To cite this article:

Walentowska, W., Severo, M. C., Moors, A., & Pourtois, G. (in press). When the outcome is different than expected: Subjective expectancy shapes reward prediction error at the FRN level. *Psychophysiology*.

WHEN THE OUTCOME IS DIFFERENT THAN EXPECTED: SUBJECTIVE EXPECTANCY SHAPES REWARD PREDICTION ERROR AT THE FRN LEVEL

Wioleta Walentowska (1,2), Mario Carlo Severo (2), Agnes Moors (3), and Gilles Pourtois (2)

(1) Psychophysiology Laboratory, Institute of Psychology, Jagiellonian University, Krakow, Poland
(2) Cognitive & Affective Psychophysiology Laboratory, Ghent University, Belgium
(3) Research Group of Quantitative Psychology and Individual Differences, Center for Social and Cultural Psychology, KU Leuven, Belgium

Running head: Unsigned RPE triggers the FRN

Corresponding address:

Wioleta Walentowska

Psychophysiology Laboratory

Jagiellonian University

Ingardena 6

30-060 Krakow

Poland

e-mail address: w.walentowska@gmail.com

phone number: +48 12 663 24 08

Impact Statement

In this study, we focused on feedback processing and the FRN, as well as subsequent P3b ERP component. We found that the FRN size was similar for worse- and better-than-expected events, suggesting that salience probably drives the detection of mismatches between outcome and expectancy. Following the FRN, the P3b was selectively increased for positive feedback when outcome and expectancy mismatched. These results are compatible with a two-stage PM model, according to which the early detection of violations between action and prediction at the FRN level is followed by an update of the positive feedback's value at the P3b level.

Abstract

Converging evidence in human electrophysiology suggests that evaluative feedback provided during performance monitoring (PM) elicits two distinctive and successive ERP components: the Feedback-Related Negativity (FRN) and the P3b. Whereas the FRN has previously been linked to reward prediction error (RPE), the P3b has been conceived as reflecting motivational or attentional processes following the early processing of the RPE, including action value updating. However, it remains unclear whether these two consecutive neurophysiological effects depend on the direction of the unexpectedness (better- or worse-than-expected outcomes; signed RPE) or instead only on the degree of unexpectedness irrespective of direction (i.e., unsigned RPE). To address this question, we devised an experiment in which we manipulated the objective reward probability and the subjective reward expectancy (via instructions) in a factorial within-subject design, and explored amplitude changes of the FRN and the P3b. Sixty-four channel EEG was recorded while 32 participants performed a speeded Go/NoGo task in which evaluative feedback based on the reward probability either violated expectancy (thereby creating a RPE) or did not. This violation corresponded either to better- or worse-than-expected events. Results showed that the FRN was larger when RPE occurred than when it did not, but irrespective of the direction of this violation. Interestingly, in these two conditions, action value was updated for the positive feedback selectively, as shown by the P3b amplitude. These results obey a two-stage model of PM assuming that unsigned RPE is first rapidly detected (FRN level) before the positive feedback's value is updated selectively (P3b effect).

Keywords: ERP; FRN; P3b, P3a, Performance Monitoring; Reward Prediction Error (RPE); Action Value Updating; Unexpectedness

1. Introduction

Performance monitoring (PM) is a very important mental ability, which is essential to foster goal-adaptive behavior and self-regulation (Botvinick & Braver, 2015; Inzlicht, Schmeichel, & Macrae, 2014). PM is fairly complex and likely involves different computations, performed at different levels within a hierarchical system implemented in the prefrontal cortex and inter-connected dopaminergic regions located deeper in the brain, including the basal ganglia and the striatum (Ullsperger, Fischer, Nigbur, & Endrass, 2014). Although this process is rather sophisticated, it is also flexible and dynamic, using either internal/motor-related cues or external/feedback information depending on the evidence available, with the aim to extract the current action value, and update it in case outcome and expectation mismatched with each other. Feedback processing guiding PM is clearly visible when these internal cues are lacking, or when processing of these cues is incomplete. For example, if uncertainty about an action's value is high at the time of response onset, then evaluative feedback provided after the response helps to reduce it and is preferentially processed. The electrophysiological correlates of these dynamic PM effects have been studied extensively in the past, and well-defined ERP components have been identified. At the response level, the Error-Related Negativity (ERN; see Falkenstein, Hohnsbein, Hoormann, & Blanke, 1991; Gehring, Goss, Coles, Meyer, & Donchin, 1993) has been put forward as the first stage following action execution that allows to detect mismatches between action and goal or prediction. At the feedback level, when external-based PM operates, the Feedback-Related Negativity (FRN) is usually considered as the counterpart of the ERN, sharing many similarities with it. The FRN is a phasic, negative-going wave, peaking around 250–300 ms after feedback onset over fronto-central locations along the midline. Its amplitude is larger for negative relative to positive performance feedback (Miltner, Braun, & Coles, 1997; Nieuwenhuis, Holroyd, Mol, & Coles, 2004; von Borries, Verkes, Bulten, Cools, & de Bruijn, 2013), and for unexpected compared to expected events (Hajcak, Moser, Holroyd, & Simons 2007; Pfabigan, Alexopoulos, Bauer, Lamm, & Sailer, 2011). Hence, *valence* (Hajcak, Moser, Holroyd, & Simons, 2006) and *expectedness* (Ferdinand, Mecklinger, Kray, & Gehring, 2012) are two main ingredients that account

for the generation of the FRN during PM (see also San Martin, 2012). Expectedness is related to *expectancy* or the degree to which a person expects to receive a certain feedback following her action: If expectancy to receive a positive feedback is high/low, the occurrence of a negative feedback or the absence of a reward will be unexpected/expected.

However, the functional meaning or specific contribution of the FRN to PM remains debated in the literature (Hajihosseini & Holroyd, 2013; Proudfit, 2015; Ullsperger, 2017; Ullsperger et al., 2014). Notably, the type of unexpectedness driving the FRN amplitude modulations remains unclear. According to the dominant reinforcement learning framework (Holroyd & Coles, 2002), the FRN is primarily generated when the outcome is worse than expected, which means that a 'signed' or directional reward prediction error (RPE) occurs. A clear FRN is usually observed when the outcome is worse than expected (negative RPE): Participants expect the outcome to be rewarding (e.g. monetary gain), but it unexpectedly turns out to be non-rewarding. A smaller and weaker FRN is observed when the outcome is better than expected (positive RPE): Participants expect the outcome to be non-rewarding but it unexpectedly turns out to be rewarding. In the former situation, it is postulated that the outcome (i.e., feedback) is especially informative for the participants because it allows to improve learning and adapt behavior accordingly, while this is less the case in the latter situation (Frank, Seeberger, & O'reilly, 2004; Sambrook & Gosslin, 2015; Walsh & Anderson, 2012). At slight variance with this theory, the salience prediction error account (SPE; Alexander & Brown, 2011; Oliveira, McDonald, & Goodman, 2007) suggests that the medial prefrontal cortex, and more specifically, the dorsal anterior cingulate cortex (ACC), which is thought to be the main intracranial generator of the FRN (Gehring & Willoughby, 2002; Miltner et al., 1997; Yeung, Holroyd, & Cohen, 2004), is sensitive to mismatches regardless of their sign, thereby responding equally strongly to better-than-expected or worse-than-expected outcomes since they are both salient (Hauser, Iannaccone, Stämpfli, Drechsler, Brandeis, Walitza, & Brem, 2014; Soder & Potts, 2018; Talmi, Atkinson, & El-Deredy, 2013; van der Veen, van der Molen, van der Molen, & Franken, 2016). The question thus remains unsolved whether the FRN codes for a signed or an unsigned RPE.

In the current study, we sought to address this question. To this end, we devised a within-subject experiment in which four main experimental conditions were created by crossing reward probability (low or high across different blocks), and expectancy based on specific instructions (low or high across different blocks as well). These two main factors were embedded in a factorial design. As a result, participants performed the exact same simple decision-making task in four different contexts that differed regarding reward probability and expectancy. More specifically, participants performed this task and expected it to be relatively easy (i.e., yielding many positive feedbacks) or more difficult instead (i.e., resulting in a lower number of positive feedbacks received). Crucially, unknown to them, this prior expectation was either confirmed or violated by adjusting reward probability in a blockwise fashion, leading eventually to better- or worse-than-expected outcomes. According to the reinforcement learning account (Holroyd & Coles, 2002), the FRN, defined as the difference wave between negative and positive feedback, should be the largest in the condition where the positive feedback is expected but the outcome eventually violates this expectation and a negative feedback is provided instead (i.e., worse-than-expected event). Alternatively, if salience is the key feature underlying the generation of prediction errors at the FRN level (Hauser et al., 2014; Oliveira et al., 2007; Soder & Potts, 2018; Talmi et al., 2013), then this ERP component should be equally large for better-than-expected and worse-than-expected outcomes. Translated to a statistical analysis applied to this factorial design, both frameworks predict a significant interaction effect between reward probability and expectancy, but the interactions take different shapes in each framework. In the reinforcement learning account, the interaction reflects the modulation (increase) of the FRN amplitude for a specific or unique combination of probability (low) and expectancy (high) for negative feedback. In the salience account, two combinations of probability and expectancy (i.e., when they mismatch with each other, but irrespective of the direction of this deviation) both lead to an equally large (and statistically undistinguishable) FRN component.

Noteworthy, when PM is based on the processing of external evaluative feedback, it does not terminate at the offset of the FRN. Following the FRN, evaluative feedback usually elicits a clear P3b at posterior parietal leads along the midline (Ullsperger et al., 2014). Whereas the FRN is thought to reflect

an early evaluation process during which the correspondence between action and goal is processed, the subsequent P3b translates this information into specific attentional, motivational, or perhaps working memory (WM) processes (Donchin & Coles, 1998; Polich, 2007; Verleger, 1997; Verleger, Jaskowski, & Wauschkuhn, 1994). A prominent proposal is that the P3b component reflects action value updating during PM after detection of a mismatch between action and goal (Ullsperger, 2017; Ullsperger et al., 2014). Many previous ERP studies on PM have focused primarily on the FRN (see Sambrook & Gosslin, 2015; San Martin, 2012) but only a few of them have also explored this subsequent action value updating process at the P3b level. Usually, a larger P3b is observed for unexpected than expected action outcomes (von Borries et al., 2013). In addition to unexpectedness, it is typically found that negative feedback gives rise to a larger P3b than positive feedback (de Bruijn, Mars, & Hulstijn, 2004; Fischer & Ullsperger, 2013; Walentowska, Moors, Paul, & Pourtois, 2016; but see Hajcak et al., 2007; Severo, Walentowska, Moors, & Pourtois, 2017, 2018 for a reversed pattern; as well as Yeung & Safney, 2004 for the lack of a valence effect), suggesting that action value updating likely depends on both expectedness and valence, as well as on the context within which this updating takes place. The second goal of our study was to assess action value updating at the P3b level when different combinations of reward probability and expectancy were created and compared. Based on earlier studies, we hypothesized a larger P3b for unexpected feedback, especially if it was negative. Hence, we surmised a stronger updating of the action value for worse compared to better-than-expected outcomes.

In the current study, participants performed a speeded Go/NoGo task that was previously used and validated extensively in different EEG studies in the past (e.g. Aarts & Pourtois, 2012; Koban, Pourtois, Bediou, & Vuilleumier, 2012; Severo et al., 2017, 2018; Vocat, Pourtois, & Vuilleumier, 2008; Walentowska et al., 2016; Walentowska, Paul, Severo, Moors, & Pourtois, 2018). We chose this specific task setting as it allowed us to introduce subtle variations in reward probability across different blocks without changing the stimuli or task demands between them. Using this procedure, we could create two experimental conditions where reward probability was either low (conservative cutoff, resulting in about 30%) or higher (lenient cutoff, reaching about 50%). In addition to reward probability, we tweaked

expectancy in one or the opposite direction. More specifically, we instructed participants before each block about its putative reward probability level, and created thereby clear expectations about the encounter of positive feedback. Participants were told that positive feedback was either hard or easy to get, leading in turn to a low or high reward expectancy, respectively. Crucially, we then created four different conditions by crossing these two independent variables (probability and expectancy), and alternated block order across participants to control for unwanted fatigue or habituation effects. As a result of this factorial design, participants eventually encountered in some blocks low reward probability with this Go/NoGo task when reward was either expected or unexpected, and likewise, a higher reward probability with the same task in different blocks when reward was either expected or unexpected. Participants' subjective ratings about the positive and negative feedback after each block were used as the main manipulation check of reward expectancy. As mentioned earlier, we pitted two sets of predictions against each other to assess whether the FRN (and P3b) captured either a signed RPE or instead an unsigned RPE during PM in this task. For each of these two components separately, we also estimated the underlying intracranial generators using a standard source localization algorithm to confirm that non-overlapping cortical regions gave rise to them.

