Modulating hierarchical learning by high-definition transcranial alternating current stimulation at theta frequency

Meng Liu[1], Wenshan Dong[1], Yiling Wu[1], Pieter Verbeke[2], Tom Verguts[2], Qi Chen[1*]

[1] Key Laboratory of Brain, Cognition and Education Sciences, Ministry of Education, China; School of Psychology, Center for Studies of Psychological Application, and Guangdong Key Laboratory of Mental Health and Cognitive Science, South China Normal University, China

[2] Department of Experimental Psychology, Ghent University, Ghent, Belgium

[*]Address Correspondence to

Qi Chen

Address: *School of Psychology, South China Normal University, 510631 Guangzhou, China.*

Tel: +86-18617353673

E-mail: chen.qi@m.scnu.edu.cn

Abstract

Considerable evidence highlights the dorsolateral prefrontal cortex (DLPFC) as a key region for hierarchical learning. In a previous electroencephalography (EEG) study we found that the low-level precision-weighted prediction errors (pwPEs) were encoded by frontal theta oscillations (4–7Hz), centered on F4 in the 10-20 EEG system corresponding to right DLPFC (rDLPFC). However, the causal relationship between frontal theta oscillations and hierarchical learning remains poorly understood. To investigate this question, in the current study, participants received theta (6Hz) and sham high-definition transcranial alternating current stimulation (HD-tACS) over the rDLPFC, while performing the probabilistic reversal learning task. Behaviorally, theta tACS induced a significant reduction in accuracy for the stable environment, but not for the volatile environment, relative to the sham condition. Computationally, we implemented a combination of a hierarchical Bayesian learning and a decision model. Theta tACS induced a significant increase in low-level (i.e., probability) learning rate and uncertainty of low-level estimation relative to sham condition. Instead, the temperature parameter of the decision model was not significantly altered due to theta stimulation. These results indicate that theta frequency may modulate the (low-level) learning rate. Furthermore, environmental features (e.g., its stability) may determine whether learning is optimized as a result.

Key words: high-definition tACS; dorsolateral prefrontal cortex; hierarchical learning; computational modeling

Introduction

Hierarchy appears widely in the anatomy and cognitive processes of the human brain, as it provides several computational advantages for adapting to the environment. Understanding hierarchical learning at the computational and the neural level is crucial for understanding adaptive learning under uncertainty. Previous studies have demonstrated that humans engage in hierarchical learning (Botvinick 2008; Botvinick et al. 2009; Mathys 2011; Diuk et al. 2013; Iglesias et al. 2013, 2021; Mathys et al. 2014; Verbeke and Verguts 2019; Verbeke et al. 2021). In this perspective, hierarchically organized prediction errors (PEs) are crucial for predicting information inputs and inferring their underlying causes. Such a hierarchical mechanism was demonstrated by Iglesias et al. (2011): This fMRI work linked low-level PEs to frontal dopamine-receptive regions like dorsolateral prefrontal cortex (DLPFC) and anterior cingulate cortex (ACC); in contrast, high-level PEs activated the cholinergic basal forebrain. These findings suggested separable roles for dopaminergic and cholinergic modulation at different hierarchical levels of learning (Iglesias et al. 2013, 2021; Powers et al. 2017; Deserno et al. 2020; Henco et al. 2020; Sevgi et al. 2020). Given the role of the DLPFC in adaptive learning, flexible regulation of activity in the DLPFC might thus adjust hierarchical learning (Bishop, 2009; Fonteneau et al., 2018).

Additionally, a parallel line of work has demonstrated functionally relevant oscillatory activity in the same network during hierarchical learning. Frontal cortical activity in the theta band (4–7Hz) is crucial in the implementation of behavioral adaptation and cognitive control (Senoussi et al., 2022). It has been proposed that

theta oscillations are modulated by unsigned PE (Bayesian surprise) which was considered to correspond to low-level PE (Oliveira et al. 2007; Cavanagh et al. 2010). Particularly, theta oscillations are considered to be involved in long-range neural communication between cortical and subcortical areas (Womelsdorf et al. 2010; Voloh et al. 2015; Babapoor-Farrokhran et al. 2017), and it is also considered key to synchronize neuronal activity to implement hierarchical learning (Verguts 2017; Verbeke and Verguts 2021; Verbeke et al. 2021; Senoussi et al. 2022). Recently, in Liu et al. (2022), we combined a hierarchical learning model and EEG technology to reveal that hierarchical PEs at low-level were encoded in frontal theta oscillations, centered on F4 in the 10-20 EEG system corresponding to right DLPFC (rDLPFC). However, the causal relationship between frontal theta oscillations and hierarchical learning is incompletely understood. To address these issues, we used non-invasive, high-definition tACS (HD-tACS) to establish the functional contribution of theta frequency in hierarchical learning and examined its potential impact on computational mechanism.

The Hierarchical Gaussian Filter (HGF) is a validated hierarchical computational framework used to probe hierarchical learning (Mathys 2011; Mathys et al. 2014). In this setting, the brain is considered to compute two hierarchical PEs for updating hierarchically coupled beliefs: a low-level PE serves to update the estimate of the probability of the stimulus, and a high-level PE serves to update the estimate of the volatility of the probability. For each level, PEs are weighted by uncertainty to modulate the magnitude of belief updates (Mathys 2011; Mathys et al. 2014).

Generally, learning rate at each level is proportional to the uncertainty about the same level; specifically, more uncertain environments promote a greater integration of new information to better predict the future (Behrens et al. 2007, 2008). Overestimating or underestimating uncertainty would thus alter adaptive learning (i.e., the learning rate) (Feldman and Friston 2010; Mathys 2011; Hein et al. 2021). Accordingly, the current study aims to examine whether this hierarchical Bayesian framework would help explain the mechanism of frontal theta oscillations regulating task performance. Specifically, we investigate whether frontal theta stimulation regulates hierarchical learning primarily through changes in uncertainty estimates.

In summary, the current study aimed to test the causal relationship between frontal theta oscillations and hierarchical learning. Participants received theta (6Hz), and sham HD-tACS over rDLPFC (within-participants, once a week), while performing the probabilistic reversal learning task. In general, behavior depends on both learning and decision. Consequently, we fitted individual behavioral data using a combination of a learning and a decision model, to reveal the potential stimulation effects on learning and decision processes respectively. Here, the learning model (i.e., HGF model) allows for obtaining individual learning characteristics at hierarchical level; while the decision model contains an individual temperature parameter which represents the noise or the degree of exploration during decision. On the one hand, if the neuromodulation effect regulates the learning process, significant changes in learning parameters would be revealed after the theta tACS. On the other hand, if the neuromodulation effect regulates the decision process, a significant change in

temperature parameter would be found after the theta tACS. Overall, based on our previous finding that low-level pwPEs were encoded by theta oscillations (Liu et al. 2022), we hypothesized that theta neuromodulation would regulate hierarchical learning primarily through increasing low-level learning rate and uncertainty.

