by rewarding outcomes are more likely to be repeated, whereas those followed by negative outcomes tend to be suppressed. In experimental settings, this principle is commonly examined using feedback-based learning tasks in which participants select responses and receive trial-by-trial evaluative feedback indicating success or failure. Adaptive behavior in such tasks is expressed as systematic changes in response repetition and switching as a function of feedback history.

A central open question concerns the role of conscious awareness in reinforcement learning. Although classical accounts implicitly assume that feedback must be consciously perceived to guide behavior, growing evidence suggests that learning-related adjustments can occur in the absence of awareness. Subliminal reward cues have been shown to modulate motivation and performance (Bijleveld et al., 2012; Pessiglione et al., 2007; Zedelius et al., 2014), and Correa et al. (2018) demonstrated behavioral adjustment following both consciously and nonconsciously processed feedback, accompanied by corresponding modulations in event-related potentials (ERPs). These findings indicate that reinforcement learning mechanisms may operate independently of conscious access. However, methodological limitations—such as incomplete masking and (resulting) residual perceptual awareness in the Correa et al. study—leave open the question of whether such effects genuinely reflect unconscious learning.

The present study addressed this issue by employing a rigorously validated continuous flash suppression (CFS) procedure (Fang and He, 2005; Tsuchiya and Koch, 2005) to render feedback stimuli invisible while preserving their temporal and informational structure. Hypothesis 1 tested whether feedback-based learning occurs independently of awareness. We predicted that positive feedback would increase response repetition on subsequent trials, whereas negative feedback would increase response switching, reflecting adaptive learning, and that this pattern would be observed both under conscious and nonconscious

* Corresponding authors at: Faculty of Psychology, Shandong Normal University, No. 88 East Wenhua Road, Jinan 250014, Shandong Province, China.
E-mail addresses: jiashiwei82@126.com (S. Jia), colzato@bhommel.onmicrosoft.com (L. Colzato), bh@bhommel.onmicrosoft.com (B. Hommel).

https://doi.org/10.1016/j.neuroimage.2026.121894
Received 2 September 2025; Received in revised form 4 March 2026; Accepted 26 March 2026
Available online 27 March 2026

feedback conditions. Confirmation of this hypothesis would provide strong evidence that core reinforcement mechanisms operate without conscious access to feedback.

Beyond behavior, feedback processing engages a well-characterized set of neural signals. One of the most prominent is the feedback-related negativity/reward positivity (FRN/RewP), a frontocentral ERP component emerging approximately 200–300 ms after feedback onset. According to reinforcement learning theory, the FRN/RewP reflects a reward prediction error signal transmitted via dopaminergic projections to medial frontal cortex (Holroyd and Coles, 2002; Holroyd and Umemoto, 2016). The magnitude of the FRN/RewP has been shown to scale with feedback valence and expectancy and to predict subsequent behavioral adjustments (Sambrook and Goslin, 2015). A later parietal positivity, the P3, has been linked to the updating of stimulus–response associations and the integration of feedback into working memory and action selection processes (Polich, 2007; Chase et al., 2011). In the frequency domain, increases in midfrontal theta-band power following negative feedback have been consistently associated with performance monitoring, cognitive control, and learning-related adaptation (Cohen, 2014; Cavanagh and Frank, 2014).

However, traditional ERP and time–frequency analyses conflate oscillatory activity with aperiodic (1/f-like) neural activity. Recent methodological work has demonstrated that aperiodic activity reflects the balance of excitation and inhibition in cortical networks and carries functionally meaningful information distinct from oscillatory components (He, 2014; Donoghue et al., 2020). Task-related changes in the aperiodic exponent have been linked to individual differences in learning, cognitive control, and behavioral flexibility (Voytek et al., 2015; Zhang et al., 2023; Yan et al., 2024), suggesting that aperiodic activity may play a key role in shaping adaptive behavior over longer timescales.

This distinction aligns closely with the metacontrol framework, which conceptualizes cognitive control as a dynamic regulation along a continuum between persistence and flexibility (Hommel and Colzato, 2017; Hommel et al., 2024; Colzato et al., 2025). Within this framework, oscillatory feedback signals such as the FRN/RewP, P3, and theta activity are thought to reflect immediate evaluative and control processes on a trial-by-trial basis, whereas aperiodic activity may index strategic adjustments of control state as feedback accumulates over time (Jia et al., 2024, 2025).

Hypothesis 2 addressed the neural mechanisms supporting feedback-based learning. We hypothesized that learning-related behavior would be jointly predicted by oscillatory and aperiodic neural dynamics. Specifically, Hypothesis 2a predicted that FRN/RewP amplitude, P3 amplitude, and midfrontal theta power would be modulated by feedback valence and would predict immediate behavioral adjustments, such as response switching following negative feedback. Hypothesis 2b assumed that the aperiodic exponent would predict gradual changes in response strategy across trials, reflecting metacontrol-based shifts toward greater persistence or flexibility, independent of trial-level feedback effects.

In summary, by combining a feedback-based learning task with rigorously masked feedback using continuous flash suppression, and by dissociating oscillatory from aperiodic neural activity, the present study aimed to clarify how feedback shapes adaptive behavior both with and without conscious awareness, and to elucidate the distinct yet interacting neural mechanisms that support learning across immediate and strategic timescales.

2. Methods

2.1. Participants

Sixty participants (29 males) aged 18–26 (Mage = 21.00, SD ± 2.00) were recruited. In data analysis, 2 females were excluded since they could see the invisible feedback according to the two-alternative forced choice test. All participants had normal or corrected-to-normal vision and were right-handed with no history of neurological problems. They provided written informed consents before experiment and received monetary compensation for their participation after experiment. The study was approved by the Ethics Committee of the Faculty of Psychology, Shandong Normal University.

2.2. Experimental design

The study used a 2 (feedback valence: positive vs. negative) × 2 (feedback visibility: visible vs. invisible) within-participants design.

2.3. Experimental materials

The Mondrian masks in this study involved a randomization process that combined different, overlapping rectangles. Since the electrophysiological activities started around 100 ms involving 10 Hz frequency, we used 20 Hz dynamic mask to minimize interference for those activities while achieving effective intraocular suppression.

The feedback stimuli were facial expression pictures,[1] which were chosen from the Chinese Affective Picture System (CAPS, Wang and Luo, 2005). The happy expression indicated positive feedback (picture number: HM95), while the fearful expression indicated negative feedback (picture number: FM1). The size of the picture was 4 cm (3°) × 5 cm (3.9°).

2.4. Task procedure

Participants were seated comfortably in an electromagnetically shielded room and performed a time estimation task. Fig. 1 shows the sequence of events during each trial. Firstly, participants saw a fixation cross (500 ms) followed by a random blank (400 to 600 ms), then a white pentagram appeared to cue the beginning of time estimation. When the participants believed that 1 s had elapsed, they should press the space bar, after an interval of 400~600 ms the 800 ms feedback was presented, followed by the 600~800 ms ITI. If participants failed to respond for more than two seconds, the feedback "response is too slow, please focus your attention" was presented, and then moved to the next trial.

The feedback included visible and invisible conditions, and the invisible condition was realized by a continuous flash suppression (CFS) paradigm. Participants viewed facial expression pictures through reflective stereoscopy. The visible condition presented the same expression to both eyes; the invisible condition presented the dynamic Mondrian pictures to the dominant eye (Porac and Coren, 1976), while the expression picture was simultaneously presented to the other eye, which resulted in a suppression effect of the feedback picture.

Feedback was conveyed by pictures showing happy or fearful faces, which were taken to indicate positive and negative feedback, respectively. The time window for correct response was determined by a sliding time window procedure (Miltner et al., 1997). The initial time window was 1000±100 ms, the size would be tightened 10 ms following each correct trial, and expanded 10 ms following each incorrect trial. With this procedure, the frequencies for positive and negative feedback were comparable.

