

BRIEF REPORT

Medial frontal negativity reflects learning from positive feedback

JURJEN VAN DER HELDEN^{a,b} AND MAARTEN A.S. BOKSEM^{a,c}

^aRadboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands

^bDepartment of Cognitive Psychology and Ergonomics, University of Twente, Enschede, The Netherlands

^cRotterdam School of Management, Erasmus University, Rotterdam, The Netherlands

Abstract

The ability to learn from the consequences of our actions is crucial for adaptive goal-directed behavior. We learn to avoid actions that lead to unfavorable outcomes and pursue actions that lead to desirable results. By recording event-related potentials (ERPs), we show that neural reinforcement learning signals associated with positive outcomes are predictive of subsequent learning of a sequence of motor actions: Positive feedback to a response that was later correctly repeated was associated with a larger Medial Frontal Negativity (MFN) compared to when it was not correctly repeated on a subsequent encounter. This finding adds to recent evidence suggesting that the function of the Anterior Cingulate Cortex is to establish associations between actions and their outcomes, both positive and negative.

It has been shown that Anterior Cingulate Cortex (ACC) activity is reflected by a negative event-related potential (ERP) on the scalp (Dehaene, Posner, & Tucker, 1994; Miltner, Braun, & Coles, 1997). Such negative amplitude shifts are prominent when subjects commit response errors (Error-Related Negativity) or when they receive negative feedback (Feedback-Related Negativity). It has been suggested that these ERP components are associated with common neural processes (Nieuwenhuis, Holroyd, Mol, & Coles, 2004), and, for convenience, we will refer to them as Medial Frontal Negativity (MFN). A prominent theory holds that ACC activity in response to such negative events reflects a reinforcement learning (RL) process by which the organism learns to associate actions with negative outcomes (Holroyd & Coles, 2002). However, recent studies using electrophysiological recordings of single ACC cells in monkeys and also human imaging studies show that the ACC can also be activated by positive action outcomes, which suggests that it supports a more general mechanism to establish action–outcome associations, by also learning from positive outcomes (see Rushworth, Noonan, Boorman, Walton, & Behrens, 2011). Indeed, the ratio of error-responding to reward-responding cells may be approximately 5:4 (Quilodran, Rothé, & Procyk, 2008), and some ACC neurons respond to both types of feedback. However, although it has been shown that learning to associate actions with negative outcomes is reflected by MFN amplitude (Van der Helden, Boksem, & Blom, 2010), evidence that learning such action–outcome associations by means of positive feedback is currently lacking.

In the present study we tested whether the MFN to positive feedback in a trial-and-error learning task is predictive of the

success of future actions. By contrasting the MFN elicited by positive feedback to an action that is subsequently correctly repeated with those that are not correctly repeated, we were able to establish whether the MFN elicited by positive feedback is also predictive of the success of future performance, providing evidence that the ACC is involved in updating action–reward contingencies through both positive and negative feedback.

Methods

Participants

The study was approved by a faculty internal review board of the University of Twente. A total of 44 participants (15 men; mean age 19.9 years old) were recruited from the university population and received course credits for their participation. To be able to contrast the MFNs elicited by positive feedback that were followed by correct performance (good learning) with those followed by incorrect performance (bad learning), a good number of trials in both of these conditions was required to achieve acceptable signal-to-noise ratios. However, such bad learning trials following positive feedback are quite rare. Therefore, we included only those participants that produced at least 15 bad positive RL learning errors in the analysis presented here ($n = 15$; 2 men; mean age 20.2 ± 2.8 years old). Written informed consent was obtained prior to the experiment.