Materials and Methods

Participants

Forty-four healthy college students participated in this study for reimbursement. All of them were right-handed, with normal or corrected-to-normal vision. This experimental procedures were approved by the Ethics Committee of the Institute of Psychology, South China Normal University (No. SCNU-PSY-2021-042). Each participant provided written consent before the experiment. Four measured participants were not included in the analysis: three participants did not perform the task according to the instruction, and one was excluded due to excessive reaction time. Accordingly, the final sample included forty effective subjects (24 females), age from 18 to 24 years old (M = 20.43 ± 1.6 years).

Experimental Procedure

We used a within-participants design, and each participant was recruited to complete the probabilistic reversal learning task for two sessions (once a week): theta (6 Hz) and sham stimulation. The order of the HD-tACS protocol was counterbalanced across participants. The study design was single blinded. As such, participants were

kept naïve about the session-specific stimulation conditions.

In each session, participants were asked to complete a probabilistic reversal learning task including 120 trials, where they need to learn the predictive strength of cues (two circular patterns) and predict a subsequent visual stimulus (horizontal or vertical grating); see Figure 1A and 1B for illustration. All procedures for each session were identical, with the exception of the cues. These cues were selected from Kool and McGuire's photo gallery (Kool et al. 2010; McGuire and Botvinick 2010). The stimuli (horizontal or vertical grating) were also used in several earlier studies (Verbeke et al. 2021; Liu et al. 2022). Critically, the cue-stimulus association strength changed over time in the current task. This learning environments contained a stable environment, the probability of horizontal given a cue A remained at 0.75 (as did the probability of vertical given a cue B), and a volatile environment, the probability of horizontal given a cue A switched between 0.2 and 0.8 (see Figure 1B). Participants were explicitly instructed that the conditional probabilities were perfectly coupled in the sense that $p(horizontal \mid cueA) = p(vertical \mid cueB)$.

At the beginning of each trial, one of two possible cues was presented in the center of the screen for 800 ms, which was followed by a 400ms fixation cross. Subsequently, participants were informed to determine whether a horizontal or vertical grating would follow, and to make their responses as accurately and as quickly as possible by pressing either "F" or "J" buttons. Horizontal and vertical gratings were randomly displayed on the left and right of the screen in order to control for handedness, where the grating on the left corresponded to "F" button and the

grating on the right corresponded to "J" button. The presentation of selection ended immediately after a response was made. Then a fixation cross lasted for 1000-1500ms followed by a 500ms feedback (i.e., correctly matched grating, horizontal or vertical grating). Each session began with a practice with 20 trials, ensuring that participants understood the experimental requirements (see Figure 1A).

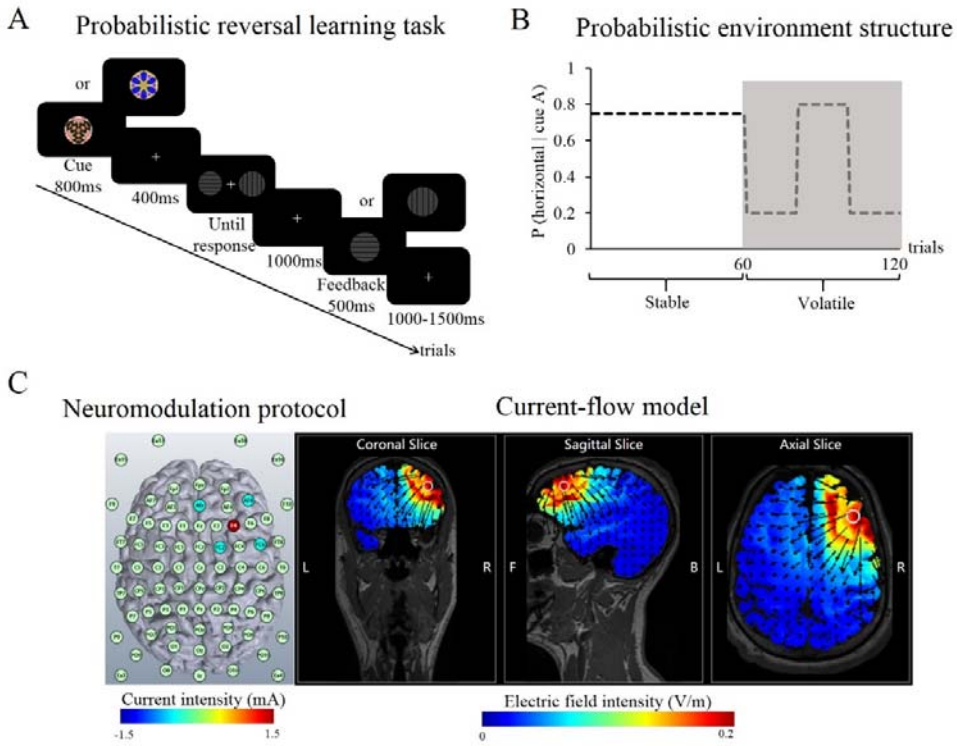

Figure 1. Task design. (A) Probabilistic reversal learning task. Each trial started with one of two cues centrally presented for 800 ms. Participants were instructed to report either horizontal or vertical grating would follow. After response, a 500ms feedback was presented in the center of the screen. (B) Time-varying cue- stimulus contingency across trials. Participants were asked to complete the session without breaks, including 120 trials. The first half, a stable environment, the probability of horizontal

given a cue A remained at 0.75; the second half, a volatile environment, the probability of horizontal given a cue A switched between 0.2 and 0.8 three times. For half of the subjects, this order was reversed. (C) RDLPFC neuromodulation protocol and current-flow models on three-dimensional reconstructions of the cortical surface.

High-definition Transcranial Alternating Current Stimulation

We applied HD-tACS using a five-channel high-definition transcranial electrical-current stimulator (Soterix Medical). In line with our hypothesis, we planned to deliver focalized current to the rDLPFC. HD-Explore and HD-Targets was applied to model electrical field to decide how and where to place electrodes. Figure 1C shows the modulation parameters. The total current strength was 1.5 mA, with AFz (-0.375 mA), AF8 (-0.375 mA), F4 (1.5 mA), FC2 (-0.375 mA), and FC6 (-0.375 mA) separately (see Figure 1C). In the theta (6 Hz) stimulation, modulation was turned on 1 minute before the start of the probabilistic reversal learning task and lasted until the end of the session. In the sham stimulation, modulation was ramped up and then ramped down within 30 seconds before the start of the task, and no real stimulation was applied in the whole session. All participants reported that the stimulation process was acceptable without intense skin pain or phosphene.

Computational Modelling

To investigate the mechanisms of modulation effects on hierarchical learning, we modelled individually learning and decision processes with a combination of

perceptual learning and decision model (Mathys 2011; Mathys et al. 2014), as in our previous study (Liu et al. 2022). Two types of models were paired in order to infer the latent states of an observer during the task.

