Published in final edited form as:

Psychophysiology. 2013 February ; 50(2): 125–133. doi:10.1111/j.1469-8986.2012.01490.x.

The Role of Outcome Expectations in the Generation of the Feedback-related Negativity

Andrew W. Bismark^{1,*}, Greg Hajcak², Nicole M. Whitworth¹, and John J.B. Allen¹

¹Department of Psychology, University of Arizona, Tucson, AZ

²Department of Psychology, Stony Brook University, Stony Brook, NY

Abstract

The Feedback-related Negativity (FRN) is thought to index activity within the midbrain dopaminergic reward-learning system, with larger FRN magnitudes observed when outcomes are worse than expected. This view holds that the FRN is an index of neural activity coding for prediction errors, and reflects activity that can be used to adaptively alter future performance. Untested to date, however, is a key prediction of this view: the FRN should not appear in response to negative outcomes when outcome expectations are not allowed to develop. The current study tests this assumption by eliciting FRNs to win and loss feedback in conditions of participant choice, participant observation of computer choice, and critically, simple presentation of win or loss feedback in the absence of a predictive choice cue. Whereas FRNs were observed in each of the conditions in which there was time for an expectation to develop, no FRN was observed in conditions without sufficient time for the development of an expectation. These results provide empirical support for an untested but central tenet of the reinforcement learning account of the genesis of the FRN.

Performance monitoring is an essential process in learning and is managed by a system that not only guides behavior, but allows for behavioral adjustments in the face of outcome errors or violations of predictions about those outcomes. This performance monitoring system is particularly responsive to negative feedback about errors and can signal a variety of conditions such as the commission of an error, the invalidity of the current “rule” of task performance, and the subsequent need to increase task vigilance in order to increase goal-oriented behavior on the next occurrence (Holroyd & Coles, 2002). In the current study, the feedback-related negativity (FRN), an event-related potential (ERP) component that may signify that an outcome was worse than expected, was examined as an index of the neural response to feedback in a choice task. The FRN is a negative deflection peaking between 250–350 milliseconds post feedback with a fronto-central scalp topography, and is thought to be one aspect of a larger performance monitoring system tasked with monitoring both internal and external feedback concerning behavior (Simons, 2010).

Research has implicated the anterior cingulate cortex (ACC) as being an integral component in this general performance monitoring system (Holroyd & Coles, 2002) as well as in reward-based selection of actions (Holroyd, Nieuwenhuis, Mars & Coles, 2004). Within this system, negative feedback serves as a negative reward, allowing the network to modify the appropriate response to the current stimulus or task and adjust task performance. A prominent neurophysiological model suggests that the ACC, as part of the reinforcement learning network, gives rise to the FRN (Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis,

*Corresponding author: Andrew W. Bismark Department of Psychology University of Arizona PO Box 210068 Tucson, AZ 85721-0068 Abismark@email.arizona.edu.

2005) via a phase-locked enhancement of theta activity (Luu, Tucker, & Makeig, 2004; Trujillo & Allen, 2007). The role of the ACC in this reinforcement learning system can be understood in terms of its role within the mesocortical dopamine system (Holroyd and Coles, 2002), which codes errors in reward prediction by phasic changes in the activity of midbrain dopamine neurons, with increased burst magnitude coding for outcomes that are better than expected, and phasic pauses in firing coding for outcomes that are worse than expected (Maia & Frank, 2011). The ACC is the recipient of reward-related predictions and evaluative information that is carried by the mesocortical dopamine system via these phasic changes in dopaminergic afferent activity (Holroyd and Coles, 2002), and upon receipt of these signals, the ACC uses them to improve task performance (Ridderinkhof et al., 2005) by training the motor system to respond to future iterations of that stimulus in ways that match its updated neural representation. Additional research has supported the role of the dopaminergic reward system in the generation of the FRN, such that the administration of a single-dose dopamine agonist enhances FRN amplitude, but that increasing tonic systemic dopamine impairs reinforcement learning (Santesso, Evins, Frank, Schetter, Bogdan & Pizzagalli, 2009).

Consistent with the idea that prediction errors engage the reinforcement learning system to give rise to the FRN, Hajcak, Moser, Holroyd, and Simons (2007) found that FRN amplitudes were larger to feedback signifying loss on trials where subjects expected gains than on those trials where they expected losses. Similarly, Holroyd, Nieuwenhuis, Yeung, and Cohen (2003) found that when rewards were frequent, and thus non-rewarded trials were unexpected, FRNs to non-reward were larger than when non-rewards were frequent and thus expected. Whereas a majority of studies examine FRN amplitudes averaged across multiple trials, this model of reinforcement learning would predict that single-trial variations in amplitude should reflect the magnitude of the negative prediction error (i.e. the extent to which things were worse than expected). By examining single-trial theta band power (which is thought to give rise to the FRN), and by utilizing an abstract computational model that provided a trial-by-trial estimate of the prediction error, Cavanagh, Frank, Klein, & Allen (2010) found that trial-by-trial midfrontal theta power tracked positive and negative prediction errors, and accounted for behavioral slowing on trials subsequent to an error.

