Worse than feared? Failure induction modulates the electrophysiological signature of error monitoring during subsequent learning

Kerstin Unger · Jutta Kray · Axel Mecklinger

Abstract This study examined how self-relevant failure influences error monitoring—as reflected in the error-related negativity (Ne/ERN) —and behavioral adaptation during subsequent feedback-based learning. We applied two phases (pre- and posttest) of a probabilistic learning task. Between pre- and posttest, participants were assigned to one of two groups receiving either failure feedback or no feedback during a visual search task described as diagnostic of intellectual abilities. To disentangle the effects of failure and motivational disengagement due to prolonged task performance, we linked the posttest to intelligence (Experiment 1) or described it in neutral terms (Experiment 2). Failure induction was associated with an increase in Ne/ERN amplitude at posttest in both experiments, although there were no differences in overall performance. In contrast, the Ne/ERN decreased from pre- to posttest in the no-failure-feedback group, particularly in Experiment 2. Furthermore, failure feedback affected error-related behavioral adjustments, suggesting a shift toward a reactive, error-driven mode of behavior control. These findings emphasize the importance of affective-motivational state in error processing and subsequent behavioral adaptation.

Keywords Negative affect · Motivation · Performance monitoring · Reinforcement learning · ERN

K. Unger (✉) · J. Kray · A. Mecklinger
Department of Psychology, Saarland University,
Campus,
66123 Saarbruecken, Germany
e-mail: k.unger@mx.uni-saarland.de

Introduction

The acquisition and maintenance of adaptive goal-directed behavior requires the ability to detect discrepancies between intended and actual responses (i.e., errors) and to adjust behavior accordingly. The significance of an error, however, can vary considerably, with some errors placing the individual in serious danger or threatening a person's self-worth, whereas others have virtually no consequences. An efficient performance-monitoring system should, therefore, take into account the affective and motivational context of an action. Previous research has shown that the action-monitoring system is sensitive to motivational influences—for example, the significance of an error (Falkenstein, Hoormann, Christ, & Hohnsbein, 2000; Hajcak, Moser, Yeung, & Simons, 2005)—and there is also evidence that the induction of short-term negative affect modulates error monitoring (Olvet & Hajcak, 2011; Wiswede, Münte, Goschke, & Rüsseler 2009; Wiswede, Münte, & Rüsseler 2009). Given that the processing of response errors plays a critical role in learning, an important, but thus far unaddressed, question concerns the influence of experimental manipulations in affective-motivational state on action-monitoring processes in error-driven learning.

So far, a large body of research indicates that uncontrollable failure experiences can severely disrupt subsequent instrumental learning (Mikulincer, 1994; Seligman, 1975; Wortman & Brehm, 1975). Detrimental consequences of failure on cognitive performance have been demonstrated in various tasks (Brunstein & Gollwitzer, 1996) and are assumed to be mediated by motivational deficits, particu-

larly an expectation of future uncontrollability (Dweck & Reppucci, 1973), or cognitive interference—for example, that caused by ruminative thoughts (Brunstein, 1994). It should be emphasized, however, that failure experiences have also been found to improve performance by enhancing effort and task engagement (Brunstein, 2000; Elliott & Dweck, 1988). Nonetheless, most researchers in the field assume that failure outcomes are aversive and trigger negative affective states that individuals have to cope with.

In the present study, we examine whether self-relevant failure affects action-monitoring processes in a subsequent learning task. To this end, we analyze modulations of the error negativity (Ne; Falkenstein, Hohnsbein, Hoormann, & Blanke, 1991) or error-related negativity (ERN; Gehring, Goss, Coles, Meyer, & Donchin, 1993), an event-related potential (ERP) component that has been linked to the activity of an internal error-monitoring system. The Ne/ERN is a negative deflection in the ERP that peaks approximately 80 ms after a participant's erroneous response and is maximal at fronto-central recording sites. Converging evidence indicates that the Ne/ERN is generated in the anterior cingulate cortex (ACC; Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004), a region believed to be involved in the regulation of both cognitive and emotional processing (Bush, Luu, & Posner, 2000). Hence, contemporary accounts of the functional significance of the Ne/ERN focus either on cognitive processes, such as error or conflict detection (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Holroyd & Coles, 2002), or on the evaluation of the affective and motivational significance of an error (Luu, Tucker, Derryberry, Reed, & Poulsen, 2003).

According to the reinforcement learning (R-L) theory, the Ne/ERN reflects the transmission of a negative reinforcement learning signal from the midbrain dopamine system to the ACC (Holroyd & Coles, 2002). This signal is based on reward expectations shaped during learning history and indicates that the outcome of an action is worse than expected—that is, that an error has occurred. The R-L theory predicts that the Ne/ERN should increase with learning, reflecting the development of an internal representation of the correct response. Consistent with this view, several studies have demonstrated learning-related increases of the Ne/ERN (e.g., Eppinger & Kray, 2011; Holroyd & Coles, 2002). Whereas the R-L theory offers a conceptualization of the Ne/ERN in terms of cognitive functions, Luu et al. (2003) proposed that the Ne/ERN might index broader activity of the action regulation circuitry in the limbic system, including the affective evaluation of an error. In line with the notion that mood disorders are associated with excessive responsivity to negative information, increased Ne/ERN amplitudes have been observed in patients suffering from depression (Eshel & Roiser, 2010) and obsessive–compulsive disorder (Gehring, Himle, & Nisenson, 2000).

Results from studies with healthy participants suggest that the Ne/ERN varies as a function of individual differences in punishment sensitivity (Boksem, Tops, Wester, Meijman, & Lorist, 2006) and is greater for participants scoring high on measures of general anxiety and worry (Hajcak, McDonald, & Simons, 2003) or trait-level negative affect (Hajcak, McDonald, & Simons, 2004). Furthermore, the Ne/ERN is related to the salience or significance of an error and is larger when accuracy is emphasized over speed (Falkenstein et al., 2000). Hajcak et al. (2005) showed that the Ne/ERN is enhanced for incorrect trials associated with high monetary value and under conditions of social evaluation. Whereas the latter findings were based on direct manipulations of the motivational impact of an error, the aforementioned studies on affect-related modulations of the Ne/ERN typically used an individual-differences approach. Therefore, it cannot be ruled out that non-affect-related variables account for the observed differences in the Ne/ERN amplitude. Moreover, it remains to be determined whether experimentally induced state variations in negative affect and trait-level negative affect are accompanied by similar changes in the functioning of the error-monitoring system. Experimental studies that examined how affective states influence the Ne/ERN are scarce. In a recent study, Wiswede, Münte, et al. (2009) induced short-term negative and positive affect by brief presentations of pleasant and unpleasant pictures prior to each stimulus in a flanker task. Ne/ERN amplitude was increased on incorrect trials following unpleasant pictures (but see Larson, Perlstein, Stigge-Kaufman, Kelly, & Dotson, 2006). Another study by this group investigated how encouraging or derogatory feedback based on participants' reaction times (RTs) influenced performance monitoring during a flanker task (Wiswede et al., 2009). The Ne/ERN was relatively enhanced for participants receiving derogatory feedback, as compared with those receiving encouraging feedback.

Taken together, a growing body of evidence suggests that the Ne/ERN is sensitive to affective and motivational influences, yet no prior study has addressed this issue using a reinforcement learning paradigm. Given that the Ne/ERN has been related to a prediction error signal indicating negative reinforcement learning (see Holroyd & Coles, 2002), one would expect affective-motivational modulations of this component to be associated with changes in adaptive response selection. Here, we induced self-relevant failure by giving negative feedback regarding participants' intellectual capability to determine the extent to which this manipulation influences the error-monitoring processes indexed by the Ne/ERN and the ability to use error signals for behavioral adaptation in a subsequent feedback-based learning task.

We applied two consecutive phases (pre- and posttest) of a probabilistic reinforcement learning task. Consistent with the R-L-theory (Holroyd & Coles, 2002), a number of studies have shown that Ne/ERN amplitude is reduced

when partly invalid feedback interferes with learning (e.g., Eppinger, Kray, Mock, & Mecklinger, 2008; Nieuwenhuis et al., 2002). In order to investigate whether the effects of failure induction are sensitive to the uncertainty of stimulus–response mappings inherent in a probabilistic learning task, we included three different conditions of feedback validity: In the deterministic learning condition, feedback was always valid; in the probabilistic learning condition, feedback was valid on 80% of the trials but invalid on 20% of the trials; and in the chance condition, feedback was delivered randomly. After the pretest, participants performed a visual search task that they were informed was diagnostic of their intellectual abilities. One half of the participants were exposed to failure feedback, while the other half received no feedback during this task. Subsequently, both groups performed the posttest, in which the effects of prior failure manipulation were assessed. A potential drawback of the pre–post design is that prolonged task performance has been shown to impair action monitoring as indexed by the Ne/ERN (Boksem, Meijman, & Lorist, 2006). Notably, this effect is suggested to be driven by motivational disengagement (Tops & Boksem, 2010), which appears to be sensitive to individual differences in negative emotionality (Luu, Collins & Tucker, 2000). In an attempt to disentangle the effects of reduced task engagement and failure induction, we manipulated participants' motivation by linking the learning task at posttest to intelligence (Experiment 1) or describing it in neutral terms (Experiment 2).

On the basis of previous findings (Hajcak et al., 2005; Wiswede, Münte, et al., 2009; Wiswede, et al., 2009), we expected failure feedback to result in an increase of the Ne/ERN at posttest. Moreover, we hypothesized that this affect-related enhancement would be more pronounced the better participants were able to represent the correctness of their responses; that is, we predicted larger pre–post Ne/ERN differences toward the end, as compared with the beginning, of the learning process. According to the findings on mental fatigue due to sustained task performance (Boksem et al., 2006; Tops & Boksem, 2010), the failure-induced Ne/ERN modulation should be larger in Experiment 1 (self-relevant posttest instruction) than in Experiment 2 (neutral posttest instruction). In addition, we expected accuracy and the Ne/ERN to decrease from pre- to posttest for participants exposed to failure feedback in Experiment 2, but not in Experiment 1.

Experiment 1

Method

Participants

Forty-two undergraduate students with normal or corrected-to-normal vision, free of neurological or psychological disorders and free from psychoactive medication use, participated in the study. Participants were selected on the basis of a screening phase at the beginning of the experiment that was used to rule out near-perfect performance in the learning task. This was done to obtain reliable measures of the Ne/ERN for the present analyses (>14 error trials in both halves of pre- and posttest). Data from 7 participants were discarded due to poor learning performance[1] (3), excessive artifacts (2), and technical problems during EEG recording (2). The final sample thus consisted of 17 participants (12 women; mean age = 22.6 years; age range = 19–33 years) in the failure feedback group and 18 participants (13 women; mean age = 21.7 years; age range = 19–27 years) in the no-failure-feedback group. They were paid 8 Euros per hour or received course credit. Informed written consent was required in accordance with the protocols approved by the local ethics committee of Saarland University, and participants were thoroughly debriefed after the experiment.