Learning model

The perceptual learning model was a validated Hierarchical Gaussian Filter (HGF). The implementation of the HGF analyses in the current study was available (TAPAS, http://www.translationalneuromodeling.org/tapas). Regarding to our probabilistic reversal learning task, the three levels of the HGF model correspond to the following: the first level represents the occurrence of the cue and grating stimuli (x_1), the second level represents the conditional probability of the grating given the cue (x_2), and the third level represents the change in this conditional probability (log-volatility, x_3). Each of these hidden states (x_2 and x_3) is assumed to evolve as a Gaussian random walk, such that its mean is centered around its previous value at trial k-1, and its variance depends on the state at the level above. As in previous work (De Berker et al. 2016; Hein et al. 2021; Liu et al. 2022), ω_2 and ω_3 were estimated in each participant, and κ was fixed to 1. Here, ω_2 represents a constant component of the step size at the second level, and ω_3 represents metavolatility parameter. In addition, κ represents the strength of the estimated environmental volatility affected the probability learning rate. Details on the exact update equations are provided in Mathys et al. (2011, 2014).

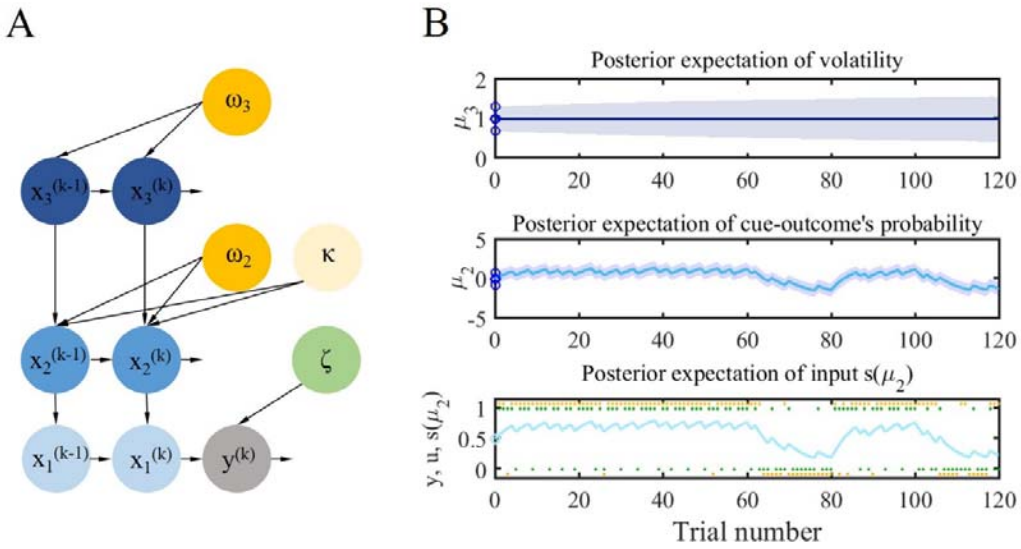

Figure 2. Computational modeling of behavior. (A) Model architecture. A combination of a three-level HGF and a decision model describe individual learning and decision processes. The blue and gray circles represent potential states changed trial-by-trial, and the yellow and green circles denote free parameters estimated from the HGF or decision models. (B) Binary outcomes and model-based hierarchical belief trajectories from a representative participant. At the first level, green dots correspond to the actual associative stimulus of each trial, and yellow dots corresponds to the participant's selections (1 = cue A-horizontal stimulus or cue B-vertical stimulus, 0 = cue A-vertical stimulus or cue B-horizontal stimulus).

For each trial, the update formula of HGF for both low-level (i = 2) and high-level (i = 3) is expressed in the following form:

$$\Delta \mu_i^{(k)} = \mu_i^{(k+1)} - \mu_i^{(k)} \propto \psi_i^{(k)} \delta_{i-1}^{(k)} = \varepsilon_i^{(k)} \quad (1)$$

This expression illustrates that the updating of the expectation of the posterior mean $\Delta \mu_i^{(k)}$ is proportional to the weighted PE $\varepsilon_i^{(k)}$. Here, the precision weighting of the

PE is a ratio of precisions $\psi_i^{(k)}$, which can be considered as a dynamic learning rate (Preuschoff and Bossaerts 2010; Mathys et al. 2014; Liu et al. 2022). This precision weighting $\psi_i^{(k)}$ equals

$$\psi_i^{(k)} = \frac{\hat{\pi}_{i-1}^{(k)}}{\pi_i^{(k)}} \quad (2)$$

where $\hat{\pi}_{i-1}^{(k)}$ is the precision of the prediction of the level below and $\pi_i^{(k)}$ is the precision of the posterior expectation on the current level. Precision (π) is defined as the inverse variance (or uncertainty, σ_i) of the posterior expectation:

$$\pi_i^{(k)} = \frac{1}{\sigma_i^{(k)}} \quad (3)$$

The intuition behind this is that the precision of its data modulate the magnitude of updating of an individual's belief, with a higher precision (or lower uncertainty) promoting a greater integration of PEs.

Decision model

The decision model is a unit-square sigmoid observation model, which links the posterior estimates of beliefs to expressed decisions $y^{(k)}$ (Mathys 2011; Mathys et al. 2014). Here, the predicted probability $m^{(k)}$ that a grating stimulus (e.g., horizontal grating) given the cue A on trial k is linked to trial-wise predictions of grating stimulus category by using the softmax (logistic sigmoid) function:

$$p\ (y|m,\zeta)\ =\ \left(\frac{(m^{\zeta})}{m^{\zeta}+(1-m)^{\zeta}}\right)^{y}\ .\ \left(\frac{(1-m)^{\zeta}}{m^{\zeta}+(1-m)^{\zeta}}\right)^{1-y} \quad (4)$$

where ζ is a subject-specific temperature parameter and can be interpreted as inverse decision noise (see Figure 2A). This temperature parameter ζ is the only free parameter in the decision model. It captures how deterministically y is associated

with m. Higher ζ values represent that participants are more likely to choose the option that is more congruent with their current belief.

Model Space

Using random effects Bayesian model selection (BMS; Stephan et al. 2009; code from the freely available MACS toolbox; Soch and Allefeld 2018), following our previous study (Liu et al. 2022), we compared three different learning models. The decision model mentioned above was combined with three different learning models. First was the three-level hierarchical learning model (with volatility on the third level: HGF3). The second was the Rescorla Wagner (RW) model, which is probably the simplest and most widely used learning model. It adopts the idea of prediction errors driving belief updating, but the fitted learning rate is not adaptive (Rescorla and Wagner 1972). Third, the Sutton K1 model (SK1) model allows adaptive learning rate to adapt with recent prediction errors, but without a hierarchy (Sutton and Barto 1998). Also, as in previous work (Iglesias et al. 2013, 2021; Hein et al. 2021; Liu et al. 2022), there was one single value tracked in all three models. Participants were explicitly informed that the probability of one stimulus (i.e., horizontal grating) given a particular cue (i.e., cue A), was the same as the probability of another stimulus (i.e., vertical grating) given another cue (i.e., cue B).

Data Analysis

Firstly, paired *t*-tests were carried out to test the neuromodulation effects on

behavioral task performance (i.e., accuracy) during the probabilistic reversal learning task. Accuracy was defined as making the optimal choice (i.e., the choice with the currently highest probability of success) on a given trial. Furthermore, paired *t*-tests were also performed for stable and volatile environments separately, in order to analyze the neuromodulation effects on different learning environments.

