Title: Theta-Band Temporal Interference Stimulation Targeting the Dorsal Anterior Cingulate Cortex Modulates Prediction Error Encoding in Food Addiction

Authors: Yatong Wen[a,b,c], Mo Wang[d,e], Xinwen Dong[a], Mateusz Gola[f,g], Quanying Liu[d], Yiheng Tu[a,b*], Yonghui Li[a,b*]

[a] State Key Laboratory of Cognitive Science and Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing, China

[b] Department of Psychology, University of Chinese Academy of Sciences, Beijing, China

[c] Department of Neuroscience, Karolinska Institute, Stockholm 171 65, Sweden

[d] Department of Biomedical Engineering, Southern University of Science and Technology, Shenzhen 518055, China

[e] Department of Computer Science, University of Warwick, Coventry, United Kingdom

[f] Institute of Psychology, Polish Academy of Sciences, Warsaw, Poland

[g] Swartz Center for Computational Neuroscience, Institute for Neural Computations, University of California, San Diego, CA, USA

*Corresponding authors:

Yiheng Tu, State Key Laboratory of Cognitive Science and Mental Health, Institute of Psychology, Chinese Academy of Sciences, No.16 Lincui Road, Chaoyang District, Beijing 100101, China. Email: tuyh@psych.ac.cn

Yonghui Li, State Key Laboratory of Cognitive Science and Mental Health, Institute of Psychology, Chinese Academy of Sciences, No.16 Lincui Road, Chaoyang District, Beijing 100101, China. Email:

liyonghui@psych.ac.cn

Running title: tTIS Targeting dACC Enhances PE Sensitivity in FA

Keywords: temporal interference stimulation; dorsal anterior cingulate cortex; electroencephalography; feedback-related negativity; prediction error; food addiction

ABSTRACT

BACKGROUND: Food addiction is characterized by persistently elevated expectations of food reward, suggesting impairments in prediction error processing. The dorsal anterior cingulate cortex (dACC) encodes prediction errors via theta-band activity, but causal evidence in humans remains limited. Transcranial temporal interference stimulation (tTIS) offers a non-invasive approach to modulate deep cortical regions including the dACC.

METHODS: Thirty-four individuals with food addiction (15 males) were randomly assigned to active theta-band tTIS or sham stimulation. Participants completed a food incentive delay task in a satiated state while EEG was recorded pre- and post-stimulation. Feedback-related negativity (FRN) was analyzed at the mean ERP and single-trial levels using mixed-effects models linking FRN amplitudes to behavioral and model-based prediction errors; stimulus preceding negativity (SPN) was analyzed at the mean ERP level.

RESULTS: Trial-level mixed-effects models revealed robust prediction error encoding in FRN amplitudes during food reward feedback. Critically, tTIS altered FRN sensitivity to prediction errors, such that pre-to-post changes in FRN–prediction error coupling differed between the active and sham groups for both behavioral prediction errors ($\beta = 1.74$, $p = .049$) and model-based prediction errors ($\beta = 1.77$, $p = .048$). At the mean ERP level, FRN during food feedback became more negative after stimulation in the active group ($t_{16} = 2.75$, $p = .014$), whereas SPN was unchanged.

CONCLUSIONS: Theta-band tTIS targeting the dACC selectively enhances sensitivity to prediction errors during food reward processing, highlighting temporal interference

stimulation as a tool to causally modulate feedback-evaluation mechanisms in humans.

INTRODUCTION

Individuals with food addiction often show persistently elevated expectations of food reward, suggesting impairments in prediction error processing (1,2). Even under satiety, these individuals show reduced sensitivity to discrepancies between expected and actual food outcomes (1). Such prediction error–related impairments have been associated with aberrant dorsal anterior cingulate cortex (dACC) function (3) and blunted reward prediction error signaling in addiction-related conditions (4), potentially disrupting the recalibration of reward expectations and thereby promoting maladaptive food anticipation independent of hunger state.

The dACC plays a central role in encoding prediction errors during reward processing (5,6), primarily via theta-band oscillatory activity (7). Neurons within the dACC dynamically track expected outcomes and update neural responses upon receipt of feedback, with outcome discrepancies driving changes in firing patterns (8). Electrophysiological evidence further suggests that the dACC supports sustained error monitoring through synchronized theta rhythms (7). In addition, coordinated temporal coding across prefrontal regions may facilitate adaptive control and error processing (9). Consistent with this framework, inactivation of the dACC reduces task-related oscillatory activity in lateral prefrontal regions, underscoring the functional importance of dACC theta dynamics in prediction error processing (10).

Despite converging evidence linking dACC theta activity to prediction errors, direct causal evidence in humans remains limited. While invasive recordings during dACC Deep

brain stimulation (DBS) procedures provide rare human electrophysiological evidence for prediction error–related signaling in the dACC, DBS is invasive, restricted to clinical populations, and does not allow frequency-specific modulation of neural activity (11). Non-invasive neuromodulation approaches, such as transcranial alternating current stimulation (tACS), offer an alternative means to probe the causal role of cortical oscillations (12,13). Although rarely applied to prediction error processing, frontal theta-band tACS has been reported to modulate dACC-related network properties (14). Nevertheless, conventional non-invasive stimulation techniques primarily affect superficial cortical regions and are poorly suited for directly targeting deeper structures such as the dACC. Accordingly, a technique capable of frequency-specific stimulation while directly targeting deep structures such as the dACC is critically needed.

Transcranial Temporal Interference Stimulation (tTIS) is a recently developed non-invasive technique that enables modulation of deep brain regions by generating low-frequency envelopes through the interaction of two high-frequency alternating currents (15). By adjusting the frequency difference between these currents, tTIS can induce oscillatory modulation at a targeted frequency while minimizing stimulation of overlying cortex, achieving effects analogous to frequency-specific stimulation such as tACS (16). Recent studies demonstrate that tTIS can modulate activity in deep and subcortical regions, including the striatum and hippocampus (17,18), highlighting its potential utility for causal interrogation of dACC theta dynamics.

In summary, the present study investigated whether theta-band tTIS targeting the dACC

modulates neural sensitivity to prediction errors during food reward processing in individuals with food addiction. To address this question, we examined feedback-stage neural activity at both the single-trial level and the mean event-related potential level, focusing on how outcome-related discrepancies are encoded following reward feedback. Given that food addiction is characterized by persistently elevated reward expectations, we hypothesized that stimulation-related modulation would be most pronounced for unfavorable (negative prediction-error) outcomes.

METHODS AND MATERIALS

Participants

The study was approved by the Ethics Committee of the Institute of Psychology, Chinese Academy of Sciences (No. H23113). Thirty-four right-handed participants meeting criteria for food addiction (15 males) were recruited. All participants had normal or corrected-to-normal vision and provided written informed consent in accordance with the Declaration of Helsinki.

Inclusion criteria required a Yale Food Addiction Scale (YFAS) symptom score ≥ 3 and a body mass index (BMI) ≥ 25, consistent with prior work (19–21). BMI was verified by measuring height and weight prior to the experiment. Participants were randomly assigned to the active tTIS group (n = 17) or the sham stimulation group (n = 17). All participants reported being in a satiated state at the time of testing, with hunger ratings below 5 on a 10-point scale, and were instructed to consume a substantial lunch before the experiment. Participants received a base payment of ¥200–¥230, with an additional performance-based bonus of up to ¥15.

Study design and procedure

The experiment comprised three sequential phases: pretest, stimulation, and posttest. The pretest and posttest each included 150 trials during which EEG was recorded. No EEG data were collected during stimulation. Before the main experiment, participants completed a practice session and were required to achieve at least 80% accuracy. The overall experimental timeline is illustrated in Figure 1, and the task structure is shown in Figure 2.

The study used a mixed 2 × 2 design with Group (active tTIS vs. sham stimulation) as a between-subject factor and Phase (pre- vs. post-stimulation) as a within-subject factor. All trials involved food rewards depicted by food images matched on arousal, pleasantness, and familiarity (Supplementary Section 1 and Table S1).

[Insert Figure 1 here]

Incentive Delay Task

Reward processing was assessed using a modified Incentive Delay Task (IDT) adapted from Knutson et al. (22) focusing on food reward outcomes. Each trial began with a cue indicating food reward probability (high: 80%; low: 20%), followed by a response phase in which participants indicated their reward expectation via button press. After a 2000-ms delay, reward feedback was presented for 1500 ms. Inter-trial intervals varied between 350 and 650 ms. Behavioral outcomes included reaction times and reward prediction rates under high- and low-probability conditions.

[Insert Figure 2 here]

Electroencephalogram (EEG) recording and processing

Continuous EEG was recorded using a 64-channel system (Neuracle NeuSen.W 64;

Neuracle, Changzhou, China), referenced to CPz with AFz as ground. Signals were sampled at 1000 Hz, electrode impedances were kept below 10 kΩ, and data were rereferenced offline to the average of the bilateral mastoids (M1/M2). Analyses focused on two ERP components: feedback-related negativity (FRN), indexing feedback evaluation, and stimulus preceding negativity (SPN), indexing anticipatory processing.

EEG preprocessing was conducted in EEGLAB/ERPLAB using custom MATLAB scripts. Data were filtered (FRN: 0.1–30 Hz; SPN: low-pass 30 Hz), epoched relative to feedback onset, and cleaned using automated rejection, manual inspection, and independent component analysis. SPN epochs were extracted from −2000 to 500 ms (baseline: −2000 to −1800 ms), and FRN epochs from −200 to 1000 ms (baseline: −200 to 0 ms). Single-trial amplitudes were extracted for trial-level analyses, and grand-average waveforms/topographies were computed for visualization.

Transcranial Temporal Interference Stimulation

tTIS was delivered using an interferential neuromodulation system (Soterix Medical Inc.) targeting the left dACC. Ag/AgCl electrodes (1.2 cm diameter; contact area = 1.13 cm^2) were placed according to an optimized unilateral montage (23) (Supplementary Section 2). Target coordinates (MNI: −5, 34, 30) were derived from prior neuromodulation studies (23) and verified using FSL and MNI–Talairach conversion tools (https://bioimagesuiteweb.github.io/bisweb-manual/tools/mni2tal.html). Four electrodes were used (AF3 positive, Fz negative, CP3 positive, F7 negative).

Finite-element simulations estimated the temporal interference envelope distribution and

confirmed field intensity in the dACC target (Figure 3). Stimulation used two high-frequency currents (2000 and 2005 Hz) producing a 5-Hz theta-band envelope. Current intensity was 1.6 mA (peak-to-baseline) per electrode for 20 minutes. Sham stimulation consisted of 30-s ramp-up/down without ongoing stimulation.