[1] Coins were usually used to represent reward conditions in past research (Correa et al., 2018). We did not use coins for two reasons. Firstly, coins are rarely used by young people in China, thus they are not familiar with coin pictures. Secondly, the CNY coins of 1 yuan and 10 cents are very similar in color and shape, and thus easy to confuse. There are two advantages for facial expressions: firstly, a study has shown that facial expressions can express feedback valence very well (Schulreich et al., 2013); secondly, happy and fearful expressions correspond to approaching and avoidance motivations, respectively, which are consistent with the meaning of positive and negative feedback.

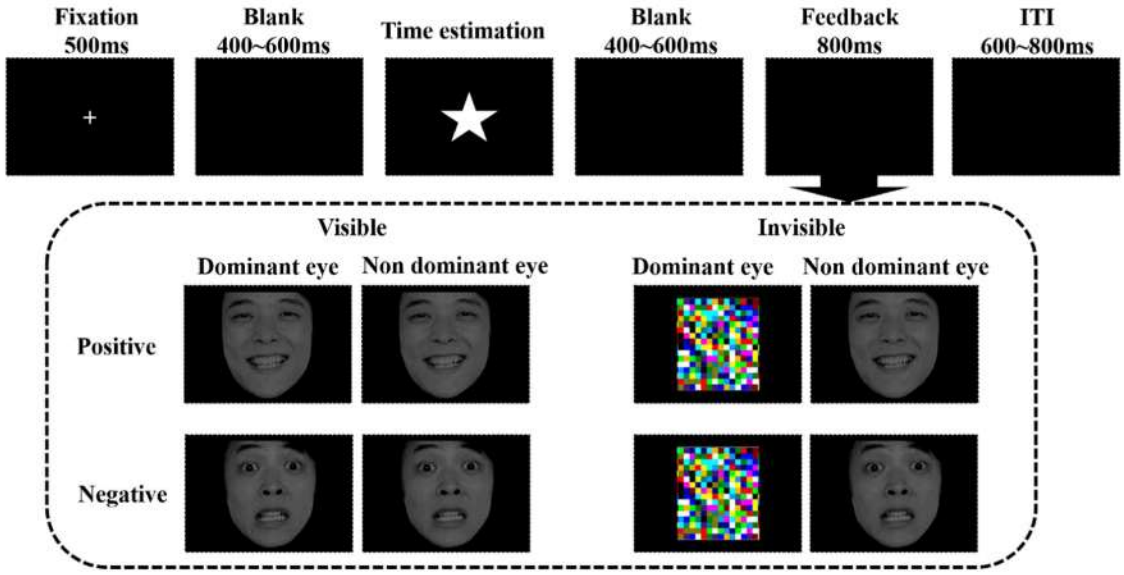

Fig. 1. Schematic illustration of a single trial of the time estimation task. In the visible feedback condition, both eyes were presented with the same expression picture. In the invisible feedback condition, the dominant eye was presented with the Mondrian picture and the nondominant eye was presented with the facial expression picture.

The experiment had four conditions: visible-positive feedback, visible-negative feedback, invisible-positive feedback, and invisible-negative feedback. The experiment consisted of 20 practice trials and 240 formal trials, in which 120 trials of visible condition and 120 trials of invisible condition were randomly presented. The formal experiment was divided into 4 blocks, and participants were allowed to take a break after each block.

To ensure that participants in the formal experiment did not see the faces, a two-alternative forced choice task (Fig. 2) was conducted prior to the EEG experiment for each participant (Jiang and He, 2006; Tsuchiya and Koch, 2005). As Fig. 2 shows, in each trial, two stimuli appeared, one after another. One was a facial expression masked with a Mondrian picture and the other was a blank masked with a Mondrian picture. Participants were required to judge whether the facial picture appeared in the former or the latter. There were 120 trials, the order of the facial picture and blank was random. A binomial distribution was utilized to analyze the accuracy of consciousness detection tasks, thereby excluding participants who could see the faces. With a mean (μ) of 60 and a standard deviation (σ) of 5.48, the upper limit of the binomial distribution accuracy was calculated as $(\mu + 1.96\sigma)/120 = 0.5896$, while the lower limit was $(\mu - 1.96\sigma)/120 = 0.4104$. Thus, the effective range is (0.4104, 0.5896), within which accuracy was at chance level. Participants with accuracy greater than 0.5896 or <0.4104 were deemed capable of seeing the faces and their data were consequently excluded. A total of 2 females were excluded based on this criterion.

2.5. Electrophysiological data recording and processing

The EEG was recorded using a 64-channel system (BrainAmp MR, Brain Products GmbH, Munich, Germany), and the electrodes were placed according to the 10/20 system. FCz was used as the online reference, and AFz served as the ground electrode, two electrodes on the left and right mastoids were also recorded. The horizontal electrooculogram (HEOG) at 1.5 cm lateral to the left eye and the vertical electrooculogram (VEOG) at 1.5 cm inferior to the right eye were recorded simultaneously. Electrode impedances were kept below 10 kΩ. Continuous EEG signals were recorded at a sampling rate of 1000 Hz using a bandpass filter of 0.016–70 Hz.

The raw EEG data were processed using EEGLAB (Delorme and Makeig, 2004) in Matlab R2021a, the MNE toolbox v.0.18.2 and the

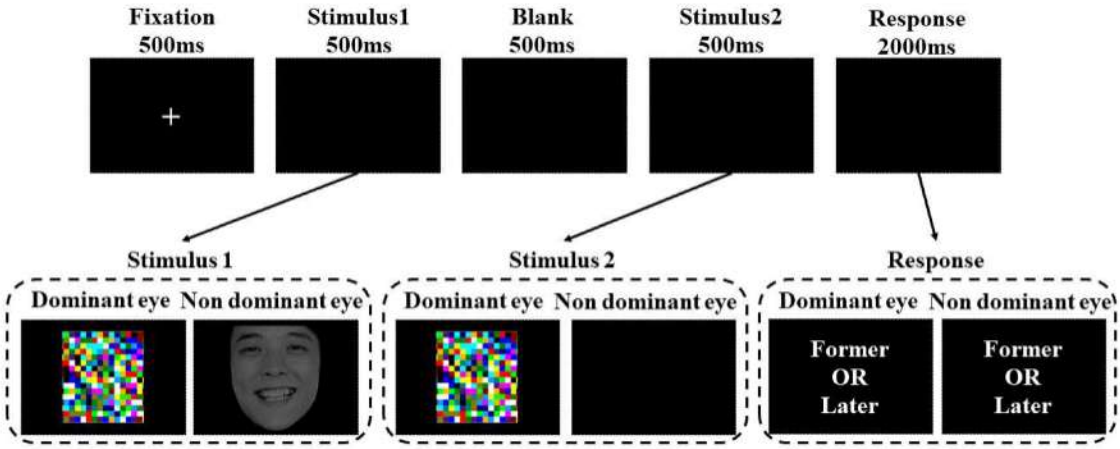

Fig. 2. Two-alternative forced choice task. Participants needed to judge whether the face appears in stimulus 1 or in stimulus 2. In this trial, the face appeared in stimulus 1 position.

FOOOF toolbox (Donoghue et al., 2020) in Python (v3.9, Python Software Foundation). Data were bandpass filtered from 0.1 to 40 Hz offline. Epochs ranging from 1000 ms before to 1000 ms after feedback onset were extracted. Using independent component analysis, artifacts caused by blinks and other events not related to brain activities were removed from the EEG data. On average, 2.77±1.18 ocular independent components (ICs) were removed (range: 0 to 6). Finally, the data re-referenced to average reference. This experiment resulted in 53.783 ±6.808 trials for the visible-positive feedback condition, 58.167±7.434 trials for the visible-negative feedback condition, 54.333±6.501 trials for the invisible-positive feedback condition, and 57.967±6.006 trials for the invisible-negative feedback condition.