Overview of experimental procedure and design

A schematic overview of the experimental procedure is outlined in Fig. 1. After a brief description of the experiment, participants filled out a consent form and a short demographic questionnaire. To control for possible group differences in cognitive as well as in motivational and emotional variables, we administered a number of psychometric tests and questionnaires that included cognitive measures (fluid and crystallized intelligence, working memory capacity), as well as motivation- and affect-related measures (reward/punishment sensitivity, positive/negative affect, action control).[2] The following experiment involved three consecutive phases: the first probabilistic learning phase (pretest), failure manipulation (visual search task), and the second probabilistic learning phase (posttest). After completion of the pretest, one half of the participants were assigned to the no-failure-feedback group, and the other half were assigned to the failure feedback group. Both

---

[1] Less than 55% correct in the deterministic learning condition.

[2] The Digit-Symbol Substitution test (DSS; adapted from Wechsler, 1981) and the Spot-a-Word test (adapted from Lehrl, 1977) served to measure perceptual speed of processing and semantic knowledge, respectively. A modified version of the Digit Ordering Test (Cooper, Sagar, Jordan, Harvey, & Sullivan, 1991) was used to measure working memory capacity. German versions of the following questionnaires were administered: The Positive Affect Negative Affect Scale (PANAS; Krohne, Egloff, Kohlmann, & Tausch, 1996; Watson, Clark, & Tellegen, 1988), the Carver & White (1994) Behavioral Inhibition Scale/Behavioral Activation Scale (BIS/BAS), and the Action Control Scales (ACS-90; Kuhl, 1994). The PANAS assesses the predisposition to experience positive or negative affective state. The BIS/BAS scales were used as measures of punishment and reward sensitivity. The ACS-90 was used to assess the general tendency toward action- versus state-oriented behavior after failure experiences.

Fig. 1 Schematic illustration of the experimental procedure
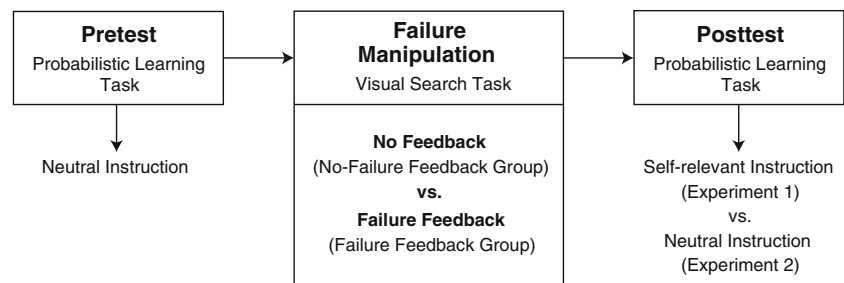
groups were matched for performance in the pretest. In order to test the effect of failure manipulation, participants' mood state was assessed with a short questionnaire immediately after the visual search task ("Befindlichkeitsskala" [BfS]; von Zerssen, 1976).[3] At posttest, participants were informed that they were going to perform the probabilistic learning task again. In contrast to the pretest, it was now stressed that the task was indicative of intellectual abilities. At the end of the experiment, participants were given a brief final questionnaire that focused on the motivational involvement and the emotional experience associated with performing the two learning phases and the visual search task. Afterward, participants were informed of the actual purpose of the study and the nature of the experimental manipulations. Particularly, participants in the failure feedback group were told that the negative feedback in the visual search task was totally independent of their actual performance.

Stimuli and tasks

Probabilistic learning task (pre- and posttest) The stimulus material consisted of 24 colored images of objects (Snodgrass & Vanderwart, 1980). On a given trial, participants were presented with a target stimulus and had to press the left or right response key. Following the response either the word "RICHTIG" ("correct"), "FALSCH" ("incorrect"), or "ZU LANGSAM" ("too slow") was presented. Participants were instructed to infer the correct stimulus–response mappings on the basis of the feedback. In order to maintain motivation throughout the experiment, participants were told that they would gain a point for each correct answer and lose a point for each incorrect or too slow response. At the end of the experiment, a monetary bonus of up to 10 Euros was awarded on the basis of the total amount of points. Stimuli were associated with three different conditions of feedback validity (100%, 80%, and 50%). Four stimuli were assigned to each learning condition, yielding a total of 12 different target stimuli in each learning phase. In the deterministic learning condition, feedback was always valid (100%

validity). Two stimuli were mapped to the left and right response keys, respectively. In the probabilistic learning condition, feedback was valid on 80% of the trials. For the two stimuli that were mapped to the left response key, participants thus received "Correct" feedback for 80% and "Incorrect" feedback for 20% of left buttonpresses (and vice versa for right buttonpresses and the other two stimuli mapped to the right response key). In the chance condition, "Correct" and "Incorrect" feedback was delivered at random (50% validity).

Each trial started with a central fixation cross presented on a light gray background for a variable interval of 250–500 ms, followed by the centrally presented target stimulus that was displayed for 500 ms. In order to obtain a sufficient number of error trials, we applied an adaptive response deadline ranging from 400 to 1,000 ms. The response window was successively incremented or decremented in steps of 100 ms on the basis of the proportion of time-out trials (for details, see Eppinger et al., 2008). After the response, a blank screen was presented for 500 ms, followed by the feedback displayed for 500 ms. A randomly jittered 1,250- to 2,000-ms interval separated each trial. Participants first worked through 60 practice trials. Pre- and posttest consisted of 20 blocks of 30 trials. During the breaks, they received feedback about the total amount of points they had collected up to that point.

Visual search task (failure manipulation) Ten pairs of natural-scene pictures served as stimuli for the failure manipulation task. Each picture pair comprised the "original" and a modified "copy" of a colored photograph that were presented on the left (original) and right (copy) sides of the screen. The "copies" were created by changing 6 to 10 subtle details in each picture (see Fig. 2 for an example stimulus pair). All photographs were scaled to 9 × 12 cm. The task was described as a mental speed test predictive of intelligence. This was incorporated to enhance self-relevance, which is assumed to be an essential feature of efficient failure induction techniques (see Brunstein, 2000). Each picture pair was presented for 60 s, and participants were asked to indicate by mouse-click all differences in the modified picture as quickly as possible. Participants in the failure feedback group received predominantly negative

---

[3] The BfS comprises 28 pairs of opposite adjectives (e.g., *self-confident/insecure*). For each pair, participants are required to indicate the adjective that better represents their current feelings.

Fig. 2 Example stimulus pair presented in the visual search task. Differences are marked in the rightmost picture

feedback indicating the number of differences that had not been detected. Feedback was delivered according to a fixed schedule that has been shown to induce a strong failure experience (Brunstein & Olbrich, 1985). An initial sequence of success and failure (0–2–0–3) was followed by continuous failure (4–4–3–5–4–6). Thereafter, a brief protocol was presented that provided participants in the failure feedback group with a spurious evaluation of their performance: the total number of differences missed by the individual (31) and the average number of differences missed by other participants (19). In the no-failure-feedback group, no feedback was provided.

Control variables Tables 1 and 2 show the results of the psychometric tests and the questionnaires for the no-feedback and the failure feedback groups in Experiments 1 and 2, respectively. Importantly, the two groups did not differ significantly with respect to psychometric measures in either experiment (all p values > .14).

EEG recording

The electroencephalogram (EEG) was recorded with Ag/AgCl electrodes from 59 sites according to the extended 10–20 system, referenced to the left mastoid, at a sampling rate of 500 Hz (filtered online from DC to 70 Hz). The horizontal and vertical electrooculograms were recorded from electrodes placed on the left and right canthi of both eyes and at the infra- and supra-orbital ridges of the left eye. Electrode impedances were kept below 5 kΩ.

Data analyses

Behavioral data analyses Response latencies of less than 244 ms (>2 *SD*) or exceeding the response deadline were

excluded from further analyses.[4] Mean accuracy rates were computed separately for each learning condition in pre- and posttest by averaging the data into six bins of 100 trials each—that is, Bin 1 contained Trials 1–100, Bin 2 contained Trials 101–200, and so on. Note that for the probabilistic learning condition, only valid trials were included. To examine error-related behavioral changes, we additionally computed posterror accuracy—that is, the proportion of correct choices on the next presentation of a particular item—separately for each learning condition and test phase.

ERP data analyses Offline, the data were rereferenced to the linked mastoid and band-pass filtered from 0.05 to 30 Hz. Eye movement artifacts were removed using independent component analysis from the BrainVision Analyzer Software Package (Brain products); remaining artifacts were eliminated with a semiautomatic artifact rejection procedure (amplitudes over ±100 μV, changing more than 50 μV between samples or more than 200 μV within single epochs, or containing baseline drifts). Artifact-free EEG data were segmented relative to response onset and were baseline corrected with respect to a -200- to -50-ms preresponse interval.

Following previous studies using probabilistic learning tasks (e.g., Eppinger et al., 2008; Frank, Woroch, & Curran,

---

[4] The deadline was exceeded by 2.2% (no-failure-feedback group) and 2.1% (failure feedback group) of the responses. At pretest, mean RTs on correct and incorrect trials were 421 ms (*SD*=25 ms) and 417 ms (*SD*=30 ms) for the no-failure-feedback group and 422 ms (*SD*=26 ms) and 416 ms (*SD*=29 ms) for the failure feedback group. At posttest, mean RTs on correct and incorrect trials were 405 ms (*SD*=30 ms) and 398 ms (*SD*=31 ms) for the no-failure-feedback group and 396 ms (*SD*=32 ms) and 388 ms (*SD*=31 ms) for the failure feedback group.