[Insert Figure 3 here]

Statistical analysis

Trial-level analysis of prediction error

To assess trial-level neural sensitivity to behavioral prediction errors, single-trial FRN amplitudes were analyzed using linear mixed-effects (LME) models implemented in MATLAB (fitlme). Behavioral prediction errors were coded as −1 (worse-than-expected), 0 (expected), or +1 (better-than-expected). The primary model included PE, Group (active vs. sham), Phase (pre vs. post), and all interactions, with a subject-specific random intercept: FRN_ amplitude ~ PE × Group × Phase + (1|subject). Fixed effects were evaluated using Wald *z* tests with normal approximation (two-tailed). Where higher-order interactions involving PE were observed, follow-up analyses were conducted to aid interpretation. First, subject-level FRN–PE slopes were estimated within each participant and phase and summarized by group to provide a descriptive visualization of Group × Phase patterns in prediction-error sensitivity. Second, simple-effects contrasts tested the Group × Phase contrast separately at each PE level (PE = −1, 0, +1) to localize interaction effects across prediction-error categories.

Computational estimation of trial-level prediction errors

To obtain continuous estimates of internal prediction errors, model-derived prediction errors were computed using a Rescorla–Wagner reinforcement learning model (24) fitted to trial-wise behavioral data (Supplementary Section 3). These prediction errors (PE_RW) were entered into an analogous LME model: FRN_ amplitude ~ PE_RW × Group × Phase + (1|subject). Inference was based on Wald *z* tests (two-tailed). Where higher-order interactions involving PE_RW were observed, follow-up analyses were conducted to facilitate interpretation. First, subject-level FRN–PE_RW slopes were estimated within each participant and phase and summarized by group to provide a descriptive visualization of Group × Phase patterns in model-based prediction-error sensitivity. Second, simple-effects contrasts tested the Group × Phase contrast within three PE_RW bins (PE_RW < 0, PE_RW = 0, PE_RW > 0) to localize interaction effects across prediction-error sign.

ERP and behavioral analyses

Mean ERP and behavioral outcomes were analyzed using repeated-measures ANOVAs with Group (active vs. sham) as a between-subject factor and Phase (pre vs. post) as a within-subject factor. FRN amplitude was quantified as peak negativity within the 250–350 ms time window at FCz; difference waves (no reward − reward) were used for waveform visualization. SPN amplitude was quantified as mean voltage within the −200 to 0 ms window preceding feedback onset. Based on the study's a priori focus on unfavorable outcomes in food addiction, complementary condition-level ERP analyses focused on negative prediction errors. Greenhouse–Geisser corrections were applied where appropriate, with statistical significance set at $p < .05$.

Sample size

Sample size was determined a priori using G*Power (version 3.1) (25) for a repeated-measures ANOVA with a within–between interaction (Group × Phase). Given the limited availability of directly comparable effect-size estimates in prior EEG and neuromodulation studies (2,14), a conservative medium effect size ($f = 0.25$) was assumed, with an alpha level of 0.05, power of 0.80, two groups, and two repeated measurements with a correlation of 0.50 ($\varepsilon = 1$). The required total sample was 34 participants (17 per group).

RESULTS

Demographic characteristics

A total of 34 participants with food addiction were recruited and randomly assigned to the active tTIS group (mean age = 26.29 years, SD = 5.84) or the sham stimulation group (mean age = 24.35 years, SD = 2.98). The two groups did not differ in Yale Food Addiction Scale scores, body mass index, age, education level, or household income (all $p > .05$; Table S2). All participants were tested in a satiated state, with hunger ratings below 3 and no group differences in hunger levels ($p > .05$). Mild transient sensations (e.g., headache or tingling) were more frequently reported in the active group, but no participant reported lasting discomfort or was able to correctly identify stimulation condition (Table S3).

Trial-level encoding of behavioral prediction errors

Single-trial FRN amplitudes scaled significantly with PE_behavior ($\beta = 2.66$, $SE = 0.44$, $z = 5.99$, $p = 2.06 \times 10^{-9}$), indicating robust trial-level encoding of prediction errors during reward feedback. Significant PE_behavior × Phase ($\beta = -1.69$, $SE = 0.61$, $z = -2.76$, p

= .006) and PE_behavior × Group interactions ($\beta = -1.51$, $SE = 0.63$, $z = -2.39$, $p = .017$) were observed. Importantly, a significant Group × Phase × PE_behavior interaction ($\beta = 1.74$, $SE = 0.89$, $z = 1.97$, $p = .049$) indicated that pre-to-post changes in prediction-error sensitivity differed between the active and sham groups.

Subject-level FRN-PE slopes are shown in Figure 4 as descriptive indices of individual sensitivity. Slopes showed a numerically larger pre-to-post increase in the active group (Δ slope = 1.72, $p = .091$, $d = .45$), with no apparent change in the sham group (Δ slope = −0.04, $p = .937$, $d = .02$). Follow-up simple-effects contrasts indicated that the Group × Phase effect was most pronounced for unfavorable outcomes (PE_behavior = −1; $\beta = -3.45$, $SE = 0.97$, $z = -3.56$, $p = 3.70 \times 10^{-4}$; Table S4). Figure S1 illustrates the trial-level FRN–PE relationship.

[Insert Figure 4 here]

Trial-level encoding of model-based prediction errors

FRN responses were further examined in relation to model-based prediction errors estimated from a Rescorla–Wagner learning model (PE_RW; Supplementary Section 3). Single-trial FRN amplitudes scaled significantly with PE_RW ($\beta = 2.94$, $SE = 0.45$, $z = 6.53$, $p = 6.62 \times 10^{-11}$), indicating robust encoding of model-derived prediction errors. Significant PE_RW × group ($\beta = -1.42$, $SE = 0.64$, $z = -2.21$, $p = .027$) and PE_RW × phase interaction ($\beta = -1.04$, $SE = 0.62$, $z = -1.69$, $p = .092$) were found. Critically, a significant Group × Phase × PE_RW interaction ($\beta = 1.77$, $SE = 0.90$, $z = 1.98$, $p = .048$) indicated differential pre-to-post changes in prediction-error sensitivity between groups.

Subject-level FRN–PE_RW slopes are shown in Figure 5 for descriptive visualization.

Slopes showed a numerically increasing trend from pre to post in the active group (Δ slope = 1.05, p = .260, d = .29) and a decreasing trend in the sham group (Δ slope = −0.71, p = .238, d = .30). Follow-up contrasts indicated that the Group × Phase effect was most pronounced when PE_RW values were negative (PE_RW < 0; $\beta = -2.46$, $SE = 0.56$, $z = -4.36$, $p = 1.29 \times 10^{-5}$; Table S5). Figure S2 illustrates the trial-level FRN–PE_RW relationship.

[Insert Figure 5 here]

ERP indices of feedback and anticipation

FRN

To complement the trial-level mixed-effects analyses, condition-averaged FRN amplitudes were additionally examined for negative prediction errors (PE_behavior = −1; PE_RW < 0). This follow-up was guided by the trial-level localization analyses indicating that stimulation-related modulation was most pronounced for unfavorable outcomes (Tables S4–S5). These analyses yielded a consistent pattern of increased post-stimulation FRN negativity in the active group, with no corresponding change in the sham group (Figures S3–S4; Supplementary Section 6 for full statistics).

Figure 6 shows grand-average FRN waveforms and scalp topographies elicited by food reward outcomes in the active and sham groups before and after stimulation. FRN was observed as a fronto-central negative deflection peaking around 300 ms after feedback. Baseline FRN amplitudes did not differ between groups during the pre-stimulation phase ($t_{32} = -0.35$, p = .730, d = .12). A significant main effect of phase was observed ($F_{1,32} = 6.13$, p = .019, η_p^2 = .16), along with a trend-level Group × Phase interaction ($F_{1,32} = 3.86$, p = .058,

η_p^2 = .12). Follow-up analyses indicated that FRN amplitudes became significantly more negative after stimulation in the active group (t_{16} = 2.75, p = .014, d = .67), whereas no significant change was observed in the sham group (t_{16} = 0.44, p = .67, d = .11) (Figure 8A).

[Insert Figure 6 here]

SPN

SPN waveforms and scalp distributions are shown in Figure 7. Baseline SPN amplitudes were comparable between groups (t_{32} = 0.60, p = .55, d = .21). No main effects of Group or Phase, nor a Group × Phase interaction, were observed (all p > .80). Quantitative SPN results are summarized in Figure 8B.

[Insert Figure 7 here]

[Insert Figure 8 here]

Behavioral data

Behavioral performance during the incentive delay task was examined to determine whether stimulation-related neural effects were accompanied by overt behavioral changes. Reaction times showed a significant main effect of Phase ($F_{1,32}$ = 10.73, p = .003, η_p^2 = .25), with faster responses post-stimulation, but no main effect of Group or Group × Phase interaction (both p > .05). Reward prediction rates did not differ as a function of Group, Phase, or their interaction for either high- or low-probability cues (all p > .05).

DISCUSSION

To our knowledge, the present study provides the first evidence that theta-band tTIS targeting the dACC can causally modulate single-trial prediction error encoding during food

reward processing in individuals with food addiction. Using a randomized, sham-controlled pre–post design, we show that tTIS altered trial-level coupling between FRN amplitude and both behavioral and model-derived prediction errors. Mean ERP analyses provided converging support, showing increased FRN negativity after stimulation in the active group, while SPN was unchanged.

Theta-band tTIS modulates feedback-stage prediction error encoding

The primary contribution of the present study lies in the trial-level analyses of prediction error encoding. Single-trial FRN amplitudes reliably scaled with both categorical behavioral prediction errors and continuous prediction errors estimated from a Rescorla–Wagner learning model, demonstrating robust feedback-stage encoding of outcome discrepancies (2,26,27). Critically, theta-band tTIS modulated this encoding, as evidenced by significant Group × Phase × prediction error interactions for both behavioral prediction errors and model-based prediction errors. These interactions indicate that stimulation altered how FRN responses tracked trial-level prediction errors from pre to post, and that this change differed between the active and sham groups.

Follow-up contrasts further localized this modulation, indicating that stimulation-related changes were most pronounced for unfavorable outcomes (PE_behavior = −1; PE_RW < 0). This pattern is consistent with the established sensitivity of the FRN to negative outcomes during feedback evaluation (2,27,28). Together, these findings suggest that theta-band tTIS altered the sensitivity (gain) of prediction error encoding (i.e., the degree to which FRN responses tracked trial-level prediction errors) rather than producing a nonspecific change in

feedback-evoked activity.

Notably, subject-level FRN–prediction error slopes (Figures 4–5) provided descriptive support for this interpretation. Although pre-to-post slope changes did not reach significance at the subject level, slopes showed a numerically larger increase following stimulation in the active group compared with sham, consistent with an enhancement of prediction error sensitivity at the single-trial level. Importantly, the key inference is derived from the trial-level mixed-effects models, which leverage the full dataset and are statistically better powered to detect interaction effects than summary slope tests.

dACC theta modulation as a candidate mechanism

The observed effects are consistent with the established role of the dACC in signaling outcome discrepancies and guiding adaptive control through theta-band oscillatory dynamics (5,7). Converging evidence indicates that dACC theta oscillations coordinate feedback evaluation and error monitoring, and that coordinated theta-band activity across medial frontal cortex and lateral prefrontal regions supports the implementation of adaptive adjustments following unfavorable outcomes (7,9,10). Given that the FRN is closely linked to fronto-central theta activity and reflects the integration of feedback with internal expectations (2,26–29), the selective modulation of FRN–prediction error coupling observed here provides mechanistic support for theta-frequency dACC involvement in prediction error encoding.