The FRN itself has been demonstrated under a number of conditions, including but not limited to, the evaluation of goodness (good/bad) (Hajcak, Moser, Holroyd & Simons, 2006), correctness (error/correct) (Miltner, Braun & Coles, 1997; Zanolie, Leijenhof, Rombouts & Crone, 2008), feedback valence (win/loss) (Gehring & Willoughby, 2002; Yeung, Holroyd & Cohen, 2005; Yeung & Sanfey, 2004), feedback magnitude (Bellebaum, Pölzli & Daum, 2010; Yeung & Sanfey, 2004) and prediction errors (Bellebaum & Daum, 2008; Cavanagh et al., 2010; Holroyd & Coles, 2002; Holroyd, Pakzad-Vaezi & Krigolson, 2008; Potts, Martin, Kamp & Donchin, 2010). Because it is unlikely that, when faced with a choice, an individual will intentionally select a clearly suboptimal option, errors in performance generally constitute violations of the expectation of relative correctness or goodness in choice situations. Such expectations reflect contingencies between behavioral or observed choices and the outcomes of those choices, and engage the mesocortical dopamine reinforcement learning system. This suggests that the FRN is unlikely to be observed in cases where good or bad outcomes occur, but are untethered to choice performance, such as when a computer's screen saver interrupts a movie, following no specific action on the user's part and no clear signal that the screen saver was impending. In this case, the outcome is clearly negative, but untethered to any action or signal that would predict that outcome. The current model suggests that such situations would not involve prediction errors or attendant alterations in dopamine burst magnitude or pause duration, as no cue indicates the likelihood of an outcome and thus no expectation about that outcome can develop.

The present study thus manipulated the time window between a choice and feedback about that choice, thus creating conditions where it would and would not be possible for expectations to develop. In the critical condition, the feedback was presented with no lag from the choice, thus making it possible to examine the impact of outcome expectations on the amplitude or appearance of the FRN. Although previous studies have manipulated various aspects of expectancy, such as the expectancy of response congruent mapping (Meckler et al, 2011) or the probability of correct feedback using probabilistically-rewarded stimuli (Cavanagh, et al., 2010; Pfabigan et al, 2010), or the presentation of feedback inconsistent with learned rules (Cohen, Elger, & Ranganath, 2007), studies to date have not examined feedback that is provided without the opportunity for an outcome expectation to develop. When subjective expectations about the outcome of an action have been assessed, the FRN is related to a negative violation of a positive outcome expectation only when the expectation is assessed close in time to the feedback (Hajcak, Moser, Holroyd, & Simons, 2007). In gambling tasks, like the one used in the present study, the outcome expectation can include predictions about both outcome valence (positive or negative) and the timing of the outcome. In the present study, outcome valence was not contingent on subject choices, so the key distinction of interest for this investigation is the *outcome timing* and whether sufficient time between a response and the outcome exists for an expectation to develop. The rationale guiding the present study follows from a central tenet of the reinforcement-learning account of FRN genesis: when expectations about an outcome are not allowed to develop, the midbrain dopamine reward-learning system will not code for prediction errors, and negative outcomes will not elicit an FRN.

Yeung, Holroyd, & Cohen (2005) designed a study roughly consistent with this logic and found that even in cases where participants made no choice (they simply pressed to initiate a roulette wheel spinning) or even no response (a roulette wheel simply began spinning), an FRN component was observed. In their study, however, there was always a clear interval between the signal that feedback would be forthcoming and delivery of that feedback, allowing for participants to develop an expectation that outcome feedback would occur. This indicates that although no direct action is necessary to produce the FRN, the critical element for its generation may be the *expectation* that one is about to receive information about the outcome, an expectation that is generated within the lag-time between the signal of impending feedback and the feedback itself. Therefore, if the FRN is produced only under conditions where information about an outcome is expected, then manipulating the response-to-outcome lag-time, during which expectations about those outcomes are produced, should have an impact on the FRN. This time-lag was the central manipulation in the present study, as the Yeung et. al. (2005) study lacked a condition in which trial-by-trial loss/gain feedback was presented in the absence of a chance to develop expectations.

The present study included three conditions within a simple four-choice gambling task (see Figure 1). A standard self-choice condition using win-loss feedback was used to replicate previous findings where larger FRNs followed negative feedback compared to positive feedback when individuals made choices. To investigate the role of reward expectations, two observation conditions were utilized, differing in the choice-to-outcome timing. This design would allow for the assessment of whether the FRN would be observed in cases where valenced feedback is presented while preventing outcome expectancies to develop. To prevent these expectancies from developing, the present study presented participants with a condition in which a computer-driven choice and its corresponding feedback were presented simultaneously, such that feedback presentation coincided with choice without an intervening lag time, so no outcome-related expectation could develop prior to receiving feedback (see Figure 2). As a control for a situation where participants do not make a choice themselves, an observational condition was included where participants observed a computer-generated choice using standard choice-outcome timing (see Figure 1). This

observation condition maintained the sequential stimulus-feedback presentation timing of the Self choice condition, providing a similar lag time to allow for outcome expectancy formation. Supporting the choice of this observational condition as a control, research has suggested that the same performance monitoring system engaged during personal task performance is also evoked when passively observing others' task performance (Bellebaum, Kobza, Thiele & Daum, 2010; Van Schie, Mars, Coles & Bekkering, 2004), although the magnitude of the FRN in these observational tasks may be moderated by the extent of task involvement (Yeung, Holroyd & Cohen, 2005), and the perceived closeness of the relationship with the other (Kang, Hirsh, & Chasteen, 2010).

