From Feedback- to Response-based Performance Monitoring in Active and Observational Learning

Christian Bellebaum[1,2] and Marco Colosio[3]

Abstract

■ Humans can adapt their behavior by learning from the consequences of their own actions or by observing others. Gradual active learning of action–outcome contingencies is accompanied by a shift from feedback- to response-based performance monitoring. This shift is reflected by complementary learning-related changes of two ACC-driven ERP components, the feedback-related negativity (FRN) and the error-related negativity (ERN), which have both been suggested to signal events "worse than expected," that is, a negative prediction error. Although recent research has identified comparable components for observed behavior and outcomes (observational ERN and FRN), it is as yet unknown, whether these components are similarly modulated by prediction errors and thus also reflect behavioral adaptation. In this study, two groups of 15 participants learned action–outcome contingencies either actively or by observation. In active learners, FRN amplitude for negative feedback decreased and ERN amplitude in response to erroneous actions increased with learning, whereas observational ERN and FRN in observational learners did not exhibit learning-related changes. Learning performance, assessed in test trials without feedback, was comparable between groups, as was the ERN following actively performed errors during test trials. In summary, the results show that action–outcome associations can be learned similarly well actively and by observation. The mechanisms involved appear to differ, with the FRN in active learning reflecting the integration of information about own actions and the accompanying outcomes. ■

INTRODUCTION

Successful adaptation to the environment requires the continuous monitoring of behavior. The dopamine (DA) system plays a key role in evaluating the consequences of actions. DA neurons code a reward prediction error with activation increases signaling better-than-expected events and firing decreases indicating worse-than-expected behavioral outcomes (Zaghloul et al., 2009; Schultz & Dickinson, 2000). DA signals are projected to the striatum and the medial pFC (Haber & Fudge, 1997; Berger, Gaspar, & Verney, 1991). In humans, the important role of the striatum and ACC in the processing of performance feedback have been demonstrated in fMRI studies showing that activity in both structures is modulated by feedback valence and/or expectancy not only during feedback processing but also in the period of feedback expectation (Rolls, McCabe, & Redoute, 2008; Knutson & Cooper, 2005; O'Doherty et al., 2004; Delgado, Locke, Stenger, & Fiez, 2003; Breiter, Aharon, Kahneman, Dale, & Shizgal, 2001). ERP research has identified a feedback-locked component, the "feedback-related negativity" (FRN; Miltner, Braun, & Coles, 1997), which emerges between 200 and 300 msec after feedback presentation, is generated in ACC (Gehring & Willoughby, 2002), and has been sug-

gested to indirectly reflect DA neuron activity in response to unfavorable events and thus a negative prediction error (Holroyd & Coles, 2002). In fact, many recent studies have found more pronounced FRN amplitudes the more a negative event deviated from what was expected, particularly in situations in which action–outcome contingencies could be learned (Holroyd, Krigolson, Baker, Lee, & Gibson, 2009; Bellebaum & Daum, 2008; Hajcak, Moser, Holroyd, & Simons, 2007). In their influential reinforcement learning theory, Holroyd and Coles (2002) suggested a close functional link between the FRN and the response-locked error negativity (Ne; Falkenstein, Hohnsbein, Hoormann, & Blanke, 1991) or error-related negativity (ERN; Gehring, Goss, Coles, Meyer, & Donchin, 1993), which is seen for error compared with correct responses within 100 msec after the response, also with a source in ACC (Dehaene, Posner, & Tucker, 1994). They consider these potentials as expressions of the same neural system, as both code events that are worse than expected and show complementary learning-related changes in the acquisition of stimulus–response–outcome associations. Initially, when contingencies are unknown, performance monitoring is feedback-guided and a strong FRN is observed for negative feedback. In later stages of learning, behavior can be evaluated based on the response alone and an ERN is elicited by errors whereas the FRN amplitude for negative feedback decreases (Eppinger, Kray, Mock, & Mecklinger, 2008; Pietschmann,

[1]Heinrich Heine University Düsseldorf, [2]Ruhr University Bochum, [3]University of Padova

Simon, Endrass, & Kathmann, 2008; Holroyd & Coles, 2002).

Behavioral adaptation may also follow the observation of behavior and outcomes in others (Canessa, Motterlini, Alemanno, Perani, & Cappa, 2011; Kobza, Thoma, Daum, & Bellebaum, 2011; Bellebaum, Kobza, Thiele, & Daum, 2010). Via observational learning, the same stimulus–action–outcome associations can be learned as with active learning, but the observed outcome is not given to the observer and does not refer to his or her own actions. Studying differences between active and observational learning is of theoretical interest, because it can provide insights into the mechanisms involved in linking own actions and the accompanying outcomes.

On the one hand, many studies suggest that the monitoring of one's own and other's behavior is mediated by similar brain structures. For example, the ventral striatum codes outcome prediction errors for active and observational learning (Burke, Tobler, Baddeley, & Schultz, 2010). Single neurons in the monkey medial frontal cortex respond to own and observed errors as well as to negative feedback for own or observed behavior (Yoshida, Saito, Iriki, & Isoda, 2012; Matsumoto, Matsumoto, Abe, & Tanaka, 2007). Both ERN- (De Bruijn & von Rhein, 2012; van Schie, Mars, Coles, & Bekkering, 2004) and FRN-like components (Koban, Pourtois, Bediou, & Vuilleumier, 2012; Bellebaum et al., 2010; Yu & Zhou, 2006) were described for the observation of others' errors or error feedback. These components are referred to as observational ERN and FRN (oERN and oFRN), respectively, and are, at least partially, also generated in ACC (Koban et al., 2012; van Schie et al., 2004). On the other hand, a closer look at these findings also reveals critical differences to the monitoring of own behavior. Burke et al. (2010) found that correlations between ventral striatal activity and outcome prediction errors were reversed in observational learning. In contrast to active feedback learning, strongest striatal activations were seen for worse-than-predicted outcomes in observational learning. Single neurons in the monkey medial frontal cortex coding for own and observed errors are anatomically segregated (Yoshida et al., 2012). FRN amplitudes for observed outcomes are typically smaller than for active experience (Bellebaum et al., 2010; Yu & Zhou, 2006), mediated by social context (Koban et al., 2012). This has led to the question whether the oFRN is mediated by expectancy at all, as has been shown for the FRN (e.g., Ferdinand, Mecklinger, Kray, & Gehring, 2012; Bellebaum & Daum, 2008). In a purely observational learning study, Kobza and colleagues (2011) showed that the oFRN was modulated by outcome probability. A larger negative feedback FRN compared with positive feedback was only seen for a very low outcome probability. According to the authors, this finding suggested reduced negative prediction error coding in observational learning. The results also corroborate other recent findings by our group and others showing reduced striatal and DA system involvement in outcome processing in observational learning (Bellebaum, Jokisch, Gizewski, Forsting, & Daum, 2012; Kobza et al., 2012) and a prominent role of DA in value-based action selection (Guitart-Masip et al., 2012; Shiner et al., 2012; Smittenaar et al., 2012).

As outlined above, FRN prediction error coding in active learning is also reflected in amplitude changes during a learning task. FRN amplitude typically becomes smaller once action outcome contingencies have been learned and performance can be judged based on the response alone (Eppinger et al., 2008; Pietschmann et al., 2008; Holroyd & Coles, 2002). In this study, we hypothesized that—because of the reduced sensitivity to negative deviations from expectations—the oFRN in observational learners, assessed as the amplitude difference between negative and positive feedback, would show less evidence of dynamic amplitude changes throughout the learning process compared with the FRN in active learners. More specifically, only the FRN amplitude in active learners and not the oFRN in observational learners was expected to reflect the learning-induced prediction error reduction for negative feedback after erroneous responses. For the analysis of the FRN/oFRN, we thus expected to see an interaction between the factors Group (active vs. observational learners) and Learning (before and after participants had gained insight into stimulus–response–reward contingencies), which was further hypothesized to interact with feedback contingency, because amplitude changes were only expected for contingent feedback in active learners. Similarly, we hypothesized that the development of the ERN and oERN would differ during the learning process. As in previous studies (e.g., Eppinger et al., 2008), the ERN was expected to complement the development of the FRN in the sense that it increased for errors after learning for the contingent feedback condition. As a consequence of the reduced amplitude changes expected for the oFRN with learning, the oERN was hypothesized to also show no or only small changes with learning.

Most studies on observational learning have shown that participants can learn similarly well from own and observed responses and the accompanying outcomes (Kobza et al., 2011, 2012; Bellebaum et al., 2010). Thus, we did not expect differences in outcome-based learning and, consequently, the strength of the resulting stimulus–action contingencies between active and observational learning. The ERN following performance errors has a larger amplitude for more contingent action–outcome associations (Eppinger et al., 2008; Holroyd & Coles, 2002) and has therefore been considered as a marker for the internal representation of an incorrect response (Eppinger et al., 2008). To compare the strength of this representation in active and observational learners, we had participants of both groups engage in test trials in which they responded actively without receiving feedback. The ERN in response to own errors in these test trials was hypothesized to be similarly pronounced in

active and observational learners. In addition to the oFRN and oERN, we also analyzed the feedback-locked P300, which has been linked to expectancy in the context of feedback processing (Wu & Zhou, 2009; Hajcak et al., 2007), and the response-locked Pe (Falkenstein et al., 1991), an indicator of error awareness (Endrass, Reuter, & Kathmann, 2007; O'Connell et al., 2007; Nieuwenhuis, Ridderinkhof, Blom, Band, & Kok, 2001).

METHODS

Participants

Two groups of subjects participated in the experiment, one learning actively and one by observation. The active learning group consisted of 20 participants (eight women) with a mean age of $M = 24.2$ years ($SD = 2.7$ years). Twenty participants (eight women, $M = 25.3$ years, $SD = 2.5$ years) learned by observation. Exclusion criteria for study participation were a history of neurological or psychiatric disorders or drug abuse and regular medication affecting the CNS. As the main aim of the study was to explore learning-related changes in reward and error processing, only participants who showed some evidence of learning entered the analysis. Fifteen participants in the active (five women, $M = 23.5$ years, $SD = 2.3$ years) and 15 participants in the observational learning group (five women, $M = 24.7$ years, $SD = 2.2$ years) fulfilled the learning criteria, reaching an accuracy level of at least 75%, and entered analysis (see below for details). All procedures carried out were in accordance with the declaration of Helsinki. The study was approved by the ethics committee of the Faculty of Psychology at the Ruhr University Bochum, Germany.

