# Distinct Processing of Selection and Execution Errors in Neural Signatures of Outcome Monitoring

Faisal Mushtaq*[1], Samuel D. McDougle[2], Matt P. Craddock[3], Darius E. Parvin[2,4], Jack Brookes[1], Alexandre Schaefer[5], Mark Mon-Williams[1], Jordan A. Taylor[6,7] & Richard B. Ivry[2,4]

[1]School of Psychology, University of Leeds, Leeds, West Yorkshire, United Kingdom
[2]Department of Psychology, University of California, Berkeley, CA, USA
[3]School of Psychology, University of Lincoln, Lincoln, United Kingdom
[4]Helen Wills Neuroscience Institute, University of California, Berkeley, CA, USA
[5]Department of Psychology, Monash University, Malaysia Campus
[6]Department of Psychology, Princeton University, Princeton, NJ, USA
[7]Princeton Neuroscience Institute, Princeton University, Princeton, NJ, USA

*Correspondence can be addressed to Faisal Mushtaq, School of Psychology, Faculty of Medicine and Health, University of Leeds, Leeds, LS2 9JT, UK; email: f.mushtaq@leeds.ac.uk

Abbreviated Title: ERP Correlates of Execution and Selection Error
Number of Pages: 33
Number of Figures: 8
Number of Words: Abstract = 238; Introduction = 648; Discussion = 1,457
Total Word Count: 9,552

**Conflicts of Interest:** The authors declare no competing financial interests.

**Acknowledgements:** The authors would like to thank Ms. Zeynep Uludag for her assistance with data collection and Dr. Assaf Breska for helpful discussions on an earlier draft of the manuscript. F.M and M.M-W were supported by Fellowships from the Alan Turing Institute and a Research Grant from the EPSRC (EP/R031193/1). J.A.T was supported by NS084948 from the National Institutes of Health (USA). R.B.I and D.E.P were supported by NS092079 and NS105839 from the National Institutes of Health (USA). S.D.M was supported by the National Science Foundation's Graduate Research Fellowship Program. Data collected as part of a BBSRC grant (BB/H001476) awarded to A.S are also reported.

# Abstract

Losing a point playing tennis may result from poor shot selection or poor stroke execution. To explore how the brain responds to these different types of errors, we examined EEG signatures of feedback-related processing while participants performed a simple decision-making task. In Experiment 1, we used a task in which unrewarded outcomes were framed as selection errors, similar to how feedback information is treated in most studies. Consistent with previous work, EEG differences between rewarded and unrewarded trials in the medial frontal negativity (MFN) correlated with behavioral adjustment. In Experiment 2, the task was modified such that unrewarded outcomes could arise from either poor execution or selection. For selection errors, the results replicated that observed in Experiment 1. However, unrewarded outcomes attributed to poor execution produced larger amplitude MFN, alongside an attenuation in activity preceding this component and a subsequent enhanced error positivity (Pe) response in posterior sites. In terms of behavioral correlates, only the degree of the early attenuation and amplitude of the Pe correlated with behavioral adjustment following execution errors relative to reward; the amplitude of the MFN did not correlate with behavioral changes related to execution errors. These results indicate the existence of distinct neural correlates of selection and execution error processing and are consistent with the hypothesis that execution errors can modulate action selection evaluation. More generally, they provide insight into how the brain responds to different classes of error that determine future action.

## Significance Statement

To learn from mistakes, we must resolve whether decisions that fail to produce rewards are due to poorly selected action plans or badly executed movements. EEG data were obtained to identify and compare the physiological correlates of selection and execution errors, and how these are related to behavioral changes. A neural signature associated with reinforcement learning, a medial frontal negative (MFN) ERP deflection, correlated with behavioral adjustment after selection errors relative to reward outcomes, but not motor execution errors. In contrast, activity preceding and following the MFN response correlated with behavioral adjustment after execution errors relative to reward. These results provide novel insight into how the brain responds to different classes of error that determine future action.

# Introduction

When an action fails to produce the desired goal, there is a "credit assignment" problem to resolve: Did the lack of reward occur because the wrong course of action was selected, or was it because the selected action was poorly executed? Consider a tennis player who, mid-game, must determine whether losing the last point was the result of selecting the wrong action or executing the action poorly. The player might have attempted a lob rather than the required passing shot, an error in action selection. Alternatively, a lob might have been appropriate but hit with insufficient force, an error in motor execution.

Reinforcement learning presents a framework for understanding adaptive behavior through trial and error interactions with the environment. According to numerous models (e.g. Sutton and Barto, 1998), the discrepancy between expected and actual outcomes, the reward prediction error, provides a learning signal that allows an agent to refine its predictions and update its action selection policy. But what happens when, as with our tennis player, a negative prediction error could arise from either poor action selection or poor response execution?

To address this, McDougle et al. (2016) used a "bandit" task in which participants chose between two stimuli to maximize reward. In one condition, choices were made using a standard button-press method, limiting the possibility that the negative prediction error on unrewarded trials could be the product of poor execution. In a second condition, choices were made by reaching to the desired bandit. Here, unrewarded trials were attributed to movement execution errors. In the latter condition, participants strongly discounted the negative prediction errors on unrewarded trials relative to the former condition. The authors hypothesized that errors credited to the motor execution system block value updating in the action selection system. Consistent with this, McDougle et al. (2019) have subsequently shown that reward prediction error coding in the human striatum is attenuated following execution versus selection errors. Differences between responses to selection and execution errors have been attributed to a greater sense of "agency" in the latter, with participants' choice biases indicating a belief that they can reduce execution errors by making more accurate movements (Parvin et al., 2018).

To further explore the idea that these errors produce qualitatively distinct behavioral and neural responses, we used scalp-recorded EEG to measure event-related potentials (ERPs) related to both outcomes. One well-studied outcome monitoring signal is a medial frontal negativity (MFN) shift observed after the commission of an error (error-related negativity [ERN]; Falkenstein et al., 1991; Gehring et al., 1993) or post-decision feedback

(Feeedback-related Nefgativity [FRN]; Gehring and Willoughby, 2002). It has been hypothesized that amplitude variation in the MFN is the manifestation of phasic activity modulation in mesencephalic dopamine neurons following unexpected outcomes (Holroyd and Coles, 2002). If we assume this activity is related to reinforcement learning (Cohen and Ranganath, 2007), then we would expect the MFN might be a neural correlate of the differential response of the brain to execution and selection errors in the service of decision-making.

In Experiment 1, we used a standard bandit task with minimal sensorimotor demands to establish a relationship between the MFN and behavioral adjustments following unrewarded trials, outcomes that would be attributed to selection errors. In Experiment 2, we used a modified bandit task where choices were selected via rapid arm movements. Unrewarded trials were either framed as errors in choosing the wrong bandit or the result of an inaccurate movement. In line with models of reinforcement learning, we expected the MFN to be larger on unrewarded trials compared to rewarded trial, but that this MFN response would be attenuated following execution errors. We further expected the magnitude of the MFN response to correlate with behavioral change. Specifically, we predicted that participants who exhibited larger MFN would be more likely to switch between the different options. Notably, we expected this brain-behavior relationship would hold for selection errors, but not for execution errors.

# Materials and Methods

## Participants

The purpose of Experiment 1 was to examine the relationship between outcome-locked ERPs and behavioral adjustment following non-rewards errors in a choice decision task where outcomes were framed as selection errors and precluded the possibility of non-rewards arising from errors of execution. To this end, we pooled the data from two studies that used a "classic" two-armed bandit task in which the choices were made by simple button presses. The data from one study have been published (n=27, Mushtaq et al., 2016) and the data from the other study are unpublished (n=21, details below). This provided a total sample of 48 right-handed participants (Edinburgh Handedness Inventory [EHI] > 40; Oldfield 1976), with normal or corrected-to-normal eyesight, and no history of psychiatric or neurological conditions for analysis (34 females, 14 males, μ age= 22.1 years, ±3.30 years).

In Experiment 2, 32 right-handed participants (EHI > 40) were tested on a novel sensorimotor variant of a three-arm bandit reinforcement learning task. Two participants were excluded due to excessive EEG artifacts, and a technical error during data collection rendered one participant's dataset unusable. All analyses were performed on the resulting sample of 29 participants (19 females, 10 males, μ age = 26.75 years, ±9.51 years).

Across both tasks, participants were told they would be remunerated based on their performance. However, due to the pseudo-veridical nature of outcomes (see Procedure), all received a fixed payment (£7.50 in Experiment 1 and £10.00 in Experiment 2). Participants signed an informed consent document, were fully debriefed, and the experiments were approved by the Ethics Committee in the School of Psychology at the University of Leeds, United Kingdom.

## Design and Procedure

## Experiment 1

Experiment 1 served to provide a benchmark dataset that would allow us to identify the "canonical" characteristics of error and reward-related processing in a situation where the demands on the movement execution system were limited. Key details relevant to this retrospective analysis of previously acquired data are provided below, and full details of the experimental procedure have been reported previously (Mushtaq et al., 2016).

Following preparation for EEG recordings, the participant was seated approximately 50 cm away from a 17″ Dell monitor (1280 × 1024 resolution, 60 Hz refresh rate) and asked to complete a multi-trial two-armed bandit task (controlled using E-Prime® v1.2; Psychology Software Tools, Pittsburgh, PA). On each trial, a fixation cross was presented at the center of the screen for 750 ms, followed by the presentation of a square and circle, the two bandits, with one bandit on the left side of the screen and the other on the right side (1500 ms). The participant chose one of the two bandits, using the index finger of the right hand (left button vs. right) on a stimulus response pad (Psychology Software Tools Serial Response Box, Pittsburgh, PA; **Figure 1** A & B).

Participants were instructed that they must select a combination of squares and circles over the course of the experiment to maximize the number of points earned, which would translate to real money. After each selection, outcome feedback indicated reward ("You Win!") or error ("You Lose!") and the amount of points to be added or subtracted from the total score. Outcomes were randomized and each outcome was equally weighted i.e., the probability of receiving a reward and punishment was 50%. As such, no effective rule-learning strategy was possible, and the expected value of each bandit was equal. There were a total of 416 trials.

We also included data from a second, unpublished study in which we had examined the impact of temporal decay on the MFN. Here the protocol was the same as described above, with one difference: On half of the trials, the inter-trial interval was 750 ms (as in the published data set) and on the other half, the inter-trial interval was 2000 ms. Only the 208 trials from the 750 ms condition were included in this analysis.

## Experiment 2

Experiment 2 employed a three-armed bandit reaching task (**Figure 1** C & D) in which we also included outcomes where non-rewards could be the product of poorly executed actions (McDougle et al., 2019). Following EEG set-up, the participant was seated in a chair approximately 50 cm away from a 24" ASUS monitor (53.2 X 30 cm [2560 x 1600 pixels], 100 Hz refresh rate). The participant was instructed to make their choices by executing a reaching movement by sliding their right arm on a graphics tablet (49.3 X 32.7 cm, Intuos 4XL; Wacom, Vancouver, WA) while holding a digitizing pen encased inside in a customized air hockey paddle. The tablet was placed below the monitor on the table and between an opaque platform that occluded the hand. The task comprised 400 trials with opportunity for self-paced breaks between trials. Trials in which the movement duration exceeded 1000 ms resulted in a "Too Slow" error message and were repeated.

