

Learning-related changes in reward expectancy are reflected in the feedback-related negativity

Christian Bellebaum and Irene Daum

Institute of Cognitive Neuroscience, Department of Neuropsychology, Ruhr-University of Bochum, Universitaetsstraße 150, D-44780 Bochum, Germany

Keywords: ACC, FRN, P300, prediction error, reward

Abstract

The feedback-related negativity (FRN) has been hypothesized to be linked to reward-based learning. While many studies have shown that the FRN only occurs in response to unexpected negative outcomes, the relationship between the magnitude of negative prediction errors and FRN amplitude remains a matter of debate. The present study aimed to elucidate this relationship with a new behavioural procedure that allowed subjects to predict precise reward probabilities by learning an explicit rule. Insight into the rule did not only influence subjects' choice behaviour, but also outcome-related event-related potentials. After subjects had learned the rule, the FRN amplitude difference between non-reward and reward mirrored the magnitude of the negative prediction error, i.e. it was larger for less likely negative outcomes. Source analysis linked this effect to the anterior cingulate cortex. P300 amplitude was also modulated by outcome valence and expectancy. It was larger for positive and unexpected outcomes. It remains to be clarified, however, whether the P300 reflects a positive prediction error.

Introduction

In human subjects, the feedback-related negativity (FRN), a negative event-related brain potential occurring between 200 and 300 ms after the presentation of a feedback stimulus, has been linked to the processing of negative performance outcomes (Miltner *et al.*, 1997). It shares some features with the error-related negativity (ERN; Falkenstein *et al.*, 1990; Gehring *et al.*, 1993). Both components appear to play a role in performance monitoring and are generated in or near the anterior cingulate cortex (ACC; Miltner *et al.*, 1997; Gehring & Willoughby, 2002). In contrast to the FRN, however, the ERN is response-locked and peaks about 100 ms after an error response.

While it is generally accepted that the ACC is involved in performance monitoring (Shima & Tanji, 1998; Knutson *et al.*, 2000; Ito *et al.*, 2003), the precise nature of its functional role remains to be clarified. The reinforcement learning (RL) theory of the FRN (Holroyd & Coles, 2002) is based on the assumption that the ACC receives an error signal from the mesencephalic dopamine system via the basal ganglia (Schultz, 1999, 2001), which reflects a negative reward prediction error and is used to optimize the acquisition of new action–outcome relations. Consistent with the RL theory, a significant FRN was only observed in response to unexpected negative outcomes (Holroyd *et al.*, 2003, 2004a; Yasuda *et al.*, 2004; but see Hajcak *et al.*, 2005), but the results concerning the relationship between FRN amplitude and the magnitude of the prediction error are mixed (e.g. Holroyd *et al.*, 2003, 2004a; Yeung & Sanfey, 2004; Hajcak *et al.*, 2006; Cohen *et al.*, 2007). Recently, Hajcak *et al.* (2007) pointed out that previous studies lacked a measure of subjects' actual expectations. Modifying a previously applied paradigm (Hajcak *et al.*, 2005), they asked subjects to indicate on each trial whether or not they expected to

receive reward. Interestingly, FRN amplitude was significantly affected by reward probability, when subjects' expectations were taken into account. These results are the first evidence of FRN amplitude modulation by prediction error magnitude.

The present study aimed to shed further light on the relationship between reward expectancy and FRN, providing evidence that expectations modulate choice behaviour and FRN amplitude in parallel. Modulations of FRN amplitude related to different reward probabilities were assessed with a new behavioural task that tapped changes in subjective reward expectations and provided precise prediction error estimates. As in previous studies, objective reward probabilities were varied. In contrast to the task used by Hajcak *et al.* (2007), expectations were inferred directly from subjects' choice behaviour. FRN amplitude was expected to vary with negative prediction errors of different sizes, being larger the more unexpected the negative outcome.

A further aim of the study was to further elucidate the role of the P300 in reward-related processing. Previous studies yielded inconsistent results, linking the P300 to reward magnitude irrespective of valence (Yeung & Sanfey, 2004; Sato *et al.*, 2005), to positive (Hajcak *et al.*, 2005, 2007) or negative feedback processing (Frank *et al.*, 2005).

Materials and methods

Participants

Twenty-eight healthy, right-handed subjects – 13 women and 15 men – participated in this study. All subjects were students of the Ruhr-University of Bochum. They had normal or corrected-to-normal vision. The data of one male subject had to be excluded from analysis because of technical data acquisition problems. The mean age of the remaining 13 women and 14 men was 25.6 years [standard deviation

Correspondence: Dr C. Bellebaum, as above.
E-mail: christian.bellebaum@rub.de

Received 18 October 2007, revised 1 February 2008, accepted 7 February 2008

(SD) = 4.4]. The study was approved by the Ethics Committee of the Faculty of Medicine of the Ruhr-University of Bochum and all participants gave written informed consent.

The learning task

Subjects were told that: (i) on each trial they could win money by guessing the location where a 5 cent coin was hidden; and (ii) the coin was hidden in one of 12 boxes. Each trial started with the presentation of a fixation cross. Then two sets of 12 rectangles (the boxes) were presented on the left and right side of a computer screen. The participants were told that different subsets of individual boxes were 'pre-selected' in the two sets. The pre-selected boxes were marked in red (grey in Fig. 1A), the non-selected boxes were white (see Fig. 1A). The task was to choose between the left or right box subset by pressing a left or right response button, depending on where they expected the 5 cent coin to be hidden. The subject thus did not have to make a guess about the individual box with the 5 cent coin, but he/she had to judge whether one of the boxes in the left or one of the boxes in the right subset of selected boxes was more likely to contain the reward. The stimuli remained on the screen until a choice was made. The maximal response time was 2700 ms. After the choice was made, the selected set was shown for 500 ms. After another 500-ms interval (black screen), the feedback was displayed for 500 ms. On rewarded trials, an icon of a 5 cent coin appeared in a large red box, signalling reward. When the selected boxes did not contain the coin, an empty red box appeared, indicating non-reward. This feedback was

also given if subjects did not respond within the given time limit. After a 600-ms interval (black screen), the next trial started (see Fig. 1A for the sequence of events on an individual trial).