Learning Tasks

A variant of a previously described learning task (Eppinger et al., 2008; Holroyd & Coles, 2002) was used and modified to yield two comparable versions, one requiring learning from feedback for own choices and one requiring learning via observed outcomes given for the actions of another person. In both versions, participants were asked to learn stimulus–response–outcome associations to adjust their own behavior.

In each trial of the active learning task, following a fixation cross, an Asian symbol appeared as imperative stimulus on the screen together with two red rectangles representing the response buttons. Participants were asked to press the left or right CTRL keys of a computer keyboard as response buttons to receive monetary feedback for their choice. As soon as they had pressed the left or right button, the corresponding rectangle on the screen turned green, indicating that a response had been made. The imperative stimulus remained on the screen for 500 msec; maximum RT was 1000 msec (i.e., participants could also respond after the stimulus had disappeared). After another 500 msec, in which the screen was black, the monetary feedback stimulus was shown for 500 msec (see Figure 1 for the sequence and timing of events in a single trial). "Correct" responses were rewarded with 20 cents, whereas "incorrect" choices were followed by a monetary punishment (−10 cents). The FRN is sensitive to both the utilitarian and performance aspect of feedback, depending on what information is emphasized in the instruction (Nieuwenhuis, Yeung, Holroyd, Schurger, & Cohen, 2004). In the instruction to our learning tasks, we stressed that response accuracy would be indicated by monetary feedback, so that both types of information

Figure 1. Sequence of events on a particular trial on the active learning task, the observational learning task, and test trials without feedback, which were used to assess learning in both learning tasks.
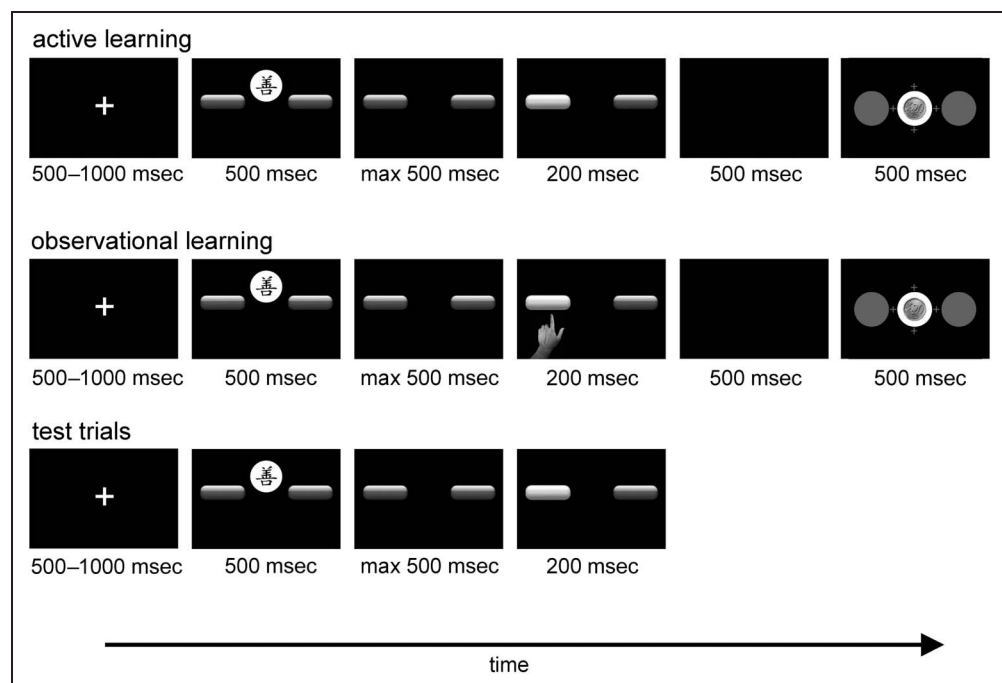
were in accordance with each other. Importantly, the validity of feedback was varied in three different conditions. For two of a total of six imperative stimuli (Stimuli A and B, 100% validity condition), feedback was entirely consistent with reward being associated with one of the two buttons (left for Stimulus A and right for Stimulus B) and punishment with the other in 100% of the trials. For Stimuli C (left button correct) and D (right button correct), outcome feedback was consistent in only 80%, that is, on 20% of the trials a "correct" response led to punishment and an incorrect response to reward (80% validity condition). Finally, monetary outcome feedback was random for two further stimuli (E and F) with either button press being associated with reward and punishment in 50% of trials, respectively (50% validity condition). The fast sequence of stimuli urged participants to respond quickly, so that erroneous responses could still occur after participants had learned the contingencies. Each block involved 60 trials, with trial types involving Stimuli A–F being randomly distributed in each block.

The observational learning task involved the same stimuli and feedback validity conditions as the active learning task. As in real-life observation conditions, the difference was that participants did not choose to press a response button themselves but observed the choices of another person and that the outcomes referred to this other person and not the observational learner. Participants were instructed that they would see the response pattern of another participant who had previously performed the task actively. They were further told that the other person was trying to earn money by responding via button press to particular stimuli. The observer's task would be to try to find out, which responses were rewarding and which were not, because they would be asked to respond themselves in subsequent test trials. Each observer indeed observed the recorded performance pattern and accompanying outcomes of one particular actively learning participant. As in the active version, each trial started with a fixation cross, followed by the presentation of an imperative stimulus together with two red rectangles (the response buttons). After a delay equaling the RT of the observed participant in that particular trial, a picture of a hand was shown on the left or right side of the screen indicating a left or right button press of the observed person, respectively. At the same time, the corresponding rectangle turned green. Then a black screen and the outcome stimulus followed (see Figure 1). To ascertain that observers paid attention to what they saw on the computer screen, they were asked to indicate the response made by the observed participant on the preceding trial on randomly interspersed "attention trials," which occurred on average after every tenth trial. As outlined above, only participants who showed evidence of learning stimulus–response–outcome contingencies were included in the analysis. Therefore, only learners among the actively learning participants were chosen as model for observational learners. As five participants from the active condition

did not learn, five active learners served as model for two observers each. For the 15 participants who succeeded in learning by observation, 13 different active learners served as models (i.e., two pairs of observational learners learned from the responses and outcomes of same active learner, respectively).

To assess whether observers actually learned, each learning phase was followed by a test phase with test trials requiring active responding. These trials involved the same stimuli and response buttons as the trials in the learning phase. They did, however, not involve feedback to prevent learning from active choices. The active learning task entailed identical test phases to provide a comparable measure of learning in active and observational learners. In the test trials, as in the active learner's learning trials, each correct response was rewarded with 20 cents and each error was penalized by subtracting 10 cents for both active and observational learners. Participants were instructed to try to gain as much money as possible also in the test phases and were told that they could keep the money of the most successful block of trials (in active learners, these entailed both the learning and test phases). However, they did not receive feedback on their earnings in the test trials. Timing of stimulus presentation and RTs on test trials were identical to those on active trials (see Figure 1).

Procedure

When participants came into the laboratory, they first filled out an informed consent form. Then a short demographic questionnaire was administered, which also asked for a history of neurological or psychiatric diseases and for regular medication affecting the CNS. After the electrodes had been attached and impedances had been checked, participants were instructed for the (active or observational) learning task, followed by a short practice session of eight (active or observational) learning trials and eight test trials. Each learning phase consisted of 60 trials, 20 per validity condition (10 per stimulus). Also the test phases consisted of 60 trials (10 per stimulus) each. Each learning phase was followed by one test phase. In total, both tasks comprised 18 learning and 18 test phases, that is, 1080 learning and 1080 test trials. In contrast to the procedure in previous studies (e.g., Eppinger et al., 2008), the same six stimuli were used throughout the whole task in all participants. Pilot testing had shown that the prelearning phase was long enough to have enough trials in all validity conditions. At the same time, a long postlearning period was necessary to have participants commit enough errors after learning.

EEG Recording

EEG was recorded from 64 scalp sites (international 10–10 system), and EOG was recorded from electrodes at

the outer canthi of both eyes and above and below the left eye with silver/silver chloride electrodes using two standard Brainamp amplifiers and the Vision Recorder software (Brainproducts, Munich, Germany). Recordings were performed with a sampling rate of 500 Hz while participants engaged in either the active or observational learning task. Electrodes were attached to an elastic cap (Easycap). The EEG signal was referenced to linked mastoids during recording. Impedances were kept below 10 kΩ.

Analysis of Behavioral and ERP Data

To compare learning performance between active and observational learners, response accuracy in the 18 test phases was analyzed. For the 100% validity condition, correct responses were always rewarded. In the 80% validity condition, those responses usually accompanied by reward were scored as correct, irrespective of the actual outcome in a particular trial. Finally, in the 50% validity condition, the left and right buttons were scored as "correct" for Stimuli E and F, respectively, although there were no stimulus–response–outcome contingencies for these trials. For the active learning task, participants' responses in the 18 learning phases were also analyzed and qualitatively compared with performance in the test phases.

Response accuracy data obtained in the test phases were also used to determine when each individual participant gained insight into stimulus–response–outcome associations. On this basis, the experiment was divided into a pre- and a postlearning phase for each participant and for each validity condition. As learning was possible only in the 100% and 80% validity conditions, a learning criterion was applied only for these conditions. Participants were considered to have learned the contingencies for one of the two conditions, when they reached a response accuracy level of 75% and managed to keep this level until the end of the experiment. A further criterion to be classified as "learner" was that participants did not show a clear preference for one of the two stimuli of the 50% validity condition. Although learning was not possible, the experiment was also divided into a pre- and a postlearning phase in this condition to control for sequence effects on the ERPs. This was done on an individual participant basis, taking the average block number in which participants had reached the learning criterion for the learnable conditions (100% and 80% validity) to provide comparable pre- and postlearning phases for all conditions. If, for example, one participant had reached the learning criterion for the 100% and 80% validity conditions in Blocks 3 and 5, respectively, the postlearning phase was considered to start after Block 4 in the 50% validity condition.

