

RESEARCH ARTICLE | *Control of Movement*

EEG correlates of physical effort and reward processing during reinforcement learning

Dimitrios J. Palidis^{1,2,3} and  Paul L. Gribble^{1,2,4,5}

¹The Brain and Mind Institute, Western University, London, Ontario, Canada; ²Department of Psychology, Western University, London, Ontario, Canada; ³Graduate Program in Neuroscience, Schulich School of Medicine and Dentistry, Western University, London, Ontario, Canada; ⁴Department of Physiology and Pharmacology, Schulich School of Medicine and Dentistry, Western University, London, Ontario, Canada; and ⁵Haskins Laboratories, New Haven, Connecticut

Submitted 22 June 2020; accepted in final form 19 July 2020

Palidis DJ, Gribble PL. EEG correlates of physical effort and reward processing during reinforcement learning. *J Neurophysiol* 124: 610–622, 2020. First published July 29, 2020; doi:10.1152/jn.00370.2020.—Effort-based decision making is often modeled using subjective value, a function of reward discounted by effort. We asked whether EEG event-related potential (ERP) correlates of reward processing are also modulated by physical effort. Human participants performed a task in which they were required to accurately produce target levels of muscle activation to receive rewards. Quadriceps muscle activation was recorded with electromyography (EMG) during isometric knee extension. On a given trial, the target muscle activation required either low or high effort. The effort was determined probabilistically according to a binary choice, such that the responses were associated with 20% and 80% probability of high effort. This contingency could only be learned through experience, and it reversed periodically. Binary reinforcement feedback depended on accurately producing the target muscle activity. Participants adaptively avoided effort by switching responses more frequently after choices that resulted in hard effort. Feedback after participants' choices that revealed the resulting effort requirement did not elicit modulation of the feedback-related negativity/reward positivity (FRN/RP). However, the neural response to reinforcement outcome after effort production was increased by preceding physical effort. Source decomposition revealed separable early and late positive deflections contributing to the ERP. The main effect of reward outcome, characteristic of the FRN/RP, loaded onto the earlier component, whereas the reward \times effort interaction was observed only in the later positivity, which resembled the P300. Thus, retrospective effort modulates reward processing. This may explain paradoxical behavioral findings whereby rewards requiring more effort to obtain can become more powerful reinforcers.

NEW & NOTEWORTHY Choices probabilistically determined the physical effort requirements for a subsequent task, and reward depended on task performance. Feedback revealing whether choices resulted in easy or hard effort did not elicit reinforcement learning signals. However, the neural responses to reinforcement were modulated by preceding effort. Thus, effort itself was not treated as loss or punishment, but it affected the responses to subsequent reinforcement outcomes. This may explain how effort can enhance the motivational effect of reward.

effort; feedback-related negativity; human; reinforcement; reward

INTRODUCTION

Humans and other animals tend to make decisions that lead to more rewarding and less physically effortful outcomes (Hartmann et al. 2013; Kennerley et al. 2009; Morel et al. 2017; Rangel and Hare 2010; Selinger et al. 2015; Shadmehr et al. 2016; Walton et al. 2006). During decision making dopaminergic, prefrontal, and striatal structures are implicated in motivating effortful behavior to obtain reward and in integrating reward and effort cost to make value-based choices (Denk et al. 2005; Hosking et al. 2015; Kurniawan et al. 2010, 2011; Rudebeck et al. 2008; Salamone et al. 2003, 2007; Schweimer et al. 2005; Walton et al. 2003). After decisions are made and actions are produced, midbrain dopaminergic neurons signal the difference between expected and obtained reward to the ventral striatum and prefrontal regions (Bayer and Glimcher 2005; Gläscher et al. 2010; Graybiel 2008; Holroyd and Coles 2002; Pessiglione et al. 2006; Puig and Miller 2012; Schultz 2006). This reward prediction error signal is thought to drive reinforcement learning by updating reward expectations, allowing for adaptive behavior in uncertain or changing environments. Although reward-processing areas have been shown to process effort costs during decision making before action selection, we know relatively little about whether motor costs associated with effort also modulate reward processing after an action is completed, in response to feedback about success or failure.

During decision making the anterior cingulate cortex (ACC) is known to encode prospective reward and effort cost and to integrate both into a unitary subjective utility signal characterized by effort-discounted reward (Croxson et al. 2009; Kennerley et al. 2011; Klein-Flügge et al. 2016; Porter et al. 2019; Prévost et al. 2010; Rudebeck et al. 2006). During outcome evaluation the ACC encodes reward prediction error and supports reinforcement learning (Amiez et al. 2005; Ito et al. 2003; Kennerley et al. 2011; Seo and Lee 2007; Walsh and Anderson 2012; Williams et al. 2004), but it remains to be shown whether reward learning signals in the ACC also integrate motor effort costs. Contrary to this idea, fMRI studies have argued that separate neural systems underlie reward and effort learning, with ACC activity reflecting prediction errors for effort but not reward (Hauser et al. 2017; Skvortsova et al. 2014). However, an event-related potential (ERP) measured by EEG called the

Correspondence: P. L. Gribble (pgribble@uwo.ca).

feedback-related negativity, or alternatively the reward positivity (FRN/RP), is a reliable neural correlate of reward prediction error and is consistently localized to the ACC (Becker et al. 2014; Cohen and Ranganath 2007; Emeric et al. 2008; Gehring and Willoughby 2002; Hauser et al. 2014; Holroyd and Coles 2002; Mathewson et al. 2008; Miltner et al. 1997; Nieuwenhuis et al. 2005; Vezoli and Procyk 2009; Walsh and Anderson 2012; Warren et al. 2015). We sought to test whether the FRN/RP not only acts as a learning signal for reward outcomes but also integrates physical effort requirements during learning. Because other ERP components have also been implicated in outcome processing, we tested for effects in a broad temporal range (Glazer et al. 2018). In particular, the P300 is affected by various properties of motivationally relevant feedback including valence, magnitude, likelihood, and other high-level variables (Ma et al. 2014; San Martín 2012; Sato et al. 2005; Wang et al. 2014; Wu and Zhou 2009; Yeung and Sanfey 2004).

Economic theories assert that effort is a cost that devalues reward and thus predict a diminished neural response to reinforcement for more costly rewards (Botvinick et al. 2009; Hartmann et al. 2013; Hauser et al. 2017; Shadmehr et al. 2016). Paradoxically, it has been found in humans and animals that effort can enhance the reinforcing quality of rewards (Clement et al. 2000; Inzlicht et al. 2018; Lydall et al. 2010; Zentall 2010). It may be that prospective effort devalues reward, while retrospective effort amplifies reinforcement. For example, when given a choice between responses requiring high and low effort, participants choose to produce less effort in the immediate future. However, when given a choice between conditioned reinforcers that follow either low or high effort, humans and other animals tend to prefer the reinforcer that followed greater effort in the past (Alessandri et al. 2008; Clement et al. 2000; Hernandez Lallement et al. 2014; Zentall 2010). Like many real-world situations, uncertain reward was obtained only after effort expenditure in the present study.

Previous EEG experiments have investigated interactions between effort and reward processing. In Gheza et al. (2018), participants were sometimes permitted to redo gambles that resulted in no reward by paying an effort cost. In trials with the opportunity to redo, reward resulted in a larger neural reward positivity signal. In this case, reward outcomes also meant that participants avoided a potential prospective effort cost, which could serve to increase the valuation of the reward. However, the effort manipulation also required additional time to complete, which delayed task progress, and the effects of effort cost vs. temporal cost could not be dissociated. In tasks that provide reinforcement feedback after performance of cognitive tasks with varying attentional or mental demands, increased effort has been shown to enhance the FRN/reward positivity (Ma et al. 2014; Schevernels et al. 2014; Wang et al. 2017). This is consistent with the notion that preceding effort enhances reward signals. However, increased effort in cognitive tasks is almost invariably associated with higher difficulty and thus lower probability of success. In this case, it is difficult to determine whether enhanced reward signals are due to increased effort or lower reward expectations, which would result in larger reward prediction error.

In the present study, physical effort is manipulated by changing the magnitude of muscle contractions required to complete the task, but the probabilities of reward and success

are equated across the different effort conditions. This allows us to assess the effects of effort in terms of motor cost without confounds related to reward expectation. Furthermore, unlike previous EEG experiments that varied cognitive effort randomly, physical effort requirements in the present study were affected by participants' choice behavior. It has been shown that the FRN/RP and other neural correlates of outcome processing are typically more sensitive when outcomes are attributable to agents' actions (Hassall et al. 2019; Martin and Potts 2011; Sambrook and Goslin 2015; Walsh and Anderson 2012; Yeung et al. 2005; Zink et al. 2004).

Participants first made binary choices, and then they received feedback about the resulting effort requirements, which were probabilistic and uncertain. Subsequently, they performed an effortful electromyographic (EMG) production task for which they received variable reward that was dependent on precisely producing a target level of EMG activity. This trial sequence allowed us to test the hypothesis that effort information is maintained during the course of an action and that this information is integrated retrospectively with reward feedback. According to this hypothesis, feedback indicating effort requirements in the present study would not elicit neural reinforcement signals such as the FRN/RP, whereas the neural response to reward feedback at the end of each trial would be modulated by both reinforcement outcome and the preceding effort. Alternatively, if effort is treated simply as an aversive stimulus or an economic loss by a standard temporal difference learning process, then feedback that predicts the upcoming effort but not the reward outcome should elicit neural reinforcement signals (Mulligan and Hajcak 2018).

MATERIALS AND METHODS

Participants

A total of $n = 18$ healthy participants were included in our study (mean age: 22.12 yr, SD: 3.66; 9 men, 9 women). Four participants underwent the experimental procedure but were excluded because of excessive EEG artifacts caused by sweat or movement associated with the task. Participants provided written informed consent to experimental procedures approved by the Research Ethics Board at The University of Western Ontario.

Experimental Setup

To allow for isometric contractions of the quadriceps muscles, participants were restrained to a chair by straps on their shoulders and waists. Participants' ankles were strapped to a rack fixed at the base of the chair, with the knees bent at $\sim 90^\circ$. Participants were seated in front of a CRT monitor with their hands resting on a table positioned to make button presses on a response box.

EMG and EEG Recording

Unreferenced EEG activity was recorded at 512 Hz with a 64-channel Biosemi ActiveTwo system (Biosemi). Electrodes were mounted in an elastic cap and distributed according to the extended 10-20 system with electrode Cz placed over the vertex. Instead of the typical ground electrode, Biosemi forms a feedback loop between an active Common Mode Sense electrode and a passive Driven Right Leg electrode. The Common Mode Sense electrode was located in the center of the area between P1, Pz, PO3, and POz. The Driven Right Leg electrode was located in the center of the area between Pz, P2, PO3, and PO4. Electrooculogram (EOG) was recorded with elec-

trodes placed above and below each eye and the outer canthus of each eye. Additional electrodes were placed on each mastoid.