As illustrated in Fig. 1, each set of 12 boxes was arranged in two columns of six boxes each. The total number of pre-selected boxes was either four or eight, but – unknown to the subject – reward probability did not depend on the total number of pre-selected boxes. The 5 cent coin was always hidden in the right column of a 12-box set, reward probability was thus determined by the number of pre-selected boxes in the right column of a set. Pre-selection could vary in the following way: all six boxes pre-selected in the right column, with reward probability = 1, referred to as a '1-stimulus'; four boxes pre-selected, reward probability = 2/3, a '2/3-stimulus'; two boxes pre-selected, reward probability = 1/3, a '1/3-stimulus'; or no boxes pre-selected, reward probability = 0, a '0-stimulus' (see Fig. 2 for the different types of stimuli used in the experiment). Choosing a 1/3-stimulus thus led to reward in 1/3 of the cases and to non-reward in 2/3 of the cases. Subjects were informed that there was a 'rule' determining the probability of reward and that they could maximize reward by applying this rule when making their choices. The experiment consisted of 660 trials (three blocks of 220 trials each). The sum of money, i.e. the rewards accumulated during the course of the task, was paid to the subjects at the end of the experiment.

Pilot testing showed that the rule was difficult to learn. Therefore, the participants were given additional feedback, following the reward/non-reward feedback, showing the individual box in which the 5 cent coin had been hidden. On rewarded trials, the coin thus

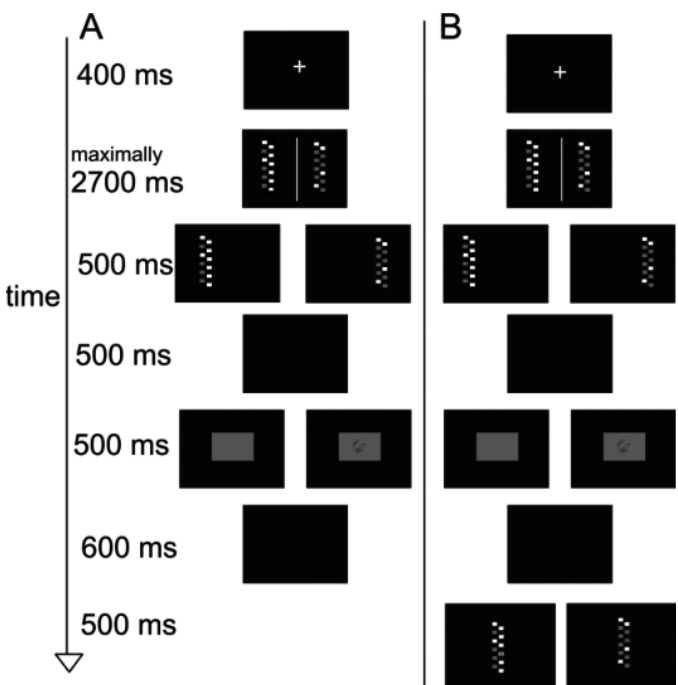


FIG. 1. Temporal sequence of events in a single trial. (A) First and last block of trials (trials 1–220 and 441–660): subjects had to choose a selection of boxes to try to find a 5 cent coin, which was hidden in one of the 12 boxes. In the example trial, choosing the stimulus on the right side leads to reward and choosing the stimulus on the left side is followed by non-reward. (B) Second block of trials (trials 221–440): after the first feedback indicating reward or non-reward, subjects were shown in which box the 5 cent had actually been hidden (see also Materials and methods). The coin was always presented in one of the boxes in the right column of the 12-box set.

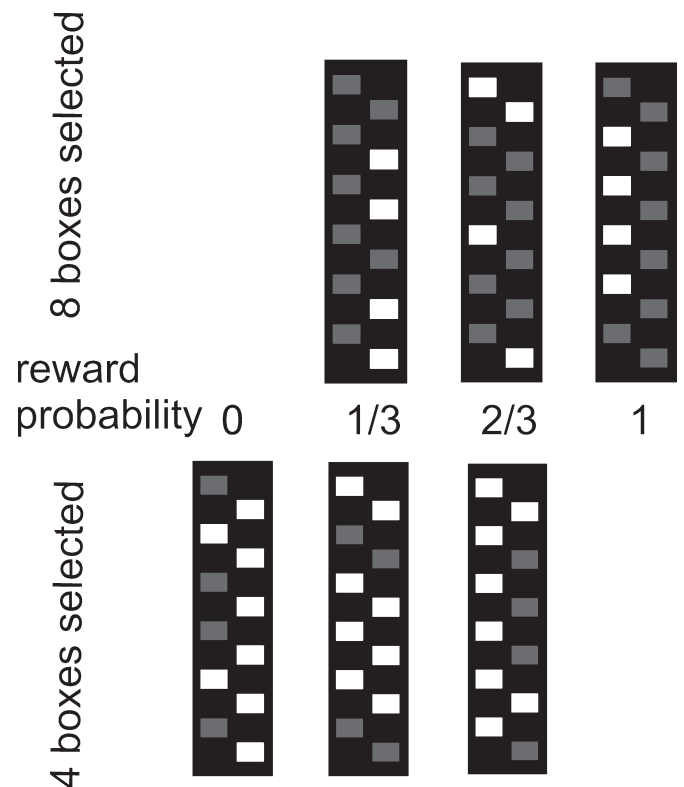


FIG. 2. Stimulus types. The stimuli either contained four or eight selected boxes. As the coin was always hidden in the right column, reward probability was determined by the number of selected boxes in the right column. Thus, for four selected boxes the reward probability was 0, 1/3 or 2/3. Stimuli with eight selected boxes were related to reward probabilities of 1/3, 2/3 or 1. There were 10 different stimuli for each stimulus type.

appeared in one of the red boxes in the right column of a stimulus set, and on non-rewarded trials in a white box of the right column. The exact location of the coin (i.e. which individual red or white box contained the reward) was determined randomly (see Fig. 1B for the additional feedback trials). The additional feedback was introduced in the second block of trials (trials 221–440), because analysis was based on event-related potentials (ERPs) in pre- and post-learning phases, and a sufficient number of pre-learning trials was required.