Table 1 Results of the psychometric measures, the mood scale (BfS), self-evaluation for the visual search task, and motivational involvement/rumination at posttest (means and standard deviations) for the no-failure-feedback group and the failure feedback group (Experiment 1)

Measure	No-Failure-Feedback Group	Failure Feedback Group
Cognitive variables		
Digit-Symbol	63.67 (10.62)	62.18 (9.70)
Spot-a-word	18.56 (6.61)	18.76 (4.89)
Digit ordering	9.33 (2.06)	8.94 (1.92)
Affect and action control		
Positive Affect	34.28 (4.38)	34.47 (5.51)
Negative Affect	23.17 (6.30)	20.06 (5.93)
BIS	2.87 (0.45)	2.91 (0.49)
BAS	2.99 (0.49)	3.11 (0.33)
Action Control (ACS-90)[a]	4.50 (3.01)	3.82 (3.36)
Mood (postmanipulation)		
BfS[b]	9.72 (3.56)	18.47 (3.34)
Final questionnaire		
Self-evaluation	3.47 (0.74)	2.71 (0.44)
Involvement	3.33 (0.78)	3.80 (0.61)
Rumination	2.03 (1.01)	2.06 (0.92)

[a] Higher scores indicate action orientation, while lower scores indicate state-orientated behavior

[b] Higher scores indicate more negative feelings

Table 2 Results of the psychometric measures, the mood scale (BfS), self-evaluation for the visual search task, and motivational involvement/rumination at posttest (means and standard deviations) for the no-failure-feedback group and the failure feedback group (Experiment 2)

Measure	No-Failure-Feedback Group	Failure Feedback Group
Cognitive variables		
Digit-Symbol	61.59 (9.01)	61.88 (8.41)
Spot-a-word	21.06 (4.49)	19.81 (3.31)
Digit ordering	8.53 (2.43)	8.94 (2.05)
Affect and action control		
Positive Affect	36.29 (4.78)	37.38 (4.76)
Negative Affect	21.47 (4.61)	23.19 (7.83)
BIS	2.88 (0.49)	2.84 (0.41)
BAS	3.16 (0.29)	3.18 (0.31)
Action Control (ACS-90)[a]	4.12 (2.00)	4.75 (2.65)
Mood (postmanipulation)		
BfS[b]	10.71 (5.55)	18.06 (5.54)
Final questionnaire		
Self-evaluation	3.53 (0.82)	2.59 (0.64)
Involvement	3.43 (0.62)	3.32 (0.69)
Rumination	2.09 (0.92)	1.59 (0.73)

[a] Higher scores indicate action orientation, while lower scores indicate state-orientated behavior

[b] Higher scores indicate more negative feelings

2005), we quantified the Ne/ERN after 15-Hz low-pass filtering at electrode FCz as the peak-to-peak difference in voltage between the most negative peak between -50 and 100 ms and the largest positive peak in the prior 100 ms. We selected FCz for analyses on the basis of visual inspection of the waveforms, which demonstrated a fronto-central maximum of the Ne/ERN. Peak-to-peak voltage was measured to determine baseline-independent amplitudes and to minimize distortions due to the positivity on which the Ne/ERN is superimposed. Since a negative peak could not be reliably determined for correct trials, mean amplitudes in a 0- to 100-ms postresponse time window at electrode FCz were computed to analyze the positivity following correct responses. We decided to analyze correct and incorrect responses separately, rather than to compute difference waves (error - correct), because former studies revealed larger learning-related modulations in the positivity on correct trials than in the ERN (e.g., Eppinger et al., 2008), whereas affective-motivational manipulations have been shown to specifically affect the Ne/ERN (e.g., Hajcak et al., 2005). To examine learning-related changes of the two components, ERP averages were computed for each learning condition, separately for the first and second halves of pre- and posttest. Thus, Bin 1 (first half) of the ERP analysis comprised Bins 1–3 of the

behavioral analysis, and Bin 2 (second half) of the ERP analysis comprised Bins 4–6 of the behavioral analysis.

Statistical analyses Accuracy and ERP data were analyzed using repeated measures analyses of variance (ANOVAs). Whenever necessary, the Geisser–Greenhouse correction was applied (Geisser & Greenhouse, 1958), and corrected p values are reported together with the uncorrected degrees of freedom and the epsilon values (ε). Pearson's correlations were calculated to examine the relation between Ne/ERN amplitude and posterror accuracy. Questionnaire data were analyzed using a multivariate analysis of variance.

Results

Manipulation check

After performing the visual search task, the failure feedback group indeed reported more negative feelings than did the no-failure-feedback group, as indicated by a higher BfS score, $F(1, 33) = 56.07$, $p < .001$. Moreover, participants' self-reports on the final questionnaire showed that they were less satisfied with their performance on the visual search task, $F(1, 33) = 15.78$, $p < .001$, and also tended to

indicate higher posttest involvement, $F(1, 33) = 3.90$, $p = .057$ (see Table 1).

Accuracy data

Accuracy data (see Fig. 3a) were analyzed using an ANOVA with the between-subjects factor *feedback group* (failure feedback vs. no feedback) and the within-subjects factors *test phase* (pretest vs. posttest), *learning condition* (deterministic, probabilistic, and chance conditions), and *bin* (Bins 1–6). As was expected, the analysis yielded a

significant main effect of learning condition, $F(2, 66) = 231.26$, $p < .001$, $\varepsilon = .80$. Contrasts revealed accuracy to be higher for the deterministic and probabilistic learning conditions than for the chance condition and higher for the deterministic than for the probabilistic learning condition (p values $< .01$).

Learning-related effects We found a reliable main effect of bin, $F(5, 165) = 22.81$, $p < .001$, $\varepsilon = .66$, that was qualified by significant interactions between learning condition and bin, $F(10, 330) = 4.83$, $p < .001$, $\varepsilon = .82$, and test phase,

Fig. 3 a Mean accuracy learning curves for the three learning conditions displayed separately for the no-failure-feedback group and failure feedback group at pretest (left) and posttest (right). Error bars indicate standard errors **b** Mean posterror accuracy rates for the three learning conditions at pretest and posttest displayed separately for the no-failure-feedback group (left) and failure feedback group (right). Error bars indicate standard errors
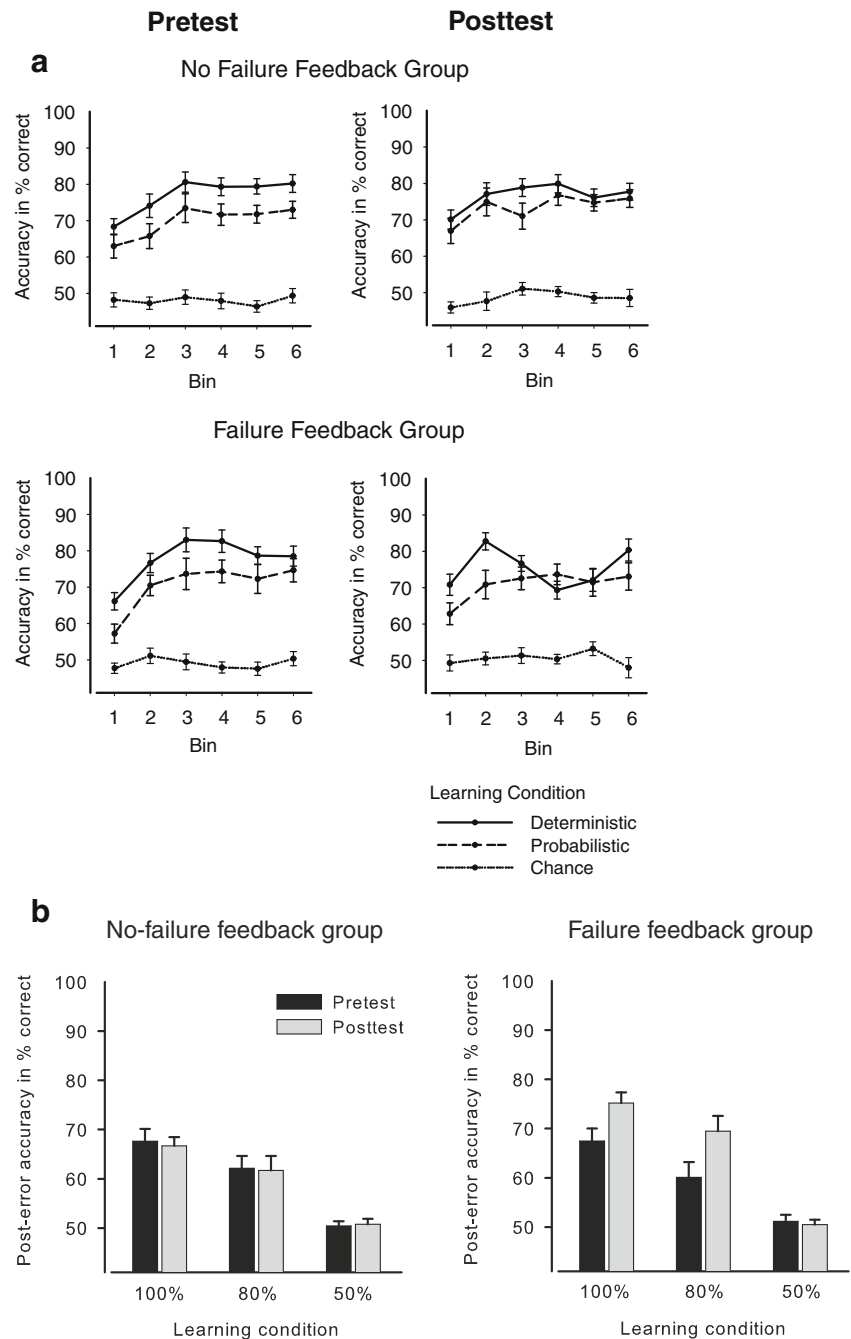
learning condition, and bin, $F(10, 330) = 3.28$, $p < .01$, $\varepsilon = .85$, suggesting that the course of learning differed between pre- and posttest as a function of feedback validity. To decompose the interaction, we ran separate analyses for pre- and posttest. At pretest, polynomial contrasts showed that accuracy increased over the course of the task, following a linear trend for the deterministic as well as the probabilistic learning conditions (p values $< .001$; see Fig. 3a). At posttest, a significant linear trend across bins was obtained only for the probabilistic learning condition ($p < .01$). In contrast, a predominantly cubic trend for the deterministic learning condition ($p < .001$) reflected that after an initial enhancement, accuracy dropped and, finally, increased again.

Effects of failure manipulation Visual inspection of the posttest data suggested that the transient decrease in accuracy emerged for the failure feedback group only (see Fig. 3a). However, the four-way interaction of feedback group, test phase, learning condition, and bin failed to reach significance ($p = .10$).

Posterror accuracy data

In order to determine whether failure induction affects the ability to learn from past mistakes and, thus, modulates error-related behavioral changes, posterror accuracy rates were subjected to an ANOVA with the between-subjects factor *feedback group* and the within-subjects factors *test phase* and *learning condition*. Consistent with the results for the total accuracy, we obtained a significant main effect of learning condition, $F(2, 66) = 78.80$, $p < .001$, indicating that posterror accuracy was lowest in the chance condition and highest in the deterministic learning condition (all p values $< .001$) (see Fig. 3b).