2. Methods

2.1. Participants

To estimate the sample size, we used G*Power 3.1.9.2 software (Faul, Erdfelder, Lang, & Buchner, 2007), and referred to the observed effect size of the interaction effect in the time window of the FRN component in our previous study as prior (Walentowska et al., 2016). A sample of 33 participants was estimated to achieve a power of 0.95, with the significance level set at $p = 0.05$ [1]. Thirty-three healthy adult subjects were therefore recruited and participated in exchange of 30 Euro compensation. One participant was removed from further analyses due to noncompliance to the task instructions and excessive movements

[1] The specific parameters chosen to run the power analysis are available in the Supplementary Materials section.

throughout the whole task. The final sample consisted of 32 participants (10 men; mean age: 20.96; SD = 2.69) [2]. All were right-handed (assessed using self-reports), and had normal or corrected-to-normal vision. They were free of neurological or psychiatric history and of psychoactive medication use. They all gave written informed consent prior to the beginning of the experiment. The study was approved by the local ethics committee (Faculty of Psychology and Educational Sciences of Ghent University).

2.2. Experimental paradigm and procedure

A speeded Go/NoGo task, which was previously validated in different EEG studies (Aarts & Pourtois, 2012; Koban et al., 2012; Severo et al., 2017, 2018; Vocat et al., 2008; Walentowska et al., 2016, 2018) was used in the current study. Cues, targets, and non-targets consisted of an arrow presented in the center of the screen against a white background. Each trial started with a fixation cross (1,000 ms). Then, a black arrow ('cue'), oriented up or down, was presented. After a variable interval (1,000−2,000 ms), this black arrow changed color and turned into either green or turquoise, while its orientation could remain either identical or switch. When the black arrow turned green and the orientation remained unchanged ('target'), participants were instructed to press a predefined key on the response box as fast as possible with the index finger of their right hand ('Go trials'). However, participants had to withhold responding when either the arrow became green but flipped orientation, or when it became turquoise and kept its initial orientation ('non-targets' in 'NoGo trials'). In the absence of motor responses, targets and non-targets remained on the screen for 1,000 ms. At the onset of the motor response (correct: 'hits'; incorrect: 'false alarms'), a colored frame (blue or magenta) appeared around the target stimulus, and was presented for 1,000 ms to indicate to participants the registration of a motor response and the imminent presentation of the evaluative feedback. Following that, an evaluative feedback was presented. It consisted of a colorful dot that was either green (for positive feedback) or red (for negative feedback) and that was displayed in the center of the screen for 1,000 ms (see Figure 1 for the trial structure).

[2] In two participants, the EEG data from one condition (out of four) were not recorded properly due to excessive movements and artefacts. Instead of simply removing the whole data sets of these two participants, we replaced the missing values with the condition-specific mean amplitude computed for the group of participants.

Participants were given positive feedback (green dot) when they responded both correctly and fast to Go trials ('fast hit') or when they correctly withheld responding to NoGo trials ('correct inhibition'). They were given negative feedback (red dot) when the response to Go trials was correct but too slow ('slow hit'), when they gave a response to NoGo trials ('false alarm'), or when there was no response to Go trials ('omission'). We used an online adaptive algorithm to set up a limit for correct and fast RTs (i.e., response deadline procedure) in Go trials. At the beginning of the experiment, the RT limit was set to 300 ms (based on previous pilot testing; Vocat et al., 2008). This limit was adjusted online (i.e., after each trial) as a function of the immediately preceding trial history, more specifically, as the mean of the current and previous RTs. Responses that were slower than this limit were classified online as slow hits (and followed by negative feedback), while responses that were faster than the limit were coded online as fast hits (and followed by positive feedback). The advantage of this algorithm is that uncertainty about current RTs is high throughout the task (given the fluctuations of RTs), which motivates participants to actively attend to the evaluative feedback presented after each response (on the Go/target stimulus) to infer whether their actions were timely (fast hits) or not (slow hits). Moreover, the response deadline was updated throughout the experiment in order to avoid habituation or fatigue, and it was set up in such a way that correct and fast responding to Go trials was fairly difficult to achieve (Aarts & Pourtois, 2012; Dhar & Pourtois, 2011; Dhar, Pourtois, & Wiersma, 2011; Koban, Pourtois, Vocat, & Vuilleumier, 2010; Koban et al., 2012; Vocat et al., 2008). When correct inhibitions, omissions, or false alarms occurred, participants could use internal PM to extract the value of their actions, with effects visible at the ERN or CRN (response) level mainly, but not at the FRN/P3 level since the evaluative feedback becomes uninformative and highly redundant in these situations (see Koban et al., 2012, for a clear demonstration). Therefore, in this study, we selectively focused on the FRN and P3 components in response to evaluative feedback following fast and slow hits (Go trials).

Similarly to our previous study (Walentowska et al., 2016), we adapted the online algorithm to determine the response cutoff in order to create two versions of the task that differed in terms of reward probability. In some blocks, a strict response cutoff was used whereby reward (i.e., fast hits) had a low

probability (about 30%). In other blocks, a more lenient cutoff was used with a reward probability that increased up to 50% (see Walentowska et al., 2016 for details). Thus, the low and high reward probability conditions corresponded to 30% and 50% positive feedback, respectively. Orthogonally to the manipulation of reward probability, we manipulated reward expectancy at the start of each block. We created specific expectations about the upcoming blockwise reward probability using written instructions presented at the beginning of each block, as well as a visual cueing technique. After a training session that gave the participants a general overview of the speeded Go/NoGo task, participants were instructed that the experiment would be divided into two consecutive parts that differed with regard to the probability of receiving positive feedback. More specifically, they were told that they could expect positive feedback to be delivered with either low or higher probability, thus creating low and high reward expectancy for these two parts, respectively. Participants were therefore not confronted with the actual number for the reward probability that was expected in the following part, but only with a rough estimate of it. In addition, participants were told that the color of the frame appearing around the target stimulus upon motor response (blue or magenta, depending on the condition) would signal which reward probability would apply. A blue frame signaled a low reward probability whereas a magenta frame signaled a high reward probability. This mapping between color frame and expectancy was counterbalanced across participants.

The experiment consisted of a training session with 32 trials, followed by 12 experimental blocks, each including 56 trials (40 Go and 16 NoGo trials). Go and NoGo trials were presented in a random order within each block. These twelve blocks were divided into two parts according to reward expectancy (low or high). Unbeknown to the participants, the 6 blocks composing one part were further divided into two sub-parts (with 3 consecutive blocks for each of them), depending on reward probability. In this way, a factorial design was devised where effects of expectancy and reward probability (as well as their possible interactions) on the FRN and the P3b could be explored systematically. The order of the two main parts was counterbalanced across participants. Further, in each part, the two reward probabilities were also alternated across participants. Hence, 16 different versions of the experimental procedure were created, and participants were randomly assigned to one of them at the beginning of the experiment, so that each

version of the procedure was used twice (given that we had 32 participants). Stimulus presentation and response recording were controlled using E-prime software (V2.0., http://www.pstnet.com/products/e-prime/).

[insert Figure 1 here]

Participants completed subjective ratings after each triplet of blocks, hence in between each of the four conditions. Participants were asked to evaluate the expectancy of positive and negative feedback in the preceding blocks, as well as how informative this feedback was by means of specific Visual Analog Scales (VASs). More specifically, they were asked to rate: (i) how expected was the positive feedback, (ii) how expected was the negative feedback, and (iii) how informative was the feedback on average in the last three blocks. Each scale ranged from 0 (not expected/informative at all) to 100 (expected/informative a lot). These subjective ratings served as indirect manipulation checks of reward expectancy.

2.3. EEG acquisition and ERP analyses

Participants were seated in a dimly lit, sound-attenuated, and electrically shielded cabin. Continuous EEG was acquired at 512 Hz using a 64-channel (pin-type) Biosemi Active Two system (http://www.biosemi.com), referenced online to the Common ModeSense (CMS)–Driven Right Leg (DRL) ground. All electrodes were placed according to the extended International 10-20 EEG system using an elastic head cap. The vertical and horizontal EOG were monitored by means of four electrodes, placed above and below the right eye and on the outer canthi of both eyes, respectively.

ERPs of interest (FRN and P3b) were computed offline following a standard sequence of data transformation (Keil, Debener, Gratton, Junghöfer, Kappenman, Luck, Luu, Miller, & Yee, 2014) using BrainVision Analyzer 2.0 software: (i) 50-Hz notch filter; (ii) rereferencing of the EEG signal using a common average reference; (iii) -500/+1,000 ms segmentation around the onset of the feedback stimulus; (iv) pre-stimulus interval baseline correction (from -500 ms to feedback onset); (v) vertical ocular correction for blinks (Gratton, Coles, & Donchin, 1983); (vi) semi-automatic artifact rejection (trials with

motor artifacts were rejected, with a fixed criterion of ±80 μV) [3]; (vii) averaging of the feedback-locked ERPs for each type of feedback separately; and (viii) low pass digital filtering of the individual average data (30 Hz) [4].

In accordance with previous ERP studies focused on feedback-based PM (Aarts & Pourtois, 2012; Bismark, Hajcak, Whitworth, & Allen, 2013; Ferdinand et al., 2012; Fischer & Ullsperger, 2013; Pfabigan et al., 2011; Severo et al., 2017, 2018; von Borries et al., 2013; Walentowska et al., 2016; Walsh & Anderson, 2012), as well as the electrophysiological properties of the current data set (see Figure 3), the FRN was defined as the mean voltage within 250–300 ms after feedback onset over frontal and fronto-central electrodes along the midline (Fz and FCz pooled together). The P3b amplitude was measured as a mean voltage between 350 and 600 ms after feedback onset at centro-parietal and parietal electrodes (CPz and Pz pooled together). Moreover, feedback-locked ERP waveforms revealed the existence of another positive component occurring prior to the P3b but immediately after the FRN, with a fronto-central scalp distribution, hence sharing many similarities with a P3a component (see Figure 4 and Results). This fronto-central positivity peaking around 400 ms (P3a) was scored and defined as the mean voltage appearing 350–470 ms after feedback onset at the same locations as used for the FRN (Fz and FCz pooled together).

2.4. Source localization

[3] After artifact rejection, the following number of trials were retained for averaging per condition: low reward probability-low reward expectancy (M = 30.23, SD = 2.52 for positive, and M = 61.11, SD = 2.44 for negative feedbacks), low reward probability-high reward expectancy (M = 29.51, SD = 1.52 for positive, and M = 59.89, SD = 2.13 for negative feedbacks), high reward probability-high reward expectancy (M = 50.32, SD = 1.92 for positive, and M = 48.12, SD = 2.12 for negative feedbacks), and high reward probability-low reward expectancy (M = 47.97, SD = 1.72 for positive, and M = 50.03, SD = 2.11 for negative feedbacks) condition. Both low reward probability conditions were matched for the number of positive, $t_{(31)}$ = -1.18, p = 0.211, and negative, $t_{(31)}$ = -0.38, p = 0.631 feedback trials retained for averaging. Comparably, in both high reward probability conditions, a similar number of trials was used after artifact rejection for positive, $t_{(31)}$ = -1.57, p = 0.317, and negative feedback, $t_{(31)}$ = 0.59, p = 0.616.

[4] Similarly to other studies conducted in the same laboratory (see Walentowska et al., 2016, 2018), we refrained from using a high-pass filter or detrend function because the EEG data were recorded using active electrodes and in a booth that was shielded from external noise and electromagnetic interference, and the raw signals were eventually not distorted.

In order to estimate the neural generators underlying previously identified ERP components, a distributed linear inverse solution was used, the standardized low-resolution brain electromagnetic tomography (sLORETA; Pascual-Marqui, 2002). SLORETA solutions are computed within a three-shell spherical head model, which is coregistered to the MNI152 template (Mazziotta, Toga, Evans, Fox, Lancaster, Zilles, & Mazoyer, 2001). SLORETA estimates the 3-D intracerebral current density distribution within a 5-mm resolution. The 3-D solution space is restricted to the cortical gray matter and hippocampus. The head model uses the electric potential field computed with a boundary element method applied to the MNI152 template (Fuchs, Kastner, Wagner, Hawes, & Ebersole, 2002), and the scalp electrode coordinates on the MNI brain are derived from the international 5% system (Jurcak, Tsuzuki, & Dan, 2007). Separately for each ERP component, positive and negative feedback were compared during the exact same interval (with a mean activity), as used for the standard ERP data analysis (see here above). We used paired-sample *t*-tests performed on the log-transformed data, and as in our previous studies, we set the level of significance for all source localization analyses at $p < 0.01$ (see also Schettino, Loeys, Delplanque, & Pourtois, 2011; Schettino, Loeys, & Pourtois, 2013).

2.5. Statistical analyses

Behavioral and ERP data were submitted to separate repeated-measures ANOVAs including the within-subject factors reward EXPECTANCY (low vs. high) and reward PROBABILITY (low vs. high). For the manipulation checks, we also used feedback VALENCE (positive vs. negative) as an additional factor, and for the behavioral data we used RESPONSE (fast vs. slow hit) as an additional variable.

For the ERP data, we computed and used difference waveforms to reduce the number of factors eventually entered in the statistical analyses (see Luck & Gaspelin, 2017). For each subject and condition separately, the ERP activity for positive feedback was subtracted from that for negative feedback. The ERP components of interest (i.e., FRN, P3a and P3b) were measured on these difference waves (see also Figure 5). The resulting amplitude values were submitted to repeated-measures ANOVAs including the within-subject factors reward EXPECTANCY (low vs. high) and reward PROBABILITY (low vs. high).

In an auxiliary analysis, we also included VALENCE as an additional factor to investigate whether positive or negative feedback underwent a systematic change depending on EXPECTANCY and PROBABILITY. For this analysis, the ERP components of interest (i.e., FRN, P3a and P3b) were scored and measured from the individual waveforms obtained for positive and negative feedback.