We assessed two ERP components evoked by feedback processing in the four experimental conditions, the REWP and the P3a. Component amplitudes were calculated using estimation routines implemented in the Letswave, time-locked to feedback onset and baseline-corrected from 200 to 0 ms pre-feedback. REWP amplitude was measured as the mean amplitude in the window from 240–300 ms post-feedback at the electrodes FCz. P3a amplitude was measured as the mean amplitude in the window from 300–450 ms post-feedback at electrodes FCz, the FCz point was chosen according to the topography of P3a (Fig. 5), and was termed as P3a in this study. The time periods of 200 ms before and 800 ms after the feedback stimulus were re-intercepted as the analysis window. To be noted, we selected the time windows and electrodes based on visual selection and previous studies; although this is a traditionally used method, it might inflate the chance of false positives positive (Brooks et al., 2017).

2.6. Separating periodic and aperiodic activity

Power spectral densities (PSDs) were first calculated separately for each participant and electrode using multitaper method in Python. Averaging the individual PSDs of each epoch resulted in a smoothed power spectrum for future analysis.

In the standard analysis approach, the total canonical theta band power was calculated by averaging power in the range fixed of 4 ~ 8 Hz.

The SpecParam algorithm (FOOOF Python toolbox; version 1.1.0; https://fooof-tools.github.io/fooof/) was then used to parameterize the spectral data through separation of the periodic and aperiodic components of the signal. Using this approach, PSDs are treated as a linear combination of both aperiodic activity and oscillatory peaks with amplitudes that extend above the aperiodic signal (for a detailed overview of this approach see: Donoghue et al., 2020). Using a model driven approach, the FOOOF algorithm is able to extract both periodic and aperiodic components within the overall power spectra (Donoghue et al., 2020).

For the present study, we extracted the aperiodic exponent across a broad frequency range between 3 and 40 Hz, and as recommended in the FOOOF documentation in order to allow for reliable estimation of the aperiodic component of the data. Spectral parameterization settings for the algorithm were: peak width limits = [1,8], maximum number of peaks = 6, minimum peak height = 0.05. The final FOOOF outputs are the aperiodic exponent and offset values of the signal. The average R^2 of spectral fits for all participants was 0.923.

Subsequently, the aperiodic signal was subtracted from the total power spectrum to receive an aperiodic-adjusted power spectrum. The total power spectrum is extracted using the 'power_spectrum' attribute and the aperiodic signal is extracted using the 'fooofed_spectrum_' attribute. Aperiodic-adjusted theta power was calculated in the frequency ranges described above.

2.7. Data processing

2.7.1. Behavioral data

Time estimation error in trial $N + 1$ refers to the absolute difference between time estimation and 1 s, which was calculated as $|\text{time estimation}_{(\text{trial } N+1)}\text{-}1000|$ ms. Time estimation adjustment in trial $N + 1$ was calculated as $|\text{time estimation}_{(\text{trial } N+1)} - \text{time estimation}_{\text{trial N}}|$ ms. Two-way repeated measures ANOVA, with the factors of feedback valence (positive vs. negative) and visibility (visible vs. invisible) in trial N, were conducted on time estimation error and time estimation adjustment in trial $N + 1$.

2.7.2. Electrophysiological data

For the ERP data, two-way repeated measures ANOVAs with the same factors were conducted on averaged REWP and P3a amplitudes, on theta and aperiodic-adjusted theta power, and on the aperiodic exponent. Statistical analyses were conducted using SPSS, version 27.0 and JASP 0.18.2.0, and Greenhouse-Geisser corrections were applied to *p*-values that did not meet the assumption of sphericity. In addition, the *F*-test report η_p^2, where η_p^2 <0.06, indicates a weak effect size; $0.06 \leq \eta_p^2 \leq 0.14$, indicates a moderate effect size; and $\eta_p^2 > 0.14$, indicates a strong effect size.

3. Results

3.1. Behavioral results

Participants received 47.7 %±3.9 % (*M*±*SD*) positive feedback on average, indicating that the sliding time window was effective. The ANOVA on trial $N + 1$ time estimation error yielded a significant main effect of feedback valence, $F(1,57)=76.250$, $p < 0.001$, $\eta_p^2 =0.572$, indicating that the error was smaller after positive feedback (159.405± 7.676 ms) than after negative feedback (206.539 ± 10.210 ms)(see Fig. 3A). The main effect of feedback visibility was also significant, $F(1,57)=5.538$, $p = 0.022$, $\eta_p^2 =0.089$, indicating a smaller error after invisible feedback (178.558± 8.739 ms) than after visible feedback (187.386 ± 8.903 ms). The absence of a significant interaction ($p = 0.375$) suggests that we did not detect evidence for an influence of visibility on the effectiveness of feedback.

The ANOVA of time estimation adjustments in trial $N + 1$ also revealed a main effect of feedback valence, $F(1,57)=220.478$, $p < 0.001$, $\eta_p^2 =0.795$, indicating greater adjustments after negative feedback (205.224±6.696 ms) than after positive feedback (144.217±5.351 ms), and a main effect of visibility, $F(1,57)=14.018$, $p < 0.001$, $\eta_p^2 =0.197$, indicating lesser adjustments after invisible feedback (170.232± 5.781 ms) than after visible feedback (179.209 ± 5.872 ms) (see Fig. 3B). Again, there was no significant interaction, $p = 0.068$, confirming that the effectiveness of feedback was independent of its visibility.

3.2. Electrophysiological results

3.2.1. REWP

The ANOVA found no significant main effect of feedback valence, $p = 0.118$, but a significant main effect of feedback visibility, $F(1,57)= 116.517$, $p < 0.001$, $\eta_p^2 =0.672$ (see Fig. 4). It was driven by more positive REWP amplitudes for the visible condition (4.132±0.345μV) than for the invisible condition (0.188±0.392μV). There was a significant interaction between feedback visibility and valence, $F(1,57) = 6.261$, $p = 0.015$, $\eta_p^2 =0.099$. Further simple effect analysis found that the valence effect was significant when the feedback was visible, $F(1,57)= 7.283$, $p = 0.009$, $\eta_p^2 =0.113$, but not significant when the feedback was invisible, $F(1,57)=0.488$, $p = 0.488$, $\eta_p^2 =0.008$. Negative feedback (3.925 ± 0.337 μV) induced more negative REWP amplitudes than positive feedback (4.339 ± 0.370 μV) in visible conditions (see Fig. 4 and Fig. 6A).

3.2.2. P3a

The ANOVA found significant main effects of feedback valence, $F(1,57) = 30.992$, $p < 0.001$, $\eta_p^2 =0.352$, and of feedback visibility, *F*

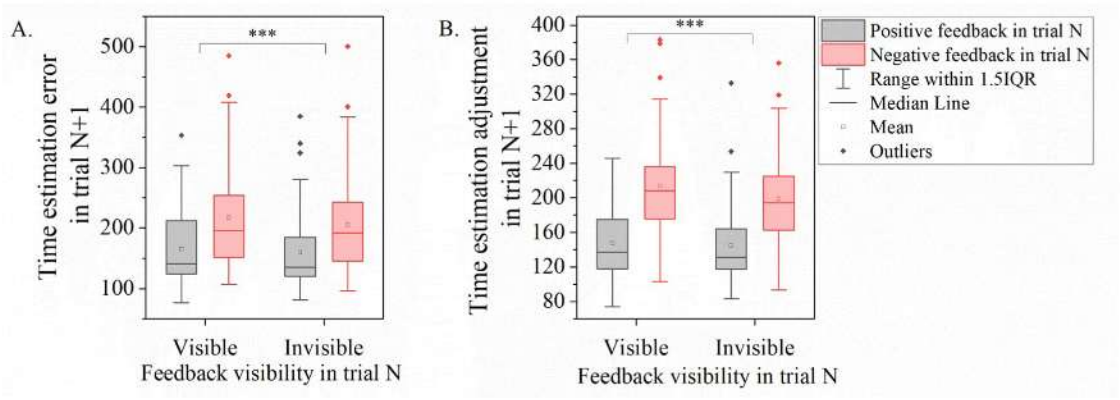

Fig. 3. The time estimation error of trial $N + 1$. After positive feedback, time estimation performance was more accurate than after negative feedback, in both visible and invisible feedback conditions (A). Time estimation adjustment in trial $N + 1$. After negative feedback, the time estimation adjustment was greater than after positive feedback, unaffected by feedback visibility (B).

