Check for updates

Received: 28 March 2022 | Revised: 17 October 2022 | Accepted: 17 November 2022

DOI: 10.1111/psyp.14235

PSYCHOPHYSIOLOGY SPR WILEY

ORIGINAL ARTICLE

Dissociable feedback valence effects on frontal midline theta during reward gain versus threat avoidance learning

Christopher Stolz[1,2,3] | **Alan D. Pickering**[3] | **Erik M. Mueller**[1]

[1]Department of Psychology, University of Marburg, Marburg, Germany

[2]Leibniz Institute for Neurobiology (LIN), Magdeburg, Germany

[3]Department of Psychology, Goldsmiths, University of London, London, UK

Correspondence
Christopher Stolz, Faculty of Psychology, Differential and Personality Psychology, University of Marburg, Gutenbergstr, 18, D-35032 Marburg, Germany.
Email: chris.stolz@staff.uni-marburg.de

Funding information
Deutsche Forschungsgemeinschaft, Grant/Award Number: 422744262–TRR 289 and 290878970–RTG 2271

Abstract

While frontal midline theta (FMθ) has been associated with threat processing, with cognitive control in the context of anxiety, and with reinforcement learning, most reinforcement learning studies on FMθ have used reward rather than threat-related stimuli as reinforcer. Accordingly, the role of FMθ in threat-related reinforcement learning is largely unknown. Here, $n = 23$ human participants underwent one reward-, and one punishment-, based reversal learning task, which differed only with regard to the kind of reinforcers that feedback was tied to (i.e., monetary gain vs. loud noise burst, respectively). In addition to single-trial EEG, we assessed single-trial feedback expectations based on both a reinforcement learning computational model and trial-by-trial subjective feedback expectation ratings. While participants' performance and feedback expectations were comparable between the reward and punishment tasks, FMθ was more reliably amplified to negative vs. positive feedback in the reward vs. punishment task. Regressions with feedback valence, computationally derived, and self-reported expectations as predictors and FMθ as criterion further revealed that trial-by-trial variations in FMθ specifically relate to reward-related feedback-valence and not to threat-related feedback or to violated expectations/prediction errors. These findings suggest that FMθ as measured in reinforcement learning tasks may be less sensitive to the processing of events with direct relevance for fear and anxiety.

KEYWORDS

computational model, EEG, prediction error, punishment avoidance, reinforcement learning, theta

1 | INTRODUCTION

Approaching rewards and avoiding punishments are fundamental aspects of reinforcement learning (Sutton & Barto, 1998) and action monitoring (Ullsperger et al., 2014). Activity in the medial prefrontal cortex (mPFC) has been proposed to be centrally involved in a neural action monitoring network (Badre & Nee, 2018; Ridderinkhof et al., 2004). Particularly, the anterior cingulate cortex (ACC) has been identified as a neural hub that implicates a wide range of processes including cognitive control, conflict processing, action selection (Rushworth et al., 2011; Shackman et al., 2011;

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
© 2022 The Authors. *Psychophysiology* published by Wiley Periodicals LLC on behalf of Society for Psychophysiological Research.

https://doi.org/10.1111/psyp.14235

14698986, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/psyp.14235 by Test, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Vassena et al., 2017), learning of reinforcement contingencies (Behrens et al., 2007; Rushworth & Behrens, 2008), and the computation of reward expectations and expectation violations, i.e. prediction errors (Alexander & Brown, 2019; Sambrook & Goslin, 2014, 2015, 2016; Silvetti et al., 2011, 2014). While frontal brain processes involved in action monitoring largely have been studied in the context of reward-related reinforcement learning, the neural dynamics of punishment avoidance learning are less well known.

The use of electroencephalography (EEG) techniques to measure mid-frontal neural processing in reinforcement learning has revealed the frontal midline 4–8 Hz theta (FMθ) rhythm to be of particular relevance for action monitoring (Cavanagh, Zambrano-Vazquez, & Allen, 2012; Wang et al., 2020). FMθ activity is amplified in response to negative feedback indicating monetary non-reward and monetary loss (Bernat et al., 2015; Cohen et al., 2007; Gruber et al., 2013; Mueller et al., 2015), and in goal-directed control (Cavanagh et al., 2013; Cavanagh, Figueroa, et al., 2012; Cohen & Cavanagh, 2011; Cooper et al., 2019; Pinner & Cavanagh, 2017). While this may be interpreted within the viewpoint that FMθ relates to processing events that are worse than expected (i.e. negatively signed prediction error), other studies have also suggested that any surprising feedback outcomes may increase feedback-locked FMθ activity (i.e. unsigned prediction error; Cavanagh, Figueroa, et al., 2012; Gheza et al., 2019; Mas-Herrero & Marco-Pallarés, 2014; Rawls et al., 2020; van de Vijver et al., 2014). Overall, FMθ seems to reflect activity of a neural system that is involved in the processing of prediction errors, cognitive control, and feedback-driven learning.

While the majority of studies on FMθ in reinforcement learning investigated monetary reward and monetary non-reward or loss (Cavanagh, Zambrano-Vazquez, & Allen, 2012), it is still unknown whether FMθ functioning in reward learning tasks is similar or different to punishment avoidance learning with primary threat reinforcers. Several theories suggest separate neural systems are involved in reward gain vs. punishment avoidance learning (Corr, 2008; Corr & McNaughton, 2012; Daw et al., 2002; Gray & McNaughton, 2000; Smillie et al., 2006). In addition, neuroimaging studies report the involvement of different systems in monetary gain vs. loss avoidance (Menegas et al., 2018; Palminteri et al., 2015; Palminteri, Justo, et al., 2012; Rigoli et al., 2016; Seymour et al., 2007; Yacubian et al., 2006). Given that reward gain and punishment avoidance learning recruit non-identical neural systems, it is an open question whether FMθ oscillations, which have been observed in response to reward-related negative feedback, are also relevant for processing punishment-related negative feedback.

In fact, FMθ has been related to processing of punishment cues. It is amplified in response to classical conditioned threat (Mueller et al., 2014; Sperl et al., 2018), unsuccessful vs. successful avoidance of unpleasant sounds (van Noordt et al., 2018), and there is evidence that dispositional anxiety moderates FMθ effects in reinforcement learning (Cavanagh & Shackman, 2015). This suggests that FMθ may be sensitive for feedback signaling aversive reinforcers and converges with the claim of the adaptive control hypothesis (Cavanagh & Shackman, 2015; Shackman et al., 2011), stating that threat scenarios demand a higher need for neural processes involved in cognitive control, which in turn are indexed by FMθ activity. However, to our knowledge, the sensitivity of FMθ to feedback valence and prediction errors in the context of threat-related punishment avoidance has not yet been investigated.

To address this question, we used time-frequency analyses to study feedback processing in one reward-related and one punishment-related two-choice reversal learning task. While feedback in the reward task signaled monetary reward (+10 Cent) vs. non-reward (+0 Cent), feedback in the punishment task signaled non-punishment (no noise burst) vs. punishment (noise burst). Importantly, to improve the comparability between the tasks, the intensity of the noise burst was individually titrated to match the aversiveness of monetary non-reward. For each task, a reinforcement learning model was fitted to behavior providing separate learning rates for gain and loss and computationally derived prediction errors (Cavanagh et al., 2010; Frank et al., 2007). For a comprehensive assessment of prediction errors, self-reported prediction errors were also calculated based on every participants' trial-by-trial subjective reward and non-punishment expectation ratings, respectively (Hajcak et al., 2007; Weismüller & Bellebaum, 2016). To investigate the relationship between prediction errors and FMθ, the computationally derived and self-reported prediction errors were entered into within-subjects single-trial regression analyses with FMθ as criterion. We hypothesized that FMθ would be amplified in response to negative vs. positive feedback, regardless of whether negative feedback indicated monetary non-reward or an aversive noise burst. Moreover, if FMθ reflected an unsigned prediction error signal, it should covary with the degree of feedback unexpectedness in both tasks. We also calculated perseverations after reversal, post-error slowing (PES), and post-correct speeding (PCS) as indices of successful reversal learning and behavioral adjustment, respectively.

2 | METHODS

2.1 | Participants

Twenty-six participants were recruited from the University of Marburg and were included only if they

14698986, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/psyp.14235 by Test, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

were right-handed, German native speakers[1] with normal, or corrected-to-normal vision. One participant was excluded due to technical issues with saving EEG data during the tasks. Two further participants were excluded because of excessive artifacts in the EEG during the punishment task. The final sample consisted of 23 participants (13 female; mean age of 24.30 years, SD = 3.31). The local ethics committee approved the study.

2.2 | Procedure

After reading and signing informed consent, participants were seated in the laboratory in front of a 24″ LED computer monitor (Dell U2412Mc; 1920 × 1080 px). The stimuli were presented using Neurobehavioral Systems Presentation 18.2 software (www.neurobs.com). At the beginning, the participants underwent a white noise titration procedure (see below) and subsequently provided a baseline valence rating of the four different feedback events and provided an unpleasantness rating of the previously titrated noise burst. Subsequently, participants underwent, in counterbalanced order, the reward and the punishment versions of a 280-trial probabilistic reversal learning task (see below). After each reversal learning task, the participants rated how rewarding the desired consequences (monetary reward/not receiving a noise burst) and how punishing the undesired consequences (monetary non-reward/receiving the noise burst was) were perceived to be, and again rated the valence of the feedback cues. To assess whether the noise burst and monetary non-reward were still perceived as similarly aversive throughout the experiment, the participants underwent another cycle of the white noise burst titration procedure after each reversal learning task. The participants were informed in the participation information that they were able to win a maximum of 10 € in the reward task, which all participants achieved.

2.3 | Self-reported feedback cue and consequences ratings

To control for initial differences on how the feedback events were perceived, participants rated the positive vs. negative valence of the feedback cues at the beginning of the experiment and after each reversal learning task on a bipolar scale (−4 to +4, unpleasant, negative vs. pleasant, positive). A unipolar scale was used to measure the unpleasantness of the individually reward-matched white noise burst (0–4, “How unpleasant was the loud noise?”). After each reversal learning task participants rated how rewarding they perceived the desired feedback consequence (monetary reward & no noise burst) and how punishing they perceived the non-desired feedback consequence to be (monetary non-reward & noise burst) on a unipolar scale (0–4, “How rewarding did you feel it was to gain 10 Cents/not experience the loud noise?” and “How punishing did you feel it was to gain no money/experience the loud noise?”).

2.4 | White noise burst titration

As the goal of this study was to compare feedback-locked theta in the context of punishment vs. non-reward, we took particular care in calibrating the subjective value of a punishment stimulus (i.e., the noise burst) to match the subjective value of a reward stimulus (i.e., a 10-cent monetary reward). To this end, we applied the following procedure: participants were presented a noise burst for 1000 ms with an intensity of 81, 84, 87, and 90 dB and were asked every time, if they would be willing to receive another burst presentation for a monetary reward of 10 Cent. If they affirmed, the respective burst was presented again. After the first habituation-cycle, this procedure was repeated, again starting with 81 dB and continuing with increasing intensities. The intensity level that was one level higher than the level at which the participant would still listen to the burst for the monetary reward during the second cycle was set as the “reward value equivalent burst intensity.” It was used for the punishment task, with the following distribution of intensities across participants: $N = 2$ (8.70%) 84 dB, $N = 6$ (26.09%) 87 dB, $N = 15$ (65.22%) 90 dB. The average unpleasantness rating (0–4) of the individually reward-matched noise burst was M = 3.04 (SD = 0.78). The maximum intensity of the noise burst was restricted to 90 dB on ethical grounds. To measure potential habituation effects throughout the task, one cycle was repeated after each reversal learning task (cycle 3 and cycle 4). Table 1 presents the percentage of “yes” choices for every cycle[2]. To investigate the stability of the last “yes” choices over the course of the experiment, we calculated a choice difference score from cycle 2 to cycle 3 and 4 (cycle 2 minus cycle 3, cycle 2 minus cycle 4). A score of

---

[1] The instructions given in the task description below are translations into English from the German original.

[2] An ANOVA including the factor cycle (one, two, three & four) on the latest “yes” choice was not significant, $F(1, 20) = 0.97$, $p = .43$, $\eta_p^2 = 0.13$. Moreover, the latest “yes” choices in the first versus fourth cycle were positively correlated ($r_\rho = .79$, $p < .001$) suggesting a relatively high stability of the reward value equivalent burst intensity over the course of the experiment.

14698986, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/psyp.14235 by Test, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

0 means no change, negative scores indicate the choice of a lower dB step and positive scores a higher dB step relative to the reward-matched noise burst selected in cycle 2. From the 23 participants, 14 (60.87%) showed no change from cycle 2 to 3 and 4 (score = 0), 3 participants (13.04%) showed a decrease (M = −1.16 steps), and 6 (26.09%) an increase (M = 1.25 steps) from cycle 2 to either cycle 3 or 4. Accordingly, the last "yes" choices in cycle 2 were positively correlated with those in cycle 3, $r_\rho = 0.90$, $p < .001$, and cycle 4, $r_\rho = .82$, $p < .001$. In general, the choices showed only slight changes over the course of the experiment, suggesting that the noise burst intensity was relatively stably related to the monetary reward value.