Effects of failure manipulation The analysis revealed a reliable main effect of test phase, $F(1, 33) = 4.27$, $p < .05$, that was qualified by a significant interaction between feedback group and test phase, $F(1, 33) = 5.53$, $p < .05$, and a marginally significant interaction between feedback group, test phase, and learning condition, $F(2, 66) = 2.85$, $p = .065$. Separate analyses for the two groups showed that posterror accuracy increased from pre- to posttest for the failure feedback group ($p < .01$), but not for the no-failure-feedback group ($p = .85$). To examine whether the posterror accuracy differences were due to the failure feedback group's adopting a more conservative response strategy—that is, more accurate at the expense of slower posterror responding—we also analyzed posterror correct versus incorrect RTs. This analysis did not reveal significant interactions involving the factors feedback group and test phase (p values $> .15$). Thus, the failure-related change in

posterror accuracy does not appear to reflect a speed–accuracy trade-off.

Summary of behavioral findings

Not surprisingly, we observed that accuracy increased with feedback validity. Moreover, accuracy increased in a linear function over the course of pretest for both learning conditions, but this was the case only for the probabilistic learning condition at posttest. For the deterministic learning condition, considerable learning occurred at the beginning of posttest, but accuracy decreased with time on task. Since participants were likely to become quickly aware of the response contingencies in the deterministic learning condition, this finding suggests a higher amount of "slips" in cognitive control, resulting in motor errors of commission—rather than weak reinforcement learning—during later stages of posttest. Consistent with this notion, RT decreased from pre- to posttest, and erroneous responses were faster in the deterministic than in the probabilistic learning condition (p values $< .001$). There were, however, no between-group differences in overall performance at either pretest or posttest. Importantly, we also did not obtain pre–post changes in overall accuracy for the two feedback groups. Instead, participants in the failure feedback group were more likely to correct their errors on the next repetition of a given stimulus, as was shown by an increase in posterror accuracy from pre- to posttest.

ERP data

As is illustrated in Fig. 4, the Ne/ERN was evident as a fronto-centrally distributed negative deflection in the deterministic and probabilistic learning conditions for both groups at pre- and posttest. Correct responses were followed by a pronounced positivity that, unlike the negativity on incorrect trials, clearly increased over the course of learning. To test for group differences in the peak-to-peak amplitude of the Ne/ERN and the mean amplitude of the correct response-related positivity, we used separate 2 (*feedback group*: failure feedback vs. no feedback) × 2 (*test phase*: pretest vs. posttest) × 3 (*learning condition*: deterministic, probabilistic, and chance conditions) × 2 (*bin*: Bin 1 vs. Bin 2) ANOVAs. Both the Ne/ERN and the correct response-related positivity increased with feedback validity, $F(2, 66) = 49.14$ and 49.50, $p < .001$ and $.01$, respectively]. Below, we will first report the failure- and learning-related effects for the Ne/ERN, followed by the results for correct responses.

General effects of failure on the Ne/ERN We obtained a reliable main effect of feedback group, $F(1, 33) = 8.30$, $p < .01$, and significant interactions between feedback group and test phase, $F(1, 33) = 10.01$, $p < .01$, and feedback
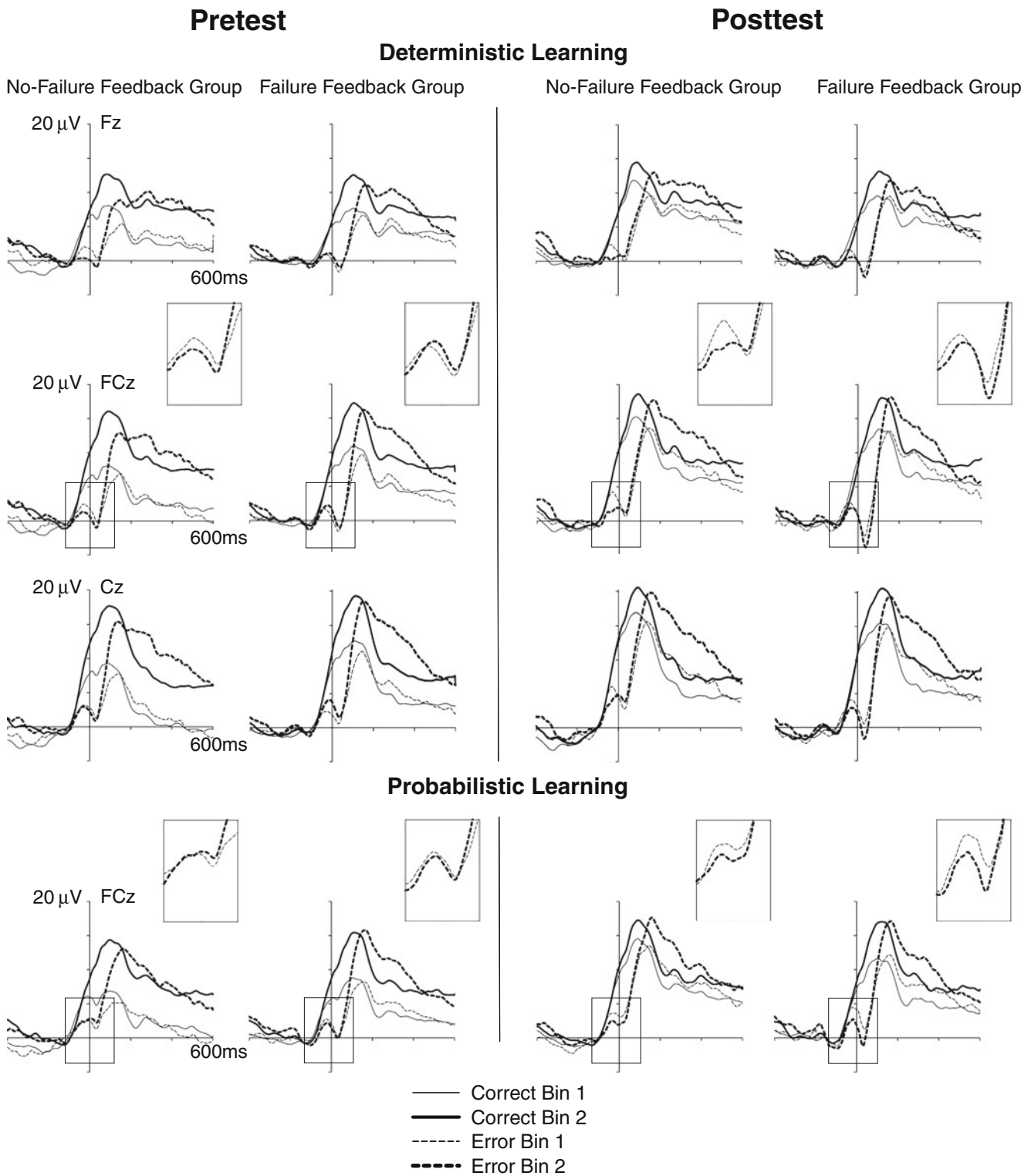
Fig. 4 Response-locked ERPs to correct (solid lines) and incorrect (dashed lines) responses displayed separately for the no-failure-feedback group and failure feedback group and the two halves of pretest (left) and posttest (right). The upper panels show the ERPs in the deterministic learning condition at electrode sites Fz, FCz, and Cz. Small boxes highlight the Ne/ERN effect at electrode FCz. The lowest panel shows the ERPs in the probabilistic learning condition at FCz. Note that the Ne/ERN amplitude was quantified peak-to-peak

Fig. 5 Bar graphs show the mean Ne/ERN amplitude at electrode FCz, separately for the three learning conditions within each group at pretest (left) and posttest (right). Error bars indicate standard errors. Note that the amplitude difference in comparison with the waveforms shown in Fig. 4 is due to the latency jitter across participants causing a reduction of the Ne/ERN in the grand average ERP
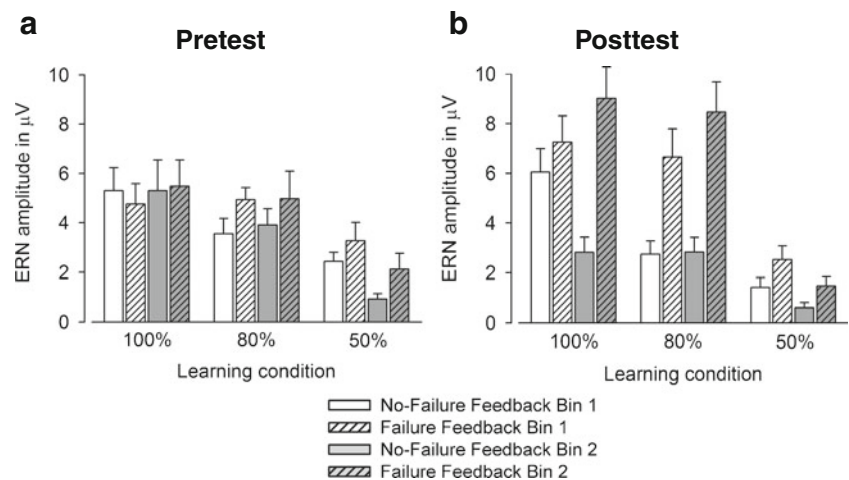
group, learning condition, and test phase, $F(2, 66) = 4.29$, $p < .05$, $\varepsilon = .75$. A follow-up ANOVA that was split by test phase confirmed that there were no group differences at pretest (p values $> .34$), whereas the Ne/ERN amplitude was larger for the failure feedback group than for the no-failure-feedback group at posttest, $F(1, 33) = 15.26$, $p < .001$. As is illustrated in Fig. 5, the failure-related Ne/ERN modulation was more pronounced in the deterministic and probabilistic learning conditions than in the chance condition, reflected in an interaction between feedback group and learning condition, $F(2, 66) = 5.64$, $p < .01$, $\varepsilon = .81$. Nonetheless, separate analyses revealed significant differences for the deterministic ($p < .01$) and probabilistic learning ($p < .001$) conditions, as well as for the chance condition ($p < .05$). Figure 5 also shows that the Ne/ERN increased from pre- to posttest for the deterministic and probabilistic learning conditions in the failure feedback group ($p < .05$ and $.01$, respectively), but not in the no-failure-feedback group (p values $> .17$).