Stimuli

Stimuli were presented on a black background. Four white squares ($1.4^\circ \times 1.4^\circ$ each) that represented the four response buttons were presented on screen for the entire duration of the task. Each trial began with the presentation of the number of the current item in the sequence in white on fixation. Participants were asked to choose between four response buttons after presentation of the current item number. If participants chose, for example, to press the leftmost

Address correspondence to: Jurjen van der Helden, Radboud University Nijmegen, Donders Institute for Brain, Cognition, and Behaviour, P.O. Box 9104, 6500 HE Nijmegen, The Netherlands. E-mail: JurjenvdHelden@gmail.com

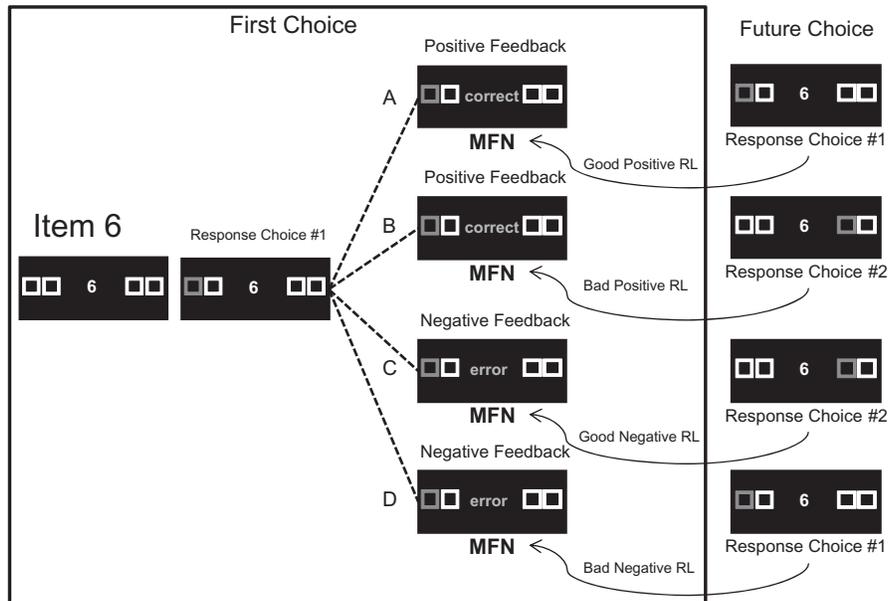


Figure 1. Labeling procedure. Left: choice (leftmost button) on Item 6 in the sequence followed by positive feedback (a and b) and by negative feedback (c and d). The future choice of the subject on Item 6 (right) may involve the same or another response. If participants choose the same response following positive feedback (a) we label the MFN as “Good Positive Reinforcement Learning” (GPRL). If, however, the participant erroneously chooses another response subsequently (b), this is labeled as Bad Positive RL (BPRL). Similarly, if participants choose another response on the next occasion than the choice that resulted in negative feedback (c), this is labeled as Good Negative RL (GNRL), but when they erroneously repeat the choice that received negative feedback (d), this is labeled as Bad Negative RL (BNRL).

button (pressed with their left middle finger), the leftmost box on the screen instantly changed from white to blue. One thousand milliseconds after the participant’s response, feedback was presented visually on fixation. If the participant’s choice was correct, he or she received positive feedback (i.e., “Correct”) presented in green; if the participant’s choice was incorrect, he or she received negative feedback (i.e., “Error”) presented in red. Response choices exceeding the time limit (1500 ms) were indicated by visual feedback “Too late” (in blue font) on fixation. The visual feedback remained on screen for 1000 ms, until the start of the following trial. Participants were instructed to fixate on the central location in the display, at the location where feedback was presented for the entire trial.

Procedure

Stimuli were presented with the E-Prime package (version 1.2., Psychology Software Tools, Inc., Pittsburgh, PA; www.psnet.com/) on a 17-in. monitor, and responses were collected through an E-Prime compatible PST Serial Response Box.

Participants were asked to learn a sequence of 12 button presses. They were instructed to learn this sequence by trial and error and that for every item only one option was correct. When subjects chose the correct response, positive feedback was always provided and the task proceeded with the next item in the sequence. If they chose an incorrect response or if they did not respond in time, the sequence restarted at Item 1. Participants thus only successfully completed a sequence if all 12 items were responded to correctly. When a sequence was completed, participants received overall feedback and a short break (30 s) before they proceeded with the next sequence. In total, participants completed a maximum of 10 and a minimum of 6 sequences.