To initiate each trial, the participant moved their arm to position a white cursor (diameter of 0.5 cm) inside the home position, indicated by a solid white circle at the center of the screen. After maintaining this position for 400 ms, the start circle turned green and three bandits appeared on the screen (positioned 8 cm radial distance from the center at 90°, 210° and 330° degrees relative to the origin). The bandits were colored light blue, dark blue, or purple and the color-position mappings were maintained for the entire experiment (randomized across participants).

Following the appearance of the bandits, participants had 1 second to make a rapid straight-line "shooting" movement through one of the bandits. Upon movement initiation, the cursor indicating hand position disappeared and did not reappear until feedback presentation. Participants were informed that there were three possible outcomes for each trial: If the movement was accurate (hand passed through the bandit) the cursor was displayed within the spatial boundary of the bandit. On these trials, there were two possible outcomes: (1) The bandit could turn green, indicating they would receive a reward for the trial (reward outcome), or (2) the bandit would turn red, indicating that, while the movement was accurate, no reward would be given on that trial (selection error). If the movement missed the bandit, a cursor would appear indicating the position when the hand was at the radial distance of the bandits, and thus indicate if the execution error was clockwise or counterclockwise relative to the target. The bandit would turn yellow, further signaling an execution error.

Following McDougle et al.(2019), each bandit had its own fixed probabilities for the three trial outcomes. All bandits had a 40% reward outcome, and thus, the expected value for

171 the three bandits were identical. However, the frequency of selection error and execution
172 error trials varied. For one bandit, 50% of the trials resulted in execution errors and 10%
173 resulted in selection errors. We refer to this as the "High Execution/Low Selection Error"
174 bandit. A second bandit resulted in execution errors on 10% of trials and 50% resulted in
175 selection errors (a "Low Execution/High Selection Error" bandit). A third, "Neutral" bandit
176 produced an equal number (30%) of execution and selection errors.

177 To enforce these probabilities, outcomes were surreptitiously manipulated so that they
178 aligned with predetermined feedback (a randomized sequence for each run) for the selected
179 bandit. On trials in which the actual movement produced the desired outcome (hit or miss the
180 bandit), the cursor was shown at its veridical position. However, if the participant's
181 movement missed the bandit, but the trial outcome was set as either a reward or selection
182 error (i.e., outcomes requiring successful motor execution), the end-point feedback showed
183 the cursor landing inside the bandit, albeit near the side consistent with the actual hand
184 position. Conversely, where a trial was set to be an execution error but the stylus successfully
185 intersected the bandit, the cursor was shifted just outside the bandit, with the side again
186 consistent with the actual hand position (e.g., if the hit was slightly clockwise to the center of
187 the bandit, the cursor appeared outside the spatial boundary of the bandit on the clockwise
188 side). For trials in which a change in feedback was required, the cursor position was shifted by
189 randomly sampling from a normal distribution (± 6.24°, equivalent to .5 cm with an 8 cm
190 reach) until a new cursor position was chosen that landed inside the bandit (for false hits) or
191 outside the bandit (false misses).

192 We included three further constraints to minimize the likelihood that participants
193 would recognize that the outcomes were not always directly reflective of their movements: (i)
194 No online movement feedback was available; (ii) end-point feedback was presented 1 second
195 after the stylus had passed the bandit location (this also helped reduce the impact of motor
196 artefacts contaminating the ERP); and (iii) if the actual reaching angle was greater than 10°
197 from the closest bandit on any trial (irrespective of the set outcome), no outcome was shown,
198 the experiment software instructed participants to "Please Reach Closer to the Bandit," and
199 the trial was repeated.

200 To increase motivation, participants were told that at the end of the experiment the
201 software would randomly select five trials, and based on the outcomes from these trials, a
202 cash bonus between £1-5 would be provided. As such, the goal was to accumulate as many
203 reward trials as possible. In actuality, all participants received a fixed payment of £10 for
204 taking part in the experiment.

The experimental task was programmed using the Psychophysics Toolbox (Brainard, 1997; Kleiner et al., 2007) and lasted approximately 35 minutes, with an additional 25-30 minutes of technical set up for EEG data acquisition.

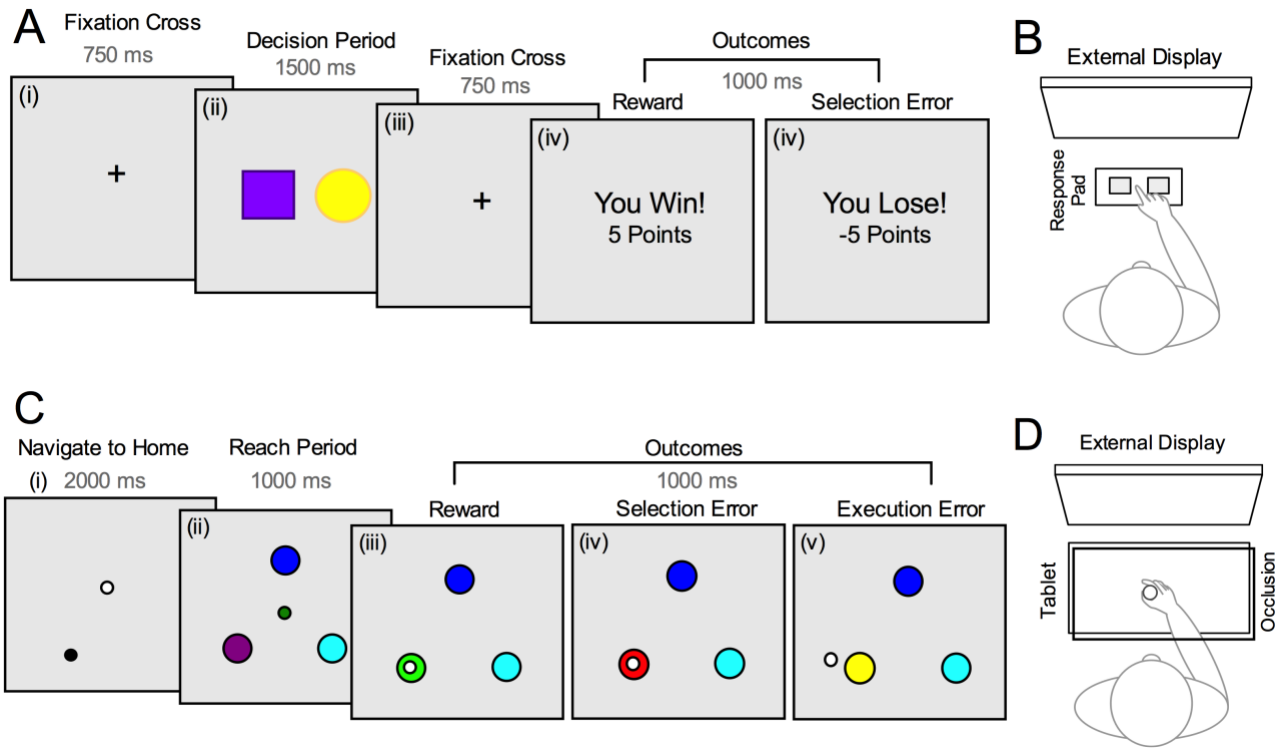


**Figure 1- Experimental Tasks:** (A) In Experiment 1, participants selected between two options, presented as squares and circles, colored yellow and purple, using a stimulus response pad (B). Outcomes were random, with feedback indicating if the participant won or lost points based on their selection on each trial. (C) In Experiment 2, participants moved a stylus on a tablet to make rapid shooting movements towards one of 3 bandits at 90°, 210° and 330° degrees relative to the home position. Following a 1000 ms delay (not pictured), pseudo-veridical feedback was provided indicating if the outcome was a reward (iii), a selection error (iv) or an execution error (v). (D) The hand was occluded throughout and stimuli were presented on a monitor positioned in front of the participants at approximately eye level.

## Electrophysiological data recording and preprocessing

### Experiment 1

The pooled EEG data available for re-analysis had been acquired with a 128-channel net connected to a high-input amplifier (Electrical Geodesics, Inc., Eugene, OR) at a rate of 500 Hz (0.01–200 Hz bandwidth), and an impedance ≤ 20 kΩ for frontocentral electrodes. The pre-processing routine for this pooled dataset has been reported previously (Mushtaq et al., 2016). These data had been recorded using a Cz reference online and digitally converted to an average mastoids reference offline. Following inspection of raw data, bad channels were replaced using a spherical spline interpolation method implemented in BESA 5.1 (MEGIS

Software GmbH, Gräfelfing, Germany). The data were filtered offline (0.1–30 Hz bandwidth) and segmented into epochs of 0–1000 ms, time-locked to the onset of feedback presentation (with an additional 200 ms pre-feedback baseline). Eye movement artifacts were corrected using a multiple source analysis method (Berg and Scherg, 1994, Ille et al., 2002) as implemented in BESA 5.1 ("surrogate method"). For each channel, epochs with a difference between the maximum and minimum amplitude > 120 μV or a maximum difference between two adjacent points > 75 μV were rejected (after eye movement artifact correction). All participants included in the analysis had a minimum of 16 artifact-free trials in each condition. The waveforms were baseline corrected using a 200 ms time window pre-feedback onset.

## Experiment 2

Whilst Experiment 1 involved the analysis of previously acquired data, Experiment 2 required the collection of a new data set, one in which the choices were made by arm movements to the selected bandit, opening up the possibility that non-rewards could arise from execution error in addition to selection error. For this experiment, EEG data were recorded continuously from 64 scalp locations at a sampling rate of 1024 Hz using a BioSemi Active-Two amplifier (BioSemi, Amsterdam). Four electrooculograms (EOG) – above and below the left eye, and at the outer canthi of each eye – were recorded to monitor eye movements. Two additional electrodes were placed at left and right mastoids. The CMS and DRL active electrodes placed close to the Cz electrode of the international 10-20 system served as reference and ground electrodes, respectively. EEG pre-processing was performed using the EEGLAB (Delorme and Makeig, 2004) and Fieldtrip (Oostenveld et al., 2011) toolboxes, combined with in-house procedures running using Matlab (The MathWorks, Inc., Natick, Massachusetts).

All data were first re-referenced offline to the average of all channels, and downsampled from 1024 Hz to 256 Hz. The continuous time series data were filtered using a high-pass filter with a cut-off at 0.1 Hz (Kaiser windowed-sinc FIR filter, beta = 5.653, transition bandwidth = .2 Hz, order = 4638) and a low-pass filter with a cut-off at 30 Hz (Kaiser windowed-sinc FIR, beta = 5.653, transition bandwidth = 10 Hz, order = 126). A second filtering of the data was performed for subsequent independent component analysis using a high-pass filter cut-off at 1 Hz (Kaiser windowed-sinc FIR filter, beta = 5.653, transition bandwidth = 2 Hz, order = 4666). ICA typically attains better decompositions on data with a 1 Hz high-pass filter (Winkler et al., 2015). The data were segmented into epochs

beginning 1s before and lasting 1s after the onset of feedback.