Consistent with the trial-level findings, theta-band tTIS appears to have affected the sensitivity with which prediction errors were represented at the single-trial level, rather than simply shifting mean feedback-evoked activity. One plausible interpretation is that tTIS

increased the responsivity (gain) of dACC-centered evaluative circuits to deviations from expected outcomes, thereby strengthening the mapping between prediction error magnitude and FRN amplitude. The localization of stimulation effects to unfavorable prediction errors is consistent with theories proposing that medial frontal theta activity is preferentially engaged by negative outcomes that require behavioral updating and enhanced cognitive control (2,7,28,29).

Although the stimulation protocol was designed to deliver a theta-frequency envelope within the dACC, the present EEG data do not allow dissociation between direct entrainment of local theta oscillations and indirect modulation via distributed cingulo-frontal networks. Given the dense anatomical and functional connectivity of the dACC, it is likely that stimulation influenced a broader feedback-evaluation network rather than a single isolated node. Notably, although the study did not include a healthy control group, the randomized sham-controlled pre–post design allows causal inference about stimulation-related modulation of prediction error encoding within the target population.

Future work combining tTIS with source-resolved EEG or connectivity analyses may help determine whether the observed enhancement in prediction error sensitivity reflects increased medial frontal theta power, strengthened theta synchronization with prefrontal control regions, or both. More broadly, these results support the view that theta-band dynamics within dACC-centered networks play a causal role in shaping the neural gain of feedback-based prediction error signals (2,5,7,9,10,26–29).

Dissociation between feedback evaluation and anticipatory processing

Complementing the feedback-stage effects on FRN-based prediction error encoding, theta-band tTIS exerted stage-specific modulation of reward processing. While stimulation reliably altered feedback-stage FRN–prediction error coupling, anticipatory activity indexed by SPN remained unchanged. SPN has been linked to anticipatory attention, motivational engagement, and preparatory processes during the pre-feedback interval (30–32). This dissociation suggests that theta-band modulation of dACC-centered networks may preferentially influence evaluative computations engaged at outcome delivery and feedback-based updating, rather than anticipatory preparatory processes (2,26–28,33,34).

Notably, this stage-specific modulation aligns with our prior findings of an imbalanced anticipatory–feedback profile in food addiction under satiation, characterized by elevated anticipatory signals (enhanced SPN) but reduced feedback-based updating (attenuated FRN) compared with healthy controls (1). In this context, the present results suggest that a single session of theta-band tTIS may acutely modulate feedback updating mechanisms supporting prediction error signaling, without substantially altering elevated anticipatory expectancy signals. Future work should test whether repeated stimulation or alternative protocols are required to modulate anticipatory components such as SPN.

More broadly, the absence of stimulation effects on SPN supports a dissociation between anticipatory and evaluative components of reward processing, consistent with models in which prediction error computations arise during feedback evaluation and guide subsequent updating of reward expectations, whereas anticipatory signals reflect preparatory processes supported by partially distinct neural mechanisms (2,26–28,30–34).

Implications for non-invasive modulation of deep reward-monitoring circuits

Transcranial temporal interference stimulation has been proposed as a promising non-invasive approach for modulating deep cortical structures that are difficult to access using conventional stimulation techniques (16,35). Although much of the early tTIS literature has relied on computational modeling or indirect behavioral readouts, recent human studies have begun to provide evidence for functional modulation of targeted circuits (17,18).

By showing that theta-band tTIS altered trial-level FRN–prediction error coupling, the present findings extend tTIS research beyond mean ERP effects and highlight its potential to causally probe reinforcement-learning computations in deep medial frontal circuits. More broadly, these results support the use of tTIS as a mechanistic tool for causal interrogation of deep reward-monitoring circuits in humans, and motivate future work to test whether repeated stimulation can produce more sustained modulation of prediction error signaling and downstream behavioral adaptation.

Limitations and future directions

Several limitations should be noted. First, although the sample size was sufficient to detect the primary interaction effects and was further supported by trial-level mixed-effects analyses, power may have been limited for detecting subtle between-group differences in subject-level indices (e.g., individual FRN–prediction error slopes). Trial-level mixed-effects models improve sensitivity by leveraging all available observations (36). Second, only short-term effects of a single tTIS session were assessed. Future studies should test whether repeated stimulation produces sustained changes in prediction error sensitivity and related behavioral

adaptation. Finally, reward processing depends on distributed circuits beyond the dACC, including the striatum and prefrontal cortex (37). Combining tTIS with multimodal imaging or connectivity analyses may help clarify network-level mechanisms of dACC theta modulation.

CONCLUSIONS

Theta-band tTIS targeting the dACC causally modulated feedback-stage prediction error encoding during food reward processing in individuals with food addiction. Trial-level mixed-effects models showed altered single-trial FRN–prediction error coupling for both behavioral and model-derived prediction errors, consistent with enhanced neural sensitivity to unfavorable outcomes. tTIS may therefore serve as a non-invasive tool for mechanistically probing deep medial frontal feedback-evaluation circuits in humans.

ACKNOWLEDGMENTS AND DISCLOSURES

This work was supported by the Chinese National Programs for Brain Science and Brain-like Intelligence Technology (Grant No. 2021ZD0202104) and the International Cooperation and Exchange Program of the National Natural Science Foundation of China (Grant No. 32161133004).

Y.T.W. had full access to all study data and takes responsibility for data integrity and analysis accuracy. Y.T.W. is the first author and contributed to study conception and design, data acquisition and analysis, and manuscript drafting. Y.H.L. and Y.H.T. obtained funding and provided supervision and critical revision of the manuscript. M.W. contributed to brain region modeling and stimulation methodology. Q.Y.L. provided methodological support and contributed to interpretation of the results. X.W.D. and M.G. contributed to interpretation of

findings and critical revision of the manuscript. All authors approved the final manuscript.

A preprint version of this manuscript has been posted on the bioRxiv preprint server and is disclosed in accordance with journal policy. The authors report no biomedical financial interests or potential conflicts of interest.

REFERENCES

1. Wen Y, Zeng Y, Zhou X, Gola M, Dong X, Li Y (2026): Satiety modulates reward processing in food addiction: Evidence from event-related potentials. *Appetite* 219: 108395.
2. Holroyd CB, Coles MGH (2002): The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review* 109: 679–709.
3. Fitzgerald KD, Welsh RC, Gehring WJ, Abelson JL, Himle JA, Liberzon I, Taylor SF (2005): Error-related hyperactivity of the anterior cingulate cortex in obsessive-compulsive disorder. *Biological Psychiatry* 57: 287–294.
4. Lei W, Liu K, Chen G, Tolomeo S, Liu C, Peng Z, *et al.* (2022): Blunted reward prediction error signals in internet gaming disorder. *Psychol Med* 52: 2124–2133.
5. Warren CM, Hyman JM, Seamans JK, Holroyd CB (2015): Feedback-related negativity observed in rodent anterior cingulate cortex. *Journal of Physiology-Paris* 109: 87–94.
6. Ide JS, Shenoy P, Yu AJ, Li CR (2013): Bayesian Prediction and Evaluation in the Anterior Cingulate Cortex. *J Neurosci* 33: 2039–2047.
7. Womelsdorf T, Johnston K, Vinck M, Everling S (2010): Theta-activity in anterior cingulate cortex predicts task rules and their adjustments following errors. *Proc Natl Acad Sci USA*

107: 5248–5253.

8. Hyman JM, Holroyd CB, Seamans JK (2017): A Novel Neural Prediction Error Found in Anterior Cingulate Cortex Ensembles. *Neuron* 95: 447-456.e3.

9. Smith EH, Horga G, Yates MJ, Mikell CB, Banks GP, Pathak YJ, *et al.* (2019): Widespread temporal coding of cognitive control in the human prefrontal cortex. *Nat Neurosci* 22: 1883–1891.

10. Ma L, Chan JL, Johnston K, Lomber SG, Everling S (2019): Macaque anterior cingulate cortex deactivation impairs performance and alters lateral prefrontal oscillatory activities in a rule-switching task. *PLoS Biol* 17: e3000045.

11. Weiss AR, Gillies MJ, Philiastides MG, Apps MA, Whittington MA, FitzGerald JJ, *et al.* (2018): Dorsal Anterior Cingulate Cortices Differentially Lateralize Prediction Errors and Outcome Valence in a Decision-Making Task. *Front Hum Neurosci* 12: 203.

12. Schutter DJLG (2014): Syncing your brain: electric currents to enhance cognition. *Trends in Cognitive Sciences* 18: 331–333.

13. Kanai R, Paulus W, Walsh V (2010): Transcranial alternating current stimulation (tACS) modulates cortical excitability as assessed by TMS-induced phosphene thresholds. *Clinical Neurophysiology* 121: 1551–1554.

14. Onoda K, Kawagoe T, Zheng H, Yamaguchi S (2017): Theta band transcranial alternating current stimulations modulates network behavior of dorsal anterior cingulate cortex. *Sci Rep* 7: 3607.

15. Guo W, He Y, Zhang W, Sun Y, Wang J, Liu S, Ming D (2023): A novel non-invasive brain

stimulation technique: “Temporally interfering electrical stimulation.” *Front Neurosci* 17: 1092539.

16. Mirzakhalili E, Barra B, Capogrosso M, Lempka SF (2020): Biophysics of Temporal Interference Stimulation. *Cell Syst* 11: 557-572.e5.
17. Violante IR, Alania K, Cassarà AM, Neufeld E, Acerbo E, Carron R, *et al.* (2023): Non-invasive temporal interference electrical stimulation of the human hippocampus. *Nat Neurosci* 26: 1994–2004.
18. Wessel MJ, Beanato E, Popa T, Windel F, Vassiliadis P, Menoud P, *et al.* (2023): Noninvasive theta-burst stimulation of the human striatum enhances striatal activity and motor skill learning. *Nat Neurosci* 26: 2005–2016.
19. Gearhardt AN, Corbin WR, Brownell KD (2009): Preliminary validation of the Yale Food Addiction Scale. *Appetite* 52: 430–436.
20. Gearhardt AN, Corbin WR, Brownell KD (2016): Development of the Yale Food Addiction Scale Version 2.0. *Psychol Addict Behav* 30: 113–121.
21. Gearhardt AN, Boswell RG, White MA (2014): The association of “food addiction” with disordered eating and body mass index. *Eating Behaviors* 15: 427–433.
22. Knutson B, Fong GW, Adams CM, Varner JL, Hommer D (2001): Dissociation of reward anticipation and outcome with event-related fMRI: *Neuroreport* 12: 3683–3687.
23. Wang M, Lou K, Liu Z, Wei P, Liu Q (2023): Multi-objective optimization via evolutionary algorithm (MOVEA) for high-definition transcranial electrical stimulation of the human brain. *Neuroimage* 280: 120331.