The study hypotheses were as follows: (1) Negative feedback will generate larger FRNs, compared to positive feedback trials; (2) larger FRNs will be generated in conditions with active participation compared to those conditions where the participant is observing the computer's performance with yoked outcomes; (3) FRNs will be generated when there is lag time between the choice (or observed choice) and feedback but no FRN will be generated when choice and feedback are simultaneous, as no outcome expectancies would develop.

Research Design and Methods

Participants and Procedure

Forty-eight, right-handed undergraduates (30 female) participated in this study. Five subjects were subsequently disqualified, two for psychotropic medication use and three for a history of anxiety or affective disorders (Pizzagalli, Peccoralo, Davidson, and Cohen, 2006; Simons, 2010). All remaining participants ($n=43$, 25 female) were psychotropic medication free, healthy individuals with no history of neurological disease, head injury, concussion or loss of consciousness ($>10\text{min}$) or mental illness as measured by self-report. Ages of the participants ranged from 18 to 26 years, with a mean age of $19.2 (\pm 1.6 \text{ S.D.})$. All participants signed informed consent and were informed they would be participating in a decision making task, for which they would receive course credit with the opportunity to win additional money based on their task performance. Upon entering the lab, after a brief study explanation and informed consent, all participants underwent an initial screening and performed a battery of questionnaires. Subjects were then prepared for electrophysiological recording with an electrode cap.

Task Conditions

Self choice—During the Self Choice condition, the participant was asked to choose one of four colored ovals appearing on the screen via button press on a standard USB number pad. Following the button press, a white square outlined the participant's choice for 750ms to allow an outcome 'expectation' to form. Following this window, feedback was given: the white outline changed to blue or yellow, signifying monetary gain or loss, respectively. Figure 1 depicts a task trial.

Observation-Delay—The Observation-Delay condition was similar to the Self condition except the participant was asked to simply sit and observe the trials as if they were watching another person playing. In this condition the computer selected an oval at a randomly determined interval ranging from 500–2500 msec following the presentation of the ovals, and then the win/loss feedback was presented following 750ms to allow an outcome 'expectation' to form. Each participant was instructed to maintain attention to the computer's performance as the participant's outcome (monetary gain or loss) was directly related to the computer's performance (i.e. the computer “wins,” the participant “wins”).

Observation-Immediate—The Observation-Immediate condition was similar to the Observation-Delay condition except it lacked the “expectation” time that had been provided by the white square indicating the choice but preceding the feedback. Instead, a blue or yellow square was immediately displayed surrounding the computer's “choice” signifying monetary gain or loss. Thus, only one square was present in this condition, yellow or blue. This condition also included longer inter-trial intervals which were jittered within pre-set constraints but used to create similar trial lengths across conditions. Figure 2 depicts this type of trial.

Trial Structure

Participants completed each of the three task conditions in a randomized order, with the entire session lasting approximately 30 minutes. Each task consisted of 144 trials, comprising three separate blocks of 48 trials. Consecutive blocks were never from the same condition. Each stimulus presentation window was between 500–2500 milliseconds within which the participant was asked to respond with a button press in the Self condition or maintain attention to the computer's choice in the Observation-Delay and Observation-Immediate conditions. If the participant didn't respond within the allotted time, the trial was skipped. After stimulus presentation, a 750ms delay occurred before feedback was presented in the Self Choice and Observation-Delay conditions. In the Observation-Immediate condition, the participant was shown the choice and outcome simultaneously without delay by the presentation of a colored square (no white square presented). Most investigations of the FRN rely on probabilistic learning or gambling tasks where the participant must learn the contingencies of the task such that increased performance leads to increased monetary gain and therefore increases incentive to do well on the task. For the current study, there was no manipulation of the contingency between the subjects' choices and the likelihood for a winning or losing outcome. Rather, the probabilities of winning or losing varied only slightly within each block but were fixed across the experiment to winning 60% and losing 40%. This was done to maximize the unpredictable nature of the feedback allowing for a large number of ERPs to losing trials while maintaining participant engagement. Each trial was then followed by a variable ITI so each trial, regardless of condition, was the same duration (see Figure 1).