EMG activity was recorded at 2,400 Hz bilaterally from the rectus lateralis muscles of the quadriceps with an active electrode system and amplifier (g.USBamp; g.tec Medical Engineering). Two electrodes were placed on each muscle belly for bipolar recordings, and a ground electrode was placed on the left shin. EMG signals were filtered at the time of recording with a 5- to 500-Hz band-pass filter and a 60-Hz notch filter.

Visual Feedback of EMG

The EMG signal used to provide online visual feedback of quadriceps muscle activity was first rectified, low-pass filtered with a 10-Hz cutoff frequency, and then downsampled to 120 Hz. At the beginning of each block, participants performed isometric knee extensions with maximum effort continuously for 4 s. All samples greater than the median value recorded during maximum effort were averaged to determine the value of maximum voluntary contraction (MVC) used throughout the block. Subsequently, participants were cued to remain completely still and keep their legs relaxed for 4 s. The mean EMG signal during this period was used as a baseline value throughout the block.

During each trial, an animation of a thermometer was displayed to participants. The fluid level of the thermometer increased in real time (monitor refresh rate: 60 Hz) as a linear function of the processed EMG signal. In “hard” effort trials, the top of the thermometer corresponded to 85% of MVC and the bottom of the thermometer corresponded to the baseline measure. In “easy” effort trials, the top of the thermometer corresponded to 15% of MVC. Because easy trials required only a small amount of muscle activity to reach the target, the gain of the visual feedback relative to the EMG was high. To reduce the gain and provide smooth feedback, the baseline measure (resting EMG) was made to correspond to the point halfway up the thermometer for the easy condition, so that the temperature moved half the distance on the display. The fluid level was calculated separately for each leg based on their respective MVC and baseline measures, and the average was used to display feedback. A running average of fluid level for the previous 60 samples was drawn to the screen to provide smooth feedback. During each trial, the maximum fluid level for that trial was continuously displayed such that the fluid level only increased, and if the participant relaxed their quadriceps muscles the feedback would remain at the same level. This allowed for smooth, ballistic isometric contractions. It also made it so that participants were not required to hold the fluid level constant without visual feedback, which often resulted in the fluid level fluctuating or drifting away from the target during pilot experiments.

Experimental Task

Participants first performed a block of 28 practice trials (see below). Participants then performed four blocks of 74 trials with self-paced rest periods between blocks. Each block consisted of 12 control condition trials, followed by 50 experimental condition trials and finally 12 additional control condition trials. At the end of each block, participants were verbally surveyed as to how physically effortful the easy and hard effort trials were with a scale of 1 to 5, with 5 corresponding to maximum effort.

Experimental condition. During each trial, participants made a binary choice that probabilistically determined whether the trial would require easy or hard physical effort. The effort contingencies had to be learned through experience. Participants then performed isometric knee extensions to control visual EMG feedback on a screen. Participants were instructed to exceed a minimum level of muscle activation indicated by a visual target while remaining as close as possible to the target. Binary reinforcement feedback was provided at the end of each

trial to indicate success or failure, which corresponded to a small monetary reward.

Visual stimuli are shown in Fig. 1. An animated thermometer was drawn on the screen throughout the task. A cross was drawn at the top of the thermometer to serve as a target for EMG feedback. Letters “A” and “B” drawn to the left and right of the thermometer represented the options for binary choices made in each trial. Participants initiated each trial by pressing either a left or right button on a response box with their left or right index finger, respectively. Immediately on each button press, the choice was indicated by a box appearing around the letter “A” or “B” for the left and right response buttons, respectively. The box remained throughout the trial.

Effort feedback. One second after the button press participants received feedback indicating the effort that they would be required to exert on the present trial. The word “easy” or “hard” replaced the target cross for 700 ms to indicate upcoming required effort. The effort condition was determined probabilistically by the participants’ response, and the effort contingencies had to be learned through experience. One of the responses led to a hard effort trial with a probability of 0.8 and an easy effort trial with a probability of 0.2. The other response led to a hard effort trial with a probability of 0.2 and an easy effort trial with a probability of 0.8. Unannounced to participants, the effort contingencies periodically reversed. Reversals occurred after the response more likely to produce easy effort was chosen a cumulative number of times, which was randomly selected to be between 5 and 9 for each reversal. Participants were instructed that their responses would affect the effort requirements in some way but were not informed of the specific nature of the task. Participants were not instructed to respond in any particular way other than to sample both choices.

After the effort feedback was removed from the display, the target cross reappeared for 800 ms. Subsequently, the effort production phase of the trial began. During this phase, the fluid level of the thermometer was drawn continuously to provide EMG feedback (see *Visual Feedback of EMG*). The fluid level increased with increasing EMG signal but represented the maximum signal for the trial, and thus never decreased. A purple circle was drawn under the target to cue the beginning of the effort production phase, and participants were instructed to keep their legs relaxed until they saw this cue. The circle shrank continuously during the course of the trial, disappearing in 2,500 ms to signal the end of the effort production phase, at which point EMG feedback disappeared. Participants were instructed that to complete the task successfully the final fluid level must exceed the target represented by the center of the cross. The target corresponded to 15% and 85% of MVC in the easy effort and hard effort conditions, respectively. Furthermore, participants were instructed to keep the fluid level as close as possible to the target; thus their goal was to always overshoot the target but to minimize the extent of overshoot. Participants were instructed to relax their legs as soon as possible after reaching the target, as the fluid level did not decrease during a trial. EMG feedback was withheld above the target by a mask drawn on the top of the thermometer. This prevented participants from seeing the extent of their overshoot errors.

Reinforcement feedback. Feedback about performance was provided at the end of the trial with binary reinforcement. At the end of the effort production phase, the EMG feedback and the mask disappeared. After 1,500 ms of fixation, the target cross was replaced with either “\$\$\$” or “XXX” to indicate a rewarded or failed trial, with a reward being indicated if the fluid level exceeded the target while remaining sufficiently close to it. Participants were instructed that they could earn up to an additional 10 CAD throughout the task according to the number of trials in which they received feedback indicating success. The error threshold for overshoot was adjusted with a 1-up-1-down adaptive staircase separately for the two effort conditions to ensure a 50% reinforcement rate overall for both conditions.

Control condition. Each block began and ended with 12 control trials, during which the task was the same as the experimental

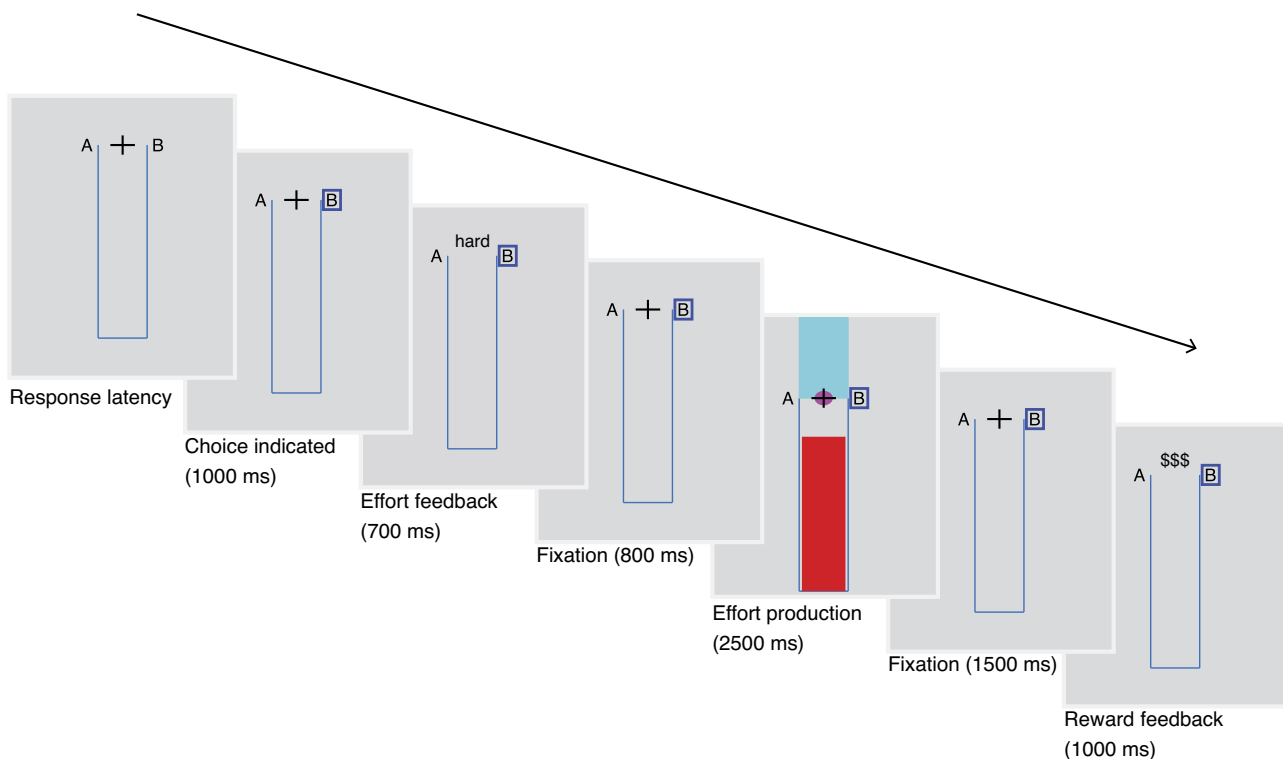


Fig. 1. Stimuli. Participants initiated each trial by indicating a binary decision through button press. A box immediately appeared around either the letter “A” or “B,” corresponding to the choice options. After 1,000 ms, feedback appeared to inform participants that their choice resulted in either easy or hard physical effort requirements for the upcoming electromyographic (EMG) production task. A purple circle appeared over the target cross to cue the onset of the EMG production phase, during which participants performed isometric knee extension and the fluid level of a thermometer indicated quadriceps muscle activation. The circle shrank continuously, disappearing in 2.5 s to cue the end of the EMG production phase. Participants attempted to bring the fluid level above the target, represented by a cross, while remaining as close as possible to the target. However, a mask drawn above the target prevented participants from seeing the extent of errors that they made in overshooting the target. Instead, binary reinforcement feedback was provided 1.5 s after the EMG production phase ended, indicating whether or not participants had successfully exceeded the target while remaining sufficiently close.

condition except no reinforcement feedback was provided and the effort condition was deterministic and independent of participants’ responses. Both runs of 12 control trials consisted of 6 easy effort trials and 6 hard effort trials, with the trials of each effort condition occurring consecutively. The text “easy effort” or “hard effort” was displayed at the top of the screen continuously to cue the effort condition for all control trials. Participants were instructed to make a button press to initiate each trial but were instructed that the choice was arbitrary and that the effort condition would always correspond to the cue at the top of the screen. In the first 12 control trials of each block, there was no mask drawn on the top of the thermometer, so participants could see their overshoot errors in order to practice the task more effectively. In the final 12 control trials of each block, the mask was drawn for each trial as in the experimental condition. The orders of easy and hard condition runs during the control trials were randomized and balanced across the four blocks for each participant.

