CORTEX 140 (2021) 26–39

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research Report

The aversion positivity: Mediofrontal cortical potentials reflect parametric aversive prediction errors and drive behavioral modification following negative reinforcement

Eric Rawls [a,*] *and Connie Lamm* [b]

[a] *Department of Psychiatry and Behavioral Sciences, University of Minnesota Health, USA*
[b] *Department of Psychological Sciences, University of Arkansas, USA*

ARTICLE INFO

Article history:
Received 4 November 2020
Reviewed 8 January 2021
Revised 7 February 2021
Accepted 17 March 2021
Action editor Veena Kumari
Published online 27 March 2021

Keywords:
Feedback-related negativity (FRN)
Reward positivity (RewP)
Prediction error
Reinforcement learning
Negative reinforcement
Value
Salience
Signed
Unsigned

ABSTRACT

Reinforcement learning capitalizes on prediction errors (PEs), representing the deviation of received outcomes from expected outcomes. Mediofrontal event-related potentials (ERPs), in particular the feedback-related negativity (FRN)/reward positivity (RewP), are related to PE signaling, but there is disagreement as to whether the FRN/RewP encode signed or unsigned PEs. PE encoding can potentially be dissected by time-frequency analysis, as frontal theta [4–8 Hz] might represent poor outcomes, while central delta [1–3 Hz] might instead represent rewarding outcomes. However, cortical PE signaling in negative reinforcement is still poorly understood, and the role of cortical PE representations in behavioral reinforcement learning following negative reinforcement is relatively unexplored. We recorded EEG while participants completed a task with matched positive and negative reinforcement outcome modalities, with parametrically manipulated single-trial outcomes producing positive and negative PEs. We first demonstrated that PEs systematically influence future behavior in both positive and negative reinforcement conditions. In negative reinforcement conditions, mediofrontal ERPs positively signaled unsigned PEs in a time window encompassing the P2 potential, and negatively signaled signed PEs for a time window encompassing the FRN/RewP and frontal P3 (an "aversion positivity"). Central delta power increased parametrically with increasingly aversive outcomes, contributing to the "aversion positivity". Finally, negative reinforcement ERPs correlated with RTs on the following trial, suggesting cortical PEs guide behavioral adaptations. Positive reinforcement PEs did not influence ERP or time-frequency activity, despite significant behavioral effects. These results demonstrate that mediofrontal PE signals are a mechanism underlying negative reinforcement learning, and that delta power increases for aversive outcomes might contribute to the "aversion positivity."

* *Corresponding author.*
E-mail address: erawls89@gmail.com (E. Rawls).
https://doi.org/10.1016/j.cortex.2021.03.012

1. Introduction

Reinforcement learning (RL) models place particular weight on prediction errors (PEs), a calculation of how closely the outcomes of our actions match our expectations (Sutton & Barto, 2018). Different RL models use distinct types of PE for optimal learning. Rescorla and Wagner (1972) suggested deviations from expectancy follow a monotonic function (a signed PE), such that low-value PEs makes behaviors less likely, and high value PEs make behaviors more likely. Pearce and Hall (1980) suggested an important role of the absolute value of the PE (an unsigned PE) in modulating the amount of attention devoted to outcomes. A later class of models, temporal difference models (Sutton & Barto, 2018), utilize both signed and unsigned PEs at once to guide behavior more optimally than earlier models. Theoretical RL accounts have sparked interest in how neural activity computes PEs.

Schultz, Dayan, Montague, (1997) and Hollerman and Schultz (1998) demonstrated that dopamine (DA) neurons encode signed PEs, initially leading to the belief that all DA cells signal signed reward PEs. These early reports of DA reward PE encoding were quickly applied to the study of human reinforcement processing. Upon observing a negativity in the scalp-recorded event-related potential (ERP) following omission of expected positive feedback (the *feedback-related negativity* [FRN]), Holroyd and Coles (2002) hypothesized that this potential encoded a dopaminergic signed negative PE (worse-than-expected outcomes). Recent work has elaborated on this result, suggesting that the mediofrontal brain response to reinforcing feedback might instead be dominated by a positive-going potential sensitive primarily to better-than-expected outcomes (the *reward positivity* [RewP]; Holroyd, Pakzad-Vaezi, & Krigolson, 2008), an effect that has been linked theoretically to DA reward signaling. Regardless of whether results are interpreted as a *reward positivity* or a *punishment negativity*, results in positive reinforcement paradigms are clear that incorrect feedback results in a more negative mediofrontal ERP amplitude than correct feedback (Sambrook & Goslin, 2015).

This pattern does not appear to hold in negative reinforcement contexts. In negative reinforcement, punishment omission evokes a more negative mediofrontal ERP than punishment delivery (an "aversion positivity"; Talmi, Atkinson, & El-Deredy, 2013), an effect that is well replicated (Hird, El-Deredy, Jones, & Talmi, 2018; Huang & Yu, 2014; Rawls, Miskovic, Moody, et al., 2020; Soder & Potts, 2018). However, results describing the brain response to graded PEs in negative reinforcement are incomplete, as the most recent meta-analysis of magnitude effects on the feedback-locked ERP did not separate negative and positive reinforcement (Sambrook & Goslin, 2015). In addition to a need for careful examination of negative reinforcement PE magnitude effects on the ERP, a detailed understanding of the brain response to graded PEs in negative reinforcement can be yielded by supplementing ERP analysis with time-frequency analysis. This approach is particularly attractive since it is argued that the FRN and RewP might be separable components of the ERP, with the FRN reflecting a theta-band response to aversive PEs, and the RewP reflecting a delta-band response to reward PEs (Bernat, Nelson, & Baskin-Sommers, 2015; Cavanagh, 2015; Cavanagh, Frank, Klein, & Allen, 2010; Sambrook & Goslin, 2016).

Given that a primary reason for encoding PEs should be to drive behavioral adaptation (Sutton & Barto, 2018), it is to be expected that PE-encoding brain activation should predict reinforcement learning. However, the relationship of the FRN/RewP to behavioral adaptation is infrequently examined, and results linking the feedback-locked ERP to behavior are inconsistent (Walsh & Anderson, 2012). For example, some studies have shown that the FRN/RewP changes amplitude over time only in participants who show behavioral learning (Bellebaum & Daum, 2008; Krigolson, Pierce, Holroyd, & Tanaka, 2009; Walsh & Anderson, 2011), yet other studies have shown behavioral reinforcement learning with no relationship to feedback-locked ERP amplitude (Cavanagh, 2015). Direct evidence of within-subject links between feedback-locked neural activation and subsequent behavioral adaptation is rarer still [but see (Fischer & Ullsperger, 2013)]. Even more critical to the current study, we are unaware of any analysis examining direct links between neural PE representations and negative reinforcement learning in humans.

The current report utilized a theory-driven task design to deliver positive and negative PEs with parametrically varying magnitudes in a task with matched positive and negative reinforcement outcome modalities, unlike previous tasks that have contrasted monetary rewards and shock avoidance (Heydari & Holroyd, 2016; Mulligan & Hajcak, 2018; Talmi et al., 2013). We randomized outcomes in each correct trial to produce single-trial PEs without being confounded by behavioral errors. We examined neural representations of PE magnitude using a single-trial ERP analysis within single subjects. Given recent evidence that delta and theta activity might serve to differentiate reward and aversion encoding, we supplemented ERP analyses using delta and theta time-frequency analysis. We hypothesized that, specific to negative reinforcement trials, the FRN/RewP would grow more positive as outcomes grew more aversive (an "aversion positivity"). We further hypothesized that this aversion PE response would predict behavioral adaptation in negative reinforcement. As neurophysiological evidence indicates that positive reinforcement learning mostly relies on mesolimbic (rather than mesocortical) DA projections, we hypothesized that positive reinforcement conditions would be accompanied by behavioral learning but no mediofrontal PE.

2. Materials and methods

2.1. Participants

We report how we determined our sample size, all data exclusions (if any), all data inclusion/exclusion criteria, whether inclusion/exclusion criteria were established prior to data analysis, all manipulations, and all measures in the study. Sample size was determined based on prior literature in the area, specifically similar single-trial studies reported in (Sambrook & Goslin, 2016; who recruited 46 subjects) and (Sambrook & Goslin, 2014; who recruited 55 subjects). We initially recruited 56 subjects with the intention of including a minimum of 50 subjects with useable data, but all subjects

were able to be included following preprocessing. No post-recruitment exclusion criteria were used, and minimal pre-recruitment exclusion criteria (detailed below) were in place as laboratory protocol prior to study design. Data collection and analyses for this study were not pre-registered.