Unbeknownst to the subjects, for each item in the sequence, the number of response choices that were considered incorrect before a choice would be positively reinforced and considered as correct was manipulated. In every sequence there were three items that were positively reinforced on the first, second, third, and fourth encounter of that item. The order in which these items occurred was randomized for every sequence. This ensured that there were no differences in performance between subjects and sequences due to better guessing.

We distinguished between four types of feedback (see Figure 1). First, we labeled positive feedback that was followed by the same response choice on that item as Good Positive Reinforcement Learning (GPRL), as this indicates that the subject learned from the positive feedback. Second, we labeled negative feedback that was followed by a different response choice on that same item in the sequence as Good Negative RL (GNRL), as this indicates that the participant has learned from the negative feedback and tried a response not chosen before. Third, we labeled positive feedback that was followed by an alternative response choice on the same item as Bad Positive RL (BPRL). Because positive feedback is only informative the first time an item is performed correctly, only these trials were included in the GPRL and BPRL ERPs. Finally, we labeled negative feedback that was followed by a response choice that subjects had tried before on that same item as Bad Negative RL (BNRL).

Electroencephalogram (EEG) Recording and Data Reduction

EEG was recorded from 61 standard channels, using Ag/AgCl ring electrodes mounted on an electrocap (EasyCap), with a forehead ground and an on-line average reference. The vertical and horizontal electro-oculograms were measured from electrodes above and