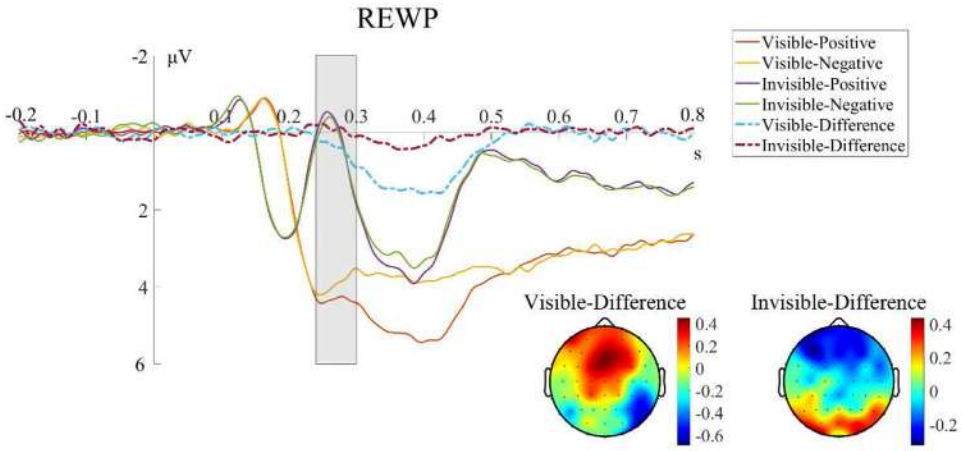

Fig. 4. The grand-averaged waveforms of the REWP at channel FCz for each condition as well as their difference waves. Topographies were drawn based on difference waves between positive and negative conditions (240 - 300 ms).

(1,57)= 19.352, $p < 0.001$, η_p^2 =0.253 (see Fig. 5). The main effect of feedback valence was driven by more positive P3a amplitudes for the positive condition (4.118±0.312μV) than for the negative condition (3.309±0.290μV). The main effect of feedback visibility was driven by more positive P3a amplitudes for the visible condition (4.421±0.357μV) than for the invisible condition (3.006±0.310μV). There was a significant interaction between feedback visibility and valence, F(1,57) = 22.495, $p < 0.001$, η_p^2 =0.283. Further simple effect analysis found that the valence effect was significant when the feedback was visible,F(1,57) = 39.032, $p < 0.001$, η_p^2 =0.406, but not significant when the feedback

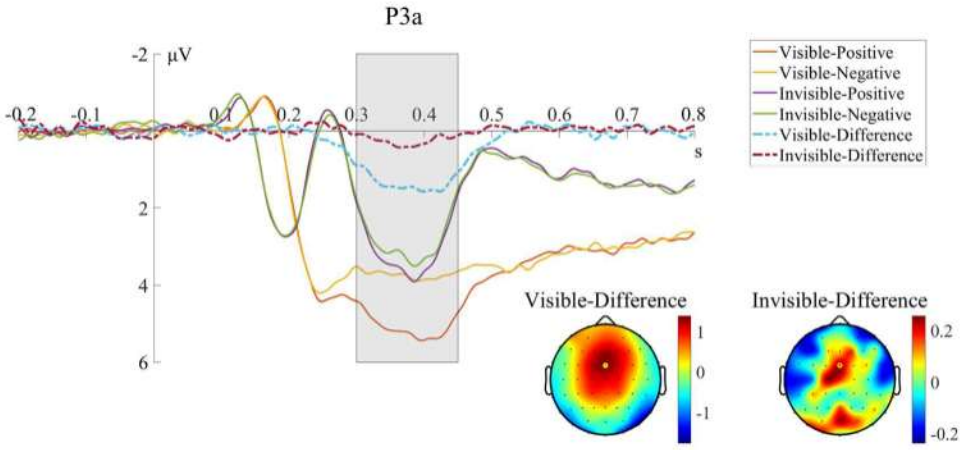

Fig. 5. Grand-averaged waveforms of the P3a at channel FCz for each condition as well as their difference. The topographies were drawn based on difference waves between positive and negative conditions (300 - 450 ms).

was invisible, $F(1,57) = 2.906$, $p = 0.094$, $\eta_p^2 = 0.049$. Positive feedback (5.104 ± 0.388 μV) induced more positive P3a amplitudes than negative feedback (3.738 ± 0.357 μV) in visible conditions (see Fig. 5 and Fig. 6B).

3.2.3. Theta

The ANOVA yielded significant main effects of feedback valence, $F(1,57)=21.435$, $p < 0.001$, $\eta_p^2 = 0.273$, and feedback visibility, $F(1,57)=23.210$, $p < 0.001$, $\eta_p^2 =0.289$ (see Fig. 7). There was also a significant interaction, $F(1,57)= 30.256$, $p < 0.001$, $\eta_p^2 =0.347$. Further simple effects analysis found that the valence effect was significant when the feedback was visible, $F(1,57)= 34.236$, $p < 0.001$, $\eta_p^2 =0.375$, but not significant when the feedback was invisible, $F(1,57)=0.624$, $p = 0.433$, $\eta_p^2 = 0.011$. Positive feedback (−12.262 ± 0.188μV) induced more negative theta oscillation than negative feedback (−12.217 ± 0. 204μV) in the visible condition (see Fig. 7).

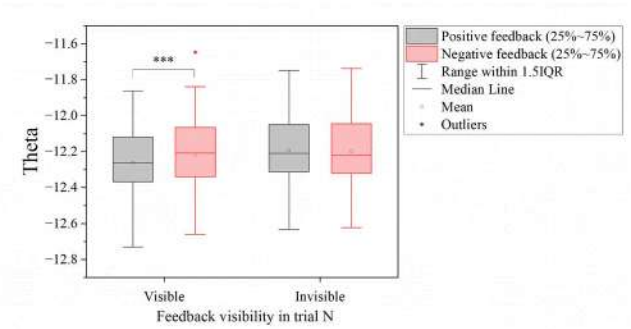

Fig. 7. Theta oscillation in positive and negative feedback conditions for visible and invisible feedback. In invisible feedback conditions, theta power was larger in the negative than in the positive feedback condition.

3.2.4. Aperiodic-adjusted theta

The ANOVA yielded a significant main effect of feedback visibility, $F(1,57)=16.914$, $p < 0.001$, $\eta_p^2 = 0.229$, indicating higher aperiodic-adjusted theta oscillation for the visible condition (−0.067±0.008) than for the invisible condition (−0.086±0.009), but no significant main effect of feedback valence, $F(1,57)= 0.101$, $p = 0.752$, $\eta_p^2 =0.002$, or interaction, $p = 0.401$ (see Fig. 8).

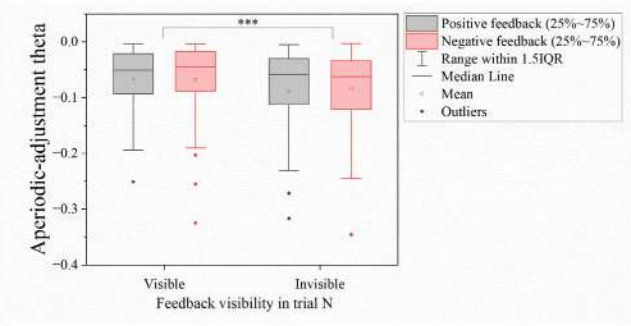

Fig. 8. The aperiodic-adjusted theta between positive and negative feedback conditions across both visible and invisible feedback conditions.