Infomax ICA, as implemented in the EEGLAB toolbox, was run on the 1 Hz high-pass-filter epoched data, and the resulting component weights were copied to the .1 Hz high-pass-filter epoched data. All subsequent steps were conducted on the .1 Hz high-pass-filtered data. Potentially artefactual components were selected automatically using SASICA (Chaumon et al., 2015), based on low autocorrelation, high channel specificity, and high correlation with the vertical and horizontal eye channels. The selections were visually inspected for verification purposes and adjusted when necessary. After removal of artefactual components, the Fully Automated Statistical Thresholding for EEG Artefact Rejection plugin for EEGLAB (Nolan et al., 2010) was used for general artefact rejection and interpolation of globally and locally artefact contaminated channels, supplemented by visual inspection for further periods of non-standard data, such as voltage jumps, blinks, and muscle noise.

There was an unequal number of trials per outcome type given that the outcome probabilities differed between the three bandits. From all artifact-free epochs (93.5% of the total sample) more reward trials ($\mu = 150, \pm 9$) were available for analysis relative to execution error trials ($\mu = 114, \pm 12$; $p < .001$) and selection error trials ($\mu = 110, \pm 11$; $p < .001$). To increase the reliability of our conclusions by addressing potential problems of distribution abnormalities and outliers, grand average waveforms were constructed for each individual by taking the bootstrapped ($n = 1{,}000$) means from the EEG time series epochs. In practice, there was little difference in this bootstrapped analysis and the grand averages constructed from the raw data. The waveforms were baseline corrected using a 200 ms time window pre-feedback onset.

**ERP Quantification**

Given that we had specific hypotheses, we did not perform a spatiotemporal mass univariate analysis for all electrodes and time points across the scalp (c.f. Groppe et al., 2011). Instead, the parameters were constrained by focusing on two locations. First, given that meta-analyses (Walsh and Anderson, 2012; Sambrook and Goslin, 2015) have shown the feedback-locked MFN effect to be maximal over the frontocentral region of the scalp, we averaged activity across electrodes e6, e7 and e106 in Experiment 1 and electrodes FC1, FCz, and FC2 in Experiment 2. Second, given that the feedback-P3 (specifically, the P3b sub-component) is commonly present in feedback-locked ERPs and typically maximal over parietal electrodes (Polich, 2007), we averaged over electrodes e62, e61 and e78 in Experiment 1, and P1, Pz, and P2 in Experiment 2. We opted to average across electrodes to

296 improve the signal-to-noise ratio of the ERP measures (Oken and Chiappa, 1986). We also
297 performed the analyses on the individual electrodes for comparison to ensure that our results
298 were not biased by the specific configurations of electrodes in the cluster. The patterns of
299 results for clustered and individual electrode analysis were essentially indistinguishable; thus,
300 only the averaged electrode data are reported here.

301 Distinguishing between the outcome-locked MFN and components that precede (P2)
302 or follow (a large P3 component comprising a frontal P3a and parietal P3b) the MFN is a
303 challenge due to spatial and temporal overlap (Glazer et al., 2018). Difference waveforms are
304 useful in such situations (Kappenman and Luck, 2017) and indeed, the majority of research on
305 the MFN has typically computed "reward prediction error" (RPE) waveforms, derived by
306 subtracting error/loss trials from reward trials (Sambrook and Goslin, 2015). Here, we create
307 an equivalent "Selection Prediction Error" difference waveform, subtracting the average
308 activity associated with Selection Error trials from the average activity related to all Reward
309 trials. These difference waveforms were subjected to a series of one-sample t tests to identify
310 where this signal was different to zero.

311 In Experiment 2, pairwise comparisons were also visualized through difference
312 waveforms: In addition to the "Selection Prediction Error" difference waveform described
313 above, we also computed an "Execution Prediction Error" waveform by subtracting the
314 average activity associated with Execution Error trials from the average activity associated
315 with Reward trials. We directly contrasted Execution and Selection Error ERPs by subtracting
316 the Execution Error waveform from the Selection Error waveform to create an "Error
317 Sensitivity" difference waveform. Outcome trials were subjected to a one-way ANOVA and
318 where main effects emerged, one-sample t tests were conducted to identify where the
319 difference waveforms were significantly different to zero.

320 To reduce the number of false positives (Luck and Gaspelin, 2017), the ERP data were
321 downsampled to 250 Hz and we only analyzed activity between 150 and 500 ms (spanning
322 the P2, MFN and P3 ERPs). For each analysis, *p* values were corrected by applying a false
323 discovery rate (FDR) control algorithm (Benjamini and Hochberg, 1995; Lage-Castellanos et
324 al., 2010). The Benjamin-Hochberg correction approach was adopted as previous studies have
325 shown it to reliably control the FDR when data are correlated, even when the number of
326 comparisons are relatively small (Hemmelmann et al., 2005). This method is also ideally
327 suited for the exploration of focally distributed effects (Groppe et al., 2011).

It is not possible to identify whether increases or decreases in the parent waveforms drive amplitude changes in the observed difference waveform. To aid the interpretation of the difference waveforms, we also visualized the grand averaged ERPs related to each outcome. For each significant effect, we show the average amplitude for individual conditions in the frontal and parietal clusters.

Differences between relevant conditions at each electrode site are visualized through topographical maps to support interpretation of underlying components: Predicated on previous research (Walsh and Anderson, 2012), we anticipated that the MFN should show a frontocentral topography and, following an early frontocentral peak, there would be a subsequent posterior maximum corresponding to the P3b (Holroyd and Krigolson, 2007).

## Brain-Behavior Relationships

A key question in this study is whether the electrophysiological signatures of outcome processing is predictive of the participants' choice behavior (reviewed in San Martín, 2012). Based on a reinforcement learning account of the MFN (Holroyd and Coles, 2002), we would expect the amplitude of the MFN to correlate with the degree of behavioral adjustment: Large differences in the MFN should be more likely to lead to changes in choice behavior compared to small differences in the MFN.

To establish a firm grounding for studying the relationship between error type and reinforcement learning processes indexed by the MFN, we used the pooled data from the classic bandit task (no execution errors) in Experiment 1 to explore brain-behavior correlations. At each time-point, the relationship between the amplitude of the "Selection Prediction Error" waveform and subsequent choice adjustment was examined. The behavioral adjustment score, or switch bias rate, was operationalized as the ratio of the percentage of trials that the participant switched following an error to the percentage of switching following a reward. Results derived from this analysis were used to inform the approach in Experiment 2, where we took statistically significant clusters of activity and examined whether the average activity in these clusters correlated with behavioral responses to the two error types.

## Statistical Analysis

For reporting purposes, time points are rounded to the nearest millisecond, amplitude (in microvolts; μV) to two decimal places and *p* values to three decimal places. The range for the scalp maps was time-interval specific and determined by the 1st and 99th percentile values across all electrodes. Spearman's rho ($r_s$) was used to examine correlations between amplitude and behavior. For correlations between behavior and neural activity, we extracted peak and mean amplitudes, report both results and visualize the strongest correlations. Where appropriate, pairs of correlations were directly compared with Hittner, May, and Silver's (2003) modification of Dunn and Clark's (1969) approach using a back-transformed average Fisher's Z procedure, as implemented in the R package Cocor v. 1.1-3 (Diedenhofen and Musch, 2015). The statistical significance threshold was set at $p < .05$. For one-way ANOVAs, we report generalized eta squared ($\eta_{G2}$) as a measure of effect size. This measure was selected over eta squared and partial eta squared because it provides comparability across between- and within-subjects designs (Olejnik and Algina, 2003; Bakeman, 2005); we considered $\eta_{G2} = 0.02$ to be small, $\eta_{G2} = 0.13$ medium and $\eta_{G2} = 0.26$ to be a large effect size. All statistical analyses were performed using R (R Core Team, 2015).

# Results

## Experiment 1

In Experiment 1, we performed new analyses on archival EEG data, pooling together data sets from one published study (Mushtaq et al., 2016) and one unpublished study. The procedure was the same in both, with participants indicating their choice between two options by making a button press, followed by feedback indicating the success or failure of that choice. The feedback associated was based on a random reward schedule (50/50 outcomes), and thus no optimal response strategy existed. Exploratory analysis of the independent datasets showed qualitatively similar relationships and thus, for purposes of statistical power, we report only pooled analysis.

In terms of behavior, the switch rate following an unrewarded outcome was not different than following a reward (**Figure 2** A; t(47)= 0.12, p = .898). Despite showing similar behavioral effects, the neural signals were dependent on the outcome type. The grand averaged ERP signal showed a negative medial frontal deflection (MFN) that peaked at approximately 270 ms, and, importantly, was larger following feedback indicating no reward compared to when the feedback indicated reward (**Figure 2** B & C). The MFN was preceded and succeeded by positive peaks, assumed to reflect P2 and P3 responses, respectively, with the latter component showing an initial response dominant in frontocentral clusters (P3a) and latter response dominant over parietal clusters (P3b). Timepoint-by-timepoint analysis of the Selection Prediction Error difference waveform (**Figure 3**A & B) revealed significantly differences from 262 to 500 ms in the frontocentral electrodes, and from 238 to 500 ms in the parietal electrodes.

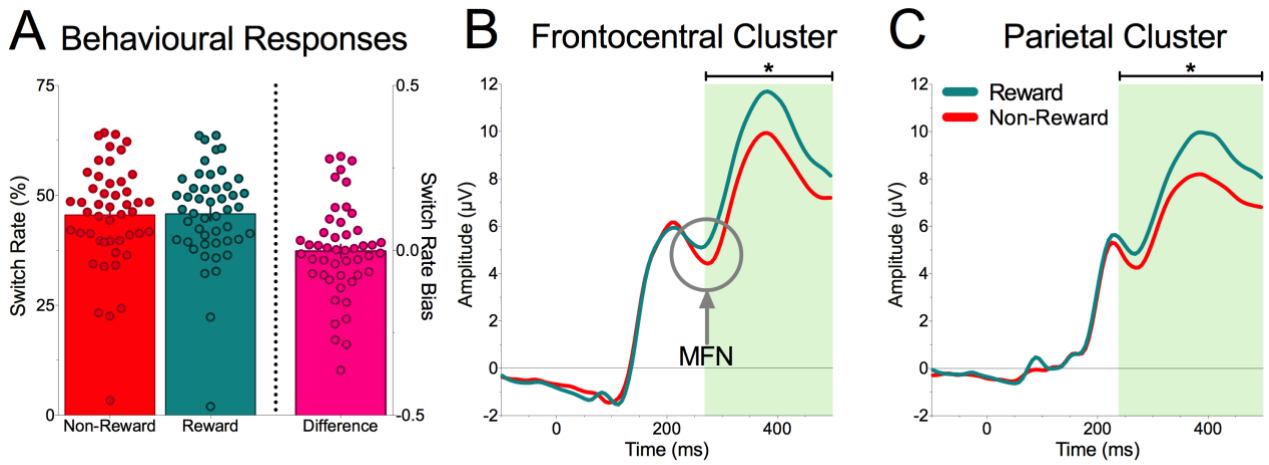

**Figure 2 Experiment 1 Behavioral and ERP Results:** (A) Participants showed similar rates of switching following selection error and reward outcomes. Each dot represents the percentage an individual switched to a different bandit following each outcome. The purple bar shows the switch rate bias (right ordinate). Positive values indicate more switching following unrewarded trials (selection errors) relative to rewarded trials; negative values indicate more switching following rewarded trials. (B) Grand averaged ERPs at frontocentral and parietal region in response to Reward (dark green) and Selection Error (red) outcomes. The green shaded region indicates the cluster of time points that show significant differences.