Each participant was instructed about the different conditions and given three practice trials of each condition before beginning the first experimental block. Before each condition and block, a brief break occurred in which the participant was asked if they had any questions and was provided instructions for the following condition. Each participant was handed a numbered keypad and asked to respond (during the appropriate condition) with numbers 1,2,3 and Enter (the four keys along the bottom row of the number pad), corresponding to the 4 balloons from left to right. Participants were instructed that the task was a gambling task and that they would be able to win a small amount of money (about \$10), depending on their performance on the task. In addition, they were instructed that in some conditions they would be playing and in others they would be observing the computer play for them but to pay close attention as their monetary outcome would be directly linked to the computer's performance.

EEG/ERP Data Collection

All EEG data were collected using a 64-channel NeuroScan Synamps2 amplifier (Charlotte, NC) and recording system, utilizing the international 10/20-system for electrode placement. Data were collected at sites: FP1, FP2, FPZ, AF3, AF4, F7, F5, F3, F1, FZ, F2, F4, F6, F8, FT7, FC5, FC3, FC1, FCZ, FC2, FC4, FC6, FT8, T7, C5, C3, C1, CZ, C2, C4, C6, T8, TP7, CP5, CP3, CP1, CPZ, CP2, CP4, CP6, TP8, P7, P5, P3, P1, PZ, P2, P4, P6, P8, PO7, PO5, PO3, POZ, PO4, PO6, PO8, O1, OZ, O1, CB1, CB2 and mastoid channels M1, M2. Four

ocular channels (two vertical: superior and inferior orbit of the left eye; two lateral: outer canthi) were also collected. Electrode impedances were kept under 10k Ω . Data were collected in DC mode using 1000hz sampling rate, amplified 2816 times, and filtered with 200Hz low pass filter prior to digitization. EEG data were online referenced to an electrode site immediately posterior to CZ and subsequently rereferenced offline to averaged mastoids. Each participant's data were collected during a single session. After EEG recording equipment preparation, participants were seated in a sound-dampened room with dim (50W bulb) illumination for task duration.

Data Reduction and Analysis

All data were manually screened for artifacts (spikes, drifts, DC shifts, and non-biological signals). Any bad channels (characterized by greater than 15% of epochs contaminated by such artifacts) were interpolated using a spherical spline interpolation (in EEGLab version 9.0.0.2b, Makeig and Delorme, 2004). Data were re-referenced to averaged-mastoids, then band-pass filtered from .1–15 Hz with a linear digital finite impulse response filter (designed using FWTGEN V3.8, Cook & Miller, 1992), following which data were epoched from –1000 ms to +1000 ms with respect to feedback presentation time, and baseline-corrected using the mean of the values from –200 to 0 ms, where time 0 represents feedback presentation. Epoched data were subjected to Independent Component Analysis (ICA) and IC(s) representing blinks were identified using the ADJUST algorithm (Mognon, Jovicich, Bruzzone, & Buiatti, 2010) and removed. Data were analyzed within the Study function of EEGLab, using statistical permutation testing, and also analyzed using parametric tests on area measures using the Mixed models procedure of SPSS (SPSS Predictive Analytics Software [PASW] Version 18.0, IBM Corporation). The primary window of FRN analysis was between 250–365ms post feedback, as previous studies have suggested windows of similar size or larger (Hajcak, Moser, Holroyd, Simons, 2006). Statistical permutation testing was conducted for all channels, and the parametric statistics for area measures were analyzed for midline channels (Fz, FCz, Cz, CPz & Pz). Epochs were truncated to –200 to +900 for visual display.

Results

Parametric Tests at Midline Sites

FRN area measures were computed using the mean area from 250–365ms post feedback. An omnibus test of whether the FRN varied as a function of feedback under the three conditions was provided using a 3 (Condition: Self Choice, Observation-Delay, Observation-Immediate) $\times$ 2 (Outcome: Win, Loss) $\times$ 5 (Channel: Fz, FCz, Cz, CPz, Pz) Mixed Linear Model. In addition to main effects of Condition ($F(2,686) = 86.9, p < .000$), of Outcome ($F(1,591) = 100.5, p < .000$) and Channel ($F(4,226) = 6.6, p < .000$), there was a predicted Condition by Outcome interaction ($F(2,686) = 32.4, p < .000$). Post hoc tests for each condition separately revealed a significant difference in FRN area between win and loss trials for the Self Choice ($F(1,298) = 89.1, p < .000$) and Observation-Delay conditions ($F(1,375) = 22.1, p < .000$), but not the Observation-Immediate condition ($F(1,412) = 0.9, ns$).ⁱ There was also a significant interaction between Condition and Channel ($F(8,256) = 2.96, p < .01$), which reflected that both the Self Choice and Observation-Delay conditions had a frontal central maxima (p 's $< .05$), whereas there were no significant differences across the midline for the Observation-Immediate condition. The Outcome $\times$ Channel interaction, and three-way interaction involving Condition, Outcome, and Channel did not reach significance.

ⁱThese results were replicated using a peak-to-peak method of FRN-P2.