24. Rescorla RA, Wagner AR (1972): A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: Black AH, Prokasy WF, editors. *Classical Conditioning II: Current Research and Theory*. New York: Appleton-Century-Crofts, pp 64–99.

25. Faul F, Erdfelder E, Lang A-G, Buchner A (2007): G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods* 39: 175–191.

26. Sambrook TD, Goslin J (2015): A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychological Bulletin* 141: 213–235.

27. Chase HW, 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: 936–946.

28. Walsh MM, Anderson JR (2012): Learning from experience: Event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience & Biobehavioral Reviews* 36: 1870–1884.

29. Cavanagh JF, Frank MJ (2014): Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences* 18: 414–421.

30. Donkers FCL, Nieuwenhuis S, Van Boxtel GJM (2005): Mediofrontal negativities in the absence of responding. *Cognitive Brain Research* 25: 777–787.

31. Brunia CHM, Hackley SA, Van Boxtel GJM, Kotani Y, Ohgami Y (2011): Waiting to perceive: Reward or punishment? *Clinical Neurophysiology* 122: 858–868.

32. Ohgami Y, Kotani Y, Hiraku S, Aihara Y, Ishii M (2004): Effects of reward and stimulus modality on stimulus-preceding negativity. *Psychophysiology* 41: 729–738.

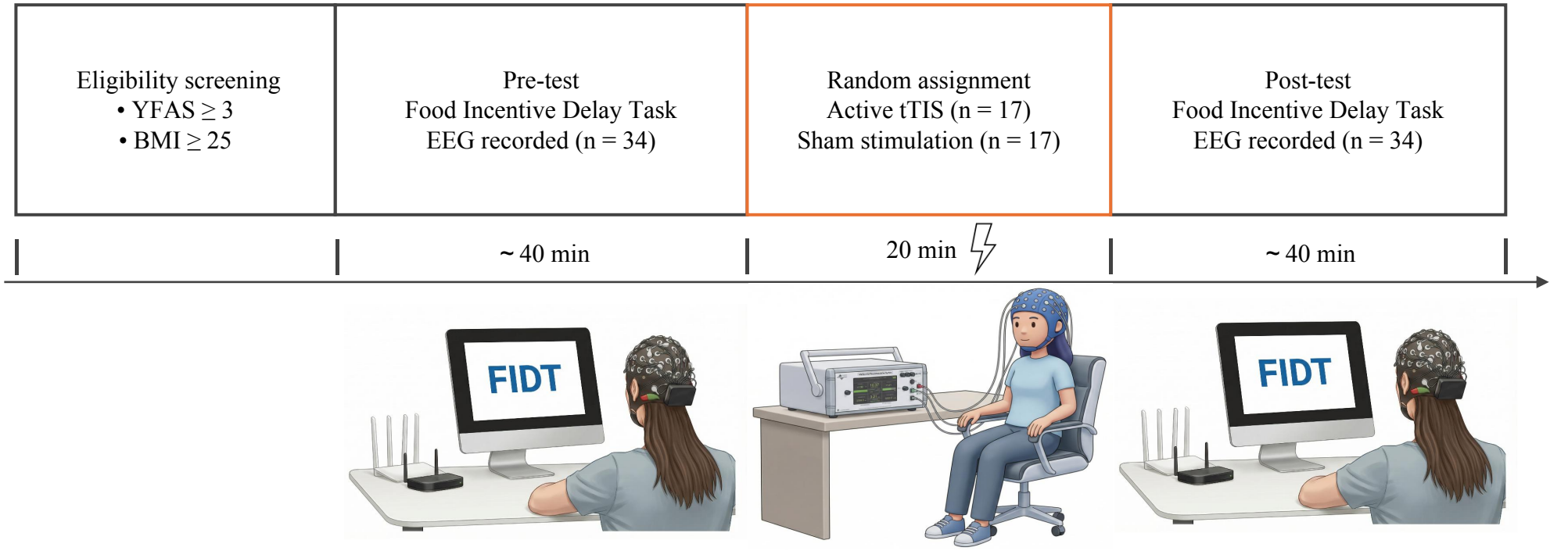
33. Reinhart RMG, Woodman GF (2014): Causal Control of Medial–Frontal Cortex Governs Electrophysiological and Behavioral Indices of Performance Monitoring and Learning. *J Neurosci* 34: 4214–4227.

34. Holroyd CB, Nieuwenhuis S, Yeung N, Nystrom L, Mars RB, Coles MGH, Cohen JD (2004): Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nat Neurosci* 7: 497–498.

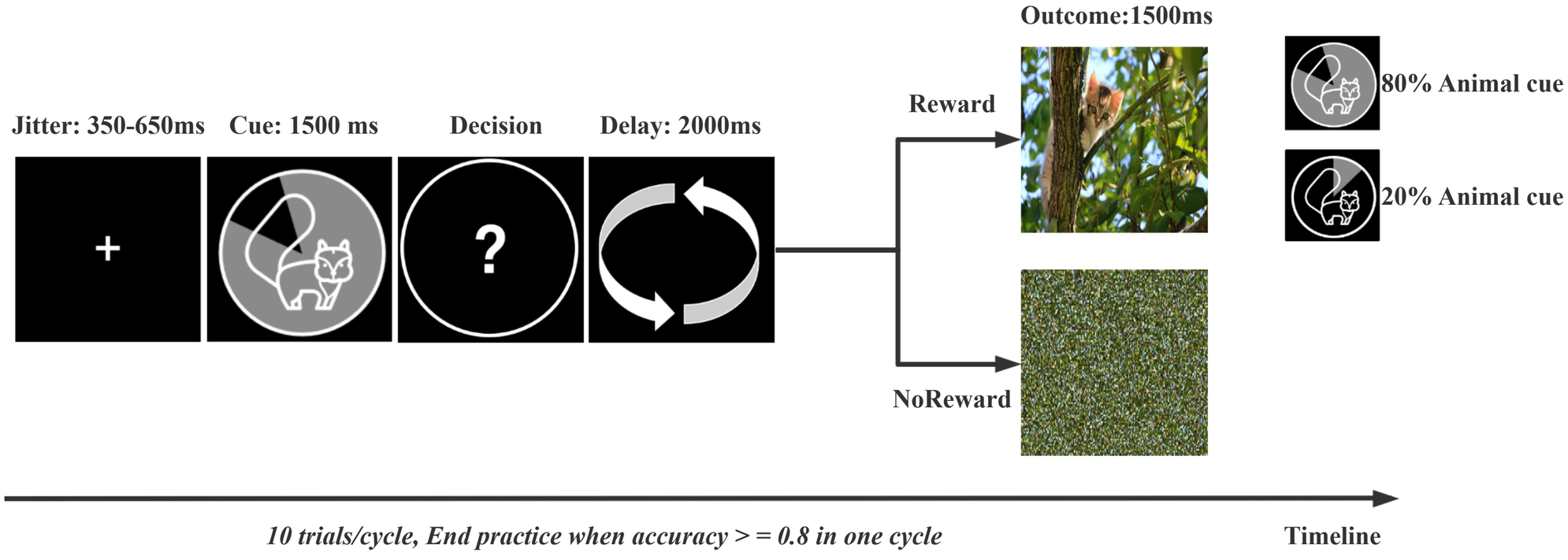
35. Grossman N, Bono D, Dedic N, Kodandaramaiah SB, Rudenko A, Suk H-J, *et al.* (2017): Noninvasive Deep Brain Stimulation via Temporally Interfering Electric Fields. *Cell* 169: 1029-1041.e16.

36. Baayen RH, Davidson DJ, Bates DM (2008): Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* 59: 390–412.

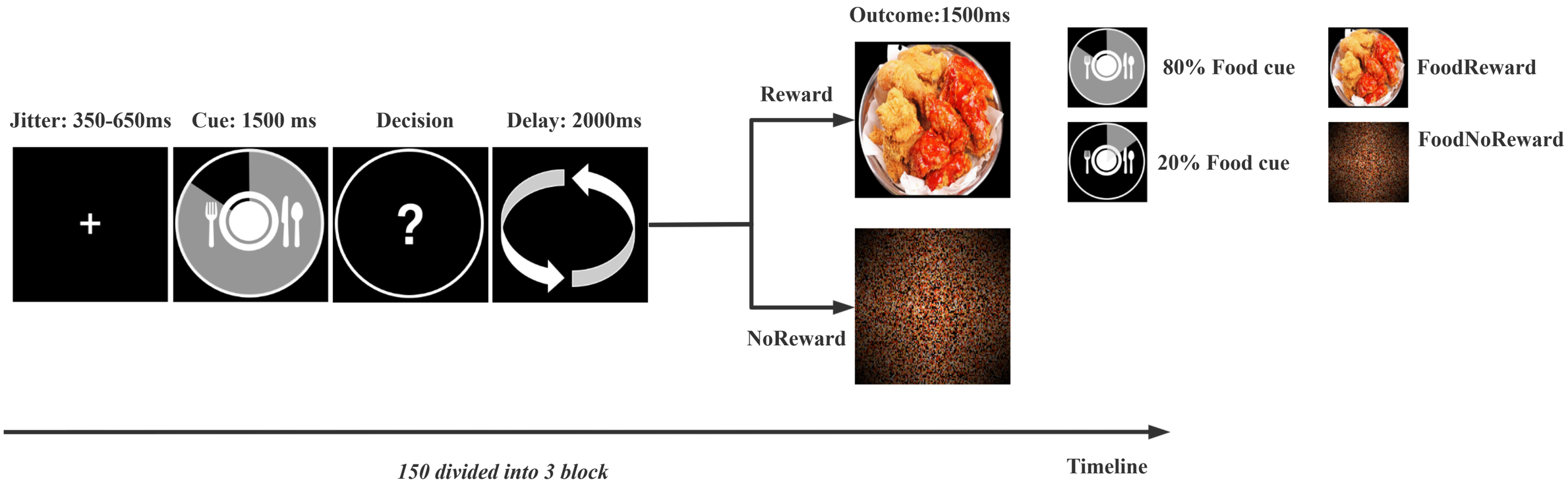
37. Haber SN, Knutson B (2010): The Reward Circuit: Linking Primate Anatomy and Human Imaging. *Neuropsychopharmacol* 35: 4–26.