Preliminary analyses did not reveal significant differences in response- or feedback-locked ERPs in the 100% and 80% validity conditions. Therefore, ERPs from these conditions were pooled in the final analysis, yielding a "contingent feedback" condition, contrasting the "noncontingent feedback" trials of the 50% validity condition.

EEG data were analyzed offline using the Brain Vision Analyser software (Brainproducts, Munich, Germany) and Matlab (Mathworks, Natick, MA). An independent component analysis (ICA) was performed on the single-subject raw data (Lee, Girolami, & Sejnowski, 1999). ICA decomposes multichannel scalp EEG recordings into a sum of temporally independent and spatially fixed components. In each individual participant of this study, the 64 components (equaling the number of channels) were screened to find one or two components with a symmetric frontally positive topography, which likely represented vertical eye movement and blink artifacts. After removal of the component(s) from the raw data by means of an ICA back transformation, the resulting EEG curves were visually checked for a significant reduction of blink artifacts. Back-transformed data were then filtered with a 0.5-Hz high-pass and a 40-Hz low-pass filter.

First, feedback-locked ERPs were considered in both active and observational learners. Segments from 200 msec before up to 600 msec after feedback presentation were created, separately for monetary reward and punishment in the contingent and noncontingent feedback conditions in the pre- and postlearning phases. Baseline correction was performed relative to the average amplitude in the 200 msec before the feedback stimulus. Trials with artifacts (i.e., with an amplitude difference of more than 150 μV between the highest and lowest data point) were excluded. Finally, the segments were averaged. Importantly, only the valid trials of the 80% validity condition were considered for feedback-locked potentials in the contingent feedback condition, that is, trials in which positive or negative feedback followed a correct or incorrect response, respectively. This was done, because only in valid trials feedback was in accordance with the expectations of the participants after learning and thus learning-related reductions of FRN amplitude could only be expected for these trials (see Eppinger et al., 2008). Similarly, for the noncontingent feedback condition, only positive and negative feedback trials following "correct" and "incorrect" choices, respectively, entered analysis (note that in this condition participants could not learn and response buttons were arbitrarily considered as "correct" or "incorrect" for a given stimulus, see above). As mentioned, data from the 80% and 100% validity conditions were pooled. Two active and three observational learners did, however, not learn stimulus–response–outcome associations in either the 100% or the 80% validity condition. Thus, the postlearning phase in these participants only contained trials of the one validity condition which was actually learned.

To analyze the temporal back-shift from feedback- to response-based performance monitoring during the learning phases, the ERN in active and the oERN in observational learners were analyzed separately, because the ERN and oERN are qualitatively different components, time-locked to an internal event the first and to an external event the

second, and they differ in latency (see De Bruijn & von Rhein, 2012; Koban, Pourtois, Vocat, & Vuilleumier, 2010; van Schie et al., 2004). In both groups, ERPs were segmented separately for (observed) erroneous and correct responses in contingent and noncontingent feedback conditions in the pre- and postlearning phases from 200 msec before to 600 msec after the (observed) response. For active learners, baseline correction was performed relative to the time window from -200 to -100 before the response (see above), because ERN onset is often seen before movement execution. The oERN could not start before an action was observed, so the 200 msec before the observed response were used as baseline. Again, the average ERP for every condition in each group was calculated after the removal of segments with artifacts (see above).

Finally, ERPs time-locked to own responses in the test phases were analyzed in active and observational learners. As for the analyses described above, separate segments were created for correct and error trials for the two contingency conditions before and after learning. Segments again comprised the period from 200 msec before to 600 msec after the response, the average signal in the first 100 msec of the segments (-200 to -100 msec relative to the response) was considered as baseline. Following artifact rejection (see above), the average response-locked ERPs were created.

All analyses were carried out on difference waves (punishment–reward or error–correct) at electrode site FCz in a first step. For feedback-locked potentials, the FRN was defined as the maximum negative peak amplitude of the difference wave in the time window from 180 to 350 msec after feedback presentation. For the analyses of ERPs time-locked to own responses (learning and test trials in active learners, test trials in observational learners), the ERN was scored as the maximum negative peak of the difference wave from 50 msec before to 100 msec after the response. As outlined above, the oERN is a qualitatively different component, time-locked to an external event, and with a less pronounced peak than the ERN (e.g., van Schie et al., 2004). In accordance with previous studies (De Bruijn & von Rhein, 2012; Koban et al., 2010; van Schie et al., 2004), the oERN was thus assessed with a mean amplitude measure, applied to the difference wave (observed error–observed correct response). We chose the time window from 180 to 280 msec after the response, because the largest difference between observed errors and observed correct responses has been described to occur around 250 msec after the observed response (van Schie et al., 2004).

In a second step, data from the original ERPs were analyzed to find out if modulations of the difference wave amplitude were caused by modulations of positive or negative feedback ERPs or by potentials following error or correct trials, respectively. For feedback-locked ERPs, the maximum negative peak amplitude between 180 and 350 msec after positive and negative feedback entered

analysis. In the case of ERPs time-locked to own responses, the negative peak amplitude between -50 and 100 msec relative to the response was analyzed. For observed responses, the mean amplitude between 180 and 280 msec was taken, as for the difference wave (see above). Finally, the P300 for feedback-locked ERPs and the Pe for response-locked ERPs (only for active responses of both groups in the test phase) were analyzed, also at electrode FCz, where they were most pronounced. Here, the peak amplitude between 300 and 500 msec after the feedback and the mean amplitude between 200 and 600 msec were used for analysis, respectively.

Statistical Analysis

Behavioral data were analyzed with a repeated-measures ANOVA including the factors Block (1–18) and Contingency (contingent vs. noncontingent) and the between-subject factor Group (active vs. observational learners). Both the oFRN in learning trials and the ERN in test trials were analyzed by means of a repeated-measures ANOVA with the factors Learning (pre vs. post), Contingency, and Group. For analysis of the original ERPs, the factor Valence (positive vs. negative) or Accuracy (error vs. correct) was added. As outlined above, ERPs following active and observed error responses during the learning phase had to be analyzed separately. Thus, separate ANOVAs were calculated for the ERN in active and the oERN in observational learners with the within-subject factors mentioned above. For all analyses, significant interactions were resolved by means of post hoc t tests (one-tailed).

RESULTS

Behavioral Data

Figure 2 shows the learning curves for the different validity conditions in the 15 active and observational learners (nonlearners excluded). Data are derived from test trials without feedback. In active learners, performance in learning trials is also shown for comparison. On average, both active and observational learners quickly gained insight into stimulus–response–outcome contingencies for the 100% and 80% validity conditions, whereas performance remained on chance level for the noncontingent feedback condition (50% validity). Accordingly, statistical analysis revealed significant main effects of Block (linear trend: $F(1, 28) = 34.894, p < .001$) and Contingency, $F(1, 28) = 124.963, p < .001$, with performance accuracy being higher for the 100% than for the 80%, $t(29) = 3.531$, $p < .01$, and for the 80% relative to the 50% validity condition, $t(29) = 10.245, p < .001$. Additionally, a significant Block × Contingency interaction emerged (linear trend: $F(1, 28) = 19.028, p < .001$). Significant performance increases were seen for the 100% (linear trend: $F(1, 29) = 25.763, p < .001$) and 80% validity conditions (linear trend: $F(1, 29) = 45.552, p < .001$), but not for

Figure 2. Learning curves in active and observational learners.
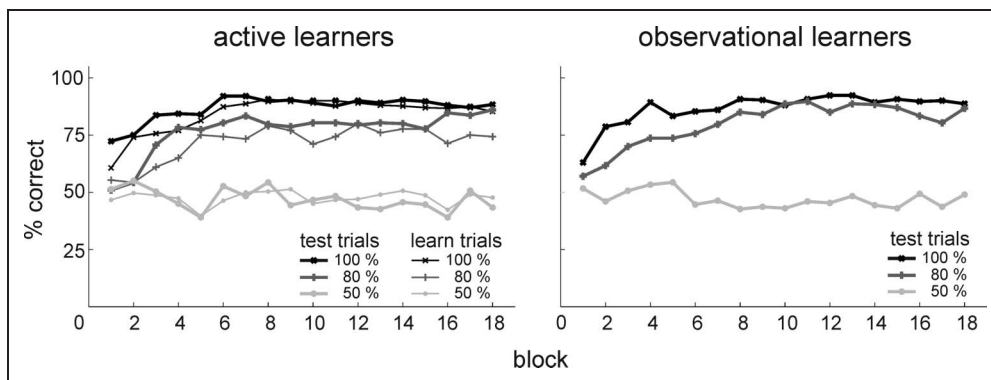
noncontingent feedback (p = .098 for the linear trend). All other effects including the main Group effect did not reach or approach significance (all ps > .320).

To exclude that observers just imitated the choices of active learners without paying attention to the outcomes, we compared each observer's learning success with the performance he or she observed. More specifically, the number of blocks to reach the learning criterion in active learners' learning trials was correlated with the blocks, in which observers reached the criterion. These correlations did not reach or approach significance for the 100% and 80% validity conditions (r = −.187 for 100% and r = .107 for 80%, both p > .500; Figure 3 shows examples of observers' learning curves together with the performance they observed). This shows that the observers did not just pay attention to observed choices and imi-

tated them but took the outcomes into account in the adaptation of their own behavior.