As outlined above, subjects had to choose between the left and right box sets, each being associated with a specific reward probability. Table 1 illustrates the different stimulus pairs (e.g. 0- vs 1/3-stimuli or 1/3- vs 2/3-stimuli). For ERP analysis, different trial types were defined based on subjects' choices, irrespective of the alternative, non-chosen stimulus. Only trials in which a 1/3- or 2/3-stimulus was chosen were considered for FRN analysis (1/3-choice or 2/3-choice), because only in these trials both reward or non-reward could be followed by the choice. The pairings 1/3 vs 1 and 0 vs 2/3 mainly served to facilitate learning, but they also entered analysis, when the 1/3 or 2/3 alternatives were chosen. Analysis of the P300 followed the same procedure. An additional, exploratory analysis that was based on rewarding outcomes only also included the trials, in which subjects chose a 1-stimulus. On 120 of the total number of 660 trials, reward probabilities for both box sets were identical (1/3 or 2/3) and the stimuli only differed with respect to the total number of pre-selected boxes (four and eight). These trials were introduced to assess whether subjects selected stimuli on the basis of the total number of selected boxes, i.e. if they preferred stimuli with eight or four selected boxes. In 360 of the remaining 540 trials, subjects' choices could clearly be attributed to learning the rule, because the two stimuli from which subjects had to choose both involved either four or eight pre-selected boxes, and thus only differed with respect to reward probability. These trials were used for the analysis of behavioural data (see below).

Rationale of the task

The present study explicitly tested the hypothesis that the magnitude of the FRN in response to non-reward compared with reward increases with increasing negative prediction errors, as suggested by the RL theory (Holroyd & Coles, 2002). More specifically, it was expected that: (i) subjects would gain insight into the rule during the course of the experiment and would adapt their choice behaviour accordingly, i.e. they would consistently choose the subset of boxes with more pre-selected boxes in the right column; and (ii) their expectations with respect to reward or non-reward would be reflected

TABLE 1. Combinations of reward probabilities in single trials and their frequency

Type of stimulus		Number of trials
1-stimulus reward probability	2-stimulus reward probability	
0	1/3	120/660
1/3	2/3	240/660
0	2/3	90/660
1/3	1/3	60/660
2/3	2/3	60/660
1/3	1	90/660

The sides on which the 1- and 2-stimulus were presented (left or right) were counterbalanced throughout the experiment.

in the ERPs. After learning the rule, subjects should be more surprised to receive non-reward following a 2/3- compared with a 1/3-choice. In both conditions, non-reward should lead to a larger FRN compared with reward, but the negative prediction error should be higher following 2/3-choices, yielding a larger FRN amplitude relative to reward.

Electroencephalography (EEG) recording

Subjects were comfortably seated approximately 70 cm in front of a computer monitor. The left and right CTRL keys of a computer keyboard were used as response keys for choosing the left or right box sets. Throughout the task, EEG was recorded from 30 scalp sites according to the International 10–20 system with silver–silver chloride electrodes: F7, F3, Fz, F4, F8, FT7, FC3, FCz, FC4, FT8, T7, C3, Cz, C4, T8, TP7, CP3, CPz, CP4, TP8, P7, P3, Pz, P4, P8, PO7, PO3, POz, PO4, PO8. Recordings were referenced to the average of two electrodes placed on the left and right mastoids. Stimulus timing was controlled by Presentation Software (Neurobehavioural Systems; <http://www.neuro-bs.com>). Data were recorded with a sample rate of 500 Hz using a Neuroscan Synamps System and the appropriate Software. Impedances were kept below 10 k Ω .

Data analysis

EEG data

EEG and electrooculography data were analysed off-line using the Brain Vision Analyser Software Package and Matlab. After applying a 0.1-Hz high-pass and a 40-Hz low-pass-filter, an independent component analysis was performed on single-subject EEG data (Lee *et al.*, 1999). Independent component analysis yields an unmixing matrix, which decomposes the multichannel scalp EEG into a sum of temporally independent and spatially fixed components. The number of components matches the number of channels. Each resulting component is characterized by a time course of activation and a topographical map. For each subject, the 30 components were screened for maps with a symmetric frontally positive topography, which could represent eye movement and blink artefacts. These components were then removed from the raw data by performing a back transformation. In most subjects, one eye movement-related component was identified and removed; for the remaining subjects, two components were related to eye movements. In all subjects, back-transformed data were screened for eye movement and blink artefacts. Application of artefact-detection techniques, which automatically exclude trials with data points exceeding an absolute amplitude value of 150 μ V, indicated that all trials were free of large artefacts after back-transformation.

To analyse feedback-related ERPs, segments were created which started 200 ms before and up to 600 ms after feedback presentation (reward or non-reward). Trials with EEG artefacts (segments with an amplitude difference of more than 150 μ V between the highest and lowest data point) were excluded. For FRN analysis, data from electrode sites FC3, FCz and FC4 were pooled by averaging, as the largest amplitudes were seen at these positions. Based on the procedure of previous studies (e.g. Potts *et al.*, 2006), a mean amplitude measure was applied as mean amplitude measures are less susceptible to differences in the number of trials between conditions (see Luck, 2005). The time window between 220 and 280 ms was used. For P300 analysis, data from three parietal electrodes were pooled (P3, Pz and P4). The mean amplitude between 300 and 500 ms after feedback presentation was analysed.

Behavioural data: pre- and post-learning periods

As significant modulations of FRN amplitude were only expected for the period after subjects had insight into the rule determining reward probability, ERP analysis focused on the learners. For every learner, the experiment was divided into pre- and post-learning periods based on the individual behavioural data. Of the total 660 trials, those 360 trials were used for behavioural analysis in which subjects' choices were unequivocally related to reward probabilities, i.e. the two stimuli only differed on reward probability and not on the number of boxes pre-selected (four or eight, see above – unequivocal trials). To assess the onset of learning as exactly as possible, response accuracy in overlapping blocks of 20 trials was analysed (trials 1–20, 2–21, 3–22, ...). The block, on which a stable performance criterion of 80% correct responses (i.e. 16 correct responses within the last 20 trials) was reached and maintained until the end of the experiment, was considered as learning onset for an individual subject. The trial number was then related to the whole series of all 660 trials. If, for example, a subject reached the learning criterion on trial 150 of the unequivocal trials, this may have corresponded to trial 290 relative to all trials. For this individual subject, the pre-learning period would entail trials 1–290, and the post-learning period trials 291–660, respectively.