The time windows over which the ERPs to the two types of feedback were statistically different are much longer than the expected temporal duration of the MFN (~100 ms); for example, a recent meta-analysis reported that the FRN generally lasts between 228 – 334 ms post-feedback (Sambrook and Goslin, 2015). The extended period observed here may reflect overlap with a difference in the P3 for the two outcomes. Although the P3 is most dominant in parietal sites (i.e., a sub-component, or P3b, observed ~300-500 ms post-feedback), there is also a component visible in frontocentral regions (referred to as the P3a). The temporal proximity of this large positive deflection can contaminate the preceding signal (Krigolson, 2017). Observation of the scalp topographies confirmed this MFN-P3 transition and overlap: The temporal evolution of amplitude differences (visualized in steps of 50 ms from 250 ms post-feedback in **Figure 3** C) across this time-range showed an early frontocentral maximum, indicative of a MFN followed by an increasingly posterior maximum, consistent with the P3b. This shift from frontocentral to posterior sites is said to reflect a transition in overlapping, but separable, outcome monitoring processes, moving from the rapid evaluation of the outcome to a context-specific updating process (von Borries et al., 2010).

Having identified when and where the ERP signal for rewards and selection errors could be reliably differentiated, we next asked whether individual differences in behavioral adjustment was reflected in the neural responses to these outcomes. This variation was examined by calculating the difference between the percentage of trials for which participants switched bandits following selection errors relative to the percentage of trials for which participants switched following reward. As such, this behavioral adjustment score, or 'switch rate bias', is bounded between -1 to +1. There were no group level differences between switch rates following selection errors relative to rewarded trials (M = 0.002, SD=0.14), but the switch rate bias scores were quite variable, ranging from -0.36 to +0.28 (Figure 2a, right bar). Positive values indicated they switched more often following errors relative to rewards, and negative values indicated participants switched more often following a reward relative to an error.

We asked whether, despite the absence of group level differences, individual variations in this switch rate bias may be linked to the difference in MFN amplitude following rewarded and selection error trials. We predicted that participants with a higher tendency to switch following errors would show larger differences in the Selection Prediction Error waveform. Specifically, we expected that these differences would be reflected in a time-interval corresponding to the MFN, with larger negative deflections following errors correlated with an increased tendency to switch choices (i.e., a positive Switch Rate Bias score) and thus show a negative correlation. To test this, the difference waveform amplitude at each time-point was correlated with the Switch Rate Bias.

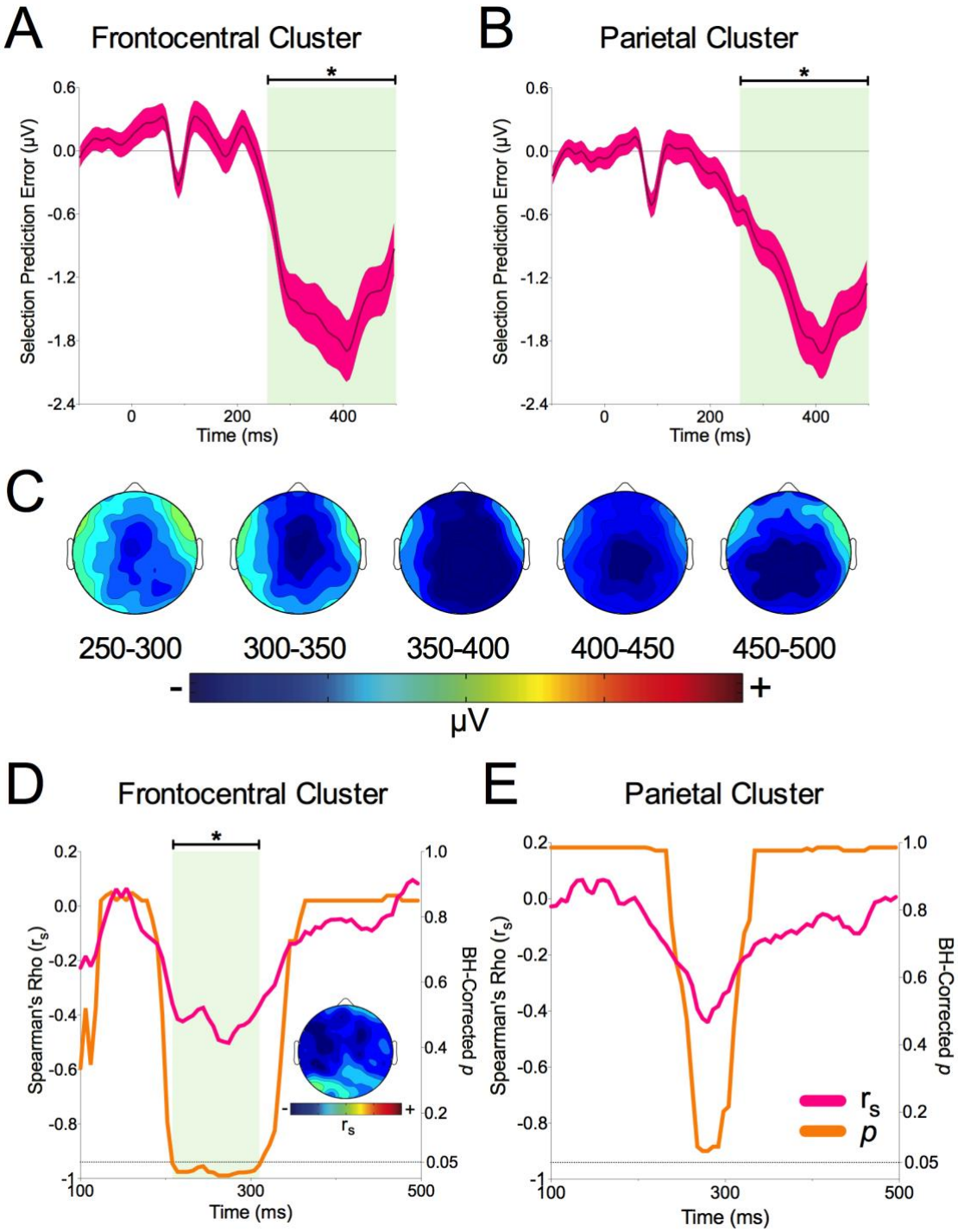


**Figure 3- Experiment 1 Brain-Behavior Correlations:** 'Selection Prediction Error' difference waveforms for the (A) frontocentral and (B) parietal cluster. (C) Temporal evolution of this difference revealed a shift from a frontocentral maximum to a parietal distribution. (D, E) Correlation between difference waveform and behavioral adjustment, calculated at each time (pink), with p value (orange) for each time point, corrected for multiple comparisons. Frontocentral cluster is shown in D, along with scalp map showing distribution of the correlation across the scalp for the mean amplitude in this time window. Parietal cluster is presented in E. Green shaded region indicate clusters of time points showing significant differences (Benjamin-Hochberg corrected $p < 0.05$).

In the frontocentral cluster, the Selection Prediction Error difference waveform correlated with behavioral sensitivity to error between 208 and 310 ms (**Figure 3** D). That is, greater negativity in response to no reward relative to reward, was associated with an increased bias to switch following errors. Parietal activity showed a qualitatively similar pattern, but the effect did not reach significance (**Figure 3** E). The scalp maps corroborated this observation, with negative correlations present at most sites, but maximal in frontocentral regions. Interestingly, the ERP-behavior correlations did not persist in the P3 time-range, consistent with the claim that the earliest portion of the difference waveform captures RPE-related processes (Sambrook and Goslin, 2015).

To directly compare the correlations for the MFN and P3a, and recapitulate the correlations in a more traditional fashion, the statistically significant cluster of differences between reward and selection error was parsed into two time windows for the MFN and P3a based on the onset of the statistical significant cluster and visual inspection of the waveform: The mean average of the activity in the time-interval from 262 to 310 ms was used to define the MFN, and the mean average of the remainder of the activity was defined as P3a (**Figure 4** A). Correlations were then computed between behavior and the mean and peak amplitude of each ERP component. We observed the same pattern of results using this more traditional approach. There was a negative correlation in the MFN time window: Participants with larger MFN amplitudes were also more likely to switch following non-rewards (mean: $r = -.46$, $p < .001$; peak: -.38, $p = .008$) (**Figure 4** B). However, there no reliable correlation for the P3a window (**Figure 4** C; $r = -.10$, $p = .482$) and these two correlations were significantly different ($z = 3.39$, $p < .001$). We again note that these point-estimate results simply re-capitulate the time-based correlations in Figure 3.

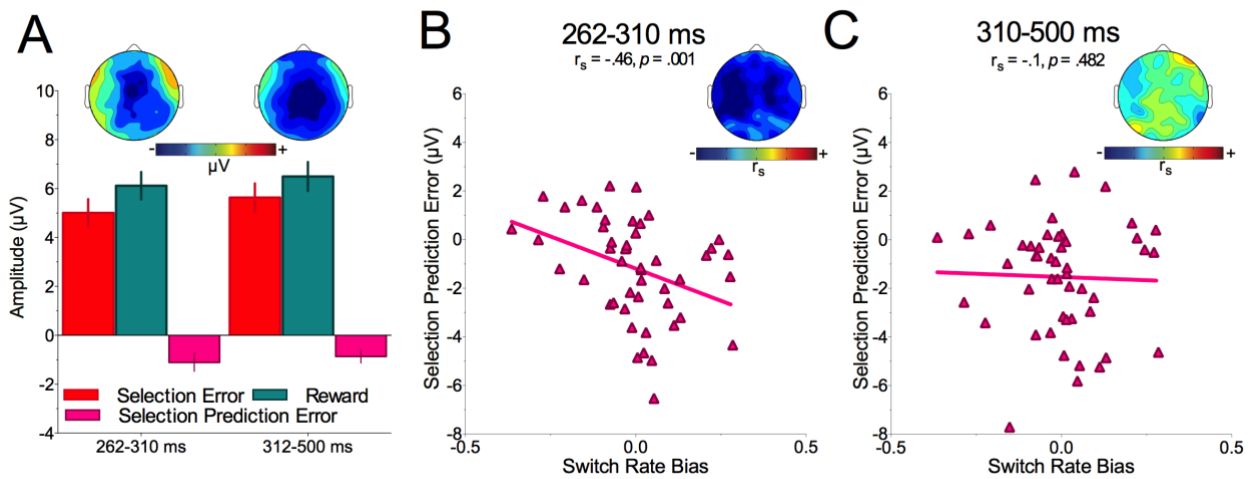

**Figure 4- Experiment 1 MFN and P3 Amplitude and Correlations.** (A) Mean amplitudes for selection errors (red) and reward (green) and their difference (purple) in the early (MFN) and late (P3a) phases of green-shaded cluster shown in Figure 3D (inset scalp maps show the topographical distribution of mean amplitude differences in these clusters). Error bars represent ±1 SEM. Correlations between these differences and the risk bias score reveal that greater negativity in the difference waveform was associated with increased behavioral adjustment following error at a time interval corresponding to the MFN (B), but not the P3a (C). The inset scalp maps show the magnitude of the correlations across electrode sites.