Nonparametric Permutation Tests Across the Scalp

To provide further tests of these effects, incorporating the entire scalp topography, the region from 250–365ms post feedback as examined using statistical permutation testing to compare ERPs for loss vs. win outcomes within each condition with a significance threshold at $p < .05$. For the Self Choice condition, in which the participant was actively choosing, there were significant differences in FRN amplitude between FRNs on win and loss trials, with larger FRNs occurring to the loss trial type (Figure 3: left column). In the Observation-Delay condition, where the participant was asked to observe the choices of the computer, there were also significant differences between win and loss trials with loss trials again producing larger FRNs compared to wins (Figure 3: middle column). In the Observation-Immediate condition, where participants are asked to again observe the computer's choices, but where choice and feedback were simultaneous, there were no significant differences in FRN amplitude between win and loss trials (Figure 3: right column). For the Self Choice condition, topographic head-maps indicate a central maxima over channel Cz for both win and loss conditions (Figure 4, top row), and significant differences across many frontal central sites (right column). Topographic head-maps for the Observation-Delay condition indicates a fronto-central maxima over FCz for both win and loss conditions (Figure 4, middle row), and also significant differences across many frontal central sites (right column). In the Observation-Immediate condition, topographic maps show a bilateral parieto-occipital maxima for both win and loss (Figure 4, bottom row), and no significant differences at any scalp sites (right column).

Discussion

Consistent with a large literature, loss feedback elicited a relative negativity compared to win feedback trials (i.e., the FRN). These effects were observed in conditions where the participant made a direct choice as well as in conditions where the participant passively observed negative feedback that was presented after a short time lag following the observed choice of a computer. Importantly, and as hypothesized, there were no significant differences between win and loss feedback for observation trials when the observed choice and feedback were presented simultaneously, which prevented an outcome expectation from developing.

The Observation-Delay condition replicates findings of previous research involving feedback during the observation of others' performance (e.g., Bellebaum, Kobza, Thiele & Daum, 2010; Van Schie et al. 2004, Yeung et al., 2005), finding that neural systems activated during action selection and error processing are also activated during observation. These neural systems may play a central role in our ability to predict and evaluate the actions of others and provide a possible pathway for observational learning. As a result of observing others' task performance, the viewer can learn what rules or contingencies are preferred in specific situations, so that when faced with the same or similar situations, the field of possible rules is defined, or at the very least constrained. The fact that the FRN is also seen in contexts where participants merely observe outcomes of others would imply that learning by self-initiated action and by observing others' actions both depend on the midbrain dopamine reward-learning system.

In the Observation-Immediate condition of the present study, by contrast, no outcome expectation was allowed to develop, so that feedback appeared in the absence of expectations about each outcome. In terms of the reward learning system, there would be no prediction error since there was no direct (or observed) action upon which to base a prediction about the outcome.

It may be that predictions can in-fact develop in the absence of an obvious action, as long as some relationship between two events can be ascertained. For instance, Yeung et al. (2005) reported FRNs during conditions where participants were passive observers of the experiment—and also when no action was taken to initiate a trial, but rather a roulette wheel simply began spinning without any initial action from the participant. This roulette wheel condition, however, still created the possibility for event-outcome expectations to develop, and Yeung et al. (2005) interpreted these findings to suggest that the monitoring system indexed by the FRN may track not only response-outcome contingencies, but track a broader array of event-outcome contingencies to facilitate learning, improving the predictive validity of outcome expectations with experience.

Studies assessing the FRN fundamentally assume that violations of expectations, specifically receiving an outcome that is worse than anticipated, engage the reinforcement learning system, which gives rise to the surface-recorded FRN (Holroyd & Coles, 2002). While several studies have tested whether manipulations of the expected valence of an outcome impacts the FRN (Hajcak, Moser, Holroyd, & Simons, 2006), none have examined the most basic assumption, that in order to see a differential impact of positive and negative outcomes, an expectation about those outcomes must have developed. The present study did not directly manipulate the likelihood of outcomes, as this was a simple gambling task where outcomes were not contingent on particular responses. What it did manipulate, importantly, was the timing of the choice cue and the resultant outcome feedback. By manipulating this lag-time, which impacted the time available for an expectation to develop, this study tested this fundamental assumption that in order to generate an FRN, an outcome expectation needs to develop to then be confirmed or violated (thus producing an FRN). As the results indicate, during conditions where expectations about outcome valence were not allowed to develop (the Observation-Immediate condition) due to the simultaneous presentation of choice cue and outcome feedback, this monitoring system was not engaged, and produced no differential FRN.

One might question why the reinforcement learning system is engaged in the present gambling task where there is no objective relationship between the outcome and the choice made, and in instances where no personal action is taken (Observation-Delay condition). Certainly this finding is not unique to the present study, as numerous gambling task studies (e.g. Yeung, et al. 2005; Holroyd, Hajcak, & Larsen, 2006) have observed FRNs to negative outcome feedback even though there is objectively nothing to learn about choice-outcome contingencies. Individuals may expect reward more often than is warranted given the objective probabilities (Hajcak, Moser, Holroyd, & Simons, 2007), and it may be the case that this reinforcement learning system is automatically engaged whenever an action precedes an outcome. Even in cases where there is no connection between the action and the outcome (e.g., pulling the slot lever does not determine whether one wins or loses), the default approach may be for this system to be engaged to monitor event-outcome contingencies, just in case the information can be used to improve future decision making.