Significant (at $p < 0.05$; see section 2.4) main or interaction effects are reported first, followed by post-hoc paired *t*-tests when applicable. Statistical analyses were run using SPSS 24 for Windows and JASP 0.7.5.6 (Love, Selker, Marsman, Jamil, Dropmann, Verhagen, & Wagenmakers, 2015) software.

3. Results

3.1. Manipulation checks

Manipulation checks confirmed that reward expectancy was effective and successful. The ANOVA revealed a significant main effect of EXPECTANCY, $F_{(1,31)} = 5.38$, $p = 0.027$, $\eta_p^2 = 0.148$, as well as an EXPECTANCY x VALENCE interaction, $F_{(1,31)} = 56.12$, $p < 0.001$, $\eta_p^2 = 0.644$. In the two low reward expectancy conditions, participants expected less positive than negative feedback, $t_{(31)} = -4.25$, $p < 0.001$ (see Figure 2A). In comparison, in the two high reward expectancy conditions, participants expected positive feedback more often than negative feedback, $t_{(31)} = 5.57$, $p < 0.001$. These two effects were not influenced by the actual reward probability encountered (neither the main effect of PROBABILITY, nor interaction effects with PROBABILITY were significant, all *ps* > 0.327).

With regard to informativeness, the ANOVA showed a significant main effect of PROBABILITY, $F_{(1,31)} = 4.12$, $p = 0.049$, $\eta_p^2 = 0.134$, as well as a significant interaction between EXPECTANCY and PROBABILITY, $F_{(1,31)} = 5.25$, $p = 0.031$, $\eta_p^2 = 0.147$. The main effect of EXPECTANCY was not significant ($p = 0.145$). Post-hoc *t*-tests showed that feedback's informativeness increased when the outcome was worse than expected ($M = 67.51$, SEM = 2.63) compared to the corresponding high expectancy-high probability condition ($M = 53.13$, SEM = 4.34), $t_{(31)} = 2.57$, $p = 0.011$. The symmetrical effect was not found: Feedback's informativeness did not increase when the

outcome was better than expected (M = 53.22, SEM = 4.26) compared to the corresponding low expectancy-low probability condition (M = 54.78, SEM = 3.95), $t_{(31)}$ = -0.15, p = 0.713.

[insert Figure 2 here]

3.2. Behavioral results

Behavioral results confirmed that the speeded Go/NoGo task yielded the expected proportion of fast relative to slow hits depending on the actual reward probability (i.e., strictness of the response deadline) used (see Figure 2B). The ANOVA showed that the main effect of RESPONSE was significant, $F_{(1,31)}$ = 11.41, p = 0.004, η_p^2 = 0.233, as well as the RESPONSE x PROBABILITY interaction, $F_{(1,31)}$ = 72.61, p < 0.001, η_p^2 = 0.712. When reward probability was low, irrespective of expectancy, slow hits clearly outnumbered fast hits, and participants had approximately 33% of fast hits followed by positive feedback and approximately 66% of slow hits followed by negative feedback, $t_{(31)}$ = -1.72, p = 0.017. In comparison, when reward probability was increased, fast and slow hits were balanced, amounting to about 50% each, irrespective of expectancy again, $t_{(31)}$ = 0.72, p = 0.477 (see Figure 2B).

3.3. ERP results

3.3.1. FRN

The ANOVA showed a significant EXPECTANCY x PROBABILITY interaction, $F_{(1,31)}$ = 9.07, p = 0.009, η_p^2 = 0.213. Neither the main effect of EXPECTANCY (p = 0.209), nor of PROBABILITY (p = 0.794) was significant. To test our a priori hypothesis, we performed post-hoc comparisons. They showed that the FRN (computed as a difference wave between positive and negative feedback; see Methods) was larger when the expectancy and outcome mismatched (M = -2.53 μV, SD = 1.52) compared to when they matched (M = -1.44 μV, SD = 1.71), $t_{(31)}$ = -3.32, p = 0.002. However and critically, the RPE captured by the FRN amplitude did not differ between trials with a worse-than-expected outcome (M = -2.72 μV, SD =

2.31) and those with a better-than-expected outcome (M = -2.16 μV, SD = 2.09), $t_{(31)}$ = 0.66, p = 0.479 [5] (see Figures 3 and 5A&B). Moreover, control analyses (see Supplementary Materials section) confirmed that these results could not easily be explained by an imbalance between conditions in the number of trials included in the averages. When signal-to-noise ratios between conditions were matched by selecting a subset of the trials corresponding to negative feedback, almost identical results were found.

When entering VALENCE as an additional factor in the ANOVA, results showed a significant main effect of VALENCE, $F_{(1,31)}$ = 57.09, p < 0.001, η_p^2 = 0.648, as well as a significant VALENCE x EXPECTANCY x PROBABILITY interaction, $F_{(1,31)}$ = 7.35, p = 0.011, η_p^2 = 0.192. To break down this significant 3-way interaction, we ran 2-way ANOVAs, separately for each valence. For the positive feedback, the EXPECTANCY x PROBABILITY interaction was significant, $F_{(1,31)}$ = 7.01, p = 0.013, η_p^2 = 0.184. This interaction translated a more positive FRN amplitude when outcome and expectancy mismatched with each other (M = 1.56 μV, SD = 2.22), compared to when they were aligned (M = 0.89 μV, SD = 2.43), although this difference failed to reach significance, $t_{(31)}$ = 1.38, p = .175. In comparison, for negative feedback, the FRN amplitude was similar regardless of whether the outcomes mismatched (M = -0.72 μV, SD = 2.11) or matched with expectancy (M = -0.55 μV, SD = 3.29), as revealed by a non-significant EXPECTANCY x PROBABILITY interaction, $F_{(1,31)}$ = 0.41, p = 0.529.

[insert Figure 3 here]

3.3.2. P3a

The ANOVA showed a significant EXPECTANCY x PROBABILITY interaction, $F_{(1,31)}$ = 8.78, p = 0.006, η_p^2 = 0.233. Main effects of PROBABILITY (p = 0.061) and EXPECTANCY (p = 0.282) were non-significant. Post-hoc t-tests showed that the P3a (defined as the difference between positive and negative feedback; see Methods) was larger when reward probability was low and compatible with

[5] To validate the lack of difference in FRN amplitudes between the two mismatching conditions (i.e., better- or worse-than-expected outcomes), we ran a complementing analysis by performing a Bayesian Paired Samples T-Test. The estimated BF_{10} was 0.221, suggesting weak support in favor of a statistical difference between them (Raftery, 1995).

expectancy (M = 2.34 μV, SD = 2.47) compared to a better-than-expected outcome (M = -0.32 μV, SD = 2.29), $t_{(31)}$ = 2.94, p = 0.006. More specifically, the P3a was larger for negative (M = 3.66 μV, SD = 3.07) compared to positive feedback (M = 1.32 μV, SD = 3.85), when reward probability was low and it was expected (see Figures 4A and 5C). By comparison, the amplitude of the P3a was negligible and did not differ between trials in which both reward probability and expectancy were high (M = 0.72 μV, SD = 2.99) and those with a worse-than-expected outcome (M = -0.05 μV, SD = 3.32), $t_{(31)}$ = -1.32, p = 0.195.

When including VALENCE as an additional factor, the ANOVA showed a significant VALENCE x EXPECTANCY x PROBABILITY interaction, $F_{(1,31)}$ = 5.44, p = 0.026, η_p^2 = 0.149. However, none of the two separate 2-way ANOVAs showed a significant EXPECTANCY x PROBABILITY interaction (for positive feedback, $F_{(1,31)}$ = 2.38, p = 0.136; for negative feedback, $F_{(1,31)}$ = 2.53, p = 0.121).

[insert Figure 4 here]

3.3.3. P3b

The ANOVA showed that the EXPECTANCY x PROBABILITY interaction was significant, $F_{(1,31)}$ = 8.43, p = 0.007, η_p^2 = 0.214. Main effects of EXPECTANCY (p = 0.257) and PROBABILITY (p = 0.712) were not significant. In agreement with the hypothesis of action value updating following early error prediction, the P3b (defined as the difference between positive and negative feedback; see Methods) was larger when outcome and expectancy mismatched (M = -1.18 μV, SD = 2.51) than when they matched (M = 0.21 μV, SD = 2.61), $t_{(31)}$ = -2.92, p = 0.006. Importantly, the P3b amplitude did not differ for trials with a worse-than-expected outcome (M = -1.06 μV, SD = 2.58) and a better-than-expected outcome (M = -1.31 μV, SD = 2.77), $t_{(31)}$ = -0.63, p = 0.598 [6] (see Figures 3 and 5A&B). Irrespective of the direction of expectancy violation, P3b amplitudes were systematically larger for positive (M = 2.85 μV, SD = 3.25) than negative feedback (M = 1.67 μV, SD = 3.18), $t_{(31)}$ = 2.12, p = 0.007.

[6] A Bayesian Paired Samples T-Test revealed that the BF_{10} was 0.249, indicating weak support for a statistical difference between these two conditions.

When including VALENCE as an additional factor, the ANOVA showed a significant VALENCE x EXPECTANCY x PROBABILITY interaction, $F_{(1,31)} = 10.79$, $p = 0.003$, $\eta_p^2 = 0.258$. To break down this significant 3-way interaction, we ran 2-way ANOVAs, separately for each valence. For positive feedback, the EXPECTANCY x PROBABILITY interaction was not significant, $F_{(1,31)} = 1.39$, $p = 0.246$. The P3b amplitude was similar irrespective of whether the outcome mismatched (M = 2.22 μV, SD = 3.31) or matched with expectancy (M = 2.03 μV, SD = 3.01). In comparison, the EXPECTANCY x PROBABILITY interaction was significant for negative feedback, $F_{(1,31)} = 6.06$, $p = .021$, $\eta_p^2 = 0.1.64$. The P3b was reduced when the outcome and expectancy mismatched (M = 1.21 μV, SD = 2.14) compared to when they matched (M = 2.71 μV, SD = 3.62), $t_{(31)} = -2.43$, $p = .023$.

[insert Figure 5 here]

3.4. Source localization results

For the FRN, the statistical comparison in the inverse-solution space between positive and negative feedback (run for both conditions with violated expectancy), revealed a widespread suprathreshold activation within the medial prefrontal cortex, which was stronger for negative than positive feedback. More specifically, a main cluster was located within the midcingulate/anterior cingulate cortex (ACC), overlapping with Brodmann areas (BAs) 24, 32, 33 (see Figure 5E, upper panel). Its maximum was located at x = 5, y = 10, z = 35 in BA 24, $t_{(31)} = 6.81$, $p < 0.0001$.

For the P3a, the comparison between positive and negative feedback (using the low reward probability condition when expectancy matched) showed a stronger activation for negative than positive feedback within the right superior frontal gyrus (BAs 8-10). This activation was maximal at x = 25, y = 55, z = 30 in BA 10, $t_{(31)} = 3.36$, $p = 0.02$ (see Figure 5E, middle panel).

For the P3b, this comparison (using both conditions with violated expectancy) revealed a stronger activation for negative than positive feedback in the superior frontal gyrus (BAs 6, 8), and was maximal at x = -3, y = 30, z = 60 in BA 6, $t_{(31)} = 2.55$, $p = 0.007$ (see Figure 5E, lower panel).

4. Discussion

PM is instrumental in fostering goal-adaptive behavior (Ullsperger et al., 2014). It entails the swift detection of mismatches between goal and action (at the FRN level), and subsequently, the updating of the action value (P3b effect) which is necessary to avoid error commission in the future and to learn new contingencies in the environment, and in this way to improve self-regulation and control (Inzlicht et al., 2014). However, whether the FRN reflects a signed RPE (i.e., worse-than-expected event mostly) or an unsigned RPE (i.e, either worse- or better-than-expected events, irrespective of the direction) remains debated in the literature. Likewise, it is still somewhat unclear how action value is updated at the P3b level after a mismatch between goal and action is processed at the FRN level.

To address these two questions, we devised an experiment in which worse-than-expected or better-than-expected outcomes were artificially created and compared to yoked conditions for which actual and expected outcome did not conflict with each other. The behavioral results showed that our manipulation of expectancy was successful: Participants' reports of feedback expectancy after each condition closely followed the instructions provided to them prior to task execution. This expectancy either matched or mismatched with the actual feedback participants received in the Go/NoGo task in different blocks and which was based on the objective reward probability (see Figures 2A and 2B). Moreover, participants reported worse-than-expected outcomes (i.e., negative feedback provided after action execution although a positive feedback was expected for it) to be overall more informative than better-than-expected outcomes (i.e., positive feedback provided after action execution whereas negative outcome was anticipated for it), suggesting an asymmetry in the strength of the mismatch detection between goal and action at the subjective level.

However and importantly, the ERP results clearly showed that the FRN was larger when outcome and expectancy mismatched with each other compared to when they did not, but the direction of this violation did not matter. This early feedback-locked ERP activity was reliably larger for negative compared to positive feedback, and this was true regardless of whether the outcome was either worse or

better than expected. This unsigned RPE captured by the FRN was followed by a differential action value updating at the P3b level, whereby positive feedback was associated with a larger activity than negative feedback, but again, this effect occurred irrespective of the direction of the violation between outcome and expectancy. Finally, we also found an unexpected P3a effect for one condition only: When reward probability was low and reward expectancy was aligned with it, negative feedback led to a larger P3a compared to positive feedback. Here after, we discuss the implications of our new findings for neurobiological models of PM in greater detail.