Learning-related effects of failure on the Ne/ERN The analysis of learning-related changes in the Ne/ERN revealed significant interactions between learning condition and bin, $F(2, 66) = 7.18$, $p < .01$, and feedback group, learning condition, and bin, $F(2, 66) = 5.29$, $p < .01$, as well as a marginally significant interaction between test phase, feedback group, learning condition, and bin, $F(2, 66) = 2.88$, $p = .065$. Follow-up ANOVAs for the two test phases yielded a reliable interaction between feedback group, learning condition, and bin for posttest, $F(2, 66) = 9.66$, $p < .001$, but not for pretest, $F < 1$, indicating group differences in the modulation of the Ne/ERN across posttest only. Decomposing the interaction within each group revealed a significant main effect of bin, $F(1, 17) = 11.57$, $p < .01$, and a significant interaction between learning condition and bin, $F(2, 34) = 7.72$, $p < .01$, $\varepsilon = .71$, for the no-failure-feedback group, as well as for the failure-feedback group, $F(1, 16) = 4.66$, $p < .05$, and $F(2, 32) =$

7.97, $p < .01$, respectively. Figure 5 illustrates that the Ne/ERN decreased from the first to the second half of posttest for the deterministic learning condition in the no-failure-feedback group ($p < .01$) but increased for deterministic and probabilistic learning conditions in the failure feedback group (p values $< .05$). It should be noted that we found no evidence for learning-related changes in the Ne/ERN for either the deterministic or the probabilistic learning condition at pretest (p values $> .61$). However, more negative Ne/ERN amplitudes reliably predicted higher posterror accuracy for the deterministic and probabilistic learning conditions at both pretest [$r(35) < −.46$ and $−.36$, $p < .01$ and $.05$, respectively] and posttest [$r(35) < −.52$ and $−.44$, respectively; p values $< .01$].

Effects of learning on the correct response-related positivity The analysis of the positivity following correct responses yielded reliable main effects of bin, $F(1, 33) = 97.03$, $p < .001$, and test phase, $F(1,33) = 21.77$, $p < .001$, indicating that the correct response-related positivity increased across the two bins and from pre- to posttest. As is shown in Fig. 4, the pre–post difference was greater in Bin 1 than in Bin 2, reflected in a significant interaction between test phase and bin, $F(1, 33) = 18.88$, $p < .001$. No main effect or interaction involving the factor feedback group approached significance (p values $> .20$).

Summary of ERP findings

Consistent with our predictions, failure induction was accompanied by an increase of the Ne/ERN for the deterministic and probabilistic learning conditions. Moreover, we found an increase in Ne/ERN amplitude across posttest for the failure feedback group, whereas the Ne/ERN decreased in the deterministic learning condition for the no-failure-feedback group. Importantly, the failure manipulation did not reliably affect the positivity following correct responses.

Summary of Experiment 1

The first experiment revealed two main findings concerning the failure feedback group and the no-failure-feedback group. First, we found that failure induction resulted in an increase of the Ne/ERN for the deterministic and probabilistic learning conditions at posttest. Importantly, the amplitude enhancement was not accompanied by higher overall accuracy and, therefore, cannot simply be explained by changes in error expectancy. Instead, the Ne/ERN increase was associated with higher posterror accuracy; that is, the participants were more likely to correct an erroneous response on the next presentation of a stimulus after a variable number of intervening items. Thus, failure induction appears to increase the impact of error signals on behavioral adaptation during subsequent feedback-based learning. Second, we found a reduced Ne/ERN at the end of posttest in the deterministic learning condition for the no-failure-feedback group. At the same time, there was no significant pre–post accuracy decrease for either learning condition in this group. The latter result suggests that we successfully motivated participants to maintain task engagement by linking posttest performance to intelligence. Moreover, the present data indicate that lower responsivity of the error-monitoring system—as reflected in the decrease of the Ne/ERN from the first to the second half of posttest—is not necessarily associated with performance deficits.

Experiment 2

The results of Experiment 1 show that failure induction results in an enhanced responsivity to performance errors during a subsequent learning task. However, it remains unknown to what extent the observed effects were modulated by the fact that we manipulated the self-relevance of the learning task at posttest by linking it to intelligence. To explore this issue, we conducted a second experiment using a different posttest instruction: Participants were simply told that they were going to perform the same task as in the pretest. On the basis of findings by Tops & Boksem (2010), we expected the participants in the no-failure-feedback group to disengage from the task over the course of the posttest. In contrast to Experiment 1, accuracy, as well as Ne/ERN amplitude, should decrease from pre- to posttest in the no-failure-feedback group, since the cognitive system becomes less efficient in monitoring ongoing behavior. Although Luu et al. (2000) reported a relationship between negative affect and motivational disengagement, it seems plausible to assume that threatening participants' self-worth by self-relevant failure is associated with an "inherent" motivational boost at posttest

(Brunstein, 2000). Hence, we expected action monitoring to be less affected by motivational disengagement in the failure feedback group than in the no-failure-feedback group. This should be indicated by between-group differences in accuracy and Ne/ERN amplitude at posttest. Moreover, the findings from Experiment 1 suggest that failure feedback specifically promotes a reactive, error-driven motivational engagement. It remains an open question, however, whether we still find pre–post increases in Ne/ERN amplitude and posterror accuracy when the posttest is described in neutral terms.

Method

Participants

Forty-two undergraduate students were recruited for participation in this experiment by applying the same criteria as in Experiment 1. None of them had participated in the first experiment. Data from 9 participants had to be excluded because of poor performance in the learning task (4), excessive artifacts (1), and technical problems during EEG recording (4). The final sample thus consisted of 16 participants (11 women; mean age = 21.0 years; age range = 19–28 years) in the failure feedback group and 17 participants (12 women; mean age = 22.4 years; age range = 18–29 years) in the no-failure-feedback group.

Stimuli, tasks, and procedure

The stimuli and tasks were the same as in the first experiment. Procedural details were also identical to those in Experiment 1, except for one important difference: Before starting with the learning task at posttest, participants received the same instruction as at pretest; that is, the task was not linked to intelligence but was described in neutral terms.

Results

Manipulation check

As in Experiment 1, participants in the failure feedback group reported more negative feelings, $F(1, 31) = 14.49$, $p < .01$, and were less satisfied with their performance on the visual search task, $F(1, 31) = 13.28$, $p < .01$ (see Table 2).

Accuracy data

Response latencies shorter than 256 ms (>2 SD) or exceeding the response deadline were excluded from

further analyses.[5] Accuracy data were analyzed using the same ANOVA design as in Experiment 1. As can be seen from Fig. 6a, a significant main effect of learning condition, $F(2, 62) = 250.66$, $p < .001$, indicated that accuracy was highest for the deterministic learning condition and lowest for the chance condition (p values < .001).

Learning-related effects The analysis yielded a significant main effect of bin, $F(5, 155) = 26.30$, $p < .001$, $\varepsilon = .68$, that was qualified by an interaction between learning condition and bin, $F(10, 310) = 4.38$, $p < .001$. Moreover, an interaction between test phase and bin, $F(5, 155) = 4.29$, $p < .01$, $\varepsilon = .82$, and a marginally significant interaction between test phase, bin, and learning condition, $F(10, 310) = 1.83$, $p = .088$, $\varepsilon = .66$, indicated that the course of learning differed between pre- and posttest. Polynomial contrasts revealed linear increases in accuracy for both learning conditions across pretest only (p values < .01). At posttest, accuracy varied across the bins following a predominantly cubic ($p < .001$) and quadratic ($p < .01$) trend for the deterministic and probabilistic learning conditions, respectively (see Fig. 6a).

Effects of failure manipulation In contrast to Experiment 1, we obtained a significant interaction between feedback group and test phase, $F(1, 31) = 8.32$, $p < .01$. Separate analyses for the two groups revealed that accuracy significantly decreased from pre- to posttest in the no-failure-feedback group, $F(1, 16) = 9.85$, $p < .01$, whereas no pre–post difference in accuracy was found for the failure feedback group ($p = .23$) (see Fig. 6a).

Posterror accuracy data

As in Experiment 1, posterror accuracy rates were subjected to an ANOVA with the factors *feedback group*, *test phase*, and *learning condition*. Figure 6b shows that posterror accuracy increased with feedback validity, $F(2, 62) = 82.86$, $p < .001$.

Effects of failure manipulation The analysis yielded a significant interaction between feedback group and test phase, $F(1, 31) = 11.14$, $p < .01$. Decomposing the interaction revealed that posterror accuracy reliably de-

creased from pre- to posttest for the no-failure-feedback group, $F(1, 16) = 5.05$, $p < .05$, but increased for the failure feedback group, $F(1, 15) = 6.25$, $p < .05$ (see Fig. 6b). In contrast to Experiment 1, the analysis of posterror RT revealed a significant interaction between feedback group and test phase, $F(1, 31) = 9.46$, $p < .01$, reflecting a smaller pre–post decrease in posterror RT for the failure feedback group than for the no-failure-feedback group. Although there were no reliable between-group differences at pre- or posttest (p values > .10), this finding suggests that the failure feedback group increased posterror accuracy at the cost of relatively longer posterror RT.