Moreover, although numerous studies have identified that FRNs are clearly present for task observers (Bellebaum et al., 2010; Itagaki and Katayama, 2007; Marco-Pallares et al., 2010; van Schie et al., 2004), Marco-Pallares et al (2010) also demonstrated that observers without an outcome affected by players' performance, nonetheless show FRNs of indistinguishable magnitude of that of the player. Thus, even in cases where the observed outcome is not personally relevant, the FRN reflects the tracking of the predictive validity of events to signify particular outcomes. The reinforcement learning system thus facilitates learning in environments that may not be personally relevant currently, but that may pay off in future situations. However, in the present study's Observation-Immediate condition, there is no predictive relationship between event and outcome. Although feedback valence was still

presented, no expectation about the predictive validity of a cue could be generated. Simple win or loss feedback is thus insufficient information to assess the predictive validity of the event-outcome representation.

Caveats and Implications

The lack of FRN differentiating feedback valence between win and loss outcomes in the Observation-Immediate condition may reflect that such a difference might be obscured by the greater stimulus processing needs for the more complex, simultaneous presentation of choice and outcome. However, ERP morphologies for both gain and loss trials were quite similar between both Observation conditions (see Figure 3). Although the topography of the difference is of course different, the raw waveforms suggest similar engagement with the two tasks, and that that stimulus complexity was unlikely to account for the present results.

By testing the importance of expectations only under conditions of observed feedback, and not those directly influenced by the player, one might argue that differential FRNs to feedback valence might be present but masked by the issue of personal relevance. However, in all conditions the outcome feedback impacted the participants' winnings. Although personal relevance does not seem to explain the present findings, the present study lacked a Self-Choice condition with immediate feedback, which would provide a test of the role of cue-outcome expectations under conditions of direct action. Although this seems a theoretically ideal condition to provide an additional test of this hypothesis, the practical aspects of making a Self-Choice Immediate condition are fraught with complexity. The implementation of a Self-Choice Immediate condition presupposes that the feedback could be presented simultaneously with a subject's decision. Given that there will be some latency between making the decision and actual motor movement of pressing the button to choose (e.g., about 200 msec if we infer this latency from the Libet, Wright Jr, & Gleason (1982) and Libet (1985) experiments), and then feedback presentation on the subsequent screen refresh interval (16.6 msec) following the button push, the feedback would be presented with roughly a 217msec delay from the time of the choice. This condition thus has an unavoidable, albeit short, delay of perhaps a quarter second, which would be time in which an expectation seems likely to develop. Perhaps a future study could parametrically manipulate the delay of feedback presentation following choice over these short intervals to assess the time-course of expectation development. But to date, the minimal amount of time needed for an expectation regarding an outcome to develop remains untested.

Interpretations concerning the meaning of the ACC-generated FRN have been mixed, with debate focused on whether it reflects the arrival of the reward signal to the ACC or whether it reflects the use of reward contingencies to alter future behavior. The current dataset support the former hypothesis, as we observe FRN's in conditions without direct response; whereas the latter hypothesis suggests the FRN should occur only in relation to specific executed responses. Our data suggest the FRN does not reflect the use of feedback by the ACC to reinforce/punish actions per se, but instead indicates the ACC uses reward signals to represent stimulus contingencies, regardless of direct or observed involvement. This is in line with recent work concluding that the FRN reflects prediction errors and associated decision values whereas the P300 reflects behavioral adjustment based on explicit task rules (Chase, Swainson, Durham, Benham and Cools, 2012). Given this framework, it is likely that the decreased FRN amplitudes seen in the Observation-Delay relative to the Self-Choice condition in the present study may indicate that the ACC tracks external reward contingencies in this condition, in the same way as conditions requiring actions, but with lower FRN amplitudes reflecting decreased motivational impact or salience leading to decreased involvement of this monitoring system. Furthermore, when expectations regarding stimulus probabilities of reward are not allowed to develop, the monitoring system is not evoked and produces no FRN, irrespective of feedback valence. This indicates the FRN

reflects the method by which feedback is used by the ACC to evaluate event-outcome contingencies via the assessment of expectations (or prediction errors) about those contingencies. These expectation assessments can then be used to update the ongoing event-outcome representations that are used to influence learning and future behavior.

Conclusion

The present study sought to examine a simple, but yet untested, prediction from the reward learning model of FRN generation. The study is novel in that it clearly demonstrates that an FRN can fail to be elicited following negative feedback if the timing is such that an expectation is not allowed to develop. This study demonstrated the necessity of expectations for the generation of the FRN, yet more research is needed to ascertain the specific nature and timing of expectations that engage the reward learning system and give rise to the FRN. Another important goal for future research will be to understand how motivational state (i.e. interest or salience) and affective state influence the mechanisms of the evaluative process in play for personal and observational learning.