TABLE 1 Percentages of the noise burst titration choices for every cycle. The intensity of the white noise burst for punishment task was set in cycle 2

	81 dB (%)	84 dB (%)	87 dB (%)	90 dB (%)
Cycle 1	86.96	86.96	65.22	56.52
Cycle 2	86.96	91.30	65.22	47.83
Cycle 3	82.61	82.61	60.87	56.52
Cycle 4	82.61	69.57	65.22	52.17

2.5 | Reversal learning task

As seen in Figure 1b, each trial of the 280-trials reversal learning tasks began with a crosshair that was displayed for 1000–4000 ms. Thereafter, two shapes (hexagon and triangle or circle and diamond) were presented left and right from a crosshair until the participant selected one of them via button press with the right hand. The left vs. right position of the shapes was randomized from trial to trial. Once a shape was chosen, it was highlighted with a white frame which was shown for 1000 ms. Participants were then asked to report the strength of their expectation to receive positive feedback on a unipolar 0 to 3 scale

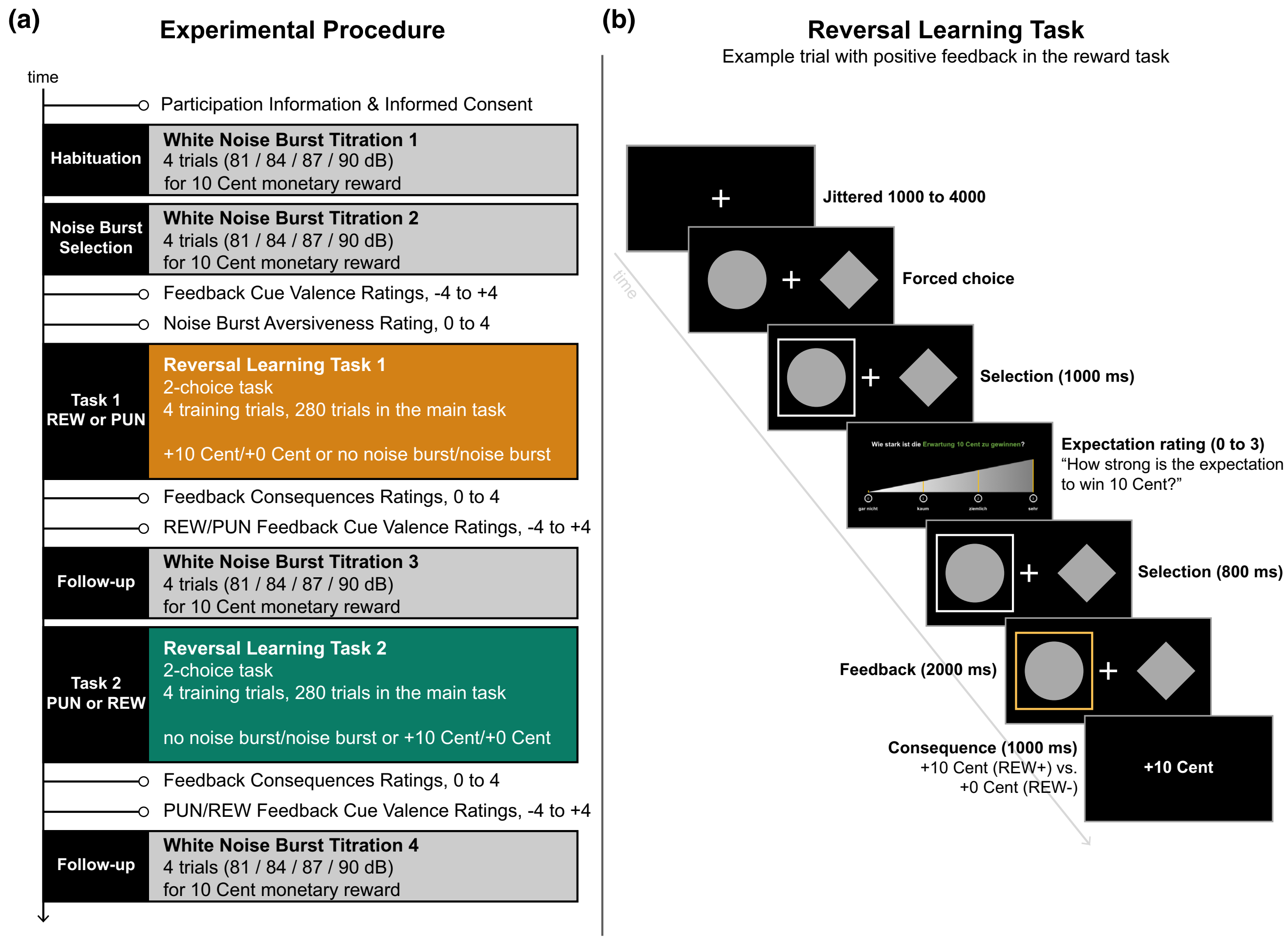

FIGURE 1 (a) Overview of the experimental procedure. The order of the reward and punishment reversal learning tasks was counterbalanced between participants. (b) Example trial for the reversal learning task in the reward context with circle vs. triangle as choices and orange vs. teal as feedback cues.

14698986, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/psyp.14235 by Test, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

("How strong is the expectation to win 10 Cent/not have to listen to the loud noise?"). The expectation rating responses were self-paced and assessed via button press of the left hand. The shapes with the selection indicated by the white frame were presented for another 800 ms. Thereafter, the frame changed its color to either orange/teal or green/rose, indicating positive/negative feedback for 2000 ms. In the reward task, positive feedback indicated reward (REW+) and negative feedback non-reward (REW−). In the punishment task, positive feedback indicated non-punishment (PUN−) and negative feedback punishment (PUN+). The actual reinforcer was presented for 1000 ms and was different for the reward and the punishment task: In the reward task, reward and non-reward consequences were visually displayed on the screen reading "+10 Cent" or "+0 Cent," respectively. In the punishment task, punishment was delivered as a white noise burst (with the reward-value equivalent burst intensity) for 1000 ms via a speaker while non-punishment was the absence of a noise burst. At the beginning of every reversal learning task, four practice trials were presented under the supervision of the experimenter with letters as selection cues instead of shapes (P vs. Y and V vs. F). See Table S7 for details on the permutation of shapes, feedback colors, and task sequence.

The contingencies were 70/30 for positive/negative feedback for selecting one shape and 30/70 positive/negative feedback for the other shape. In the first reversal learning task that was performed, contingencies reversed after 82, 150, and 225 trials, in the second reversal learning task after 86, 160, and 223 trials (see Figure 2a). Participants were instructed that the frequencies of positive and negative feedback of the shapes would reverse several times during the task. Participants were able to pause after the ongoing trial whenever they pressed the pause button (reward task: M = 1.87 pauses, SD = 1.69; punishment task: M = 2.13 pauses, SD = 1.66; no significant difference between the tasks, $t(22) = 0.84$, $p = .41$, $d = 0.16$). Participants were instructed to choose as fast as possible.

The shapes were hexagon vs. triangle or circle vs. diamond (shape size: 8.47 × 8.47 cm) starting with hexagon or circle as the 70/30 cue. The frame colors indicating positive and negative feedback were orange vs. teal and green vs. rose (frame size: 10.37 × 10.09 cm, contour width: 0.42 cm). An ANOVA including the factor frame color (orange, teal, green, & rose) on the feedback cue valence ratings before the reversal learning tasks was not significant (orange: M = 0.48; teal: M = 0.91; green: M = 0.14; rose: M = 0.62), $F(1, 22) = 1.54$, $p = .21$, $\eta_p^2 = 0.07$. The shape pairs and feedback pairs, as well as the sequence of the reward task and punishment task, were permuted randomly across the participants (48% started with the reward task).

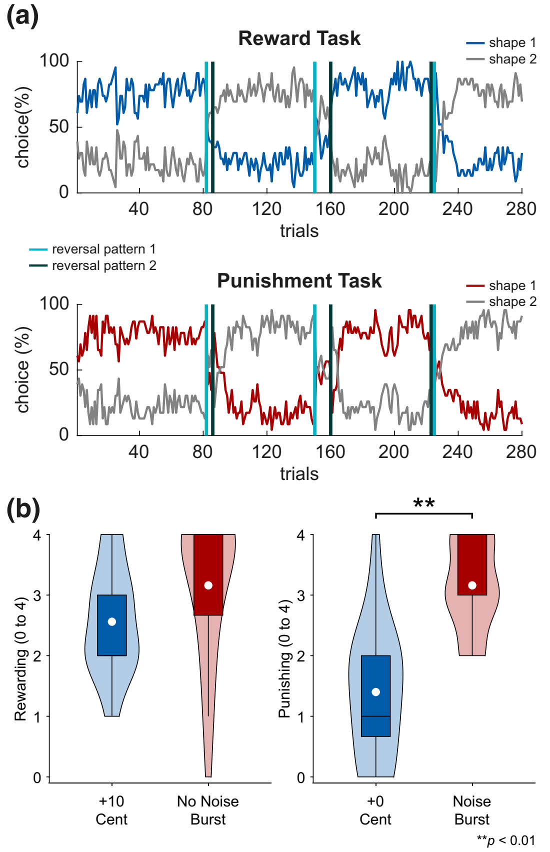

FIGURE 2 (a) Trial-by-trial percentages of choices between the shape 1 and shape 2 before and after reversals for the reward and punishment tasks. The vertical lines denote the position of the reversals which varied slightly depending on whether the task was given first (pattern 1) or second (pattern 2). (b) Violin plot with mean values depicted as white circles for ratings on how rewarding the desired (left panel) and how punishing the undesired (right panel) feedback consequences (+10 cent/no noise burst, +0 cent/noise burst) were subjectively perceived.

2.6 | Computational modeling and derivation of prediction error variables

2.6.1 | Computational modeling

The trial-by-trial choices were fit for every subject and task by a reinforcement learning model (Q-learning; Sutton & Barto, 1998). That is, for every choice option (choosing shape 1 or shape 2), an expected value (*Q* value) was calculated which was updated based on previous learning experiences with the choice, the current feedback outcome, and a learning rate (α_{REW} & α_{PUN}). The *Q* values for the shape *i* chosen on trial *t* were updated on each trial based on the *Q* value and feedback in trial *t*. The updated values were used in trial $t+1$. Since the reward task demanded reward maximization and the punishment task punishment minimization, the reinforcements were

14698986, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/psyp.14235 by Test, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

set to 1 (REW+) and 0 (REW−) for the reward task and 0 (PUN−) and −1 (PUN+) for the punishment task (e.g. Betts et al., 2020)[3]:

$$Q_i(t+1) = Q_i(t) + \alpha \times \left[R(t) - Q_i(t)\right]$$

The trial-by-trial probabilities of choices were calculated using a softmax function with an inverse gain parameter (β) which determines a more exploratory (lower beta) vs. exploitative (higher beta) choice strategy. The probabilities were used to calculate the log likelihood estimate (LLE) over all the trials of the task:

$$\mathrm{LLE} = \sum_t \ln\left(\frac{\exp\left(\beta \times Q_i(t)\right)}{\sum_i \exp\left(\beta \times Q_i(t)\right)}\right)$$

To minimize the negative LLE, we used the MatLab 2015a (MathWorks) function *fmincon* (Bai et al., 2014; Collins & Frank, 2016; Fischer & Ullsperger, 2013) and started the minimization from ten random starting points. The lower and upper bounds for β were [0, 10] and [0, 1] for every other parameter. The maximum of iterations until convergence was set to 10,000 (Palminteri et al., 2015).

2.6.2 | Prediction errors

The learning rates of the fitted model for each subject were used to calculate an individual's trial-by-trial *Q* values and model-estimated prediction errors (ePEs):

$$\mathrm{ePE}(t) = \mathrm{R}(t) - Q(t)$$

The self-reported prediction errors (sPE) were calculated as the trial-by-trial difference between the feedback R and the self-reported expectation SR. To set the boundaries of the sPEs to −1 and 1, the SRs which ranged from 0 to 3 were divided by three [0, 1/3, 2/3, 1]:

$$\mathrm{sPE}(t) = R(t) - \mathrm{SR}(t)$$

2.6.3 | Model comparisons

The first model was a single learning rate model (M1: Vanilla) which included α and β. Since the participants showed an exploitative behavioral strategy (see performance results), we also fit variants of the M1 model which updated not only the *Q* value of the selected but also the non-selected option. Model two (M2: DualUpdate) included α, β, and κ and has already been used in similar reversal learning tasks (Marković et al., 2019; Reiter et al., 2016, 2017). In this model, the *Q* value of the non-selected option was updated as if the outcome was the opposite of the selected option (e.g., if R(t) = 1 then R_f(t) = 0 in the reward task) with κ denoting the coupling strength of the PE:

$$\begin{aligned} &Q(t+1)_{\text{non-selected}} \\ &= Q(t)_{\text{non-selected}} + \kappa\alpha \times \left[\mathrm{R\_f}(t) - Q(t)_{\text{non-selected}}\right] \end{aligned}$$

In the present study, self-paced expectation ratings were assessed which might have influenced the time interval between action and feedback onset. Previous studies demonstrated that the duration of this time interval may affect electrocortical feedback processing (Arbel et al., 2017; Krigolson, 2018; Weinberg et al., 2012). To assess the influence of action-feedback time interval, a third model was fit to the data (M3: Time), including a time decaying parameter λ which scaled the influence of the single-trial time interval between action and feedback onset (in ms) on *Q* values. The upper and lower bounds for λ were [1, Inf]:

$$Q_i(t+1) = Q_i(t) + \alpha \times (\lambda \times 1/\mathrm{time}) \times \left[\mathrm{R}(t) - Q_i(t)\right]$$

In a fourth model (M4: Decay), the non-selected option was updated with a free decay parameter (γ) which reduced the *Q* values for the non-selected option in every positive feedback trial (Cavanagh, 2015). Finally, in the fifth model (M5: DecayAll) γ updated the *Q* values for the non-chosen shape in every trial regardless of the feedback outcome:

$$\begin{aligned} &\text{Reward task: } Q(t+1)_{\text{non-selected}} \\ &= (1-\gamma) \times \left(1 + Q(t)_{\text{non-selected}}\right) \end{aligned}$$

$$\begin{aligned} &\text{Punishment task: } Q(t+1)_{\text{non-selected}} \\ &= (1-\gamma) \times \left(1 + Q(t)_{\text{non-selected}}\right) - 1 \end{aligned}$$

As indices for model fits, we computed pseudo-R^2 statistics for every participant by comparing the individual LLE to chance (e.g. Cavanagh, 2015; Collins & Frank, 2014; Gershman & Tzovaras, 2018), and the Akaike information criterion (AIC) and Bayesian Information Criterion (BIC) for comparisons between different models.