Although the sequence of events on an individual trial was slightly different in the second block of trials, which included an additional feedback stimulus, all trials were pooled in the analysis. The second feedback stimulus was not expected to significantly affect the ERPs in response to the first feedback stimulus, as the first feedback stimulus still served to signal reward or non-reward. An exploratory analysis based only on the trials of the blocks without second feedback (i.e. blocks 1 and 3) yielded the same pattern of ERP findings as the analysis based on all trials. Therefore, all trials were included in ERP analysis.

To further explore a specific link of the ERP data and learning, the data of the non-learners were also analysed. For each individual non-learner, the experiment was also divided into a first and a second part, based on the distribution of learning onset trials in the learners. For each subject, the minimum number of trials that entered analysis of feedback-related ERPs in the 1/3- and 2/3-conditions was 10 trials. For the learners, analysis was on average based on 80 non-rewarded and 38 rewarded 1/3-trials, and on 34 non-rewarded and 66 rewarded 2/3-trials in the pre-learning period. For the post-learning period, 67 non-rewarded and 34 rewarded 1/3-trials entered analysis, whereas the number of trials for 2/3-choices was higher: 68 non-rewarded and 138 rewarded trials.

In the non-learners, the average frequencies of the different trial types were as follows: first part (corresponding to 'pre-learning'): 79 non-rewarded and 38 rewarded 1/3-trials, 30 non-rewarded and 60 rewarded 2/3-trials; second part ('post-learning'): 95 non-rewarded and 45 rewarded 1/3-trials, 38 non-rewarded and 80 rewarded 2/3-trials.

As the total number of 1-choice-trials was generally low, the minimum number of trials to reach the additional exploratory P300 analysis including 1-stimuli was set to 5. On average, 20 trials entered analysis for the pre-learning period and 52 trials for the post-learning period in the learners. In the non-learners, 18 trials entered analysis for the first part and 25 trials for the second part.

Source analysis

LORETA (low-resolution brain electromagnetic tomography) source analysis (Pascual-Marqui *et al.*, 1994) was performed for FRN amplitude contrasts between non-rewarded and rewarded trials,

separately for low- and high-probability outcomes. It has consistently been shown that the FRN is generated in the ACC. Source analysis in the present study aimed to elucidate differences in FRN sources related to different reward probabilities. As the FRN was largest in a latency range between 230 and 270 ms in all conditions, this time window was chosen for analysis. LORETA is based on computation of the current density at each grey matter voxel of a reference brain as a linear, weighted sum of the scalp electrical potentials. The smoothest of all possible current density configurations is chosen, with the only constraint that neighbouring voxels should have maximally similar activity. LORETA images represent electrical activity at each voxel as squared current density. The solution space consists of 2394 voxels with a size of $7 \times 7 \times 7 \text{ mm}^3$, covering the cortical grey matter and the hippocampi.

For every learner, FRN-related LORETA images were created for 1/3- and 2/3-conditions in the post-learning period, based on the global field power (no band-specific power). The images were converted (http://www.ihb.spb.ru/~pet_lab/L2S/L2SMain.htm) and further analysed using SPM99 (<http://www.fil.ion.ucl.ac.uk/spm/>). A PET/SPECT design with a two-sample *t*-test was performed with the following parameters: global normalization with proportional scaling to a mean of 50, absolute threshold masking with a threshold of 5 and global calculation of mean voxel value (within per image). The level of significance was set to $P = 0.05$. In accordance with other studies using this technique, an uncorrected significance level was applied (e.g. Wills *et al.*, 2007), but only those sources were considered that consisted of at least five voxels. The coordinates of the foci of significant differences between conditions were transformed into Talairach coordinates (Talairach & Tournoux, 1988) with the algorithm suggested by Brett (<http://www.mrc-cbu.cam.ac.uk/Imaging/Common/mnispace.shtml>).

Results

Behavioural data

Of the 27 subjects entering behavioural data analysis, only 18 showed convincing evidence of learning, i.e. gaining insight into the rule determining reward. Figure 3 shows behavioural data of three single subjects to illustrate different courses of learning, as well as group data for the 18 learners and the nine non-learners. In relation to the 360 unequivocal trials (see Materials and methods), the learners reached the learning criterion on average after trial 142 (SD = 77), corresponding to trial 268 (SD = 137) relative to all 660 trials. As expected, most subjects learned about reward probabilities during the second block of trials, when feedback about the exact location of the hidden coin was given (e.g. subject RR in Fig. 3A). However, two subjects only learned in the last block of trials, whereas four subjects already gained insight in the first block, before the additional feedback was introduced. One of these subjects had to be excluded from ERP analysis, because the number of pre-learning trials was insufficient for ERP analysis (subject BK, Fig. 3B). The remaining learners reached the learning criterion on trial 282 relative to all 660 trials (SD = 129). Figure 3C illustrates the data of a non-learner (subject IM).

It might be surprising that nine subjects did not learn the rule, although they received information about the exact location of the coin in the second block of trials. As revealed by the post-experimental interviews, these subjects were looking for a more complex rule, overlooking that the coin was always hidden in the right column.