56 undergraduates (37 female, mean age 19.2 [SD 2.06], 2 left handed) completed the study after giving informed consent. Participants were excluded from participating in the study if they had a self-reported psychiatric diagnosis, uncorrected visual impairments, or were currently using psychoactive medication (exclusion established prior to participation). All procedures were approved by the University of Arkansas Institutional Review Board (Protocol # 1708016049). Participants were compensated with course credit, as well as a monetary bonus based on their performance (mean: $7, min: $2, max: $18). As our study included two left handed subjects, we ran all described analyses after excluding the 2 left-handed subjects to ensure that our results are not confounded by participant handedness. In all cases, results were virtually unchanged; therefore, we present results including both right- and left-handed participants.

During this study, participants completed the reported reinforcement learning task in addition to a number of questionnaires, which are not of interest to the current report. These questionnaires included demographics information, the State-Trait Anxiety Inventory (Spielberger, 1983), the WHO AUDIT and ASSIST questionnaires (Saunders, Aasland, Babor, de la Fuente, & Grant, 1993), the Tobacco Craving Questionnaire Short Form (Heishman, Singleton, & Pickworth, 2008), the Minnesota Nicotine Withdrawal Scale (Hughes, Gust, Skoog, Keenan, & Fenwick, 1991), the Fagerström Tobacco and Nicotine Dependence survey (Fagerström, 1978), the UPPS-P Impulsive Behavior Scale (Lynam, Whiteside, Smith, & Cyders, 2006), the Adult Temperament Questionnaire Short Form (Evans & Rothbart, 2007), the BIS-11 (Patton, Stanford, & Barratt, 1995), and the Five Facets of Mindfulness Questionnaire (Baer, Smith, Hopkins, Krietemeyer, & Toney, 2006). None of these self-report instruments were used in this report, which was concerned only with behavioral and brain responses in reinforcement learning.

2.2. Reinforcement learning task

Task stimuli were presented in E-Prime 2.0 on a black background. See Fig. 1 for a graphical depiction of the behavioral task and analysis strategy. At the beginning of every trial, a fixation cross was presented. Participants were then shown a white circle or square for 500 msec, signifying positive or negative reinforcement respectively. Participants were informed during the task instruction block what each cue shape signified, and the task practice did not begin until each participant verbally demonstrated an understanding of the meaning of the two task cues. Negative and positive reinforcement were presented in pseudo-random order, with all participants receiving the same task order. This shape was followed by a fixation cross (+) that lasted 500–700 msec. A set of congruent (<< < < < or > > > > >) or incongruent (<< > < < or > > < > >) flanker arrow stimuli were then shown for 100 msec, followed by a fixation cross lasting from 900 to 1100 msec. Valenced feedback (correct or incorrect) was shown for 500 msec, followed by a fixation cross lasting 500–700 msec. Finally, point feedback was shown for 1000 msec.

The critical task manipulation lay in the final presentation of point feedback. The average (thus expected) return for correct positive reinforcement trials was 50 points, but the actual return fluctuated within 30 points of that on a trial-by-trial basis (i.e., for positive reinforcement wins, participants could earn anywhere from 20 points to 80 points, with an average win of 50 points). Likewise, the average (thus expected) return for correct negative reinforcement trials was zero points, but the actual return varied from −30 to 30 points. Critically, participants were aware that for both positive and negative reinforcement trials, correct answers always resulted in better outcomes (50 points higher) than incorrect answers. Before participants began the main task, they completed a minimum of 50 trials for practice. Participants could not move on to the main task until they demonstrated understanding of the task during the practice, measured as an accuracy rate of 80% or above during practice. During practice trials, the outcome was always as expected (i.e., for positive reinforcement correct answers resulted in a gain of 50 points while incorrect answers resulted in an outcome of 0 points, and for negative reinforcement correct answers resulted in an outcome of 0 points and incorrect answers resulted in a loss of 50 points). Participants responded using their left thumb (for left arrow targets) or their right thumb (for right arrow targets). Participants completed 960 trials of the main task, divided into 16 blocks of 60 trials each, with rest breaks between blocks. Participants required approximately 90 min to complete the task. By creating an expectation of average outcome in the practice and then systematically providing more or less points than expected on correct trials, we are able to separate the effect of PE (was it better or worse than expected?) from the effect of error monitoring (was the outcome correct or incorrect?). The current analysis was restricted to sets of two correct trials in a row, in order to remove confounding effects of error/performance monitoring on PE signaling (for the first trial), and in order to allow interpretation of reaction time speeding as improved behavior (for the second trial).

2.3. EEG processing

EEG were sampled at 1000 Hz using a 129-channel EGI sensor array referenced to vertex (Philips EGI, Inc.). Recording began after impedances were reduced below 50 kΩ. Data were processed using EEGLAB 15 (Delorme & Makeig, 2004) and MATLAB 2018b. Continuous data were downsampled to 125 Hz with anti-aliasing, low-pass filtered at 35 Hz using a zero-phase FIR filter, and high-pass filtered at .1 Hz using a zero-phase FIR filter. Bad channels were removed using joint probabilities with built-in EEGLAB functions (cutoff of 4 standard deviations). Copies were made of each dataset, which were high-pass filtered at 1 Hz using a zero-phase FIR filter, in preparation for computing independent components analysis (ICA) (Makeig, Bell, Jung, & Sejnowski, 1996). All data were epoched into 3 sec windows surrounding point feedback onset (from −1000 msec before to 2000 msec after). Infomax ICA was computed on the 1 Hz filtered dataset (Winkler, Debener, Muller, & Tangermann, 2015, pp.

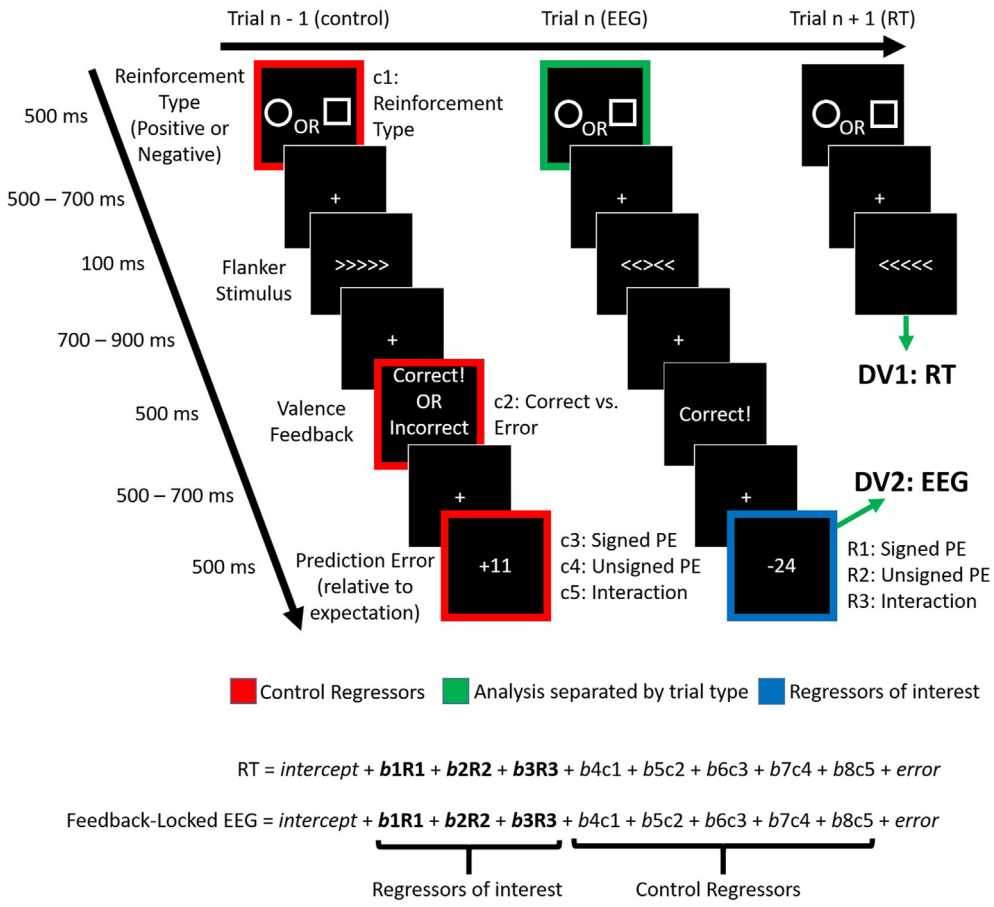

Fig. 1 – Task diagram and analysis for the modified reinforcement flanker paradigm. Participants were cued as to whether the current trial was to be positive or negative reinforcement with a white square or a white circle, respectively. Participants then had to respond to a flanker arrow stimulus that was either congruent or incongruent (congruent and incongruent trials were considered equivalent for the current manuscript). Only correct trials were analyzed (trials n and n+1). Finally, participants were given some amount of points (on average + 50 for positive reinforcement and +0 for negative reinforcement). Every trial, the amount of points given deviated slightly from the overall mean expectation, generating outcomes that were worse-than-expected or better-than-expected. All task events outlined in red are used as control variables of no interest in analysis of RTs, ERP, and time-frequency regressions. Regressions were run separately for positive and negative reinforcement conditions. Behavior and brain analysis proceeded with three regressors: 1) signed PE, 2) unsigned PE (absolute value of signed PE), 3) weighted PE (interaction of signed and unsigned PE). Regressors of interest are outlined in green. The dependent variable for behavioral analysis was the RT immediately following reinforcement outcomes, and the dependent variable for EEG analysis was the single-trial ERP amplitude (or delta/theta-band power) time-locked to reinforcement outcomes (PEs). Dependent variables are outlined in blue.