---

[3] The use of different reinforcement sets for the reward vs punishment task does not affect the calculation of the trial-by-trial probabilities, and thus not the model fitting. It just reflects a more natural valuation: Q values in the reward task ranged from 0 to 1 with increasing values after monetary reward. The Q values in the punishment task ranged from −1 to 0 with an increase in Q values moving closer to 0 (i.e., a decrease in the size of the negative values) after successful punishment avoidance.

14698986, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/psyp.14235 by Test, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

2.7 | Behavioral measures

Win-stay behavior was defined as two subsequent choices of the same shape after positive feedback and lose-shift as a shift from one choice to another after receiving negative feedback. The frequency of win-stay/lose-shift strategies was calculated as percentage of stay choices after positive feedback and the percentage of shift choices after negative feedback, respectively (Cavanagh, 2015; Frank et al., 2007; Pizzagalli et al., 2008). Post-error slowing (PES) and post-correct speeding (PCS) reaction times (RTs) were computed as pairwise comparisons around each error/correct trial, i.e. the average RTs of the trial after positive/negative feedback minus the average of the RTs in the trial before (Dutilh et al., 2012). Perseverative errors after a reversal were measured as every sequence of more than two subsequent choices of the disadvantageous shape (i.e. 30% positive feedback) leading to negative feedback (den Ouden et al., 2013).

2.8 | EEG recording and pre-processing

The EEG was recorded using a 64-channel ActiveTwo system (BioSemi) at 1024 Hz sampling rate. EEG preprocessing was performed using BrainVision Analyzer 2 (Brain Products). EEG was downsampled to 512 Hz and filtered using a 1 Hz high-pass filter and 50 Hz notch-filter to optimize the independent component analysis (ICA) solution in detecting components reflecting eye-blinks. The EEG was re-referenced to the average reference and then manually screened for muscle artifacts before and after an infomax ICA. ICA components reflecting eye-blinks were removed by a trained rater (author CS). The artifact markers of the raw data inspections and the resulting ICA weight matrices were saved and applied to the data with the same preprocessing steps as described above, but with a 0.1–100 Hz bandpass-filter (see Winkler et al., 2015). Trials including artifacts marked in the raw data inspection were rejected. On average, there were 14% bad epochs in the reward task (13.60% positive feedback, 14.40% negative feedback) and 17.20% bad epochs in the punishment task (15.60% positive feedback, 18.90% negative feedback). An ANOVA on the percentage of bad epochs revealed a significant effect of task (reward vs. punishment task), $F(1, 22) = 7.57$, $p < .05$, $\eta_p^2 = 0.26$. The main effect feedback valence (positive vs. negative feedback) and the interaction of feedback valence and task did not reach significance (Fs ≤ 4.28, ps $\geq .05$, η_ps ≤ 0.16; larger effect for the interaction). The final EEG data were segmented into epochs ranging from −1500 to 3000 ms relative to feedback onset.

2.9 | Wavelet analysis and event-related potentials

The wavelet analysis was performed by using established MatLab custom scripts (Keil et al., 2007; Panitz et al., 2019). EEG segments were baseline corrected in the time domain from −1500 to −1400 ms and cosine-square-tapered (20 samples). Power was computed by applying complex morlet wavelets with $m = f/\sigma_f = 7$, resulting in power for the frequency bands from 0.22 to 50 Hz in steps of 0.22 Hz. Thereafter, each epoch was cut in length ranging from −300 to 2000 ms relative to feedback onset. Finally, the power was normalized to a decibel (dB) scale for every frequency using the averaged power from −300 to −200 ms. Based on previous studies, the mean theta power was scored as average power in the time-frequency region of 4–8 Hz and a time window 250–500 ms after feedback onset (Cavanagh, Figueroa, et al., 2012; Padrão et al., 2012). As grand average theta power was maximal at frontocentral electrodes, average FMθ was computed across Fz, F1, F2, FCz, FC1, and FC2 electrode sites.

2.10 | Single-trial regressions

To examine the associations between FMθ and prediction errors, single-trial regression analyses with either the standardized ePEs or the sPEs as predictor and the standardized theta power as criterion were conducted for the reward and punishment task, respectively (Cohen & Cavanagh, 2011; Cooper et al., 2019). The prediction errors were transformed to absolute values and thus ranged from 0 to 1. For every participant, single-trial regressions were calculated with positive prediction errors (positive feedback trials; PE+), and negative prediction errors (negative feedback trials; PE-) as predictor and every data point of single-trial power as criterion. Since there was an a priori defined time-frequency window (4–8 Hz, 250–500 ms), this analysis was resulting in a channel (Fz, F1, F2, FCz, FC1, & FC2) × frequency (4–8 Hz) × time (250–500 ms) matrix of beta slopes. This 6 × 6 × 129 matrix then was averaged, resulting in one average slope for every participant. This procedure was conducted for both the reward task and the punishment task resulting in 2 (reward & punishment task) × 2 (ePE & sPE) × 2 (positive & negative prediction error) mean beta values for every participant. Following the same protocol, single-trial analyses were conducted with FMθ as regressor and RT slowing (RT current minus previous trial) in the subsequent trial as criterion, and logistic regressions analysis with FMθ as regressor and switch vs. stay behavior in the subsequent trial as criterion. This was separately done for

14698986, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/psyp.14235 by Test, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

positive and negative feedback trials and the reward and punishment task.

2.11 | Statistical analysis

ANOVAs including the factors task (reward vs. punishment) and feedback valence (positive vs. negative) were performed on feedback cue valence ratings and EEG measures. Between-task comparisons of α, γ, β model fit indices, win-stay and lose-shift proportions, feedback consequences ratings, PES, and PCS were conducted using t tests. To further analyze the feedback valence effects in the reward and punishment task, Bayes factors (BF_{10}) were computed using the bayesFactor toolbox (Krekelberg, 2022); see Kass and Raftery (1995); Morey and Wagenmakers (2014); Rouder et al. (2012); Schönbrodt and Wagenmakers (2018). Perseverative errors were subjected to an ANOVA with the factors task (reward vs. punishment) and reversal block (1, 2, & 3). Mean beta slopes reflecting the association between prediction errors and theta power were tested against zero using t tests. These beta slopes were further subjected to an ANOVA including the factors task (reward vs. punishment), sign (PE+ vs. PE−), and prediction error type (ePE vs. sPE). The repeated measures ANOVA statistics are reported with Greenhouse–Geisser correction where appropriate. Post hoc t tests were performed in case of significant interaction effects and were reported with Bonferroni adjusted p values for multiple comparisons ($p = .05/4 = .0125$). All reported correlation coefficients are Spearman's rho rank correlations (r_ρ). The reported effect sizes were partial eta squared (η_p^2) and Cohen's d. For all of these statistical analyses, SPSS Statistics Version 25 (IBM) was used.

3 | RESULTS

3.1 | Performance

There was an average of 168.91 (SD = 10.92) positive feedback and 111.09 negative feedback trials in the reward task and 169.22 (SD = 9.22) positive feedback and 110.78 negative feedback trials in the punishment task. As seen in Figure 2a, the choices with regard to shapes (choosing shape 1 or shape 2) and behavioral adjustments after reversals were comparable between the reward and punishment tasks. Overall, participants chose the same shape after positive feedback in 79.48% (SD = 8.02) of the cases in the reward task and 82.53% (SD = 8.31) in the punishment task (win-stay). Participants shifted after negative feedback in only 22.79% (SD = 10.62) of the cases in the reward task and 20.00% (SD = 9.10) in the punishment task (lose-shift). The percentage of win-stay or lose-shifts did not differ between the tasks (ts$\leq$1.84, ps$\geq$.08, ds$\leq$0.32). Overall, this suggests that, regardless of the task, participants did not tend to follow a simple win-stay/lose-shift strategy but were persistent with their choice even after receiving negative feedback (the lose-shift percentages were similarly low in either task). Thus, we also fit dual updating models, which have been demonstrated to better fit to exploitative action selection strategies (Cavanagh, 2015; Marković et al., 2019; Reiter et al., 2017).

3.2 | Statistical power analysis

Sensitivity power analyses were conducted using GPower with $N = 23$, $\alpha = 0.05$ (two-tailed), and a power = 0.80 ($1-\beta$). These analyses revealed that the present study's sample size was suited to detect medium to large effect sizes using ANOVAs with four repeated measures ($\eta_p^2 \geq 0.09$) that were calculated for the task and feedback valence effects on FMθ. For one sample t tests, i.e. within-subjects comparisons and t tests against zero, the present study's sample size was suited to detect medium to large effect sizes ($d \geq 0.61$).

As indicated by the AIC, BIC, and pseudo-R^2, the models fit increasingly better (M1 to M5); see Table S1 in the supplements. The best fitting model (M5) suggests that the participants' expectation levels for the desired outcome (Q values) of non-chosen options diminished regardless of feedback valence. The Q values and ePEs were calculated using the best fitting parameters of each participant in M5; see Table 2. The model M5 showed better fitting indices in the punishment vs. reward task, $t(22) = 2.13$, $p < .05$, $d = 0.50$. The free parameters were not significantly different between the tasks, (ts$\leq$0.81, ps$\geq$.43, ds$\leq$0.17), suggesting that learning performance was comparable for reward gain learning and punishment avoidance learning.

To assess the similarity of the computationally derived parameter for expectations (Q-values) and the trial-by-trial self-reported positive feedback expectation we assessed

TABLE 2 Mean (SD) model parameter and fit indices for the best fitting model (M5: DecayAll)

	α	β	γ	AIC	BIC	Pseudo-R^2
Reward task	0.44 (0.31)	3.85 (1.82)	0.61 (0.37)	−240.00 (56.68)	−229.09 (56.68)	0.37 (0.15)
Punishment task	0.43 (0.28)	4.14 (1.22)	0.58 (0.37)	−211.56 (58.63)	−200.66 (58.63)	0.44 (0.15)

14698986, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/psyp.14235 by Test, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

within-subject correlations of the respective time-series for each participant. The average within-subject correlation was $r_\rho = .24$ (Min = −0.19, Max = 0.77, SD = 0.27) for the reward task and $r_\rho = .25$ (Min = −0.23, Max = 0.76, SD = 0.20) for the punishment task. The within-subject correlations for the reward (Fisher's $z = 0.27$) and punishment tasks (Fisher's $z = 0.27$) did not differ from each other, $t(22) = 0.01$, $p = .99$, $d < 0.01$, and showed a between-task correlation across participants, $r_\rho = .58$ ($p < .01$). Together, this indicates that (a) computational indices and self-report markers reflect two substantially different perspectives on expectations with only a very modest degree of overlap and (b) subjects with relatively higher convergence of subjective and computationally derived expectations in the reward task also have a relatively higher convergence of subjective and computationally derived expectations in the punishment task and vice versa.

3.3 | Reaction times

Overall, RTs in the reward task (M = 761.76 ms, SD = 37.94) were numerically faster than those in the punishment task (M = 832.46 ms, SD = 181.96), although the difference was at the borderline for statistical significance, $t(22) = 2.10$, $p = .05$, $d = 0.43$. As given in Table 3, PES was significantly increased in the punishment vs. reward task, $t(22) = 3.34$, $p < .01$, $d = 0.79$, suggesting slower RTs after negative feedback in the punishment task. Moreover, PCS was greater in the punishment vs. reward task, $t(22) = 3.25$, $p < .01$, $d = 0.71$, i.e. participants reacted faster after positive feedback in the punishment task. There were no significant effects of task or reversal block on the amount of perseverative choices after reversal ($F\text{s} \leq 1.38$, $p\text{s} \geq .27$, $\eta_p^2 \leq 0.12$); see Table 3.

3.4 | Self-report feedback cue and consequences ratings

3.4.1 | Feedback cue ratings

There was a significant main effect of feedback valence, $F(1, 22) = 30.37$, $p < .01$, $\eta_p^2 = 0.58$ as positive feedback cues were perceived as more appetitive than negative feedback cues[4]. The task × feedback valence interaction also reached significance, $F(1, 22) = 12.29$, $p < .01$, $\eta_p^2 = 0.36$, indicating that the difference between positive and negative feedback was larger for the punishment (M = 3.76, SD = 2.63) vs. reward (M = 1.81, SD = 2.90) task; see Table 3. The main effect of task on self-reported valence ratings of the feedback cue did not reach significance, $F(1, 22) = 2.06$, $p = .17$, $\eta_p^2 = 0.09$.

TABLE 3 Mean (SD) values of behavioral and rating variables in the reward and punishment task

	Reward task	Punishment task
Post-error slowing (PES; ms)	27.01 (71.26)	100.28 (107.27)
Post-correct speeding (PCS; ms)	23.47 (47.99)	68.08 (72.29)
Perseverations reversal block 1	3.09 (1.65)	3.13 (2.22)
Perseverations reversal block 2	3.04 (2.44)	3.52 (1.70)
Perseverations reversal block 3	3.35 (2.06)	2.83 (1.83)
Positive feedback cue ratings	1.43 (1.64)	2.19 (1.47)
Negative feedback cue ratings	−0.38 (1.74)	−1.57 (1.67)

3.4.2 | Feedback consequences ratings

The absence of the noise burst (M = 3.13, SD = 1.22) was numerically rated as more rewarding than the reception of a monetary reward (+10 Cent; M = 2.52, SD = 0. 85), although the difference was at the borderline for statistical significance, $t(22) = 2.13$, $p = .05$, $d = 0.58$; see Figure 2b. Being confronted with the noise burst was rated as significantly more punishing (M = 3.32, SD = 0.78) than experiencing monetary non-reward (+0 Cent; M = 1.26, SD = 0.96), $t(22) = 7.40$, $p < .001$, $d = 2.18$; see Figure 2b.