Secondly, to investigate the neuromodulatory effects on hierarchical learning specifically, we focused on individual computational parameters from the learning and decision models separately. On the one hand, we carried out paired *t*-tests to test the neuromodulatory effects on hierarchically learning model parameters: learning rate (ψ_2, ψ_3), estimation of uncertainty (σ_2, σ_3), and precision-weighted PEs ($|\varepsilon_2|$, ε_3) on both low-level (here, level 2) and high-level (here, level 3). Consistent with previous studies (Iglesias et al. 2013; Hein et al. 2021; Liu et al. 2022), the absolute value of ε_2 was chosen. Indeed, it does not matter if we code the probability of horizontal grating given the cue A or instead the probability of vertical grating given the cue A. On the other hand, a paired *t*-test was used to test the neuromodulation effects on temperature parameter (ζ) from decision model. Furthermore, as for the analysis of accuracy, the same paired *t*-tests were carried out for stable and volatile environments separately.

Finally, to investigate the relationship between task performance and internal computational mechanisms, we tested whether the behavioral change induced by HD-tACS stimulation was related to the computational learning parameter. Here, we calculated the (Pearson) correlations between the change in accuracy (Δaccuracy, the

difference between theta and sham stimulation regarding accuracy) and the change in low-level probability learning rate ($\Delta\psi_2$, expected computational mechanism regulated by neuromodulation effect) for stable and volatile environments respectively.

Results

Stimulation Effects on Accuracy

Figure 3 illustrates the average accuracy for both theta and sham tACS, respectively. As a whole, the theta HD-tACS condition (M = 0.809 ± 0.010) showed a significant reduction in accuracy (averaged across all 120 trials) relative to the sham condition (M = 0.832 ± 0.008), $p < 0.01$, d = -0.474. Furthermore, the theta HD-tACS condition showed a significant reduction in accuracy relative to the sham condition only for the stable environment (M = 0.872 ± 0.014 vs. M = 0.920 ± 0.012), $p < 0.001$, d = -0.778, but not for the volatile environment (M = 0.746 ± 0.012 vs. M = 0.745 ± 0.010) (p = 0.946).

For exploratory purposes, we investigated whether and found that the theta HD-tACS condition showed a significant increase in accuracy relative to the sham condition only for the post-reversal-1 (the five trials right after the first reversal were labelled post-reversal-1) (M = 0.255 ± 0.030 vs. M = 0.170 ± 0.020), $p < 0.05$, d = 0.376, but not for the post-reversal-2 (M = 0.505 ± 0.023 vs. M = 0.455 ± 0.023), p = 0.086, and also not for the post-reversal-3 (M = 0.380 ± 0.045 vs. M = 0.310 ± 0.040), p = 0.217 (see Figure 3). These data suggest that the characteristics of the environment (e.g., its stability) may determine whether learning is optimized.

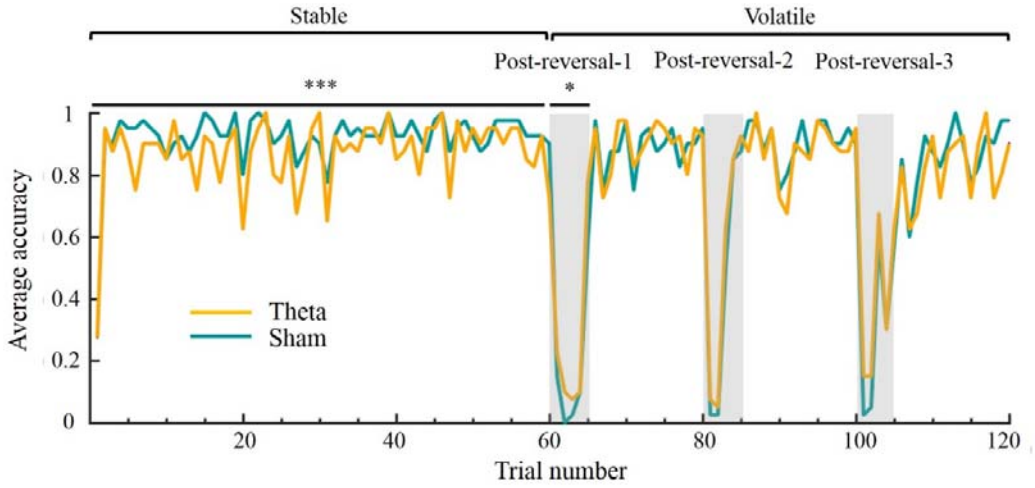

Figure 3. Neuromodulation effects on average accuracy. Average accuracy per trial of both theta and sham HD-tACS conditions. The theta HD-tACS condition showed a significant reduction in accuracy relative to the sham condition for the stable environment. However, the theta HD-tACS condition showed a significant increase in accuracy relative to the sham condition for post-reversal-1.

Model Comparison and Selection

BMS were applied to quantify the protected exceedance probability (PXP) of three competing models. The model with the highest PXP was selected as the optimal model. Comparisons of the three models are summarized in Table 1. Here, the HGF3 model was obviously superior to the other two alternative models. These suggest that participants were indeed likely to implement hierarchical learning. Having identified the optimal model (HGF3), we proceeded to testing whether there were significant differences in parameter estimates between theta and sham tACS stimulation.

Table 1. Model comparison results.

	HGF3	RW	SK1
EXP	0.9608	0.0233	0.0159
XP	1.000	0.000	0.000
PXP	1.000	0.000	0.000

Note: HGF3, hierarchical Gaussian filter with 3 levels; RW, Rescorla Wagner model; SK1; Sutton K1 model. EXPs, expected posterior probabilities; XPs, exceedance probabilities; PXPs, protected exceedance probabilities.

Stimulation Effects on Low-level Learning about Cue-stimulus Probabilities

The probability learning rate (ψ_2) in theta HD-tACS condition (M = 1.017 ± 0.048) was significantly higher than in the sham condition (M = 0.788 ± 0.038), p < 0.001, d = 0.889 (see Figure 4A). Similarity, the probability learning rate in theta HD-tACS condition was significantly higher than sham condition for both stable (M = 1.001 ± 0.045 vs. 0.789 ± 0.036), p < 0.001, d = 0.882, and volatile environments (M = 1.032 ± 0.052 vs. 0.787 ± 0.041), p < 0.001, d = 0.894.

Although learning rate, uncertainty, and prediction error are statistically related, we report results on all three variables for completeness. The uncertainty of probability estimation (σ_2) in the theta HD-tACS condition (M = 1.301 ± 0.077) was significantly higher than in the sham condition (M = 0.954 ± 0.055), p < 0.001, d = 0.861 (see Figure 4B). Similarly, the uncertainty of probability estimation in the theta HD-tACS condition was significantly higher than in the sham condition for both

stable (M = 1.280 ± 0.073 vs. 0.954 ± 0.052), p < 0.001, d = 0.856, and volatile environments (M = 1.322 ± 0.082 vs. 0.955 ± 0.058), p < 0.001, d = 0.865.