4101–4105). Bad channels were not interpolated before running ICA. Likely artifactual independent components were detected using the SASICA plugin (Chaumon, Bishop, & Busch, 2015) using a combination of three methods: 1) autocorrelation statistics, 2) focal component activity, and 3) routines from the ADJUST plugin (Mognon, Jovicich, Bruzzone, & Buiatti, 2011). ICs were additionally labeled using the ICLabel plugin (Pion-Tonachini, Kreutz-Delgado, & Makeig, 2019), and ICs labeled as eye, heart, muscle, line noise, or channel noise with greater than 70% confidence were marked as artifactual. ICA weights and artifact components calculated in the 1 Hz high-pass filtered dataset were copied to the .1 Hz high-pass filtered dataset, and detected artifactual ICs were visually verified and removed from the data; all further analyses were performed on the .1 Hz filtered dataset. For ERP analyses, data were epoched from −200 prior to point feedback to 800 msec following point feedback, and for time-frequency analysis continuous data were epoched from −2500 msec prior to point feedback to 3500 msec after point feedback to eliminate edge artifacts resulting from wavelet convolution (Cohen, 2014). Epochs containing fluctuations with voltage exceeding ±140 μV were detected and removed as well. Removed channels were interpolated using spherical splines and data were

rereferenced to the montage average. Finally, artifact-free single trials of EEG were baseline corrected for the 200 msec preceding feedback presentation in preparation for single-trial ERP analysis.

To more finely examine neural processing of reinforcing feedback, we also examined time-frequency activation by convolving single trials of EEG with a family of Morlet wavelets (EEGLAB *newtimef()* function). We extracted single trials of delta power by using wavelets centered at 1, 2, and 3 Hz, and we extracted single trials of theta power by using wavelets centered at 4, 5, 6, 7, and 8 Hz. All wavelets had a width of 3 cycles. Instantaneous power was extracted by taking the absolute value of the squared result of the convolution, then averaging over all frequencies within the frequency band (delta or theta). Single trials of time-frequency power data were not baseline corrected prior to single-trial analysis, as any baseline offsets only contribute to the intercept of the regression (Cohen, 2014).

2.4. Control analyses for trial expectancy

Our task conditions did not differ by difficulty, and we presented the same number of trials in each category and for each possible outcome. Note that the current analysis did not examine error trials; however, examination of error rates is still necessary in order to ensure our analyses controlled for expectancy. It is possible that some systematic relationship between PE and accuracy might have existed, since we did not explicitly control accuracy rates. However, if this were the case, it would have meant some trials were experienced more frequently than others. For our results to be unambiguously interpreted, it is essential that all task categories were equally expected over the length of the task, that is, that accuracy did not differ depending on reinforcement type or prediction error. To test whether or not all trial types had equivalent expectancies, we used single-trial logistic regression within individual subjects to predict accuracy (correct/error) using reinforcement type (positive/negative) and PE (signed PE, unsigned PE, and signed × unsigned PE) (MATLAB *glmfit()* function with binomial distribution). We then tested the coefficients against a null hypothesis mean of zero using one-sample *t*-tests. From this test we are able to conclude whether or not accuracy rates, and therefore expectancy, differed by condition.

2.5. Analysis of task effects on behavioral adaptation using reaction times

We used within-subject regression models to analyze whether RTs on trial n+1 differed systematically according to prediction errors in trial n. A robust regression (MATLAB *robustfit()* function) predicted trial-by-trial RTs (trial n+1) separately following positive and negative reinforcement trials using prior trial (trial n) signed PE, prior trial unsigned PE, and the interaction of prior trial signed × unsigned PE. All regressions controlled for the influence of current trial (trial n+1) reinforcement type, as this was not of interest in measuring behavioral adaptation following specific types of reinforcement. Regressions also included potential influences of trial n-1 signed PE, unsigned PE, weighted PE (signed × unsigned interaction), and accuracy to remove potential carryover effects from prior trials and to ensure that behavioral adaptation was due only to the immediately preceding outcome. Note that our task design centered signed prediction errors on zero, ensuring that signed and unsigned PE terms were orthogonal and therefore suitable for use in simultaneous regression analysis. Previous analyses have orthogonalized correlated signed and unsigned PE terms for regression (Hauser et al., 2014), but in line with (Sambrook & Goslin, 2016), we instead used signed PEs centered on zero resulting in uncorrelated signed and unsigned PEs.

This analysis returned within-subject *b* weights for the impact of signed PE, unsigned PE, and signed × unsigned PE on behavior, separately following positive and negative reinforcement. All independent variables in the regression were z-scored; therefore, *b* weights represent units of milliseconds per regressor SD. As *b* weights are normally distributed under the null hypothesis, regression coefficients were tested against a null hypothesis of zero (no impact of PE on following trial RT) using one-sample *t*-tests. From this analysis we are able to conclude whether PEs influence behavioral adaptation in the following trial, and whether this adaptation effect differs between positive and negative reinforcement. This analysis formed our primary measure of behavioral reinforcement learning.

2.6. Single-trial analysis of signed, unsigned, and weighted PEs in the ERP and time-frequency power

Within each participant, single trials of raw artifact-free EEG (as well as single trials of theta-band and delta-band EEG) were separated by reinforcement type (positive/negative reinforcement) and analyzed at every sensor and sample from 0 to 800 msec post-stimulus. For every sensor and sample, we fit a robust regression equation that modeled EEG as a function of signed PE, unsigned PE, and an interaction term of signed × unsigned PE, as well as control regressors for prior trial characteristics (reinforcement type, signed PE, unsigned PE, signed × unsigned PE, and correct/error). This regression returned a series of sensor (129) × time point (101) *b* weights for signed PE, unsigned PE, and signed × unsigned PE. All independent variables were z-scored prior to fitting the regressions, and brain activity was left in its native scaling; therefore, *b* weights represent units of μV (for single-trial ERP analysis) or μV^2 (for single-trial delta/theta power analysis) change per SD of the regressor. Significance of regression coefficients were tested using mass univariate one-sample *t*-tests implemented in the Mass Univariate ERP Toolbox (Groppe et al., 2011a, 2011b). This analysis included only the 91 sensors located on the scalp (i.e., face electrodes were not considered). *p*-values were corrected for multiple comparisons using Benjamini and Hochberg's (1995) false discovery rate (FDR).

2.7. Single-trial analysis of relationship between brain activity and behavior

To test whether brain activity on a single trial basis predicts behavioral modification on the following trial over and above task characteristics, we used single-trial partial Spearman

correlations to predict single trial RTs (trial n+1) using single trials of raw EEG, delta power EEG, and theta power EEG (trial n) at every sensor (91 scalp sensors) and sample (101 time points), while controlling for 1) following trial (trial n+1) effects (reinforcement type), 2) current trial (trial n) effects (PE [signed, unsigned, and interaction]) and 3) prior trial (trial n-1) effects (reinforcement type, PE [signed, unsigned, and interaction], and correct/error outcome). Whole-scalp correlations were tested against a null hypothesis mean of zero (no correlation between brain activity and the following behavior) using mass univariate one-sample *t*-tests implemented in the Mass Univariate ERP Toolbox (Groppe et al., 2011a, 2011b). *p*-values were corrected for multiple comparisons using Benjamini and Hochberg's (1995) FDR. This allowed a well-controlled analysis of whether or not EEG activity predicts behavioral modification on the following trial, as predicted by reinforcement learning models.

3. Results

3.1. Control analysis of expectancies

Control analyses (within-subject logistic regression of accuracy on reinforcement type [positive or negative] and prediction errors) indicated that accuracy did not differ systematically depending on reinforcement type or PE (all $p > .16$). We conclude that all outcomes were encountered equal numbers of times and therefore equally expected, so expectancy is unlikely to be a factor in any of our reported results. As correct outcomes were equally expected across all reinforcement types and PE magnitudes, we examine only correct trials (n) followed by correct trials (n+1) in the remainder of our analyses. As accuracy rates were high in this task (mean = 80%, SD = 10%), this allowed sufficient trials for single-trial analysis within every participant.