Summary of behavioral findings

Whereas the no-failure-feedback group showed worse performance at posttest, overall accuracy did not differ between pre- and posttest in the failure feedback group. Similar to Experiment 1, non-learning-related dynamics of performance became prevalent for both groups at posttest. While these findings are likely to reflect reduced task engagement toward the end of posttest in the no-failure-feedback group, the pattern of performance in the failure feedback group suggests that participants produced relatively more errors of commission during posttest, particularly in the deterministic learning condition. In support of this, RTs decreased from pre- to posttest and were shorter in the deterministic than in the probabilistic learning condition on error trials (p values < .001). Consistent with Experiment 1, failure induction was specifically associated with increased error correction rates.
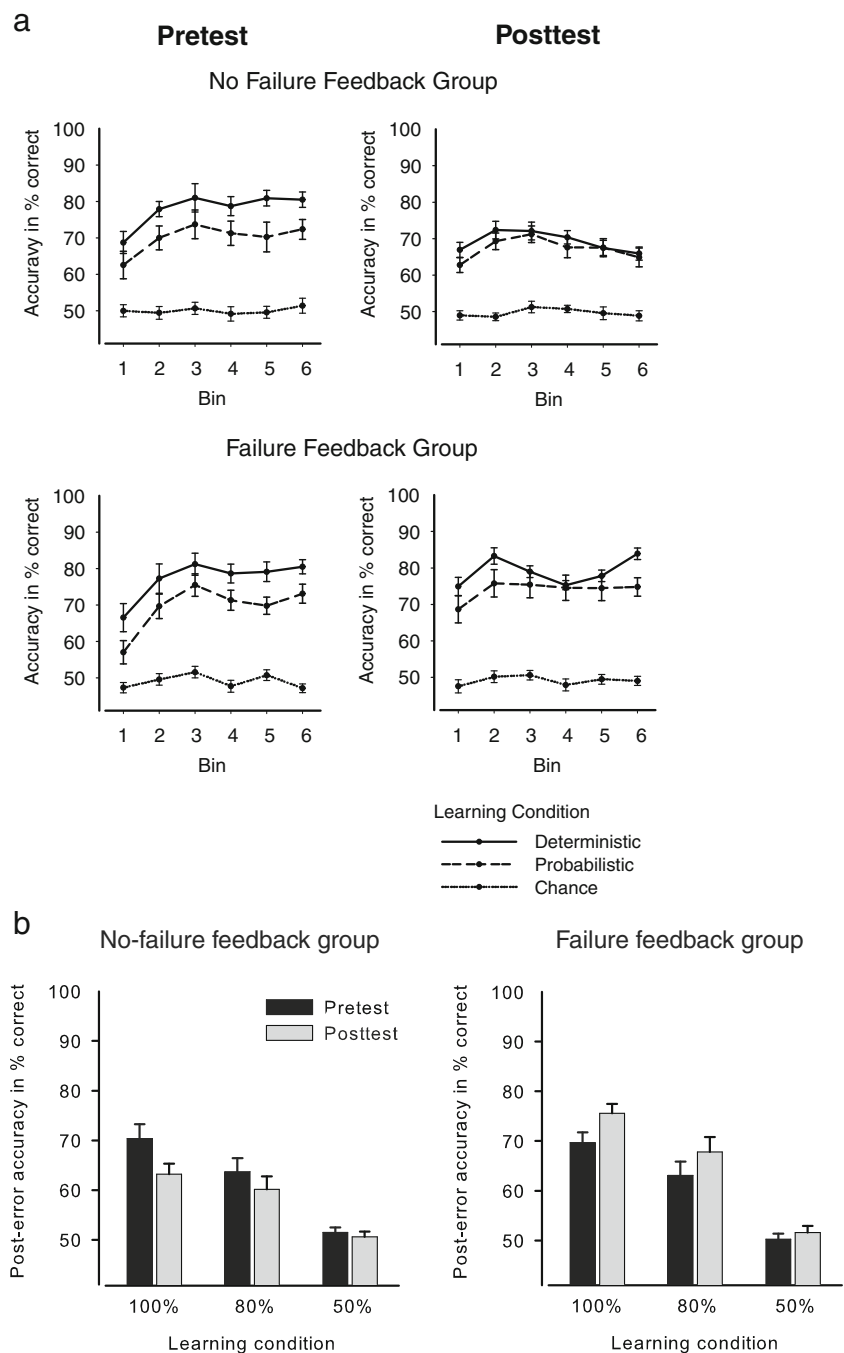
ERP data

Figure 7 shows the ERPs to correct and incorrect responses, separately for both feedback groups in the first and second halves of pre- and posttest (see also Fig. 8). The peak-to-peak measures of the Ne/ERN and the mean amplitude of the correct response-related positivity were analyzed using the same ANOVA design as in Experiment 1. Consistent with Experiment 1, both the Ne/ERN and the correct response-related positivity increased with feedback validity, $F(2, 62) = 47.89$ and 61.42, respectively; p values < .001]. Below, the failure-related effects on the Ne/ERN are reported first, followed by the results for the correct-response-related positivity.

General effects of failure on the Ne/ERN The analysis revealed a reliable main effect of feedback group, $F(1, 31) = 4.74$, $p < .05$, that was qualified by a significant interaction between feedback group and test phase, $F(1, 31) = 17.41$, $p < .001$. As is illustrated in Fig. 8, the Ne/ERN decreased

---

[5] The deadline was exceeded by 2.2% (no-failure-feedback group) and 2.3% (failure feedback group) of the responses. At pretest, mean RTs on correct and incorrect trials were 446 ms ($SD=33$ ms) and 440 ms ($SD=30$ ms) for the no-failure-feedback group and 427 ms ($SD=23$ ms) and 423 ms ($SD=22$ ms) for the failure feedback group. At posttest, mean RTs on correct and incorrect trials were 404 ms ($SD=32$ ms) and 393 ms ($SD=26$ ms) for the no-failure-feedback group and 411 ms ($SD=27$ ms) and 403 ms ($SD=26$ ms) for the failure feedback group.

from pre- to posttest for the no-failure-feedback group, $F(1, 16) = 10.97$, $p < .01$, but increased from pre- to posttest for the failure feedback group, $F(1, 15) = 6.14$, $p < .05$.

Learning-related effects of failure on the Ne/ERN As in Experiment 1, we obtained significant interactions between learning condition and bin, $F(2, 62) = 11.88$, $p < .001$, and test phase, feedback group, learning condition, and bin, $F(2, 62) = 3.34$, $p < .05$. Separate pre- and posttest analyses yielded a significant interaction between feedback group, learning condition, and bin for posttest, $F(2, 62) = 4.93$,

$p < .05$, $\varepsilon = .84$, but not for pretest, $F < 1$. At posttest, a significant interaction between learning condition and bin was obtained for the failure feedback group only, $F(2, 30) = 6.56$, $p < .01$ (no-failure-feedback group: $p = .30$). Figure 8 shows that the Ne/ERN increased over the course of posttest for the deterministic and probabilistic learning conditions (p values $< .05$) but tended to decrease for the chance condition ($p = .056$). Consistent with Experiment 1, greater Ne/ERN amplitudes predicted higher posterror accuracy for deterministic and probabilistic learning conditions at both pretest [$r(33) = -.38$ and $-.40$, respectively;

Fig. 7 Response-locked ERPs to correct (solid lines) and incorrect responses (dashed lines) displayed separately for the no-failure feedback group and failure feedback group and the two halves of pretest (left) and posttest (right). The upper panels show the ERPs in the deterministic learning condition at electrode sites Fz, FCz, and Cz. Small boxes highlight the Ne/ERN effect at electrode FCz. The lowest panel shows the ERPs in the probabilistic learning condition at FCz. Note that the Ne/ERN amplitude was quantified peak-to-peak
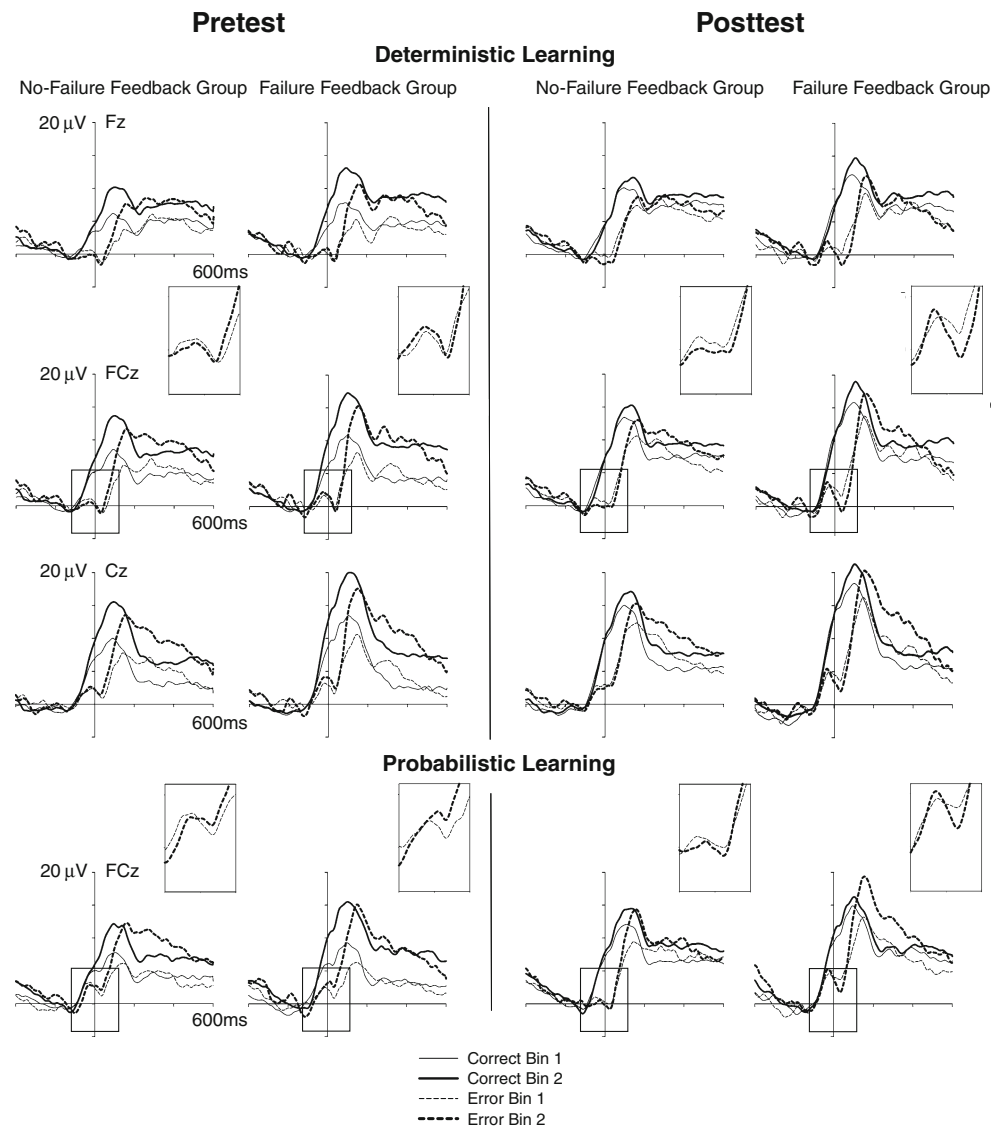
p values < .05] and posttest [*r*(33) = –.54 and –.67, respectively; *p* values < .01].

Effects of learning on the correct response-related positivity The positivity on correct trials increased from Bin 1 to Bin 2, $F(1, 31) = 70.61$, $p < .001$, and from pre- to posttest, $F(1, 31) = 16.13$, $p < .001$. As was indicated by a significant interaction between bin and test phase, $F(1, 31) = 20.83$, $p < .001$, the amplitude difference between the two bins was larger at pretest (see Fig. 7). No effect of failure manipulation was found (all *p* values > .18).

Summary of ERP findings

In line with Experiment 1, these findings demonstrate that failure feedback resulted in an enhancement of the Ne/ERN. Moreover, we observed an increase of the Ne/ERN over the course of posttest for the deterministic and probabilistic learning conditions in the failure feedback group. In contrast, the Ne/ERN decreased from pre- to posttest for the no-failure-feedback group. This decrease was much more pronounced than in Experiment 1, where the no-failure group showed a reduced Ne/ERN for the deterministic learning condition in the second half of posttest only.