The absolute probability pwPE ($|\varepsilon_2|$) in the theta HD-tACS condition (M = 0.451 ± 0.022) was significantly higher than in the sham condition (M = 0.348 ± 0.017), p < 0.001, d = 0.881 (see Figure 4C). Furthermore, the distinction between stable and volatile environments revealed similar results. That is, the absolute probability pwPE in the theta HD-tACS condition was significantly higher than in the sham condition for both stable (M = 0.443 ± 0.022 vs. 0.340 ± 0.017), p < 0.001, d = 0.874, and volatile environment (M = 0.460 ± 0.022 vs. 0.356 ± 0.017), p < 0.001, d = 0.887.

In sum, we found a significant increase in probability learning rate (ψ_2), uncertainty of probability estimation (σ_2), and prediction error (epsilon) regulated by theta stimulation, in both stable and volatile learning environments.

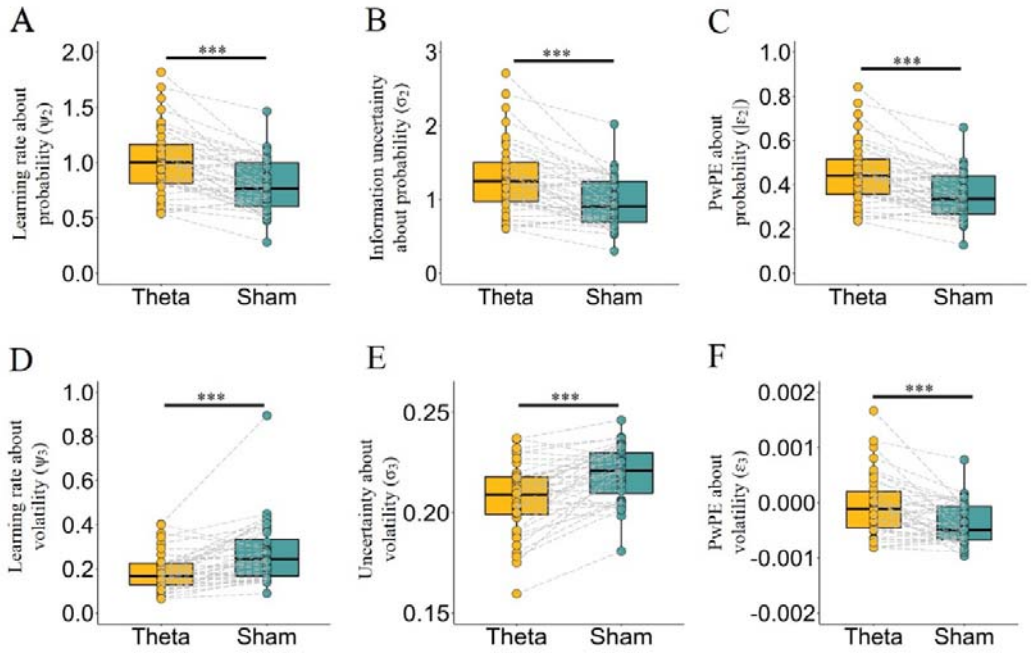

Figure 4. Neuromodulation effects on hierarchical computational learning parameters. (A-C) Stimulation effects on learning about low-level (probability). (A) Probability learning rate (ψ_2) in theta HD-tACS condition was significantly higher than in the sham condition. (B) Uncertainty of probability estimation (σ_2) in theta HD-tACS condition was significantly higher than in the sham condition. (C) Absolute probability pwPE ($|\varepsilon_2|$) in theta HD-tACS condition was significantly higher than sham condition. (D-F) Stimulation effects on learning about high-level (volatility). (D) Volatility learning rate (ψ_3) in the theta HD-tACS condition was significantly lower than in the sham condition. (E) Uncertainty of volatility estimation (σ_3) in the theta HD-tACS condition was significantly lower than in the sham condition. (F) Theta HD-tACS condition showed a significant increase in volatility pwPE (ε_3) relative to the sham condition.

Stimulation Effects on High-level Learning about Volatility

The volatility learning rate (ψ_3) in theta HD-tACS condition (M = 0.187 ± 0.014) was significantly lower than in the sham condition (M = 0.272 ± 0.022), $p < 0.001$, d = -0.747 (see Figure 4D). Similarly, the volatility learning rate in the theta HD-tACS condition was significantly lower than in the sham condition for both stable (M = 0.143 ± 0.009 vs. 0.199 ± 0.014), $p < 0.001$, d = -0.772, and volatile environments (M = 0.230 ± 0.018 vs. 0.346 ± 0.031), $p < 0.001$, d = -0.734.

The uncertainty of volatility estimation (σ_3) in the theta HD-tACS condition (M = 0.207 ± 0.003) was significantly lower than in the sham condition (M = 0.220 ±

0.002), $p < 0.001$, $d = -0.890$ (see Figure 4E). Similarly, the uncertainty of volatility in the theta HD-tACS condition was significantly lower than in the sham condition for both stable ($M = 0.161 \pm 0.001$ vs. 0.166 ± 0.001), $p < 0.001$, $d = -0.850$, and volatile environments ($M = 0.254 \pm 0.004$ vs. 0.274 ± 0.003), $p < 0.001$, $d = -0.892$.

The theta HD-tACS condition showed a significant increase in volatility pwPE (ε_3) relative to the sham condition ($M = 0.0000 \pm 0.00001$ vs. $M = -0.0004 \pm 0.0001$), $p < 0.001$, $d = 0.693$ (see Figure 4F). Similarly, the theta HD-tACS condition showed a significant increase in volatility pwPE relative to the sham condition for both stable ($M = -0.0005 \pm 0.00001$ vs. $M = -0.0008 \pm 0.0001$), $p < 0.001$, $d = 0.796$, and volatile environments ($M = 0.0004 \pm 0.00001$ vs. $M = 0.0000 \pm 0.0001$), $p < 0.01$, $d = 0.529$.

In sum, we found a significant decrease in volatility learning rate (ψ_3) and uncertainty of volatility estimation (σ_3) regulated by the theta stimulation, in both stable and volatile learning environments.

Stimulation Effects on Temperature Parameter of Decision Model

We tested whether the neuromodulation also affected the decision process. The temperature parameter is the only free parameter in the decision model. No significant difference was found between the theta condition ($M = 4.350 \pm 1.001$) and sham condition ($M = 4.493 \pm 0.969$) ($p = 0.670$) on the temperature parameter.

Correlation between the Changes of Accuracy and Low-level Learning Rates

In line with our hypothesis, theta stimulation modulated the low-level learning rate.

Consequently, we next correlated task performance (i.e., accuracy) and the low-level learning rate across subjects. Considering that the effect of neuromodulation on accuracy varies across environments, the correlations were calculated for stable and volatile environment separately. As can be seen in Figure 5A, the change in accuracy induced by HD-tACS stimulation (Δaccuracy, the difference between theta and sham condition) was significantly negatively correlated with the change in probability learning rate ($\Delta\psi_2$, the difference between theta and sham condition) in the stable environment, r = -0.355, p = 0.024 (see Figure 5A). However, no significant correlation was reported for the volatile environment, r = 0.186, p = 0.249 (see Figure 5B).