Acknowledgments

The authors would like to thank Loran Kelly and Jamie Velo for their tireless efforts on this project and invaluable assistance with electrophysiological data acquisition and screening. This work was supported in part by the infrastructure provided by National Institute of Mental Health (R01MH066902).

References

- Bellebaum, Christian; Daum, I. Learning-related changes in reward expectancy are reflected in the feedback-related negativity. *The European journal of neuroscience*. 2008; 27(7):1823–1835. [PubMed: 18380674]
- Bellebaum C, Poldz D. It is less than you expected: The feedback-related negativity reflects violations of reward magnitude expectations. *Neuropsychologia*. 2010
- Bellebaum, Christian; Kobza, S.; Thiele, S.; Daum, I. It Was Not MY Fault: Event-Related Brain Potentials in Active and Observational Learning from Feedback. *Cerebral cortex* (New York, NY : 1991). 2010
- Cavanagh JF, Frank MJ, Klein TJ, Allen JJB. Frontal theta links prediction errors to behavioral adaptation in reinforcement learning. *NeuroImage*. 2010; 49(4):3198–3209. [PubMed: 19969093]
- Chase HW, Swainson R, Durham L, Benham L, Cools R. Feedback-related Negativity Codes Prediction Error but Not Behavioral Adjustment during Probabilistic Reversal Learning. *Journal of Cognitive Neuroscience*. 2011; 23(4):936–946. [PubMed: 20146610]
- Cohen MX, Elger CE, Ranganath C. Reward expectation modulates feedback-related negativity and EEG spectra. *NeuroImage*. 2007; 35(2):968–978. doi:10.1016/j.neuroimage.2006.11.056. [PubMed: 17257860]
- Cook E III, Miller GA. Digital filtering: Background and tutorial for psychophysicologists. *Psychophysiology*. 1992; 29(3):350–367. [PubMed: 1626044]
- Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*. 2004; 134(1):9–21. [PubMed: 15102499]
- Gehring WJ, Willoughby AR. The medial frontal cortex and the rapid processing of monetary gains and losses. *Science* (New York, NY). 2002; 295(5563):2279–2282.
- Hajcak G, Moser JS, Holroyd CB, Simons RF. The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. *Biological Psychology*. 2006; 71(2):148–154. [PubMed: 16005561]
- Hajcak G, Moser JS, Holroyd CB, Simons RF. It's worse than you thought: The feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology*. 2007; 44(6):905–912. doi:10.1111/j.1469-8986.2007.00567.x. [PubMed: 17666029]

- Holroyd CB, Coles MGH. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychological review*. 2002; 109(4):679–709. [PubMed: 12374324]
- Holroyd CB, Nieuwenhuis S, Yeung N, Cohen JD. Errors in reward prediction are reflected in the event-related brain potential. *Neuroreport*. 2003; 14(18):2481–2484. [PubMed: 14663214]
- Holroyd CB, Pakzad-Vaezi KL, Krigolson OE. The feedback correct-related positivity: sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology*. 2008; 45(5): 688–697. [PubMed: 18513364]
- Holroyd CB, Hajcak G, Larsen JT. The good, the bad and the neutral: electrophysiological responses to feedback stimuli. *Brain Research*. 2006; 1105(1):93–101. [PubMed: 16427615]
- Itagaki S, Katayama J. Self-relevant criteria determine the evaluation of outcomes induced by others. *Neuroreport*. 2008; 19(3):383–387. [PubMed: 18303586]
- Kang SK, Hirsh J, Chasteen A. Your mistakes are mine: Self-other overlap predicts neural response to observed errors. *Journal of Experimental Social Psychology*. 2010; (46):229–232.
- Libet B, Wright EW Jr, Gleason CA. Readiness-potentials preceding unrestricted 'spontaneous' vs. pre-planned voluntary acts. *Electroencephalography and Clinical Neurophysiology*. Sep; 1982 54(3): 322–35. [PubMed: 6179759]
- Libet B. Unconscious cerebral initiative and the role of conscious will in voluntary action. *The Behavioral and Brain Sciences*. 1985; 8:529–566.
- Luu P, Tucker DM, Makeig S. Frontal midline theta and the error-related negativity: neurophysiological mechanisms of action regulation. *Clinical Neurophysiology*. 2004; 115(8): 1821–1835. [PubMed: 15261861]
- Maia TV, Frank MJ. From reinforcement learning models to psychiatric and neurological disorders. *Nature Neuroscience*. 2011; 14(2):154–162.
- Marco-Pallares J, Krämer UM, Strehl S, Schröder A, Münte TF. When decisions of others matter to me: an electrophysiological analysis. *BMC neuroscience*. 2010; 11:86. [PubMed: 20670398]
- Meckler C, Allain S, Carbonnell L, Hasbroucq T, Burle B, Vidal F. Executive control and response expectancy: a Laplacian ERP study. *Psychophysiology*. 2011; 48(3):303–311. [PubMed: 20667036]
- Miltner WHR, Braun CH, Coles MGH. 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*. 1997; 9(6):788–798.
- Mognon A, Jovicich J, Bruzzone L, Buiatti M. ADJUST: An automatic EEG artifact detector based on the joint use of spatial and temporal features. *Psychophysiology*. 2010; 48:229–240.
- Opitz B, Ferdinand NK, Mecklinger A. Timing Matters: The Impact of Immediate and Delayed Feedback on Artificial Language Learning. *Frontiers in Human Neuroscience*. 2011; 5
- 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*. 2011; 48(5):656–664. [PubMed: 21039585]
- Potts GF, Martin LE, Kamp S-M, Donchin E. Neural response to action and reward prediction errors: Comparing the error-related negativity to behavioral errors and the feedback-related negativity to reward prediction violations. *Psychophysiology*. 2010
- Ridderinkhof KR, Ullsperger M, Crone EA, Nieuwenhuis S. The Role of the Medial Frontal Cortex in Cognitive Control. *Science*. 2004; 306(5695)
- Santesso DL, Evins AE, Frank MJ, Schetter EC, Bogdan R, Pizzagalli DA. Single dose of a dopamine agonist impairs reinforcement learning in humans: Evidence from event-related potentials and computational modeling of striatal-cortical function. *Human brain mapping*. 2009; 30(7):1963–1976. [PubMed: 18726908]
- Simons RF. The way of our errors: Theme and variations. *Psychophysiology*. 2010; 47(1):1–14. [PubMed: 19929897]
- Tzur G, Berger A. When things look wrong: theta activity in rule violation. *Neuropsychologia*. 2007; 45(13):3122–3126. [PubMed: 17583758]