Practice trials. Participants first performed a practice block to learn how to control the EMG feedback. As in the control trials, no reinforcement feedback was provided and the effort condition was cued to participants before each trial and independent of participants’ responses. The practice block began with seven easy effort trials followed by seven hard effort trials without the mask drawn at the top of the thermometer. Participants then performed seven easy effort trials followed by seven hard effort trials with the mask.

Behavioral analysis. The effect of effort and reinforcement outcomes on behavioral choice was analyzed with logistic regression performed with the *Glmnet* package in R. The dependent variable was whether the participants’ choice on trial n corresponded to staying or switching from the choice on trial $n - 1$, coded as 0 or 1. The

independent variables were determined by the effort and reinforcement outcomes on trial $n - 1$:

Effort: -1 for easy effort, 1 for hard effort

Reward: -1 for nonreward, 1 for reward

Effort/Reward interaction: Effort \times Reward

Switch: 1 for all trials

Logistic regression was calculated separately for each participant. Regularization was applied with an L2-norm penalty. The penalty constant, λ , was selected by leave-out cross-validation. A value of 0.04297 was chosen as it is the largest value that minimizes the cross-validated misclassification error, averaged across subjects. The coefficients for Effort, Reward, and the interaction term were each submitted to one-sample t tests against zero.

EEG Preprocessing and Denoising

EEG data were preprocessed with the EEGLAB toolbox (see Delorme and Makeig 2004 for details), except for filtering, which was performed with the MATLAB *filtfilt* function. Data, initially referenced to linked mastoids, were band-pass filtered with a second-order Butterworth filter with a passband of 0.1–45 Hz. Channels with poor recording quality or excessive artifacts were identified with visual inspection and interpolated with spherical interpolation. EEG data were then rereferenced to the average scalp potential, and interpolated electrodes were subsequently removed from the data before independent component analysis (ICA). Two epochs were extracted for each trial corresponding to effort condition feedback following the button press response and reinforcement feedback following the effort production phase. Continuous data were segmented into 2.5-s epochs time-locked to stimulus onset at 0 ms (time range: $-1,000$ to $+1,500$

ms). Data epochs containing artifacts other than blinks were removed by visual inspection. Subsequently, extended infomax ICA was performed on each participant's data (Delorme and Makeig 2004). Components reflecting eye movements and blink artifacts were identified by visual inspection and subtracted by projection of the remaining components back to the voltage time series.

Event-Related Potential Analysis

Trial averaging. We computed event-related potentials (ERPs) on an individual participant basis by trial-averaging EEG time series epochs recorded from electrode FCz after artifact removal. ERPs were analyzed after time-locking signals to two points in time: effort feedback and reinforcement (performance) feedback. The FRN/RP is typically maximal at electrode FCz, and this selection is consistent with previous work including our own (Holroyd and Krigolson 2007; Miltner et al. 1997; Palidis et al. 2019; Pfabigan et al. 2011). We selected trials corresponding to various feedback conditions in each task.

For ERPs time locked to reinforcement feedback, we computed ERPs corresponding to “easy nonreward” (45.4 ± 6.9 trials) “easy reward” (46.3 ± 6.2 trials), “hard nonreward” (37.9 ± 8.4 trials), and “hard reward” (45.6 ± 7.7 trials) conditions. In the control condition, participants performed the effort production task but did not receive any reinforcement feedback. We computed ERPs for the “control easy” (41.5 ± 5.9 trials) and “control hard” (38.5 ± 5.5 trials) conditions time locked to the moment when reinforcement feedback would have been delivered in the experimental condition. For ERPs corresponding to reinforcement feedback and the control condition, we excluded all trials in which the visual EMG feedback did not reach the target, as in this case a nonreward outcome was evident before the reinforcement feedback was delivered.

We also extracted ERPs time locked to the effort condition feedback, which indicated the upcoming effort requirements after each button press but before the participant performed the EMG production task (“easy feedback” 94.9 ± 10.2 trials and “hard feedback” 92.4 ± 13.5 trials). All ERPs were baseline corrected by subtracting the average voltage in the 100-ms period immediately before stimulus onset. Finally, ERPs were low-pass filtered with a cutoff frequency of 30 Hz.

Statistical analysis. We performed statistical tests on each sample between 100 and 600 ms after feedback onset. We selected this time window as it is wide enough to capture effects outside of the FRN/RP yet constrained to a range during which ERPs are likely to be affected by feedback processing (Glazer et al. 2018). We corrected significance values for multiple comparisons across time with the Benjamini–Hochberg procedure for estimating the false discovery rate (FDR), implemented by the MATLAB *mafdr* function.

To analyze the neural response to reinforcement feedback, we performed 2×2 repeated-measures ANOVA with the MATLAB *ranova* function. The factors were reward outcome (levels: nonreward, reward) and effort condition (levels: easy, hard). We used one-sample *t* tests against zero on the difference waves computed between easy feedback and hard feedback ERPs, aligned to feedback indicating effort condition after each button press but before the EMG production phase. To test for artifacts related to the isometric leg extension, we used one-sample *t* tests against zero on the difference waves computed between easy control and hard control ERPs. These ERPs were aligned to the moment when reinforcement feedback would have been delivered in the experimental condition, but instead the target cross simply disappeared briefly. Participants were told that they would not receive feedback in this condition and thus did not expect a possible reward.

Scalp distributions. Scalp distributions were plotted with the EEGLAB *topoplot* function using the mean amplitude of difference waves within specified time windows, averaged across subjects.

Source Separation

Because of volume conduction, potential differences between any particular scalp electrodes contain mixed contributions from nearly all active neural sources and artifacts. Measurements of any event-related potential (ERP) component using scalp regions of interest, such as the measurements of the FRN/RP described above, are thus prone to contamination by other ERP components with distinct neural sources. Independent component analysis (ICA) can be used to produce spatial filters that isolate activity measured from separate cortical sources (Onton and Makeig 2006). Each component returned by ICA is a linear weighting of all electrodes, computed to produce signals with maximal temporal independence (Delorme and Makeig 2004).

Because ICA decomposition of neural activity sources can be particularly sensitive to signal properties and noise, we preprocessed the data using a modified procedure to produce ICA weightings for source separation. Except for the differences described below, we followed the same preprocessing used for the ERP analysis, including rejection of the same channels and epochs for artifact removal, as described in *EEG Preprocessing and Denoising*. EEG data were downsampled to 256 Hz and high-pass filtered with a second-order Butterworth filter with a cutoff of 0.25 Hz instead of 0.1 Hz. Aggressive high-pass cutoff frequencies of 1–2 Hz have been shown to improve ICA decomposition (Winkler et al. 2015). However, high-pass filtering at or above 0.3 Hz has been shown to attenuate and distort long-latency ERP components, and 0.1-Hz cutoff is generally recommended for ERP analysis (Acunzo et al. 2012; Bougrain et al. 2012; Holinger et al. 2000; Tanner et al. 2015). We chose to compute ICA weights with data high-pass filtered with a 0.25-Hz cutoff as a suitable compromise. Sixty-hertz power line noise was removed with the *CleanLine* EEGLAB plugin (Mullen 2012). Data were selected with the time range -100 to $+600$ ms centered around effort feedback and reinforcement feedback, instead of $-1,000$ to $+1,500$, so that the ICA would primarily account for variance in the time window of interest. The same epochs previously identified to contain artifacts, as described in *EEG Preprocessing and Denoising*, were rejected before extended infomax ICA was performed on each participant's data. Subsequently, the time series of the independent components' activities were visually inspected, and additional epochs containing artifacts were flagged and removed before recomputing the ICA weights. According to the tutorial wiki maintained by the developers of EEGLAB, ICA can “concentrate” artifacts for easier rejection, and recomputing ICA after such a rejection “may improve the quality of the ICA decomposition, revealing more independent components accounting for neural, as opposed to mixed artifactual activity” (see https://seccn.ucsd.edu/wiki/Chapter_01:_Rejecting_Artifacts).

To analyze the contribution of individual independent components (ICs) to the ERPs (IC-ERPs), the ICA weights computed for each participant's data were then applied to the data originally preprocessed for ERP analysis as described in *EEG Preprocessing and Denoising*. Thus, identical preprocessed data were used for traditional ERP analysis and the IC-ERP analysis, along with identical procedures described above in *Trial averaging* and *Statistical analysis*. The only difference was that the traditional ERP analysis used mixed data recorded from electrode FCz, whereas the IC-ERP analysis used activity from selected ICs back-projected onto electrode FCz. ICs corresponding to brain activity as opposed to artifacts were identified by stereotyped properties including scalp topographies resembling dipolar projections and spectral peaks at frequencies typical of EEG activity. The traditional ERP analysis revealed multiple ERP components peaking at different latencies. ICs corresponding to particular ERP components were identified by the presence of maximal peaks in the IC-ERPs in corresponding time windows (see *IC-ERP Results*).

RESULTS

Behavioral Results

Participants made binary decisions that probabilistically determined the effort requirements for each trial. Participants underwent the hard effort condition in 49.6% (SD: 4.8%) of trials. Reward was delivered if EMG feedback exceeded a target level while staying sufficiently close to the target. Participants received reward in 49.4% (SD: 0.01%) of trials. We performed logistic regression for each subject to predict switching of responses between trials $n - 1$ and n , with the effort condition and reward outcome on trial $n - 1$ as the predictors. Figure 2B shows the coefficients estimated for each subject, and Fig. 2A shows the proportion of trials after which participants switched responses for the different reward and effort outcomes. We found that the coefficients for the effect of effort on switching were significantly greater than zero [1-sample t test; $t(17) = 2.263$, $P = 0.037$]. The coefficients for the effect of reward were not reliably different from zero [$t(17) = -0.871$, $P = 0.3959$], nor were the coefficients for the interaction term [$t(17) = 0.252$, $P = 0.8043$].

ERP Results

Figure 3A shows the ERPs elicited by reinforcement feedback. We analyzed the neural response to reinforcement feedback by performing 2×2 repeated-measures ANOVA for each individual time point 100–600 ms after feedback onset. P values are corrected for multiple comparisons across time points with FDR. In response to reinforcement feedback, that ERP amplitude was larger in response to reward compared with nonreward between 184 and 336 ms after feedback onset (Fig. 3B; reward main effect, ranges for significant time points: $F = [7.99 \text{ } 52.99]$, $P = [0.045 \text{ } 0.0001]$, uncorrected $P = [0.012 < 0.0001]$). We also found that voltage was lower in the hard effort condition than in the easy effort condition 238–254 ms after feedback onset (Fig. 3C; main effect effort, ranges for significant time points: $F = [13.74 \text{ } 14.45]$, $P = [0.050 \text{ } 0.050]$, uncorrected $P = [0.0018 \text{ } 0.0011]$). We found effort/reward interaction effects starting 250 ms after feedback onset and up to 600 ms, the end of our time window for statistical testing (Fig. 3D; ranges for significant time points: $F = [5.73 \text{ } 38.94]$, $P = [0.048 \text{ } 0.0008]$, uncorrected $P = [0.028 < 0.0001]$).