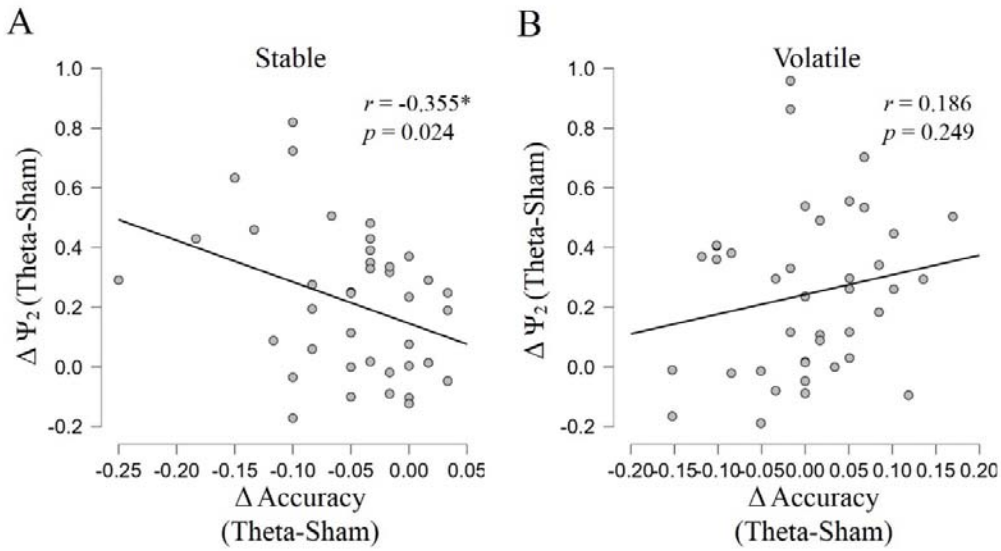

Figure 5. Correlation between behavioral changes and computational learning parameter. The change on accuracy (between theta and sham condition) and the change on probability learning rate (A) in the stable environment and (B) in the volatile environment.

Discussion

The current study combined HD-tACS and the HGF model to assess the modulatory effects of frontal theta frequency band on hierarchical learning. Behaviorally, theta tACS modified participants' adaptive behavior such that theta stimulation induced a significantly decreased accuracy in an environment suitable for stable learning, but a significantly increased accuracy in an environment suitable for fast learning (relative to the sham condition). Computationally, the HGF model revealed that the low-level learning rates were significantly increased relative to sham condition, which may explain the change in accuracy. Moreover, the HGF model also revealed that the theta condition increased the estimation of low-level uncertainty compared to the sham condition. This suggests that theta stimulation may bias the estimation of uncertainty and then adjust hierarchical learning. Furthermore, no stimulation effects on the temperature of the decision model were observed, suggesting that the effect of theta stimulation is specific to the learning process but not to the decision process. Finally, those subjects with worse performance (in stimulation relative to sham) also had a relatively more increased learning rate (in stimulation relative to sham).

Adaptive learning requires a dynamic balance between stability and flexibility (Behrens et al. 2007; Nassar et al. 2013; Massi et al. 2018). For example, in the current probabilistic reversal learning task, an ideal learner needs to adjust learning strategies to adapt to different learning environments: more stable in a stable environment and more flexible in a volatile environment. However, the current results

revealed that accuracy was significantly decreased (presumably due to increased low-level learning rate) after theta stimulation in the stable environment, but not significantly altered in the volatile environment. This suggests that theta stimulation does not promote adaptive learning, but makes learners more inclined to faster learning. This matches previous findings about frontal neuromodulation in probabilistic reversal learning (Wischnewski et al. 2016; Borwick et al. 2020; Panitz et al. 2022), which demonstrated that neuromodulation promotes faster learning to volatile environment. Our study supplements these earlier results by demonstrating that faster learning reduces task performance in the stable environment. Furthermore, our exploratory analysis found that theta tACS induced a significantly increased accuracy in post-reversal-1 relative to sham condition, which may also support accelerating learning in theta stimulation. Collectively, theta tACS seemed to speed up (lower-level) learning, improving accuracy in the environment suitable for faster learning, but impairing accuracy in the environment suitable for more stable learning. These findings suggest that environmental features (e.g., its stability) may determine whether learning is optimized.

The use of (hierarchical) computational modeling in the current study allowed for a specific focus on (low-level) learning rate, to understand the underlying mechanisms of the changes of accuracy. Consistent with previous studies, the empirical advantage of the HGF framework suggested that a hierarchical framework is the key to learning (Mathys 2011; Iglesias et al. 2013, 2021; Mathys et al. 2014). The current results, combined with those of previous theta oscillations studies (Cohen

et al. 2007; Cavanagh et al. 2010; Cavanagh and Frank 2014; Soutschek et al. 2021), suggested that frontal theta oscillations increase an individual's low-level (probability) learning rate from feedback information. Previous EEG studies suggested that frontal theta oscillations underlie neural processes associated with reinforcement learning and decision making. For example, frontal theta oscillations have been linked to response to an incorrect trial, and then predict subsequent adaptive behavior (Cohen et al. 2007; Cohen 2011). In further support, theta oscillations in a probabilistic reversal learning task can improve reinforcement learning and speed up rule learning after reversal (Wischnewski et al. 2016). In the current study, theta tACS significantly increased low-level learning rate, which could explain the change of behavioral task performance. These results suggest that frontal theta increases learning from feedback, and makes humans over-react to environmental change (i.e., they “chase the noise”).

Previous studies established links between theta frequency and uncertainty (Cavanagh et al. 2012; Wischnewski and Compen 2022). For example, the increase in reaction time during a decision-making task when theta tACS was applied, has been suggested to reflect an increase in perceived uncertainty (Wischnewski and Compen 2022). Consistently, our findings further provide insight into how theta tACS may influence subjective uncertainty from a computational perspective. Examination of uncertainty estimates using Bayesian framework in probabilistic reversal learning task is increasingly used to provide mechanistic explanations for hierarchical learning (Mathys 2011; Lawson et al. 2017; Pulcu and Browning 2019). The HGF model

updates its parameters trial-by-trial via hierarchical pwPEs to reduce multiple uncertainty. The less confident (more uncertain) an individual is about its estimate of a parameter at some level, the higher it should weigh the information about parameter updates conveyed by PEs (higher learning rate). In the same line, our results also revealed that uncertainty and learning rate have the same correspondence for each level. Specifically, theta stimulation increased both the uncertainty of low-level estimation and low-level learning rates; while decreasing both the uncertainty of high-level estimation and its corresponding learning rate, compared with sham condition.

Notably, impaired uncertainty estimation may be a characteristic of a wide range of neuropsychiatric disorders, including anxiety disorder, autism, and schizophrenia (Kaplan et al. 2016; Lawson et al. 2017; Powers et al. 2017; Pulcu and Browning 2019; Cole et al. 2020; Deserno et al. 2020). For example, Lawson et al. (2017) reported that people with autism spectrum disorder significantly increased volatility learning rates, resulting in worse performance in probabilistic reversal learning. They proposed that these patients may be more willing to be uncertain about the volatility of the world itself, and difficult to tolerate the uncertainty of probability information (Lawson et al. 2017; Pulcu and Browning 2019). In contrast, the present results revealed that enhanced coding of low-level pwPEs in theta stimulation increased informational uncertainty about the probability and increased willingness to update the probability learning at low-level.