4.1. Unsigned RPE at the FRN level

An important contribution of the current study to the existing literature is the demonstration that the FRN component is involved in the rapid detection of mismatches between expectancy and outcome, but irrespective of their direction: If the outcome deviates from expectancy, then a large FRN is elicited (see Figures 3, 5A&B). Previous EEG studies already reported larger FRN for negative relative to positive performance feedback (Miltner et al., 1997; Nieuwenhuis et al., 2004; von Borries et al., 2013), and for unexpected compared to expected events (Hajcak, Moser, Holroyd, & Simons, 2007; Pfabigan et al., 2011). Our new results confirm the roles of valence and expectedness but suggest that expectedness must be understood as unsigned expectedness in line with the salience processing account (Hauser et al., 2014; Soder & Potts, 2018; Talmi et al., 2013) and not as signed expectedness in line with the reinforcement learning framework (Holroyd & Coles, 2002).

This salience effect during PM is probably supported by dopaminergic activity that arises within the basal ganglia and the striatum (Schultz, 2016; Schultz, Dayan, & Montague, 1997; Ullsperger et al., 2014) and spreads to the dACC where the FRN is eventually generated (Gehring & Willoughby, 2002; Miltner et al., 1997). This, in turn, enables the rapid trial-by-trial detection of a potential mismatch between outcome and expectancy. This detection is thus deemed ‘low level’, and carries certain features of automaticity (Moors & De Houwer, 2006). The assumption that the FRN may reflect unsigned RPEs during PM, and is thereby mostly driven by salience or surprise, is not new, but backed up by a series of

neuroscientific studies (e.g. Alexander & Brown, 2011; Oliveira et al., 2007), including in animals (Hayden, Heilbronner, Pearson, & Platt, 2011). For example, using a probabilistic learning task, Hauser et al. (2014) concluded that the FRN is associated with surprise signals and absolute (and not signed) RPEs. Likewise, Soder and Potts (2018) and Talmi et al. (2013) both reported ERP results suggesting that the FRN reflects an unsigned RPE rather than a signed RPE, and that it has similar spatio-temporal properties and is even functionally equivalent for both worse- and better-than-expected outcomes. Recently, van der Veen and collaborators (2016) showed that unexpected social judgments yielded larger FRNs when compared to correctly predicted ones. This result likewise suggests that the FRN is sensitive to salience rather than signed RPEs. In sum, our new ERP findings are compatible with this broad literature. However, they also add to it because the experimental design that we devised was able to manipulate the relevant factors in a clean and transparent manner. Unlike these earlier ERP studies, we used a simple Go/NoGo task in which reward probability and reward expectancy were manipulated using a stringent within-subjects design. As a result, the four main conditions had similar stimuli and task demands but nevertheless led to different behavioral performances and expectancies. Our design was also devoid of learning effects and specific incentives (such as monetary gain or loss). Thus, the systematic amplitude variations of the FRN (and P3b) captured across the four conditions in our design could not easily be explained by uncontrolled factors, such as motivation, learning or task involvement for instance.

However, a striking observation is that subjective ratings regarding feedback's informativeness did not align with these FRN results. As it turned out, participants judged unexpected negative feedback as the most informative, whereas feedback's informativeness was lower and balanced for the other experimental conditions, including unexpected positive feedback. This dissociation is intriguing as it suggests indirectly that these ratings were likely based on different evaluation or monitoring processes compared to those involved in the FRN. On the other hand, this dissociation is perhaps not so surprising considering that this ERP component is thought to reflect a rather automatic evaluation of the feedback performed by the dACC, right after its onset (Holroyd & Coles, 2002), and presumably, this effect may not be readily accessible to introspection. Our supplementing source localization results also clearly

corroborated the involvement of the dACC in the generation of the FRN recorded in our study, but not in the generation of the P3a or P3b. The involvement of the dACC in the generation of the FRN as observed here, accords well with previous EEG studies (Gehring & Willoughby, 2002; Miltner et al., 1997; Yeung et al., 2004), as well as with studies using combined EEG-fMRI techniques (Hauser et al., 2014), or neuroimaging only (Oliveira et al., 2007).

Moreover, a set of auxiliary analyses showed that this modulation of the FRN with expectancy was mostly driven by the positive rather than the negative feedback. Because positive feedback and more specifically, reward is associated with a distinctive ERP component, known as the Reward Positivity (Proudfit, 2015), our results suggest indirectly that this reward-related activity could eventually be the one that was influenced by expectancy in our study. However, additional studies are needed to assess whether it is reward expectancy or rather reward sensitivity that is altered when the outcome (i.e., positive feedback) and expectancy mismatch, especially if this outcome denotes either a better- or a worse-than-expected event. In this context, ERP studies that seek to better characterize the fine-grained spatio-temporal dynamics of reward processing during feedback-based PM could help disentangle effects of valence from expectancy (see Gheza, Paul, & Pourtois, 2018 for a recent attempt).

A valid objection could be that expectancy or probability already shaped feedback processing before the onset of the FRN. To rule out this alternative account, we ran a set of auxiliary analyses during the pre-feedback interval (see Supplementary Materials section) where we focused on the Stimulus-Preceding Negativity (SPN) component, which is sensitive to reward anticipation (Brunia, 1988; Brunia, Hackley, van Boxtel, Kotani, & Ohgami, 2011; Brunia & van Boxtel, 2001) and feedback informativeness (Walentowska et al., 2018). These supplementary results clearly showed that these FRN effects were not simply mirrored by earlier SPN effects occurring prior to feedback onset. In line with our previous study (Walentowska et al., 2018), we found that the SPN was larger (i.e., more negative) in anticipation of positive than negative feedback, but exclusively when positive feedback had a high probability. These control analyses therefore support the interpretation that salience influenced feedback processing after its onset (at the FRN level) but not prior to it (at the SPN level).

4.2. Action value updating at the P3b level

Our new ERP results also inform about the subsequent updating process, following salience detection at the FRN level. A larger P3b for positive than negative feedback (following mismatches between outcome and expectancy) suggests that the former was better or more strongly processed than the latter (see Figures 3, 5A&B). Moreover, when entering VALENCE as additional factor in the statistical model, we found that the P3b was selectively reduced for negative feedback when outcome and expectancy mismatched, suggesting impaired updating for this specific combination and outcome. In line with previous EEG studies that linked the P3b to closure, WM updating, or attention more generally (Donchin & Coles, 1998; Polich, 2007; Verleger, 1997; Verleger et al., 1994), we can conclude that positive feedback provided in a context of violations between expectancy and outcomes received eventually more weight or attention (e.g. facilitated closure and updating) than negative feedback. An important additional contribution of our study is to show that this gating effect at the P3b level happens to occur regardless of whether the positive feedback was a worse- or better-than-expected event. Hence, similar to the FRN, our results for the P3b suggest a decoupling between the specific information value carried by the evaluative feedback at the subjective level (being presumably larger for worse-than-expected compared to better-than-expected outcome; see Holroyd & Coles, 2002) and action value updating at the neural level. To be noted, the involvement of the P3b in this specific process during PM was already demonstrated in previous EEG studies that linked the amplitude of the P3b to information/value updating at the computational level (Fischer & Ullsperger, 2013; Ullsperger, 2017; Ullsperger et al., 2014; von Borries et al., 2013).

Alternatively, this enhanced P3b for positive compared to negative feedback in two contexts in which expectancy and outcome mismatched could reflect enhanced goal relevance processing of this specific feedback, especially if positive. More specifically, in a recent EEG study, we found that the P3b component was systematically larger for evaluative feedback deemed goal relevant for participants (Walentowska et al., 2016). Relevance was understood in this study as the degree to which a stimulus was informative about the satisfaction status of pursued goals (see also Moors, 2007). Moreover, in another

recent EEG study, we found evidence that goal relevance understood as impact increased the P3b during PM (Severo et al., 2017, 2018). Impact corresponded to the amount of goal satisfaction that was signaled by the feedback stimulus. Hence, it is possible that the P3b was larger for positive feedback in these situations where violations occurred because it was more relevant for participants, in the sense of informing them swiftly about something important happening, or perhaps having a greater impact for their goals. In this framework, action value updating and relevance processing at the P3b level are not necessarily mutually exclusive. Presumably, action value updating could be facilitated by the enhanced relevance assigned to the positive feedback when it was delivered in a context where outcome and expectancy mismatched with each other.

Unexpectedly, apart from the P3b updating effect, another and earlier valence effect was observed after the FRN at the same fronto-central locations, and thereby sharing many similarities with a P3a ERP component. Specifically, when reward probability was low and there was no mismatch between outcome and expectancy, negative feedback gave rise to a much larger P3a than positive feedback in this context. Obviously, this effect cannot translate a simple oddball effect (Polich, 2007) given that negative feedback was not deviant in this condition, on the contrary. Unlike the FRN and P3b components, this P3a effect was not 'unsigned' or explained by salience, but actually found for a single combination of probability and expectancy only, namely when low reward probability and low reward expectancy converged (see Figures 4A and 5C). Some hints on the possible role of the P3a during PM were already provided by Ullsperger and colleagues (2014), suggesting that the P3a could be associated with an attention orienting to potentially goal-relevant stimuli, before its exact motivational meaning (i.e., whether the stimulus matches or mismatches with goals) is extracted later at the P3b level (see here above). When estimating the intracranial generators of this P3a effect, we found the superior frontal gyrus to be the main source – a result which is in line with older neurophysiological findings that have linked frontal lobe activity to the P3 (Baudena, Halgren, Heit, & Clarke, 1995; Halgren, Marinkovic, & Chauvel, 1998).

Presumably, negative feedback provided in a context where it dominated and was expected by the participants (with as a result, a small FRN) unlocked additional attentional or motivational processes that

may explain this selective P3a effect. In this condition, participants' self-efficacy (Bandura, 1982, 1993; Bandura & Cervone, 1983) was challenged extensively as they were expecting negative feedback after their Go/NoGo decisions, and the outcome confirmed this (i.e., low self-efficacy). This situation, although not associated with a violation between expectancy and outcome, was probably associated with a high level of negative affect and/or frustration, facilitating in turn the processing of negative feedback at the P3a level in this condition. Although this assumption awaits validation at the empirical level, our new results for the P3a and P3b suggest that PM brain mechanisms are highly flexible, exploiting in a dynamic and context-sensitive manner the evidence available after feedback onset to monitor and update action value. As our new results indirectly suggest, this process could very well occur earlier in time for negative information when it met expectancy (at the P3a level), compared to the processing of positive information, which was globally enhanced at later stage when expectancy was violated (P3b level).

4.3. Limitations

A few limitations warrant comment. First, we considered reward probabilities of 33% as low and of 50% as high. It could be argued that these percentages reflect low or intermediate levels instead. Moreover, it could be objected that uncertainty (which is larger for 50% than a 33% reward probability), rather than reward probability (see Mushtaq, Bland, & Schaefer, 2011; Yu & Dayan, 2005), was the main difference between our conditions. However, a previous ERP study already reported that the FRN and P300 scaled up with unexpectedness rather than uncertainty (Kogler, Sailer, Derntl, & Pfabigan, 2017), suggesting that the former variable likely accounted for the systematic amplitude modulation of the FRN (and P3) seen in the current study. Moreover, we chose these specific reward probabilities (and specific task setting for the RT deadline) based on many previous EEG studies that already validated them using the same Go/NoGo task (see Aarts & Pourtois, 2012; Koban et al., 2012; Severo et al., 2017, 2018; Vocat et al., 2008; Walentowska et al., 2016), and reported clear-cut FRN and P3b effects. In addition, the behavioral results and subjective ratings (see Figure 2) both confirmed that we created four different conditions that could indeed be discriminated from one another based on the match or mismatch between

reward probability and reward expectancy. Notwithstanding these elements, it appears important to explore in future studies the neural processing of matches vs. mismatches between reward probability and expectancy, with tasks yielding higher reward probabilities than in the present experiment if possible.

Second, in the current study, we did not include a baseline condition without instructions about the reward probabilities to be expected because we did not want to increase the already long duration of our experiment. However, in order to replicate and extend the amplitude change of the FRN (and P3b) with expectancy, it would be important in future ERP studies to add such a control condition. This would probably allow to more easily disentangle the contribution of subjective (i.e., expectancy) from objective (i.e., probability) effects on the FRN and P3b levels.

Third, our main manipulation check for reward expectancy was probably not 'pure' and contaminated by social desirability or memory effects. It was important to confirm, via specific questions asked to the participants, that they actually followed the instructions given to them beforehand, and hence eventually 'remembered' afterwards the reward probability (either low or high) that was expected in a specific condition. However, this procedure may have hindered the possibility to find a significant effect of reward probability, besides expectancy, on these subjective ratings. To overcome this limitation, we suggest using different and more implicit manipulation checks in future studies if possible, which would probably allow to more objectively capture effects of both expectancy and reward probability during the monitoring of positive vs. negative performance feedback.