4 | EEG MEASURES

4.1 | Event-related FMθ power

For the event-related theta power, the main effect of task did not reach significance, $F(1, 22) = 1.00$, $p = .33$, $\eta_p^2 = 0.04$. There was a significant effect of Feedback Valence (positive vs. negative), $F(1, 22) = 12.72$, $p < .01$, $\eta_p^2 = 0.37$, with increased power following negative (M = 1.16 dB, SEM = 0.09) compared with positive feedback (M = 0.85 dB, SEM = 0.06). Importantly, this main effect was further qualified by a significant interaction of task × feedback valence, $F(1, 22) = 5.38$, $p < .05$, $\eta_p^2 = 0.20$; Figure 3. Post hoc comparisons revealed that feedback-locked FMθ power was potentiated in response to REW− vs. REW+, $t(22) = 3.93$, $p < .01$, $d = 1.01$. This feedback valence effect was substantially weaker for the comparison PUN+ vs. PUN−, $t(22) = 2.15$, $p < .05$, $d = 0.49$ and would not reach statistical significance when correcting for multiple comparisons of post hoc tests. Moreover,

---

[4] The missing feedback cue ratings of two participants, which arose through technical issues with the stimulus presentation software, were replaced by the mean values of the other participants' ratings.

14698986, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/psyp.14235 by Test, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

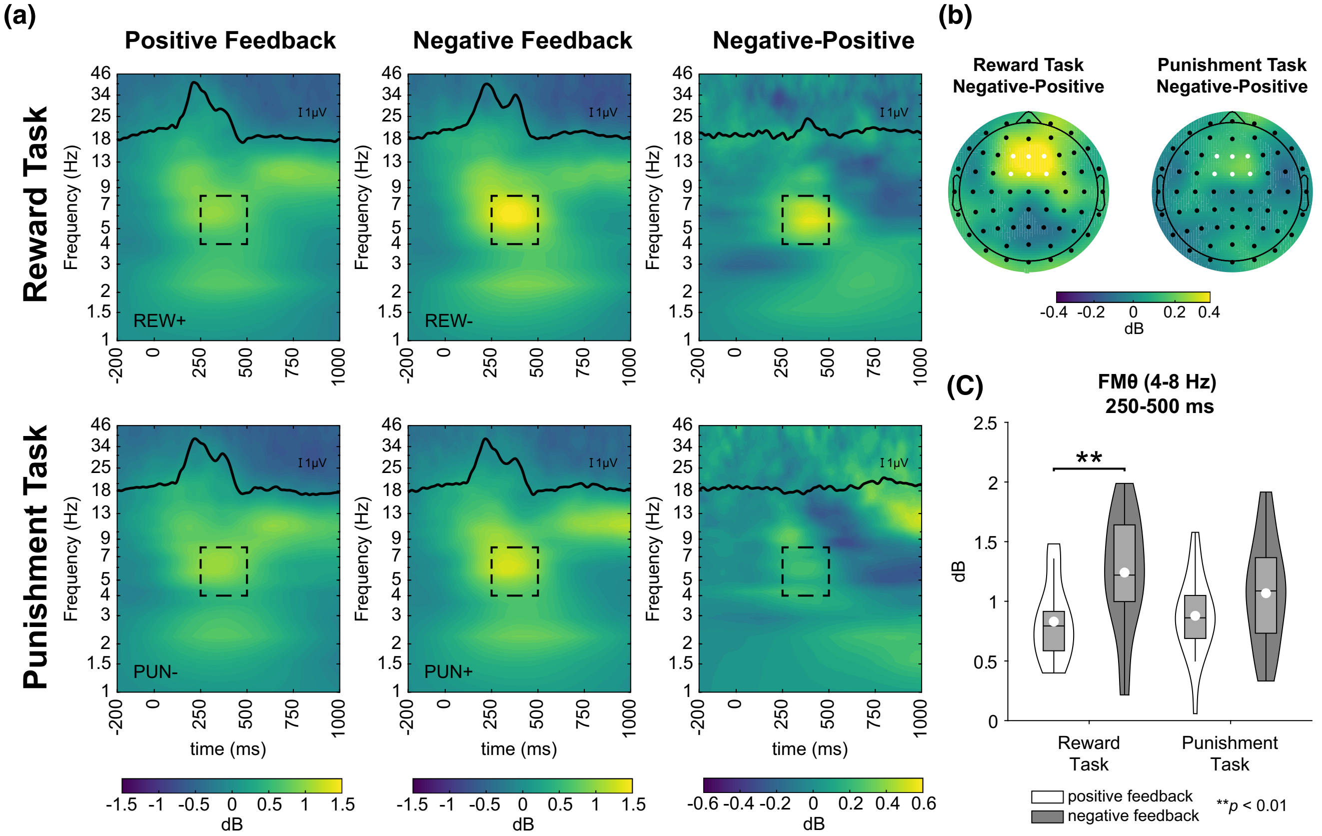

FIGURE 3 (a) Feedback-locked time-frequency plots for positive and negative feedback at frontocentral electrode sites (averaged across Fz, F1, F2, FCz, FC1, & FC2). The ERPs are superimposed over the power plots in black. The time-frequency regions of interest are shown in black boxes. (b) Scalp topographies of the mean FMθ power negative–positive differences in the time-frequency region of interest. (c) Violin plot with mean values depicted as white circles for FMθ power in response to positive feedback (REW+/PUN−) and negative feedback (REW−/PUN+).

there was no significant difference in FMθ to REW− vs. PUN+, $t(22) = 2.02$, $p = .06$, $d = 0.38$, and REW+ vs. PUN−, $t(22) = 0.69$, $p = .50$, $d = 0.15$. For a similar analysis of delta (1–4 Hz) power, see Appendix S1.

To provide a more comparative hypothesis testing for the compelling evidence of smaller feedback valence effect in the punishment vs. reward task, we calculated the Bayes factor (BF) for the comparison of REW+ vs. REW− and PUN− vs, PUN+. These analyses revealed a $BF_{10} = 47.75$ in the reward task (REW+ vs. REW−), and $BF_{10} = 1.46$ in the punishment task (PUN− vs. PUN+). Following the suggestions of Wagenmakers et al. (2011), the BF_{10} indicates very strong evidence in favor of the H_1 (REW+ and REW− are different from each other) in the reward task and anecdotal evidence in favor of the H_1 in the punishment task. This further supports the notion of a smaller feedback valence effect in the punishment vs. reward task.

An ANOVA including the factors task and feedback valence on uncorrected baseline (−300 to −200 ms) in FMθ did not reveal any significant effect (Fs≤ 0.30, ps≥.59, η_p^2s≤0.01), which overall does not provide evidence for a difference in tonic FMθ activity. Moreover, an ANOVA including the factors task and half (first half vs. second half) on uncorrected baseline (−300 to −200 ms) FMθ power did not reveal any significant effect (Fs≤ 1.68, ps≥ .21, η_p^2s≤0.07). These results overall do not provide evidence for global tonic differences in the tonic FMθ in the reward vs. punishment task.

4.2 | Single-trial regressions with signed prediction errors as regressors and FMθ as criterion

None of the eight slopes resulting from the single-trial regressions with computationally derived estimated positive (ePE+) and negative (ePE−) prediction errors (|t|s≤1.27, ps≥.22, ds≤0.27) or self-reported positive (sPE+) and negative (sPE−) prediction errors (|t|s≤1.28, ps≥.21, ds≤0.27) as regressor and FMθ as criterion was significantly different from zero; see Table 4. Consistent with the power analysis and the absence of any significant effects, BF_{01}s comparing ePE/sPE against zero ranged between 1.93

14698986, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/psyp.14235 by Test, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 4 Mean (SD) values of standardized beta slopes with single-trial estimated and self-reported prediction errors as regressors and FMθ as the criterion

	ePE+	epE−	sPE+	sPE−
Reward task	0.007 (0.04)	−0.005 (0.07)	0.013 (0.05)	−0.001 (0.05)
Punishment task	0.011 (0.04)	0.007 (0.07)	0.011 (0.05)	0.013 (0.08)

and 4.54, suggesting anecdotal to substantial evidence in favor of the H_0 (no difference from zero). This indicates that the present data do not provide a robust association between FMθ and either a signed or an unsigned prediction error. The relationships between FMθ and the ePE of every fit computational model (Table S2) as well as the results of ANOVAs including the factors of prediction error type (ePE vs. sPE), sign (PE+ & PE−), and task (reward vs. punishment) on the resulting beta values are presented in the Tables S3 and S4.

Since the chosen time window for the analysis of prediction errors and FMθ includes both early (e.g., P200 and the Reward Positivity) and late phasic activity (e.g., P3a), we calculated regression analysis with ePE/sPE as predictor and early (250–375 ms) and late (375–500 ms) FMθ power as criterion. None of the slopes with ePE as predictor and early ($0.006 \leq Ms \leq 0.019$, $ts \leq 1.01$, $ps \geq .32$, $ds \leq 0.21$) and late FMθ ($-0.022 \leq Ms \leq 0.004$, $|t|s \leq 1.02$, $ps \geq 0.32$, $ds \leq 0.21$) as criterion, and with sPE as predictor and early ($-0.026 \leq Ms \leq 0.005$, $|t|s \leq 1.09$, $ps \geq 0.29$, $ds \leq 0.23$) and late FMθ ($-0.032 \leq Ms \leq -0.001$, $|t|s \leq 1.91$, $ps \geq 0.07$, $ds \leq 0.19$) as criterion were statistically different from zero; see Table S4. This suggests an absence of evidence for a significant association between prediction errors and early as well as late FMθ.

In addition to fitting a model which quantifies the influence of the time interval between action and feedback onset (M3: Time) on action selection, we also calculated regression analyses with the single-trial time interval as regressor and FMθ as criterion for every participant. The resulting slopes did not significantly differ from zero for both the reward ($M = 0.007$, $SD = 0.03$) and punishment ($M = 0.003$, $SD = 0.04$) task ($ts \leq 1.02$, $ps \geq 0.32$, $ds \leq 0.21$). Together with the relatively poor fit of the M3 model (see Table S1) this overall suggests little evidence in support of a modulation of FMθ by the time interval between action and feedback onset.

It is conceivable that cognitive control was of particular relevance at stages of the experiment where participants had to learn and relearn feedback contingencies (i.e., at the beginning of the task and after a reversal). To evaluate whether the relationship between PEs and FMθ depends on the learning stage in the experiment, we ran single-trial multiple regression analyses with every signed prediction error (ePE+/ePE−/sPE+/sPE−), the trial numbers within learning/reversal blocks (i.e. 1, 2, 3, ...), and the interaction term PE×trial as predictor and FMθ as criterion. None of the slopes of ePE+/ePE− ($-0.006 \leq Ms \leq 0.011$, $|t|s \leq 1.17$, $ps \geq .25$, $ds \leq 0.24$), trial ($-0.003 \leq Ms \leq 0.017$, $|t|s \leq 1.05$, $ps \geq .30$, $ds \leq 0.22$), and ePE+/ePE−×trial ($-0.010 \leq Ms \leq 0.011$, $|t|s \leq 0.89$, $ps \geq 0.38$, $ds \leq 0.19$) was significant different from zero. None of the slopes of sPE+/sPE− ($Ms \leq 0.017$, $ts \leq 1.33$, $ps \geq .20$, $ds \leq 0.28$), trial ($-0.006 \leq Ms \leq 0.017$, $|t|s \leq 1.18$, $ps \geq .25$, $ds \leq 0.24$), and sPE+/sPE−×trial ($-0.004 \leq Ms \leq 0.009$, $|t|s \leq 0.84$, $ps \geq 0.41$, $ds \leq 0.17$) was significantly different from zero. Overall, these results do not provide supporting evidence for a substantial influence of the trial within the block on prediction error processing indexed by FMθ.

4.3 | Exploratory analyses: Effects of feedback valence vs. unsigned prediction errors on FMθ

Because negative feedback was less frequent than positive feedback, the observed increases of FMθ to negative vs. positive reward feedback may still be driven by surprise/prediction error processing (Rawls et al., 2020). To probe this interpretation of the observed feedback effect on FMθ, we computed within-subject multiple regressions with single-trial feedback valence (positive vs. negative) and (unsigned) single-trial prediction error as predictors and FMθ as criterion on each data point in the 250–500 ms time window on frontocentral cluster (see methods section). Finally, the resulting beta matrix was averaged for every participant. The mean standardized beta values across subjects for each analysis are shown in Figure 4.