Interestingly, our model-based behavioral data found that learning rate at

different hierarchical level showed opposing patterns of change after stimulation. Specifically, frontal theta stimulation increased the low-level learning rate and decreased high-level learning rate relative to sham. Collectively, these results are consistent with previous EEG and fMRI studies. That is, the brain uses hierarchical learning under dynamic environments, and learning signals at different hierarchies can be separately represented in the brain (Iglesias et al. 2013, 2021; Lawson et al. 2017; Deserno et al. 2020; Henco et al. 2020; Liu et al. 2022). For example, low-level pwPEs (served to update the estimate of stimulus probabilities) were encoded by DLPFC, and ACC, while high-level pwPEs (served to update the estimate of volatility) were not. Further, in Liu et al. (2022), we revealed that the trajectories of low-level but not high-level pwPEs were associated with increases in theta power, centered on F4 in the 10-20 EEG system corresponding to rDLPFC. Based on these data, the current results further support the separation of different hierarchical levels of learning. Theta stimulation over rDLPFC strengthens low-level learning, thus leading humans to over-react to low-level probabilistic changes. Concurrently, chasing the environmental probability noise means increasing resistance against environmental volatility's noise. Thereby, theta stimulation leads humans to learn less from the change of high-level volatility.

The “Inverted-U-shaped” function hypothesis proposes that cognitive control requires a dynamic balance between flexible updating and cognitive stabilization (Cools and D’Esposito 2011). The relationship between brain dopamine and performance is inverted-U-shaped, where both excessive and insufficient levels

impair performance (Murphy et al. 1996; Vijayraghavan et al. 2007; Cools and D'Esposito 2011). It is worthy to note that patients with psychiatric disorders had smaller uncertainty about probability contingencies and were thus more resistant to updating the probability level, ultimately leading to worse performance in probabilistic reversal learning (Lawson et al. 2017; Pulcu and Browning 2019; Hein et al. 2021). At the other end of the spectrum, in the current theta stimulation, excessively increased probabilistic updating also impaired performance in this task. Specifically, the theta stimulation significantly increased absolute probability pwPEs and reduced accuracy in the stable environment. Consistent with the inverted-U-shape hypothesis, the present results may indicate that individuals with excessively stable or unstable updates of probability information have worse performance in probabilistic reversal learning. This novel hypothesis may provide a plausible starting point for further empirical work.

One limitation of the current study is that we did not have an active control (e.g., stimulation at the gamma frequency band), in order to pinpoint the functionally specific effect of frontal theta on hierarchical learning. Moreover, we used no personalized neuromodulation frequencies to apply stimulation, which emphasizes that we cannot draw strong conclusions regarding the observed stimulation effects. Finally, to further investigate the neural mechanism underlying the behavioral changes, future research will combine non-invasive brain stimulation with concurrent neuroimaging techniques.

To summarize, the present study combined HD-tACS and a hierarchical

Bayesian model to reveal a causal relationship between probabilistic learning rate during probabilistic reversal learning and theta oscillations over rDLPFC. Behaviorally, the effects of theta stimulation on task performance were influenced by the stability of the environment. Computationally, a hierarchical Bayesian model provided a possible explanation for the change of accuracy, that is, theta tACS induced a significant increase in low-level learning rate and the uncertainty of low-level estimation relative to sham condition. We conclude that learning rates at different levels, primarily through changes in uncertainty estimates, are directly malleable by theta tACS over rDLPFC.

Disclosure Statement

No potential conflict of interest was reported by the authors

Funding

This work was supported by National Science and Technology Innovation 2030 Major Program (no. 2021ZD0203800), National Natural Science Foundation of China (no. 32071049), and Guangdong Basic and Applied Basic Research Foundation, China (no. 2022A1515012185). PV was supported by Research Foundation Flanders grant 1102519; TV and PV were supported by Ghent University Research Council grant BOF17/GOA/004, and by FWO /FNRS EOS Grant GOF3818N.

References

Babapoor-Farrokhran S, Vinck M, Womelsdorf T, Everling S. 2017. Theta and beta synchrony coordinate frontal eye fields and anterior cingulate cortex during sensorimotor mapping. Nat Commun. 8:1–14.

Behrens TEJ, Hunt LT, Woolrich MW, Rushworth MFS. 2008. Associative learning of social value. Nature. 456:245–249.

Behrens TEJ, Woolrich MW, Walton ME, Rushworth MFS. 2007. Learning the value of information in an uncertain world. Nat Neurosci. 10:1214–1221.

Bishop SJ. 2009. Trait anxiety and impoverished prefrontal control of attention. Nat Neurosci. 12:92–98.

Borwick C, Lal R, Lim LW, Stagg CJ, Aquili L. 2020. Dopamine depletion effects on cognitive flexibility as modulated by tDCS of the dlPFC. Brain Stimul. 13:105–108.

Botvinick MM. 2008. Hierarchical models of behavior and prefrontal function. Trends Cogn Sci. 12:201–208.

Botvinick MM, Niv Y, Barto AC. 2009. Hierarchically organized behavior and its neural foundations: A reinforcement learning perspective. Cognition. 113:262–280.

Cavanagh JF, Figueroa CM, Cohen MX, Frank MJ. 2012. Frontal theta reflects uncertainty and unexpectedness during exploration and exploitation. Cereb Cortex. 22:2575–2586.

Cavanagh JF, Frank MJ. 2014. Frontal theta as a mechanism for cognitive control.

Trends in Cognitive Sciences. 18:414–421.

Cavanagh JF, Frank MJ, Klein TJ, Allen JJB. 2010. Frontal theta links prediction errors to behavioral adaptation in reinforcement learning. NeuroImage. 49:3198–3209.

Cohen MX. 2011. Error-related medial frontal theta activity predicts cingulate-related structural connectivity. NeuroImage. 55:1373–1383.

Cohen MX, Elger CE, Ranganath C. 2007. Reward expectation modulates feedback-related negativity and EEG spectra. NeuroImage. 35:968–978.

Cole DM, Diaconescu AO, Pfeiffer UJ, Brodersen KH, Stephan KE. 2020. Atypical processing of uncertainty in individuals at risk for psychosis. NeuroImage Clin. 26:102239.

Cools R, D'Esposito M. 2011. Inverted-U-shaped dopamine actions on human working memory and cognitive control. Biol Psychiatry. 69:e113–e125.

De Berker AO, Rutledge RB, Mathys C, Marshall L, Cross GF, Dolan RJ, Bestmann S. 2016. Computations of uncertainty mediate acute stress responses in humans. Nat Commun. 7:1–11.

Deserno L, Boehme R, Mathys C, Katthagen T, Kaminski J, Stephan KE, Heinz A, Schlagenhauf F. 2020. Volatility Estimates Increase Choice Switching and Relate to Prefrontal Activity in Schizophrenia. Biol Psychiatry. 5:173–183.