ERP Data

Feedback-locked Potentials

In Figure 4, feedback-locked potentials for monetary reward and punishment in the contingent and non-contingent feedback conditions in the pre- and postlearning phase are shown, separately for active and observational learners. The mean numbers of trials entering analysis are given in Table 1. Analyses are based on the punishment–reward difference waves, which are also depicted (see Figure 8 for topographical maps of difference wave amplitudes in two representative conditions). Because of interindividual differences in FRN and oFRN latencies, a clear negative

Figure 3. Individual learning curves of four observational learners and the corresponding observed performance. The first example (upper row on the left) is taken from the 100% validity condition, the remaining three from the 80% validity condition.
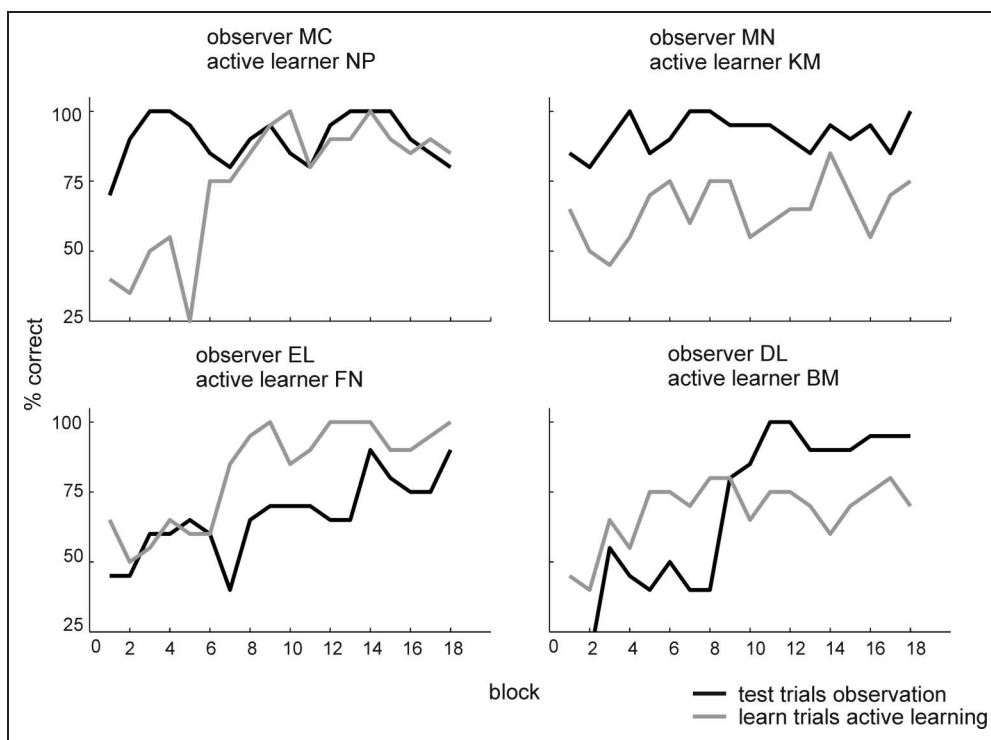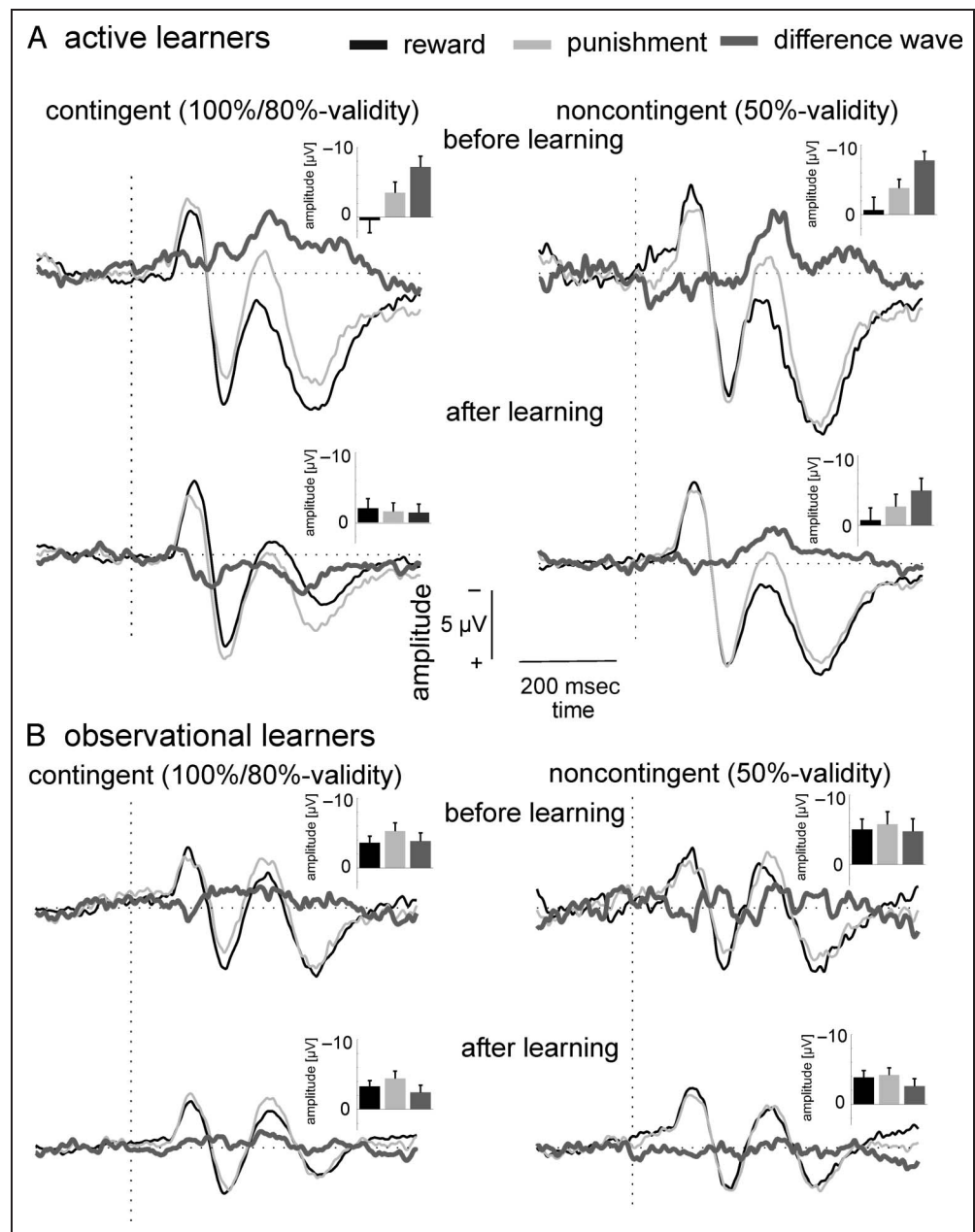
Figure 4. Feedback-locked ERPs from electrode site FCz in active and observational learners' learning trials for contingent and noncontingent feedback conditions in the pre- and postlearning phases. Bars in the upper right-hand corner of each graph show group means and *SE*s of the negative peaks in each condition and for the difference wave.

difference wave peak is not visible in all conditions. To illustrate the group's peak amplitudes, difference waves were synchronized according to each individual participant's peak's latency (Bellebaum, Kobza, Thiele, & Daum, 2011). Figure 5 shows the grand average of the synchronized difference waves. Analysis revealed that FRN amplitude was significantly larger before than after learning, $F(1, 28) = 33.982, p < .001$, for the noncontingent feedback compared with the contingent feedback condition, $F(1, 28) = 9.622, p < .01$, and for active compared with observational learners, $F(1, 28) = 4.500, p < .05$. An interaction between the factors Learning and Group, $F(1, 28) = 5.840, p < .05$, indicated that the amplitude reduction from before to after learning was more pronounced in active, $t(14) = 6.561, p < .001$, than observational learners, $t(14) = 2.194, p < .05$. Crucially, a significant three-way

interaction between Group, Learning, and Contingency, $F(1, 28) = 5.098, p < .05$, suggested between-group differences in the learning-based alteration of feedback processing. A follow-up ANOVA only in active learners indeed revealed an interaction between the factors Learning and Contingency, $F(1, 14) = 5.980, p < .05$, with a significantly larger FRN amplitude for noncontingent than contingent feedback after, $t(14) = 5.857, p < .001$, but not before learning ($p = .311$). For observational learners, the interaction was not significant ($p = .512$).

To explore if and to what extent significant "FRN effects," that is, larger amplitudes for negative than positive feedback, were seen in the different experimental conditions, peak amplitudes of the positive and negative feedback ERPs (see Methods section) entered an ANOVA with the factors Learning, Contingency, and Group and the

Table 1. Mean Number of Trials Entering Analysis of Feedback-locked Potentials in the Two Groups and the Different Conditions (*SE* in Brackets)

| Group	Condition		Contingent Feedback	Noncontingent Feedback
Active learners	Before learning	Punishment	59 (10)	34 (7)
		Reward	78 (12)	31 (6)
	After learning	Punishment	68 (5)	81 (4)
		Reward	392 (20)	73 (5)
Observational learners	Before learning	Punishment	47 (7)	37 (7)
		Reward	104 (20)	31 (6)
	After learning	Punishment	85 (8)	83 (5)
		Reward	391 (15)	71 (7)

additional factor Valence (positive vs. negative). In this analysis, FRN amplitudes were generally more negative in the negative feedback condition, $F(1, 28) = 9.834$, $p < .01$. Significant interactions between Learning and Valence, $F(1, 28) = 12.707$, $p < .01$, and between Learning, Valence, and Group, $F(1, 28) = 6.506$, $p < .05$, indicated that the amplitude difference between negative and positive feedback decreased from pre- to postlearning and that this reduction was more pronounced in active than observational learners. No general group difference in negative peak amplitudes was found ($p = .150$).

For the four-way interaction between all factors, which corresponds to the three-way interaction in the difference wave analysis, a trend toward significance was found ($p = .082$). The comparison of the values between reward and punishment separately for each condition and in active and observational learners yielded significantly more negative amplitudes for punishment in both the contingent, $t(14) = 2.934$, $p < .01$, and noncontingent feedback con-

dition before learning in active learners, $t(14) = 2.499$, $p < .05$. After learning, this effect was only seen for noncontingent, $t(14) = 2.086$, $p < .05$, but not for contingent feedback ($p = .216$). In observational learners, the amplitude difference was significant for contingent feedback both prelearning, $t(14) = 2.895$, $p < .01$, and postlearning, $t(14) = 2.010$, $p < .05$, whereas no significant differences were seen for noncontingent feedback ($p = .181$ and $p = .298$ for pre- and postlearning, respectively).