In the control condition, participants performed the EMG production task, but the effort condition was predetermined and no reinforcement feedback was provided. We found no reliable differences between the control easy and control hard ERPs (ranges for all time points between 100 and 600 ms: $t = [-1.56 \text{ } 1.42]$, uncorrected $P = [1.00 \text{ } 0.139]$).

After participants produced binary decisions by button press, feedback was provided to indicate the resulting effort condition for the current trial. Figure 4A shows the ERPs elicited on each trial by feedback indicating the effort condition. We found significant differences between the easy feedback and hard feedback ERPs between 373 and 522 ms after effort feedback (Fig. 4B; ranges for significant time points: $t = [-4.31 \text{ } -2.81]$, $P = [0.049 \text{ } 0.015]$, uncorrected $P = [0.012 \text{ } 0.0005]$). Although we only performed statistical testing up to 600 ms to avoid sacrificing statistical power, we observed that the easy feedback – hard feedback difference wave remains at least 1 standard error below zero until 1,018 ms after feedback.

IC-ERP Results

The ERP responses to reinforcement feedback at electrode FCz clearly contained multiple components (Fig. 3), with the most obvious contributions being from a relatively early positive deflection peaking between 164 and 172 ms after feedback onset depending on the feedback condition and a later positive deflection peaking between 334 and 359 ms. Independent component analysis applied to individual participants' data consistently outputted ICs corresponding to these ERP components. We identified ICs that resembled neural EEG activity and computed the average ERPs of IC activity projected onto electrode FCz. For each participant, we identified the early ERP component by selecting the IC with the largest maximal positive peak occurring before 200 ms. We identified the later component by selecting the IC with the largest maximal positive peak after 200 ms. We identified neural ICs containing maximal peaks in the early time window for 16/18 participants and in the late time window for 17/18 participants.

Early ERP component. Figure 5A shows the IC-ERPs elicited by reinforcement feedback for the early ERP components. We performed 2×2 repeated-measures ANOVA on IC-ERP amplitude at each individual time point 100–600 ms after feedback onset. P values are corrected for multiple compari-

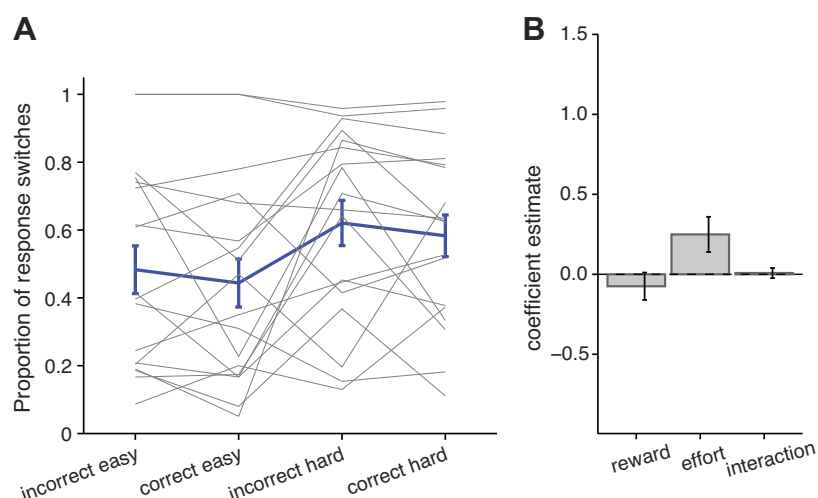


Fig. 2. Participants switch responses more frequently after hard physical effort than easy effort. **A**: the proportion of trials on which participants switched responses between trial $n - 1$ and trial n for the different reward outcomes and effort conditions on trial $n - 1$ (error bars: \pm SE). **B**: coefficients estimated by using logistic regression to predict response switching for each participant (bars indicate mean \pm 1 SE). Predictors were reward outcome, effort condition, and reward/effort interaction. The effort term was significantly greater than 0 ($P = 0.037$), indicating that participants were more likely to switch responses after hard effort than easy effort.

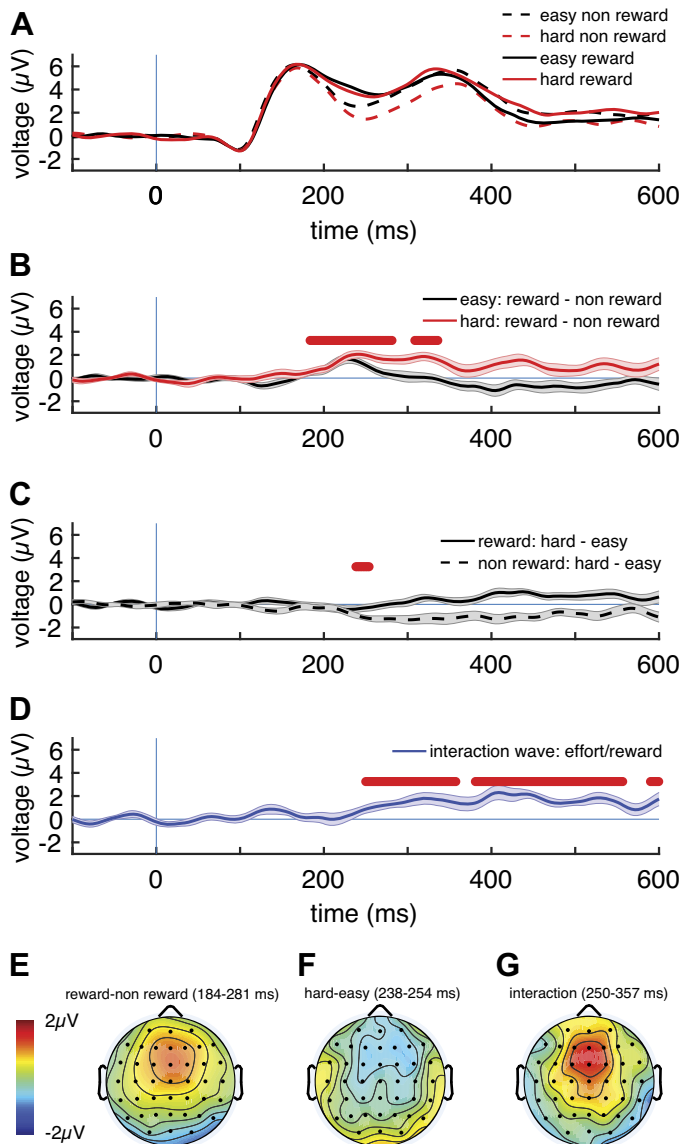


Fig. 3. Event-related potential (ERPs) elicited by reinforcement feedback. **A**: trial-averaged ERPs recorded from electrode FCz aligned to reinforcement feedback presentation (0 ms: vertical blue line), selected for reinforcement outcome (reward or nonreward) and the physical effort requirement on that trial (easy or hard). **B**: mean difference waves computed as reward ERP – nonreward ERP separately for the easy and hard effort conditions (shaded region: \pm SE). Red markers indicate time points between 100 and 600 ms with significant main effect of reward outcome [$P < 0.05$, false discovery rate (FDR) corrected]. **C**: mean difference waves computed as hard effort ERP – easy effort ERP separately for the reward and nonreward effort conditions (shaded region: \pm SE). Red marker indicates time points between 100 and 600 ms with significant main effect of effort condition ($P < 0.05$, FDR corrected). **D**: mean interaction wave computed as (hard reward ERP – hard nonreward ERP) – (easy reward ERP – easy nonreward ERP). Shaded region: \pm SE. Red markers indicate time points between 100 and 600 ms with significant interaction effect of reward outcome ($P < 0.05$, FDR corrected). **E**: scalp distribution of reward – nonreward ERPs, irrespective of effort condition, between 184 and 281 ms (1st cluster of significant time points shown in **B**). **F**: scalp distribution of hard effort – easy effort ERPs, irrespective of reward outcome, between 238 and 254 ms (cluster of significant time points shown in **C**). **G**: scalp distribution of interaction wave between 250 and 357 ms (1st cluster of significant time points shown in **D**).

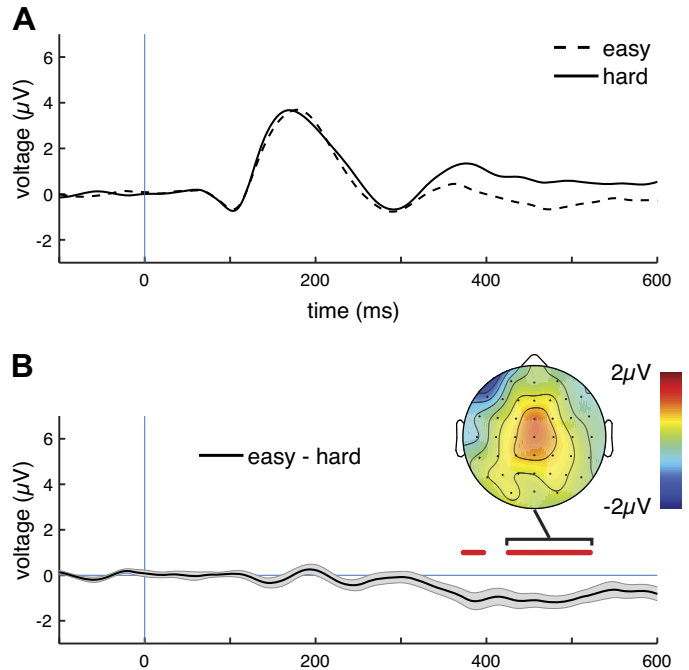


Fig. 4. Event-related potentials (ERPs) elicited by effort condition feedback. **A**: trial-averaged ERPs recorded from electrode FCz aligned to effort condition feedback presentation (0 ms: vertical blue line), selected for physical effort condition (easy or hard). **B**: mean difference waves computed as easy feedback ERP – hard feedback ERP (shaded region: \pm SE). Red markers indicate time points between 100 and 600 ms significantly different from 0 [$P < 0.05$, false discovery rate (FDR) corrected]. *Inset*: scalp distribution of difference wave between 426 and 521 ms.

sons across time points with FDR. We found reliable main effects of reward outcome between 164 and 342 ms after feedback, with larger IC-ERP amplitude for reward compared with nonreward (Fig. 5C; ranges for significant time points: $F = [8.10\ 34.86]$, $P = [0.047\ 0.0011]$, uncorrected $P = [0.012 < 0.0001]$). We found no reliable main effects of effort on the response to reinforcement feedback (Fig. 5E; ranges for all time points 100–600 ms: $F = [0.00\ 4.97]$, uncorrected $P = [1.00\ 0.0416]$). We also found no reliable interaction effects between effort and reward (Fig. 5G; ranges for all time points 100–600 ms: $F = [0.00\ 5.58]$, uncorrected $P = [0.99\ 0.032]$). Figure 6A shows the IC-ERPs elicited by effort feedback for the early ERP components. We observed no difference between the response to easy effort and hard effort feedback (Fig. 6C; ranges for all time points between 100 and 600 ms: $t = [-0.94\ 1.67]$, uncorrected $P = [1.00\ 0.12]$). We observed no differences between the control easy and control hard IC-ERPs (ranges for all time points between 100 and 600 ms: $t = [-1.38\ 2.46]$, uncorrected $P = [0.027\ 0.99]$).