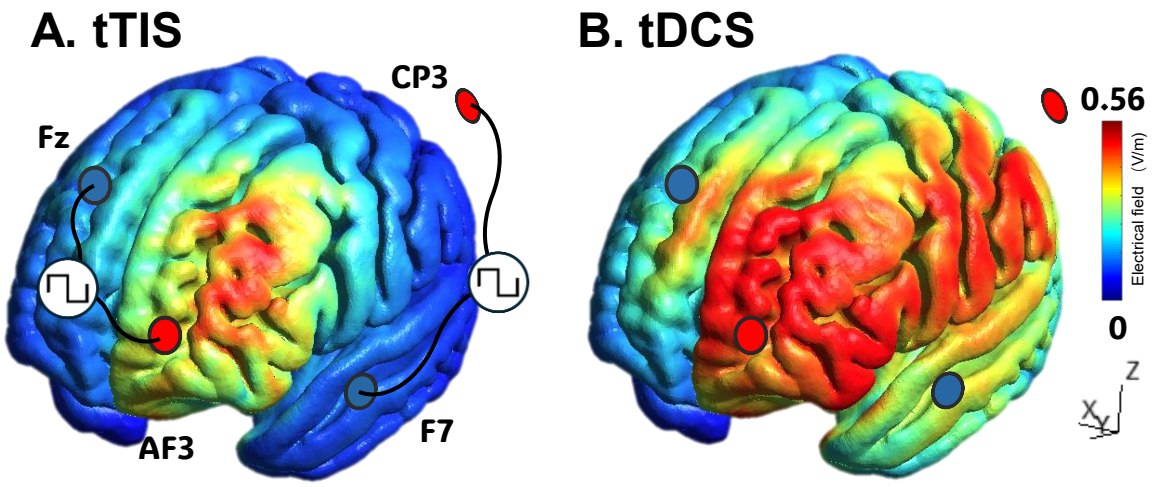

Eligibility screening
• YFAS ≥ 3
• BMI ≥ 25
Pre-test
Food Incentive Delay Task
EEG recorded (n = 34)
Random assignment
Active tTIS (n = 17)
Sham stimulation (n = 17)
Post-test
Food Incentive Delay Task
EEG recorded (n = 34)
~ 40 min
20 min
~ 40 min
FIDT
FIDT