3.2.5. Aperiodic exponent

The ANOVA yielded significant main effects of feedback visibility, $F(1,57)=11.294$, $p = 0.001$, $\eta_p^2 = 0.165$, indicating that the exponent was lower in the visible condition (0.926±0.045) than in the invisible condition (0.961±0.043), and feedback valence, $F(1,57)=4.403$, $p = 0.040$, $\eta_p^2 = 0.072$, showing that the exponent was lower for the positive feedback condition (0.936±0.043) than for the negative feedback condition (0.952±0.045)(see Fig. 9). The interaction was also significant, $F(1,57)= 20.794$, $p < 0.001$, $\eta_p^2 =0.267$. Further simple effect analysis found that the valence effect was significant when the feedback was visible, $F(1,57)= 12.601$, $p = 0.001$, $\eta_p^2 =0.181$, due to a higher exponent for negative feedback (0.948 ± 0.357) than for positive feedback (0.905 ± 0.330), but not when the feedback was invisible, $p = 0.172$.

3.3. Correlations with performance

For each of the four experimental conditions, correlations were computed between the two behavioral variables, time estimation error and time estimation adjustment, and the five electrophysiological indicators, REWP, P3a, Theta, adjusted Theta, and Aperiodic Exponent. In the visible-positive feedback condition, no significant correlation was obtained, even though the correlations between time estimation error and P3a, $r^2=-0.27$, $p = 0.07$, and between time estimation error and aperiodic exponent, $r^2=-0.23$, $p = 0.08$, approached significance. In all three other conditions, only the correlations between time estimation error and aperiodic exponent were significant: with visible-negative feedback, $r^2=-0.32$, $p = 0.02$, invisible-positive feedback, $r^2=-0.27$, $p = 0.04$, and invisible-negative feedback, $r^2=-0.33$, $p = 0.01$. Fig. 10 provides an overview.

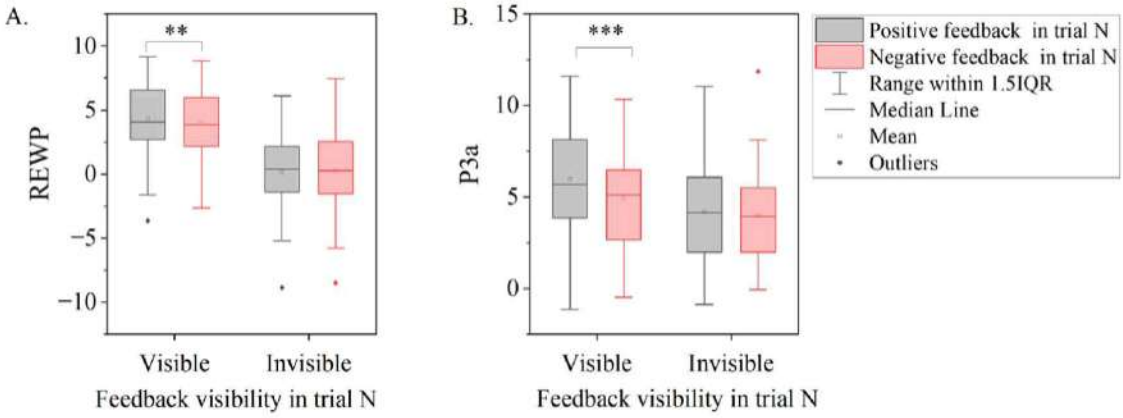

Fig. 6. REWP. Negative feedback induced more negative REWP than positive feedback only in the visible feedback condition (A). P3a. Positive feedback induced more positive P3a than negative feedback only in the visible feedback condition (B).

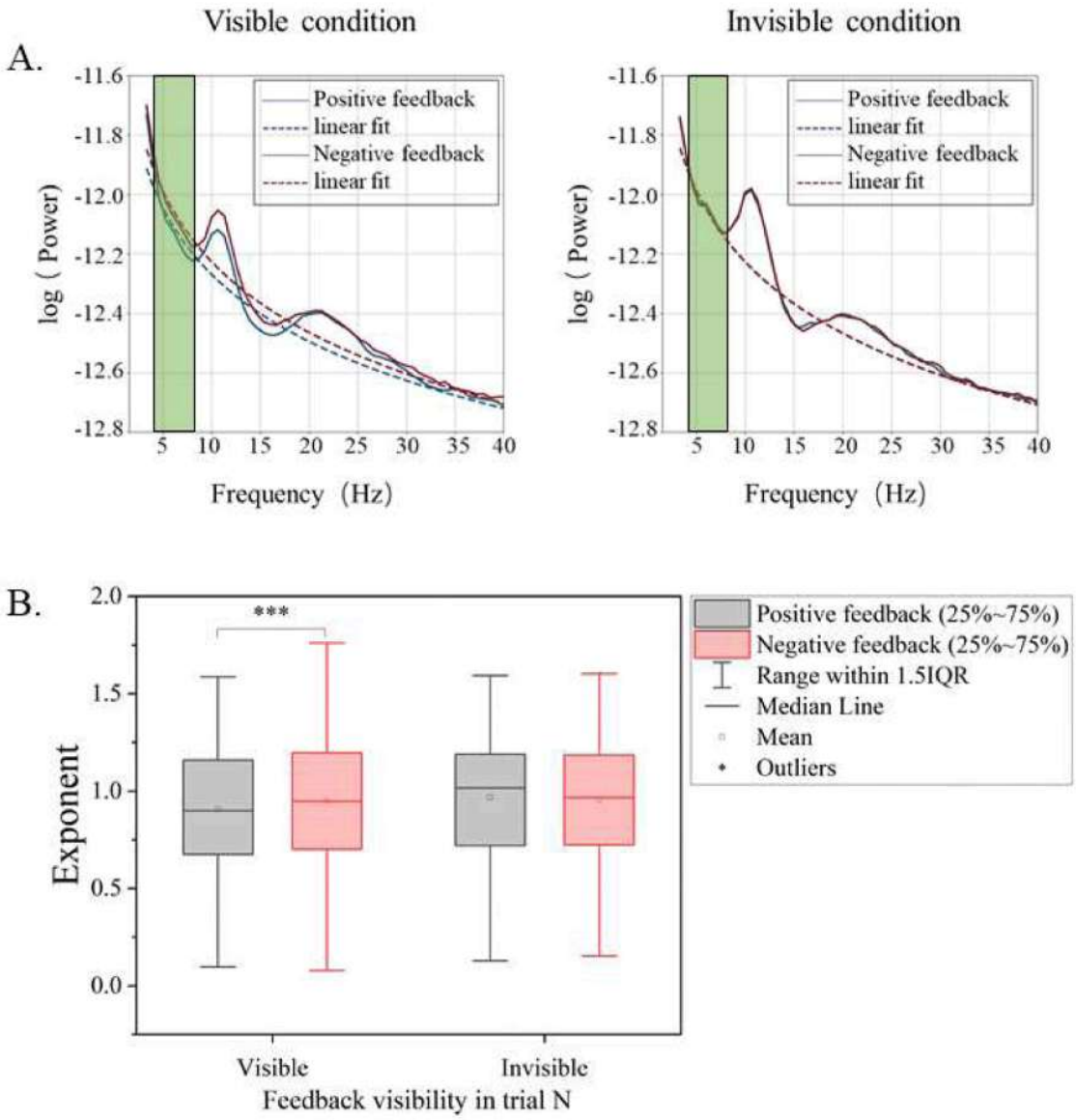

Fig. 9. Exponents as a function of feedback visibility and feedback valence. Green color shows the theta frequency. A. depicts the results of the model fit, and B. shows a boxplot of the descriptive statistics.

4. Discussion

The present study investigated whether feedback can drive adaptive behavioral learning even when it is rendered inaccessible to conscious awareness, and how both oscillatory and aperiodic components of the EEG signal contribute to this process. Our goal was to determine not only whether masked feedback influences behavior, but also whether oscillatory markers traditionally linked to reward processing and the aperiodic exponent associated play dissociable roles in shaping feedback-guided adaptation. The aim of this study was twofold.