Diuk C, Tsai K, Wallis J, Botvinick M, Niv Y. 2013. Hierarchical learning induces two simultaneous, but separable, prediction errors in human basal ganglia. J Neurosci. 33:5797–5805.

Feldman H, Friston K. 2010. Attention, uncertainty, and free-energy. Front Hum Neurosci. 4:215.

Fonteneau C, Redoute J, Haesebaert F, Le Bars D, Costes N, Suaud-Chagny MF, Brunelin J. 2018. Frontal transcranial direct current stimulation induces dopamine release in the ventral striatum in human. Cereb Cortex. 28:2636–2646.

Hein TP, de Fockert J, Ruiz MH. 2021. State anxiety biases estimates of uncertainty and impairs reward learning in volatile environments. NeuroImage. 224:117424.

Henco L, Brandi ML, Lahnakoski JM, Diaconescu AO, Mathys C, Schilbach L. 2020. Bayesian modelling captures inter-individual differences in social belief computations in the putamen and insula. Cortex. 131:221–236.

Iglesias S, Kasper L, Harrison SJ, Manka R, Mathys C, Stephan KE. 2021. Cholinergic and dopaminergic effects on prediction error and uncertainty responses during sensory associative learning. NeuroImage. 226.

Iglesias S, Mathys C, Brodersen KH, Kasper L, Piccirelli M, denOuden HEM, Stephan KE. 2013. Hierarchical Prediction Errors in Midbrain and Basal Forebrain during Sensory Learning. Neuron. 80:519–530.

Kaplan CM, Debjani S, Molina JL, Hockeimer WD, Postell EM, Apud JA, Weinberger DR, Yang TH. 2016. Estimating changing contexts in schizophrenia. Brain. 139:2082–2095.

Kool W, McGuire JT, Rosen ZB, Botvinick MM. 2010. Decision Making and the Avoidance of Cognitive Demand. J Exp Psychol Gen. 139:665–682.

Lawson RP, Mathys C, Rees G. 2017. Adults with autism overestimate the volatility

of the sensory environment. Nat Neurosci. 20:1293–1299.

Liu M, Dong W, Qin S, Verguts T, Chen Q. 2022. Electrophysiological signatures of hierarchical learning. Cereb Cortex. 32:626–639.

Massi B, Donahue CH, Lee D. 2018. Volatility Facilitates Value Updating in the Prefrontal Cortex. Neuron. 99:598-608.e4.

Mathys C. 2011. A Bayesian Foundation for Individual Learning Under Uncertainty. Front Hum Neurosci. 5:39.

Mathys CD, Lomakina EI, Daunizeau J, Iglesias S, Brodersen KH, Friston KJ, Stephan KE. 2014. Uncertainty in perception and the Hierarchical Gaussian filter. Front Hum Neurosci. 8:1–25.

McGuire JT, Botvinick MM. 2010. Prefrontal cortex, cognitive control, and the registration of decision costs. Proc Natl Acad Sci. 107:7922–7926.

Murphy BL, Arnsten AF, Goldman-Rakic PS, Roth R. 1996. Increased dopamine turnover in the prefrontal cortex impairs spatial working memory performance in rats and monkeys. Proc Natl Acad Sci. 93:1325–1329.

Nassar MR, Rumsey KM, Wilson RC, Parikh K, Heasly B, Gold JI. 2013. Rational regulation of learning dynamics by pupil-linked arousal systerms. Nat Neurosci. 15:1040–1046.

Oliveira FTP, Mcdonald JJ, Goodman D. 2007. Performance Monitoring in the Anterior Cingulate is Not All Error Related: Expectancy Deviation and the Representation of Action-Outcome Associations. J Cogn Neurosci. 19:1994–2004.

Panitz M, Deserno L, Kaminski E, Villringer A, Sehm B, Schlagenhauf F. 2022. Anodal tDCS over the medial prefrontal cortex enhances behavioral adaptation after punishments during reversal learning through increased updating of unchosen choice options. Cereb Cortex Commun. 3:1–15.

Powers AR, Mathys C, Corlett PR. 2017. Pavlovian conditioning–induced hallucinations result from overweighting of perceptual priors. Science. 357:596.

Preuschoff K, Bossaerts P. 2010. Adding Prediction Risk to the Theory of Reward Learning. Ann N Y Acad. 1104:135–146.

Pulcu E, Browning M. 2019. The Misestimation of Uncertainty in Affective Disorders. Trends Cogn Sci. 23:865–875.

Rescorla R, Wagner A. 1972. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement.

Senoussi M, Verbeke P, Desender K, De Loof E, Talsma D, Verguts T. 2022. Theta oscillations shift towards optimal frequency for cognitive control. Nat Hum Behav.

Sevgi M, Diaconescu AO, Henco L, Tittgemeyer M, Schilbach L. 2020. Social Bayes: Using Bayesian Modeling to Study Autistic Trait–Related Differences in Social Cognition. Biol Psychiatry. 87:185–193.

Soch J, Allefeld C. 2018. MACS – a new SPM toolbox for model assessment, comparison and selection. J Neurosci Meth. 306:19–31.

Soutschek A, Moisa M, Ruff CC, Tobler PN. 2021. Frontopolar theta oscillations link metacognition with prospective decision making. Nat Commun. 12:1–8.

Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ. 2009. Bayesian Model Selection for Group Studies. NeuroImage. 46.

Sutton R, Barto A. 1998. Reinforcement Learning:An Introduction. MIT Press.

Verbeke P, Ergo K, De Loof E, Verguts T. 2021. Learning to synchronize: Midfrontal theta dynamics during rule switching. J Neurosci. 41:1516–1528.

Verbeke P, Verguts T. 2019. Learning to synchronize: How biological agents can couple neural task modules for dealing with the stability-plasticity dilemma. PLOS Comput Biol. 15:e1006604.

Verbeke P, Verguts T. 2021. Neural synchrony for adaptive control. J Cogn Neurosci. 33:2394–2412.

Verguts T. 2017. Binding by Random Bursts: A Computational Model of Cognitive Control. J Cognitive Neurosci. 29:1103–1118.

Vijayraghavan S, Wang M, Birnbaum SG, Williams G V, Arnsten AFT. 2007. Inverted-U dopamine D1 receptor actions on prefrontal neurons engaged in working memory. Nat Neurosci. 10:376–384.

Voloh B, Valiante TA, Everling S, Womelsdorf T. 2015. Theta-gamma coordination between anterior cingulate and prefrontal cortex indexes correct attention shifts. Proc Natl Acad Sci. 112:8457–8462.

Wischnewski M, Compen B. 2022. Effects of theta transcranial alternating current stimulation (tACS) on exploration and exploitation during uncertain decision-making. Behav Brain Res. 426:113840.

Wischnewski M, Zerr P, Schutter DJLG. 2016. Effects of Theta Transcranial

Alternating Current Stimulation Over the Frontal Cortex on Reversal Learning. Brain Stimul. 9:705–711.

Womelsdorf T, Vinck M, Stan Leung L, Everling S. 2010. Selective theta-synchronization of choice-relevant information subserves goal-directed behavior. Front Hum Neurosci. 4.