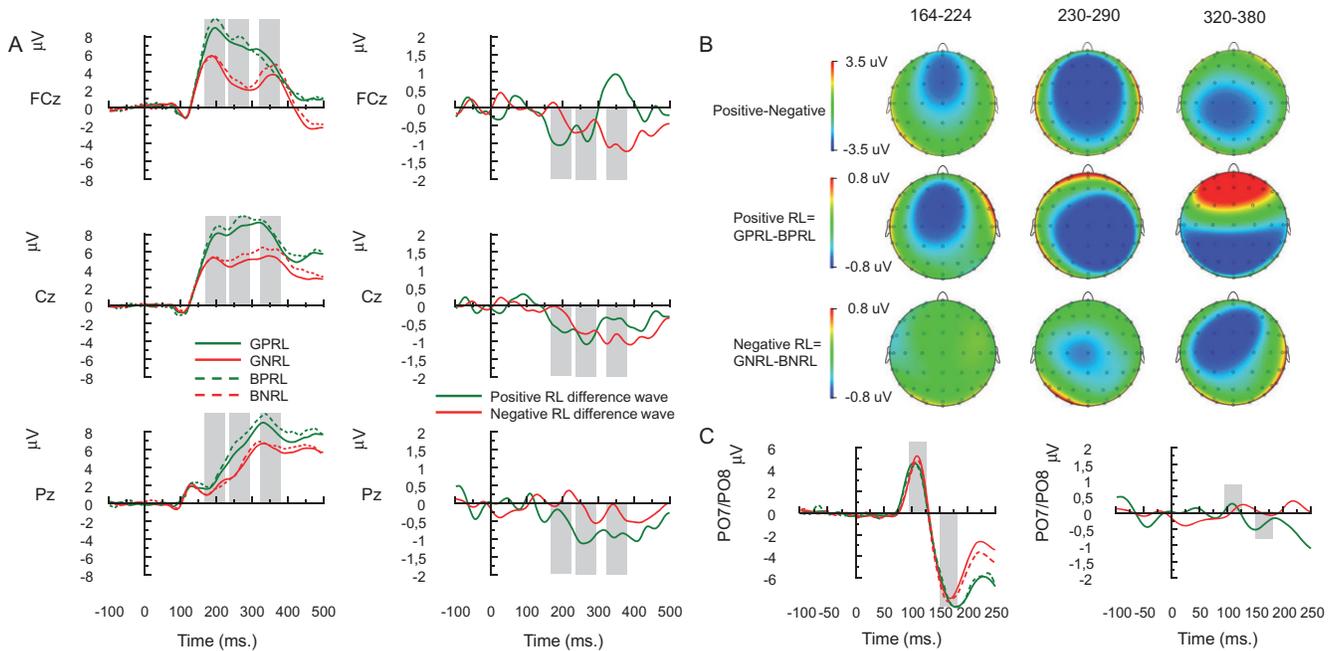


Figure 2. a (left): MFN waveforms associated with the different feedback–performance contingencies at channels FCz, Cz, and Pz. The green solid lines represent Good Positive Reinforcement Learning (GPRL), the green dashed lines Bad Positive RL (BPRL), Good Negative RL are represented in red solid lines (GNRL), and the red dashed lines represent Bad Negative RL (BNRL). Right: The Positive RL difference wave (GPRL – BPRL) is shown as the solid green line, and Negative RL difference wave (GNRL – BNRL) is shown as the red solid line. GPRL elicited a larger negativity than BPRL, and similarly, GNRL elicited a larger MFN than BNRL. The statistically tested epochs are highlighted in gray. b: Topographical distributions of the difference waves of these samples (shown from left to right). Top graphs: topographical distribution of the difference wave of Positive and Negative Feedback is shown. Middle graphs: Topographical distributions of the Positive RL difference wave. Lower graphs: Topographical distribution of the Negative RL difference waves. Note the larger scale (3.5 to $-3.5 \mu\text{V}$) for the topographical distribution Feedback Type effect in the top graph than the scale in the middle and lower graphs (0.8 to $-0.8 \mu\text{V}$). c: Early visual components pooled for PO7 and PO8. Left: ERPs of GPRL, GNRL, BPRL, and BNRL. Right: Positive RL (GPRL-BPRL) and Negative RL (GNRL-BNRL) difference waves. Labeling of the curves is the same as in panel a. Statistical tests on P1 and N1 peaks (represented by the gray bars) showed no effects of Learning Effectiveness.

below the left eye and from the outer canthi of both eyes, respectively. Electrode impedance was kept below $5 \text{ k}\Omega$. Signals were passed through a BrainAmp amplifier (Brain Products GmbH, Munich, Germany; www.brainproducts.com) and recorded online at a sample rate of 500 Hz and off-line filtered with a 200-Hz low-pass filter and a notch filter of 50 Hz.

EEG segments containing artifacts ($\pm 120 \mu\text{V}$) and eye movements ($\pm 120 \mu\text{V}$) were rejected. In addition, ocular artifacts were further corrected with the procedure described by Gratton, Coles, and Donchin (1983). ERPs of feedback signals associated with GPRL (mean number of trials = 85.5, $SD = 15.2$), GNRL ($M = 147.3$, $SD = 25.2$), BPRL ($M = 18.5$, $SD = 4.0$), and BNRL ($M = 34$, $SD = 11.4$) were analyzed and averaged separately. A baseline voltage averaged over the 100-ms interval preceding the onset of the feedback signal was subtracted from the averages.

Data Analysis

We analyzed the data from electrodes FCz, Cz, and Pz. Visual inspection of Figure 2A shows that we can distinguish three components in feedback ERPs: an early component (164–224 ms) that has its maximum at FCz, an intermediate component (230–290 ms) that was maximal at Cz, and a late component (320–380 ms), maximal at Pz. These phases appear to roughly correspond to the P2, N2, and P3 ERP components, respectively (Luck, 2005). Data from these sites at these latencies were submitted to statistical analyses. We submitted averaged ERP amplitudes of these time

windows in a repeated measures 2×2 GLM design with the factors Feedback Type (following positive or negative feedback) and Learning Effectiveness (good or bad subsequent learning). Because a potential confound that bad learning could result from relative inattention or other perceptual factors, we additionally tested the posterior P1 and N1 that are modulated by visual attention (Luck, 2005). The pooled activity of PO7 and PO8 showed the P1 and N1 peaks between 94 and 124 ms and 150 and 180 ms respectively, which were analyzed using the same GLM as with the MFN.