First, we wanted to test whether the findings of Correa et al. (2018) can be replicated with a more effective masking technique. These original findings showed that more adjustments were made after negative feedback, even in the masked feedback condition. Although the post-error adjustment was statistically significant in the masked condition, the effect was much smaller than in the unmasked condition. That is to say, masking did weaken reinforcement learning in the study by Correa and colleagues (2018)). In contrast, in the present study, the behavioral learning effect did not show evidence of modulation by masking. The particular technique used by Correa et al. permitted feedback presentation for just 17 ms in the masking condition, whereas our paradigm allowed for a presentation of 800 ms in both visible and invisible conditions. This difference may very well account for the equivalence of learning in visible and invisible conditions in the present study. Hence, we successfully replicated the previous finding that learning can occur with invisible feedback.

Moreover, using a different, highly effective masking technique, we found that when the invisible feedback was presented for a sufficiently long duration, learning performance was comparable to that with visible feedback. Behavioral analyses showed that positive feedback reinforced more accurate estimations: after positive feedback, the time estimation in trial $N + 1$ was more accurate than after negative feedback. This reinforcement effect did not differ between the visible and invisible feedback conditions. In addition, participants made greater time estimation adjustments after incorrect feedback than after correct feedback, and this post-error adjustment also did not differ between the visible and invisible conditions. Although adjustments were larger following negative feedback, performance was better after positive feedback. This apparent contradiction is understandable: negative feedback provides only correctness information, without specifying the direction or extent

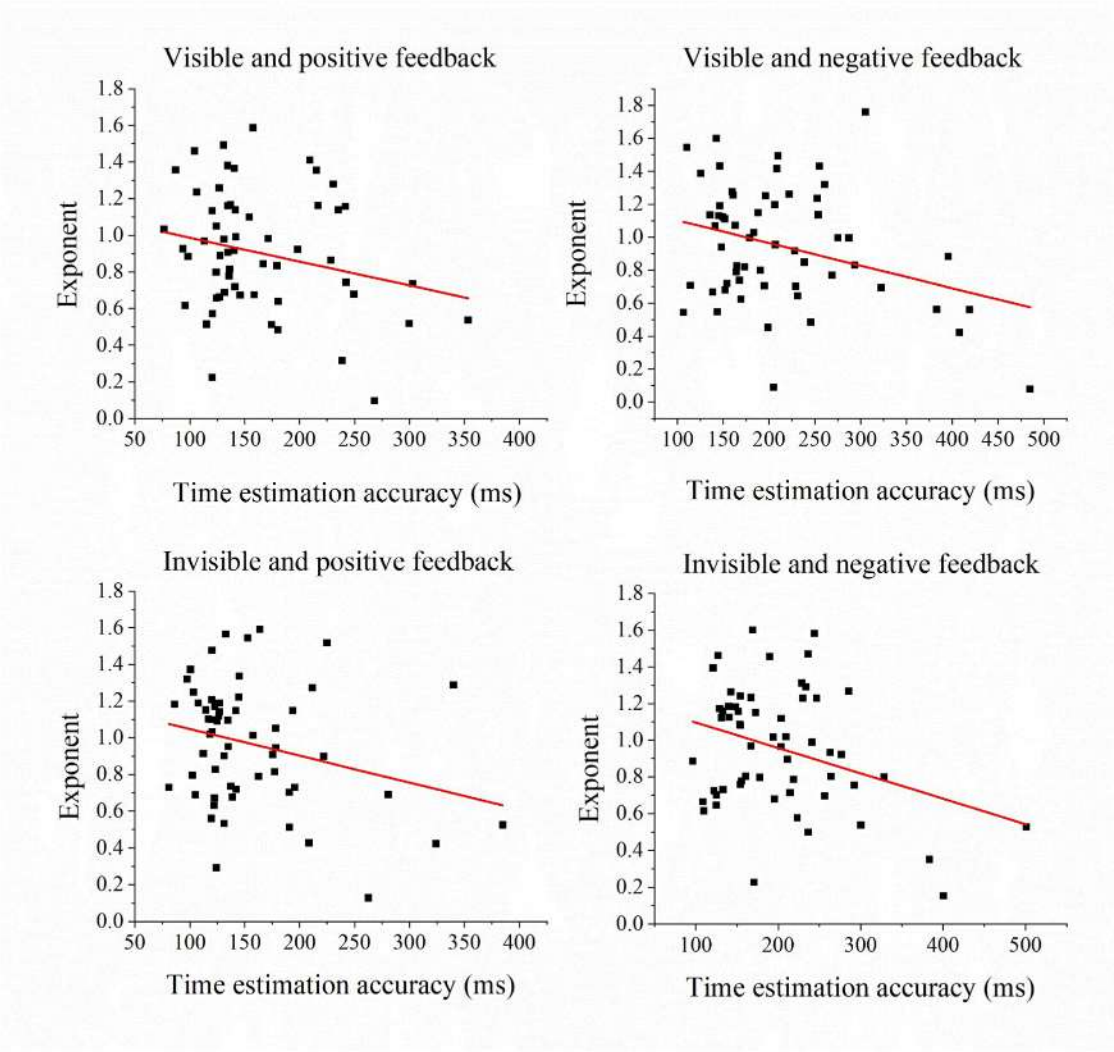

Fig. 10. Scatter plots of the correlations between aperiodic exponent and time estimation error as a function of feedback visibility and valence.

of the error. Thus, a greater adjustment does not necessarily mean a more accurate adjustment—it simply reflects a stronger response to the error. By contrast, reinforcement after correct performance provides a clear directional signal, which supports more accurate estimations. Consequently, making larger adjustments after errors does not guarantee improved accuracy. This may also explain the lower accuracy observed under the visible condition, which could be due to stronger post-error adjustments following negative feedback. Therefore, for negative feedback, future studies should include information about both the direction and the extent of the error to further examine whether such detailed behavioral adjustments can also be achieved based on invisible feedback.

Although we observed robust valence effects in the electrophysiological data for visible feedback, we did not find significant correlations between feedback-locked ERPs (FRN/REWP, P3) and behavioral learning outcomes. While some prior studies have reported links between feedback-locked ERP amplitude and learning outcomes (Bellebaum and Daum, 2008; Kobza et al., 2011), there is also consistent evidence that such brain–behavior correlations are not guaranteed. For example, Arbel et al. (2013) found that FRN amplitude elicited by negative feedback did not predict long-term learning, although positive-feedback-related FRN-like activity did. Similarly, Arbel & Fox (2021) observed that FRN responses to negative feedback in children did not correlate with learning retention, with associations emerging only for positive feedback. A review by San Martín (2012) further highlights that, across paradigms, outcome-related ERP modulations are reliable, but predictive relationships with behavior are often weak or inconsistent. In addition, Walsh & Anderson (2011) demonstrated that feedback-locked ERP amplitudes can vary with task instructions and context, limiting their predictive value for subsequent adjustments. Taken together, these findings suggest that the absence of ERP–behavior correlations in our study is not anomalous. Instead, it may reflect boundary conditions such as masked feedback visibility, adaptive sliding-window feedback, and the continuous, skill-based nature of the time-estimation task, in which immediate electrophysiological responses do not directly translate into overt behavioral adjustments.

We intentionally focused on fronto-central positivity (often referred to generically as the P3a) rather than a separate parietal P3b analysis because in our paradigm, signal characteristics and hypotheses pointed most directly to attention-orienting processes rather than memory-updating mechanisms. According to Polich's integrative model (2007), the P3a component—characterized by fronto-central scalp distribution and earlier latency—is generated by stimulus-driven frontal attention systems, whereas the P3b—with centro-parietal maxima and later latency—is associated with temporal-parietal activity and memory processing. In our task the feedback was salient and immediately visible, promoting rapid valence discrimination rather than extended memory processing; the resultant positivity peaked around 300 ms at fronto-central sites, consistent with the canonical profile of the P3a (i.e. earlier latency and frontal/central maximum), rather than a classic

parietal P3b (which typically peaks later and exhibits a centro-parietal distribution) (Polich, 2007; Polich and Criado, 2006). Hence, given our hypothesis framework focused on early valence processing and attentional allocation rather than longer latency encoding, we considered the fronto-central component to be the most theoretically appropriate index.