Taken together, the correlation analyses show a relationship between choice biases and MFN point estimates, but with this relationship restricted to a particular time window and limited to the frontocentral cluster. We use this observation as a basis for examining the impact of execution errors on reinforcement learning in Experiment 2.

## Experiment 2: Behavioral Responses

Experiment 2 introduced the possibility that unrewarded actions might arise from a failure in movement execution rather than an error in action selection. To this end, we employed a 3-arm bandit task, with each bandit having the same probability of a rewarded outcome, but different ratios of execution and selection errors. We examined participants' choice biases between these three bandits and whether selection and execution errors would elicit different patterns of behavioral adjustment.

A one-way ANOVA revealed a significant difference in selection preference ($F [2, 56] = 8.27$, $p < .001$, $\eta_{2g} = .228$), with participants exhibiting a preference for the High Execution/Low Selection Error bandit. Overall, this bandit was chosen on 39.49% (±9.38) of the trials, which was significantly greater than the Low Execution/High Selection error bandit (29.01%; ±6.89%, $p < .001$) and Neutral bandit (31.50%; ±8.64%, $p = .046$), with no difference for the latter two ($p = .877$). Consistent with previous work, when expected value is equal, participants prefer choices in which unrewarded trials are attributed to errors in movement execution rather than errors in action selection (Wu et al., 2009; Green et al., 2010; McDougle et al., 2016; Parvin et al., 2018; McDougle et al., 2019).

As in Experiment 1, we examined the effect of the different outcomes on the next choice, asking how they influenced switching behavior (**Figure 5** A). Participants exhibited high switching rates overall (53.71%; ±24.1), but the rate differed according to outcome type ($F [2, 56] = 10.23$, $p < .001$, $\eta_{2g} = .11$). Switching was highest following selection errors (66.29%; ±27.85) and markedly lower following execution errors (41.87%, ±25.50%, $p < .001$). This difference is consistent with the hypothesis that motor errors attenuate value updating, perhaps because participants believe they have more control to correct for execution errors (Parvin et al., 2018).

Interestingly, switching rates following rewarded trials fell between the other two outcome types (55.12%, ±32.13%). This value was not significantly different that than observed following selection errors ($p = .227$) or execution errors, although the latter approached significance ($p = .062$, following Bonferroni correction). The fact that many participants (18 of 29) were so prone to switch after a rewarded outcome and even more so (numerically) than after an execution error was unexpected. This high switching rates would suggest a bias towards exploratory behavior in this task (Gittins, 1979), something that might be promoted by the relatively low reward rates and/or highly probabilistic nature of the outcomes (Daw et al., 2006; Cohen et al., 2007). This bias might have been offset when a movement failed to reveal information about a chosen bandit. Notably, there were very large individual differences in the treatment of the outcomes: Switch rates ranged from 3% to 98% following rewards, 7%-99% following selection errors and 4%-81% following execution errors.

## Experiment 2: ERP Responses

Our primary aim was to examine whether selection and execution errors could be reliably distinguished in outcome-locked ERPs. To start, we ran an exploratory 3 (Bandit Type) X 3 (Outcome) ANOVA at each time point for the frontocentral and parietal clusters. The main effect of Bandit Type was not significant (corrected $p$ values => .702) and there was no Bandit Type X Outcome interaction (corrected $p$ values => .671). Thus, we collapsed across the three bandits in our primary analyses of the three outcomes, allowing us to avoid increasing the family-wise error rate.

The grand averaged ERPs related to each outcome are shown in **Figure 5** B & C. F tests revealed two significant clusters in the frontocentral region between 156 -180 ms and 210-336 ms, and three clusters in the parietal region (176-196 ms; 218-239 ms; and 355-438 ms). Descriptively, the first cluster in the frontocentral region was driven by a delay in the onset of an initial P200-like signal following an execution error, and the second cluster incorporated MFN deflections following selection and execution errors, along with subsequent positive P3a deflections. The early two clusters in the parietal region (**Figure 5** C) reflect shifts in the latency and amplitude of the execution error ERP, with the third cluster driven by the attenuation of the P3b response following selection errors.

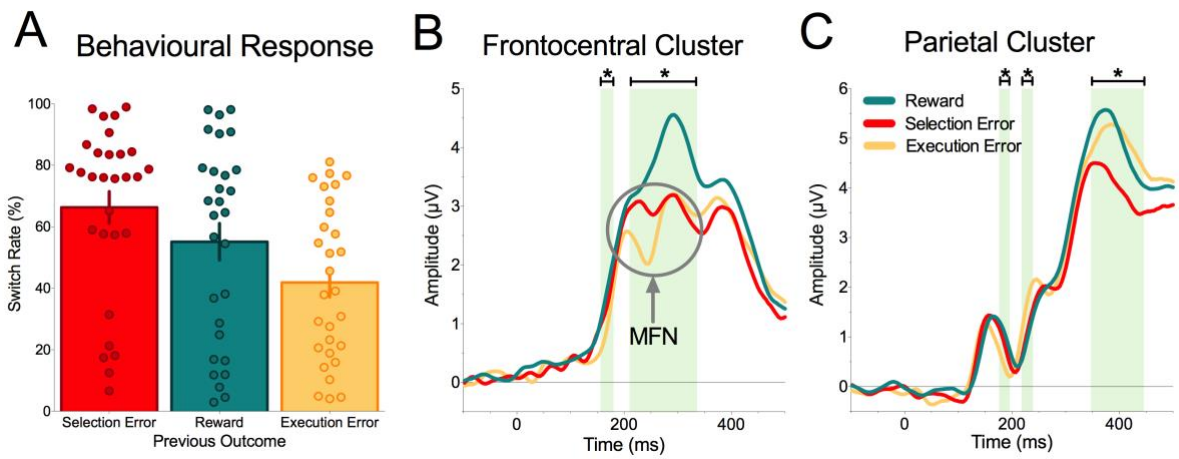


**Figure 5- Experiment 2 Behavioral and ERP Results.** (A) Switching rates following the three trial outcomes. Participants were more likely to repeat a choice following execution errors. Error bars represent ±1 SEM. Feedback-locked ERPs for each outcome type, recorded from (B) frontocentral and (C) parietal electrode clusters. The medial frontal negative deflections observed for the two errors conditions in the frontocentral cluster are highlighted. The shaded regions indicate the statistically significant time clusters identified by the mass univariate analysis.

Figure 6a depicts the Selection Prediction Error difference waveform from the frontocentral cluster, which closely resembled the results observed in Experiment 1. This waveform was significantly different from 0 between 242-336 ms (one-sample t-tests). An examination of the scalp topography of the first (242-289 ms) and second half of this window (289- 336 ms) indicated a clear frontocentral maximum in the early phase, followed by a shift towards centroparietal maximum in the later part of the window (Figure 6B), replicating Experiment 1. The magnitude and temporal duration of the P3a was not as pronounced as in the first study, an effect that may be related to the substantial methodological differences between the two experiments (e.g., feedback format, presence or absence of magnitude information).

We also replicated the pattern observed in Experiment 1 (see **Figure 4** B) for the relationship between neural activity and behavior. Specifically, amplitude (mean: $r_s = -.43$, $p = .021$; peak : $r = -0.36$, $p = .052$; **Figure 6** C) from the early part of the cluster negatively correlated with behavioral adjustment: The larger the difference waveform (i.e., greater negative deflection for selection errors relative to rewards), the greater the bias for the participant to switch in their choice following a selection error outcome relative to a reward outcome. We note that one participant had a switch rate score of -0.87, which was 2.97 standard deviations away from the mean. Re-running the analysis without this participant showed a weaker relationship, but the pattern remained (mean $r_s = -.39$, $p = .042$; peak $r_s = -.34$, $p = .074$).

The topographical map (**Figure 6** C inset) demonstrates that this effect was localized to the frontocentral region. We again found no evidence for such a relationship in the later part of the time window ($r = -.11$, $p = .567$; **Figure 6** D). The mean MFN and P3a correlations were significantly different from one another ($z = 1.96$, $p = .05$), providing further support that the MFN, but not the P3, is a reliable predictor of behavioral changes.