Late ERP component. Figure 5B shows the IC-ERPs elicited by reinforcement feedback for the late ERP components. We found no reliable main effects of reward outcome on the response to reinforcement feedback (Fig. 5D; ranges for all time points 100–600 ms: $F = [0.00\ 16.40]$, uncorrected $P = [1.00\ 0.001]$). We also found no reliable main effects of effort on the response to reinforcement feedback (Fig. 5F; ranges for all time points 100–600 ms: $F = [0.00\ 4.70]$, uncorrected $P = [0.99\ 0.046]$). We did observe an interaction effect between effort and reward, with significant time points between 242 and 549 ms (Fig. 5H; ranges for significant time points: $F = [5.78$

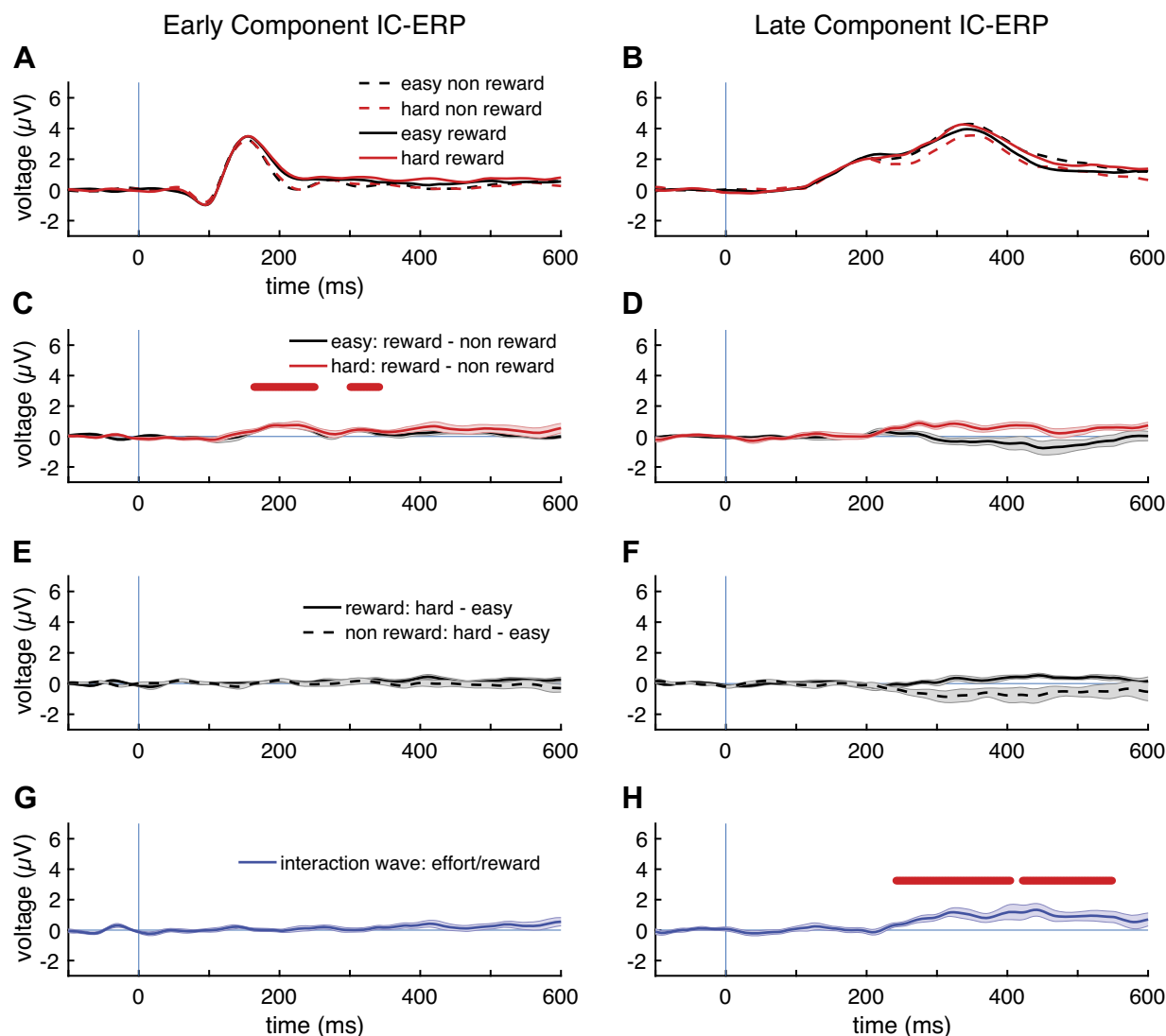


Fig. 5. Independent components to event-related potentials (IC-ERPs) elicited by reinforcement feedback. *A* and *B*: trial-averaged IC-ERPs for the early and late IC-ERP components back-projected to electrode FCz, aligned to reinforcement feedback presentation (0 ms: vertical blue line), selected for reinforcement outcome (reward or nonreward) and the physical effort requirement on that trial (easy or hard). *C* and *D*: mean difference waves computed as reward IC-ERP – nonreward IC-ERP separately for the easy and hard effort conditions (shaded region: \pm SE). Red markers indicate time points between 100 and 600 ms with significant main effect of reward outcome [$P < 0.05$, false discovery rate (FDR) corrected]. *E* and *F*: mean difference waves computed as hard effort IC-ERP – easy effort IC-ERP separately for the reward and nonreward effort conditions (shaded region: \pm SE). *G* and *H*: mean interaction wave computed as (hard reward IC-ERP – hard nonreward IC-ERP) – (easy reward IC-ERP – easy nonreward IC-ERP). Shaded region: \pm SE. Red markers indicate time points between 100 and 600 ms with significant interaction effect of reward outcome ($P < 0.05$, FDR corrected).

24.06], $P = [0.0492\ 0.010]$, uncorrected $P = [0.0287\ 0.0002]$). Figure 6*B* shows the IC-ERPs elicited by effort feedback for the late ERP components. We observed no difference between the response to easy effort and hard effort feedback (Fig. 6*D*; ranges for all time points between 100 and 600 ms: $t = [-2.89\ 1.29]$, uncorrected $P = [1.00\ 0.011]$). We observed no differences between the control easy and control hard IC-ERPs (ranges for all time points between 100 and 600 ms: $t = [-0.27\ 2.01]$, uncorrected $P = [0.97\ 0.06]$).

To directly compare the IC-ERP responses of the early and late components, we also analyzed both components with a single statistical model. We performed three-way repeated-measures ANOVA at each time point between 100 and 600 ms, with factors reward outcome, effort, and IC (early vs. late ERP component). We included the 15 participants for whom both early and late components were identified. P values are cor-

rected for multiple comparisons across time points with FDR. We found reliable main effects of reward outcome between 170 and 342 ms (ranges for significant time points: $F = [7.39\ 29.87]$, $P = [0.0451\ 0.0030]$, uncorrected $P = [0.0167\ 0.0001]$). We found reliable interaction effects between effort and reward between 238 and 535 ms (ranges for significant time points: $F = [6.071\ 14.13]$, $P = [0.0494\ 0.0425]$, uncorrected $P = [0.0273\ 0.0021]$). We found reliable three-way interaction effects between reward, effort, and IC between 299 and 325 ms (ranges for significant time points: $F = [13.75\ 22.24]$, $P = [0.0429\ 0.0192]$, uncorrected $P = [0.0023\ 0.0003]$). The three-way interaction reflects a larger interaction between reward and effort in the late component IC-ERP than the early component IC-ERP. Separate analyses revealed a reliable effect of reward in the early component but not the late component. This was reflected in a trend toward interaction

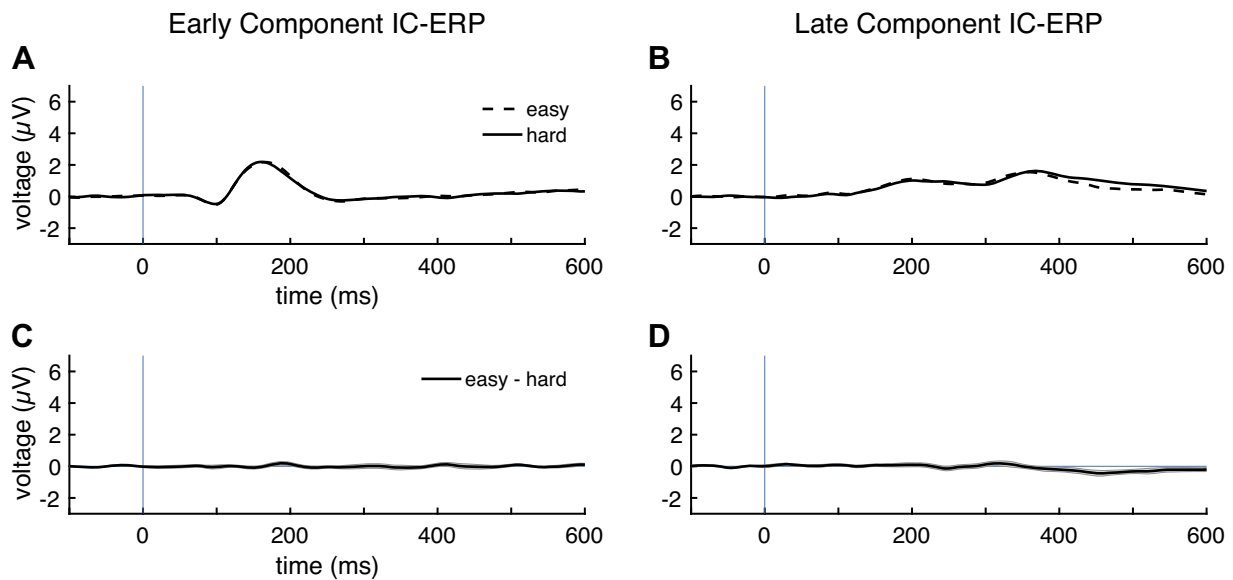


Fig. 6. Independent components to event-related potentials (IC-ERPs) elicited by effort condition feedback. *A* and *B*: trial-averaged IC-ERPs for the early and late IC-ERP components back-projected to electrode FCz, aligned to effort condition feedback presentation (0 ms: vertical blue line), selected for physical effort condition (easy or hard). *C* and *D*: mean difference waves computed as easy feedback IC-ERP – hard feedback IC-ERP (shaded region: \pm SE).

effects between IC and reward outcome, although these effects were not reliable after correction for multiple comparisons across time (ranges for significant time points uncorrected: time = [172 221 ms], F = [4.65 10.65], uncorrected P = [0.0489 0.0057]).