Converging with the pattern of results for the trial-averaged data, the single-trial regression slopes with feedback valence as predictor and FMθ as criterion differed significantly from zero in the reward task ($ts \geq 3.53$, $ps \leq 0.002$, $ds \geq 0.74$) and this was not the case in the punishment task ($ts \leq 1.74$, $ps \geq 0.10$, $ds \leq 0.36$). An ANOVA including the factors task (reward & punishment task) and prediction error type (ePE & sPE) on the slopes of feedback valence and FMθ revealed a significant main effect of task, $F(1, 22) = 4.49$, $p < .05$, $\eta_p^2 = 0.17$. The other effects were not statistically significant ($Fs \leq 0.15$, $ps \geq .71$, $\eta_p^2 s \leq 0.01$). For the relationships between the unsigned prediction errors and FMθ, regression slopes did not differ significantly from

14698986, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/psyp.14235 by Test, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

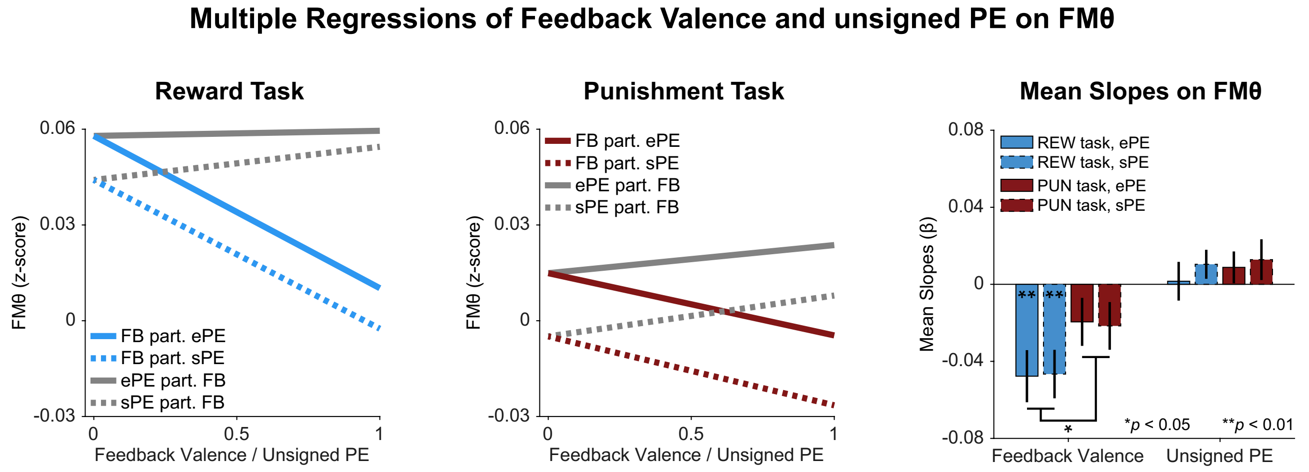

FIGURE 4 Slopes and intercepts of the multiple regression analysis including the predictors of feedback valence (here FB) and unsigned PEs (ePE & sPE) for the reward (left panel) and punishment task (middle panel); 'part.' denotes the predictor partialled out in the multiple regressions. The right panel shows the mean slopes, demonstrating that the feedback valence effect in the reward task was not affected by trial-by-trial unsigned prediction errors. $^*p<.05$ $^{**}p<.01$.

TABLE 5 Mean (SD) values of standardized beta slopes with FMθ as predictor and subsequent behavioral adjustment (RT slowing & switching) as criterion

	RT slowing positive fb.	RT slowing negative fb.	Switch positive fb.	Switch negative fb.
Reward task	0.015 (0.05)	−0.004 (0.06)	−0.006 (0.05)	0.017 (0.04)
Punishment task	−0.010 (0.05)	0.001 (0.05)	−0.001 (0.04)	−0.014 (0.08)

zero (ts ≤ 1.38, ps ≥ .18, ds ≤ 0.29) and ANOVA s on the slopes did not reveal any significant main or interaction effects (Fs ≤ 0.66, ps ≥ 0.43, η_p^2s ≤ 0.03). Overall, these results do not provide evidence for an association between unsigned prediction errors and FMθ nor do they support the idea that the observed effects involving Feedback Valence are substantially driven by unsigned prediction error processing.

4.4 | Single-trial regressions withFMθas predictor and behavioral adjustment as criterion

The single-trial regressions with FMθ as predictor and RT adjustment (i.e., slowing in the current vs. previous trial) as criterion revealed that none of the eight slopes was significantly different from zero for positive feedback ($|t|$s ≤ 1.54, ps ≥ 0.13, ds ≤ 0.32) and negative feedback trials ($|t|$s ≤ 0.28, ps ≥ 0.78, ds ≤ 0.06); see Table 5.

For the single-trial regressions with FMθ as predictor and switching (vs. staying) behavior as criterion, the slope of FMθ responses to monetary non-reward (REW−) on subsequent switching was at the borderline of significance differing from zero, $t(22) = 2.07$, $p = .05$, $d = 0.43$. None of the other slopes was different from zero ($|t|$s ≤ 0.88, ps ≥ 0.39, ds ≤ 0.18).

5 | DISCUSSION

The goal of the present study was to investigate if FMθ bursts are sensitive to negative feedback indicative of subsequent punishment. To this end, participants performed both a reward-, and a punishment-, related reinforcement learning task. While the learning performance did not depend on the type of reinforcer, the response of FMθ to negative c.f. positive feedback was significantly greater in the reward condition than in the punishment condition; specifically, FMθ was significantly elevated to negative feedback in the reward but not in the punishment task. Moreover, in contrast to feedback valence, we were not able to find any significant associations between either computationally derived or self-reported signed or unsigned prediction errors and FMθ. Together, this indicates that phasic FMθ is more sensitive to feedback valence in reward than in punishment reinforcement learning.

FMθ was amplified to negative vs. positive feedback in the reward task, i.e., to feedback signaling non-reward vs. reward. This result converges with previous findings that FMθ is modulated by monetary non-reward and loss in reward-related reinforcement learning (Andreou et al., 2017; Cavanagh et al., 2010; Cavanagh, Zambrano-Vazquez, & Allen, 2012; Marco-Pallares et al., 2008; Mas-Herrero & Marco-Pallarés, 2016; Mueller et al., 2015).

14698986, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/psyp.14235 by Test, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Together with previous studies which reported increased FMθ to negative feedback (Mueller et al., 2015; van de Vijver et al., 2011), conflict (Cavanagh et al., 2014; Neo et al., 2020; Pinner & Cavanagh, 2017), novelty (Brown & Cavanagh, 2020), and the need for cognitive control (Cavanagh & Shackman, 2015; Cooper et al., 2019), our results are in line with the assumption of FMθ as an index of ACC activity involved in cognitive control processing (Cavanagh, Zambrano-Vazquez, & Allen, 2012; Domic-Siede et al., 2021).

In contrast, there was no significant feedback valence effect on FMθ in the punishment task, i.e., to punishment vs. non-punishment. While this is in line with a recent study showing no effect of monetary loss on FMθ (Rawls et al., 2020), it is surprising from the perspective that FMθ reflects a signal sensitive to cognitive and motivational salience (Cavanagh & Shackman, 2015; Cooper et al., 2019). Self-reports and behavioral results of the present study would suggest an even higher salience of reinforcers in the punishment context: the punishment feedback cue was rated as more negatively valent than non-reward and the noise burst as a more punishing feedback consequence than monetary non-reward. Moreover, PES and PCS were greater in the punishment vs. reward context, suggesting greater cognitive control and behavioral adjustment (Cavanagh & Shackman, 2015;Danielmeier & Ullsperger, 2011; Fischer et al., 2016). As suggested by Rawls et al. (2020), FMθ may be generally elevated for threat-related reinforcement conditions leading to increased tonic FMθ levels during the punishment vs. reward task and thereby attenuating potential valence effects of phasic theta bursts. However, we did not find significant task differences in baseline FMθ, which does not provide evidence for higher tonic levels of FMθ in the punishment vs. reward task. Additionally, it is conceivable that our punishment task induced state fear, which might have moderated feedback processing indexed by FMθ. As it has been reported that induced stress alters feedback processing (for a review see Starcke & Brand, 2012), it may be that state fear, which involves an overall increased hypervigilance and alertness toward threat (Grillon, 2008; Maren et al., 2013; Panitz et al., 2018; Stolz et al., 2019), altered feedback processing in the present punishment task. This could be addressed in future studies by using state and trait measures or mood inductions to see whether feedback-locked FMθ is moderated by the levels of fear and anxiety in punishment avoidance learning.

Importantly, the present results on FMθ contrast with the adaptive control hypothesis which ascribes FMθ a more global role, namely reflecting unspecific cognitive control processing, such as in threat scenarios which elicit state fear and negative affect (Cavanagh & Shackman, 2015; Shackman et al., 2011). Our findings however suggest differential effects of FMθ in reinforcement learning depending on the learning context, which may engage separate neural systems during tasks demanding reward maximization vs. punishment avoidance, respectively. While there are studies linking activation in the striatum to both reward and avoidance learning (Häusler et al., 2016; Schlund et al., 2011; Smith et al., 2016), there is also evidence in line with the assumption of two different systems for reward and punishment domains (Kim & Anderson, 2020; Palminteri et al., 2015; Palminteri, Clair, et al., 2012; Palminteri, Justo, et al., 2012; Robinson et al., 2010). First, there are reports of intra-striatal dissociation with a more dorsal involvement in punishment cf. reward processing (Schlund et al., 2016; Seymour et al., 2007). Second, other studies stressed the anterior insula (Palminteri et al., 2015; Palminteri, Justo, et al., 2012), and the amygdala (Schlund & Cataldo, 2010; Yacubian et al., 2006) as potential substrates of a neural system specific to punishment avoidance learning. Finally, it has been argued that reward and punishment learning are differentially modulated by dopamine and serotonin (Cools et al., 2011; Daw et al., 2002; Fischer & Ullsperger, 2017; Jocham & Ullsperger, 2009; Macoveanu, 2014; Pickering & Pesola, 2014; Richter et al., 2014; Seymour et al., 2012). Thus, it may be that the present results on punishment-related feedback processing reflect the recruitment of a neural threat network whose activity is less well indexed by theta activity in the frontal EEG cluster. This overall sheds a different light on FMθ as a global index for cognitive control and provides evidence for context-dependent modulation of FMθ by feedback valence and PEs.

Even though we put considerable effort into the operationalization and analyses of PEs (Cohen & Cavanagh, 2011), we did not find any correlations between single-trial FMθ and the degree of computationally derived or self-reported PEs. In addition, a further analysis showed that unsigned PEs did not affect the observed feedback valence effect. This result is in contrast to previous findings which suggested that FMθ is sensitive to unexpected, i.e. salient events, and thus does index an unsigned PE (Cavanagh, Figueroa, et al., 2012; Mas-Herrero & Marco-Pallarés, 2014). Additionally, there are studies involving tasks, which are unlikely to engage explicit learning, that demonstrate main effects of feedback expectancies on reward- and punishment-related frontocentral event-related potentials (Hird et al., 2018; Soder & Potts, 2018; Talmi et al., 2013) as well as reward-related FMθ power (Gheza et al., 2018). The non-significant modulations of FMθ by PEs in the present study may be due to the relatively small number of reversals over the experiment, which provoked an exploitative strategy with participants commonly following win-stay and rarely following lose-shift choice patterns. Based on the

14698986, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/psyp.14235 by Test, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

behavioral results it is conceivable that participants experienced strong positive PEs at the beginning of an exploitative choice streak, which then declined after repeated positive feedback for the exploitative choices being made (Cavanagh, 2015). Moreover, successful learning in the reversal learning tasks demanded choice persistence, particularly after strong negative PEs, which presumably occurred during an exploitation streak (because positive feedback occurred on average only 70% of the time for the better response choice). Strong negative PEs in these circumstances are generally not good predictors of a reversal of feedback contingencies. This suggestion is in line with the idea of a stronger utilization of more explicit (model-based) expectations in the present study, masking expectations which are based on trial-by-trial value associations as derived by model-free Q learning (Botvinick & Weinstein, 2014; Daw et al., 2011; Wang, Kurth-Nelson, et al., 2018). Besides the task demands, it might also be that assessing self-reported expectations within each trial further contributed to participants' overall use of more explicit predictions (Akam & Walton, 2021; Doll et al., 2009; Smittenaar et al., 2013). Overall, since negative PEs were not indicative of the need for behavioral adjustment, it may be that FMθ was less affected by trial-by-trial predictions but more by implicit learning about the reversals, decoupling the often-replicated association between PEs and FMθ (Cooper et al., 2019; Kaufman et al., 2010; Pinner & Cavanagh, 2017; Reber, 2013). In the context of prior work, our results thus suggest that FMθ does not indicate unsigned or signed PEs per se but that the correlation between FMθ and prediction error magnitude may be strongly influenced by task-specific demands and learning goals.

There was a hint for a predictive value of increased FMθ power in response to feedback indicating monetary non-reward and subsequent switching behavior. This converges with previous studies, showing that increased electrocortical responses to feedback indicating non-reward/monetary loss are associated with subsequent switching behavior (Cavanagh & Shackman, 2015; van de Vijver et al., 2011). In contrast, there was no significant effect for the association between FMθ and post-punishment switching in the punishment task. This result pattern is informative with regard to the observation of a larger feedback valence effect in the reward vs. punishment task, suggesting that the increased FMθ responses to non-reward vs. reward here may reflect behavioral adaptation. In contrast to switching behavior, there was no significant relationship between FMθ and PES/PCS, presumably due to the lack of an RT limit in the current task.

There are some limitations of the present study. The noise burst titration procedure might have not been entirely successful and resulted in a slightly higher unpleasantness of the noise burst intensity compared with the pleasantness of monetary reward. This is indicated by feedback cue and consequences ratings, and the distribution of the chosen noise burst intensity. Still, even though the punishment reinforcer seemed to be more salient, we did not observe a significant FMθ valence effect in the punishment task. In addition, there were overall more bad epochs in the punishment vs. reward task, which is in line with the suggestion of higher unpleasantness and salience of the noise burst. Since the number of bad epochs were comparable across feedback valence types within the tasks, and differed significantly only between the tasks, we assume that this did not affect the feedback-valence effects on FMθ, but rather is an example of how threat experiments can induce method-related effects. Although we calculated PES and PCS measured as pairwise comparisons around error and correct trials, which is more robust than traditional approaches (see Dutilh et al., 2012), the RT-based measures should be interpreted with caution since there was no time limit for the behavioral responses in the reversal learning tasks.