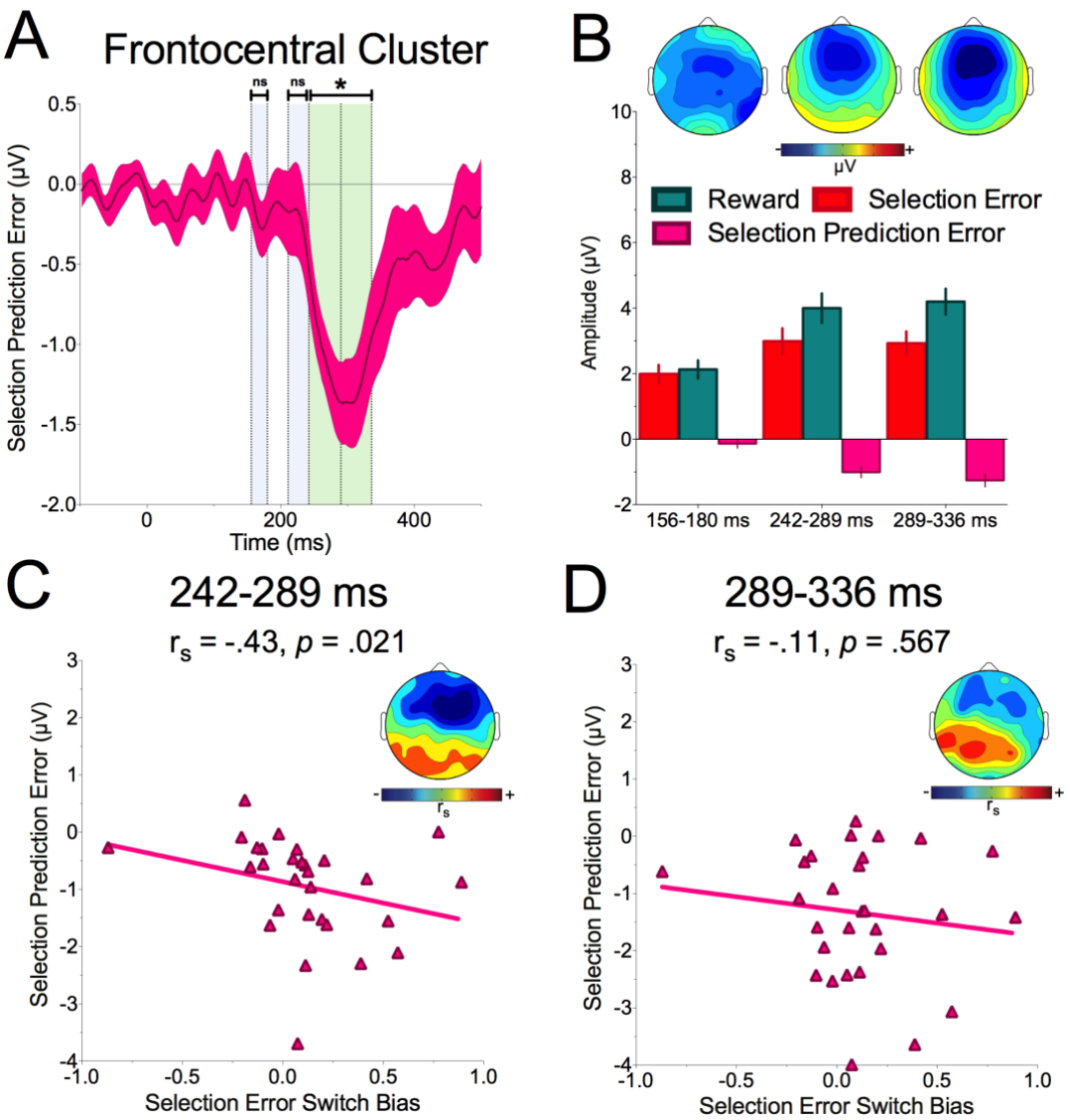


**Figure 6- Selection Prediction Error in the Frontocentral Cluster:** (A) The Selection Prediction Error difference waveform. The green shaded regions indicate the clusters showing statistically significant differences for this contrast and the grey shaded regions indicate where the clusters identified in the original time-series analysis did not reach statistical significance for this difference waveform. (B) Mean amplitudes for the early and late phases of the statistically significant clusters, with inset scalp maps showing the distribution of differences across sites for each time interval. This difference waveform correlated with an increase in behavioral adjustment following selection error at a time interval corresponding to the MFN (C), but not the P3 (D). The inset scalp maps showing the distribution of amplitude differences across sites each time interval reveal a frontocentral maxima for the MFN correlation.

Activity in the parietal region also mirrored the results from Experiment 1: While there was a clear posterior distribution, indicating the presence of a feedback related P3, this was not correlated with behavior (r = .11, p = .576).

## Execution Prediction Errors

To examine the electrophysiological correlates associated with unrewarded outcomes attributed to motor execution errors, we performed similar analyses, but now focus on the comparison between execution error trials and reward trials (the Execution Prediction Error difference waveform). This comparison revealed two statistically significant clusters- one ranging from 156-180 ms and a second between 207-325 ms (**Figure 7** A).

The first cluster showed an amplitude reduction in response to Execution Errors relative to reward trials. Similar to what we observed for the Selection Prediction Error difference waveform, we expected the second cluster would be contaminated by a P3a signal. Thus, we followed the same protocol, splitting this cluster into two equal intervals – (i) an early phase marked by the time interval 207-266 ms; and (ii) a later phase for activity between 266-325 ms. There was a clear frontocentral distribution for the early phase, and in the later time window there was a shift towards centroparietal electrodes (**Figure 7** B).

We next examined the relationship between these three epochs (156-180 ms; 207-266 ms; 266-325 ms) and behavioral adjustment (**Figure 7** C-E). The peak amplitude difference in the earliest interval (156-180 ms) correlated positively ($r_s = 0.37$, $p = .05$) with switching rates following an execution error relative to reward. This pattern is opposite to that observed between the amplitude of the MFN and behavioral adjustments following selection prediction errors. Following execution errors, smaller peaks in the 156-180 ms time window were associated with a lower tendency to switch. The mean amplitude showed a similar pattern of results, but was not significant ($r_s = 0.35$, $p = .065$). An examination of topography revealed this correlation to be maximal in the frontocentral cluster. These correlations suggest that an early attenuation of amplitude in the MFN after execution errors predicts how tolerant a participant is to those outcomes. That is, smaller amplitudes in the processing of Execution Errors early on in the feedback processing stream are associated with a higher tolerance to Execution Errors.

In contrast to the Selection Prediction Error results, the MFN captured in the 207-266 ms time window did not correlate with behavioral adjustment ($r_s = .07$, $p = .722$). We tested, and confirmed, that this correlation was reliably different to the correlation observed for Selection Prediction Errors in the MFN time interval ($z = 2.40$ $p = .016$). In the P3a, we expected and found no relationship this time window (266-325 ms) with behavioral adjustment ($r_s = -.22$, $p = .258$).

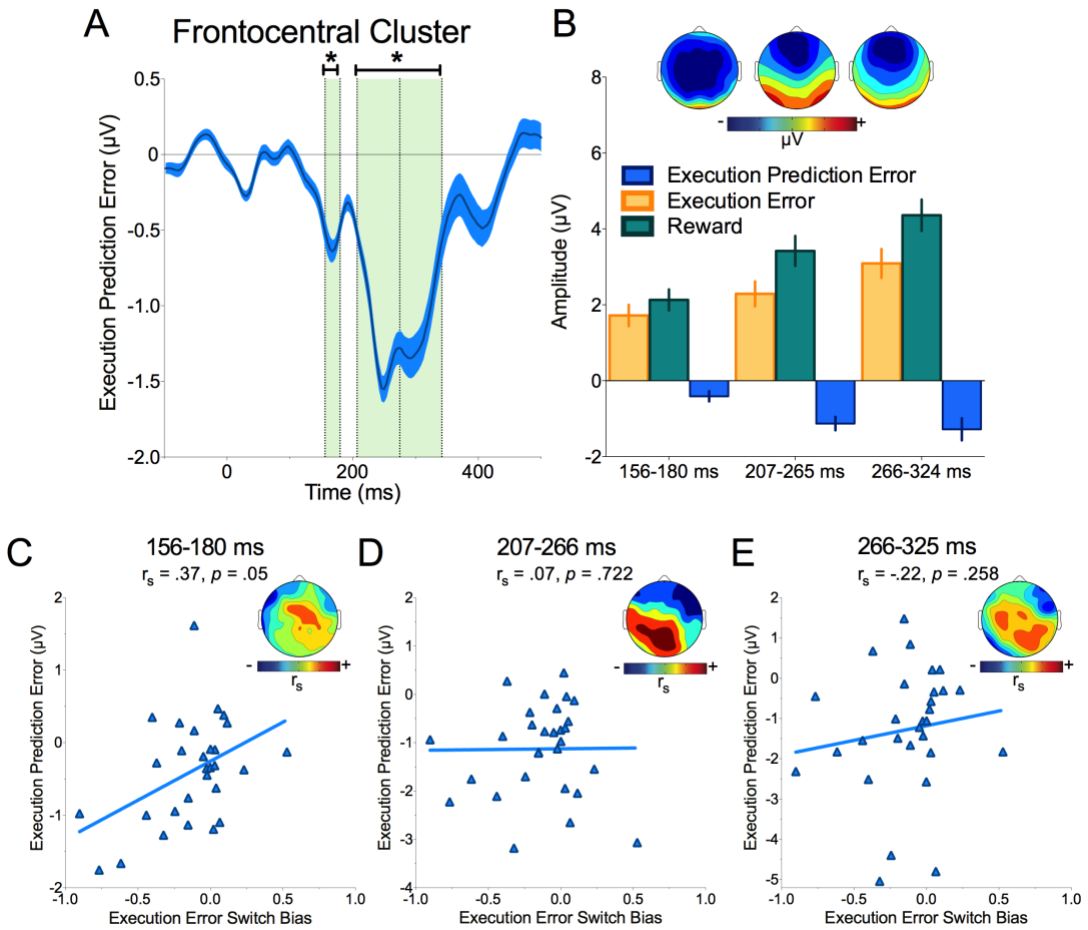


**Figure 7- Execution Prediction Error in the Frontocentral Cluster:** (A) The Execution Prediction Error difference. The green shaded regions indicate clusters showing statistically significant differences. (B) Mean amplitudes for the early and late phases of the statistically significant clusters. (C-E) Amplitude differences were positively correlated with tolerance to execution errors only in the early epoch.

We conducted the same analysis for the Execution Prediction Error difference waveform in the parietal cluster of electrodes. Execution Errors elicited smaller amplitude responses relative to Rewards in an early time window (176-196 ms) but showed larger amplitude at 218-239 ms post feedback. There was a positive correlation between amplitude and behavior ($r_s$ = .47, p = .01) in the posterior region in this later time window, indicating the processes driving behavioural adjustment involve a shift from frontocentral to parietal regions (Overbeek et al., 2005; Dhar and Pourtois, 2011). Interestingly, and unexpected, was the finding that in the later time window, the amplitude of the P3b-- often implicated in processing reward magnitude (Yeung and Sanfey, 2004; San Martín, 2012)-- showed no difference in the processing of Execution Errors and Rewards (see **Figure 5** C) and there was no relationship with behavioral adjustment ($r_s$ = -0.01, p = .946).

## Error Sensitivity Difference Waveform

As described in the previous two sections, when using a common baseline (rewarded trials), we observed differences in both the ERP results and correlational analysis between unrewarded trials that were attributed to failures in either movement execution or action selection. We performed a direct comparison between these two types of unrewarded outcomes through the "Error Sensitivity" difference waveform.

In the frontocentral cluster there was a significant difference in the range of the MFN (222-250 ms; **Figure 8** A & B). We had anticipated that the amplitude of the MFN would be attenuated to execution errors, assuming a lower response would be reflective of reduced value updating (McDougle et al., 2019). However, the observed effect was in the opposite direction: Execution Errors elicited a larger MFN deflection, relative to Selection Errors.

Finally, we examined whether the magnitude of this difference correlated with switching biases. We note that although the parent waveforms for this correlation are included in the previous analyses, the EEG activity being correlated with switch bias rate in this analysis is specific to the time range (220-250 ms), the window in which these error outcome signals significantly differed.

In the behavioral data, no participants in our sample showed more switching following a motor execution error relative to a selection error. Some participants treated the unrewarded outcomes equivalently (i.e., no difference in switch rates) whilst others showed higher switch rates following selection errors. Could these differences be linked to the magnitude of the MFN difference in this time window to these two types of unrewarded trials?

There was no relationship between mean amplitude in the 220-250 ms time window and switch bias ($r = .23$, $p = .23$). However, the peak negative amplitude revealed a positive correlation with error sensitivity switch bias ($r = .41$, $p = .026$; **Figure 8** C). Participants who showed smaller behavioral biases also showed smaller MFN differences while individuals with a greater tolerance towards motor execution errors also exhibited larger MFN amplitudes for motor execution errors relative to selection errors. This correlation was maximal in frontocentral sites (**Figure 8** C inset).