DISCUSSION

Participants were more likely to switch responses after choices that led to hard effort than easy effort, suggesting that they adapted behavior to reduce physical effort in response to uncertain outcomes. At the end of each trial, binary reinforcement feedback indicated whether participants achieved a monetary reward, which depended on precisely producing a target level of EMG activity. Unsurprisingly, reinforcement feedback elicited a robust feedback related negativity/reward positivity (FRN/RP) response, measured as a relative positivity in the ERPs elicited by reward compared to those elicited by nonreward feedback over the medial frontal scalp.

Samplewise analysis revealed interesting temporal dynamics of effort and reward processing in the midfrontal EEG during outcome evaluation. After reinforcement feedback was delivered, an effect of reward outcome first emerged with a latency of 184 ms, which remained significant while an additional main effect of preceding effort emerged at 238 ms. Finally, a sustained interaction between reward and effort first occurred around 250 ms after feedback onset. These dynamics suggest a process whereby upon receiving reward feedback the brain first encodes the immediate reward outcome and subsequently integrates signals related to the preceding effort. This process culminates in an interaction whereby the effect of reward outcome depends on the preceding effort.

The main effect of reward outcome, which is generally the definitive feature of the FRN/RP, occurred with a typical spatial and temporal distribution. The interaction effect showed similar medial frontal scalp topography and substantially overlapping temporal properties, suggesting that it may originate in the same neural generator. However, the interaction effect

persisted upwards of 600 ms. Although the FRN/RP is not typically measured beyond 400 ms, meta-analysis has shown sensitivity in medial frontal ERPs to reward prediction error upwards of 500 ms (Sambrook and Goslin 2015). These long-latency effects may be attributed to the P300, which is modulated by various features of reinforcement processing and can overlap with the FRN/RP (Glazer et al. 2018). We used independent component analysis to separate the contributions of various neural sources to the scalp ERP. We found that the main effect of reward and the interaction between effort and reward reliably load onto separate sets of independent components. We take this as evidence that these effects arise in distinct neural sources, as a single source cannot produce multiple effects with separable scalp projections. However, we do not claim that the early and late IC-ERPs necessarily represent purely isolated single ERP components, as our ICA procedure could fail to separate distinct sources. Although ICA should ameliorate issues due to component overlap, they cannot be verifiably ruled out. Nonetheless, the reward \times effort interaction effect showed latencies and IC loadings consistent with P300 effects.

Although the FRN/RP is reliably sensitive to outcome valence and likelihood aspects of reward prediction error, reward magnitude seems to be coded independently in the P300 (Sambrook and Goslin 2015; Sato et al. 2005; Walsh and Anderson 2012; Wu and Zhou 2009; Yeung and Sanfey 2004). Thus, physical effort may modulate the subjective magnitude of reinforcement outcomes. The P300 elicited by outcome processing has also been shown to be affected by various high-level properties such as motivational salience, temporal waiting cost, and cognitive effort (Glazer et al. 2018; Ma et al. 2014; San Martín 2012; Wang et al. 2014). Multiple variants of the P300 have been reported with medial frontal or posterior scalp distributions, and widespread cortical association networks are implicated including parietal, temporal, and prefrontal regions (Polich 2007; Soltani and Knight 2000).

Although these results show a clear effect of the physical effort associated with an action on the neural processing of subsequent reinforcement outcomes, specific interpretation of the effects depends on the theoretical understanding of the underlying components. An early and influential theory of the FRN/RP described it as a cortical neural correlate of negative reward prediction error (worse-than-expected outcomes), functioning to deter unsuccessful actions (Holroyd et al. 2003; Holroyd and Coles 2002; Miltner et al. 1997; Yasuda et al. 2004). However, multiple converging lines of evidence suggest that some or all of the variance described by the FRN/RP is actually generated by a positive ERP deflection in response to positive reward prediction error (better-than-expected outcomes), functioning to reinforce valuable actions (Becker et al. 2014; Carlson et al. 2011; Eppinger et al. 2008; Foti et al. 2011, 2015; Holroyd et al. 2008; Proudfit 2015; Sambrook and Goslin 2016). These findings resulted in the more recent term “reward positivity” (RP).

It is also possible that the FRN/RP is a bidirectional response to positive and negative outcomes, as are canonical signed reward prediction error signals in midbrain dopamine neurons (Walsh and Anderson 2012). The FRN/RP clearly discriminates outcome valence (good vs. bad). The most fundamental evidence for encoding reward prediction error is an interaction effect of outcome valence and frequency whereby the effect of valence is larger for unexpected outcomes. However, because of the possibility of overlapping ERP components sensitive purely to expectancy, analysis of simple effects cannot disambiguate encoding of positive, negative, or bidirectional reward prediction errors (Sambrook and Goslin 2016). Similarly, the prominent interaction effect of effort and reward observed in the present study shows that effort increases the neural response to reinforcement outcomes; however, the directionality and valence of the effect are ambiguous. An fMRI study by Hernandez Lallement et al. (2014) found that cognitive effort increased neural sensitivity to both reward and loss, with reward sensitivity being modulated in the anterior cingulate and nucleus accumbens and loss sensitivity being increased in the anterior insula.

An increased neural response to nonreward is consistent with economic theories whereby a motor cost would further devalue a nonreward outcome, leading to a larger negative reward prediction error. Although normative economic models of behavior predict that effort costs should devalue reward, it has often been reported in humans and other species that rewards produce stronger reinforcement when they require more effort to obtain (Clement et al. 2000; Inzlicht et al. 2018; Lydall et al. 2010; Zentall 2010). Unfortunately, we were not able to assess such a behavioral interaction in the present study as there was no effect of reward outcome on the binary decisions that participants made on each trial. This was not surprising, as reward outcome was not determined by these decisions but rather by performance on the EMG production task. The binary decisions only determined the effort required, and the reinforcement threshold was controlled to produce approximately equal reward rate in both effort conditions.

After participants produced binary responses, feedback indicated the resulting physical effort condition for the subsequent EMG production portion of the trial. In line with theoretical accounts of the FRN/RP as a temporal difference learning prediction error signal, stimuli that predict aversive

outcomes or economic loss typically elicit FRN/RP responses (Mulligan and Hajcak 2018). Thus, we predicted that effort feedback might elicit an FRN/RP component as a learning signal for effort minimization. However, we observed no FRN/RP modulation when comparing the ERP responses to feedback indicating easy or hard effort trials. Rather, effort modulated the response to reinforcement feedback at the end of the trial. This suggests that physical effort is not immediately treated by the reinforcement learning system as a loss or a punishing stimulus. Rather, effort information can be maintained during the course of an action and incorporated with reward information at the time of outcome evaluation.

We often undertake protracted tasks for which the effort requirements and ultimate payoffs are uncertain. It may not be efficient to punish the value representation of a task every time an unexpected effort is encountered, as the eventual payoff may be well worth the effort. Instead, it may be more efficient to integrate effort over the entire course of an undertaking and evaluate the cost and benefit simultaneously when the final outcome is observed. This process can also support interactions in which the effect of effort depends on reward that is only delivered later. Alternatively, some work suggests that we learn about effort requirements and reward separately and integrate them at the time of decision making (Hauser et al. 2017; Skvortsova et al. 2014). It is likely that economic decision making and learning involves distributed hierarchical computations and that it is possible to observe a distribution of signals with varying dependencies on effort, reward, and integrated utility throughout the brain (Hunt and Hayden 2017).

Some limitations of this study should be noted. Participants adapted their behavior to reduce physical effort, but the behavioral effect of effort was variable and relatively weak. Participants were more likely to switch responses after choices that led to hard effort than easy effort. However, participants often switched responses after easy trials or stayed with responses that produced hard effort: on average, participants switched responses after 46.5% of easy trials and 60% of hard trials. Furthermore, negative coefficients for the effect of effort on switching were estimated for several participants. The relatively weak and highly variable effects of effort are consistent with the notion that although effort is generally treated as a cost that is minimized, in many cases people are undeterred by effort or even purposefully select more effortful options (Eisenberger 1992; Inzlicht et al. 2018). These effects are often attributed to state-dependent learning in which reinforcement outcomes are evaluated relative to the value of the current state. In the present study, variable reinforcement outcomes were only evaluated after effort production and thus may have been more valuable when received after a costly, high-effort action. Other details of the task may have affected effort-related choice. Unlike some previous studies of effort minimization, participants were not instructed to avoid effort. Furthermore, success in the task was not dependent on exerting effort that exceeded an unknown criterion. These features may enhance effort minimization, but they could also conflate effort prediction errors with errors relative to the goals of the task at hand, which are also strongly represented in the ACC (Fu et al. 2019; Krigolson and Holroyd 2007; Swick and Turken 2002; Ullsperger et al. 2014).

Although the excellent temporal resolution offered by EEG proved instrumental in uncovering the dynamics of effort and

reward processing in the brain, it invariably measures a mixture of signals from neurons with different response properties. Kennerley et al. (2011) identified diverse tuning to economic value across ACC, orbitofrontal cortex, and lateral prefrontal cortex, such that many neurons that are selective to value with opposite tunings will cancel out at the population level measured by EEG of fMRI. We report effects with midfrontal scalp topographies. However, EEG measured at the scalp is difficult to localize and can represent mixtures of activity from entirely separate brain regions. Although the FRN/RP is a well-characterized response and convergent lines of evidence suggest a source in the ACC, we did not attempt any source localization. The spatial localization of the measured signals remains speculative.

ACKNOWLEDGMENTS

We thank Anca Puscasu for assistance during data collection. All data and code are available from the authors upon reasonable request.

GRANTS

This work was supported by the Natural Sciences and Engineering Research Council of Canada (Conseil de Recherches en Sciences Naturelles et en Génie du Canada) and the Canadian Institutes of Health Research (Instituts de recherche en santé du Canada).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

D.J.P. conceived and designed research; D.J.P. performed experiments; D.J.P. analyzed data; D.J.P. and P.L.G. interpreted results of experiments; D.J.P. prepared figures; D.J.P. drafted manuscript; D.J.P. and P.L.G. edited and revised manuscript; D.J.P. and P.L.G. approved final version of manuscript.