3.2. Behavioral adaptation following specific types of reinforcement

Response times following negative reinforcement outcomes were negatively predicted by unsigned PEs (higher unsigned PE drove faster RTs on the following trial), $b = -2.79$, $t(55) = -6.38$, $p = 3.9e\text{-}8$, and positively predicted by the interaction term (signed × unsigned), $b = 1.47$, $t(55) = 2.41$, $p = .02$. For responses following positive reinforcement outcomes, RTs were negatively predicted by the interaction term (signed × unsigned), $b = -2.52$, $t(55) = -4.95$, $p = 7.5e\text{-}6$. This result indicated significant moderation effects based on the interaction of signed and unsigned PE on RTs. Based on the theoretical interpretation of unsigned PE as reflecting a learning rate (Sutton & Barto, 2018), we examined these interaction effects using unsigned PE as a factor moderating the influence of signed PE on RTs. For this analysis, we examined the influence of signed PE on following RTs at unsigned PE levels of −1 SD and +1 SD.

For negative reinforcement, at low levels of unsigned PE, signed PE did not influence RT on the following trial, but at high levels of unsigned PE, signed PE positively influenced RT on the following trial. This moderation indicates an RT speeding effect following high-salience, low-value punishments, and an RT slowing effect following high-salience, high-value punishment avoidance. For positive reinforcement, the interaction indicated that at low levels of unsigned PE, signed PE modestly influenced RT on the following trial, but at high levels of unsigned PE, signed PE negatively influenced RT on the following trial. This moderation indicates an RT speeding effect following high-salience, high-value rewards.

These results indicate that behavioral modifications follow both salient positive reinforcement rewards, and salient negative reinforcement punishments. Effects for both negative and positive reinforcement are in line with reinforcement learning models suggesting that the unsigned salience of an outcome weights the importance of that outcome's signed value in predicting behavioral adaptation. Statistical results for all behavioral analyses are displayed in Fig. 2.

3.3. Single-trial analysis of signed, unsigned, and weighted PEs in the event-related potential

Results of single-trial ERP regression analysis are shown in Fig. 3. Results for negative reinforcement indicated a strong negative signed value representation in the mediofrontal ERP starting at ~200 msec and continuing until ~550 msec post-feedback. This time period encompassed the traditional time window and scalp topography of the FRN/RewP, and indicated that ERP amplitude grew more negative as the value of outcomes grew higher (i.e., an "Aversion Positivity"). A positive unsigned PE was present in the mediofrontal ERP, but at an earlier time period (~150 msec). This positive salience effect moved posteriorly over time until ~250 msec. A negative unsigned PE effect briefly reached significance at anterior sensors at ~250 msec, indicating that the ERP grew more negative as outcomes grew more salient. An interaction of signed PE × unsigned PE was briefly significant at frontal sensors at ~200 msec, which might be a mechanism that aids in early recognition of binary good/bad outcome evaluation by magnifying the representation of values that deviate from zero but are not highly salient.

Other than a brief positive signed PE representation over parietal sensors (~150 msec), no regression results were significant for positive reinforcement conditions. Note that positive reinforcement conditions did result in strong behavioral indices of reinforcement learning, but no discernible scalp representation of PEs. This is in line with neurophysiological evidence that punishing events are largely represented by mesocortical dopamine projections, but rewarding events are largely represented by mesolimbic dopamine projections.

3.4. Single-trial analysis of signed, unsigned, and weighted PEs in delta and theta bands

As previous evidence has suggested that time-frequency analysis might provide a way to dissect overlapping brain potentials, and in particular that theta activity signaling aversive PEs might underlie the FRN, while delta activity signaling rewards but underlie the reward positivity, we repeated our single-trial raw EEG analyses using delta [1–3 Hz] and theta [4–8 Hz] EEG. Results of single-trial delta power

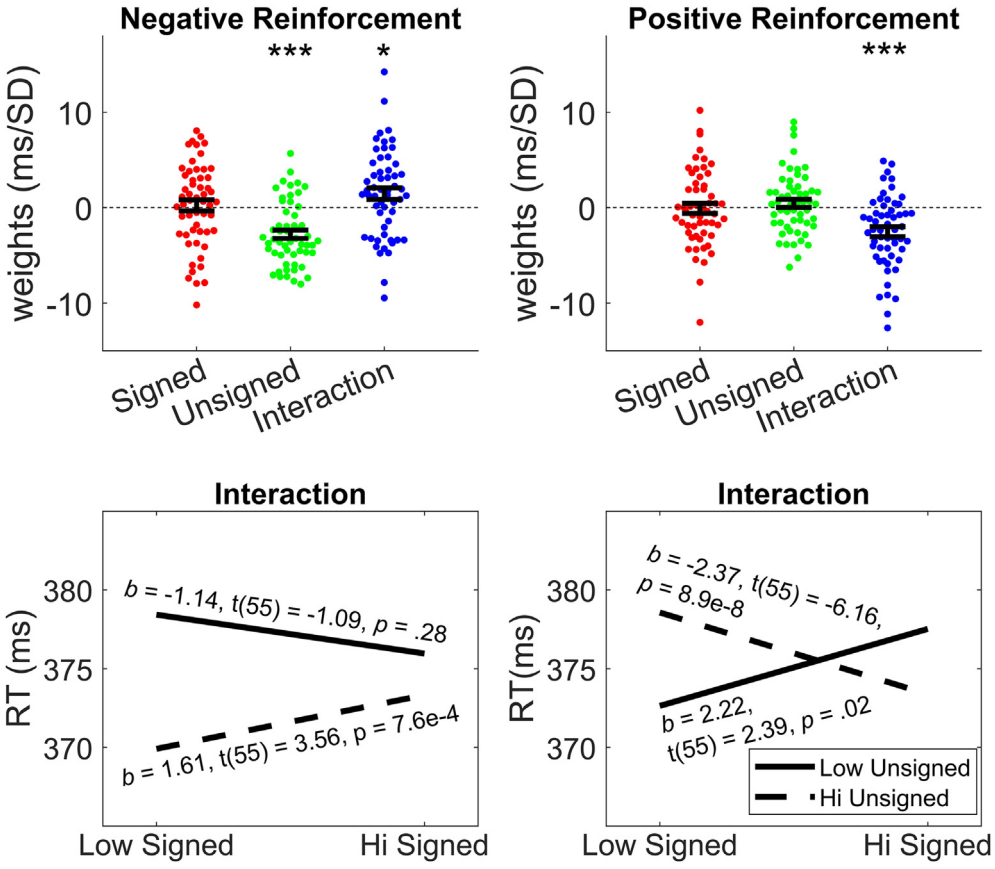

Fig. 2 – Reaction times immediately following reinforcing outcomes (trial n+1) were analyzed using single-trial regression within subjects, with trial n signed PE, unsigned PE, and the interaction (signed × unsigned PE) as regressors. This analysis also included trial n-1 reinforcement type (positive or negative), signed PE, unsigned PE, and the interaction (signed × unsigned PE), as well as accuracy, as regressors of no interest. Within-subject regression weights were tested for significance against a null-hypothesis mean of zero using one-sample t-tests. Distribution dot-plots show individual subject regression weights, with error bars corresponding to ±SEM. Distribution dot-plots were produced using the plotSpread.m function from the MATLAB file exchange (https://www.mathworks.com/matlabcentral/fileexchange/37105-plot-spread-points-beeswarm-plot). Significant moderation effects were probed at −1 SD and +1 SD as suggested by (Baron & Kenny, 1986).

regression and theta power regression are shown in Figs. 4 and 5, respectively.

Intriguingly, we found a highly significant effect of signed PE on central delta power in negative reinforcement conditions, indicating that delta power increased with more aversive outcomes in negative reinforcement. Previous results have found that delta power increases with increasing rewards in positive reinforcement conditions. This suggests a potential mechanism underlying the "aversion positivity" observed in negative reinforcement conditions – as the RewP occurs primarily in delta-band over central sensors, the observed aversion positivity in negative reinforcement might be a Reward Positivity that flips in sign for negative reinforcement (compared to positive).

Likewise, we also found a significant effect of signed PE on frontal theta power in negative reinforcement at a relatively early time period (~150–200 msec). This effect also indicated that frontal theta power increased for more aversive stimuli, in line with previous evidence that theta power increases for aversive stimuli. We also noted an effect of signed PE on occipital theta from ~100 to 400 msec, which was qualified by an interaction of signed and unsigned PE. Similar to the interaction effect present in the ERP, this might be a mechanism that aids in early recognition of binary good/bad outcome evaluation.