- Tzur G, Berger A. Fast and slow brain rhythms in rule/expectation violation tasks: focusing on evaluation processes by excluding motor action. *Behavioural brain research*. 2009; 198(2):420–428. [PubMed: 19100781]
- van der Helden J, van Schie HT, Rombouts C. Observational learning of new movement sequences is reflected in fronto-parietal coherence. *PloS one*. 2010; 5(12):e14482. [PubMed: 21217815]
- van Schie HT, Mars RB, Coles MGH, Bekkering H. Modulation of activity in medial frontal and motor cortices during error observation. *Nature Neuroscience*. 2004; 7(5):549–554.
- van Veen V, Carter CS. The anterior cingulate as a conflict monitor: fMRI and ERP studies. *Physiology & behavior*. 2002; 77(4–5):477–482. [PubMed: 12526986]
- Yeung N, Sanfey AG. Independent coding of reward magnitude and valence in the human brain. *The Journal of neuroscience : the official journal of the Society for Neuroscience*. 2004; 24(28):6258–6264. [PubMed: 15254080]
- Yeung N, Holroyd CB, Cohen JD. ERP correlates of feedback and reward processing in the presence and absence of response choice. *Cerebral cortex (New York, N.Y. : 1991)*. 2005; 15(5):535–544.
- Zanolie K, Van Leijenhorst L, Rombouts SARB, Crone EA. Separable neural mechanisms contribute to feedback processing in a rule-learning task. *Neuropsychologia*. 2008; 46(1):117–126. [PubMed: 17900633]

Figure 1.

Typical trial structure for Self and Observation-Delay conditions. The fixation cross was presented for 1000ms followed by the choice options window between 500–2500ms. During this time, the participant made a choice (for the choice condition) or observed the choice made by the computer (for the Observation-Delay condition) that was then highlighted with a white square for 750ms allowing for the outcome expectation to form. The white square then changed color to blue (win) or yellow (loss) signifying monetary gain or loss. All trials were followed by the monetary amount either won or lost followed by their experimental cumulative total.

Figure 2.

Typical trial structure for the Observation-Immediate condition during which the four choice stimuli appeared simultaneously with a blue or yellow square highlighting one of the ovals. The fixation cross was presented for 1000ms followed by the presentation of the ovals and square for 1000 ms. There was no intermediate white square or delay time for this condition. The ITI was then elongated to create equal trials length across conditions. All trials were followed by the monetary amount either won or lost followed by their experimental cumulative total.

Figure 3.

Grand averaged ERPs for each condition by midline channel. Left Column is Self Choice, the middle column is the Observation-Delay condition and the right column is the Observation-Immediate condition. Loss trials are depicted in Red, win trials in Blue. Shaded areas depict permutation test-derived areas of significant differences between win & loss ($p < .05$).

Figure 4.

Topographic maps for difference waves and grand averages of win & loss trials by condition. Topographic maps of the difference waves of Loss-Win are located in the left-most column. Loss topographic maps are located in the second column from the left, win topographies in the second column from the right. The right-most column depicts permutation test-derived statistical differences between win and loss trials where red dots indicate channels with statistically significant differences ($p < .05$).