The P300 was also analyzed as a further indicator of feedback processing. Figure 8 shows the topographies of active and observational learners' P300 peak amplitudes for positive feedback before learning. Because of the frontally pronounced topography, the P300 in this study can be considered as P3a component (Squires, Squires, & Hillyard, 1975). Analysis revealed larger amplitudes before than after learning, $F(1, 28) = 39.078$, $p < .001$, for noncontingent than contingent feedback, $F(1, 28) = 22.526$, $p < .001$, and in active than observational learners,

Figure 5. Grand average of participants' difference waves, temporally synchronized according to the individual FRN or oFRN peak in the respective condition.
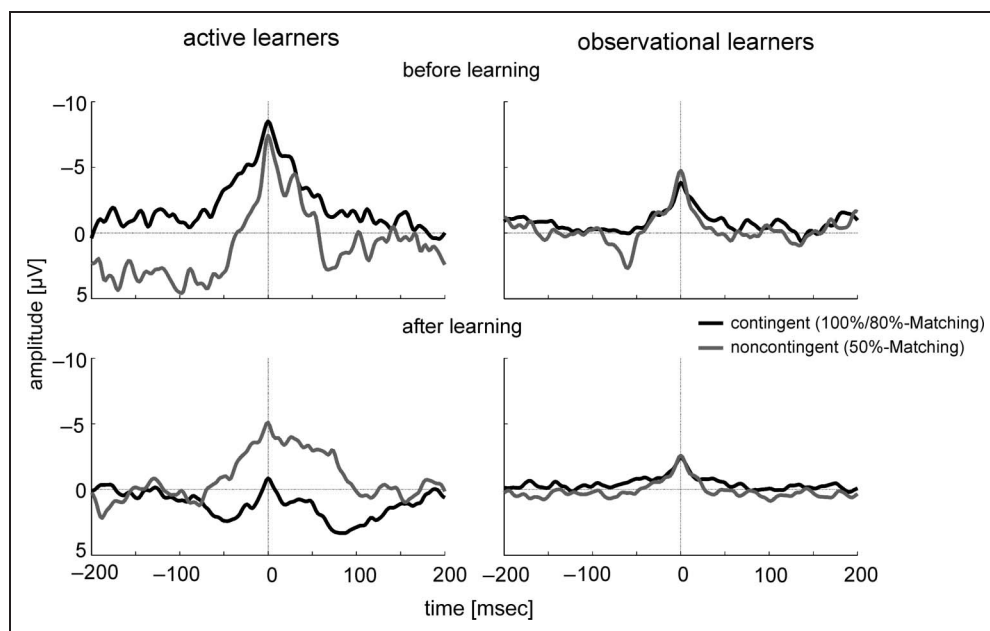
$F(1, 28) = 7.732$, $p < .05$. Moreover, a significant interaction between Group and Contingency, $F(1, 28) = 5.949$, $p < .05$, indicated that there was a significant effect of feedback contingency only in active learners, $t(14) = 5.130$, $p < .001$ ($p = .064$ for observers). A significant interaction between Learning and Valence, $F(1, 28) = 6.826$, $p < .05$, was further specified by a three-way interaction involving the factor Group in addition, $F(1, 28) = 4.525$, $p < .05$. This effect revealed that the amplitude reduction from pre- to postlearning was comparably strong for positive and negative feedback in observational learning,

$t(14) = 3.164$, $p < .01$ and $t(14) = 2.596$, $p < .05$, respectively, whereas the reduction was more pronounced for positive, $t(14) = 6.532$, $p < .001$, than negative feedback, $t(14) = 3.833$, $p < .01$, in active learners.

Response-locked Potentials during the Learning Phase

Figure 6 shows ERPs following actively performed or observed errors and correct responses and the corresponding difference waves in active and observational learners, respectively (see Table 2 for the mean numbers of trials

Figure 6. ERPs for actively performed and observed errors during learning trials in active and observational learners, respectively, from electrode site FCz. Data are shown for contingent and noncontingent feedback conditions in the pre- and postlearning phases. Bars in the upper right-hand corner of each graph show group means and *SE*s of the negative peaks in each condition and for the difference wave.
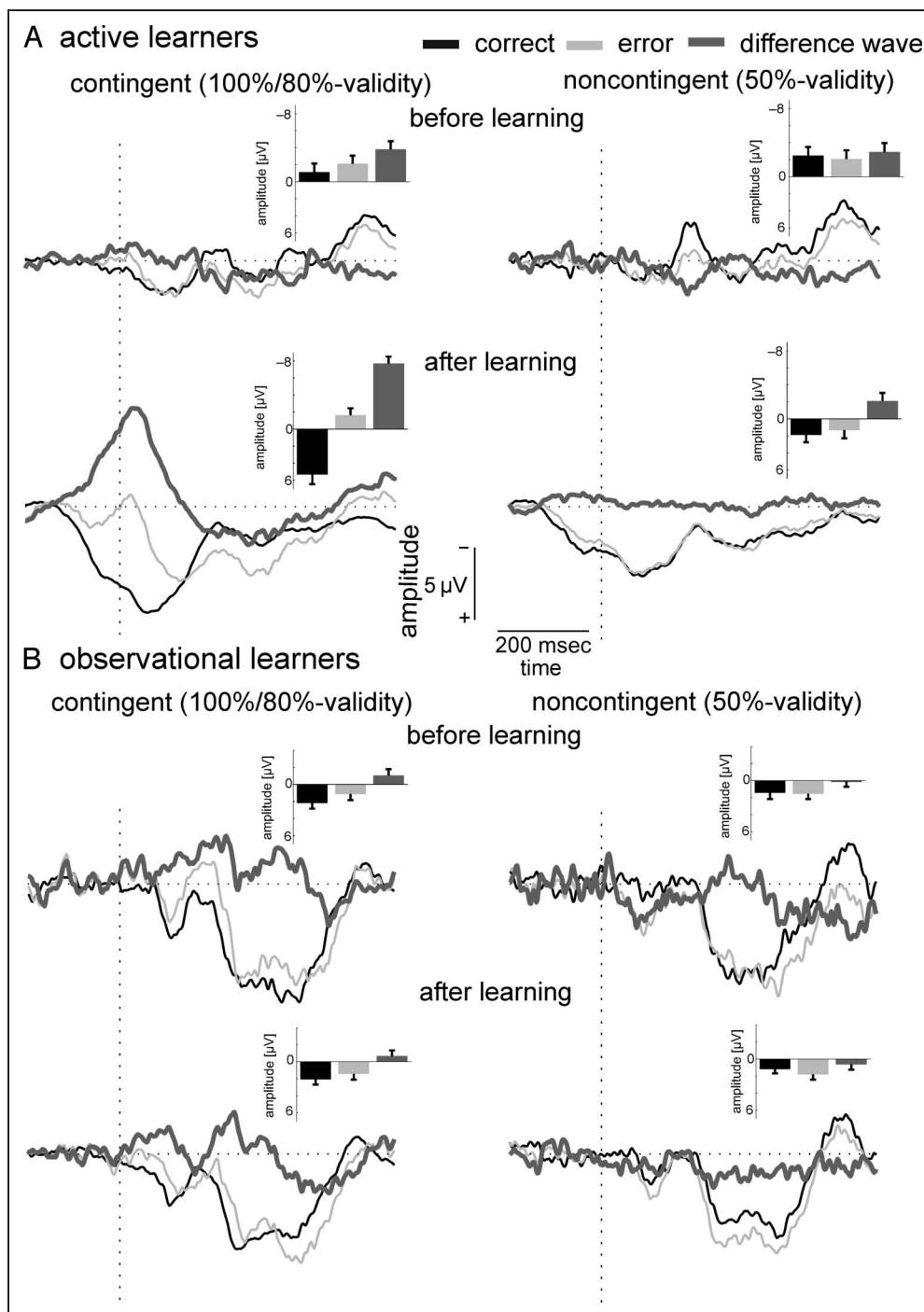
Table 2. Mean Number of Trials Entering Analysis of ERPs Locked to Responses or Observed Responses during Learning Trials in the Two Groups and the Different Conditions (*SE* in Brackets)

Group	Condition		Contingent Feedback	Noncontingent Feedback
Active learners	Before learning	Errors	64 (10)	46 (7)
		Correct	84 (12)	42 (6)
	After learning	Errors	81 (5)	130 (7)
		Correct	432 (22)	117 (7)
Observational learners	Before learning	Errors	51 (7)	47 (7)
		Correct	114 (21)	42 (6)
	After learning	Errors	92 (9)	129 (6)
		Correct	410 (17)	115 (6)

entering analysis). As was explained in the Methods section, the potentials for the two groups of participants had to be analyzed separately. In active learners, main effects of Learning, $F(1, 14) = 5.964$, $p < .05$, and Contingency, $F(1, 14) = 30.821$, $p < .001$, were found, indicating larger ERN amplitudes after than before learning and for the contingent than the noncontingent feedback condition. As for the FRN, an interaction between the factors emerged, $F(1, 14) = 8.245$, $p < .05$. Complementing the FRN findings in active learners, ERN amplitude was larger for the contingent compared with the noncontingent feedback condition after, $t(14) = 6.006$, $p < .001$, but not before learning ($p = .210$). As revealed by analysis of the negative peak amplitudes in the original waveforms, for which a significant interaction between Learning, Contingency and the additional factor Accuracy (error vs. correct) emerged, $F(1, 14) = 9.716$, $p < .01$, amplitudes for error and correct trials did not differ before learning (both $p > .125$ for contingent and noncontingent response–feedback associations). After learning, amplitudes were enhanced for error relative to correct trials only in the contingent feedback condition, $t(14) = 7.439$, $p < .001$ ($p = .101$ for the noncontingent feedback condition).