While we used an established model-free algorithm which is suitable for single-trial EEG analyses (Cavanagh, 2015), it does not provide information regarding participants' knowledge about the reversal learning task structures. Comparisons between model-based and model-free reinforcement learning algorithms are better suited to estimate individual differences in the knowledge about the reversal learning task structure (Daw et al., 2011; Lucantonio et al., 2014; Wang, Lee, et al., 2018; Wunderlich et al., 2012). For instance, there is a probabilistic model recently reported by Marković et al. (2019) that uses hidden Markov models to estimate state dynamics and could be used to examine the neural basis of behavioral transition after a contingency reversal (Marković et al., 2020). Assessing expectations via self-report may have added further factors influencing learning during the current tasks. The influence of these additional factors may have weakened the relationship between computationally derived PEs and FMθ. On the one hand, assessing self-reported expectations might have altered PE processing by increasing the time interval between action and feedback onset. However, additional analyses including a separate model and single-trial analyses on FMθ accounting for the time interval between action and feedback onset interval did not support this hypothesis. On the other hand, assessing self-reported expectations might have changed the overall nature of participants' expectations, contributing to the decoupling of trial-by-trial PEs and FMθ. In contrast, there are studies which demonstrated that the frontocentral feedback-related negativity event-related potential was associated with both computationally derived and self-reported PEs (Hajcak

14698986, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/psyp.14235 by Test, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

et al., 2005; Ichikawa et al., 2010; Soder & Potts, 2018; Yeung et al., 2005). The effects were even stronger when expectations were assessed directly after the action vs. at the end of the trial (Hajcak et al., 2007). We thus would argue that the present null findings on the association between PEs and FMθ might be due to the implicit nature of the learning task rather than the presence of self-report ratings. It could be interesting for further studies, however, to investigate if the levels of implicit learning together with manipulations of the action-feedback intervals influence feedback processing.

There are two further peculiarities of the present study that deserve attention. First, aversive noise bursts are considered a primary reinforcer while monetary reward is a secondary reinforcer. Although monetary reward is highly overlearned and may even sometimes have stronger effects on behavior than primary reinforcers such as food (Beck et al., 2010; Delgado et al., 2011), the reward delivery was briefly delayed while the noise bursts were presented immediately as a consequence of performance. Even though we cannot rule out effects of slightly delayed vs. more immediate presentation of reinforcer, the results of the present study are in line with those of Rawls et al. (2020) who similarly found no effects of negative feedback on FMθ with monetary loss as negative feedback reinforcer. Here, we showed that a similar lack of FMθ modulation is observed in the context of actual threat stimuli. Second, the tasks demanded different motivational directions. While the participants approached monetary reward in the reward task, they had to avoid an aversive noise burst in punishment task. It is possible that the reported modulations of FMθ reflect different systems that are more related to negative and positive reinforcement rather than negative and positive valence (Richter et al., 2014). Based on: (A) the similarity of the two reversal learning tasks used here; (B) the participants' comparable performance across tasks; (C) and the ratings on non-punishment consequences, we would argue that successful avoidance in the punishment task was perceived as a positive outcome (Kohls et al., 2013; Palminteri et al., 2015; Sescousse et al., 2013). Second, some authors expressed concerns regarding the statistical comparison of categorically different reinforcer types (Talmi et al., 2013; but see Heydari & Holroyd, 2016). In contrast to Talmi et al. (2013) and Heydari and Holroyd (2016), the present reinforcer magnitudes underwent a titration procedure and thus may have been qualitatively closer to each other than in previous studies. However, to accomplish the best possible matching between reinforcer magnitudes, future research should consider examining FMθ in response to primary reinforcers from same categories (e.g. juice or odors) in positive vs. negative reinforcement learning.

In summary, our study provides first evidence that feedback-locked FMθ is more sensitive to feedback valence in monetary reward learning tasks compared with punishment tasks involving threat-related primary reinforcers. Moreover, we did not find any correlation between signed and unsigned PEs and FMθ. Overall, this suggests that FMθ does not reflect a global index for cognitive control processing but rather demonstrates that FMθ activity depends strongly on the learning context. We argue that punishment avoidance vs. reward learning depend upon non-identical neural systems, and that the activity of the system supporting punishment avoidance affects measures of FMθ to a lesser degree. We would finally suggest that correlations between FMθ and PEs strongly depend on whether PE processing is crucial for achieving task-specific learning demands. Our study also raised a number of further questions, which need to be investigated in future studies in order to provide a more comprehensive insight into the role of FMθ during punishment avoidance learning.

AUTHOR CONTRIBUTIONS

Christopher Stolz: Conceptualization; data curation; formal analysis; investigation; methodology; writing – original draft; writing – review and editing. **Alan D. Pickering:** Formal analysis; methodology; supervision; writing – review and editing. **Erik M. Mueller:** Conceptualization; funding acquisition; project administration; supervision; writing – review and editing.

ACKNOWLEDGMENTS

The present study was supported by the Deutsche Forschungsgemeinschaft (DFG; Project-ID 290878970–RTG 2271 and Project-ID 422744262–TRR 289). Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

The authors declare that the research was conducted in absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

DATA AVAILABILITY STATEMENT

https://openneuro.org/datasets/ds004295/

ETHIC ISSUES

The study was approved by the Ethics Committee of the University of Marburg, Department of Psychology.

ORCID

Christopher Stolz https://orcid.org/0000-0003-0127-9229
Alan D. Pickering https://orcid.org/0000-0002-7301-5321
Erik M. Mueller https://orcid.org/0000-0002-8721-8963

14698986, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/psyp.14235 by Test, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

REFERENCES

Akam, T., & Walton, M. E. (2021). What is dopamine doing in model-based reinforcement learning? *Current Opinion in Behavioral Sciences*, *38*, 74–82. https://doi.org/10.1016/j.cobeha.2020.10.010

Alexander, W. H., & Brown, J. W. (2019). The role of the anterior cingulate cortex in prediction error and signaling surprise. *Topics in Cognitive Science*, *11*(1), 119–135. https://doi.org/10.1111/tops.12307

Andreou, C., Frielinghaus, H., Rauh, J., Mußmann, M., Vauth, S., Braun, P., Leicht, G., & Mulert, C. (2017). Theta and high-beta networks for feedback processing: A simultaneous EEG-fMRI study in healthy male subjects. *Translational Psychiatry*, *7*(1), e1016. https://doi.org/10.1038/tp.2016.287

Arbel, Y., Hong, L., Baker, T. E., & Holroyd, C. B. (2017). It's all about timing: An electrophysiological examination of feedback-based learning with immediate and delayed feedback. *Neuropsychologia*, *99*, 179–186. https://doi.org/10.1016/j.neuropsychologia.2017.03.003

Badre, D., & Nee, D. E. (2018). Frontal cortex and the hierarchical control of behavior. *Trends in Cognitive Sciences*, *22*(2), 170–188. https://doi.org/10.1016/j.tics.2017.11.005

Bai, Y., Katahira, K., & Ohira, H. (2014). Dual learning processes underlying human decision-making in reversal learning tasks: Functional significance and evidence from the model fit to human behavior. *Frontiers in Psychology*, *5*, 871. https://doi.org/10.3389/fpsyg.2014.00871

Beck, S. M., Locke, H. S., Savine, A. C., Jimura, K., & Braver, T. S. (2010). Primary and secondary rewards differentially modulate neural activity dynamics during working memory. *PLoS One*, *5*(2), e9251. https://doi.org/10.1371/journal.pone.0009251

Behrens, T. E. J., Woolrich, M. W., Walton, M. E., & Rushworth, M. F. S. (2007). Learning the value of information in an uncertain world. *Nature Neuroscience*, *10*(9), 1214–1221. https://doi.org/10.1038/nn1954

Bernat, E. M., Nelson, L. D., & Baskin-Sommers, A. R. (2015). Time-frequency theta and delta measures index separable components of feedback processing in a gambling task. *Psychophysiology*, *52*(5), 626–637. https://doi.org/10.1111/psyp.12390

Betts, M. J., Richter, A., deBoer, L., Tegelbeckers, J., Perosa, V., Baumann, V., Chowdhury, R., Dolan, R. J., Seidenbecher, C., Schott, B. H., Düzel, E., Guitart-Masip, M., & Krauel, K. (2020). Learning in anticipation of reward and punishment: Perspectives across the human lifespan. *Neurobiology of Aging*, *96*, 49–57. https://doi.org/10.1016/j.neurobiolaging.2020.08.011

Botvinick, M., & Weinstein, A. (2014). Model-based hierarchical reinforcement learning and human action control. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1655), 20130480. https://doi.org/10.1098/rstb.2013.0480

Brown, D. R., & Cavanagh, J. F. (2020). Novel rewards occlude the reward positivity, and what to do about it. *Biological Psychology*, *151*, 107841. https://doi.org/10.1016/j.biopsycho.2020.107841

Cavanagh, J. F. (2015). Cortical delta activity reflects reward prediction error and related behavioral adjustments, but at different times. *NeuroImage*, *110*, 205–216. https://doi.org/10.1016/j.neuroimage.2015.02.007

Cavanagh, J. F., Eisenberg, I., Guitart-Masip, M., Huys, Q., & Frank, M. J. (2013). Frontal theta overrides Pavlovian learning biases. *Journal of Neuroscience*, *33*(19), 8541–8548. https://doi.org/10.1523/JNEUROSCI.5754-12.2013

Cavanagh, J. F., Figueroa, C. M., Cohen, M. X., & Frank, M. J. (2012). Frontal theta reflects uncertainty and unexpectedness during exploration and exploitation. *Cerebral Cortex*, *22*(11), 2575–2586. https://doi.org/10.1093/cercor/bhr332

Cavanagh, J. F., Frank, M. J., Klein, T. J., & Allen, J. J. B. (2010). Frontal theta links prediction errors to behavioral adaptation in reinforcement learning. *NeuroImage*, *49*(4), 3198–3209. https://doi.org/10.1016/j.neuroimage.2009.11.080

Cavanagh, J. F., Masters, S. E., Bath, K., & Frank, M. J. (2014). Conflict acts as an implicit cost in reinforcement learning. *Nature Communications*, *5*, 5394. https://doi.org/10.1038/ncomms6394

Cavanagh, J. F., & Shackman, A. J. (2015). Frontal midline theta reflects anxiety and cognitive control: Meta-analytic evidence. *Journal of Physiology Paris*, *109*(1–3), 3–15. https://doi.org/10.1016/j.jphysparis.2014.04.003

Cavanagh, J. F., Zambrano-Vazquez, L., & Allen, J. J. B. (2012). Theta lingua franca: A common mid-frontal substrate for action monitoring processes. *Psychophysiology*, *49*(2), 220–238. https://doi.org/10.1111/j.1469-8986.2011.01293.x

Cohen, M. X., & Cavanagh, J. F. (2011). Single-trial regression elucidates the role of prefrontal theta oscillations in response conflict. *Frontiers in Psychology*, *2*, 30. https://doi.org/10.3389/fpsyg.2011.00030

Cohen, M. X., Elger, C. E., & Ranganath, C. (2007). Reward expectation modulates feedback-related negativity and EEG spectra. *NeuroImage*, *35*(2), 968–978. https://doi.org/10.1016/j.neuroimage.2006.11.056

Collins, A. G. E., & Frank, M. J. (2014). Cognitive control over learning: Creating, clustering and generalizing task-set structure. *Psychological Review*, *120*(1), 190–229. https://doi.org/10.1037/a0030852.Cognitive

Collins, A. G. E., & Frank, M. J. (2016). Neural signature of hierarchically structured expectations predicts clustering and transfer of rule sets in reinforcement learning. *Cognition*, *152*, 160–169. https://doi.org/10.1016/j.cognition.2016.04.002

Cools, R., Nakamura, K., & Daw, N. D. (2011). Serotonin and dopamine: Unifying affective, activational, and decision functions. *Neuropsychopharmacology*, *36*(1), 98–113. https://doi.org/10.1038/npp.2010.121

Cooper, P. S., Karayanidis, F., McKewen, M., McLellan-Hall, S., Wong, A. S. W., Skippen, P., & Cavanagh, J. F. (2019). Frontal theta predicts specific cognitive control-induced behavioural changes beyond general reaction time slowing. *NeuroImage*, *189*, 130–140. https://doi.org/10.1016/j.neuroimage.2019.01.022

Corr, P. J.(Ed.). (2008). *The reinforcement sensitivity theory of personality*. Cambridge University Press. https://doi.org/10.1017/CBO9780511819384

Corr, P. J., & McNaughton, N. (2012). Neuroscience and approach/avoidance personality traits: A two stage (valuation-motivation) approach. *Neuroscience and Biobehavioral Reviews*, *36*(10), 2339–2354. https://doi.org/10.1016/j.neubiorev.2012.09.013

Danielmeier, C., & Ullsperger, M. (2011). Post-error adjustments. *Frontiers in Psychology*, *2*, 233. https://doi.org/10.3389/fpsyg.2011.00233

Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and

14698986, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/psyp.14235 by Test, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

striatal prediction errors. *Neuron*, *69*(6), 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027

Daw, N. D., Kakade, S., & Dayan, P. (2002). Opponent interactions between serotonin and dopamine. *Neural Networks*, *15*(4–6), 603–616. https://doi.org/10.1016/S0893-6080(02)00052-7