3.5. *Mediofrontal Brain Mechanisms of reinforcement learning*

We tested whether PE-encoding EEG signatures (trial n) predicted response times in the following trial (trial n+1) using single-trial partial Spearman correlations between full-scalp EEG and reaction times on every trial. Results indicated that central ERP activity (peaking over sensor E30 at ~248 msec) significantly positively predicted following trial reaction times (Fig. 6). This result was only significant for negative reinforcement trials. For positive reinforcement trials, no ERP-RT correlations passed significance, and no time-frequency power results (delta and theta band) were capable of

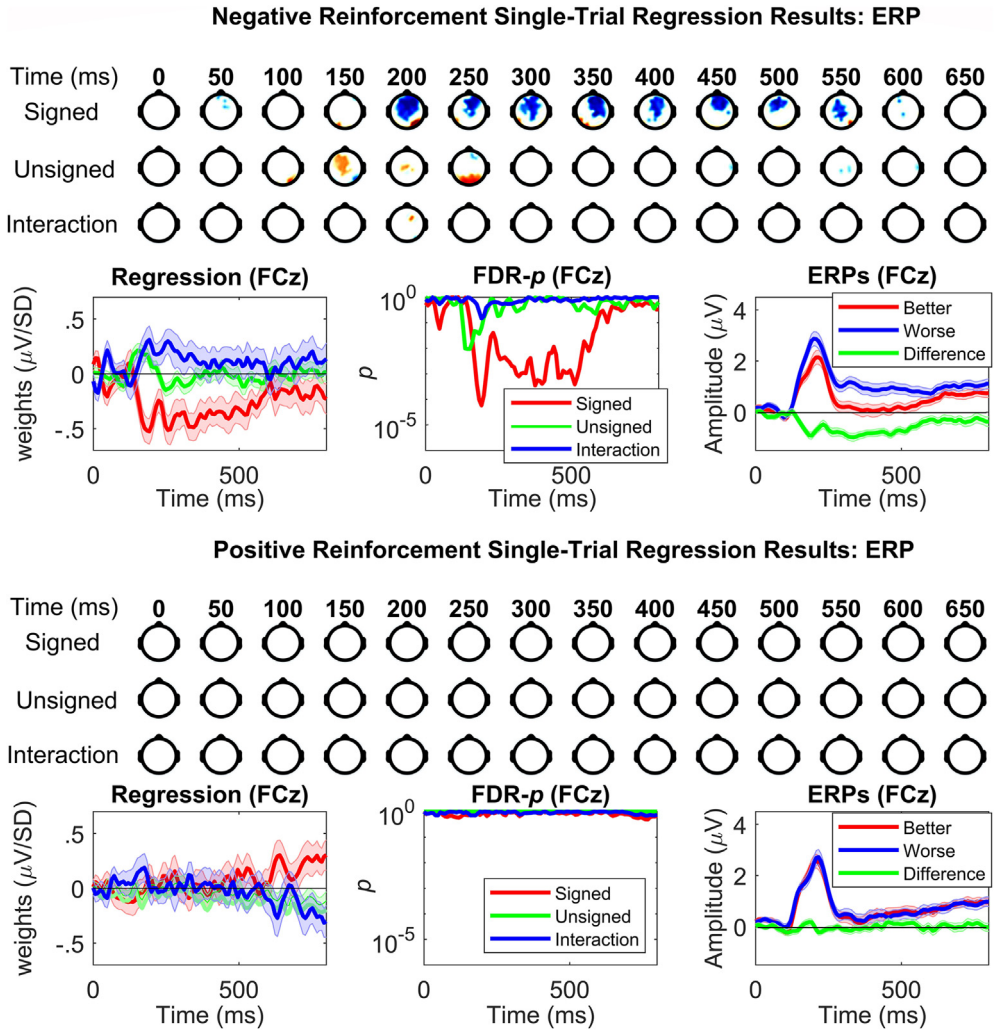

Fig. 3 – Results of single-trial ERP regression analysis. All topographic plots were masked using an alpha of .05, corrected for multiple comparisons using the false discovery rate. All line plots (regression weights and ERPs) are plotted with shading corresponding to ±SEM. For negative reinforcement conditions, a strong negative signed PE was present over mediofrontal sensors from ~200 - 550 msec. An unsigned PE was present over mediofrontal sensors at ~150 msec, and this effect shifted to be more posterior over time. Other than a brief signed PE representation over parietal sensors (~150 msec), no regression effects were significant for positive reinforcement conditions.

predicting future behavioral adaptations for any trial type. These analyses controlled for trial n-1 accuracy (correct or incorrect); trial n (ERP measurement) reinforcement type (positive or negative), signed PE, unsigned PE, and interaction (signed × unsigned PE); and trial n+1 reinforcement type (positive or negative), therefore we can conclude that ERP activity predicted future behavioral adaptation, reflected in reaction times, above-and-beyond any task effects.

4. Discussion

4.1. *General discussion*

In this study we delivered continuous prediction errors (PEs) in correct positive and negative reinforcement trials, and quantified brain and behavioral responses to PEs using single-trial reaction time (RT), ERP, and time-frequency regression analysis. We found that prediction errors (PEs) generated robust behavioral effects in both positive and negative reinforcement, whereas brain potentials signaled PEs in negative reinforcement conditions only. This replicates previous reports of an "aversion positivity" in negative reinforcement conditions (Hird et al., 2018; Huang & Yu, 2014; Rawls, Miskovic, Moody, et al., 2020; Soder & Potts, 2018; Talmi et al., 2013), and extends these results by demonstrating parametric effects of graded PEs on mediofrontal ERPs and delta-band time-frequency activity in negative reinforcement. As delta-band activity is thought to underlie the Reward Positivity (RewP; Bernat et al., 2015; Cavanagh, 2015), this suggests the negative reinforcement aversion positivity might result from a flipped RewP (increased delta activity for aversive negative reinforcement outcomes).

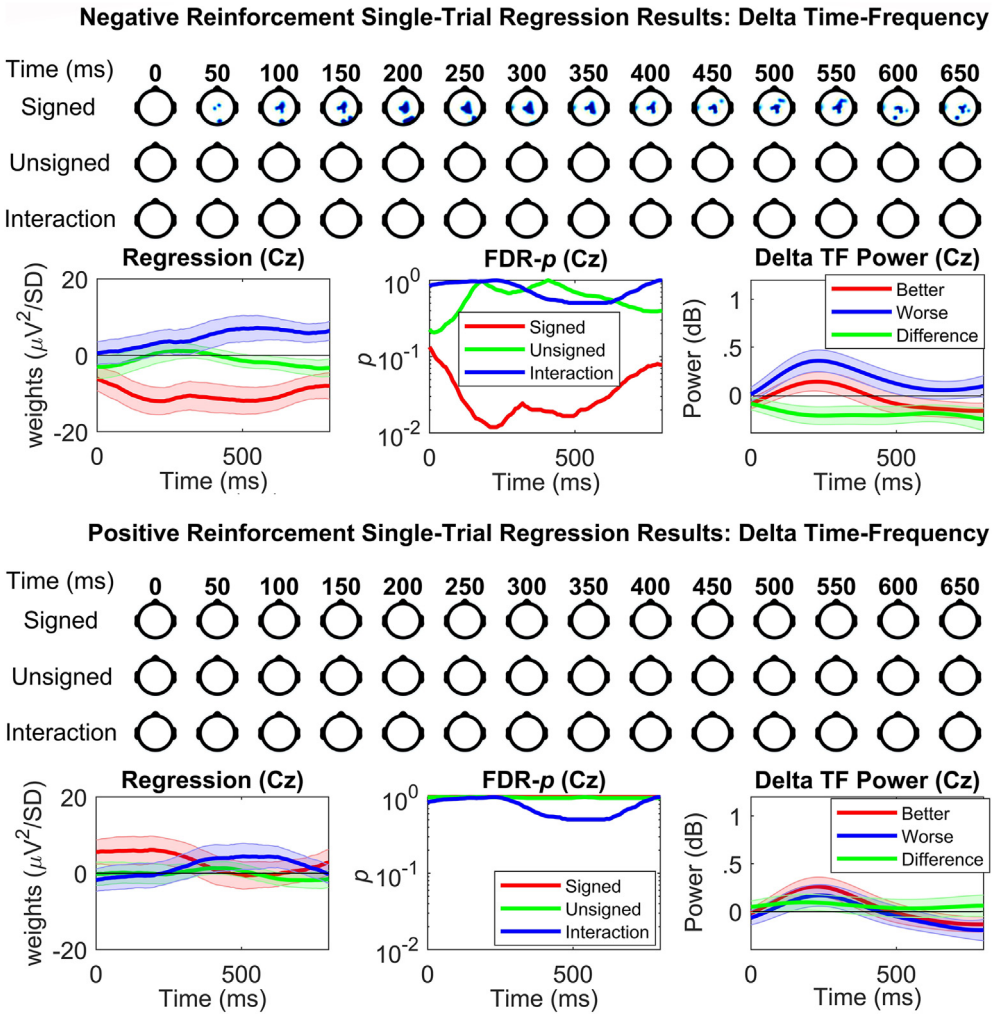

Fig. 4 – Results of single-trial delta time-frequency regression analysis. All topographic plots were masked using an alpha of .05, corrected for multiple comparisons using the false discovery rate. All line plots (regression weights and power) are plotted with shading corresponding to ±SEM. For negative reinforcement conditions, a strong negative signed PE was present over central sensors from ~100 to 650 msec. No regression effects were significant for positive reinforcement conditions.

Finally, we found that single-trial central ERPs during the FRN/RewP time window positively predicted RTs on the following trial, demonstrating a neural basis for negative reinforcement learning.