Results

Performance

As a result of our manipulation that controlled the number of erroneous choices before a choice is correct, the lowest number of errors that subjects could make on a particular sequence was 18. In addition to these 18 errors, participants produced 19.8 errors ($SD = 7.5$) on average per sequence. These errors consisted of failures to refrain from repeating negatively reinforced responses (negative RL failures; $M = 4.1$, $SD = 1.9$) and of failures in reproducing positively reinforced responses ($M = 15.6$, $SD = 6.6$). Of these latter learning failures, only a small proportion ($M = 2.2$, $SD = 0.7$) consisted of true positive RL failures (that is, on average 2.2 times per sequence subjects chose a response other than the response to which they received positive feedback on the previous encounter of that item); the remaining failures to reproduce posi-

tively reinforced responses involved errors that were made after the subject had already responded correctly to that item repeatedly ($M = 13.4$, $SD = 6.0$). Because these errors are difficult to relate to learning processes, they will not be considered further.

ERPs

ERPs for positive and negative feedback, plotted separately by subsequent learning, are shown in Figure 2A. With the exception of the 320–380-ms epoch at FCz, Feedback Type showed more negative amplitudes following negative compared to positive feedback on every electrode in every epoch, $F(1,14) > 9.9$, $p < .01$. Learning Effectiveness showed significant effects in the intermediate epoch at FCz, $F(1,14) = 5.0$, $p < .05$, Cz, $F(1,14) = 12.6$, $p < .005$, and Pz, $F(1,14) = 4.6$, $p < .05$. In the early epoch this effect was marginally significant at FCz, $F(1,14) = 4.2$, $p = .06$, and at Cz, $F(1,14) = 4.0$, $p = .066$. In the late epoch, Learning Effectiveness showed marginally significant effects at Cz, $F(1,14) = 3.9$, $p = .067$, and at Pz, $F(1,14) = 4.2$, $p = .059$. In addition, Feedback Type and Learning Effectiveness interacted at FCz in the late epoch, $F(1,14) = 6.0$, $p \leq .05$. No other significant interactions were found between Feedback Type and Learning Effectiveness, indicating that comparable negative shifts in the ERPs for positive and negative feedback are associated with subsequent learning.

Learning Effectiveness from positive feedback was marginally significant at FCz, $t(14) = -2.0$, $p = .069$, and at Cz, $t(14) = -2.1$, $p = .054$, in the early epoch and significant at Cz and Pz in the intermediate epoch, $t(14) = -2.4$, $p < .05$, and $t(14) = -2.7$, $p < .05$, respectively (see Figure 2B).

Learning Effectiveness from negative feedback was significant at Cz, $t(14) = -2.3$, $p < .05$, and marginally significant at FCz, $t(14) = -1.9$, $p = .073$, in the intermediate epoch, replicating previous findings (Van der Helden et al., 2010). In the late epoch, this effect was significant at FCz, $t(14) = -2.7$, $p < .05$, and at Cz, $t(14) = -2.5$, $p < .05$. This late effect of Learning Effectiveness from negative feedback, in the absence of a Learning Effectiveness effect from positive feedback at FCz accounts for the interaction effect observed in this epoch (see Figure 2A). No effects of Learning Effectiveness from negative feedback were observed at Pz.

To control for effects of attention that could potentially confound our results, we also analyzed the early visual attention components (P1 and N1; see Figure 2C). No significant effects were found on the P1 or N1, suggesting that ERP effects associated with learning from feedback were not due to potential differences in visual attention.