Delgado, M. R., Jou, R. L., Phelps, E. A., Seymour, B., & Centre, W. T. (2011). Neural systems underlying aversive conditioning in humans with primary and secondary reinforcers. *Frontiers in Neuroscience*, *5*, 71. https://doi.org/10.3389/fnins.2011.00071

denOuden, H. E. M., Daw, N. D., Fernandez, G., Elshout, J. A., Rijpkema, M., Hoogman, M., Franke, B., & Cools, R. (2013). Dissociable effects of dopamine and serotonin on reversal learning. *Neuron*, *80*(4), 1090–1100. https://doi.org/10.1016/j.neuron.2013.08.030

Doll, B. B., Jacobs, W. J., Sanfey, A. G., & Frank, M. J. (2009). Instructional control of reinforcement learning: A behavioral and neurocomputational investigation. *Brain Research*, *1299*, 74–94. https://doi.org/10.1016/j.brainres.2009.07.007

Domic-Siede, M., Irani, M., Valdés, J., Perrone-Bertolotti, M., & Ossandón, T. (2021). Theta activity from frontopolar cortex, mid-cingulate cortex and anterior cingulate cortex shows different roles in cognitive planning performance. *NeuroImage*, *226*, 117557. https://doi.org/10.1016/j.neuroimage.2020.117557

Dutilh, G., Van Ravenzwaaij, D., Nieuwenhuis, S., Van der Maas, H. L. J., Forstmann, B. U., & Wagenmakers, E. J. (2012). How to measure post-error slowing: A confound and a simple solution. *Journal of Mathematical Psychology*, *56*(3), 208–216. https://doi.org/10.1016/j.jmp.2012.04.001

Fischer, A. G., Danielmeier, C., Villringer, A., Klein, T. A., & Ullsperger, M. (2016). Gender influences on brain responses to errors and post-error adjustments. *Scientific Reports*, *6*, 24435. https://doi.org/10.1038/srep24435

Fischer, A. G., & Ullsperger, M. (2013). Real and fictive outcomes are processed differently but converge on a common adaptive mechanism. *Neuron*, *79*(6), 1243–1255. https://doi.org/10.1016/j.neuron.2013.07.006

Fischer, A. G., & Ullsperger, M. (2017). An update on the role of serotonin and its interplay with dopamine for reward. *Frontiers in Human Neuroscience*, *11*, 484. https://doi.org/10.3389/FNHUM.2017.00484

Frank, M. J., Moustafa, A. A., Haughey, H. M., Curran, T., & Hutchison, K. E. (2007). Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(41), 16311–16316. https://doi.org/10.1073/pnas.0706111104

Gershman, S., & Tzovaras, B. G. (2018). Dopaminergic genes are associated with both directed and random exploration. *BioRxiv*, *120*, 357251. https://doi.org/10.1101/357251

Gheza, D., Bakic, J., Baeken, C., De Raedt, R., & Pourtois, G. (2019). Abnormal approach-related motivation but spared reinforcement learning in MDD: Evidence from fronto-midline theta oscillations and frontal alpha asymmetry. *Cognitive, Affective, & Behavioral Neuroscience*, *19*(3), 759–777. https://doi.org/10.3758/s13415-019-00693-4

Gheza, D., De Raedt, R., Baeken, C., & Pourtois, G. (2018). Integration of reward with cost anticipation during performance monitoring revealed by ERPs and EEG spectral perturbations. *NeuroImage*, *173*, 153–164. https://doi.org/10.1016/j.neuroimage.2018.02.049

Gray, J. A., & McNaughton, N. (2000). *The neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system*(2nd ed.). Oxford University Press.

Grillon, C. (2008). Models and mechanisms of anxiety: Evidence from startle studies. *Psychopharmacology*, *199*(3), 421–437. https://doi.org/10.1007/s00213-007-1019-1

Gruber, M. J., Watrous, A. J., Ekstrom, A. D., Ranganath, C., & Otten, L. J. (2013). Expected reward modulates encoding-related theta activity before an event. *NeuroImage*, *64*(1), 68–74. https://doi.org/10.1016/j.neuroimage.2012.07.064

Hajcak, G., Holroyd, C. B., Moser, J. S., & Simons, R. F. (2005). Brain potentials associated with expected and unexpected good and bad outcomes. *Psychophysiology*, *42*(2), 161–170. https://doi.org/10.1111/j.1469-8986.2005.00278.x

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2007). It's worse than you thought: The feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology*, *44*(6), 905–912. https://doi.org/10.1111/j.1469-8986.2007.00567.x

Häusler, A. N., Artigas, S. O., Trautner, P., & Weber, B. (2016). Gain- and loss-related brain activation are associated with information search differences in risky gambles: An fMRI and eye-tracking study. *ENeuro*, *3*(5), 1–13. https://doi.org/10.1523/ENEURO.0189-16.2016

Heydari, S., & Holroyd, C. B. (2016). Reward positivity: Reward prediction error or salience prediction error?*Psychophysiology*, *53*(8), 1185–1192. https://doi.org/10.1111/psyp.12673

Hird, E. J., El-Deredy, W., Jones, A., & Talmi, D. (2018). Temporal dissociation of salience and prediction error responses to appetitive and aversive taste. *Psychophysiology*, *55*(2), 1–13. https://doi.org/10.1111/psyp.12976

Ichikawa, N., Siegle, G. J., Dombrovski, A., & Ohira, H. (2010). Subjective and model-estimated reward prediction: Association with the feedback-related negativity (FRN) and reward prediction error in a reinforcement learning task. *International Journal of Psychophysiology*, *78*(3), 273–283. https://doi.org/10.1016/j.ijpsycho.2010.09.001

Jocham, G., & Ullsperger, M. (2009). Neuropharmacology of performance monitoring. *Neuroscience and Biobehavioral Reviews*, *33*(1), 48–60. https://doi.org/10.1016/j.neubiorev.2008.08.011

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572

Kaufman, S. B., DeYoung, C. G., Gray, J. R., Jiménez, L., Brown, J., & Mackintosh, N. (2010). Implicit learning as an ability. *Cognition*, *116*(3), 321–340. https://doi.org/10.1016/j.cognition.2010.05.011

Keil, A., Stolarova, M., Moratti, S., & Ray, W. J. (2007). Adaptation in human visual cortex as a mechanism for rapid discrimination of aversive stimuli. *NeuroImage*, *36*(2), 472–479. https://doi.org/10.1016/j.neuroimage.2007.02.048

Kim, A. J., & Anderson, B. A. (2020). The effect of concurrent reward on aversive information processing in the brain. *NeuroImage*, *217*, 116890. https://doi.org/10.1016/j.neuroimage.2020.116890

Kohls, G., Perino, M. T., Taylor, J. M., Madva, E. N., Cayless, S. J., Troiani, V., Price, E., Faja, S., Herrington, J. D., & Schultz, R. T. (2013). The nucleus accumbens is involved in both the pursuit of social reward and the avoidance of social punishment. *Neuropsychologia*, *51*(11), 2062–2069. https://doi.org/10.1016/j.neuropsychologia.2013.07.020

14698986, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/psyp.14235 by Test, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Krekelberg, B. (2022). BayesFactor: Release 2022 (v2.3.0). https://doi.org/10.5281/ZENODO.7006300

Krigolson, O. E. (2018). Event-related brain potentials and the study of reward processing: Methodological considerations. *International Journal of Psychophysiology*, *132*, 175–183. https://doi.org/10.1016/j.ijpsycho.2017.11.007

Lucantonio, F., Caprioli, D., & Schoenbaum, G. (2014). Transition from "model-based" to "model-free" behavioral control in addiction: Involvement of the orbitofrontal cortex and dorsolateral striatum. *Neuropharmacology*, *76*(Part B), 407–415. https://doi.org/10.1016/j.neuropharm.2013.05.033

Macoveanu, J. (2014). Serotonergic modulation of reward and punishment: Evidence from pharmacological fMRI studies. *Brain Research*, *1556*, 19–27. https://doi.org/10.1016/j.brainres.2014.02.003

Marco-Pallares, J., Cucurell, D., Cunillera, T., García, R., Andrés-Pueyo, A., Münte, T. F., & Rodríguez-Fornells, A. (2008). Human oscillatory activity associated to reward processing in a gambling task. *Neuropsychologia*, *46*(1), 241–248. https://doi.org/10.1016/j.neuropsychologia.2007.07.016

Maren, S., Phan, K. L., & Liberzon, I. (2013). The contextual brain: Implications for fear conditioning, extinction and psychopathology. *Nature Reviews Neuroscience*, *14*, 417–428. https://doi.org/10.1038/nrn3492

Marković, D., Goschke, T., & Kiebel, S. J. (2020). Meta-control of the exploration-exploitation dilemma emerges from probabilistic inference over a hierarchy of time scales. *Cognitive, Affective, & Behavioral Neuroscience*, *21*, 509–533. https://doi.org/10.3758/s13415-020-00837-x

Marković, D., Reiter, A. M. F., & Kiebel, S. J. (2019). Predicting change: Approximate inference under explicit representation of temporal structure in changing environments. *PLoS Computational Biology*, *15*, e1006707. https://doi.org/10.1371/journal.pcbi.1006707

Mas-Herrero, E., & Marco-Pallarés, J. (2014). Frontal theta oscillatory activity is a common mechanism for the computation of unexpected outcomes and learning rate. *Journal of Cognitive Neuroscience*, *26*(3), 447–458. https://doi.org/10.1162/jocn_a_00516

Mas-Herrero, E., & Marco-Pallarés, J. (2016). Theta oscillations integrate functionally segregated sub-regions of the medial prefrontal cortex. *NeuroImage*, *143*, 166–174. https://doi.org/10.1016/j.neuroimage.2016.08.024

Menegas, W., Akiti, K., Amo, R., Uchida, N., & Watabe-Uchida, M. (2018). Dopamine neurons projecting to the posterior striatum reinforce avoidance of threatening stimuli. *Nature Neuroscience*, *21*(10), 1421–1430. https://doi.org/10.1038/s41593-018-0222-1

Morey, R. D., & Wagenmakers, E. J. (2014). Simple relation between Bayesian order-restricted and point-null hypothesis tests. *Statistics and Probability Letters*, *92*, 121–124. https://doi.org/10.1016/j.spl.2014.05.010

Mueller, E. M., Panitz, C., Hermann, C., & Pizzagalli, D. A. (2014). Prefrontal oscillations during recall of conditioned and extinguished fear in humans. *The Journal of Neuroscience*, *34*(21), 7059–7066. https://doi.org/10.1523/jneurosci.3427-13.2014

Mueller, E. M., Panitz, C., Pizzagalli, D. A., Hermann, C., & Wacker, J. (2015). Midline theta dissociates agentic extraversion and anhedonic depression. *Personality and Individual Differences*, *79*, 172–177. https://doi.org/10.1016/j.paid.2014.10.043

Neo, P. S. H., Tinker, J., & McNaughton, N. (2020). Goal-conflict EEG theta and biased economic decisions: A role for a second negative motivation system. *Frontiers in Neuroscience*, *14*, 342. https://doi.org/10.3389/fnins.2020.00342

Padrão, G., Mallorquí, A., Cucurell, D., Marco-Pallarés, J., & Rodriguez-Fornells, A. (2012). Neurophysiological differences in reward processing in anhedonics. *Cognitive, Affective, & Behavioral Neuroscience*, *13*(1), 102–115. https://doi.org/10.3758/s13415-012-0119-5

Palminteri, S., Clair, A. H., Mallet, L., & Pessiglione, M. (2012). Similar improvement of reward and punishment learning by serotonin reuptake inhibitors in obsessive-compulsive disorder. *Biological Psychiatry*, *72*(3), 244–250. https://doi.org/10.1016/j.biopsych.2011.12.028

Palminteri, S., Justo, D., Jauffret, C., Pavlicek, B., Dauta, A., Delmaire, C., Czernecki, V., Karachi, C., Capelle, L., Durr, A., & Pessiglione, M. (2012). Critical roles for anterior insula and dorsal striatum in punishment-based avoidance learning. *Neuron*, *76*(5), 998–1009. https://doi.org/10.1016/j.neuron.2012.10.017

Palminteri, S., Khamassi, M., Joffily, M., & Coricelli, G. (2015). Contextual modulation of value signals in reward and punishment learning. *Nature Communications*, *6*, 8096. https://doi.org/10.1038/ncomms9096

Panitz, C., Keil, A., & Mueller, E. M. (2019). Extinction-resistant attention to long-term conditioned threat is indexed by selective visuocortical alpha suppression in humans. *Scientific Reports*, *9*(1), 15809. https://doi.org/10.1038/s41598-019-52315-1

Panitz, C., Sperl, M. F. J., Hennig, J., Klucken, T., Hermann, C., & Mueller, E. M. (2018). Fearfulness, neuroticism/anxiety, and COMT Val158Met in long-term fear conditioning and extinction. *Neurobiology of Learning and Memory*, *155*, 7–20. https://doi.org/10.1016/j.nlm.2018.06.001

Pickering, A. D., & Pesola, F. (2014). Modeling dopaminergic and other processes involved in learning from reward prediction error: Contributions from an individual differences perspective. *Frontiers in Human Neuroscience*, *8*, 740. https://doi.org/10.3389/fnhum.2014.00740

Pinner, J. F. L., & Cavanagh, J. F. (2017). Frontal theta accounts for individual differences in the cost of conflict on decision making. *Brain Research*, *1672*, 73–80. https://doi.org/10.1016/j.brainres.2017.07.026

Pizzagalli, D. A., Evins, A. E., Schetter, E. C., Frank, M. J., Pajtas, P. E., Santesso, D. L., & Culhane, M. (2008). Single dose of a dopamine agonist impairs reinforcement learning in humans: Behavioral evidence from a laboratory-based measure of reward responsiveness. *Psychopharmacology*, *196*(2), 221–232. https://doi.org/10.1007/s00213-007-0957-y