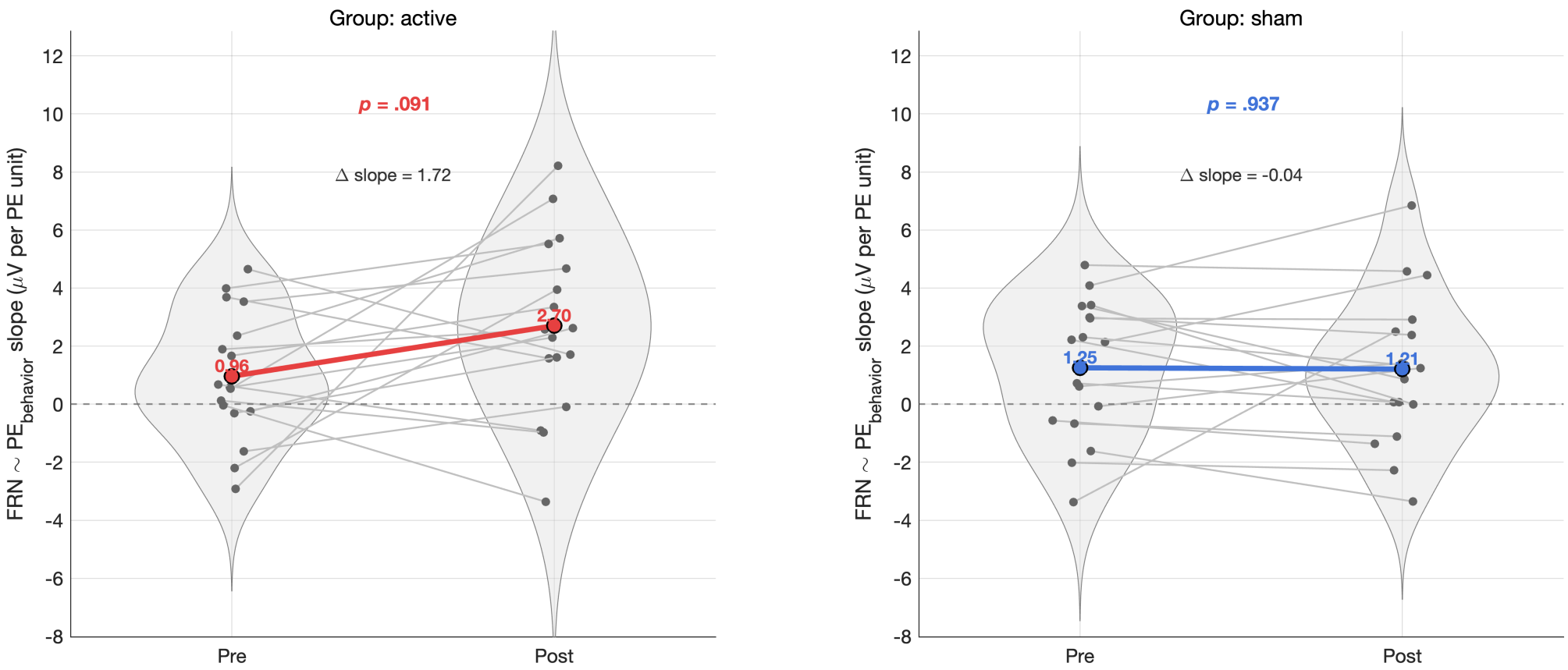
A: Practice

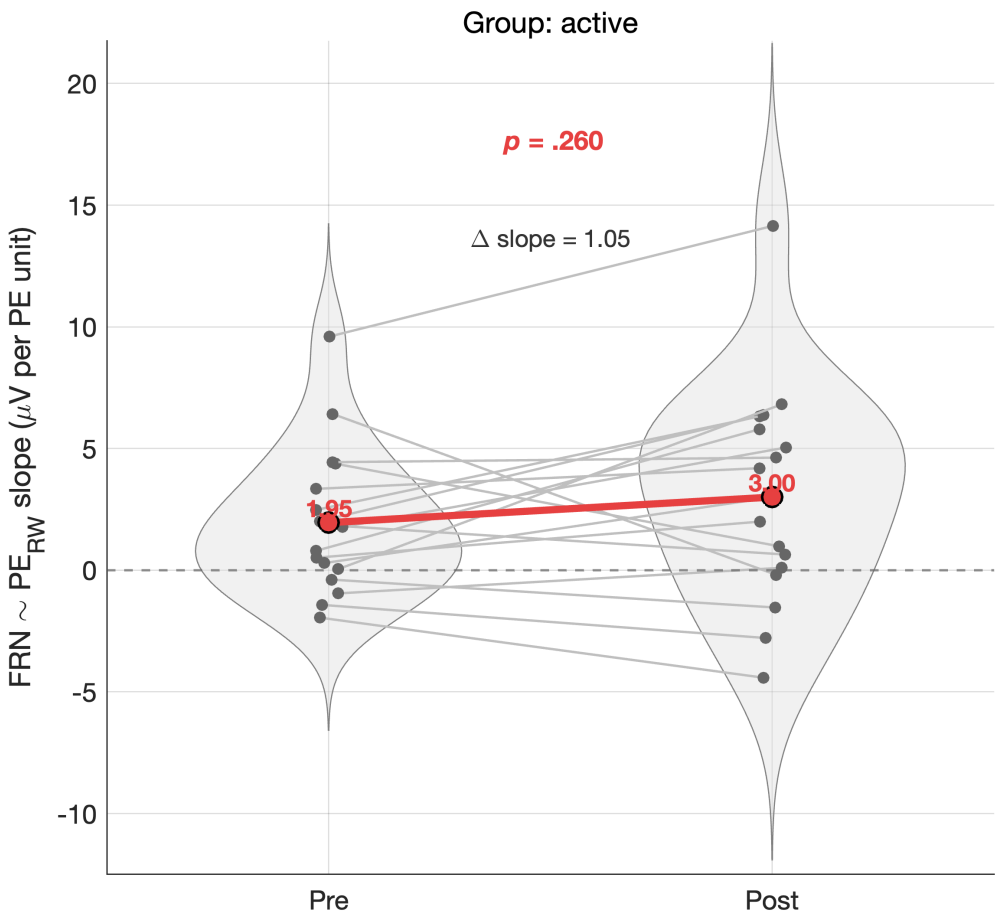


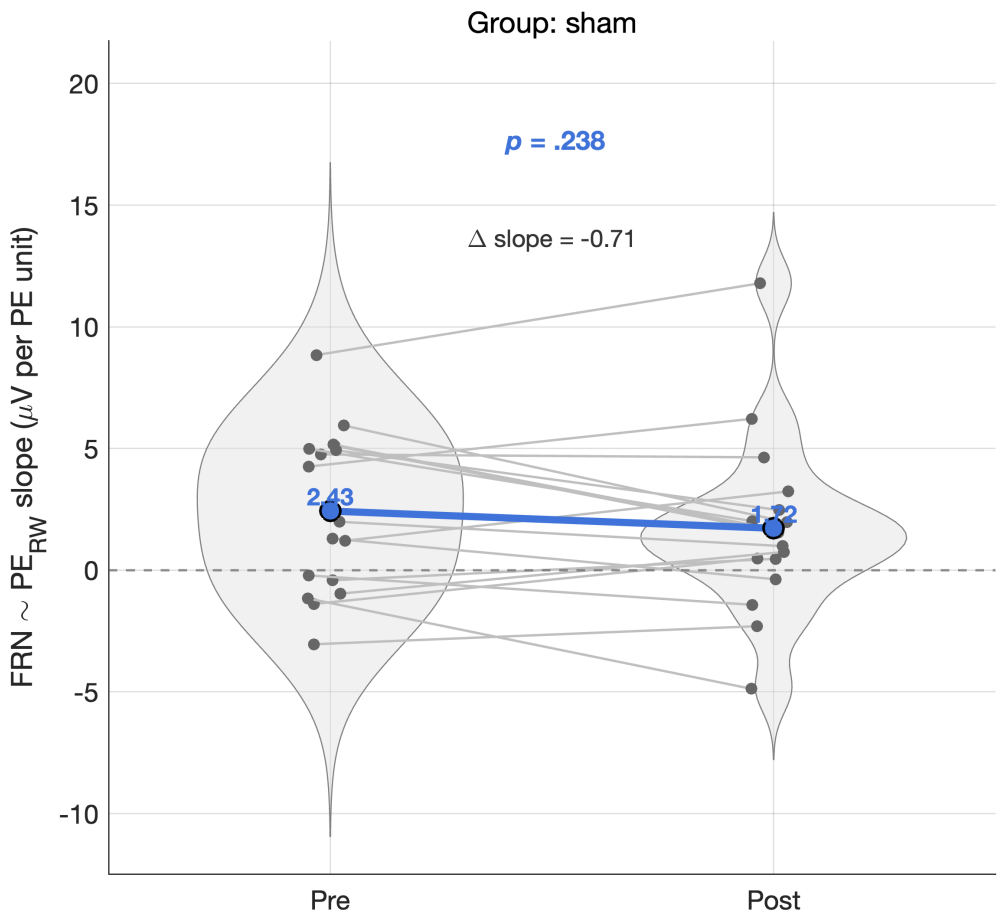
B: Food incentive delay task

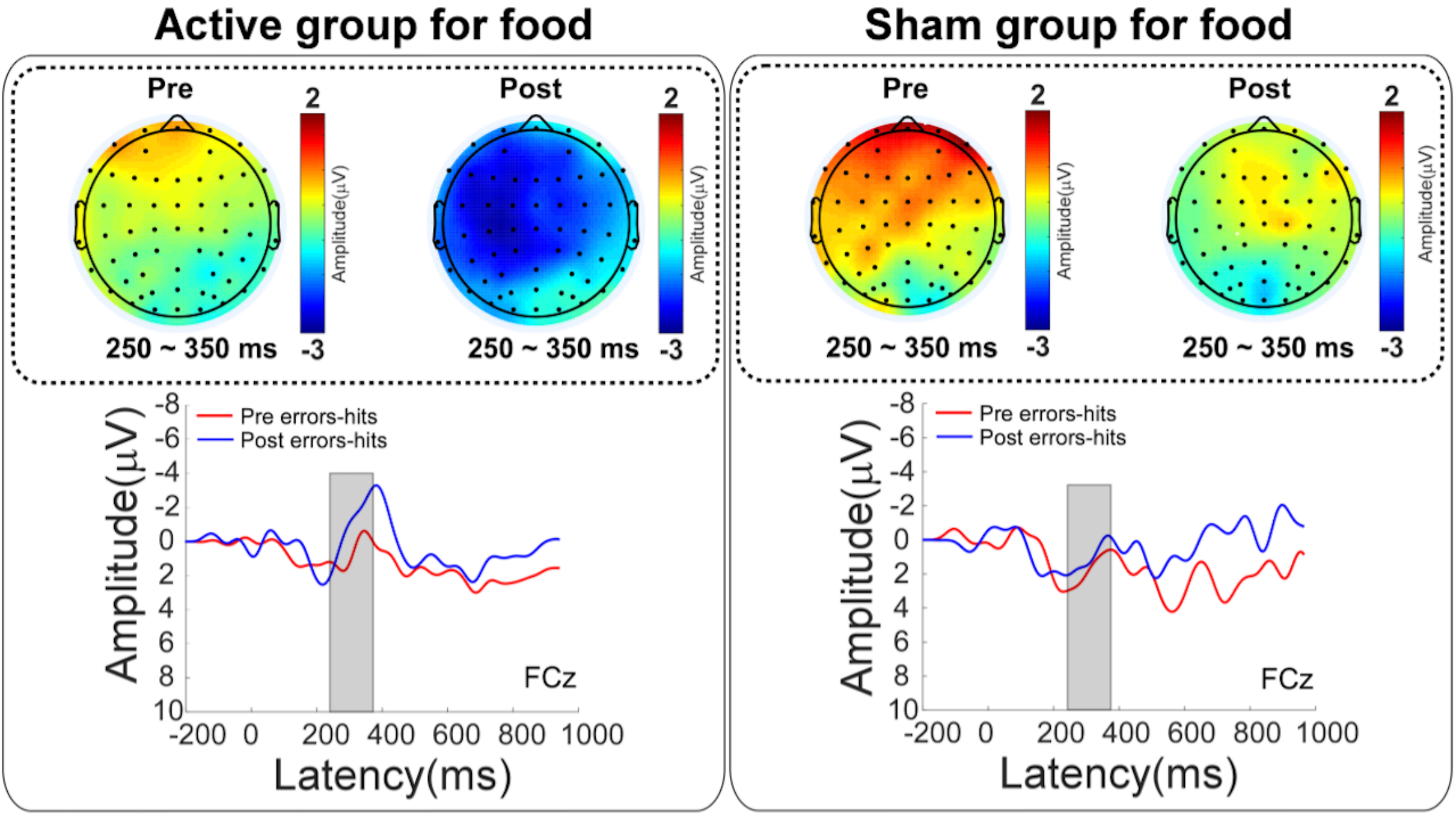


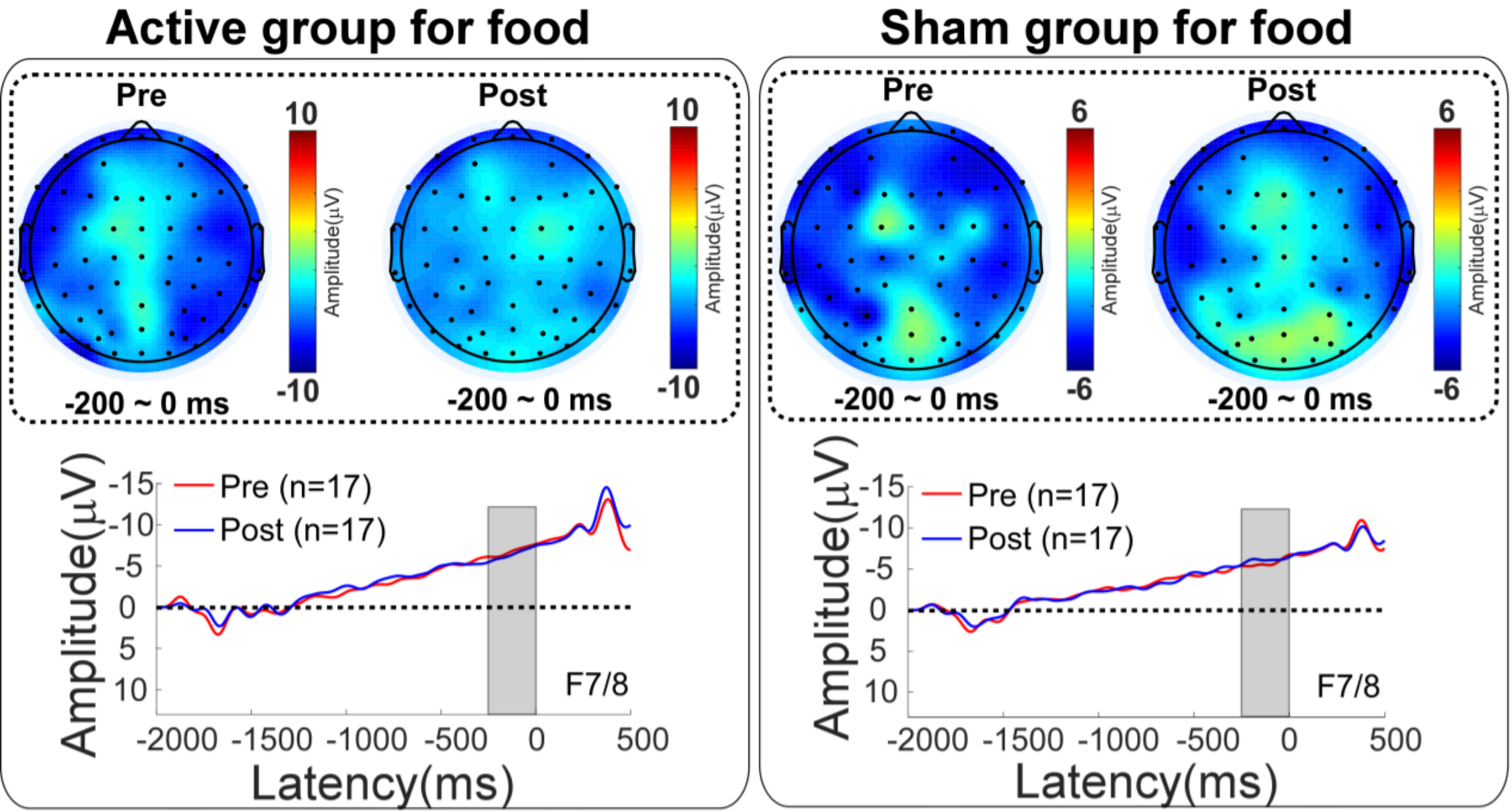

A. tTIS
CP3
Fz
AF3
F7
B. tDCS
0.56
Electrical field (V/m)
0
Z
X
Y

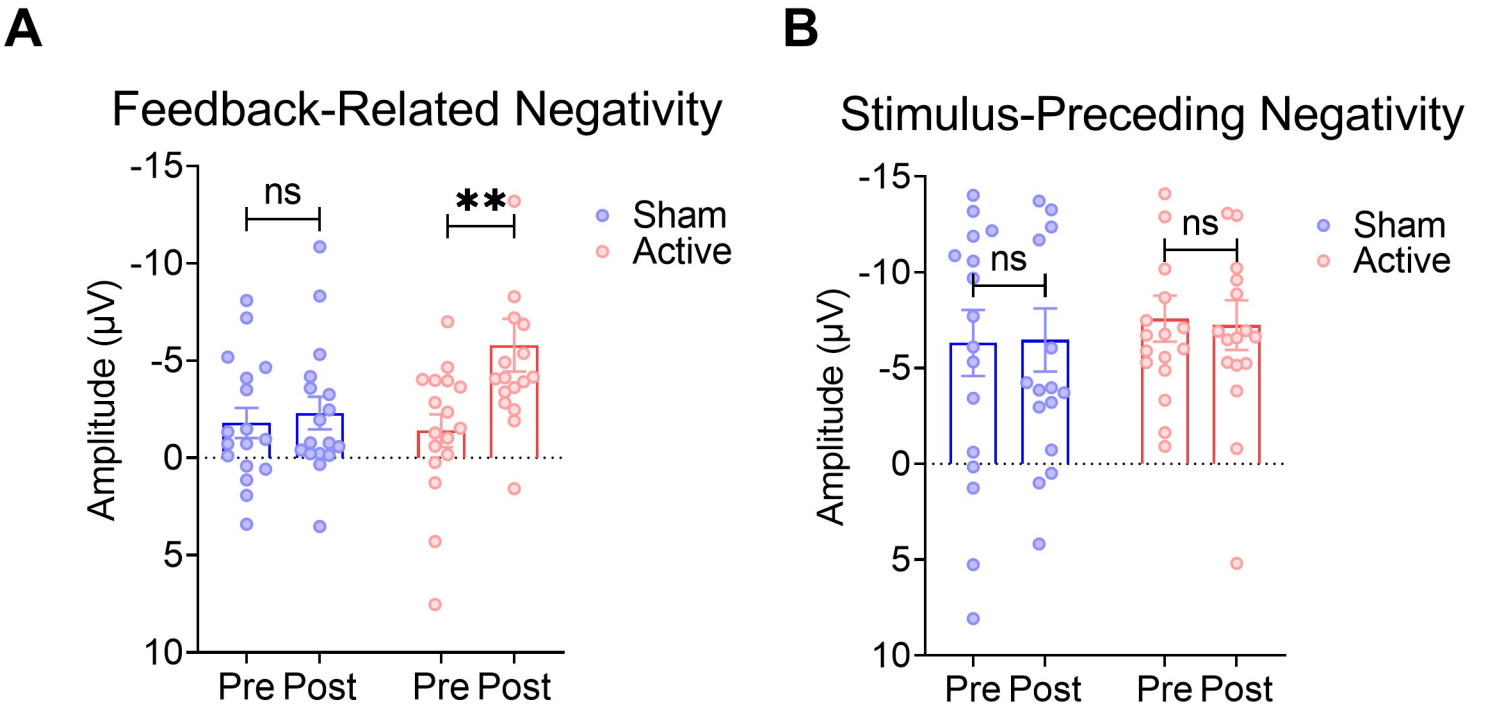

Group: active
p = .091
Δ slope = 1.72
0.96
2.70
Group: sham
p = .937
Δ slope = -0.04
1.25
1.21
FRN ~ PE_behavior slope (μV per PE unit)
12
10
8
6
4
2
0
-2
-4
-6
-8
Pre
Post