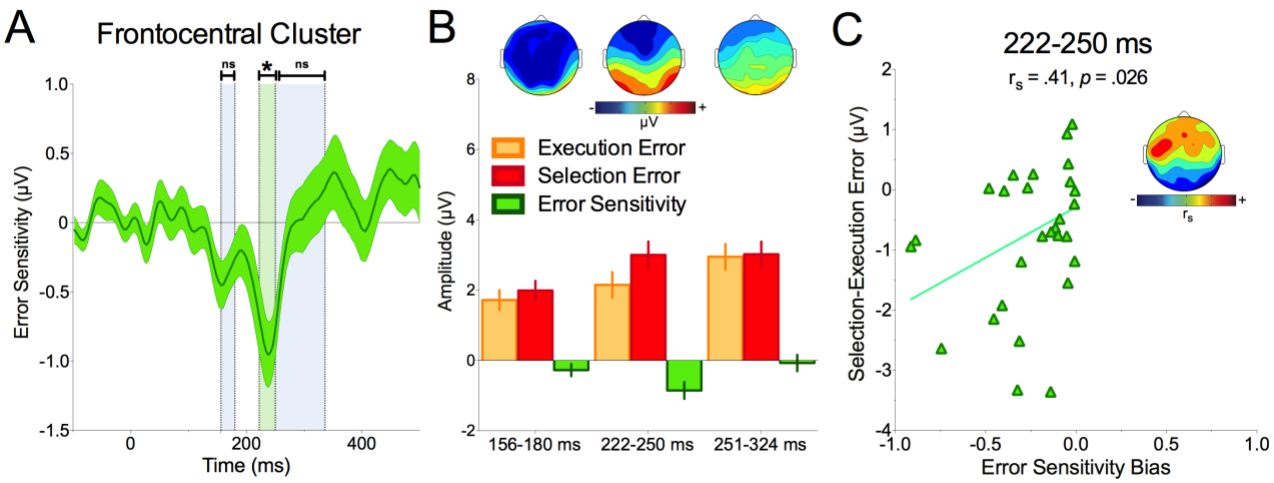


**Figure 8- Error Sensitivity in the Frontocentral Cluster:** (A) The Error Sensitivity Difference Waveform. The green shaded region indicates the cluster showing statistically significant differences for this contrast and the grey shaded regions indicate where the clusters identified in the original time-series analysis did not reach statistical significance for this difference waveform. (B) Mean amplitudes for the early and late clusters indicated by shaded regions in panel A. Inset scalp maps show topographical distribution for each cluster. (C) Peak amplitude difference in the MFN correlated with switching rate following a selection error relative to an execution error. Scalp distribution indicates that this correlation was primarily distributed in frontocentral sites.

Examining the parietal cluster revealed no differences in the earliest interval (176-196 ms), but there were differences emerged in the 218-239 ms and 359-445 ms epochs, with larger positive amplitudes for execution errors relative to selection errors. The mean amplitude across each of these statistically significant clusters (218-239 ms and 359-445 ms) was not correlated with the behavioral adjustment scores.

# Discussion

Adaptive behavior necessitates distinguishing between outcomes that fail to produce an expected reward due to either selection of the wrong action plan or poor motor execution. Although the majority of neuroscience decision-making research has focused almost exclusively on the former, a number of studies have shown that failed outcomes attributed to sensorimotor errors can markedly alter choice preferences (Green et al., 2010; McDougle et al., 2016; Parvin et al., 2018; McDougle et al., 2019). It has been postulated that these shifts arise from error signals credited to the sensorimotor system attenuating the operation of reinforcement learning processes (McDougle et al., 2016). Here, we examined whether a putative ERP signature of reinforcement learning, the medial frontal negativity (MFN), varied in response to selection and motor errors.

## Differential Error Processing indexed by the MFN

Consistent with previous work, selection errors elicited a larger MFN relative to reward outcomes (Gehring and Willoughby, 2002; Holroyd and Coles, 2002). Moreover, the amplitude of the MFN following selection errors was negatively correlated with the participants likelihood of changing their choice behavior (Holroyd & Coles, 2002, but see San Martín, 2012). We observed this relationship across two experiments that entailed distinct modes for responding and feedback presentation, along with different analysis pipelines.

Behaviorally, participants showed lower switch rates following execution errors, a pattern consistent with the hypothesis that the reinforcement learning system discounts these errors (McDougle et al., 2016; 2019). Contrary to our prediction that the MFN would be attenuated following execution errors, these errors produced the largest MFNs. However, a striking difference between the Selection and Execution Prediction Error difference waveforms was that the amplitude of the MFN following selection errors was predictive of behavioral biases, whereas this ERP measure following execution errors was not predictive of behavioral biases. These results indicate that these two types of feedback signals, both indicating the absence of a reward, are processed differentially.

While almost all participants were more likely to switch after a selection error compared to an execution error, the differential response to these two error outcomes varied considerably across participants. This behavioral difference was correlated with the physiological response to the two types of feedback: The more similarly participants treated the two outcomes at a behavioral level, the smaller the difference in MFN amplitude in response to these outcomes.

We propose that these findings may be reconciled by considering the potential top-down mechanisms driving execution error processing (Parvin et al., 2018). A recent study demonstrated that a sense of agency, operationalized through the perceived ability to correct for motor errors, biases choice behavior. Complementary evidence has shown the MFN's sensitivity to agency: Outcomes that can be controlled lead to larger negativities than those that cannot (Sidarus et al., 2017), with corroborating results showing attenuated MFNs in the absence of actively performed actions (Donkers and van Boxtel, 2005; Donkers et al., 2005). We reason that if execution errors elicit an increased sense of agency, then this could account for the observed large negativity for these outcomes.

A recent fMRI experiment using a similar 3-arm bandit task to the one employed here, revealed an attenuation of the signal associated with negative reward prediction error in the striatum following execution failures (McDougle et al., 2019). Our observation of a larger negative deflection for execution error trials in the MFN may appear contrary to these previously reported striatal results. However, the fMRI investigation did show increased ACC activity in response to execution versus selection errors, suggesting that execution errors have their own neural signature. Relatedly, the FRN has been shown to vary in response to reward prediction errors but not sensory prediction errors when comparing signals from reward and visuomotor rotation learning tasks (Palidis et al., 2019). Together with the fMRI results, the present data pose the possibility that execution error processing may be distinct from dopamine-related reinforcement learning processes.

We propose that the selection prediction error signal observed in these data represents a classic reward prediction error signal (FRN; Miltner et al., 1997; Nieuwenhuis et al., 2004), while the MFN elicited in response to execution prediction error is a manifestation of a delayed error related negativity (ERN) - an endogenous signal usually present ~100 ms following an overt incorrect responses (Falkenstein et al., 1991; Gehring et al., 1993).

The debate on whether the FRN and ERN share common neural circuitry, reflecting a general error monitoring mechanism is decades old (Zubarev and Parkkonen, 2018). The ERN-RL theory, which considers the signals equivalent, is based on converging lines of evidence showing that the ACC plays a key role in generating both components (Dehaene et al., 1994; Carter et al., 1998). However, recent research has shown that, whilst there is considerable overlap, the FRN includes an additional later anterior sub-component related to explicit reward (Potts et al., 2011). In this way, a general error monitoring system appears to comprise a rapid examination of error commission, followed by, if necessary, an evaluation of the reward.

## Outcome Processing Beyond the MFN

We also observed two distinct patterns of activity in time windows preceding and following the MFN. First, smaller amplitude responses were observed following execution errors relative to rewards in frontocentral sites 156-180 ms post-feedback, and the amplitude of this component correlated with switch rates. Second, in posterior sites (218-239 ms), larger amplitude responses were observed following execution errors relative to reward and this difference was also correlated with switch rates. The former may correspond to a system that rapidly evaluates the validity of the implemented choice. The latter is remarkably similar to the error positivity (Pe) signal that follows the ERN, and has been hypothesized to reflect the hierarchical organization of error monitoring and evaluation processing systems (van Veen and Carter, 2002). Importantly, in a reversal of the MFN pattern, magnitude differences in these early frontal and late parietal signals correlated with behavioral adjustment for Execution Prediction Error, but not for the Selection Prediction Error. This pattern provides further evidence of dissociable neural systems responsible for the processing of non-rewards attributed to execution and selection errors.

## Limitations and Future Directions

While we have hypothesized that execution errors might impact choice behavior, either by attenuating the operation of reinforcement learning processes or via an enhanced sense of agency, it is also important to consider alternative hypotheses. In Experiment 2 each bandit had the same expected value, but different frequencies of execution and selection errors. The observed preference for the bandit associated with high execution/low selection error rates may reflect an underlying drive to reduce uncertainty. That is, this bandit has a high level of uncertainty given that the potential reward, or lack thereof, is unknown on trials in which there is an execution error. Repeating that choice may reflect a bias to reduce that uncertainty, assuming the subsequent movement is successfully executed. Future work could disentangle these explanations by, for instance, assigning lower expected value to high execution/low selection error bandits and/or through the presentation of fictive outcomes for motor errors.

Finally, it is not clear if our findings are specific to execution errors, or to any endogenous or exogenous event that results in a unrewarded trial, but one that does not provide information about the reward probability associated with the selected object (Green et al., 2010). For example, if an unexpected gust of wind blew a tennis lob out-of-bounds, would that be treated as an “execution error”? Or, if after pulling the lever on a slot machine, a

power failure in the casino caused the game to terminate without a payoff, would this affect how the choice is judged? A future study could test endogenous execution errors (e.g., reaching error) and exogenous errors (e.g., the task screen goes blank randomly before an outcome is delivered). If similar results are found in both settings, elements of the early activity observed in frontocentral sites may indicate the establishment of a sensory "state", representing that the intended action plan was not properly implemented, irrespective of whether this mismatch was due to endogenous or exogenous factors, even before the prediction error is evaluated. This echoes the sequential ordering in models of temporal difference learning, where first the agent perceives its state, and then computes reward prediction errors relevant to that state (Sutton and Barto, 1998).

## Conclusion

We observed a robust MFN in response to both selection and execution errors, but only the former correlated with behavioral adjustment. Instead, the amplitude of a positive deflection in the ERP, both prior and subsequent to the MFN, correlated with choice behavior following execution errors. These results indicate a need for a more refined interpretation of what the MFN represents, and how it may be shaped by contextual information. More generally, they provide an insight into how the brain discriminates between different classes of error to determine future action.

# References


Bakeman R (2005) Recommended effect size statistics for repeated measures designs. Behav Res Methods 37:379–384 Available at: http://www.springerlink.com/index/10.3758/BF03192707.

Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Source J R Stat Soc Ser B J R Stat Soc Ser B J R Stat Soc B 57:289–300 Available at: http://www.jstor.org/stable/2346101%5Cnhttp://about.jstor.org/terms.

Brainard DH (1997) The Psychophysics Toolbox. Spat Vis 10:433–436.

Carter CS, Braver TS, Barch DM, Botvinick MM, Noll D, Cohen JD (1998) Anterior cingulate cortex, error detection, and the online monitoring of performance. Science (80- ) 280:747–749 Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9563953.