Rawls, E., Miskovic, V., Moody, S. N., Lee, Y., Shirtcliff, E. A., & Lamm, C. (2020). Feedback-related negativity and frontal midline theta reflect dissociable processing of reinforcement. *Frontiers in Human Neuroscience*, *13*, 452. https://doi.org/10.3389/fnhum.2019.00452

Reber, P. J. (2013). The neural basis of implicit learning and memory: A review of neuropsychological and neuroimaging research. *Neuropsychologia*, *51*(10), 2026–2042. https://doi.org/10.1016/j.neuropsychologia.2013.06.019

Reiter, A. M. F., Deserno, L., Kallert, T., Heinze, H. J., Heinz, A., & Schlagenhauf, F. (2016). Behavioral and neural signatures of reduced updating of alternative options in alcohol-dependent patients during flexible decision-making. *Journal*

14698986, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/psyp.14235 by Test, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

of Neuroscience, *36*(43), 10935–10948. https://doi.org/10.1523/JNEUROSCI.4322-15.2016

Reiter, A. M. F., Heinze, H. J., Schlagenhauf, F., & Deserno, L. (2017). Impaired flexible reward-based decision-making in binge eating disorder: Evidence from computational modeling and functional neuroimaging. *Neuropsychopharmacology*, *42*(3), 628–637. https://doi.org/10.1038/npp.2016.95

Richter, A., Guitart-Masip, M., Barman, A., Libeau, C., Behnisch, G., Czerney, S., Schanze, D., Assmann, A., Klein, M., Düzel, E., Zenker, M., Seidenbecher, C. I., & Schott, B. H. (2014). Valenced action/inhibition learning in humans is modulated by a genetic variant linked to dopamine D2 receptor expression. *Frontiers in Systems Neuroscience*, *8*, 140. https://doi.org/10.3389/fnsys.2014.00140

Ridderinkhof, K. R., Van Den Wildenberg, W. P. M., Segalowitz, S. J., & Carter, C. S. (2004). Neurocognitive mechanisms of cognitive control: The role of prefrontal cortex in action selection, response inhibition, performance monitoring, and reward-based learning. *Brain and Cognition*, *56*(2), 129–140. https://doi.org/10.1016/j.bandc.2004.09.016

Rigoli, F., Chew, B., Dayan, P., & Dolan, R. J. (2016). Multiple value signals in dopaminergic midbrain and their role in avoidance contexts. *NeuroImage*, *135*, 197–203. https://doi.org/10.1016/j.neuroimage.2016.04.062

Robinson, O. J., Frank, M. J., Sahakian, B. J., & Cools, R. (2010). Dissociable responses to punishment in distinct striatal regions during reversal learning. *NeuroImage*, *51*(4), 1459–1467. https://doi.org/10.1016/j.neuroimage.2010.03.036

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374. https://doi.org/10.1016/j.jmp.2012.08.001

Rushworth, M. F. S., & Behrens, T. E. J. (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nature Neuroscience*, *11*(4), 389–397. https://doi.org/10.1038/nn2066

Rushworth, M. F. S., Noonan, M. A. P., Boorman, E. D., Walton, M. E., & Behrens, T. E. (2011). Frontal cortex and reward-guided learning and decision-making. *Neuron*, *70*(6), 1054–1069. https://doi.org/10.1016/j.neuron.2011.05.014

Sambrook, T. D., & Goslin, J. (2014). Mediofrontal event-related potentials in response to positive, negative and unsigned prediction errors. *Neuropsychologia*, *61*(1), 1–10. https://doi.org/10.1016/j.neuropsychologia.2014.06.004

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

Sambrook, T. D., & Goslin, J. (2016). Principal components analysis of reward prediction errors in a reinforcement learning task. *NeuroImage*, *124*, 276–286. https://doi.org/10.1016/j.neuroimage.2015.07.032

Schlund, M. W., Brewer, A. T., Magee, S. K., Richman, D. M., Solomon, S., Ludlum, M. D., & Dymond, S. (2016). The tipping point: Value differences and parallel dorsal-ventral frontal circuits gating human approach-avoidance behavior. *NeuroImage*, *136*, 94–105. https://doi.org/10.1016/j.neuroimage.2016.04.070

Schlund, M. W., & Cataldo, M. F. (2010). Amygdala involvement in human avoidance, escape and approach behavior. *NeuroImage*, *53*(2), 769–776. https://doi.org/10.1016/j.neuroimage.2010.06.058

Schlund, M. W., Magee, S., & Hudgins, C. D. (2011). Human avoidance and approach learning: Evidence for overlapping neural systems and experiential avoidance modulation of avoidance neurocircuitry. *Behavioural Brain Research*, *225*(2), 437–448. https://doi.org/10.1016/j.bbr.2011.07.054

Schönbrodt, F. D., & Wagenmakers, E. J. (2018). Bayes factor design analysis: Planning for compelling evidence. *Psychonomic Bulletin and Review*, *25*(1), 128–142. https://doi.org/10.3758/s13423-017-1230-y

Sescousse, G., Caldú, X., Segura, B., & Dreher, J. C. (2013). Processing of primary and secondary rewards: A quantitative meta-analysis and review of human functional neuroimaging studies. *Neuroscience and Biobehavioral Reviews*, *37*(4), 681–696. https://doi.org/10.1016/j.neubiorev.2013.02.002

Seymour, B., Daw, N., Dayan, P., Singer, T., & Dolan, R. (2007). Differential encoding of losses and gains in the human striatum. *Journal of Neuroscience*, *27*(18), 4826–4831. https://doi.org/10.1523/JNEUROSCI.0400-07.2007

Seymour, B., Daw, N. D., Roiser, J. P., Dayan, P., & Dolan, R. (2012). Serotonin selectively modulates reward value in human decision-making. *Journal of Neuroscience*, *32*(17), 5833–5842. https://doi.org/10.1523/jneurosci.0053-12.2012

Shackman, A. J., Salomons, T. V., Slagter, H. A., Fox, A. S., Winter, J. J., & Davidson, R. J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nature Reviews Neuroscience*, *12*(3), 154–167. https://doi.org/10.1038/nrn2994

Silvetti, M., Nuñez Castellar, E., Roger, C., & Verguts, T. (2014). Reward expectation and prediction error in human medial frontal cortex: An EEG study. *NeuroImage*, *84*, 376–382. https://doi.org/10.1016/j.neuroimage.2013.08.058

Silvetti, M., Seurinck, R., & Verguts, T. (2011). Value and prediction error in medial frontal cortex: Integrating the single-unit and systems levels of analysis. *Frontiers in Human Neuroscience*, *5*, 75. https://doi.org/10.3389/fnhum.2011.00075

Smillie, L. D., Pickering, A. D., & Jackson, C. J. (2006). The new reinforcement sensitivity theory: Implications for personality measurement. *Personality and Social Psychology Review*, *10*(4), 320–335. https://doi.org/10.1207/s15327957pspr1004_3

Smith, D. V., Rigney, A. E., & Delgado, M. R. (2016). Distinct reward properties are encoded via corticostriatal interactions. *Scientific Reports*, *6*, 20093. https://doi.org/10.1038/srep20093

Smittenaar, P., FitzGerald, T. H. B., Romei, V., Wright, N. D., & Dolan, R. J. (2013). Disruption of dorsolateral prefrontal cortex decreases model-based in favor of model-free control in humans. *Neuron*, *80*(4), 914–919. https://doi.org/10.1016/j.neuron.2013.08.009

Soder, H. E., & Potts, G. F. (2018). Medial frontal cortex response to unexpected motivationally salient outcomes. *International Journal of Psychophysiology*, *132*, 268–276. https://doi.org/10.1016/j.ijpsycho.2017.11.003

Sperl, M. F. J., Panitz, C., Rosso, I. M., Dillon, D. G., Kumar, P., Hermann, A., Whitton, A. E., Hermann, C., Pizzagalli, D. A., & Mueller, E. M. (2018). Fear extinction recall modulates human Frontomedial theta and amygdala activity. *Cerebral Cortex*, *29*(2), 701–715. https://doi.org/10.1093/cercor/bhx353

Starcke, K., & Brand, M. (2012). Decision making under stress: A selective review. *Neuroscience and Biobehavioral Reviews*, *36*(4), 1228–1248. https://doi.org/10.1016/j.neubiorev.2012.02.003

14698986, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/psyp.14235 by Test, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Stolz, C., Endres, D., & Mueller, E. M. (2019). Threat-conditioned contexts modulate the late positive potential to faces—A mobile EEG/virtual reality study. *Psychophysiology*, *56*(4), e13308. https://doi.org/10.1111/psyp.13308

Sutton, R. S., & Barto, A. G. (1998). *Introduction to reinforcement learning*. MIT Press.

Talmi, D., Atkinson, R., & El-Deredy, W. (2013). The feedback-related negativity signals salience prediction errors, not reward prediction errors. *Journal of Neuroscience*, *33*(19), 8264–8269. https://doi.org/10.1523/JNEUROSCI.5695-12.2013

Ullsperger, M., Danielmeier, C., & Jocham, G. (2014). Neurophysiology of performance monitoring and adaptive behavior. *Physiological Reviews*, *94*(1), 35–79. https://doi.org/10.1152/physrev.00041.2012

van deVijver, I., Cohen, M. X., & Ridderinkhof, K. R. (2014). Aging affects medial but not anterior frontal learning-related theta oscillations. *Neurobiology of Aging*, *35*(3), 692–704. https://doi.org/10.1016/j.neurobiolaging.2013.09.006

van deVijver, I., Richard Ridderinkhof, K., & Cohen, M. X. (2011). Frontal oscillatory dynamics predict feedback learning and action adjustment. *Journal of Cognitive Neuroscience*, *23*(12), 4106–4121. https://doi.org/10.1162/jocn_a_00110

vanNoordt, S. J. R., Wu, J., Thomas, C., Schlund, M. W., Mayes, L. C., & Crowley, M. J. (2018). Medial frontal theta dissociates unsuccessful from successful avoidance and is modulated by lack of perseverance. *Brain Research*, *1694*, 29–37. https://doi.org/10.1016/j.brainres.2018.04.021

Vassena, E., Holroyd, C. B., & Alexander, W. H. (2017). Computational models of anterior cingulate cortex: At the crossroads between prediction and effort. *Frontiers in Neuroscience*, *11*, 316. https://doi.org/10.3389/fnins.2017.00316

Wagenmakers, E. J., Wetzels, R., Borsboom, D., & van derMaas, H. L. J. (2011). Why psychologists must change the way they analyze their data: The case of psi: Comment on Bem (2011). *Journal of Personality and Social Psychology*, *100*(3), 426–432. https://doi.org/10.1037/a0022790

Wang, J. X., Kurth-Nelson, Z., Kumaran, D., Tirumala, D., Soyer, H., Leibo, J. Z., Hassabis, D., & Botvinick, M. (2018). Prefrontal cortex as a meta-reinforcement learning system. *Nature Neuroscience*, *21*(6), 860–868. https://doi.org/10.1038/s41593-018-0147-8

Wang, O., Lee, S. W., O'Doherty, J., Seymour, B., & Yoshida, W. (2018). Model-based and model-free pain avoidance learning. *Brain and Neuroscience Advances*, *2*, 239821281877296. https://doi.org/10.1177/2398212818772964

Wang, Y. Y., Sun, L., Liu, Y. W., Pan, J. H., Zheng, Y. M., Wang, Y. F., Zang, Y. F., & Zhang, H. (2020). The low-frequency fluctuation of trial-by-trial frontal theta activity and its correlation with reaction-time variability in sustained attention. *Frontiers in Psychology*, *11*, 1555. https://doi.org/10.3389/fpsyg.2020.01555

Weinberg, A., Luhmann, C. C., Bress, J. N., & Hajcak, G. (2012). Better late than never? The effect of feedback delay on ERP indices of reward processing. *Cognitive, Affective, & Behavioral Neuroscience*, *12*(4), 671–677. https://doi.org/10.3758/s13415-012-0104-z

Weismüller, B., & Bellebaum, C. (2016). Expectancy affects the feedback-related negativity (FRN) for delayed feedback in probabilistic learning. *Psychophysiology*, *53*(11), 1739–1750. https://doi.org/10.1111/psyp.12738

Winkler, I., Debener, S., Müller, K. R., & Tangermann, M. (2015). On the influence of high-pass filtering on ICA-based artifact reduction in EEG-ERP. Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 2015-November, 4101–4105. https://doi.org/10.1109/EMBC.2015.7319296

Wunderlich, K., Smittenaar, P., & Dolan, R. J. (2012). Dopamine enhances model-based over model-free choice behavior. *Neuron*, *75*(3), 418–424. https://doi.org/10.1016/j.neuron.2012.03.042

Yacubian, J., Gläscher, J., Schroeder, K., Sommer, T., Braus, D. F., & Büchel, C. (2006). Dissociable systems for gain- and loss-related value predictions and errors of prediction in the human brain. *Journal of Neuroscience*, *26*(37), 9530–9537. https://doi.org/10.1523/JNEUROSCI.2915-06.2006

Yeung, N., Holroyd, C. B., & Cohen, J. D. (2005). ERP correlates of feedback and reward processing in the presence and absence of response choice. *Cerebral Cortex*, *15*(5), 535–544. https://doi.org/10.1093/cercor/bhh153

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Stolz, C., Pickering, A. D., & Mueller, E. M. (2022). Dissociable feedback valence effects on frontal midline theta during reward gain versus threat avoidance learning. *Psychophysiology*, *00*, e14235. https://doi.org/10.1111/psyp.14235

14698986, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/psyp.14235 by Test, Wiley Online Library on [10/01/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License