4.2. *Behavioral results*

Most EEG reinforcement studies are without corresponding behavioral analysis, but to understand brain computations behavior should be analyzed (Krakauer, Ghazanfar, Gomez-Marin, MacIver, & Poeppel, 2017). In negative reinforcement, we observed response speeding following both better-than-expected and worse-than-expected negative reinforcement outcomes; the effect of PEs on RT speeding was most notable for low value punishments. For positive reinforcement, results indicated RT speeding effect following high-value rewards. Similar effects following rewarding outcomes were reported by Sedaghat-Nejad, Herzfeld, and Shadmehr (2019), who demonstrated that human saccades toward a target were faster following high reward PEs. Interestingly, while previous studies have examined the effects of positive reinforcement outcomes on response times, our study appears to be the first to use RTs to explicitly examine negative reinforcement learning in humans. However, previous reports have examined the impact of punishments on following response times; for example, Steel, Silson, Stagg, and Baker (2016) demonstrated that punishment training resulted in response speeding. Our study extends the previous training effects for training to an examination of immediate trial-by-trial behavioral adaptation. Overall, our behavioral results demonstrate that high rewards in positive reinforcement drive behavioral modification, and both punishment and avoidance drive behavioral modification in negative reinforcement learning.

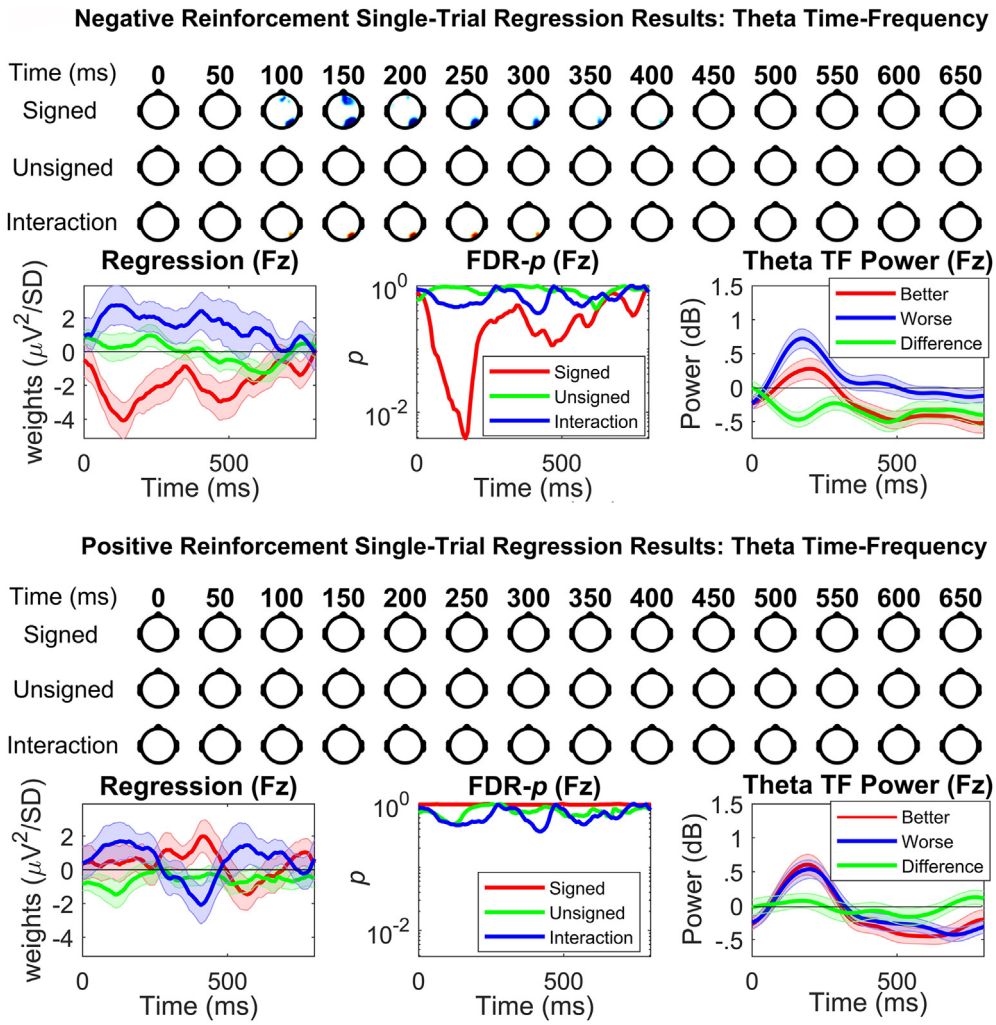

Fig. 5 – Results of single-trial theta time-frequency regression analysis. All topographic plots were masked using an alpha of .05, corrected for multiple comparisons using the false discovery rate. All line plots (regression weights and power) are plotted with shading corresponding to ±SEM. For negative reinforcement conditions, a negative signed PE was present over frontal sensors from ~100 to 200 msec, and a negative signed PE effect was significant over occipital sensors from 100 to 400 msec. This occipital effect was qualified by an interaction of signed and unsigned PE. No regression effects were significant for positive reinforcement conditions.

4.3. *Cortical representations of prediction errors*

Typical ERP studies average trials to make categorical comparisons between conditions (Luck, 2014), but real world outcomes are graded and do not map onto binary "good" versus "bad" valences. Instead, we examined parametric neural representations of reinforcing outcomes, using a whole-scalp, component-free analysis (Fischer & Ullsperger, 2013; Rawls, Miskovic, & Lamm, 2020; Rousselet et al., 2008, 2009). While this analysis did not rely on measurement of "peaks" in the waveform, our hypotheses nevertheless rested largely on a feedback-locked brain potential variously called the feedback-locked negativity (FRN) or the reward positivity (RewP). Initial reports primarily characterized the FRN/RewP using difference wave techniques – such that the FRN and RewP were in fact the same component, just mirror images of each other (Proudfit, 2015). More recent evidence suggests that the RewP and FRN might in fact be separable components of the evoked potential, with the FRN encompassing a topography and time period similar to the mediofrontal N2, and the RewP encompassing a more central-parietal topography. Time-frequency analysis might be able to separate these overlapping components, as the FRN is believed to reflect a theta-band response to aversive PEs, while the RewP instead reflects a delta-band response to reward PEs (Bernat et al., 2015; Cavanagh, 2015; Cavanagh et al., 2010; Sambrook & Goslin, 2016).

In line with previous analyses, in negative reinforcement we found that the mediofrontal ERP covaried negatively with signed PEs. While this "aversion positivity" effect is well replicated (Hird et al., 2018; Huang & Yu, 2014; Pfabigan et al., 2015; Rawls, Miskovic, Moody, et al., 2020; Soder & Potts, 2018; Talmi et al., 2013), the representation of graded PEs during

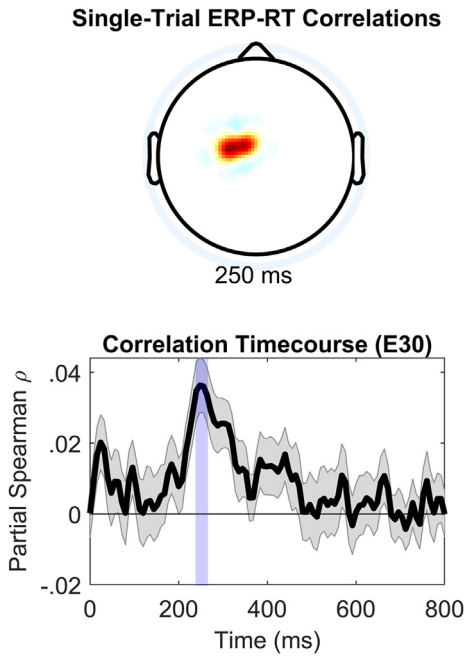

Fig. 6 – Single-trial partial Spearman correlations between ERP amplitude and following-trial reaction times. For negative reinforcement, results indicated that central ERP activity during the FRN/RewP time period (maximal at ~250 msec) significantly positively predicted response times on the following (correct) trial. No correlation results passed significance for positive reinforcement conditions. Nonsignificant results are masked in the topographic plot (top), and red shading in the bottom plot indicates ±SEM around the Spearman correlation coefficients. The blue shaded rectangle in the timecourse plot indicates timepoints that were significant following correction for multiple comparisons.

negative reinforcement in the FRN/RewP is still unknown, as a prior meta-analysis of PE magnitude effects on the FRN/RewP did not separate positive and negative reinforcement (Sambrook & Goslin, 2015). As such, our study appears to be the first to demonstrate parametric effects of PE magnitude in shaping the negative reinforcement aversion positivity. More importantly, however, our results require a different theoretical interpretation than previous reports of the aversion positivity. Previous discussions of the aversion positivity in negative reinforcement have interpreted this effect as evidence of salience coding, but we were able to dissociate salience coding from value coding within the ERP using orthogonal regressors. In fact, we found that signed and unsigned PEs were both expressed in the negative reinforcement feedback-locked ERP, but were largely dissociable – unsigned PEs were reflected in the mediofrontal ERP in an earlier time period than the usual FRN/RewP (Sambrook & Goslin, 2014, 2015, 2016). This suggests that, while we were able to replicate previous reports of a negative reinforcement aversion positivity, this potential cannot be interpreted as a salience signal. Instead, the aversion positivity must be interpreted as a signed PE signal, or a flipped reward positivity.