Summary of Experiment 2

The results of Experiment 2 confirm our prediction that prolonged task performance results in motivational disengagement for the no-failure-feedback group and provide further support for the notion that failure feedback amplifies sensitivity to internal indicators of errors. First, in contrast to Experiment 1, participants in the no-failure-feedback group showed worse performance at posttest. Consistent with former studies (Boksem et al., 2006; Tops & Boksem,

Fig. 8 Bar graphs show the mean Ne/ERN amplitude at electrode FCz, separately for the three learning conditions within each group at pretest (left) and posttest (right). Error bars indicate standard errors. Note that the amplitude difference in comparison with the waveforms shown in Fig. 7 is due to the latency jitter across participants, resulting in a reduction of the Ne/ERN in the grand average ERP
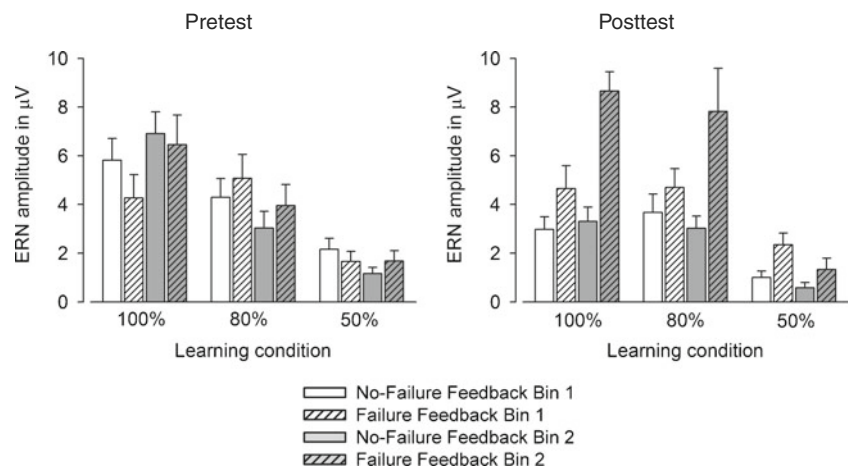
2010), the performance impairments were accompanied by a decrease of the Ne/ERN in the deterministic learning condition. Second, the results for the failure feedback group replicate the findings from Experiment 1. In both experiments, failure induction was associated with an increase in Ne/ERN amplitude, as well as higher posterror accuracy in the deterministic and probabilistic learning conditions.

General discussion

The aim of the present study was to examine whether self-relevant failure experiences affect the ability to use negative response outcomes (errors) for behavioral adaptation in a subsequent probabilistic learning task and whether the Ne/ERN is related to these differences in error processing. Consistent with previous research that has established a link between the Ne/ERN and the evaluation of the affective and motivational significance of an error (e.g., Gehring et al., 1993; Hajcak et al., 2005; Luu et al., 2003), we found that failure feedback resulted in an increase of the Ne/ERN. Extending these findings, we observed a correspondence between failure-induced Ne/ERN amplitude enhancement and error-related behavioral adjustments during learning. Notably, the increase in Ne/ERN amplitude at posttest was associated not with better overall performance, but with higher posterror accuracy—that is, a higher proportion of correct choices on the next presentation of the target on which an erroneous response occurred. It is important to note that this error-related change in behavior is unlikely to reflect an unspecific increase of attention or arousal, since stimuli were presented in random order. Instead, the behavioral adjustment appears to specifically relate to a higher impact of negative reinforcement learning signals at posttest (Frank et al., 2005; Holroyd & Coles, 2002). In line with this view, the failure-induced Ne/ERN increase was more pronounced in the second half of the posttest when the participants were better able to internally represent an

incorrect response. Moreover, failure feedback affected the Ne/ERN amplitude in the two learning conditions, but not in the chance condition.

Corroborating prior findings concerning the effects of prolonged task performance on action monitoring (Boksem et al., 2006; Tops & Boksem, 2010), Ne/ERN amplitude and accuracy decreased with time on task for the no-failure-feedback group in Experiment 2. In contrast, both groups in Experiment 1, as well as the failure feedback group in Experiment 2, showed comparable overall performance at pre- and posttest. Thus, increasing the motivational significance of posttest by linking the learning task to intelligence (Experiment 1) or giving prior negative feedback as to participants' intellectual abilities (Experiments 1 and 2) preserved task engagement. However, we observed different fluctuations of the Ne/ERN across posttest: The Ne/ERN amplitude decreased for the deterministic learning condition in the no-failure-feedback group but increased for deterministic and probabilistic learning conditions in the failure feedback groups.

Importantly, the observed effects of failure feedback on the Ne/ERN are not attributable to preexperimentally existent individual differences in trait-level negative affect or punishment sensitivity, which have been related to increased reactivity of the error-monitoring system (e.g., Boksem et al., 2006; Hajcak et al., 2004; Luu et al., 2000). Instead, our results are based on direct manipulations of affective-motivational state. As was indicated by participants' self-reports, the experimental manipulation was successful in inducing self-relevant failure and, by this, negative feelings. Importantly and as was expected, there was no evidence for between-group differences in accuracy or Ne/ERN amplitude at pretest in either experiment, whereas clear group differences were obtained at posttest. Furthermore, the observed Ne/ERN modulations cannot simply be explained by within- or between-group differences in overall performance (see Yeung, 2004). Our findings thus provide strong evidence that the functioning

of the ACC, as reflected by the Ne/ERN, is sensitive to the affective and motivational context of an action.

In particular, the present study confirms and extends the finding by Wiswede, Münte, & Rüsseler, (2009) that short-term manipulations of negative affect are reflected in modulations of the Ne/ERN. Whereas in the study by Wiswede, Münte, and Rüsseler, the Ne/ERN was measured during the affective manipulation, our data show that the effects of failure feedback generalize to a different task. This is an important new finding, suggesting that failure-induced negative affective state can bias information processing at a broader task-unspecific level. This view is largely consistent with previous studies that have established a relation between the Ne/ERN and individual differences in negative affective dispositions (Boksem et al., 2006; Hajcak et al., 2003, 2004). The present results suggest that state and trait variations in negative affect might be associated with similar changes in the functioning of the internal error-monitoring system.

According to the R-L theory, the Ne/ERN constitutes a predictive error signal that is used by the ACC to select and reinforce appropriate actions (Holroyd & Coles, 2002). Our finding that the failure-related Ne/ERN enhancement was accompanied by higher posterror accuracy supports this view. Similar conclusions have been drawn from previous studies demonstrating that activity in the medial prefrontal error-processing system predicted the correctness of future responses (Hester, Barre, Murphy, Silk, & Mattingley, 2008; van der Helden, Boksem, & Blom, 2010). Moreover, findings by Frank et al. (2005) suggest a specific relation between Ne/ERN amplitude and a bias to learn more from bad than from good choices. Applied to the present study, this idea implies that individuals exposed to self-relevant failure use error signals more efficiently to determine which response to avoid on subsequent presentations of a stimulus. While the Ne/ERN was primarily sensitive to the failure manipulation, the correct-response-related positivity increased over the course of learning and from pre- to posttest for all participants. This dissociation seems at odds with a recent proposal according to which the correct-response-related positivity, rather than the negativity on incorrect trials, is subject to experimentally induced change (e.g., Holroyd, Pakzad-Vaezi, & Krigolson, 2008) and, instead, suggests that both components reflect separable processes. Furthermore, it should be noted that alternative theoretical approaches to the Ne/ERN can likewise account for the failure-related amplitude increase. The conflict-monitoring theory (Botvinick et al., 2001; Yeung, Botvinick, & Cohen, 2004) holds that the Ne/ERN reflects posterror conflict due to the simultaneous activation of incorrect and correct responses, with the latter arising from continued processing of the target stimulus after the erroneous response is produced. Given that participants in the failure feedback group are likely to be highly motivated to perform well at posttest, larger Ne/ERN amplitudes might indicate an increased posterror activation of the correct response as a consequence of more efficient target processing (see Yeung, 2004). Similarly, more efficient continued processing of the imperative stimulus would facilitate the detection of mismatch between the actual and the intended responses (mismatch theory; Falkenstein, Hohnsbein, Hoormann, & Blanke, 1990; Gehring et al., 1993).

Although behavioral and ERP data in the first half of posttest imply that the initial effects of the experimental manipulations were comparable for the no-failure-feedback group in Experiment 1 and the failure feedback group in Experiment 2, the two groups were characterized by distinct changes in Ne/ERN amplitude in the second half of the posttest. Notably, a pattern of results similar to that observed for the no-failure-feedback group—that is, large initial Ne/ERN amplitudes followed by fast reductions—has been reported for individuals characterized by high habitual intrinsic engagement (Tops & Boksem, 2010). The susceptibility of intrinsic engagement to increasing boredom during prolonged performance of monotonous tasks fits our finding that the Ne/ERN decreased in the deterministic condition, but not in the more challenging probabilistic learning condition. In contrast, the pattern of sustained monitoring in the failure feedback groups parallels findings for individuals scoring high on constraint (Tops & Boksem, 2010). From this perspective, the Ne/ERN enhancement across posttest might reflect negatively motivated engagement resulting from worry and concerns about mistakes that become more salient with learning.

In sum, the findings of this study suggest that failure induction results in a shift toward reactive control, denoting the tendency to recruit control processes when a (negative) event has already occurred (as opposed to *proactive control*; Braver, Gray, & Burgess, 2007; Tops, Boksem, Luu, & Tucker, 2010). Failure feedback thus appears to induce a state in which participants are particularly vigilant to potential threats and negative response outcomes. In support of this, several studies have established a relation between the activity of the medial prefrontal action-monitoring system and sensitivity to negative stimuli and events. For example, larger Ne/ERN amplitudes have been linked to punishment sensitivity (Boksem et al., 2006), learning from errors (Frank et al., 2005), or defensive motivation (Hajcak & Foti, 2008). Similarly, social stress reactivity has been found to increase the sensitivity to internal indicators of errors (Cavanagh, Frank, & Allen, 2011). Hence, it is not surprising that in the present study, effects of the failure manipulation were observed for incorrect but not for correct responses.

Although the present findings highlight the need to explore the performance-monitoring system in terms of both cognitive and affective/motivational mechanisms, the

precise nature of the processes that mediate the observed effects of failure feedback remains to be determined. Whereas previous EEG/ERP source localization studies and functional imaging (fMRI) studies on error processing typically have identified a cluster of subregions within the medial prefrontal cortex—specifically, the ACC—as critically involved in monitoring ongoing behavior (for a review, see Taylor, Stern, & Gehring, 2007), broader cognitive-affective control circuits have been suggested incorporating areas commonly linked to affect (e.g., amygdala) and motivation (e.g., nucleus accumbens; Frank & Claus, 2006; Pessoa, 2008). Crucially, the proposed architecture and functioning of those networks imply that cognitive, affective, and motivational influences on executive control operate in concert to regulate behavior in accordance with specific situational demands.