Our second aim was to better understand the feedback-induced cognitive/neural processes that manage to adjust behavioral control according to the valence of this feedback. Here we took two different analytical angles. One was to look into electrophysiological indicators of processes that are initiated by the presentation of feedback. With this angle, we found an interesting dissociation from the behavioral observations, as the valence effect of all five indicators was significant in the visible condition but statistically disappeared in the invisible condition. This pattern suggests the following scenario: on the one hand, all these indicators apparently point at processes that are systematically triggered by the presence of feedback, suggesting that they are either busy with processing aspects of this feedback or triggered by the presence of feedback for other, more general purposes—an issue we will get back to in the case of the aperiodic exponent. However, the particular outcome of the initiated process does not seem to directly inform future performance, which is obvious from the fact that behavioral performance was sensitive to valence even under conditions where the electrophysiological indicators (and, presumably, the processes they indicate) were unaffected.

Our other angle was interested in the degree to which the electrophysiological indicators, and the processes they indicate, contribute to behavioral control and performance. Correlations between indicators and performance revealed that only one indicator showed effective contributions to performance (the time estimation error), namely, the aperiodic exponent. More specifically, low error rates were associated with a high exponent, suggesting that performance was best in the presence of low cortical noise. Interestingly, this correlation was independent from valence and visibility. This suggests that the variability of the aperiodic exponent does not reflect a systematic, strategic choice that was informed by the type or visibility of the feedback. Rather, the aperiodic exponent seems to capture some aspect of the general neural state under which feedback-induced adjustments are made, with a less noisy state being more supportive of effective learning. Note that this scenario is consistent with previous considerations that the aperiodic exponent might capture metacontrol states (i.e., states that enable a particular processing style, rather than a particular cognitive operation), which in turn can be characterized by variations of neural noisiness (Hommel and Colzato, 2017; Zhang et al., 2023).

Two methodological considerations should be taken into account when interpreting the electrophysiological findings of the present study. First, the FRN/RewP is known to overlap temporally and spatially with other feedback-locked components, including early P3-related positivities as well as stimulus-specific perceptual and emotional responses (Holroyd and Coles, 2002; Sambrook and Goslin, 2015). While difference-wave approaches are commonly used to isolate FRN/RewP activity, their effectiveness depends strongly on the salience, simplicity, and symbolic clarity of the feedback. When feedback is non-monetary and perceptually complex, as in the case of emotional facial expressions, overlapping processes may obscure a distinct FRN/RewP peak even when reward-related processing is present.

Second, although faces are highly salient and rapidly discriminable, they are not a standard or uniformly reliable method for eliciting FRN/RewP effects. A prior study has demonstrated that faces can reliably elicit FRN/RewP signals, particularly when the faces convey evaluative information (e.g., happy vs. angry) and are contingent on performance (Zhang et al., 2012). However, it also showed that reward-related activity may be delayed relative to the canonical FRN/RewP time window and, in some cases, can only be isolated using multivariate decomposition techniques such as principal component analysis rather than conventional difference waves (Zhang et al., 2012). Moreover, one other study has reported weak or absent FRN/RewP effects when feedback is conveyed via faces, suggesting that facial expressions may preferentially engage social, emotional, and perceptual processing streams rather than core reinforcement-learning mechanisms (Redden et al., 2021). Importantly, early electrophysiological differences between happy and fearful faces do not necessarily reflect feedback evaluation or reward prediction error processing. Such effects may instead index rapid face perception or emotion categorization processes, which can occur prior to or in parallel with feedback-related evaluation. In the present study, the meanings of happy and fearful faces were not counterbalanced, which further limits the ability to disentangle feedback valence processing from stimulus-specific emotional responses. We therefore caution against interpreting early valence effects as unambiguous markers of FRN/RewP activity.

These considerations do not undermine the central behavioral findings of the study, nor the observed dissociation between behavioral learning and electrophysiological valence sensitivity. However, they place important constraints on the interpretation of the ERP results and underscore that emotional faces should not be regarded as a standard or optimal stimulus for isolating FRN/RewP components. Future work would benefit from counterbalancing stimulus–outcome mappings and from combining traditional difference-wave analyses with decomposition methods to better separate overlapping feedback-related and emotion-related neural processes.

Taken altogether, our findings suggest that people can improve their behavior based on feedback, and they can do so irrespective of whether this feedback can be consciously perceived or not. In contrast, all the electrophysiological indicators we considered in this study show a much greater sensitivity to conscious perception, in the sense that they do not show any valence effect in the absence of conscious feedback information. This implies that the processes that the electrophysiological measures indicate may contribute to behavioral control, but the fact that they can be sensitive to the valence of visible feedback does not seem to be important for the degree to which valence can inform and help initiating behavioral adjustments. In other words, we did not observe a statistically reliable relationship between electrophysiological valence sensitivity and behavioral outcomes. Our findings also suggest that, while the processes indicated by REWP, P3a, and theta activity may well be important in the chain of processes that lead from feedback to adjustments in behavioral control, their intra-individual variability is not systematically related to the quality of behavioral performance. This was different for the aperiodic exponent, however. Even though this exponent was not any more sensitive to valence in the invisible condition than the other indicators were, it was the only indicator that systematically predicted behavioral control. The outcome pattern obtained suggests that the aperiodic exponent may capture the characteristics of some general neural state that has the potential of supporting effective adjustments of behavioral control. As our findings suggest, low cortical noisiness seems particularly suitable for this purpose.

Data and code availability statement

All data can be obtained from the corresponding author Shiwei Jia upon reasonable request.

CRediT authorship contribution statement

DanDan Liu: Writing – review & editing, Visualization, Investigation, Formal analysis. **Shiwei Jia:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Yanliang Sun:** Writing – review & editing, Visualization, Investigation, Formal analysis. **Lorenza Colzato:** Writing – review & editing. **Bernhard Hommel:** Writing – review & editing, Writing – original draft, Supervision, Funding acquisition.

Declaration of competing interest

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Acknowledgements

The study was funded by the "One case, one policy" grant from Shandong Province (China) awarded to B.H. and the Natural Science Foundation of Shandong Province (grant number: ZR2022MC115) awarded to S.J.

References

Arbel, Y., Fox, A.B., 2021. Electrophysiological examination of feedback-based learning in 8-11-year-old children. Front. Psychol. 12, 640270. https://doi.org/10.3389/fpsyg.2021.640270.

Arbel, Y., Goforth, K., Donchin, E., 2013. The good, the bad, or the useful? The examination of the relationship between the feedback-related negativity (FRN) and long-term learning outcomes. J. Cogn. Neurosci. 25 (8), 1249–1260. https://doi.org/10.1162/jocn_a_00385.

Bellebaum, C., Daum, I., 2008. Learning-related changes in reward expectancy are reflected in the feedback-related negativity. Eur. J. Neurosci. 27 (7), 1823–1835. https://doi.org/10.1111/j.1460-9568.2008.06138.x.

Bijleveld, E., Custers, R., Aarts, H., 2012. Human reward pursuit: from rudimentary to higher-level functions. Curr. Dir. Psychol. Sci. 21 (3), 194–199. https://doi.org/10.1177/0963721412438463.

Brooks, J.L., Zoumpoulaki, A., Bowman, H., 2017. Data-driven region-of-interest selection without inflating type I error rate. Psychophysiology 54 (1), 100–113. https://doi.org/10.1111/psyp.12682.

Cavanagh, J.F., Frank, M.J., 2014. Frontal theta as a mechanism for cognitive control. Trends Cogn. Sci. 18 (8), 414–421. https://doi.org/10.1016/j.tics.2014.04.012.