Discussion

Consistent with the Reinforcement Learning theory of the MFN (Holroyd & Coles, 2002), we found that larger MFNs elicited by

feedback are predictive of successful learning. When subjects received negative feedback, larger MFNs were observed when they would not choose the associated response option again as compared to when they would choose the same incorrect option again on the subsequent trial, replicating earlier findings (Van der Helden et al., 2010). Importantly, however, we show here for the first time that the MFN following positive feedback is similarly predictive of future learning: Positive feedback to an action that was correctly repeated on the next encounter with that item was associated with a larger MFN than positive feedback followed by a subsequent incorrect response. Indeed, midbrain dopaminergic cells that project to the ACC have been found to produce dips in firing rate following errors or erroneous feedback, but increases in firing rate following rewards (e.g., Schultz, 2002), which are predicted to result in activation of the ACC (Holroyd, & Coles, 2002). These findings provide support for the notion that the ACC, the putative source of the MFN, supports a general mechanism for the learning of action–outcome associations, irrespective of whether these outcomes are positive or negative. In addition, this fits well with recent findings that show that the MFN effect is mainly caused by differences in brain activity following positive feedback (see Holroyd, Pakzad-Vaezi, & Krigolson, 2008).

Even though many studies suggest that the ACC is more responsive to errors and negative feedback than to positive outcomes, neurons responding to either positive or negative outcomes have been found to be about equally represented in the ACC (Quilodran et al., 2008). Nevertheless, whereas many studies have shown involvement of the ACC in the processing of negative feedback and subsequent performance adjustments, ACC responses to positive feedback and relationships between such positive outcomes and adaptive performance adjustments and learning are not often observed. The reason for this may be that, in many studies, negative feedback is much more salient and informative for subsequent behavior than positive feedback. In the present design, positive feedback was at least as informative as negative feedback. Indeed, positive feedback was so informative here that only relatively few subjects made a sufficient number of incorrect choices after such feedback. This may have influenced our results: We only analyzed data from those subjects that made relatively many bad choices after positive feedback (so, we may have collected data from “bad learners”). Although there is no obvious reason why findings would be any different for subjects who are better learners, these findings have to be interpreted with caution and must remain preliminary.

Nevertheless, our results converge with findings from fMRI and single-cell ACC recordings (Quilodran et al., 2008) that show that positive feedback is also processed in this neural structure in order to form action–outcome associations that are imperative for adaptive learning.

References

- Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a neural system for error detection and compensation. *Psychological Science*, *5*, 303–305.
- Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifacts. *Electroencephalography and Clinical Neurophysiology*, *55*, 468–484.
- Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*, 679–709. doi:10.1037//0033-295X.109.4.679
- Holroyd, C. B., Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-related positivity: Sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology*, *45*, 688–697.
- Luck, S. J. (2005). *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.
- Miltner, W. H. R., Braun, C. H., & Coles, M. G. H. (1997). Event-related brain potentials following incorrect feedback in a time-estimation task: Evidence for a “generic” neural system for error detection. *Journal of Cognitive Neuroscience*, *9*, 788–798.

- Nieuwenhuis, S., Holroyd, C. B., Mol, N., & Coles, M. G. H. (2004). Reinforcement-related brain potentials from medial frontal cortex: Origins and functional significance. *Neuroscience and Biobehavioral Reviews*, *28*, 441–448. doi:10.1016/j.neubiorev.2004.05.003
- Quilodran, R., Rothé, M., & Procyk, E. (2008). Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron*, *57*, 314–325. doi:10.1016/j.neuron.2007.11.031
- Rushworth, M. F. S., Noonan, M. P., Boorman, E. D., Walton, M. E., & Behrens, T. E. J. (2011). Frontal cortex and reward-guided learning and decision-making. *Neuron*, *70*, 1054–1069. doi:10.1016/j.neuron.2011.05.014
- Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron*, *36*, 241–263.
- Van der Helden, J., Boksem, M. A. S., & Blom, J. H. G. (2010). The importance of failure: Feedback-related negativity predicts motor learning efficiency. *Cerebral Cortex*, *20*, 1596–1603. doi:10.1093/cercor/bhp224

(RECEIVED August 10, 2011; ACCEPTED March 30, 2012)