One non-learner showed a consistent preference for choice of stimuli with eight pre-selected boxes over stimuli with four pre-selected

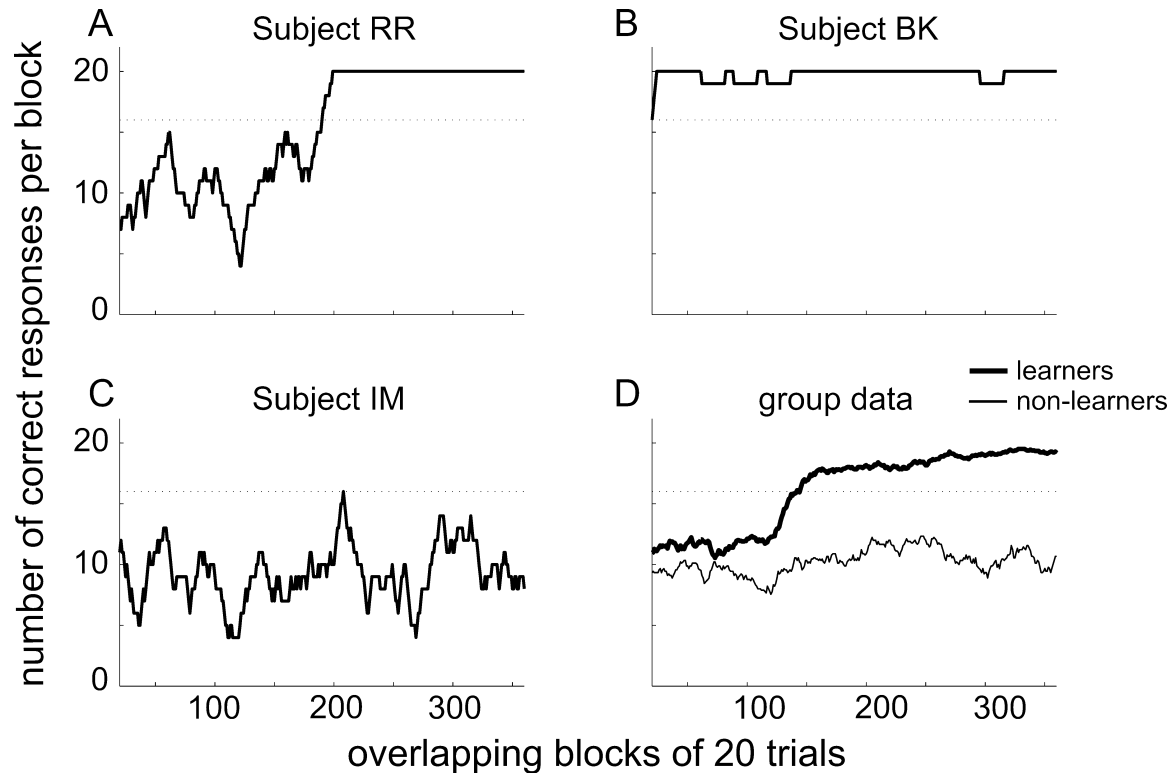


FIG. 3. Behavioural data. Number of correct responses per block for two learners (A and B), one non-learner (C) and for the groups of learners and non-learners (D). The experiment consisted of 660 trials, but only in 360 trials subjects' choices were unequivocally linked to reward probabilities and not to the total number of selected boxes.

boxes. This subject was excluded from ERP analysis, as the preference bias might have affected feedback-locked ERPs. For the remaining subjects, none of whom showed such a preference pattern, the total number of selected boxes was not considered further in the analysis.

As outlined in the Materials and methods, the experiment was divided into a first and a second part for the non-learners, in analogy to the pre- and post-learning periods in the learners. For the eight non-learners, who entered ERP analysis, the second part of the experiment started on average in trial 281 (SD = 190).

ERP data

Learners – FRN

Figure 4 illustrates feedback-locked ERPs from three pooled frontocentral electrode sites associated with 1/3- and 2/3-choices in the learners for the pre- and post-learning periods. On the left side, ERPs related to high-probability outcomes are shown (i.e. non-rewarded choices of 1/3-stimuli and rewarded choices of 2/3-stimuli). On the right side, ERPs related to low-probability outcomes are shown.

For FRN amplitude, repeated-measures ANOVA with the factors OUTCOME (non-reward vs reward), PROBABILITY (high-probability vs low-probability outcomes) and PERIOD (pre- vs post-learning) yielded main effects of OUTCOME ($F_{1,16} = 10.360$; $P = 0.005$) – reflecting a significantly larger FRN in response to non-reward compared with reward – and a main effect of PERIOD ($F_{1,16} = 14.457$; $P = 0.002$), indicating that amplitudes were generally more negative in the post-learning period. This amplitude decrease was significantly modulated by the factor PROBABILITY, with low-probability (i.e. unexpected) outcomes showing a larger

decrease than high-probability outcomes (significant two-way interaction between PERIOD and PROBABILITY; $F_{1,16} = 5.878$; $P = 0.028$). The main effect of PROBABILITY as well as the remaining two-way interactions did not reach significance (all $P > 0.420$). A near-significant three-way interaction ($F_{1,16} = 3.267$; $P = 0.090$) yields some evidence that the large negativity in the post-learning period was not only modulated by outcome PROBABILITY, but also by the OUTCOME itself.

As we had a specific *a priori* hypothesis concerning an interaction between the factors PROBABILITY and OUTCOME after, but not before, subjects had learned the rule, planned follow-up repeated-measures ANOVAs with these two factors were conducted separately for the pre- and post-learning periods (see, e.g. Weisberg *et al.*, 2007 for a similar procedure). For both pre- ($F_{1,16} = 8.031$; $P = 0.012$) and post-learning ($F_{1,16} = 10.708$; $P = 0.005$) significant main effects of OUTCOME were observed: following non-reward the amplitude was significantly more negative. A significant main effect of PROBABILITY ($F_{1,16} = 4.833$; $P = 0.043$) and a significant interaction between both factors ($F_{1,16} = 6.405$; $P = 0.022$) were, however, exclusively observed in the post-learning period (both $P > 0.250$ for the pre-learning period).