Conclusions

To conclude, the present results show that self-relevant failure results in an enhancement of Ne/ERN amplitude, reflecting an increased sensitivity to errors. Failure-related Ne/ERN modulations were not accompanied by changes in overall performance but were specifically related to posterror behavioral adjustments, suggesting a shift toward a reactive, error-driven mode of behavior control. Furthermore, the Ne/ERN increased across the learning task after, but not before, failure induction. Thus, failure feedback appeared to amplify learning-related changes of the Ne/ERN, possibly reflecting more efficient reactive monitoring during later stages of learning when participants are better able to internally represent incorrect responses. In line with a growing body of evidence indicating a close interaction between cognition and emotion in executive control (Pessoa, 2008, 2009), the present findings emphasize the importance of factors related to affective and motivational state in elucidating the neural mechanisms of performance monitoring.

References

Boksem, M. A. S., Meijman, T. F., & Lorist, M. M. (2006a). Mental fatigue, motivation and action monitoring. *Biological Psychology, 72,* 123–132.

Boksem, M. A. S., Tops, M., Wester, A. E., Meijman, T. F., & Lorist, M. M. (2006b). Error-related ERP components and individual differences in punishment and reward sensitivity. *Brain Research, 1101,* 92–101.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108,* 624–652.

Braver, T. S., Gray, J. R., & Burgess, G. C. (2007). Explaining the many varieties of working memory variation: Dual mechanisms of cognitive control. In A. R. A. Conway, C. Jarrold, M. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 76–106). Oxford: Oxford University Press.

Brunstein, J. C. (1994). Dispositional action control as a predictor of how people cope with academic failure. In J. Kuhl & J. Beckmann (Eds.), *Volition and personality: Action versus state orientation* (pp. 341–350). Seattle, WA: Hogrefe & Huber.

Brunstein, J. C. (2000). Motivation and performance following failure: The effortful pursuit of self defining goals. *Applied Psychology, 49,* 340–356.

Brunstein, J. C., & Gollwitzer, P. M. (1996). Effects of failure on subsequent performance: The importance of self-defining goals. *Journal of Personality and Social Psychology, 70,* 395–407.

Brunstein, J. C., & Olbrich, E. (1985). Personal helplessness and action control: Analysis of achievement-related cognitions, self-assessments, and performance. *Journal of Personality and Social Psychology, 48,* 1540–1551.

Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences, 4,* 215–222.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology, 67,* 319–333.

Cavanagh, J. F., Frank, M. J., & Allen, J. J. B. (2011). Social stress reactivity alters reward and punishment learning. *Social Cognitive and Affective Neuroscience, 6,* 311–320.

Cooper, J. A., Sagar, H. J., Jordan, N., Harvey, N. S., & Sullivan, E. V. (1991). Cognitive impairment in early, untreated Parkinson's disease and its relationship to motor disability. *Brain, 114,* 2095–2122.

Dweck, C. S., & Reppucci, N. D. (1973). Learned helplessness and reinforcement responsibility in children. *Journal of Personality and Social Psychology, 25,* 109–116.

Elliott, E. S., & Dweck, C. S. (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology, 54,* 5–12.

Eppinger, B., & Kray, J. (2011). To choose or to avoid: Age differences in learning from positive and negative feedback. *Journal of Cognitive Neuroscience, 23,* 41–52.

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

Eshel, N., & Roiser, J. P. (2010). Reward and punishment processing in depression. *Biological Psychiatry, 68,* 118–124.

Falkenstein, M., Hohnsbein, J., Hoormann, J., & Blanke, L. (1990). Effects of errors in choice reaction tasks on the ERP under focused and divided attention. In C. H. M. Brunia, A. W. K. Gaillard, & A. Kok (Eds.), *Psychophysiological brain research* (pp. 192–195). Tilburg, The Netherlands: Tilburg University Press.

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.

Falkenstein, M., Hoormann, J., Christ, S., & Hohnsbein, J. (2000). ERP components on reaction errors and their functional significance: A tutorial. *Biological Psychology, 51,* 87–107.

Frank, M. J., & Claus, E. D. (2006). Anatomy of a decision: Striato-orbitofrontal interactions in reinforcement learning, decision making and reversal. *Psychological Review, 113,* 300–326.

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

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

Gehring, W. J., Himle, J., & Nisenson, L. G. (2000). Action-monitoring dysfunction in obsessive–compulsive disorder. *Psychological Science, 11,* 1–6.

Geisser, S., & Greenhouse, S. W. (1958). An extension of box's results on the use of the f-distribution in multivariate analysis. *Annals of Mathematical Statistics, 29,* 885–891.

Hajcak, G., & Foti, D. (2008). Errors are aversive: Defensive motivation and the error-related negativity. *Psychological Science, 19,* 103–108.

Hajcak, G., McDonald, N., & Simons, R. F. (2003). Anxiety and error-related brain activity. *Biological Psychology, 64,* 77–90.

Hajcak, G., McDonald, N., & Simons, R. F. (2004). Error-related psychophysiology and negative affect. *Brain and Cognition, 56,* 189–197.

Hajcak, G., Moser, J. S., Yeung, N., & Simons, R. F. (2005). On the ERN and the significance of errors. *Psychophysiology, 42,* 151–160.

Hester, R., Barre, N., Murphy, K., Silk, T. J., & Mattingley, J. B. (2008). Human medial frontal cortex activity predicts learning from errors. *Cerebral Cortex, 18,* 1933–1940.

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., Pakzad-Vaezi, K. L., & Krigolson, O. E. (2008). The feedback correct-response related positvity: Sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology, 45,* 688–697.

Krohne, H. W., Egloff, B., Kohlmann, C.-W., & Tausch, A. (1996). Untersuchung mit einer deutschen Form der Positive and Negative Affect Schedule (PANAS). *Diagnostica, 42,* 139–156.

Kuhl, J. (1994). Action and state orientation: Psychometric properties of the action control scales (ACS-90). In J. Kuhl & J. Beckmann (Eds.), *Volition and personality: Action versus state orientation* (pp. 47–59). Göttingen, Germany: Hogrefe.

Larson, M. J., Perlstein, W. M., Stigge-Kaufman, D., Kelly, K. G., & Dotson, V. M. (2006). Affective context-induced modulation of the error-related negativity. *NeuroReport, 17,* 329–333.

Lehrl, S. (1977). *Mehrfachwahl-Wortschatz-Test B (MWT-B).* Erlangen: Straube.

Luu, P., Collins, P., & Tucker, D. M. (2000). Mood, personality, and self-monitoring: Negative affect and emotionality in relation to frontal lobe mechanisms of error monitoring. *Journal of Experimental Psychology. General, 129,* 43–60.

Luu, P., Tucker, D. M., Derryberry, D., Reed, M., & Poulsen, C. (2003). Electrophysiologic responses to errors and feedback in the process of action regulation. *Psychological Science, 14,* 47–53.

Mikulincer, M. (1994). *Human learned helplessness: A coping perspective.* New York: Plenum.

Nieuwenhuis, S., Ridderinkhof, K. R., Talsma, D., Coles, M. G. H., Holroyd, C. B., Kok, A., et al. (2002). A computational account of altered error processing in older age: Dopamine and the error-related negativity. *Cognitive, Affective, & Behavioral Neuroscience, 2,* 19–36.

Olvet, D. M., & Hajcak, G. (2011). The error-related negativity relates to sadness following mood induction among individuals with high neuroticism. *Social, Cognitive, and Affective Neuroscience.* Epub ahead of print. doi:10.1093/scan/nsr007

Pessoa, L. (2008). On the relationship between emotion and cognition. *Nature Reviews Neuroscience, 9,* 148–158.

Pessoa, L. (2009). How do emotion and motivation direct executive control? *Trends in Cognitive Sciences, 13,* 160–166.

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., & Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science, 306,* 443–447.

Seligman, M. E. P. (1975). *Helplessness: On depression, development, and death.* San Francisco: Freeman.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory, 6,* 174–215.

Taylor, S. F., Stern, E. R., & Gehring, W. J. (2007). Neural systems for error monitoring: Recent findings and theoretical perspectives. *The Neuroscientist, 13,* 160–172.

Tops, M., & Boksem, M. A. S. (2010). Absorbed in the task: Personality measures predict engagement during task performance as tracked by error negativity and asymmetrical frontal activity. *Cognitive, Affective, & Behavioral Neuroscience, 20,* 441–453.

Tops, M., Boksem, M. A. S., Luu, P., & Tucker, D. M. (2010). Brain substrates of behavioral programs associated with self-regulation. *Frontiers in Psychology, 1*(Art. 152), 1–14. doi:10.3389/fpsyg.2010.0152

van der Helden, J., Boksem, M. A. S., & Blom, J. H. G. (2010). The importance of failure: Feedback related negativity predicts motor learning efficiency. *Cerebral Cortex, 20,* 1596–1603.

Von Zerssen, D. (1976). *Klinische Selbstbeurteilungsskalen (KSb-S) aus dem Münchner Psychiatrischen Informations-System (PSYCHIS München). Allgemeiner Teil.* Weinheim: Beltz Test.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology, 54,* 1063–1070.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale–R.* New York: Psychological Corporation.

Wiswede, D., Münte, T. F., Goschke, T., & Rüsseler, J. (2009). Modulation of the error-related negativity by induction of short-term negative affect. *Neuropsychologia, 47,* 83–90.

Wiswede, D., Münte, T. F., & Rüsseler, J. (2009). Negative affect induced by derogatory verbal feedback modulates the neural signature of error detection. *Social Cognitive and Affective Neuroscience, 4,* 227–237.

Wortman, C. B., & Brehm, J. W. (1975). Responses to uncontrollable outcomes: An integration of reactance and the learned helplessness model. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 8, pp. 277–336). New York: Academic Press.

Yeung, N. (2004). Relating cognitive and affective theories of the error-related negativity. In M. Ullsperger & M. Falkenstein (Eds.), *Errors, conflicts, and the brain: Current opinions on response monitoring* (pp. 63–70). Leipzig: MPI of Cognitive Neuroscience.

Yeung, N., Botvinick, M. M., & Cohen, J. D. (2004). The neural basis of error detection: Conflict monitoring and the error-related negativity. *Psychological Review, 111,* 931–959.