Chase, H.W., Swainson, R., Durham, L., Benham, L., Cools, R., 2011. Feedback-relatedNegativity codes prediction error but not behavioral adjustment duringProbabilistic reversal learning. J. Cogn. Neurosci. 23 (4), 936–946. https://doi.org/10.1162/jocn.2010.21456.

Cohen, M.X., 2014. A neural microcircuit for cognitive conflict detection and signaling. Trends Neurosci. 37 (9), 480–490. https://doi.org/10.1016/j.tins.2014.06.004.

Colzato, L., Zhang, H., Roessner, V., Beste, C., Hommel, B., 2025. Non-bivalent psychopathology: rethinking mental disorders through metacontrol. Neurosci. Biobehav. Rev. 176, 106297. https://doi.org/10.1016/j.neubiorev.2025.106297.

Correa, C.M.C., Noorman, S., Jiang, J., Palminteri, S., Cohen, M.X., Lebreton, M., van Gaal, S., 2018. How the level of reward awareness changes the computational and electrophysiological signatures of reinforcement learning. J. Neurosci. 38 (48), 10338–10348. https://doi.org/10.1523/JNEUROSCI.0457-18.2018.

Delorme, A., Makeig, S., 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J. Neurosci. Methods 134 (1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009.

Donoghue, T., Dominguez, J., Voytek, B., 2020. Electrophysiological frequency band ratio measures conflate periodic and aperiodic neural activity. Eneuro 7 (6). https://doi.org/10.1523/ENEURO.0192-20.2020. ENEURO.0192-20. 2020.

Fang, F., He, S., 2005. Cortical responses to invisible objects in the human dorsal and ventral pathways. Nat. Neurosci. 8 (10), 1380–1385. https://doi.org/10.1038/nn1537.

He, B.J., 2014. Scale-free brain activity: past, present, and future. Trends Cogn. Sci. 18 (9), 480–487. https://doi.org/10.1016/j.tics.2014.04.003.

Holroyd, C.B., Coles, M.G.H., 2002. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. Psychol. Rev. 109 (4), 679–709. https://doi.org/10.1037/0033-295X.109.4.679.

Holroyd, C.B., Umemoto, A., 2016. The research domain criteria framework: the case for anterior cingulate cortex. Neurosci. Biobehav. Rev. 71, 418–443. https://doi.org/10.1016/j.neubiorev.2016.09.021.

Hommel, B., Colzato, L., Beste, C., 2024. No convincing evidence for the independence of persistence and flexibility. Nat. Rev. Psychol. 3 (9), 638. https://doi.org/10.1038/s44159-024-00353-6. –638.

Hommel, B., Colzato, L.S., 2017. The social transmission of metacontrol policies: mechanisms underlying the interpersonal transfer of persistence and flexibility. Neurosci. Biobehav. Rev. 81, 43–58. https://doi.org/10.1016/j.neubiorev.2017.01.009.

Jia, S., Liu, D., Song, W., Beste, C., Colzato, L., Hommel, B., 2024. Tracing conflict-induced cognitive-control adjustments over time using aperiodic EEG activity. Cereb. Cortex 34 (5), bhae185. https://doi.org/10.1093/cercor/bhae185.

Jia, S., Liu, D., Yue, Y., Colzato, L., Hommel, B., Beste, C., 2025. Periodic and aperiodic neural activity contribute to the microgenesis of learning from mistakes. NeuroImage 317, 121341. https://doi.org/10.1016/j.neuroimage.2025.121341.

Jiang, Y., He, S., 2006. Cortical responses to invisible faces: dissociating subsystems for facial-information processing. Curr. Biol. 16 (20), 2023–2029. https://doi.org/10.1016/j.cub.2006.08.084.

Kobza, S., Thoma, P., Daum, I., Bellebaum, C., 2011. The feedback-related negativity is modulated by feedback probability in observational learning. Behav. Brain Res. 225 (2), 396–404. https://doi.org/10.1016/j.bbr.2011.07.059.

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. J. Cogn. Neurosci. 9 (6), 788–798. https://doi.org/10.1162/jocn.1997.9.6.788.

Pessiglione, M., Schmidt, L., Draganski, B., Kalisch, R., Lau, H., Dolan, R.J., Frith, C.D., 2007. How the brain translates money into force: a neuroimaging study of subliminal motivation. Science 316 (5826), 904–906. https://doi.org/10.1126/science.1140459.

Polich, J., 2007. Updating P300: an integrative theory of P3a and P3b. Clin. Neurophysiol. : Off. J. Int. Fed. Clin. Neurophysiol. 118 (10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019.

Polich, J., Criado, J.R., 2006. Neuropsychology and neuropharmacology of P3a and P3b. Int. J. Psychophysiol.: Off. J. Int. Organ. Psychophysiol. 60 (2), 172–185. https://doi.org/10.1016/j.ijpsycho.2005.12.012.

Porac, C., Coren, S., 1976. The dominant eye. Psychol. Bull. 83 (5), 880–897.

Redden, R.S., Gagliardi, G.A., Williams, C.C., Hassall, C.D., Krigolson, O.E., 2021. Champ versus Chump: viewing an opponent's face engages attention but not reward systems. Games 12 (3), 62. https://doi.org/10.3390/g12030062.

San Martín, R., 2012. Event-related potential studies of outcome processing and feedback-guided learning. Front. Hum. Neurosci. 6, 304. https://doi.org/10.3389/fnhum.2012.00304.

Sambrook, T.D., Goslin, J., 2015. A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. Psychol. bull. 141 (1), 213–235. https://doi.org/10.1037/bul0000006.

Schulreich, S., Pfabigan, D.M., Derntl, B., Sailer, U., 2013. Fearless dominance and reduced feedback-related negativity amplitudes in a time-estimation task – Further neuroscientific evidence for dual-process models of psychopathy. Biol. Psychol. 93 (3), 352–363. https://doi.org/10.1016/j.biopsycho.2013.04.004.

Tsuchiya, N., Koch, C., 2005. Continuous flash suppression reduces negative afterimages. Nat. Neurosci. 8 (8), 1096–1101. https://doi.org/10.1038/nn1500.

Voytek, B., Kramer, M.A., Case, J., Lepage, K.Q., Tempesta, Z.R., Knight, R.T., Gazzaley, A., 2015. Age-related changes in 1/*f* neural electrophysiological noise. J. Neurosci. 35 (38), 13257–13265. https://doi.org/10.1523/JNEUROSCI.2332-14.2015.

Walsh, M.M., Anderson, J.R., 2011. Modulation of the feedback-related negativity by instruction and experience. Proc. Natl. Acad. Sci. 108 (47), 19048–19053. https://doi.org/10.1073/pnas.1117189108.

Wang, Y., Luo, Y.J., 2005. Standardization and assessment of college students' facial expression of emotion. Chin. J. Clin. Psychol. 13 (4), 3. https://doi.org/10.3969/j.issn.1005-3611.2005.04.006.

Yan, J., Yu, S., Mückschel, M., Colzato, L., Hommel, B., Beste, C., 2024. Aperiodic neural activity reflects metacontrol in task-switching. Sci. Rep. 14 (1), 24088. https://doi.org/10.1038/s41598-024-74867-7.

Zedelius, C.M., Veling, H., Custers, R., Bijleveld, E., Chiew, K.S., Aarts, H., 2014. A new perspective on human reward research: how consciously and unconsciously perceived reward information influences performance. Cogn. Affect. Behav. Neurosci. 14 (2), 493–508. https://doi.org/10.3758/s13415-013-0241-z.

Zhang, Y., Li, X., Qian, X., Zhou, X., 2012. Brain responses in evaluating feedback stimuli with a social dimension. Front. Hum. Neurosci. 6, 29. https://doi.org/10.3389/fnhum.2012.00029.

Zhang, C., Stock, A.-K., Mückschel, M., Hommel, B., Beste, C., 2023. Aperiodic neural activity reflects metacontrol. Cereb. Cortex 33 (12), 7941–7951. https://doi.org/10.1093/cercor/bhad089.