To resolve the PROBABILITY–OUTCOME interaction in the post-learning period, paired-samples *t*-tests (one-tailed) were conducted. For both high- and low-probability outcomes, non-reward elicited a significantly more negative amplitude in the FRN time window compared with reward. This difference was more pronounced for low-probability (i.e. less expected) outcomes ($t_{16} = -4.224$; $P < 0.001$) compared with expected outcomes ($t_{16} = -1.895$; $P = 0.038$). As indicated by further *t*-tests, the interaction was clearly driven by the FRN in response to non-rewarding outcomes. The FRN following the presentation of non-reward was significantly larger for low- compared

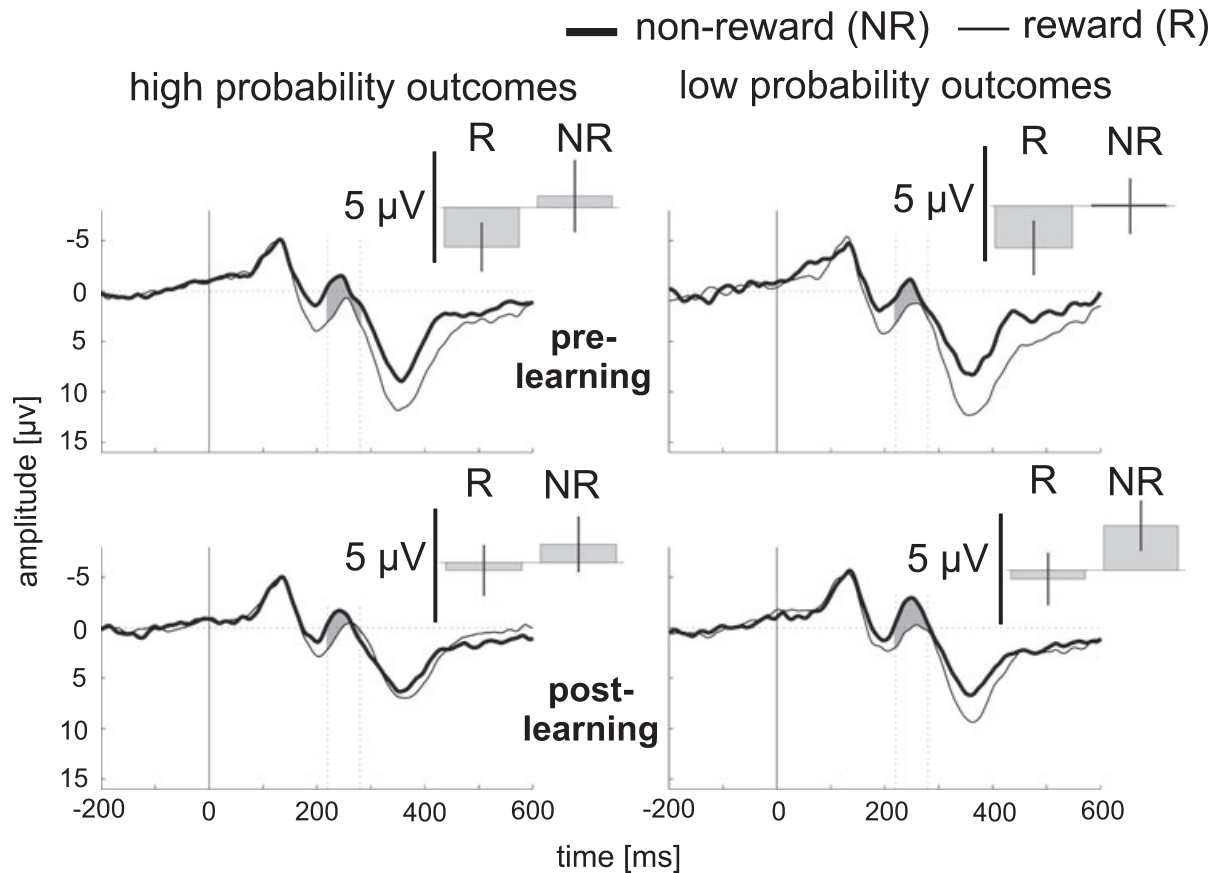


FIG. 4. Frontal ERPs for the pre- and post-learning periods in the learners (pooled data from electrode positions FC3, FCz and FC4). On the left, ERPs for highly probable rewarding and non-rewarding outcomes (corresponding to expected outcomes in the post-learning period) are shown. The right side shows ERPs for improbable (unexpected) outcomes. The time window for the FRN is shaded in grey. Bars above the ERP traces show average amplitudes for rewarded (R) and non-rewarded (NR) trials in the analysed time window.

with high-probability (unexpected vs expected) outcomes ($t_{16} = -2.694$; $P = 0.008$), whereas the FRNs following reward did not differ significantly in the two conditions ($t_{16} = -0.200$; $P = 0.422$).

LORETA source analyses were performed for the post-learning period, to explore the neural sources of outcome- and probability-dependent modulations of the FRN amplitude. Figure 5 illustrates the locations of the FRN sources for the contrast between non-rewarded and rewarded trials.

High-probability outcomes were associated with a source in the bilateral, ventral and dorsal ACC in Brodmann areas (BA) 24 and 32. For low-probability outcomes, four sources were observed. The largest source was again observed in the bilateral, ventral and dorsal ACC (BA 24 and 32). In contrast to high-probability outcomes, the source covered a larger region and extended into the medial and superior frontal gyrus (BA 8 and 9). A smaller source emerged in more lateral parts of the superior frontal gyrus in BA 9 and 10 in the right hemisphere. Two further, small sources were found in the right temporal lobe, one with a centre on the dorsal surface of the temporal lobe (BA 42), the second being located in the inferior temporal gyrus (BA 20).

Learners – P300

Figure 6 illustrates feedback-locked ERPs from three pooled parietal electrode sites. Repeated-measures ANOVA of P300 amplitude yielded significant main effects for all three factors: the main effect of

OUTCOME ($F_{1,16} = 16.801$; $P < 0.001$) showed that the P300 was generally larger following positive outcomes; The main effects of PERIOD ($F_{1,16} = 29.404$; $P < 0.001$) and PROBABILITY ($F_{1,16} = 24.860$; $P < 0.001$) indicated that the P300 was larger in the pre-learning period and larger for low-probability outcomes, respectively. There were significant interactions between OUTCOME and PERIOD ($F_{1,16} = 12.896$; $P = 0.002$) and OUTCOME and PROBABILITY ($F_{1,16} = 13.916$; $P = 0.002$). For the first interaction, *post hoc t*-tests indicated that the amplitude difference between reward and non-reward was larger in the pre-learning period ($t_{16} = 4.916$; $P < 0.001$) compared with post-learning ($t_{16} = 2.990$; $P = 0.009$). The second interaction indicated that the P300 amplitude difference was larger following low- compared with high-probability outcomes, as revealed by *post hoc t*-tests ($t_{16} = 3.135$; $P = 0.006$ for high-probability outcomes and $t_{16} = 4.749$; $P < 0.001$ for low-probability outcomes). There was, however, no significant or near-significant interaction between the factors PERIOD and PROBABILITY or three-way interaction (both $P > 0.338$).