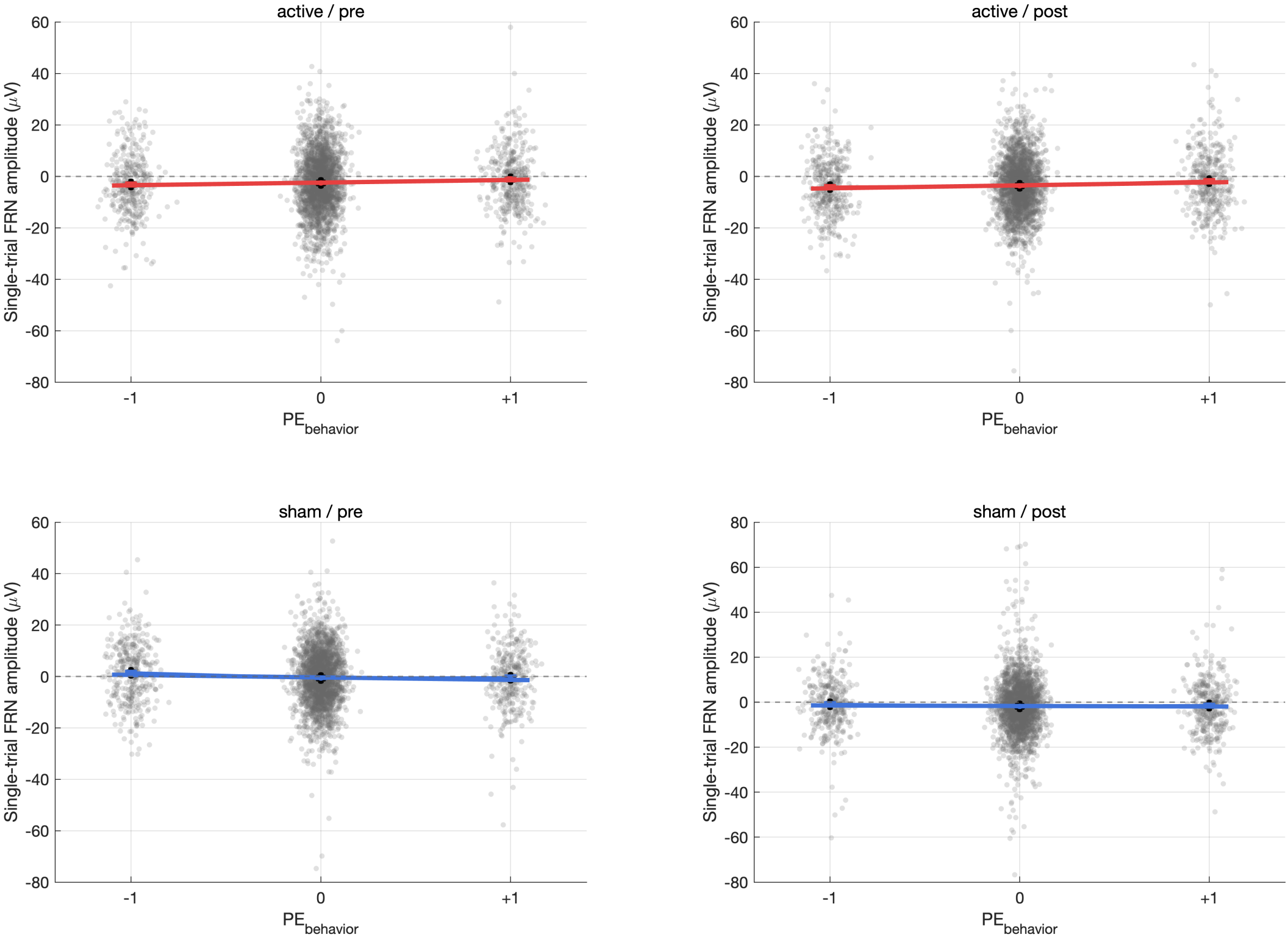

Group: active
p = .260
Δ slope = 1.05
1.95
3.00
FRN ~ PE_RW slope (μV per PE unit)
20
15
10
5
0
-5
-10
Pre
Post

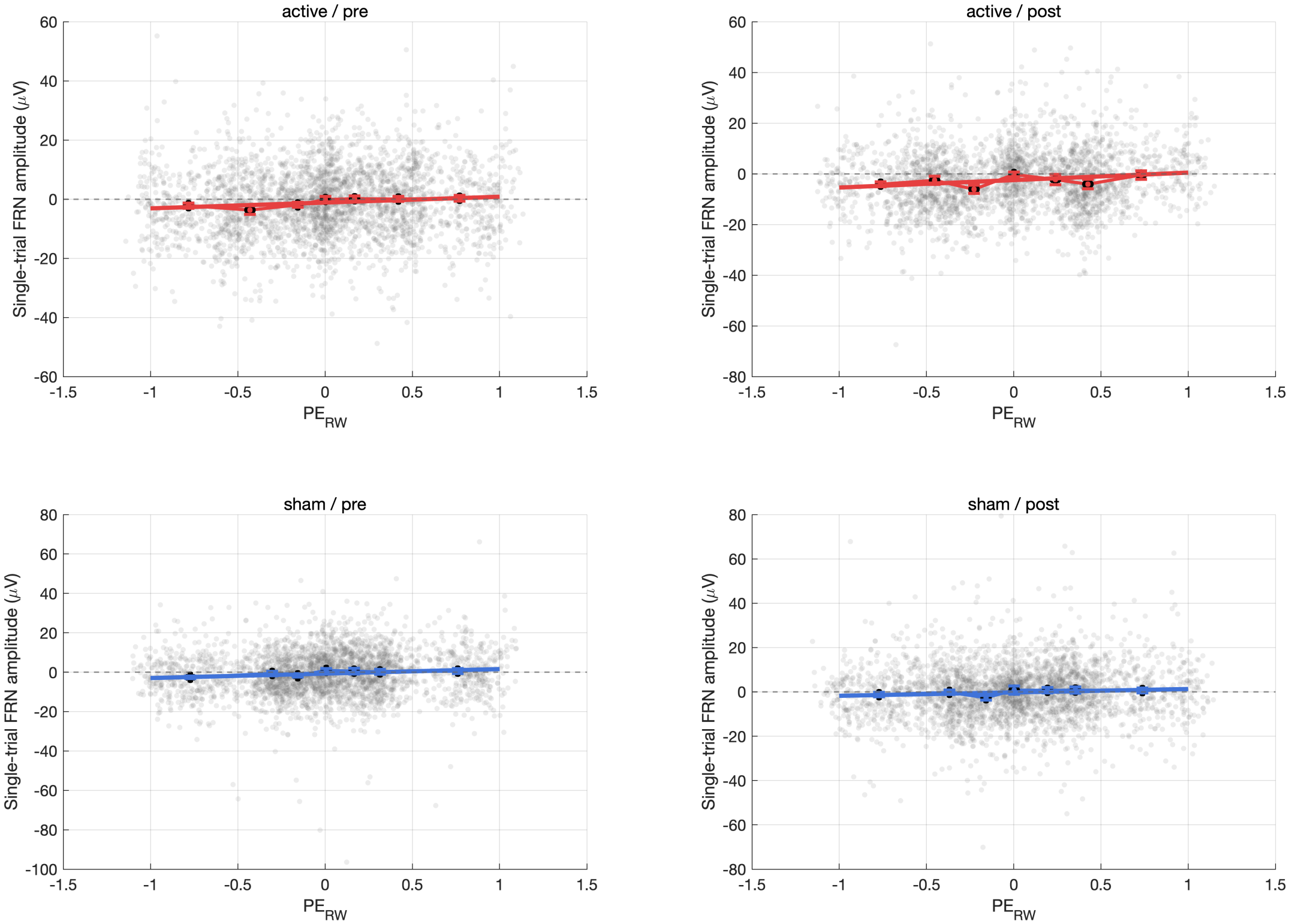

Group: sham
p = .238
Δ slope = -0.71
2.43
1.72
FRN ~ PE_RW slope (μV per PE unit)
20
15
10
5
0
-5
-10
Pre
Post

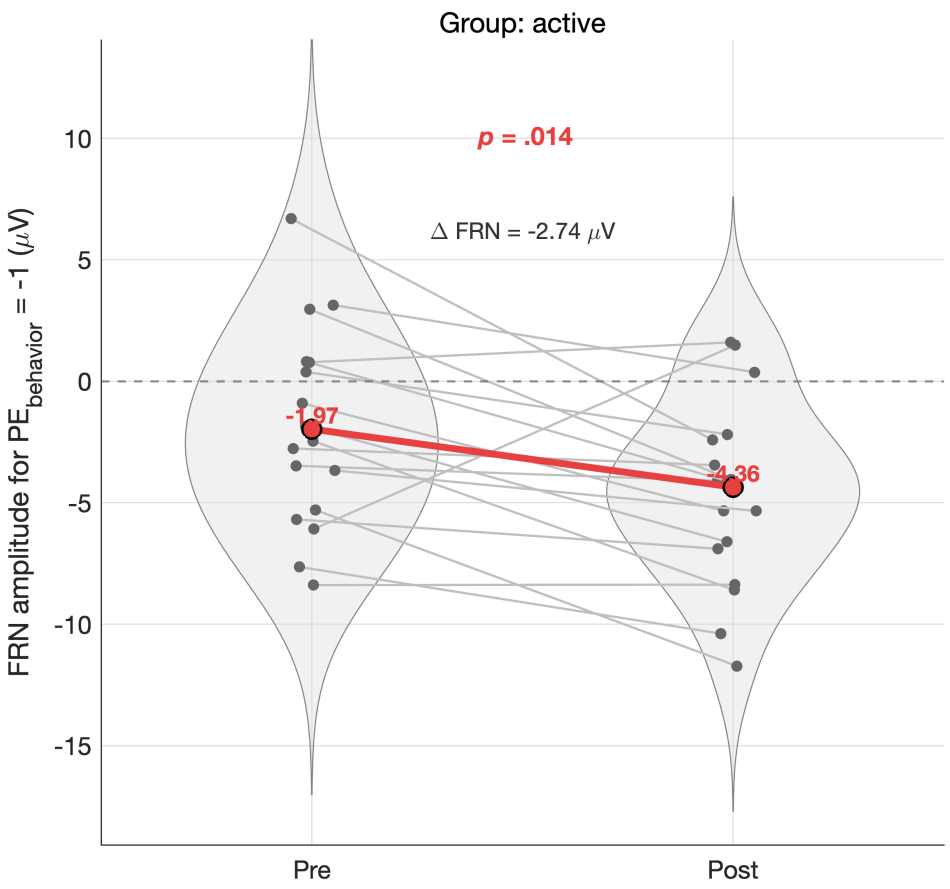

Active group for food
Sham group for food
Pre
Post
2
-3
Amplitude(μV)
250 ~ 350 ms
Pre errors-hits
Post errors-hits
FCz
Amplitude(μV)
Latency(ms)
-8
-6
-4
-2
0
2
4
6
8
10
-200
0
200
400
600
800
1000