Analysis of oERN amplitudes only revealed generally larger amplitudes for the contingent feedback condition, $F(1, 14) = 11.860$, $p < .01$. Neither the main effect of Learning nor the Learning by Contingency interaction reached significance (both $p > .220$). Accordingly, analysis of the original waveforms' mean amplitudes yielded an interaction between Contingency and Accuracy. Observed errors elicited a less positive amplitude than observed correct responses for contingent response feedback associations, $t(14) = 2.673$, $p < .01$, whereas there was no significant difference for the noncontingent condition ($p = .055$).

*Response-locked Potentials during the Test Phase*

For response-locked potentials during the test phases of the experiment (see Figure 7, Table 3 lists the numbers of trials entering analysis, Figure 8 shows topographical maps of error–correct difference waves in two postlearning conditions), a significant main effect of Contingency, $F(1, 28) = 49.129$, $p < .001$, and a Learning by Contingency interaction were found, $F(1, 28) = 66.953$, $p < .001$. Overall, ERN amplitudes were larger for the contingent feedback condition. The resolution of the interaction further revealed that ERN amplitude was significantly larger for those stimuli associated with contingent than noncontingent feedback (during the learning phase) only after learning, $t(29) = 11.619$, $p < .001$ (for the difference of ERN amplitudes before learning $p = .470$). All other main effects or interactions did not reach significance (all ps $> .050$).

As for the feedback-locked potentials, ANOVA on the original ERPs aimed to identify whether modulations of the difference wave amplitude were caused by error or correct response ERPs. In accordance with the interaction between Learning and Contingency for difference wave amplitudes, a three-way interaction between Learning, Contingency, and Accuracy (error vs. correct) was found, $F(1, 28) = 41.485$, $p < .001$. Across groups, negative peaks were larger for errors than for correct responses for the contingent feedback condition both before, $t(29) = 4.054$, $p < .001$, and after learning, $t(29) = 12.383$, $p < .001$, with the effect being much stronger after learning. For the noncontingent feedback condition, negative response-locked peaks did not differ in amplitude between erroneous and correct responses (both $p > .210$). The only effect involving the factor Group was an interaction between Group and Contingency, $F(1, 28) = 4.446$, $p < .05$, which was caused by generally more negative amplitudes for the noncontingent than the contingent feedback condition in active, $t(14) = 2.478$, $p < .05$, but not observational learners ($p = .151$).

To examine a further neural correlate of error processing, the Pe was analyzed (see Methods section). Main effects of Learning, $F(1, 28) = 9.848$, $p < .01$, Contingency, $F(1, 28) = 4.964$, $p < .05$, and Accuracy, $F(1, 28) = 5.716$, $p < .05$, indicated generally larger Pe amplitudes for errors

Figure 7. Response-locked ERPs from electrode site FCz in active and observational learners' test trials in the pre- and postlearning phases. Data are shown for stimuli that were associated with contingent and noncontingent feedback during learning. Bars in the upper right-hand corner of each graph show group means and *SE*s of the negative peaks in each condition and for the difference wave.
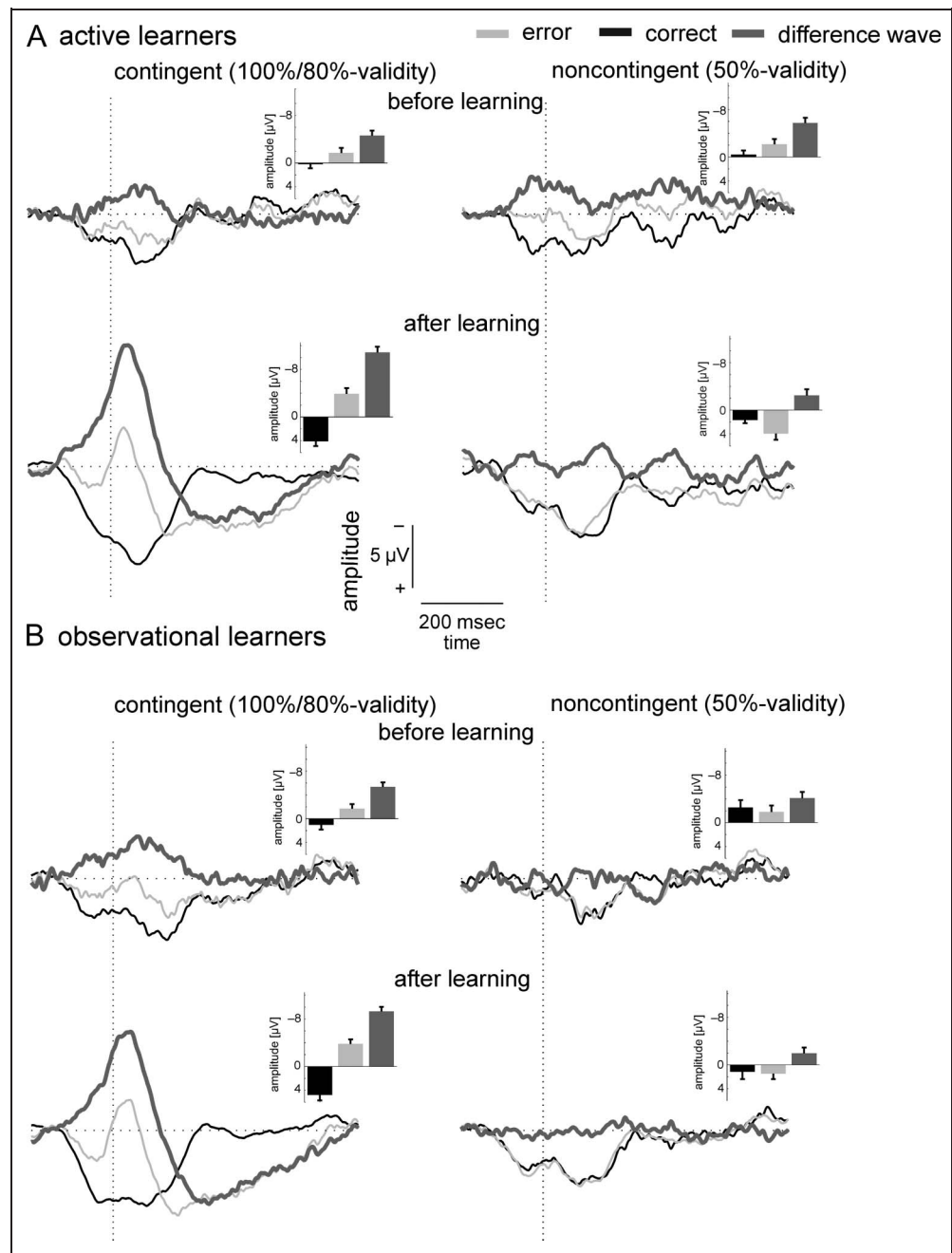
in the contingent feedback condition after learning. These three factors also interacted, $F(1, 28) = 7.625, p < .05$: A significant difference between error and correct trial Pe amplitudes only emerged for the contingent feedback condition after learning, $t(29) = 5.000, p < .001$ (all $ps > .080$ for the other comparisons between error and correct trials). The main effect of Group was not significant ($p = .723$), but an interaction between Contingency and Group emerged, $F(1, 28) = 8.959, p < .01$. The resolution revealed that there was a generally larger amplitude for responses in the contingent than noncontingent feedback condition in observers, $t(29) = 3.405, p < .01$, but not in active learners ($p = 281$).

DISCUSSION

In this study, we compared electrophysiological correlates of performance monitoring and behavioral adaptation in active and observational learning. Two groups of participants acquired stimulus–response–outcome associations, one learning actively by their own actions and outcomes and one by observing actions and outcomes of another person. In active learners, the ERP pattern was characterized by learning-related increases in ERN and decreases in FRN amplitudes following erroneous choices and the accompanying negative feedback, respectively, as had been reported previously (Eppinger

Table 3. Mean Number of Trials Entering Analysis of ERPs Locked to Own Responses in Test Trials without Feedback in the Two Groups and the Different Conditions (*SE* in Brackets)

Group	Condition		Contingent Feedback	Noncontingent Feedback
Active learners	Before learning	Errors	55 (9)	37 (5)
		Correct	78 (11)	37 (4)
	After learning	Errors	58 (5)	129 (10)
		Correct	436 (31)	113 (10)
Observational learners	Before learning	Errors	65 (11)	50 (10)
		Correct	101 (15)	39 (5)
	After learning	Errors	46 (3)	128 (11)
		Correct	459 (24)	118 (11)

et al., 2008; Pietschmann et al., 2008; Holroyd & Coles, 2002). When learning was not possible because of inconsistent action–outcome relationships, FRN amplitude remained high until the end of the experiment. Despite

Figure 8. Topographical maps of the feedback-locked FRN or oFRN and P300 during learning trials and the ERN during test trials in active and observational learners at peak latencies in representative conditions (latencies are given to the left of each map). For the oFRN and P300, the topographies associated with difference wave and positive feedback potentials before learning in the 50% validity condition are shown, respectively. For the ERN, the maps represent potentials after learning for contingent feedback (error–correct difference wave).

comparable learning performance in observational learners, the oFRN was generally reduced relative to the FRN in active learners and, as the oERN, not modulated by learning. Errors in test trials without feedback did, however, elicit strong ERN and Pe components of comparable amplitude in active and observational learners.