Non-learners

In the non-learners, repeated-measures ANOVA of mean FRN amplitude yielded a significant interaction between PERIOD and OUTCOME ($F_{1,7} = 18.266$; $P = 0.004$; see Fig. 7). *Post hoc t*-tests showed that FRN amplitudes following rewarded and non-rewarded trials only differed significantly in the first part of the experiment ($t_7 = 2.386$; $P = 0.048$; $P = 0.582$ for the second part). This effect

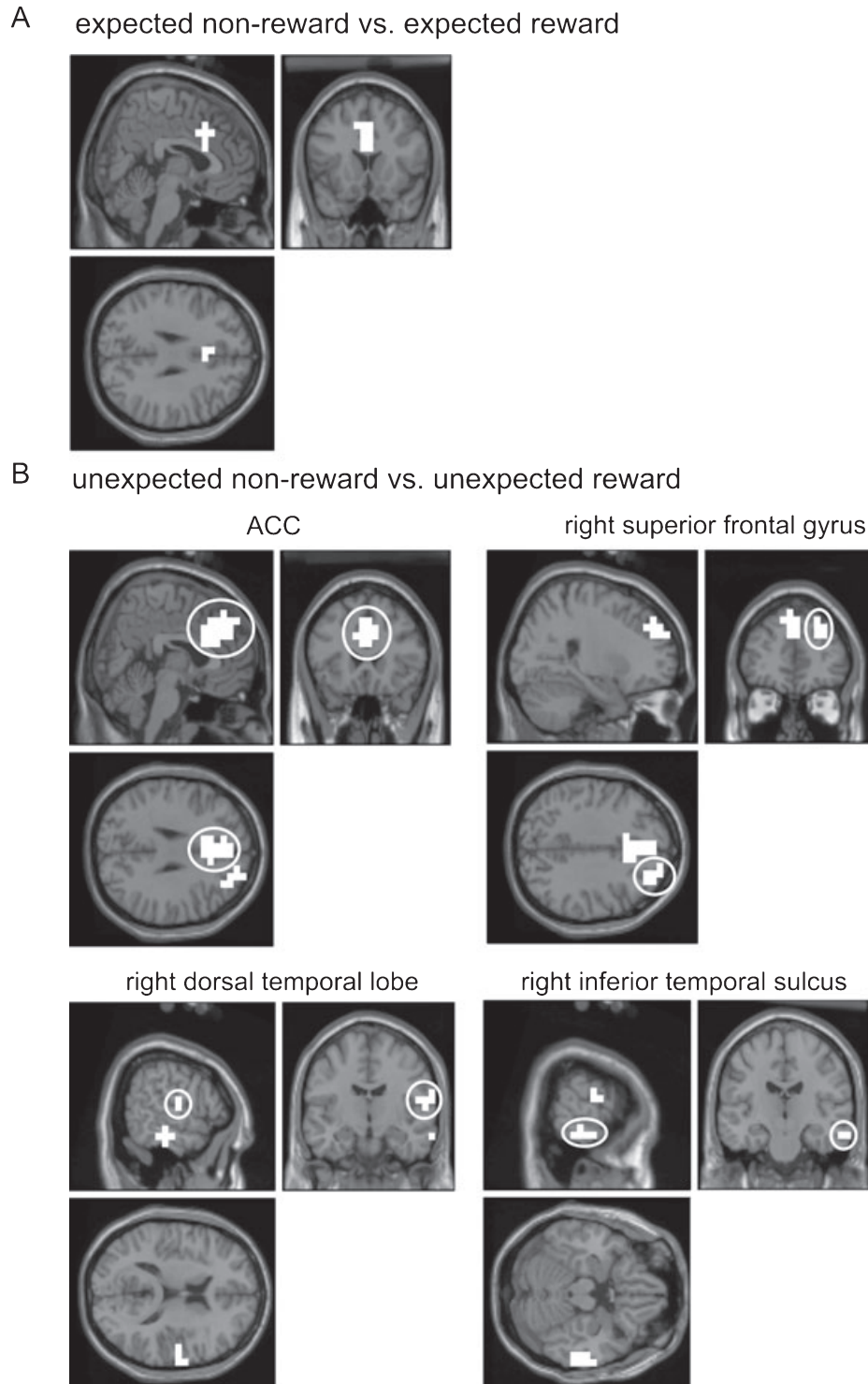


FIG. 5. Results of LORETA source analysis in the learners for the contrast non-reward–reward in the time window 230–270 ms for (A) highly probable (expected) and (B) low probable (unexpected) outcomes in the post-learning period. Each source of significant activations is projected on T1-slices of a reference brain. ACC, anterior cingulate cortex.

appeared to be driven by the responses to non-reward, which showed a tendency to be more positive in the second compared with the first part ($t_7 = 2.201$; $P = 0.064$), whereas the ERPs following reward did not change in the course of the experiment ($P = 0.194$). None of the other main effects or interactions reached significance (all $P > 0.143$).

The P300 (see Fig. 8) was significantly larger for rewarding compared with non-rewarding outcomes ($F_{1,7} = 36.489$; $P < 0.001$). The amplitude difference between rewarded and non-rewarded trials was again larger in the pre-learning period, as was shown by a significant PERIOD–OUTCOME interaction ($F_{1,7} = 9.958$; $P = 0.016$). Pre-learning ($t_7 = 5.916$; $P < 0.001$) as well as post-