REFERENCES

- Acunzo DJ, Mackenzie G, van Rossum MC. Systematic biases in early ERP and ERF components as a result of high-pass filtering. *J Neurosci Methods* 209: 212–218, 2012. doi:10.1016/j.jneumeth.2012.06.011.
- Alessandri J, Darcheville JC, Delevoeye-Turrell Y, Zentall TR. Preference for rewards that follow greater effort and greater delay. *Learn Behav* 36: 352–358, 2008. doi:10.3758/LB.36.4.352.
- Amiez C, Joseph JP, Procyk E. Anterior cingulate error-related activity is modulated by predicted reward. *Eur J Neurosci* 21: 3447–3452, 2005. doi:10.1111/j.1460-9568.2005.04170.x.
- Bayer HM, Glimcher PW. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47: 129–141, 2005. doi:10.1016/j.neuron.2005.05.020.
- Becker MP, Nitsch AM, Miltner WH, Straube T. A single-trial estimation of the feedback-related negativity and its relation to BOLD responses in a time-estimation task. *J Neurosci* 34: 3005–3012, 2014. doi:10.1523/JNEUROSCI.3684-13.2014.
- Botvinick MM, Huffstetler S, McGuire JT. Effort discounting in human nucleus accumbens. *Cogn Affect Behav Neurosci* 9: 16–27, 2009. doi:10.3758/CABN.9.1.16.
- Bougrain L, Saavedra C, Ranta R. Finally, what is the best filter for P300 detection? TOBI Workshop III—Tools for Brain-Computer Interaction—2012. Würzburg, Germany. March 2012. <https://hal.inria.fr/hal-00756669/>.
- Carlson JM, Foti D, Mujica-Parodi LR, Harmon-Jones E, Hajcak G. Ventral striatal and medial prefrontal BOLD activation is correlated with reward-related electrocortical activity: a combined ERP and fMRI study. *Neuroimage* 57: 1608–1616, 2011. doi:10.1016/j.neuroimage.2011.05.037.
- Clement TS, Feltus JR, Kaiser DH, Zentall TR. “Work ethic” in pigeons: reward value is directly related to the effort or time required to obtain the reward. *Psychon Bull Rev* 7: 100–106, 2000. doi:10.3758/BF03210727.
- Cohen MX, Ranganath C. Reinforcement learning signals predict future decisions. *J Neurosci* 27: 371–378, 2007. doi:10.1523/JNEUROSCI.4421-06.2007.
- Crosson PL, Walton ME, O’Reilly JX, Behrens TE, Rushworth MF. Effort-based cost-benefit valuation and the human brain. *J Neurosci* 29: 4531–4541, 2009. doi:10.1523/JNEUROSCI.4515-08.2009.
- Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* 134: 9–21, 2004. doi:10.1016/j.jneumeth.2003.10.009.
- Denk F, Walton ME, Jennings KA, Sharp T, Rushworth MF, Bannerman DM. Differential involvement of serotonin and dopamine systems in cost-benefit decisions about delay or effort. *Psychopharmacology (Berl)* 179: 587–596, 2005. doi:10.1007/s00213-004-2059-4.
- Eisenberger R. Learned industriousness. *Psychol Rev* 99: 248–267, 1992. doi:10.1037/0033-295X.99.2.248.
- Emeric EE, Brown JW, Leslie M, Pouget P, Stuphorn V, Schall JD. Performance monitoring local field potentials in the medial frontal cortex of primates: anterior cingulate cortex. *J Neurophysiol* 99: 759–772, 2008. doi:10.1152/jn.00896.2006.
- Eppinger B, Kray J, Mock B, Mecklinger A. Better or worse than expected? Aging, learning, and the ERN. *Neuropsychologia* 46: 521–539, 2008. doi:10.1016/j.neuropsychologia.2007.09.001.
- Foti D, Weinberg A, Bernat EM, Proudfit GH. Anterior cingulate activity to monetary loss and basal ganglia activity to monetary gain uniquely contribute to the feedback negativity. *Clin Neurophysiol* 126: 1338–1347, 2015. doi:10.1016/j.clinph.2014.08.025.
- Foti D, Weinberg A, Dien J, Hajcak G. Event-related potential activity in the basal ganglia differentiates rewards from nonrewards: temporospatial principal components analysis and source localization of the feedback negativity. *Hum Brain Mapp* 32: 2207–2216, 2011. doi:10.1002/hbm.21182.
- Fu Z, Wu DA, Ross I, Chung JM, Mamelak AN, Adolphs R, Rutishauser U. Single-neuron correlates of error monitoring and post-error adjustments in human medial frontal cortex. *Neuron* 101: 165–177.e5, 2019. doi:10.1016/j.neuron.2018.11.016.
- Gehring WJ, Willoughby AR. The medial frontal cortex and the rapid processing of monetary gains and losses. *Science* 295: 2279–2282, 2002. doi:10.1126/science.1066893.
- Gheza D, De Raedt R, Baeken C, Pourtois G. Integration of reward with cost anticipation during performance monitoring revealed by ERPs and EEG spectral perturbations. *Neuroimage* 173: 153–164, 2018. doi:10.1016/j.neuroimage.2018.02.049.
- Gläscher J, Daw N, Dayan P, O’Doherty JP. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron* 66: 585–595, 2010. doi:10.1016/j.neuron.2010.04.016.
- Glazer JE, Kelley NJ, Pornpattananangkul N, Mittal VA, Nusslock R. Beyond the FRN: Broadening the time-course of EEG and ERP components implicated in reward processing. *Int J Psychophysiol* 132: 184–202, 2018. doi:10.1016/j.ijpsycho.2018.02.002.
- Graybiel AM. Habits, rituals, and the evaluative brain. *Annu Rev Neurosci* 31: 359–387, 2008. doi:10.1146/annurev.neuro.29.051605.112851.
- Hartmann MN, Hager OM, Tobler PN, Kaiser S. Parabolic discounting of monetary rewards by physical effort. *Behav Processes* 100: 192–196, 2013. doi:10.1016/j.beproc.2013.09.014.
- Hassall CD, Hajcak G, Krigolson OE. The importance of agency in human reward processing. *Cogn Affect Behav Neurosci* 19: 1458–1466, 2019. doi:10.3758/s13415-019-00730-2.
- Hauser TU, Eldar E, Dolan RJ. Separate mesocortical and mesolimbic pathways encode effort and reward learning signals. *Proc Natl Acad Sci USA* 114: E7395–E7404, 2017. doi:10.1073/pnas.1705643114.
- Hauser TU, Iannaccone R, Stämpfli P, Drechsler R, Brandeis D, Walitza S, Brem S. The feedback-related negativity (FRN) revisited: new insights into the localization, meaning and network organization. *Neuroimage* 84: 159–168, 2014. doi:10.1016/j.neuroimage.2013.08.028.
- Hernandez Lallement J, Kuss K, Trautner P, Weber B, Falk A, Fliessbach K. Effort increases sensitivity to reward and loss magnitude in the human brain. *Soc Cogn Affect Neurosci* 9: 342–349, 2014. doi:10.1093/scan/nss147.
- Holinger DP, Hill SA, Martin DL, Faux SF, Ives JR, Schomer DL. Reappraisal of filter effects on P300 voltage and latency. *J Clin Neurophysiol* 17: 331–335, 2000. doi:10.1097/00004691-200005000-00011.
- Holroyd CB, Coles MG. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol Rev* 109: 679–709, 2002. doi:10.1037/0033-295X.109.4.679.