4.4. Conclusions

The current study showed that when a simple Go/NoGo task was used, the FRN component, generated in the dACC, captured unsigned RPEs during PM and could be interpreted as driven by salience or surprise. Intriguingly, participants rated the worse-than-expected outcome as the most informative one, suggesting a dissociation between self-report and ERP results. Moreover, the P3b component was larger for positive than negative feedback, if and only if mismatches between expectancy and outcome occurred, irrespective of their direction. These results suggest that action value updating at the P3b level was stronger for

positive than negative outcomes, and probably generic. However, when negative feedback prevailed and this dominance was anticipated by the participants, we found that it was processed more strongly than positive feedback at the P3a level, suggesting that negative information can guide action value updating or monitoring at an earlier stage than positive information, for which a P3b effect was found. All in all, these new ERP results are not only compatible with current neurobiological models of PM (e.g. Ullsperger et al., 2014), but they can also serve to inform them about the specific combination of task-specific (i.e., reward probability) and subject-dependent (i.e., reward expectancy) factors that can influence the speed and efficiency of PM.

References

Aarts, K., & Pourtois, G. (2012). Anxiety disrupts the evaluative component of performance monitoring: An ERP study. *Neuropsychologia, 50,* 1286–1296. https://doi.org/10.1016/j.neuropsychologia. 2012.02.012

Alexander, W. H., & Brown, J. W. (2011). Medial prefrontal cortex as an action-outcome predictor. *Nature Neuroscience, 14,* 1338–1344. https://doi.org/10.1038/nn.2921

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37,* 122–147. http://dx.doi.org/10.1037/0003-066X.37.2.122

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational Psychologist, 28,* 117–148. https://doi.org/10.1207/s15326985ep2802_3

Bandura, A., & Cervone, D. (1983). Self-evaluative and self-efficacy mechanisms governing the motivational effects of goal systems. *Journal of Personality and Social Psychology, 45,* 1017–1028. http://dx.doi.org/10.1037/0022-3514.45.5.1017

Baudena, P., Halgren, E., Heit, G., & Clarke, J. M. (1995). Intracerebral potentials to rare target and distractor auditory and visual stimuli. III. Frontal cortex. *Electroencephalography and clinical neurophysiology, 94,* 251–264. https://doi.org/10.1016/0013-4694(95)98476-O

Bismark, A. W., Hajcak, G., Whitworth, N. M., & Allen, J. J. (2013). The role of outcome expectations in the generation of the feedback-related negativity. *Psychophysiology, 50,* 125–133. https://doi.org/10.1111/j.1469-8986.2012.01490.x

Botvinick, M., & Braver, T. (2015). Motivation and cognitive control: from behavior to neural mechanism. *Annual Review of Psychology, 66,* 83–113. https://doi.org/10.1146/annurev-psych-010814-015044

Brunia, C. H. M. (1988). Movement and stimulus preceding negativity. *Biological Psychology, 26,* 165–178. https://doi.org/10.1016/0301-0511(88)90018-X

Brunia, C. H. M., Hackley, S. A., van Boxtel, G. J. M., Kotani, Y., & Ohgami, Y. (2011). Waiting to perceive: reward or punishment? *Clinical Neurophysiology, 122,* 858–868. https://doi.org/10.1016/j.clinph.2010.12.039

Brunia, C. H. M., & van Boxtel, G. J. M. (2001). Wait and see. *International Journal of Psychophysiology, 43,* 59–75. https://doi.org/10.1016/S0167-8760(01)00179-9

de Bruijn, E. R. A., Mars, R. B., & Hulstijn, W. (2004). 'It wasn't me... or was it?' How false feedback effects performance. In M. Ullsperger & M. Falkenstein (Eds.), *Errors, conflicts, and the brain. Current opinions on performance monitoring* (pp. 118–124). Leipzig, Germany: Max Planck Institute of Cognitive Neuroscience.

Dhar, M., & Pourtois, G. (2011). Early error detection is generic, but subsequent adaption to errors is not: evidence from ERPs. *Neuropsychologia, 49,* 1236–1245. https://doi.org/10.1016/j.neuropsychologia.2011.01.006

Dhar, M., Wiersema, J. R., & Pourtois, G. (2011). Cascade of neural events leading from error commission to subsequent awareness revealed using EEG source imaging. *PLoS One, 6,* e19578. https://doi.org/10.1371/journal.pone.0019578

Donchin, E., & Coles, M. G. (1998). Context updating and the P300. *Behavioral and Brain Sciences, 21*, 152–154. https://doi.org/10.1017/S0140525X98230950

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. https://doi.org/10.1016/0013-4694(91)90062-9

Faul, F., Erdfelder, E., Lang, A. G. & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39,* 175–191. https://doi.org/org/10.3758/BF03193146

Ferdinand, N. K., Mecklinger, A., Kray, J., & Gehring, W. J. (2012). The processing of unexpected positive response outcomes in the mediofrontal cortex. *The Journal of Neuroscience, 32,* 12087–12092. https://doi.org/10.1523/JNEUROSCI.1410-12.2012

Fischer, A. G., & Ullsperger, M. (2013). Real and fictive outcomes are processed differently but converge on a common adaptive mechanism. *Neuron*, *79*, 1243–1255. https://doi.org/10.1016/j.neuron.2013.07.006

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. https://doi.org/10.1126/science.1102941

Fuchs, M., Kastner, J., Wagner, M., Hawes, S., & Ebersole, J. S. (2002). A standardized boundary element method volume conductor model. *Clinical Neurophysiology*, *113*, 702–712. https://doi.org/10.1016/S1388-2457(02)00030-5

Gehring, W. J., Goss, B., Coles, M. G., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science*, *4*, 385–390. https://doi.org/10.1111/j.1467-9280.1993.tb00586.x

Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, *295*, 2279–2282. https://doi.org/10.1126/science.1066893

Gheza, D., Paul., K., & Pourtois, G. (2018). Dissociable effects of reward and expectancy during evaluative feedback processing revealed by topographic ERP mapping analysis. *International Journal of Psychophysiology*, *132*, 213–225. https://doi.org/10.1016/j.ijpsycho.2017.11.013

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*, 468–484. https://doi.org/10.1016/0013-4694(83)90135-9

Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2006). The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. *Biological Psychology*, *71*, 148–154. https://doi.org/10.1016/j.biopsycho.2005.04.001

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. https://doi.org/10.1111/j.1469-8986.2007.00567.x

Hajihosseini, A., & Holroyd, C. B. (2013). Frontal midline theta and N 200 amplitude reflect complementary information about expectancy and outcome evaluation. *Psychophysiology, 50*, 550–562. https://doi.org/10.1111/psyp.12040

Halgren, E., Marinkovic, K., & Chauvel, P. (1998). Generators of the late cognitive potentials in auditory and visual oddball tasks. *Electroencephalography and clinical neurophysiology*, *106*, 156–164. https://doi.org/10.1016/S0013-4694(97)00119-3

Hauser, T. U., Iannaccone, R., Stämpfli, P., Drechsler, R., Brandeis, D., Walitza, S., & Brem, S. (2014). The feedback-related negativity (FRN) revisited: New insights into the localization, meaning and network organization. *NeuroImage, 84*, 159–168. https://doi.org/org/10.1016/j.neuroimage.2013.08.028

Hayden, B. Y., Heilbronner, S. R., Pearson, J. M., & Platt, M. L. (2011). Surprise signals in anterior cingulate cortex: neuronal encoding of unsigned reward prediction errors driving adjustment in behavior. *Journal of Neuroscience, 31*, 4178–4187. https://doi.org/10.1523/JNEUROSCI.4652-10.2011

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. https://doi.org/http://dx.doi.org/.org/10.1037/0033-295X.109.4.679

Inzlicht, M., Schmeichel, B. J., & Macrae, C. N. (2014). Why self-control seems (but may not be) limited. *Trends in Cognitive Sciences*, *18*, 127–133. https://doi.org/10.1016/j.tics.2013.12.009

Jurcak, V., Tsuzuki, D., & Dan, I. (2007). 10/20, 10/10, and 10/5 systems revisited: Their validity as relative head-surface-based positioning systems. *NeuroImage*, *34*, 1600–1611. https://doi.org/10.1016/j.neuroimage.2006.09.024

Keil, A., Debener, S., Gratton, G., Junghöfer, M., Kappenman, E. S., Luck, S. J., Luu, P., Miller, G. A., & Yee, C. M. (2014). Committee report: Publication guidelines and recommendations for studies using electroencephalography and magnetoencephalography. *Psychophysiology*, *51*, 1–21. https://doi.org/10.1111/psyp.12147

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. *Social Neuroscience,5,* 360–374. https://doi.org/10.1080/17470911003651547

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. https://doi.org/10.3758/s13415-012-0091-0

Kogler, L., Sailer, U., Derntl, B., & Pfabigan, D. M. (2017). Processing expected and unexpected uncertainty is modulated by fearless-dominance personality traits – An exploratory ERP study on feedback processing. *Physiology & Behavior, 168,* 74–83. https://dx.doi.org/10.1016/j.physbeh.2016.10.016

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, A. J., & Wagenmakers, E. J. (2015). JASP (Version 0.7). Amsterdam free shareware. https://jasp-stats.org/

Luck, S. J., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology, 54,* 146–157. https://doi.org/10.1111/psyp.12639

Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., & Mazoyer, B. (2001). A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philosophical Transactions of the Royal Society B, 356*, 1293–1322. https://doi.org/10.1098/rstb.2001.0915

Miltner, W. H., Braun, C. H., & Coles, M. G. (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. https://doi.org/10.1162/jocn.1997.9.6.788

Moors, A. (2007). Can cognitive methods be used to study the unique aspect of emotion: An appraisal theorist's answer. *Cognition and Emotion, 21,* 1238–1269. https://doi.org/10.1080/02699930701438061

Moors, A., & De Houwer, J. (2006). Automaticity: a theoretical and conceptual analysis. *Psychological Bulletin, 132,* 297–326. https://dx.doi.org/10.1037/0033-2909.132.2.297

Mushtaq, F., Bland, A. R., & Schaefer, A. (2011). Uncertainty and cognitive control. *Frontiers in Psychology, 249*. https://doi.org/10.3389/fpsyg.2011.00249

Nieuwenhuis, S., Holroyd, C. B., Mol, N., & Coles, M. G. (2004). Reinforcement-related brain potentials from medial frontal cortex: origins and functional significance. *Neuroscience & Biobehavioral Reviews, 28,* 441–448. https://doi.org/10.1016/j.neubiorev.2004.05.003

Oliveira, F. T. P., McDonald, J. J., & Goodman, D. (2007). Performance Monitoring in the Anterior Cingulate is Not All Error Related: Expectancy Deviation and the Representation of Action-Outcome Associations. *Journal of Cognitive Neuroscience, 19,* 1994–2004. https://doi.org/org/10.1162/jocn.2007.19.12.1994

Pascual-Marqui, R. (2002). Standardized low-resolution brain electromagnetic tomography (sLORETA): Technical details. *Methods and Findings in Experimental and Clinical Pharmacology, 24,* 5–12.

Pfabigan, D. M., Alexopoulos, J., Bauer, H., Lamm, C., & Sailer, U. (2011). All about the money - external performance monitoring is affected by monetary, but not by socially conveyed feedback cues in more antisocial individuals. *Frontiers in Human Neuroscience, 5*. https://doi.org/10.3389/fnhum.2011.00100

Polich, J. (2007). Updating P300: an integrative theory of P3a and P3b. *Clinical Neurophysiology, 118,* 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Proudfit, G. H. (2015). The reward positivity: From basic research on reward to a biomarker for depression. *Psychophysiology, 52,* 449–459. https://doi.org/10.1111/psyp.12370

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology, 25,* 111–164.

Sambrook, T. D., & Goslin, J. (2015). A neural reward prediction error revealed by a meta-analysis of ERPs using great grand averages. *Psychological Bulletin, 141,* 213–235. https://doi.org/10.1037/bul0000006

San Martin, R. (2012). Event-related potential studies of outcome processing and feedback-guided learning. *Frontiers in Human Neuroscience, 6,* 1–17. https://doi.org/10.3389/fnhum.2012.00304

Schettino, A., Loeys, T., Delplanque, S., & Pourtois, G. (2011). Brain dynamics of upstream perceptual processes leading to visual object recognition: A high density ERP topographic mapping study. *NeuroImage, 55,* 1227–1241. https://doi.org/10.1016/j.neuroimage.2011.01.009

Schettino, A., Loeys, T., & Pourtois, G. (2013). Multiple synergistic effects of emotion and memory on proactive processes leading to scene recognition. *NeuroImage, 81,* 81–95. https://doi.org/10.1016/j.neuroimage.2013.04.115

Schultz, W. (2016). Dopamine reward prediction-error signalling: A two-component response. *Nature Reviews Neuroscience, 17,* 183–195. https://doi.org/https://doi.org/10.1038/nrn.2015.26

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275,* 1593–1599. https://doi.org/10.1126/science.275.5306.1593

Severo, M. C., Walentowska, W., Moors, A. & Pourtois, G. (2017). Goal impact influences the evaluative component of performance monitoring: Evidence from ERPs. *Biological Psychoogy, 129,* 90–102. https://doi.org/org/10.1016/j.biopsycho.2017.08.052

Severo, M. C., Walentowska, W., Moors, A. & Pourtois, G. (2018). Goals matter: Amplification of the motivational significance of the feedback when goal impact is increased. *Brain and Cognition, 128,* 56–72. https://doi.org/10.1016/j.bandc.2018.11.002

Soder, H. E, & Potts, G. F. (2018). Medial frontal cortex response to unexpected motivationally salient outcomes. *International Journal of Psychophysiology, 128,* 268–276, https://doi.org/org/10.1016/j.ijpsycho.2017.11.003

Talmi, D., Atkinson, R., & El-Deredy, W. (2013). The Feedback-Related Negativity Signals Salience Prediction Errors, Not Reward Prediction Errors. *Journal of Neuroscience, 33,* 8264–8269. https://doi.org/org/10.1523/JNEUROSCI.5695-12.2013

Ullsperger, M. (2017). Neural bases of performance monitoring. In T. Egner (Ed.), *The Wiley handbook of cognitive control* (pp. 292–313). Chichester, UK: John Wiley & Sons, Ltd.