To cement our interpretation of these ERP effects, we decomposed single trials of EEG into delta (1–3 Hz) and theta (4–8 Hz) bands (thought to underlie the RewP and the FRN, respectively). Indeed, in delta-band activity we found evidence that the aversion positivity previously noted in negative reinforcement conditions can be interpreted as a flipped reward positivity – in negative reinforcement conditions, delta activity covaried negatively with the magnitude of the signed PE regressor, indicating that delta-band activity increased with increasingly aversive outcomes. Our results indicated that in negative reinforcement, theta band activity over frontal sensors also increased in power with increasingly aversive outcomes. Thus, our results demonstrate that in negative reinforcement, theta band activity follows the expected pattern (i.e., encoding aversive information), whereas delta-band activity seems to give rise to the apparent "flipped" reward positivity in negative reinforcement. This is the opposite effect attributed to delta band activity, and the RewP, in positive reinforcement conditions (Cavanagh, 2015), and suggests the need for a new theoretical interpretation of the RewP component in terms of brain responses in both positive and negative reinforcement – as this component can clearly not be interpreted as unilaterally reflecting reward processing under the current results.

Single-trial ERP results were not significant following positive reinforcement, other than a brief signed PE effect over parietal regions (~150 msec). This might point to a flaw in the interpretation of previous study designs. Typical ERP studies of reinforcement not only either deliver or omit rewards, but also deliver win/loss feedback. This presents a confounding factor, because mediofrontal ERPs are sensitive to errors and error feedback. Prediction errors do not require behavioral errors, but only that outcomes are better or worse than expected. Following correct feedback, we gave participants a number of points that was better than, worse than, or as expected, isolating recognition of PEs from recognition of errors. Mediofrontal responses to omitted rewards might be a response to outcomes that are perceived as "errors," rather than a response to prediction errors specifically.

4.4. *Mediofrontal predictors of behavioral modification*

We demonstrated that the FRN/RewP can definitively be interpreted as reflecting signed aversion signals in negative reinforcement. Our results demonstrate that this aversion positivity can also be unambiguously interpreted as a reinforcement learning signal. Our single-trial approach enabled us to relate ERP amplitudes during reinforcement processing to immediate (post-feedback) behavioral modification, thus providing a stringent test of brain–behavior relationships (Bridwell et al., 2018). We found a positive relationship between central ERP amplitude in negative reinforcement trials and RT on the following trial, controlling for task-related influences on behavior. This demonstrates that the ERP predicts negative reinforcement learning above-and-beyond any task effects.

Intriguingly, previous results attempting to link the feedback-locked ERP to behavioral modification have generated inconclusive evidence. Feedback-locked ERPs can differentiate learners from non-learners in reinforcement learning

tasks (Bellebaum & Daum, 2008; Krigolson et al., 2009; Walsh & Anderson, 2011), but sophisticated single-trial analyses (Cavanagh, 2015) failed to find any within-subject link between feedback-locked brain activity and future behavioral modification [but see (Fischer & Ullsperger, 2013), although this study implicated parietal P3 rather than FRN/RewP]. Our finding linking the FRN/RewP to behavioral modification leads to an intuitive comparison of these reported effects with a related ERP component that robustly predicts behavioral modification (Cavanagh & Shackman, 2015) – the error-related negativity (ERN). Indeed, the initial descriptions of the FRN/RewP likened it to the previously described ERN component, a mediofrontal/central component peaking within the first 100 msec following commission of a behavioral error (Gehring, Goss, Coles, Meyer, & Donchin, 1993; Miltner, Braun, & Coles, 1997), and suggested the existence of a feedback-locked negativity provided evidence of a generic brain system for error monitoring (Holroyd & Coles, 2002). Similar to our result for the FRN/RewP, the ERN has been shown to predict post-error behavioral adjustments in a meta-analysis (Cavanagh & Shackman, 2015) and many individual reports (Beatty, Buzzell, Roberts, & Mcdonald, 2020; Debener et al., 2005; Fischer, Danielmeier, Villringer, Klein, & Ullsperger, 2016; Gehring et al., 1993; Kalfaoğlu, Stafford, & Milne, 2018). While it is important to note that our current analyses did not include any error trials (or even post-error trials, in the case of RT measurement), and therefore is unconfounded by error monitoring, prior error monitoring results are nevertheless important to interpreting the results of the current study. In particular, it is possible that the feedback-locked ERP predicts behavioral modification only following aversive outcomes, or outcomes indicating some "generic" error – whether that is an error in prediction, or an error in performance.

4.5. *Advantages of our experiment over previous experiments*

Our task design and analysis advances the literature on human reinforcement learning in several ways. As the FRN is context-dependent (Pfabigan et al., 2015), we used random ordering of positive and negative reinforcement trials coupled with single-trial regression of previous trial modality, therefore controlling for any sequence effects. Second, our task design presented PEs without errors or error feedback. Typical FRN studies conflate correct answers with "good" outcomes, and incorrect answers with "bad" outcomes, confounding reinforcement processing and performance monitoring. Prior experiments examining positive and negative reinforcement are also confounded by failure to match outcome modality. Mulligan and Hajcak (2018) compared reward/loss to punishment/omission, and Heydari and Holroyd (2016) and Talmi et al. (2013) compared reward/omission with punishment/omission. However, in these tasks, conditions differed in outcome modality (money *vs* shock) and reinforcement timing (money is paid at the end of the study, but shocks are delivered immediately), making it impossible to draw direct comparisons between positive and negative reinforcement. Most critical to the theoretical implications of this report, previous reports of the aversion positivity have interpreted this potential as evidence of salience coding in the ERP. Our task design ensured that signed and unsigned PE terms were orthogonal and therefore suitable for use in simultaneous regression analysis. Being able to include both terms in a single model allowed us to generate a novel conclusion – that the negative reinforcement aversion positivity is not, in fact, evidence of salience coding, but is instead aversive value coding.

Author contributions

Eric Rawls: Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Data Curation, Writing – Original Draft. Writing – Review & Editing, Visualization, Project Administration, Funding Acquisition.

Connie Lamm: Conceptualization, Resources, Writing – Review and Editing, Supervision, Project Administration, Funding Acquisition.

Funding

ER received support for this work from the National Institutes of Mental Health (T32-MH115866), and ER and CL received support for this work from the Arkansas Biosciences Institute (0402-27504-21-0216). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Mental Health or the Arkansas Biosciences Institute.

Data and code availability

All raw data analyzed in the current study, sufficient summary data to recreate every figure in this manuscript, stimulus presentation scripts, and all analysis code, are archived and publicly available at (https://osf.io/wu37k/).

Open practices

The study in this article earned Open Data and Open Materials badges for transparent practices. Data and Materials for this study are available at on request.

Declaration of competing interest

The authors declare no competing financial interests.

REFERENCES

Baer, R. A., Smith, G. T., Hopkins, J., Krietemeyer, J., & Toney, L. (2006). Using self-report assessment methods to explore facets of mindfulness. *Assessment, 13*(1), 27–45. https://doi.org/10.1177/1073191105283504

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173. https://doi.org/10.1037/0022-3514.51.6.1173

Beatty, P., Buzzell, G., Roberts, D., & Mcdonald, C. (2020). Contrasting time and frequency domains: ERN and induced theta oscillations differentially predict post-error behavior. *Cognitive, Affective & Behavioral Neuroscience, 20*. https://doi.org/10.3758/s13415-020-00792-7

Bellebaum, C., & Daum, I. (2008). Learning-related changes in reward expectancy are reflected in the feedback-related negativity. *The European Journal of Neuroscience, 27*(7), 1823–1835. https://doi.org/10.1111/j.1460-9568.2008.06138.x

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological), 57*(1), 289–300. JSTOR.

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

Bridwell, D. A., Cavanagh, J. F., Collins, A. G. E., Nunez, M. D., Srinivasan, R., Stober, S., et al. (2018). Moving beyond ERP components: A selective review of approaches to integrate EEG and behavior. *Frontiers in Human Neuroscience, 12*, 106. https://doi.org/10.3389/fnhum.2018.00106

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

Chaumon, M., Bishop, D. V. M., & Busch, N. A. (2015). A practical guide to the selection of independent components of the electroencephalogram for artifact correction. *Journal of Neuroscience Methods, 250*, 47–63. https://doi.org/10.1016/j.jneumeth.2015.02.025

Cohen, M. X. (2014). *Analyzing neural time series data: Theory and practice*. MIT Press.