A reduction of oFRN amplitude in observational compared with active learning in terms of a reduced difference between ERPs for negative and positive feedback has been reported by most previous studies examining neural correlates of observational reward processing (Koban et al., 2012; Bellebaum et al., 2010; Yu & Zhou, 2006). The performance monitoring differences found in this study go, however, clearly beyond a mere reduction of the oFRN. To the best of our knowledge, this study is the first to compare the relationship between feedback- and error processing and potential learning-related changes in the relative contribution of these processes to performance monitoring in active and observational learners. As suggested by the reinforcement learning theory (Holroyd & Coles, 2002), the FRN amplitude difference between negative and positive feedback has been shown to reflect a negative reward prediction error in active learning (Hajcak et al., 2007), especially when feedback can be used for the optimization of response selection (Holroyd et al., 2009; Bellebaum & Daum, 2008). Accordingly, the FRN in active learners was absent in this study, with no significant difference between amplitudes for negative and positive feedback, when the outcome did not provide useful information about performance accuracy, that is, when negative feedback was perfectly predicted and thus not associated with a negative reward prediction error. After learning, a pronounced ERN signaled performance errors already at the time of responding. In observational learning, neither oFRN nor oERN amplitude changed as learning progressed. For the oFRN, this result suggests that it does not code prediction errors in the same way as the FRN. However, both oFRN and oERN showed effects of Contingency, with generally more negative amplitudes for negative events (performed errors or error feedback) in
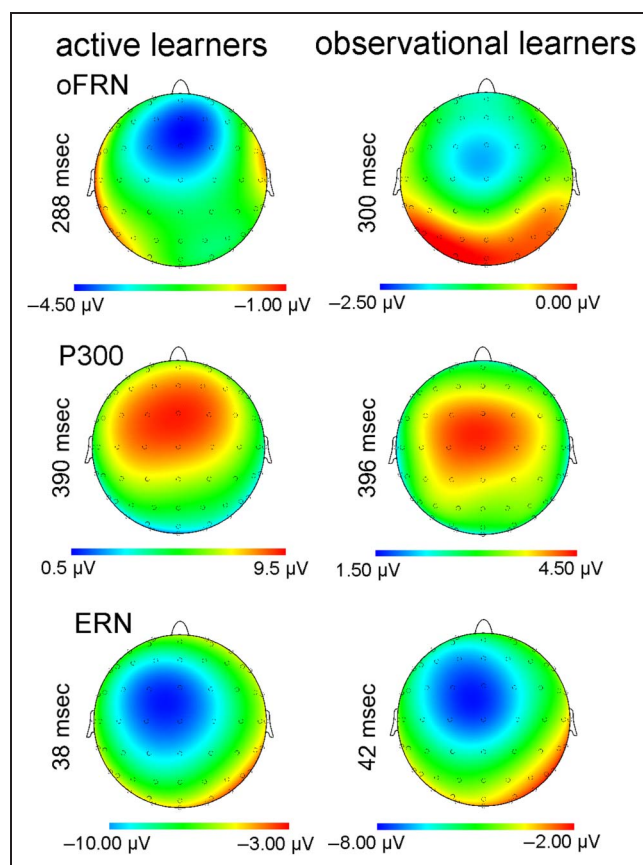
the contingent, but not the noncontingent feedback condition, irrespective of learning phase (pre- or postlearning). Importantly, the error and feedback processing differences between active and observational learning were not accompanied by differences in behavioral adaptation. Both groups showed comparable learning performance as well as comparable learning-related modulations of ERN and Pe amplitudes elicited by own errors in trials without feedback. Both types of learning thus led to similarly strong internal representations of erroneous responses.

Given the comparable learning performance, the reduced oFRN appears to suggest that the role of ACC differs in active and observational feedback learning. It is, however, not entirely clear, what this difference means. Observational feedback learning, as it was implemented in this study, differs from active feedback learning in two ways. First, the presented outcome is not given to the observer and may thus be considered less relevant for the observer than for an active learner. Second, outcome feedback does not refer to the observer's behavior and thus no direct link between (own) behavior and outcome is established. The feedback processing differences between active and observational learners may thus in principle be caused by reduced feedback relevance or by the lacking integration of action and outcome information or both.

As will be outlined in the following, it appears likely, however, that ACC is critically involved in the integration of action and outcome information, suggesting that this process is mainly responsible for the observed differences between active and observational learning. In this study, ACC-driven mediofrontal negative ERP components (ERN and FRN) signal the detection of "worse than expected" events only in active learning, corroborating the view that the dorsal ACC mediates active response selection based on prediction errors, assigning reward value to specific motor actions (Holroyd & Coles, 2008; Yeung, Holroyd, & Cohen, 2005). Functional imaging studies have revealed that the dorsal ACC is activated when participants make choices based on action values (Glascher, Hampton, & O'Doherty, 2009; Wunderlich, Rangel, & O'Doherty, 2009), and lesions to ACC have recently been reported to impair the learning of action but not stimulus values (Camille, Tsuchida, & Fellows, 2011). ERN and FRN are thought to indirectly reflect changes in DA activity associated with negative events (Beste, Saft, Andrich, Gold, & Falkenstein, 2006; Frank, Woroch, & Curran, 2005; Holroyd & Coles, 2002; Falkenstein et al., 2001). Higher FRN amplitudes in active learning compared with oFRN amplitudes in observational learning thus support a key role of the DA system in linking own actions and outcomes as opposed to a more abstract coding of associations between stimuli and/or responses on the one hand and outcomes on the other hand. In further support of this view, modulations of the DA level as in medicated and unmedicated Parkinson's disease patients do not only influence learning from positive and negative feedback,

respectively (Frank, Seeberger, & O'Reilly, 2004), but affect response selection based on action values (Shiner et al., 2012; Smittenaar et al., 2012). Pharmacological enhancement of the DA level specifically strengthens the neural representation of rewarding actions (Guitart-Masip et al., 2012). Also in other parts of the so-called "reward system," which receive dopaminergic projections, neural responses to outcome stimuli depend on associations with own actions. Both for instrumental versus classical conditioning (O'Doherty et al., 2004) and for active versus observational reward learning (Bellebaum et al., 2012), stronger prediction error responses were found in the anterior caudate nucleus.

As mentioned, the reduced personal relevance of the feedback may also play a role in the reduction of the oFRN relative to the FRN (Koban et al., 2012; Bellebaum et al., 2010). With respect to differences between active experience and observation, however, Koban and colleagues (2012) also found a reduction in oFRN amplitude in conditions in which active performer and observer shared a reward. The mere fact that the observers of this study did not receive the observed reward can thus not account for the reduction in oFRN amplitude. Furthermore, the oFRN amplitude reduction in the present as well as in the Koban et al. (2012) study was seen in the difference in neural responses between punishment and reward. It has to be noted, however, that Koban et al. (2012) used a go/no-go task, in which feedback could not be used for learning. Moreover, they varied the social context, having participants observe cooperators or competitors, with oFRN amplitudes being reduced for the latter condition. Thus, the present results and the results of the Koban et al. (2012) study cannot be compared directly. In this study, the analysis of the original ERP amplitudes did not reveal a general group difference in the negative peak amplitudes in the oFRN time window, suggesting that the strength of the neural responses to reward and punishment as such were comparable in active and observational learners. Finally, the fact that both groups of participants learned equally well suggests similar relevance of feedback for active and observational learners. In summary, the outcome processing differences between active and observational learning in this study specifically affected the amplitude difference between punishment and reward, which has been linked to ACC (e.g., Bellebaum & Daum, 2008). Together with the above-mentioned evidence, this finding indicates that ACC is more strongly involved in the processing of negative feedback for own than observed actions, because it links own actions and (negative) outcomes. Potentially reduced personal relevance of the feedback for observational learners, if at all, presumably plays a minor role.

Interestingly, Koban et al. (2012) reported that oFRN amplitude reflects a reward prediction error (Koban et al., 2012), which appears to contradict the findings of this study. In that study, participants took turns with partners in active responding and observing responses

in a go/no-go task, whereas in this study active and observed choices were clearly separated in different groups of participants. It is conceivable that own action representations in observing others influence feedback processing more strongly, when active responding is required on the next trial. In a previous purely observational learning study, we found evidence of reduced prediction error coding in observers. Larger amplitude oFRNs for negative than positive feedback were only seen for outcomes with a very low probability of 20%, but not for outcomes with probabilities between 30% and 80% (Kobza et al., 2011).

With respect to the ERPs in the observer condition (oERN and oFRN), we found a relative negativity for observed errors or error feedback, which was present in both the pre- and postlearning phase for the contingent feedback conditions and thus not modulated by observational learning. A potential explanation could be the difference in learning pace between observers and the participants they learned from. After learning, erroneous actions, that is, specific button presses for a particular stimulus, and the resulting negative outcomes were on average much less frequently observed than correct actions. For some observers this was already the case in the prelearning phase, because the person they observed had learned earlier than the observer himself/herself and thus responded correctly more often than incorrectly. In the prelearning phase, observers did not yet know which responses were wrong, but they might have noticed that the person they observed pressed a particular button and received the respective feedback, more often than the other in a particular context. Thus, the oERN and oFRN might in fact have coded unexpected actions and outcomes. As suggested by Alexander and Brown (2011), ACC acts as a general action outcome predictor and should thus be modulated by any type of expectation violation, positive or negative. In accordance with this view, Ferdinand et al. (2012) reported similar FRN amplitudes following unexpected positive or negative feedback for active responses. We also found evidence for this notion in a recent study, in which a mediofrontal negative ERP component coded unexpected rather than "wrong" actions in observing another's choices (Kobza & Bellebaum, 2013). It has to be noted, though, that exploratory correlation analyses we performed did not reveal a relationship between the strength of the oERN and oFRN Contingency effects and the difference in learning speed between observer and observed person. The questions in how far differences in outcome expectation lead to differences in outcome processing between active and observational learning and to what extent the performance of the observed person plays a role in observational learning are interesting topics for future research.

Taken together, this study shows that behavioral adaptation based on feedback leads to equivalent action–outcome associations in active and observational learning, as revealed by comparable ERN and Pe amplitudes fol-lowing own errors in participants learning actively or by observation. At the same time, the neural signatures of performance evaluation clearly differ between the two types of learning, with learning-related changes being pronounced in active and absent in observational error and feedback processing, showing that prediction errors modulate ACC activity much more in active learning. The fact that similar mediofrontal ERP components associated with performance monitoring are seen in active and observational learning does, however, suggest that the underlying processes also overlap to some extent and that the components as such and the punishment–reward amplitude difference may reflect different neural processes. Future research will have to further explore the specific mechanisms of observational learning.

Acknowledgments

REFERENCES

Alexander, W. H., & Brown, J. W. (2011). Medial prefrontal cortex as an action–outcome predictor. *Nature Neuroscience, 14,* 1338–1344.

Bellebaum, C., & Daum, I. (2008). Learning-related changes in reward expectancy are reflected in the feedback-related negativity. *European Journal of Neuroscience, 27,* 1823–1835.