Ullsperger, M., Fischer, A. G., Nigbur, R., & Endrass, T. (2014). Neural mechanisms and temporal dynamics of performance monitoring. *Trends in Cognitive Sciences, 18*, 259–267. https://doi.org/10.1016/j.tics.2014.02.009

van der Veen, F. M., van der Molen, M. J. W., van der Molen, M. W., & Franken, I. H. A. (2016). Thumbs up or thumbs down? Effects of neuroticism and depressive symptoms on psychophysiological responses to social evaluation in healthy students. *Cognitive, Affective, and Behavioral Neuroscience, 16,* 836–847. https://doi.org/org/10.3758/s13415-016-0435-2

Verleger, R. (1997). On the utility of P3 latency as an index of mental chronometry. *Psychophysiology, 34,* 131–156. https://doi.org/10.1111/j.1469-8986.1997.tb02125.x

Verleger, R., Jaskowski, P., & Wauschkuhn, B. (1994). Suspense and surprise: On the relationship between expectancies and P3. *Psychophysiology, 31,* 359–369. https://doi.org/10.1111/j.1469-8986.1994.tb02444.x

Vocat, R., Pourtois, G., & Vuilleumier, P. (2008). Unavoidable errors: a spatio-temporal analysis of time-course and neural sources of evoked potentials associated with error processing in a speeded task. *Neuropsychologia, 46,* 2545–2555, https://doi.org/10.1016/j.neuropsychologia.2008.04.006

von Borries, A. K. L., Verkes, R. J., Bulten, B. H., Cools, R., & de Bruijn, E. R. A. (2013). Feedback-related negativity codes outcome valence, but not outcome expectancy, during reversal learning. *Cognitive, Affective, & Behavioral Neuroscience, 13,* 737–746. https://doi.org/10.3758/s13415-013-0150-1

Walentowska, W., Moors, A., Paul, K., & Pourtois, G. (2016). Goal relevance influences performance monitoring at the level of the FRN and P3 components. *Psychophysiology, 53,* 1020–1033. https://doi.org/10.1111/psyp.12651

Walentowska, W., Paul, K., Severo, M. C., Moors, A., & Pourtois, G. (2018). Relevance and uncertainty jointly influence reward anticipation at the level of the SPN ERP component. *International Journal of Psychophysiology, 132,* 287–297. https://doi.org/org/10.1016/j.ijpsycho.2017.11.005

Walsh, M. M., & Anderson, J. R. (2012). Learning from experience: event-related potential correlates of reward processing, neural adaptation, and behavioral choice. *Neuroscience & Biobehavioral Reviews, 36,* 1870–1884. https://doi.org/10.1016/j.neubiorev.2012.05.008

Yeung, N., Holroyd, C. B., & Cohen, J. D. (2004). ERP correlates of feedback and reward processing in the presence and absence of response choice. *Cerebral Cortex, 15,* 535–544. https://doi.org/10.1093/cercor/bhh153

Yeung, N., & Sanfey, A. G. (2004). Independent coding of reward magnitude and valence in the human brain. *Journal of Neuroscience, 24,* 6258–6264. https://doi.org/10.1523/JNEUROSCI.4537-03.2004

Yu, A. J., & Dayan, P. (2005). Uncertainty, neuromodulation, and attention. *Neuron, 46,* 681–692. https://dx.doi.org/10.1016/j.neuron.2005.04.026

Authors Note

This study and position of WW at Jagiellonian University are supported by funding from the National Science Centre of Poland (NCN grant 2015/19/B/HS6/01259). Research staying of WW at Ghent University is supported by the Polish National Agency for Academic Exchange (Bekker Programme signature: PPN/BEK/2018/1/00257). GP is supported by the Belgian Science Policy, Interuniversity Attraction Poles program (P7/11), and by a Concerted Research Action Grant from Ghent University. He is the recipient of an (2015) independent investigator grant awarded by the NARSAD foundation. GP and AM are also supported by a research grant from the Research Foundation Flanders (FWO grant G024716N).

Conflict of interest: none declared.

Correspondence should be sent to: Wioleta Walentowska, Psychophysiology Laboratory, Jagiellonian University, Ingardena 6, 30-060 Krakow, Poland.

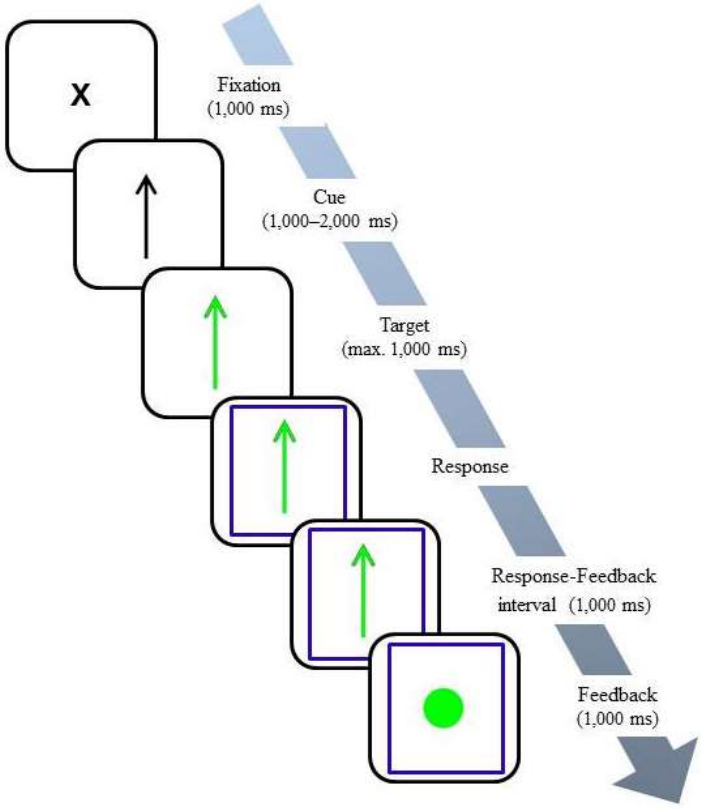


Figure 1. Trial structure. At response onset (speeded Go/NoGo task), a colored frame (blue or magenta) appeared around the target and stayed on screen until feedback offset, signaling either low or high reward expectancy. Irrespective of reward expectancy, reward probability was also either low or high, resulting in four conditions in our study when crossing these two factors. In two conditions, reward expectancy was violated with action outcomes, while in the two other ones – action outcomes matched reward expectancy. Note that the current trial illustrates a correct and fast response to a Go stimulus (i.e., below the arbitrary response deadline), followed by a positive feedback (green dot). If the response was correct but too slow (i.e., above the arbitrary response deadline), then a negative feedback (red dot) was shown instead.

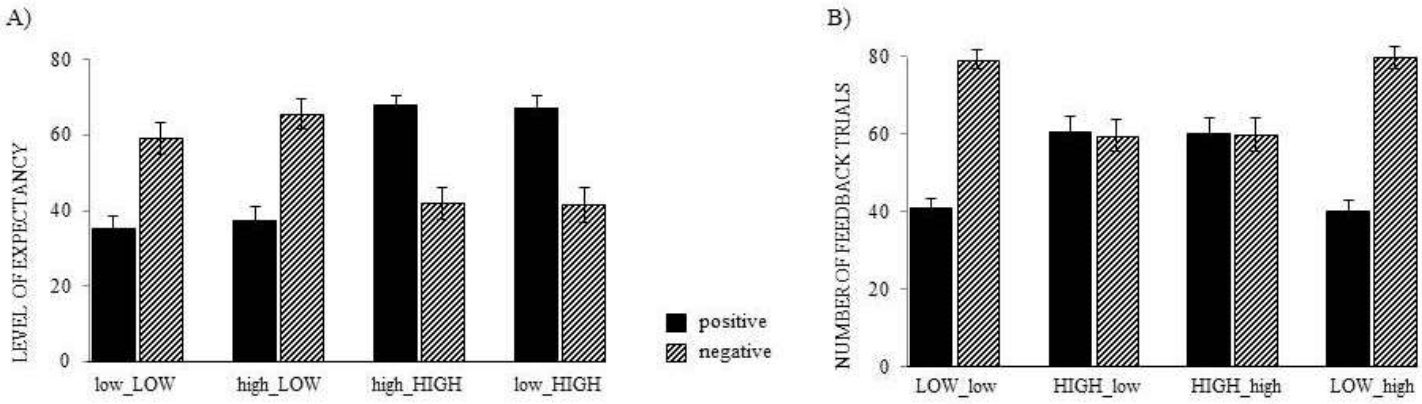


Figure 2. Behavioral results. (A) Manipulation checks of reward expectancy (based on VAS scales). After each condition, participants reported the proportion of positive vs. negative feedback received during the last three blocks. Results showed that these estimations closely followed the instructions and hence reward expectancy, but were not influenced by reward probability. (B) In comparison, when looking at the actual task data, results showed that the number of positive and negative feedbacks (following fast and slow hits, respectively), closely followed the specific reward probability (being either low or high) used in each of the two conditions. Reward expectancy did not influence this outcome. Note that for (A) scores can vary from 0 (not expected at all) to 100 (expected a lot), and for (A) and (B) error bars represent standard errors of the mean (SEM). The following rule was used in condition naming in (A) probability_EXPECTANCY; and in (B) PROBABILITY_expectancy.

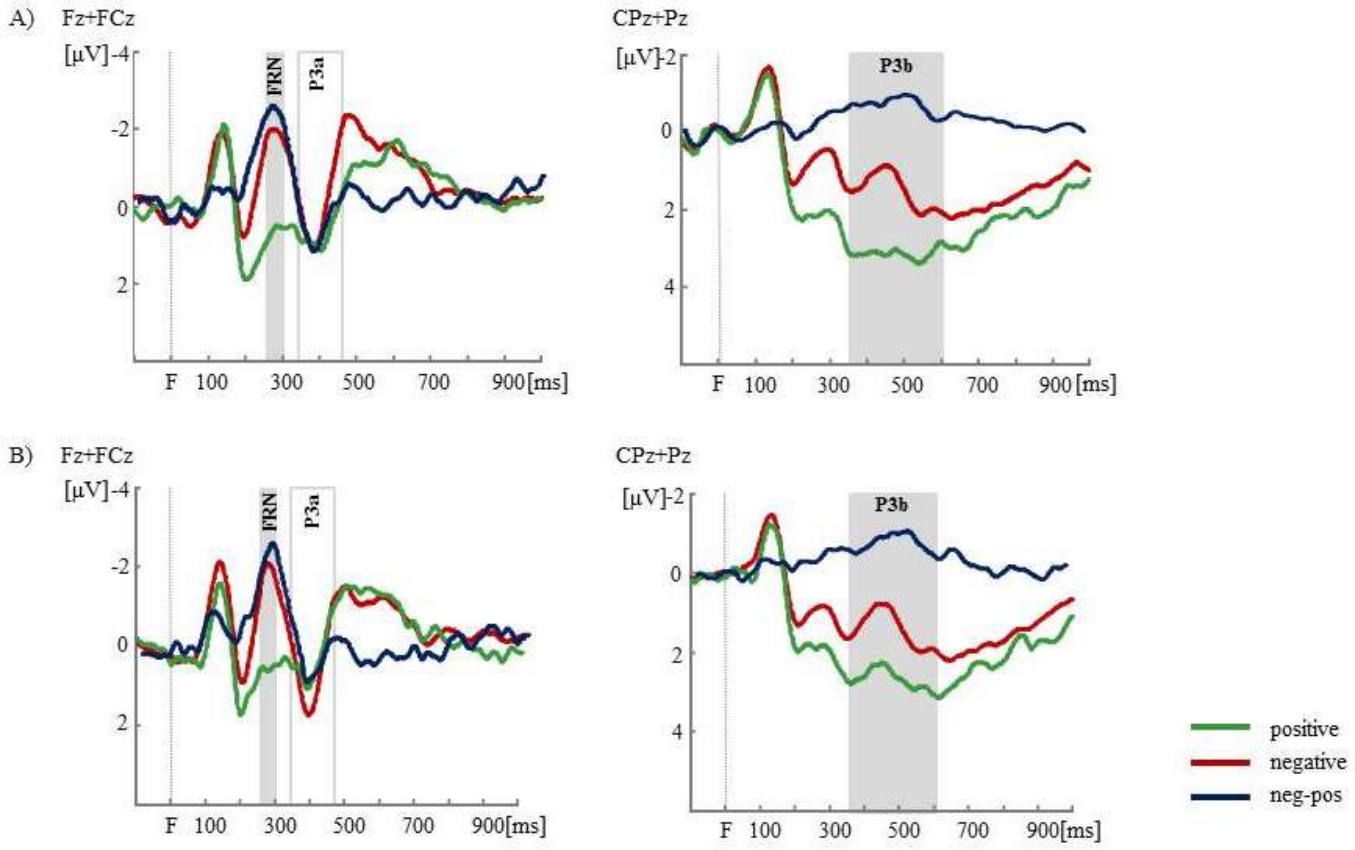


Figure 3. Feedback-locked grand average ERP waveforms, recorded from fronto-central electrodes (Fz and FCz pooled together; left panel) for the FRN (250–300ms) and P3a (350–470 ms), and non-overlapping centro-parietal locations (CPz and Pz pooled together; right panel) for the P3b (350–600 ms post-feedback onset). (A) ERPs when low reward expectancy was violated (i.e., outcome was better than expected). (B) ERPs when high expectancy was violated (i.e., outcome was worse than expected). For both conditions, a large and similar FRN (being more negative for negative than positive feedback) and a large P3b (being more positive for positive than negative feedback) were recorded. Note that F stands for feedback onset, significant effects are highlighted in grey, and negativity is plotted upwards.