Chaumon M, Bishop DVM, Busch NA (2015) A practical guide to the selection of independent components of the electroencephalogram for artifact correction. J Neurosci Methods 250:47–63.

Cohen JD, McClure SM, Yu AJ (2007) Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. Philos Trans R Soc B Biol Sci 362:933–942.

Cohen MX, Ranganath C (2007) Reinforcement Learning Signals Predict Future Decisions. J Neurosci 27:371–378 Available at: http://www.jneurosci.org/cgi/content/abstract/27/2/371 [Accessed April 14, 2011].

Daw ND, O'Doherty JP, Dayan P, Dolan RJ, Seymour B (2006) Cortical substrates for exploratory decisions in humans. Nature 441:876–879.

Dehaene S, Posner MI, Tucker DM (1994) Localization of a Neural System for Error Detection and Compensation. Psychol Sci 5:303–305 Available at: http://pss.sagepub.com/content/5/5/303.short.

Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. J Neurosci Methods 134:9–21 Available at: http://www.ncbi.nlm.nih.gov/pubmed/15102499.

Dhar M, Pourtois G (2011) Early error detection is generic, but subsequent adaption to errors is not: Evidence from ERPs. Neuropsychologia 49:1236–1245 Available at: http://dx.doi.org/10.1016/j.neuropsychologia.2011.01.006.

Diedenhofen B, Musch J (2015) Cocor: A comprehensive solution for the statistical comparison of correlations. PLoS One 10:1–12.

Donkers FC, Nieuwenhuis S, van Boxtel GJ (2005) Mediofrontal negativities in the absence of responding. Cogn Brain Res 25:777–787.

Donkers FC, van Boxtel GJ (2005) Mediofrontal negativities to averted gains and losses in the slot-machine task - A further investigation. J Psychophysiol 19:256–262.

Dunn OJ, Clark V (1969) Correlation coefficients on the same sets of data. J Am Stat Assoc 64:366–377.

Falkenstein M, Hohnsbein J, Hoormann J, Blanke L (1991) Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. Electroencephalogr Clin Neurophysiol 78:447–455 Available at: http://www.sciencedirect.com/science/article/pii/0013469491900629.

Gehring WJ, Goss B, Coles MGH, Meyer DE, Donchin E (1993) A Neural System for Error-Detection and Compensation. Psychol Sci 4:385–390.

Gehring WJ, Willoughby A (2002) The medial frontal cortex and the rapid processing of monetary gains and losses. Science (80- ) 295:2279–2282 Available at: http://www.ncbi.nlm.nih.gov/pubmed/11910116 [Accessed July 15, 2010].

Gittins J (1979) Bandit processes and dynamic allocation indices. Statistics (Ber) 41:148–177 Available at: http://www.jstor.org/stable/10.2307/2985029.

Glazer JE, Kelley NJ, Pornpattananangkul N, Mittal VA, Nusslock R (2018) Beyond the FRN: Broadening the time-course of EEG and ERP components implicated in reward processing. Int J Psychophysiol.

Green CS, Benson C, Kersten D, Schrater P (2010) Alterations in choice behavior by manipulations of world model. Proc Natl Acad Sci 107:16401–16406 Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.1001709107.

Groppe DM, Urbach TP, Kutas M (2011) Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. Psychophysiology 48:1711–1725.

Hemmelmann C, Horn M, Süsse T, Vollandt R, Weiss S (2005) New concepts of multiple tests and their use for evaluating high-dimensional EEG data. J Neurosci Methods 142:209–217.

Hittner JB, May K, Silver NC (2003) A monte carlo evaluation of tests for comparing dependent correlations. J Gen Psychol 130:149–168.

Holroyd CB, Coles MGH (2002) The Neural Basis of Human Error Processing: Reinforcement Learning, Dopamine, and the Error-Related Negativity. Psychol Rev 109:679–709 Available at: http://www.sciencedirect.com/science/article/pii/S0033295X02005704.

Holroyd CB, Krigolson OE (2007) Reward prediction error signals associated with a modified time estimation task. Psychophysiology 44:913–917 Available at: http://www.ncbi.nlm.nih.gov/pubmed/17640267.

Kappenman ES, Luck SJ (2017) Best Practices for Event-Related Potential Research in Clinical Populations. Biol Psychiatry Cogn Neurosci Neuroimaging 4425:1–13.

Kleiner M, Brainard D, Pelli D, Ingling A, Murray R, Broussard C (2007) What's new in psychtoolbox-3? Perception 36:1–16.

Krigolson OE (2017) Event-related brain potentials and the study of reward processing: Methodological considerations. Int J Psychophysiol:0–1 Available at: http://dx.doi.org/10.1016/j.ijpsycho.2017.11.007.

Lage-Castellanos A, Martínez-Montes E, Hernández-Cabrera J a, Galán L (2010) False discovery rate and permutation test: an evaluation in ERP data analysis. Stat Med 29:63–74 Available at: http://www.ncbi.nlm.nih.gov/pubmed/19941298 [Accessed March 4, 2013].

Loonis RF, Brincat SL, Antzoulatos EG, Miller EK (2017) A Meta-Analysis Suggests Different Neural Correlates for Implicit and Explicit Learning. Neuron 96:521-534.e7 Available at: https://doi.org/10.1016/j.neuron.2017.09.032.

Luck SJ, Gaspelin N (2017) How to get statistically significant effects in any ERP experiment (and why you shouldn't). Psychophysiology 54:146–157.

McDougle SD, Boggess MJ, Crossley MJ, Parvin D, Ivry RB, Taylor JA (2016) Credit assignment in movement-dependent reinforcement learning. Proc Natl Acad Sci 113:6797–6802 Available at: http://www.pnas.org/lookup/doi/10.1073/pnas.1523669113.

McDougle SD, Butcher PA, Parvin D, Mushtaq F, Niv Y, Ivry RB, Taylor JA (2019) Neural Signatures of Prediction Errors in a Decision-making Task are Modulated by Action Execution Failures. Curr Biol 29:1606–1613 Available at: https://www.biorxiv.org/content/early/2018/11/20/474361.

Miltner WHR, Braun CH, Coles MGH (1997) Event-Related Brain Potentials Following Incorrect Feedback in a Time-Estimation Task: Evidence for a {\textquotedblleft}Generic{\textquotedblright} Neural System for Error Detection. J Cogn Neurosci 9:788–798 Available at: http://www.mitpressjournals.org/doi/abs/10.1162/jocn.1997.9.6.788.

Mushtaq F, Wilkie RM, Mon-Williams MA, Schaefer A (2016) Randomised prior feedback modulates neural signals of outcome monitoring. Neuroimage 125:868–879.

Nieuwenhuis S, Holroyd CB, Mol N, Coles MGH (2004) Reinforcement-related brain potentials from medial frontal cortex: origins and functional significance. Neurosci Biobehav Rev 28:441–448 Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=15289008 [Accessed April 25, 2011].

Nolan H, Whelan R, Reilly RB (2010) FASTER: Fully Automated Statistical Thresholding for EEG artifact Rejection. J Neurosci Methods 192:152–162.

Oken BS, Chiappa KH (1986) Statistical issues concerning computerized analysis of brainwave topography. Ann Neurol 19:493–497 Available at: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=3487282.

Olejnik S, Algina J (2003) Generalized Eta and Omega Squared Statistics: Measures of Effect Size for Some Common Research Designs. Psychol Methods 8:434–447.

Oostenveld R, Fries P, Maris E, Schoffelen J-M (2011) FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Intell Neurosci 2011:1:1–1:9.

Overbeek TJM, Nieuwenhuis S, Ridderinkhof KR (2005) Dissociable components of error processing - On the functional significance of the Pe Vis-a-vis the ERN/Ne. J Psychophysiol 19:319–329.

Palidis DJ, Cashaback J, Gribble P (2019) Neural Signatures of Reward and Sensory Prediction Error in Motor Learning. J Neurophysiol:262576 Available at: https://www.biorxiv.org/content/early/2018/02/09/262576.

Parvin DE, McDougle SD, Taylor JA, Ivry RB (2018) Credit assignment in a motor decision making task is influenced by agency and not sensorimotor prediction errors. J Neurosci 38:3601–3617 Available at: http://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.3601-17.2018.

Polich J (2007) Updating P300: an integrative theory of P3a and P3b. Clin Neurophysiol 118:2128–2148 Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2715154&tool=pmcentrez&rendertype=abstract [Accessed January 29, 2013].

Potts GF, Martin LE, Kamp S-MM, Donchin E (2011) Neural response to action and reward prediction errors: Comparing the error-related negativity to behavioral errors and the feedback-related negativity to reward prediction violations. Psychophysiology 48:218–228 Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2965315&tool=pmcentrez&rendertype=abstract [Accessed May 23, 2013].

R Core Team (2015) R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Sambrook TD, Goslin J (2015) A Neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. Psychol Bull 141:213–235.

San Martín R (2012) Event-related potential studies of outcome processing and feedback-guided learning. Front Hum Neurosci 6 Available at: http://www.frontiersin.org/Journal/Abstract.aspx?s=537&name=human_neuroscience&ART_DOI=10.3389/fnhum.2012.00304.

Sidarus N, Vuorre M, Haggard P (2017) How action selection influences the sense of agency: An ERP study. Neuroimage 150:1–13.

Sutton RS, Barto AG (1998) Reinforcement Learning: An Introduction. Cambridge, MA: MIT Press.

van Veen V, Carter CS (2002) The anterior cingulate as a conflict monitor: fMRI and ERP studies. Physiol Behav 77:477–482 Available at: http://www.sciencedirect.com/science/article/pii/S0031938402009307.

von Borries a KL, Brazil I a, Bulten BH, Buitelaar JK, Verkes RJ, de Bruijn ER a (2010) Neural correlates of error-related learning deficits in individuals with psychopathy. Psychol Med 40:1559–1568 Available at: http://www.ncbi.nlm.nih.gov/pubmed/19995478 [Accessed January 7, 2015].

Walsh MM, Anderson JR (2012) Learning from experience: Event-related potential correlates of reward processing, neural adaptation, and behavioral choice. Neurosci Biobehav Rev 36:1870–1884 Available at: http://dx.doi.org/10.1016/j.neubiorev.2012.05.008.

Winkler I, Debener S, Muller KR, Tangermann M (2015) On the influence of high-pass filtering on ICA-based artifact reduction in EEG-ERP. Proc Annu Int Conf IEEE Eng Med Biol Soc EMBS 2015-Novem:4101–4105.

Wu S-W, Delgado MR, Maloney LT (2009) Economic decision-making compared with an equivalent motor task. Proc Natl Acad Sci U S A 106:6088–6093.

Yeung N, Sanfey AG (2004) Independent coding of reward magnitude and valence in the human brain. J Neurosci 24:6258–6264 Available at: http://www.ncbi.nlm.nih.gov/pubmed/15254080 [Accessed January 31, 2013].

Zubarev I, Parkkonen L (2018) Evidence for a general performance-monitoring system in the human brain. Hum Brain Mapp 39:4322–4333.