Bellebaum, C., Jokisch, D., Gizewski, E. R., Forsting, M., & Daum, I. (2012). The neural coding of expected and unexpected monetary performance outcomes: Dissociations between active and observational learning. *Behavioural Brain Research, 227,* 241–251.

Bellebaum, C., Kobza, S., Thiele, S., & Daum, I. (2010). It was not MY fault: Event-related brain potentials in active and observational learning from feedback. *Cerebral Cortex, 20,* 2874–2883.

Bellebaum, C., Kobza, S., Thiele, S., & Daum, I. (2011). Processing of expected and unexpected monetary performance outcomes in healthy older subjects. *Behavioral Neuroscience, 125,* 241–251.

Berger, B., Gaspar, P., & Verney, C. (1991). Dopaminergic innervation of the cerebral cortex: Unexpected differences between rodents and primates. *Trends in Neurosciences, 14,* 21–27.

Beste, C., Saft, C., Andrich, J., Gold, R., & Falkenstein, M. (2006). Error processing in Huntington's disease. *PLoS One, 1,* e86.

Breiter, H. C., Aharon, I., Kahneman, D., Dale, A., & Shizgal, P. (2001). Functional imaging of neural responses to expectancy and experience of monetary gains and losses. *Neuron, 30,* 619–639.

Burke, C. J., Tobler, P. N., Baddeley, M., & Schultz, W. (2010). Neural mechanisms of observational learning. *Proceedings of the National Academy of Sciences, U.S.A., 107,* 14431–14436.

Camille, N., Tsuchida, A., & Fellows, L. K. (2011). Double dissociation of stimulus-value and action-value learning in humans with orbitofrontal or anterior cingulate cortex damage. *Journal of Neuroscience, 31,* 15048–15052.

Canessa, N., Motterlini, M., Alemanno, F., Perani, D., & Cappa, S. F. (2011). Learning from other people's experience: A neuroimaging study of decisional interactive-learning. *Neuroimage, 55,* 353–362.

De Bruijn, E. R., & von Rhein, D. T. (2012). Is your error my concern? An event-related potential study on own and observed error detection in cooperation and competition. *Frontiers in Neuroscience, 6,* 8.

Dehaene, S., Posner, M. I., & Tucker, D. M. (1994). Localization of a neural system for error detection and compensation. *Psychological Science, 5,* 303–305.

Delgado, M. R., Locke, H. M., Stenger, V. A., & Fiez, J. A. (2003). Dorsal striatum responses to reward and punishment: Effects of valence and magnitude manipulations. *Cognitive Affective & Behavioral Neuroscience, 3,* 27–38.

Endrass, T., Reuter, B., & Kathmann, N. (2007). ERP correlates of conscious error recognition: Aware and unaware errors in an antisaccade task. *European Journal of Neuroscience, 26,* 1714–1720.

Eppinger, B., Kray, J., Mock, B., & Mecklinger, A. (2008). Better or worse than expected? Aging, learning, and the ERN. *Neuropsychologia, 46,* 521–539.

Falkenstein, M., Hielscher, H., Dziobek, I., Schwarzenau, P., Hoormann, J., Sunderman, B., et al. (2001). Action monitoring, error detection, and the basal ganglia: An ERP study. *NeuroReport, 12,* 157–161.

Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components. II. Error processing in choice reaction tasks. *Electroencephalography and Clinical Neurophysiology, 78,* 447–455.

Ferdinand, N. K., Mecklinger, A., Kray, J., & Gehring, W. J. (2012). The processing of unexpected positive response outcomes in the mediofrontal cortex. *Journal of Neuroscience, 32,* 12087–12092.

Frank, M. J., Seeberger, L. C., & O'Reilly, R. C. (2004). By carrot or by stick: Cognitive reinforcement learning in parkinsonism. *Science, 306,* 1940–1943.

Frank, M. J., Woroch, B. S., & Curran, T. (2005). Error-related negativity predicts reinforcement learning and conflict biases. *Neuron, 47,* 495–501.

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

Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science, 295,* 2279–2282.

Glascher, J., Hampton, A. N., & O'Doherty, J. P. (2009). Determining a role for ventromedial prefrontal cortex in encoding action-based value signals during reward-related decision making. *Cerebral Cortex, 19,* 483–495.

Guitart-Masip, M., Chowdhury, R., Sharot, T., Dayan, P., Duzel, E., & Dolan, R. J. (2012). Action controls dopaminergic enhancement of reward representations. *Proceedings of the National Academy of Sciences, U.S.A., 109,* 7511–7516.

Haber, S. N., & Fudge, J. L. (1997). The primate substantia nigra and VTA: Integrative circuitry and function. *Critical Reviews in Neurobiology, 11,* 323–342.

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2007). It's worse than you thought: The feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology, 44,* 905–912.

Holroyd, C. B., & Coles, M. G. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review, 109,* 679–709.

Holroyd, C. B., & Coles, M. G. (2008). Dorsal anterior cingulate cortex integrates reinforcement history to guide voluntary behavior. *Cortex, 44,* 548–559.

Holroyd, C. B., Krigolson, O. E., Baker, R., Lee, S., & Gibson, J. (2009). When is an error not a prediction error? An electrophysiological investigation. *Cognitive Affective & Behavioral Neuroscience, 9,* 59–70.

Knutson, B., & Cooper, J. C. (2005). Functional magnetic resonance imaging of reward prediction. *Current Opinion in Neurology, 18,* 411–417.

Koban, L., Pourtois, G., Bediou, B., & Vuilleumier, P. (2012). Effects of social context and predictive relevance on action outcome monitoring. *Cognitive Affective & Behavioral Neuroscience, 12,* 460–478.

Koban, L., Pourtois, G., Vocat, R., & Vuilleumier, P. (2010). When your errors make me lose or win: Event-related potentials to observed errors of cooperators and competitors. *Society for Neuroscience, 5,* 360–374.

Kobza, S., & Bellebaum, C. (2013). Mediofrontal event-related potentials following observed actions reflect an action prediction error. *European Journal of Neuroscience, 37,* 1435–1440.

Kobza, S., Ferrea, S., Schnitzler, A., Pollok, B., Sudmeyer, M., & Bellebaum, C. (2012). Dissociation between active and observational learning from positive and negative feedback in parkinsonism. *PLoS One, 7,* e50250.

Kobza, S., Thoma, P., Daum, I., & Bellebaum, C. (2011). The feedback-related negativity is modulated by feedback probability in observational learning. *Behavioural Brain Research, 225,* 396–404.

Lee, T. W., Girolami, M., & Sejnowski, T. J. (1999). Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Computation, 11,* 417–441.

Matsumoto, M., Matsumoto, K., Abe, H., & Tanaka, K. (2007). Medial prefrontal cell activity signaling prediction errors of action values. *Nature Neuroscience, 10,* 647–656.

Miltner, W. H., 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., Ridderinkhof, K. R., Blom, J., Band, G. P., & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: Evidence from an antisaccade task. *Psychophysiology, 38,* 752–760.

Nieuwenhuis, S., Yeung, N., Holroyd, C. B., Schurger, A., & Cohen, J. D. (2004). Sensitivity of electrophysiological activity from medial frontal cortex to utilitarian and performance feedback. *Cerebral Cortex, 14,* 741–747.

O'Connell, R. G., Dockree, P. M., Bellgrove, M. A., Kelly, S. P., Hester, R., Garavan, H., et al. (2007). The role of cingulate cortex in the detection of errors with and without awareness: A high-density electrical mapping study. *European Journal of Neuroscience, 25,* 2571–2579.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science, 304,* 452–454.

Pietschmann, M., Simon, K., Endrass, T., & Kathmann, N. (2008). Changes of performance monitoring with learning in older and younger adults. *Psychophysiology, 45,* 559–568.

Rolls, E. T., McCabe, C., & Redoute, J. (2008). Expected value, reward outcome, and temporal difference error representations in a probabilistic decision task. *Cerebral Cortex, 18,* 652–663.

Schultz, W., & Dickinson, A. (2000). Neuronal coding of prediction errors. *Annual Review of Neuroscience, 23,* 473–500.

Shiner, T., Seymour, B., Wunderlich, K., Hill, C., Bhatia, K. P., Dayan, P., et al. (2012). Dopamine and performance in a reinforcement learning task: Evidence from Parkinson's disease. *Brain, 135,* 1871–1883.

Smittenaar, P., Chase, H. W., Aarts, E., Nusselein, B., Bloem, B. R., & Cools, R. (2012). Decomposing effects of dopaminergic medication in Parkinson's disease on probabilistic action selection—Learning or performance? *European Journal of Neuroscience, 35,* 1144–1151.

Squires, N. K., Squires, K. C., & Hillyard, S. A. (1975). Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man. *Electroencephalography and Clinical Neurophysiology, 38,* 387–401.

van Schie, H. T., Mars, R. B., Coles, M. G., & Bekkering, H. (2004). Modulation of activity in medial frontal and motor cortices during error observation. *Nature Neuroscience, 7,* 549–554.

Wu, Y., & Zhou, X. (2009). The P300 and reward valence, magnitude, and expectancy in outcome evaluation. *Brain Research, 1286,* 114–122.

Wunderlich, K., Rangel, A., & O'Doherty, J. P. (2009). Neural computations underlying action-based decision making in the human brain. *Proceedings of the National Academy of Sciences, U.S.A., 106,* 17199–17204.

Yeung, N., Holroyd, C. B., & Cohen, J. D. (2005). ERP correlates of feedback and reward processing in the presence and absence of response choice. *Cerebral Cortex, 15,* 535–544.

Yoshida, K., Saito, N., Iriki, A., & Isoda, M. (2012). Social error monitoring in macaque frontal cortex. *Nature Neuroscience, 15,* 1307–1312.

Yu, R., & Zhou, X. (2006). Brain responses to outcomes of one's own and other's performance in a gambling task. *NeuroReport, 17,* 1747–1751.

Zaghloul, K. A., Blanco, J. A., Weidemann, C. T., McGill, K., Jaggi, J. L., Baltuch, G. H., et al. (2009). Human substantia nigra neurons encode unexpected financial rewards. *Science, 323,* 1496–1499.