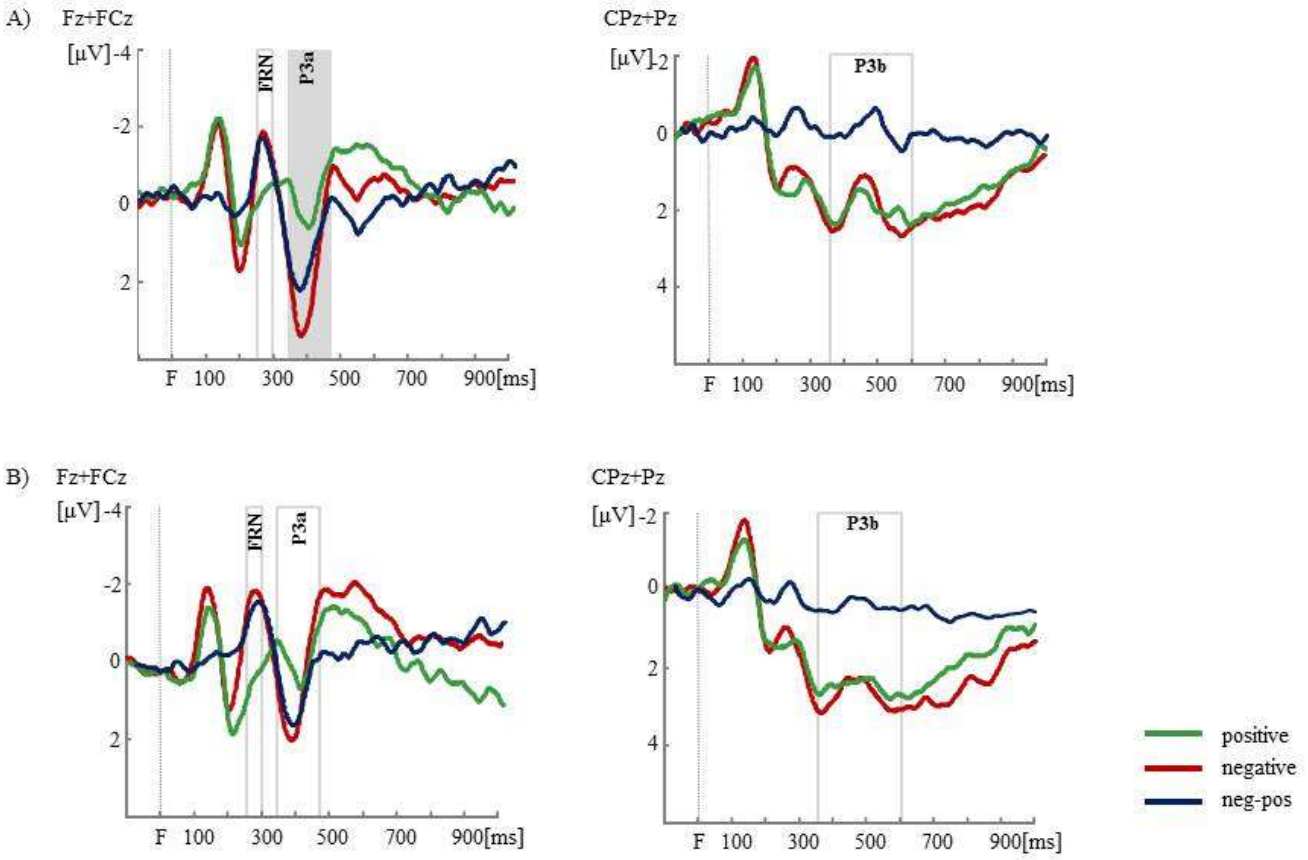


Figure 4. Feedback-locked grand average ERP waveforms, recorded from fronto-central electrodes (Fz and FCz pooled together; left panel) for the FRN (250–300 ms) and P3a (350–470 ms), and centro-parietal locations (CPz and Pz pooled together; right panel) for the P3b (350–600 ms post-feedback onset). (A) ERPs when low reward expectancy was aligned with low reward probability. (B) ERPs when high reward expectancy was matched with high reward probability. Only when both reward probability and reward expectancy were low (see panel A), a distinctive P3a was elicited for negative compared to positive feedback. Note that F stands for feedback onset, significant effects are highlighted in grey, and negativity is plotted upwards.

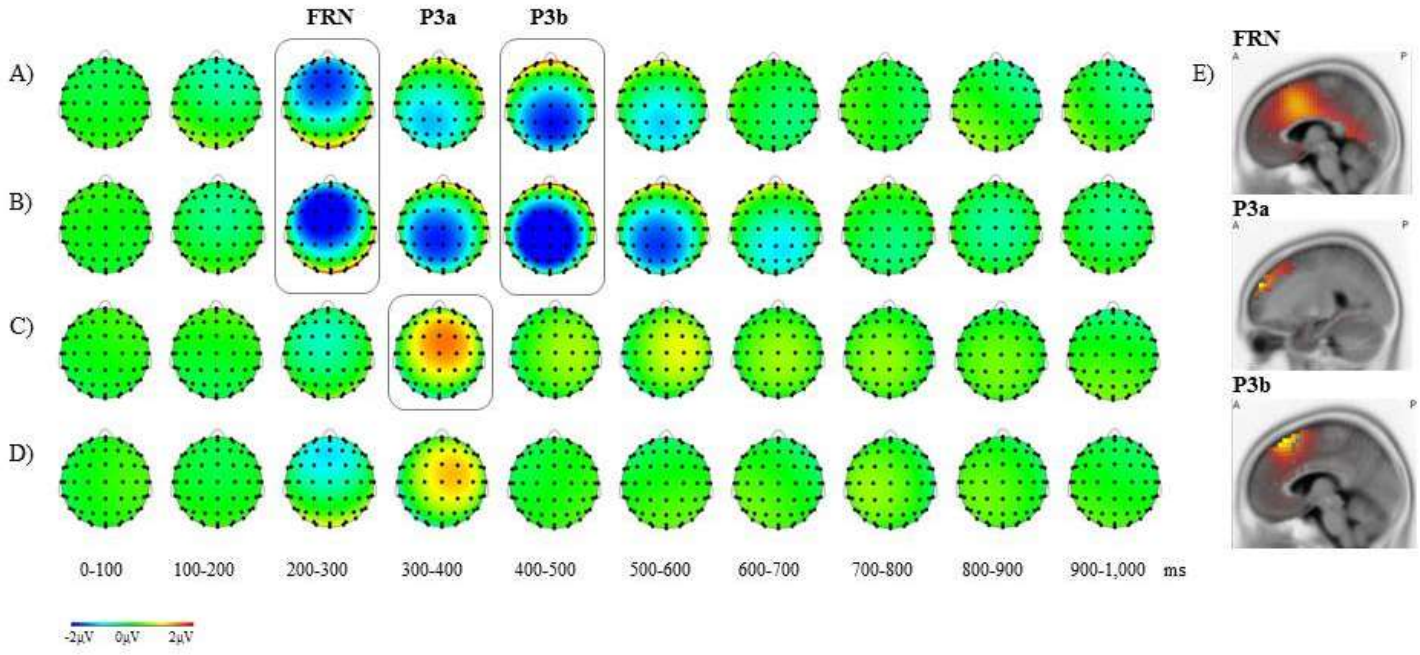
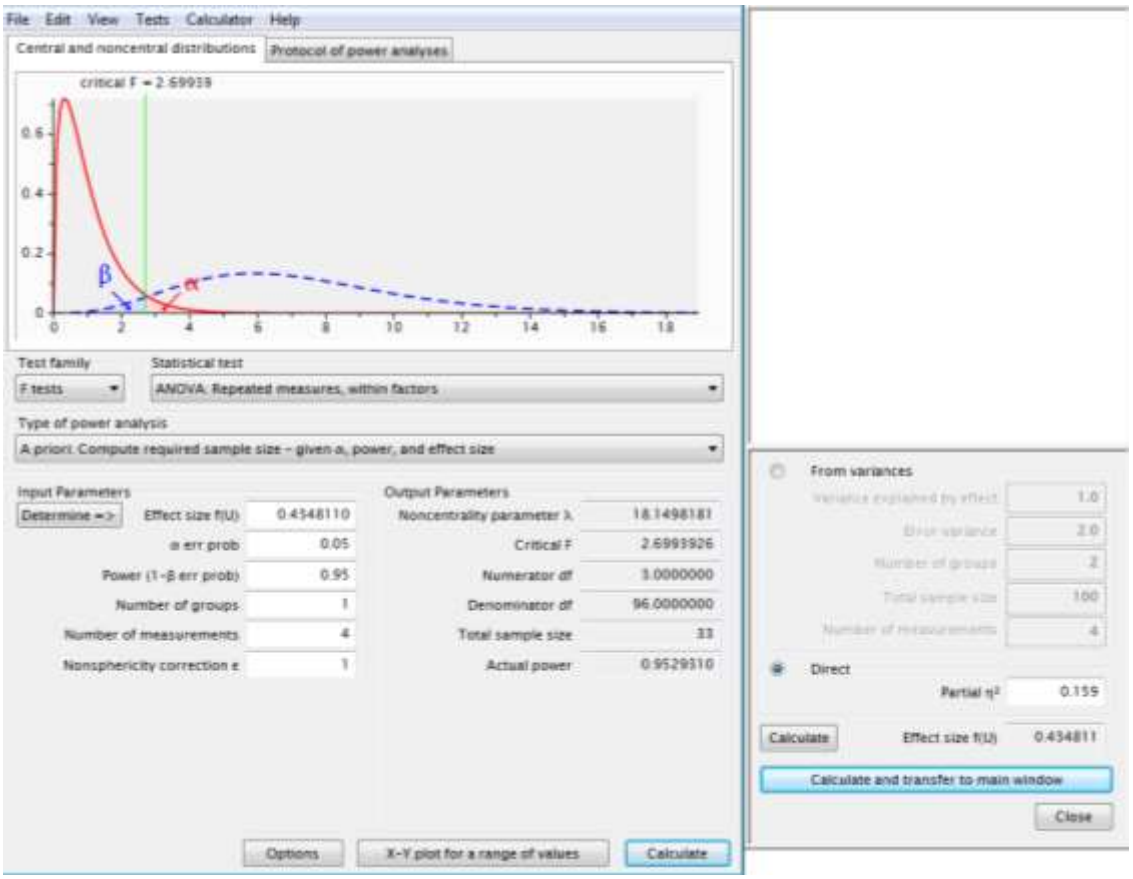


Figure 5. Horizontal topographical maps of the feedback-locked ERP data, separately for each condition. Each map shows the ERP difference wave obtained after subtracting positive from negative feedback (see also blue line in Figures 3 and 4) during a 100-milisecond time interval (mean activity) (A) when low reward expectancy was violated, (B) when high reward expectancy was violated, (C) when low reward expectancy was confirmed, and (D) when high reward expectancy was confirmed. Note that the main ERP effects are marked with a frame. (E) Source-localization results (computed using sLORETA) for the three main ERP components recorded in this study: FRN (upper panel), P3a (middle panel), and P3b (lower panel; see the text for details).

Supplementary Materials

Power analysis

Using G*Power, the following parameters were selected: (i) '*F* tests' were chosen as Test family; (ii) 'ANOVA: Repeated measures, within factors' was chosen as Statistical test; (iii) 'A priori: Compute required sample size – given α, power, and effect size' was chosen as a Type of power analyses; (iv) 'as in SPSS' was chosen during Effect size specification in Options; (v) Effect size f(U) was calculated and transferred to the main window using the direct option (η_p^2) after inserting a value of 0.159 (as based on Walentowska et al., 2016); (vi) α was set up at 0.05 level; (vii) Power was set up at 0.95 level; (viii) number of groups was defined as 1; (ix) number of measurements was defined as 4; (x) after calculations, a total sample size of 33 was estimated. Original G*Power outputs are presented below.