Debener, S., Ullsperger, M., Siegel, M., Fiehler, K., Cramon, D. Y. von, & Engel, A. K. (2005). Trial-by-Trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *Journal of Neuroscience, 25*(50), 11730–11737. https://doi.org/10.1523/JNEUROSCI.3286-05.2005

Delorme, A., & Makeig, S. (2004). Eeglab: An open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods, 134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

Evans, D., & Rothbart, M. (2007). Developing a model of adult temperament. *Journal of Radiological Protection: Official Journal of the Society for Radiological Protection, 41*, 868–888. https://doi.org/10.1016/j.jrp.2006.11.002

Fagerström, K. O. (1978). Measuring degree of physical dependence to tobacco smoking with reference to individualization of treatment. *Addictive Behaviors, 3*(3–4), 235–241. https://doi.org/10.1016/0306-4603(78)90024-2

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*(1), 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

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science, 4*(6), 385–390.

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011a). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology, 48*(12), 1711–1725. https://doi.org/10.1111/j.1469-8986.2011.01273.x

Groppe, D. M., Urbach, T. P., & Kutas, M. (2011b). Mass univariate analysis of event-related brain potentials/fields II: Simulation studies. *Psychophysiology, 48*(12), 1726–1737. https://doi.org/10.1111/j.1469-8986.2011.01272.x

Hauser, T. U., Iannaccone, R., Stämpfli, P., Drechsler, R., Brandeis, D., Walitza, S., et al. (2014). The feedback-related negativity (FRN) revisited: New insights into the localization, meaning and network organization. *Neuroimage, 84*, 159–168. https://doi.org/10.1016/j.neuroimage.2013.08.028

Heishman, S. J., Singleton, E. G., & Pickworth, W. B. (2008). Reliability and validity of a Short Form of the tobacco craving questionnaire. *Nicotine & Tobacco Research: Official Journal of the Society for Research on Nicotine and Tobacco, 10*(4), 643–651. https://doi.org/10.1080/14622200801908174

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), Article e12976. https://doi.org/10.1111/psyp.12976

Hollerman, J. R., & Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nature Neuroscience, 1*(4), 304–309. https://doi.org/10.1038/1124

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

Holroyd, C. B., Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-related positivity: Sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology, 45*(5), 688–697. https://doi.org/10.1111/j.1469-8986.2008.00668.x

Huang, Y., & Yu, R. (2014). The feedback-related negativity reflects "more or less" prediction error in appetitive and aversive conditions. *The Florida Nurse, 8*. https://doi.org/10.3389/fnins.2014.00108

Hughes, J. R., Gust, S. W., Skoog, K., Keenan, R. M., & Fenwick, J. W. (1991). Symptoms of tobacco withdrawal. A replication and extension. *Archives of General Psychiatry, 48*(1), 52–59. https://doi.org/10.1001/archpsyc.1991.01810250054007

Kalfaoğlu, Ç., Stafford, T., & Milne, E. (2018). Frontal theta band oscillations predict error correction and posterror slowing in typing. *Journal of Experimental Psychology. Human Perception and Performance, 44*(1), 69–88. https://doi.org/10.1037/xhp0000417

Krakauer, J. W., Ghazanfar, A. A., Gomez-Marin, A., MacIver, M. A., & Poeppel, D. (2017). Neuroscience needs behavior: Correcting a reductionist bias. *Neuron, 93*(3), 480–490. https://doi.org/10.1016/j.neuron.2016.12.041

Krigolson, O. E., Pierce, L. J., Holroyd, C. B., & Tanaka, J. W. (2009). Learning to become an expert: Reinforcement learning and the acquisition of perceptual expertise. *Journal of Cognitive Neuroscience, 21*(9), 1834–1841. https://doi.org/10.1162/jocn.2009.21128

Luck, S. J. (2014). *An introduction to the event-related potential technique*. MIT Press.

Lynam, D. R., Whiteside, S. P., Smith, G. T., & Cyders, M. A. (2006). *The UPPS-P: Assessing five personality pathways to impulsive behavior*. Purdue University.

Makeig, S., Bell, A. J., Jung, T.-P., & Sejnowski, T. J. (1996). Independent component analysis of electroencephalographic data. *Advances in Neural Information Processing Systems, 7*.

Miltner, W. H., Braun, C. H., & Coles, M. G. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a "generic" neural system for error detection. *Journal of Cognitive Neuroscience, 9*(6), 788–798. https://doi.org/10.1162/jocn.1997.9.6.788

Mognon, A., Jovicich, J., Bruzzone, L., & Buiatti, M. (2011). Adjust: An automatic EEG artifact detector based on the joint use of spatial and temporal features: Automatic spatio-temporal EEG artifact detection. *Psychophysiology, 48*(2), 229–240. https://doi.org/10.1111/j.1469-8986.2010.01061.x

Mulligan, E. M., & Hajcak, G. (2018). The electrocortical response to rewarding and aversive feedback: The reward positivity does not reflect salience in simple gambling tasks. *International Journal of Psychophysiology, 132*, 262–267. https://doi.org/10.1016/j.ijpsycho.2017.11.015

Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the barratt impulsiveness scale. *Journal of Clinical Psychology, 51*(6), 768–774. https://doi.org/10.1002/1097-4679(199511)51:6<768::AID-JCLP2270510607>3.0.CO;2-1

Pearce, J. M., & Hall, G. (1980). *A model for pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli* (Vol. 21).

Pfabigan, D. M., Seidel, E.-M., Paul, K., Grahl, A., Sailer, U., Lanzenberger, R., et al. (2015). Context-sensitivity of the feedback-related negativity for zero-value feedback outcomes. *Biological Psychology, 104*, 184–192. https://doi.org/10.1016/j.biopsycho.2014.12.007

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. *Neuroimage, 198*, 181–197. https://doi.org/10.1016/j.neuroimage.2019.05.026

Proudfit, G. H. (2015). The reward positivity: From basic research on reward to a biomarker for depression. *Psychophysiology, 52*(4), 449–459. https://doi.org/10.1111/psyp.12370

Rawls, E., Miskovic, V., & Lamm, C. (2020a). Delta phase reset predicts conflict-related changes in P3 amplitude and behavior. *Brain Research, 1730*, 146662. https://doi.org/10.1016/j.brainres.2020.146662

Rawls, E., Miskovic, V., Moody, S. N., Lee, Y., Shirtcliff, E. A., & Lamm, C. (2020b). 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

Rescorla, R. A., & Wagner, A. R. (1972). *3 A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement* (Vol. 18).

Rousselet, G. A., Husk, J. S., Pernet, C. R., Gaspar, C. M., Bennett, P. J., & Sekuler, A. B. (2009). Age-related delay in information accrual for faces: Evidence from a parametric, single-trial EEG approach. *BMC Neuroscience, 10*(1), 114. https://doi.org/10.1186/1471-2202-10-114

Rousselet, G. A., Pernet, C. R., Bennett, P. J., & Sekuler, A. B. (2008). Parametric study of EEG sensitivity to phase noise during face processing. *BMC Neuroscience, 9*(1), 98. https://doi.org/10.1186/1471-2202-9-98

Sambrook, T. D., & Goslin, J. (2014). Mediofrontal event-related potentials in response to positive, negative and unsigned prediction errors. *Neuropsychologia, 61*, 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

Saunders, J. B., Aasland, O. G., Babor, T. F., de la Fuente, J. R., & Grant, M. (1993). Development of the alcohol use disorders identification test (AUDIT): WHO collaborative Project on early detection of persons with harmful alcohol consumption–II. *Addiction (Abingdon, England), 88*(6), 791–804. https://doi.org/10.1111/j.1360-0443.1993.tb02093.x

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275*(5306), 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Sedaghat-Nejad, E., Herzfeld, D. J., & Shadmehr, R. (2019). Reward prediction error modulates saccade Vigor. *The Journal of Neuroscience, 39*(25), 5010–5017. https://doi.org/10.1523/JNEUROSCI.0432-19.2019

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

Spielberger, C. D. (1983). *State-Trait Anxiety Inventory for Adults*. https://doi.org/10.1037/t06496-000

Steel, A., Silson, E. H., Stagg, C. J., & Baker, C. I. (2016). The impact of reward and punishment on skill learning depends on task demands. *Scientific Reports, 6*(1), 36056. https://doi.org/10.1038/srep36056

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. The 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

Walsh, M. M., & Anderson, J. R. (2011). Learning from delayed feedback: Neural responses in temporal credit assignment. *Cognitive, Affective & Behavioral Neuroscience, 11*(2), 131–143. https://doi.org/10.3758/s13415-011-0027-0

Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: Event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience and Biobehavioral Reviews, 36*(8), 1870–1884. https://doi.org/10.1016/j.neubiorev.2012.05.008

Winkler, I., Debener, S., Muller, K.-R., & Tangermann, M. (2015). On the influence of high-pass filtering on ICA-based artifact reduction in EEG-ERP. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. https://doi.org/10.1109/EMBC.2015.7319296