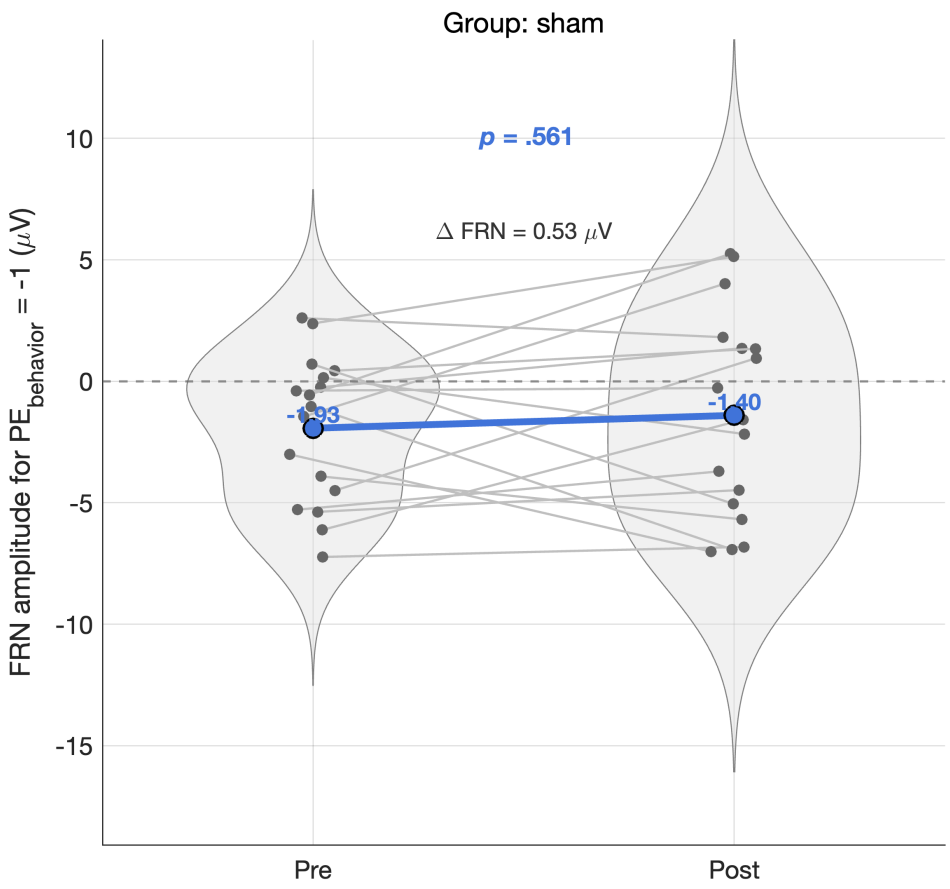

Active group for food
Pre
Post
10
-10
Amplitude(μV)
-200 ~ 0 ms
Pre (n=17)
Post (n=17)
F7/8
Amplitude(μV)
Latency(ms)
Sham group for food
6
-6

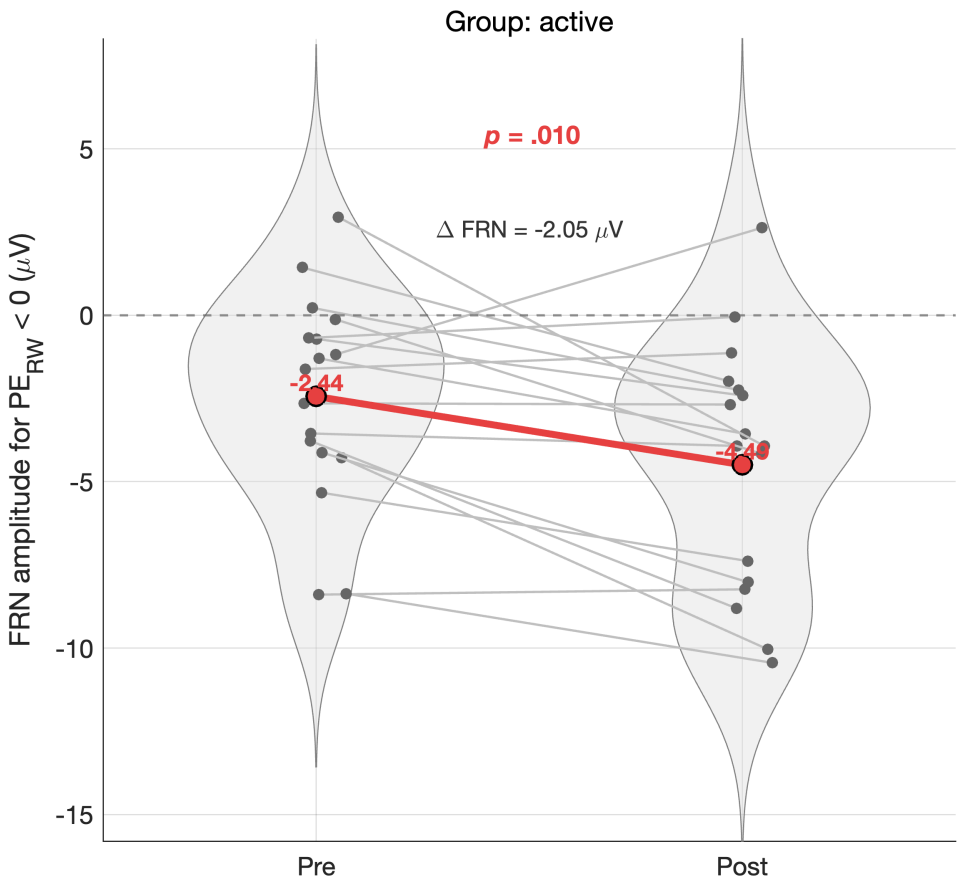

A
Feedback-Related Negativity
Amplitude (μV)
-15
-10
-5
0
5
10
ns
**
Sham
Active
Pre Post
Pre Post
B
Stimulus-Preceding Negativity
Amplitude (μV)
-15
-10
-5
0
5
10
ns
ns
Sham
Active
Pre Post
Pre Post

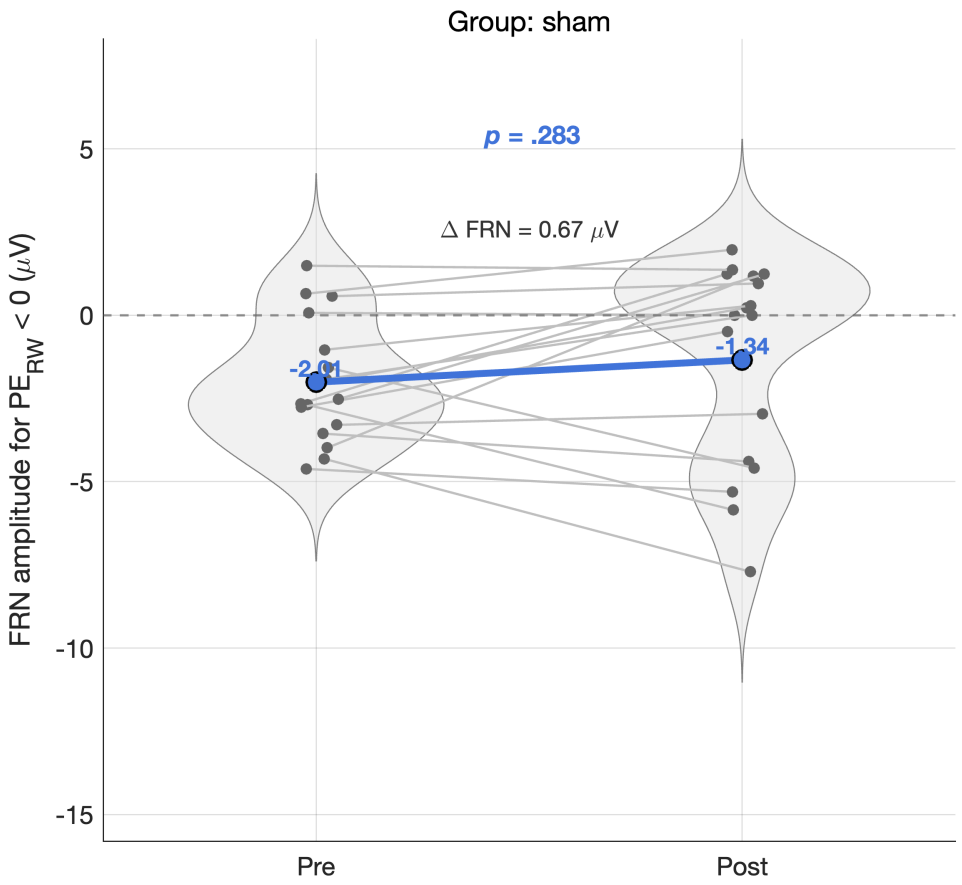

active / pre
active / post
sham / pre
sham / post
Single-trial FRN amplitude (μV)
PE$_{behavior}$
-1
0
+1
60
40
20
0
-20
-40
-60
-80
80


active / pre
active / post
sham / pre
sham / post
Single-trial FRN amplitude (μV)
PE_{RW}
60
40
20
0
-20
-40
-60
-80
-100
-1.5
-1
-0.5
0
0.5
1
1.5


Group: active
p = .014
Δ FRN = -2.74 μV
-1.97
-4.36
10
5
0
-5
-10
-15
FRN amplitude for PE behavior = -1 (μV)
Pre
Post


Group: sham
p = .561
Δ FRN = 0.53 μV
-1.93
-1.40
10
5
0
-5
-10
-15
FRN amplitude for PE behavior = -1 (μV)
Pre
Post


Group: active
p = .010
Δ FRN = -2.05 μV
-2.44
-4.49
FRN amplitude for PE_RW < 0 (μV)
5
0
-5
-10
-15
Pre
Post


Group: sham
p = .283
Δ FRN = 0.67 μV
-2.01
-1.34
FRN amplitude for PE_RW < 0 (μV)
5
0
-5
-10
-15
Pre
Post