[1] -- *Wednesday, June 26, 2019 -- 21:50:29*

F tests - ANOVA: Repeated measures, within factors

Analysis:	A priori: Compute required sample size		
Input:	Effect size f(U)	=	0.4348110
	α err prob	=	0.05
	Power (1−β err prob)	=	0.95
	Number of groups	=	1
	Number of measurements	=	4
	Nonsphericity correction ϵ	=	1
Output:	Noncentrality parameter λ	=	18.1498181
	Critical F	=	2.6993926
	Numerator df	=	3.0000000
	Denominator df	=	96.0000000
	Total sample size	=	33
	Actual power	=	0.9529310

ERP analysis controlling for the imbalance in the trial count between conditions

We performed an auxiliary analysis where we balanced trial number between conditions to ascertain that amplitude variations of the FRN as a function of valence, probability and expectancy could not be explained by asymmetries in the signal-to-noise ratio between them. More specifically, during EEG data pre-processing (BrainVision Analyzer 2.0 software), we selected a subset of the (clean) epochs available for the negative feedback condition (by means of a standard odd-even selection procedure) and to be included in the averages, eventually matching this trial number with that of the positive feedback condition.

As a result, when reward expectancy was high but the outcome deviated (i.e., worse-than-expected event), the number of negative feedback included in the averages decreased from *M* = 59.89 (SD = 2.13) to *M* = 30.13 (SD = 3.12), eventually reaching a similar amount of trials than for the positive feedback condition, *M* = 29.51 (SD = 1.52). In a similar vein, when reward expectancy was low and reward probability too, this selection procedure made it possible to decrease the number of negative feedback included in the averages from *M* = 61.11 (SD = 2.44) to *M* = 31.81 (SD = 2.74), thereby

matching trial count with the positive feedback condition, M = 30.23 (SD = 2.52). The corresponding waveforms are presented in Figures S1 and S2. As can be seen from them, balancing trial number between conditions did not change the morphology and amplitude of the FRN, where a clear modulation by valence and expectancy was still observed, and was highly comparable to the ones reported when this matching was not performed (compare panel A and B).

[insert Figure S1 and S2 here]

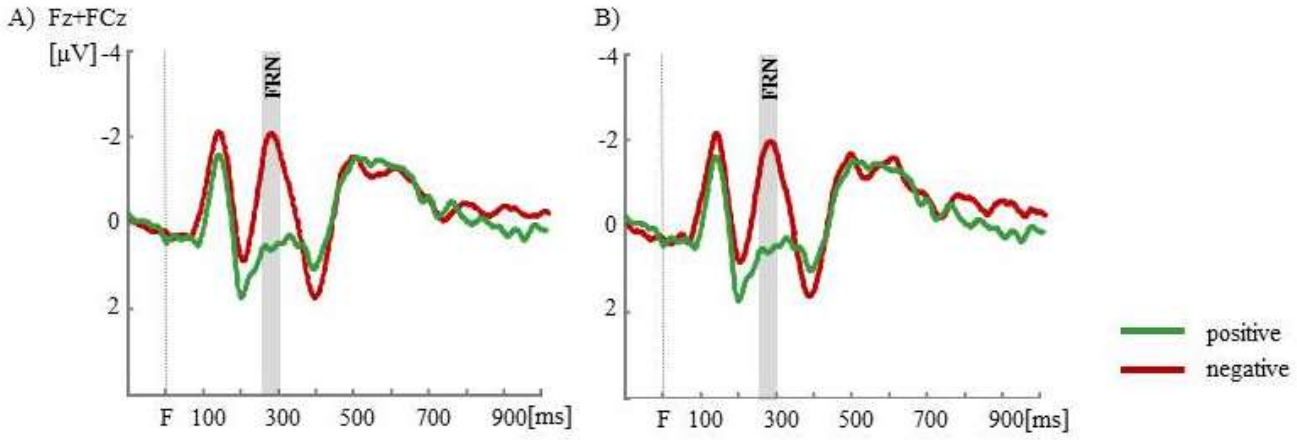


Figure S1. Feedback-locked grand average ERPs, recoded from electrodes Fz and FCz (pooled together), for the condition where high reward expectancy was violated. (A) The processing of negative feedback is compared with positive feedback, when all trials available for each condition are considered and included in the averages. For negative feedback, the averages were based on twice more trials than those for positive feedback. (B) A very similar waveform is obtained for negative feedback when a subset of trials is considered in order to match the trial count with the positive feedback condition.

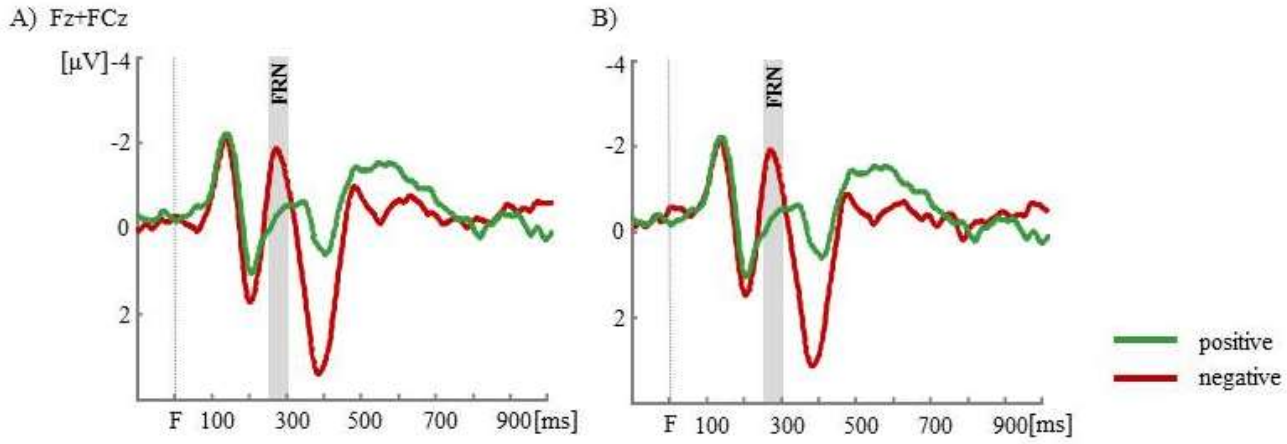


Figure S2. Feedback-locked grand average ERPs, recoded from electrodes (Fz and FCz pooled together), for the low reward expectancy-low reward probability condition. (A) The processing of negative feedback is compared with positive feedback, when all trials available for each condition are considered and included in the averages. For negative feedback, the averages were based on twice more trials than those for positive feedback. (B) A very similar waveform is obtained for negative feedback when a subset of trials is considered in order to match the trial count with the positive feedback condition.

Analysis of the Stimulus Preceding Negativity (SPN) recorded prior to the feedback onset

The SPN was computed offline following a standard procedure of data transformations (Keil et al., 2014): (i) 50-Hz notch filter; (ii) EEG offline referencing using linked mastoids; (iii) -1,000/+1,000 ms segmentation around the onset of the feedback stimulus; (iv) pre-feedback interval baseline correction (from -1000 ms to -900 ms); (v) vertical ocular correction for blinks (Gratton et al., 1983); (vi) semi-automatic artifact rejection (trials with motor artifacts were rejected, with a fixed criterion of ±80 μV) [7];

[7] Number of trials kept after artefact rejection and used for averaging were as follows: in the low probability-low expectancy (*M* = 27.12, SD = 2.77 for positive, and *M* = 58.81, SD = 2.67 for negative feedbacks), in the low probability-high expectancy (*M* = 26.98, SD = 2.52 for positive, and *M* = 57.89, SD = 2.78 for negative feedbacks), in the high probability-high expectancy (*M* = 48.22, SD = 2.92 for positive, and *M* = 46.42, SD = 2.72 for negative feedbacks), and in the high probability-low expectancy (*M* = 46.95, SD = 2.62 for positive, and *M* = 47.07, SD = 2.43 for negative feedbacks) conditions. The two low reward probability conditions did not differ regarding segments

(vii) averaging for each feedback type separately (i.e., positive vs. negative); and (viii) low pass digital filtering of the individual average data (30 Hz). The SPN was later defined as the mean voltage within 200 ms prior to feedback onset, over the left (F3) and right (F4) frontal electrodes. SPN amplitudes for the negative feedback were subtracted from the positive feedback (see Luck & Gaspelin, 2017), and the resulting/differential SPN amplitude values (positive-negative) were submitted to mixed model ANOVAs with the within-subject factors reward EXPECTANCY (low vs. high), reward PROBABILITY (low vs. high), and ELECTRODE (F3 vs. F4).

The ANOVA only showed significant main effects of PROBABILITY, $F_{(1,31)} = 4.92$, $p = 0.038$, $\eta_p^2 = 0.186$, and ELECTRODE, $F_{(1,31)} = 7.89$, $p = 0.009$, $\eta_p^2 = 0.203$. All other main or interaction effects were non-significant (all *ps* > 0.295). The main effect of PROBABILITY showed a larger SPN when reward probability was high (M = -1.81 μV, SD = 1.98) in comparison to when it was low (M = -0.87 μV, SD = 2.11). Importantly, as indicated by the non-significant EXPECTANCY x PROBABILITY interaction, reward probability modulated SPN amplitudes in a general manner, irrespective of reward expectancy. The main effect of ELECTRODE indicated a larger SPN component over right (M = -1.82 μV, SD = 2.12) versus left frontal location (M = -0.86 μV, SD = 1.18) (see Figure S3).

[insert Figure S3 here]

kept for positive, $t_{(31)} = -1.22$, $p = 0.414$, nor negative feedback, $t_{(31)} = -1.03$, $p = 0.551$. Likewise, the two high reward probability conditions had similar number of trials retained after artifact rejection for positive, $t_{(31)} = -0.97$, $p = 0.457$, and negative feedback, $t_{(31)} = 0.67$, $p = 0.634$.

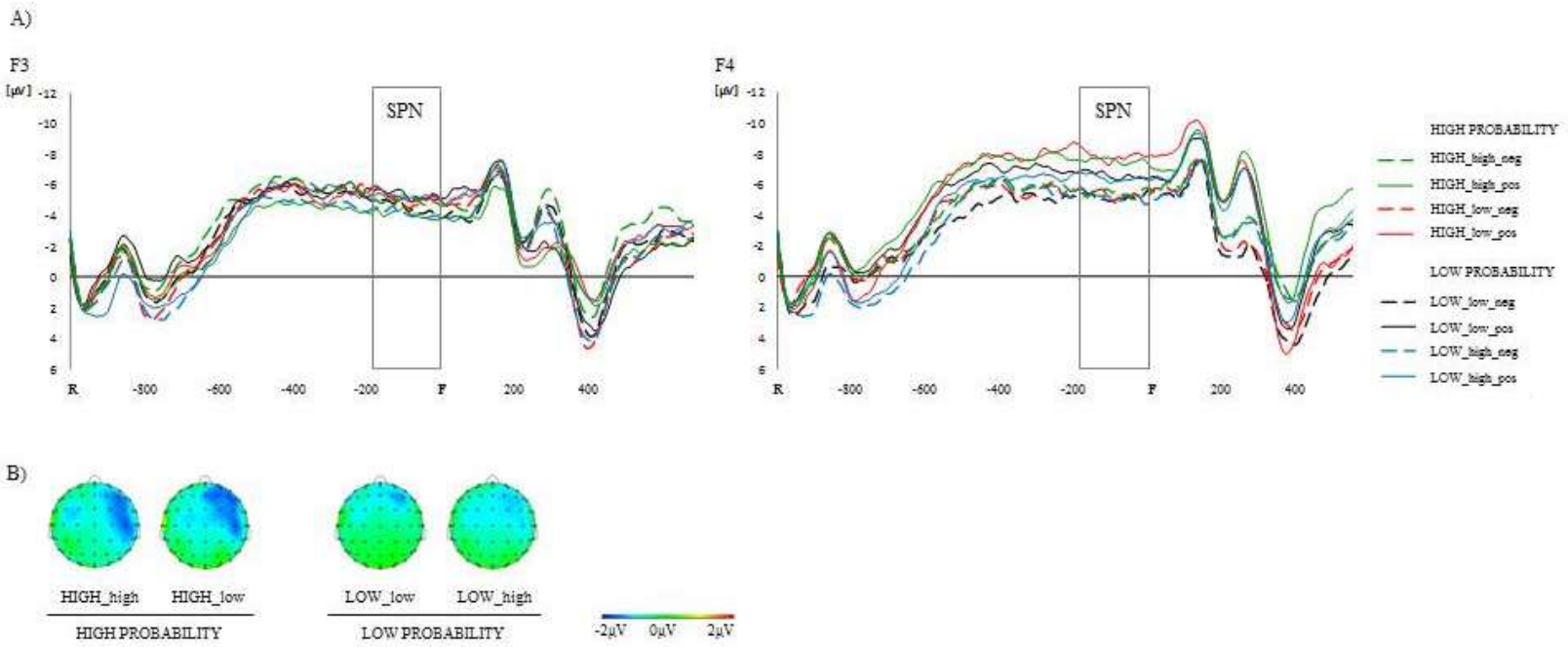


Figure S3. Pre-feedback onset ERP effects, with a focus on the SPN, from F3 and F4 electrodes. (A) Only when reward probability was high (regardless of reward expectancy), the SPN component was significantly larger (more negative) in anticipation of positive than negative feedback. Note that R stands for response onset, and F for feedback onset. The time window of interest (-200/0 ms) is framed, and negativity is plotted upwards. (B) Topographical maps (horizontal view) of the SPN (mean ERP activity computed during 200 ms prior to the feedback onset). Each map shows the ERP difference wave obtained when negative feedback was subtracted from positive feedback. Note that a clear SPN was only visible for high compared to low reward probability conditions, and was larger over right than left frontal locations.