- Holroyd CB, Krigolson OE.** Reward prediction error signals associated with a modified time estimation task. *Psychophysiology* 44: 913–917, 2007. doi:10.1111/j.1469-8986.2007.00561.x.
- Holroyd CB, Nieuwenhuis S, Yeung N, Cohen JD.** Errors in reward prediction are reflected in the event-related brain potential. *Neuroreport* 14: 2481–2484, 2003. doi:10.1097/00001756-200312190-00037.
- Holroyd CB, Pakzad-Vaezi KL, Krigolson OE.** The feedback correct-related positivity: sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology* 45: 688–697, 2008. doi:10.1111/j.1469-8986.2008.00668.x.
- Hosking JG, Floresco SB, Winstanley CA.** Dopamine antagonism decreases willingness to expend physical, but not cognitive, effort: a comparison of two rodent cost/benefit decision-making tasks. *Neuropsychopharmacology* 40: 1005–1015, 2015. doi:10.1038/npp.2014.285.
- Hunt LT, Hayden BY.** A distributed, hierarchical and recurrent framework for reward-based choice. *Nat Rev Neurosci* 18: 172–182, 2017. doi:10.1038/nrn.2017.7.
- Inzlicht M, Shenhav A, Olivola CY.** The effort paradox: effort is both costly and valued. *Trends Cogn Sci* 22: 337–349, 2018. doi:10.1016/j.tics.2018.01.007.
- Ito S, Stuphorn V, Brown JW, Schall JD.** Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science* 302: 120–122, 2003. doi:10.1126/science.1087847.
- Kennerley SW, Behrens TE, Wallis JD.** Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nat Neurosci* 14: 1581–1589, 2011. doi:10.1038/nn.2961.
- Kennerley SW, Dahmubed AF, Lara AH, Wallis JD.** Neurons in the frontal lobe encode the value of multiple decision variables. *J Cogn Neurosci* 21: 1162–1178, 2009. doi:10.1162/jocn.2009.21100.
- Klein-Flügge MC, Kennerley SW, Friston K, Bestmann S.** Neural signatures of value comparison in human cingulate cortex during decisions requiring an effort-reward trade-off. *J Neurosci* 36: 10002–10015, 2016. doi:10.1523/JNEUROSCI.0292-16.2016.
- Krigolson OE, Holroyd CB.** Hierarchical error processing: different errors, different systems. *Brain Res* 1155: 70–80, 2007. doi:10.1016/j.brainres.2007.04.024.
- Kurniawan IT, Guitart-Masip M, Dolan RJ.** Dopamine and effort-based decision making. *Front Neurosci* 5: 81, 2011. doi:10.3389/fnins.2011.00081.
- Kurniawan IT, Seymour B, Talmi D, Yoshida W, Chater N, Dolan RJ.** Choosing to make an effort: the role of striatum in signaling physical effort of a chosen action. *J Neurophysiol* 104: 313–321, 2010. doi:10.1152/jn.00027.2010.
- Lydall ES, Gilmour G, Dwyer DM.** Rats place greater value on rewards produced by high effort: an animal analogue of the “effort justification” effect. *J Exp Soc Psychol* 46: 1134–1137, 2010. doi:10.1016/j.jesp.2010.05.011.
- Ma Q, Meng L, Wang L, Shen Q.** I endeavor to make it: effort increases valuation of subsequent monetary reward. *Behav Brain Res* 261: 1–7, 2014. doi:10.1016/j.bbr.2013.11.045.
- Martin LE, Potts GF.** Medial frontal event-related potentials and reward prediction: do responses matter? *Brain Cogn* 77: 128–134, 2011. doi:10.1016/j.bandc.2011.04.001.
- Mathewson KJ, Dywan J, Snyder PJ, Tays WJ, Segalowitz SJ.** Aging and electrocortical response to error feedback during a spatial learning task. *Psychophysiology* 45: 936–948, 2008. doi:10.1111/j.1469-8986.2008.00699.x.
- Miltner WH, Braun CH, Coles MG.** Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a “generic” neural system for error detection. *J Cogn Neurosci* 9: 788–798, 1997. doi:10.1162/jocn.1997.9.6.788.
- Morel P, Ulbrich P, Gail A.** What makes a reach movement effortful? Physical effort discounting supports common minimization principles in decision making and motor control. *PLoS Biol* 15: e2001323, 2017. doi:10.1371/journal.pbio.2001323.
- Mullen T.** *CleanLine EEGLAB Plugin*. San Diego, CA: Neuroimaging Informatics Tools and Resources Clearinghouse (NITRC), 2012.
- Mulligan EM, Hajcak G.** The electrocortical response to rewarding and aversive feedback: The reward positivity does not reflect salience in simple gambling tasks. *Int J Psychophysiol* 132: 262–267, 2018. doi:10.1016/j.ijpsycho.2017.11.015.
- Nieuwenhuis S, Slagter HA, von Geusau NJ, Heslenfeld DJ, Holroyd CB.** Knowing good from bad: differential activation of human cortical areas by positive and negative outcomes. *Eur J Neurosci* 21: 3161–3168, 2005. doi:10.1111/j.1460-9568.2005.04152.x.
- Onton J, Makeig S.** Information-based modeling of event-related brain dynamics. *Prog Brain Res* 159: 99–120, 2006. doi:10.1016/S0079-6123(06)59007-7.
- Palidis DJ, Cashaback JG, Gribble PL.** Neural signatures of reward and sensory error feedback processing in motor learning. *J Neurophysiol* 121: 1561–1574, 2019. doi:10.1152/jn.00792.2018.
- Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD.** Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* 442: 1042–1045, 2006. doi:10.1038/nature05051.
- Pfabigan DM, Alexopoulos J, Bauer H, Sailer U.** Manipulation of feedback expectancy and valence induces negative and positive reward prediction error signals manifest in event-related brain potentials. *Psychophysiology* 48: 656–664, 2011. doi:10.1111/j.1469-8986.2010.01136.x.
- Polich J.** Updating P300: an integrative theory of P3a and P3b. *Clin Neurophysiol* 118: 2128–2148, 2007. doi:10.1016/j.clinph.2007.04.019.
- Porter BS, Hillman KL, Bilkey DK.** Anterior cingulate cortex encoding of effortful behavior. *J Neurophysiol* 121: 701–714, 2019. doi:10.1152/jn.00654.2018.
- Prévost C, Pessiglione M, Météreau E, Cléry-Melin ML, Dreher JC.** Separate valuation subsystems for delay and effort decision costs. *J Neurosci* 30: 14080–14090, 2010. doi:10.1523/JNEUROSCI.2752-10.2010.
- Proudfit GH.** The reward positivity: from basic research on reward to a biomarker for depression. *Psychophysiology* 52: 449–459, 2015. doi:10.1111/psyp.12370.
- Puig MV, Miller EK.** The role of prefrontal dopamine D1 receptors in the neural mechanisms of associative learning. *Neuron* 74: 874–886, 2012. doi:10.1016/j.neuron.2012.04.018.
- Rangel A, Hare T.** Neural computations associated with goal-directed choice. *Curr Opin Neurobiol* 20: 262–270, 2010. doi:10.1016/j.conb.2010.03.001.
- Rudebeck PH, Behrens TE, Kennerley SW, Baxter MG, Buckley MJ, Walton ME, Rushworth MF.** Frontal cortex subregions play distinct roles in choices between actions and stimuli. *J Neurosci* 28: 13775–13785, 2008. doi:10.1523/JNEUROSCI.3541-08.2008.
- Rudebeck PH, Walton ME, Smyth AN, Bannerman DM, Rushworth MF.** Separate neural pathways process different decision costs. *Nat Neurosci* 9: 1161–1168, 2006. doi:10.1038/nn1756.
- Salamone JD, Correa M, Farrar A, Mingote SM.** Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology (Berl)* 191: 461–482, 2007. doi:10.1007/s00213-006-0668-9.
- Salamone JD, Correa M, Mingote S, Weber SM.** Nucleus accumbens dopamine and the regulation of effort in food-seeking behavior: implications for studies of natural motivation, psychiatry, and drug abuse. *J Pharmacol Exp Ther* 305: 1–8, 2003. doi:10.1124/jpet.102.035063.
- Sambrook TD, Goslin J.** A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychol Bull* 141: 213–235, 2015. doi:10.1037/bul0000006.
- Sambrook TD, Goslin J.** Principal components analysis of reward prediction errors in a reinforcement learning task. *Neuroimage* 124: 276–286, 2016. doi:10.1016/j.neuroimage.2015.07.032.
- San Martín R.** Event-related potential studies of outcome processing and feedback-guided learning. *Front Hum Neurosci* 6: 304, 2012. doi:10.3389/fnhum.2012.00304.
- Sato A, Yasuda A, Ohira H, Miyawaki K, Nishikawa M, Kumano H, Kuboki T.** Effects of value and reward magnitude on feedback negativity and P300. *Neuroreport* 16: 407–411, 2005. doi:10.1097/00001756-200503150-00020.
- Schevernels H, Krebs RM, Santens P, Woldorff MG, Boehler CN.** Task preparation processes related to reward prediction precede those related to task-difficulty expectation. *Neuroimage* 84: 639–647, 2014. doi:10.1016/j.neuroimage.2013.09.039.
- Schultz W.** Behavioral theories and the neurophysiology of reward. *Annu Rev Psychol* 57: 87–115, 2006. doi:10.1146/annurev.psych.56.091103.070229.
- Schweimer J, Saft S, Hauber W.** Involvement of catecholamine neurotransmission in the rat anterior cingulate in effort-related decision making. *Behav Neurosci* 119: 1687–1692, 2005. doi:10.1037/0735-7044.119.6.1687.
- Selinger JC, O’Connor SM, Wong JD, Donelan JM.** Humans can continuously optimize energetic cost during walking. *Curr Biol* 25: 2452–2456, 2015. doi:10.1016/j.cub.2015.08.016.
- Seo H, Lee D.** Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J Neurosci* 27: 8366–8377, 2007. doi:10.1523/JNEUROSCI.2369-07.2007.
- Shadmehr R, Huang HJ, Ahmed AA.** A representation of effort in decision-making and motor control. *Curr Biol* 26: 1929–1934, 2016. doi:10.1016/j.cub.2016.05.065.

- Skvortsova V, Palminteri S, Pessiglione M.** Learning to minimize efforts versus maximizing rewards: computational principles and neural correlates. *J Neurosci* 34: 15621–15630, 2014. doi:10.1523/JNEUROSCI.1350-14.2014.
- Soltani M, Knight RT.** Neural origins of the P300. *Crit Rev Neurobiol* 14: 199–224, 2000. doi:10.1615/CritRevNeurobiol.v14.i3-4.20.
- Swick D, Turken AU.** Dissociation between conflict detection and error monitoring in the human anterior cingulate cortex. *Proc Natl Acad Sci USA* 99: 16354–16359, 2002. doi:10.1073/pnas.252521499.
- Tanner D, Morgan-Short K, Luck SJ.** How inappropriate high-pass filters can produce artifactual effects and incorrect conclusions in ERP studies of language and cognition. *Psychophysiology* 52: 997–1009, 2015. doi:10.1111/psyp.12437.
- Ullsperger M, Danielmeier C, Jocham G.** Neurophysiology of performance monitoring and adaptive behavior. *Physiol Rev* 94: 35–79, 2014. doi:10.1152/physrev.00041.2012.
- Vezoli J, Procyk E.** Frontal feedback-related potentials in nonhuman primates: modulation during learning and under haloperidol. *J Neurosci* 29: 15675–15683, 2009. doi:10.1523/JNEUROSCI.4943-09.2009.
- Walsh MM, Anderson JR.** Learning from experience: event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neurosci Biobehav Rev* 36: 1870–1884, 2012. doi:10.1016/j.neubiorev.2012.05.008.
- Walton ME, Bannerman DM, Alterescu K, Rushworth MFS.** Functional specialization within medial frontal cortex of the anterior cingulate for evaluating effort-related decisions. *J Neurosci* 23: 6475–6479, 2003. doi:10.1523/JNEUROSCI.23-16-06475.2003.
- Walton ME, Kennerley SW, Bannerman DM, Phillips PE, Rushworth MF.** Weighing up the benefits of work: behavioral and neural analyses of effort-related decision making. *Neural Netw* 19: 1302–1314, 2006. doi:10.1016/j.neunet.2006.03.005.
- Wang J, Chen J, Lei Y, Li P.** P300, not feedback error-related negativity, manifests the waiting cost of receiving reward information. *Neuroreport* 25: 1044–1048, 2014. doi:10.1097/WNR.0000000000000226.
- Wang L, Zheng J, Meng L.** Effort provides its own reward: endeavors reinforce subjective expectation and evaluation of task performance. *Exp Brain Res* 235: 1107–1118, 2017. doi:10.1007/s00221-017-4873-z.
- Warren CM, Hyman JM, Seamans JK, Holroyd CB.** Feedback-related negativity observed in rodent anterior cingulate cortex. *J Physiol Paris* 109: 87–94, 2015. doi:10.1016/j.jphysparis.2014.08.008.
- Williams ZM, Bush G, Rauch SL, Cosgrove GR, Eskandar EN.** Human anterior cingulate neurons and the integration of monetary reward with motor responses. *Nat Neurosci* 7: 1370–1375, 2004. [Erratum in *Nat Neurosci* 8: 121, 2005.] doi:10.1038/nn1354.
- Winkler I, Debener S, Müller KR, Tangermann M.** On the influence of high-pass filtering on ICA-based artifact reduction in EEG-ERP. *Conf Proc IEEE Eng Med Biol Soc* 2015: 4101–4105, 2015. doi:10.1109/EMBC.2015.7319296.
- Wu Y, Zhou X.** The P300 and reward valence, magnitude, and expectancy in outcome evaluation. *Brain Res* 1286: 114–122, 2009. doi:10.1016/j.brainres.2009.06.032.
- Yasuda A, Sato A, Miyawaki K, Kumano H, Kuboki T.** Error-related negativity reflects detection of negative reward prediction error. *Neuroreport* 15: 2561–2565, 2004. doi:10.1097/00001756-200411150-00027.
- Yeung N, Holroyd CB, Cohen JD.** ERP correlates of feedback and reward processing in the presence and absence of response choice. *Cereb Cortex* 15: 535–544, 2005. doi:10.1093/cercor/bhh153.
- Yeung N, Sanfey AG.** Independent coding of reward magnitude and valence in the human brain. *J Neurosci* 24: 6258–6264, 2004. doi:10.1523/JNEUROSCI.4537-03.2004.
- Zentall TR.** Justification of effort by humans and pigeons: cognitive dissonance or contrast? *Curr Dir Psychol Sci* 19: 296–300, 2010. doi:10.1177/0963721410383381.
- Zink CF, Pagnoni G, Martin-Skurski ME, Chappelow JC, Berns GS.** Human striatal responses to monetary reward depend on saliency. *Neuron* 42: 509–517, 2004. doi:10.1016/S0896-6273(04)00183-7.