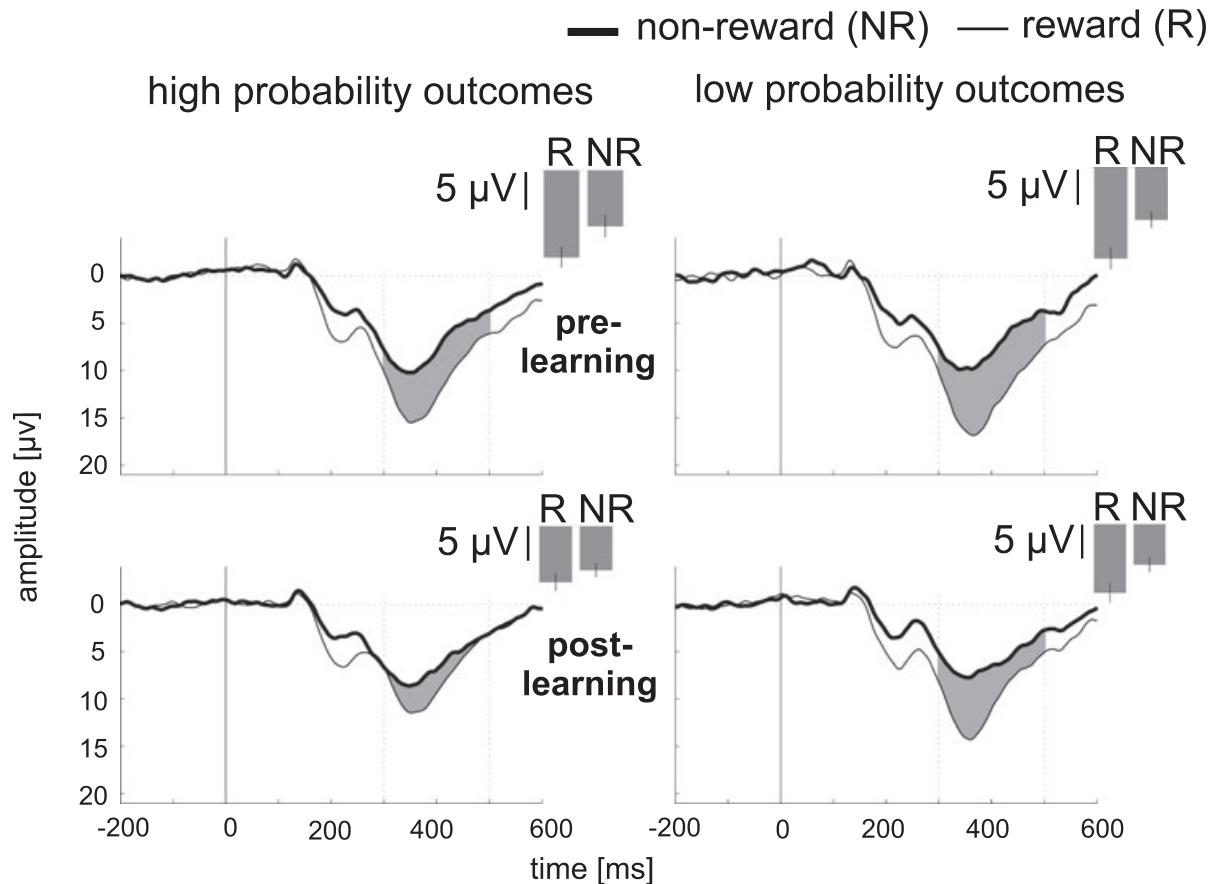


FIG. 6. Parietal ERPs for the pre- and post-learning periods in the learners (pooled data from electrode positions P3, Pz and P4). On the left, ERPs for highly probable (expected) rewarding and non-rewarding outcomes are shown. The right side shows ERPs for low-probability (unexpected) outcomes. The time window for the P300 is shaded in grey. Bars above the ERP traces show average amplitudes for rewarded (R) and non-rewarded (NR) trials in the analysed time window.

learning ($t_7 = 5.964$; $P < 0.001$), the amplitudes were significantly different.

The P300 and positive outcomes

To further analyse the role of the P300 in the coding of positive outcomes, an exploratory analysis was performed that compared P300 amplitudes following rewards after 1/3- and 2/3-choices with trials in which subjects chose a stimulus with a reward probability of 1 (Fig. 9).

For the learners, ANOVA with the factors PROBABILITY (1/3, 2/3 and 1) and PERIOD (before and after learning) yielded significant main effects of both factors (factor PERIOD: $F_{1,16} = 104.814$; $P < 0.001$; factor PROBABILITY: $F_{2,16} = 24.664$; $P < 0.001$) and a significant interaction ($F_{2,16} = 5.415$; $P = 0.009$). *Post hoc* ANOVAs for the pre- and post-learning periods showed that P300 amplitude was significantly affected by probability in both periods. For pre-learning, the amplitudes differed significantly ($F_{2,16} = 5.007$; $P = 0.013$), but the effect was stronger for post-learning ($F_{2,16} = 27.981$; $P < 0.001$). A series of *post hoc* *t*-tests indicated that the P300 amplitude following rewarded 1/3-choices was larger compared with 2/3-choices ($t_{16} = 3.310$; $P = 0.004$) and 1-choices ($t_{16} = 2.620$; $P = 0.018$) in the pre-learning period, whereas there was no significant difference between 2/3- and 1-choices ($P = 0.340$). In the post-learning-period, the P300 following 1/3-choices was significantly larger than the P300 following 2/3-choices ($t_{16} = 4.236$; $P < 0.001$), which in turn was larger than the potential following 1-choices ($t_{16} = 3.749$; $P = 0.002$).

For the non-learners, the ANOVA with the factors PERIOD and PROBABILITY did not reveal any significant effects (all $P > 0.080$). The only near-significant effect emerged for the factor PERIOD, reflecting that P300 amplitudes tended to be smaller in the second part of the experiment.

Discussion

Reward-related information processing was investigated using a new behavioural task, which induced precise prediction errors in those subjects who gained insight into a rule that determined reward probabilities. With respect to ACC activity – as reflected in the FRN – the present results extend earlier reports by showing that the FRN does not only code unexpected negative outcomes but also the magnitude of the difference between actual and expected outcome. The results are thus in line with recent reports, which showed a similar modulation with different behavioural paradigms (Hajcak *et al.*, 2007; Holroyd & Krigolson, 2007). In addition, the present study clearly relates this effect to learning about reward probabilities, as it is neither observed before subjects gained insight into reward probabilities nor in non-learners. The present data also suggest that the P300 might be more closely related to coding positive than negative outcomes.

The FRN and negative prediction errors

Holroyd & Coles (2002) hypothesized that ACC activity mirrors the activity of midbrain dopaminergic neurons, which code errors in reward

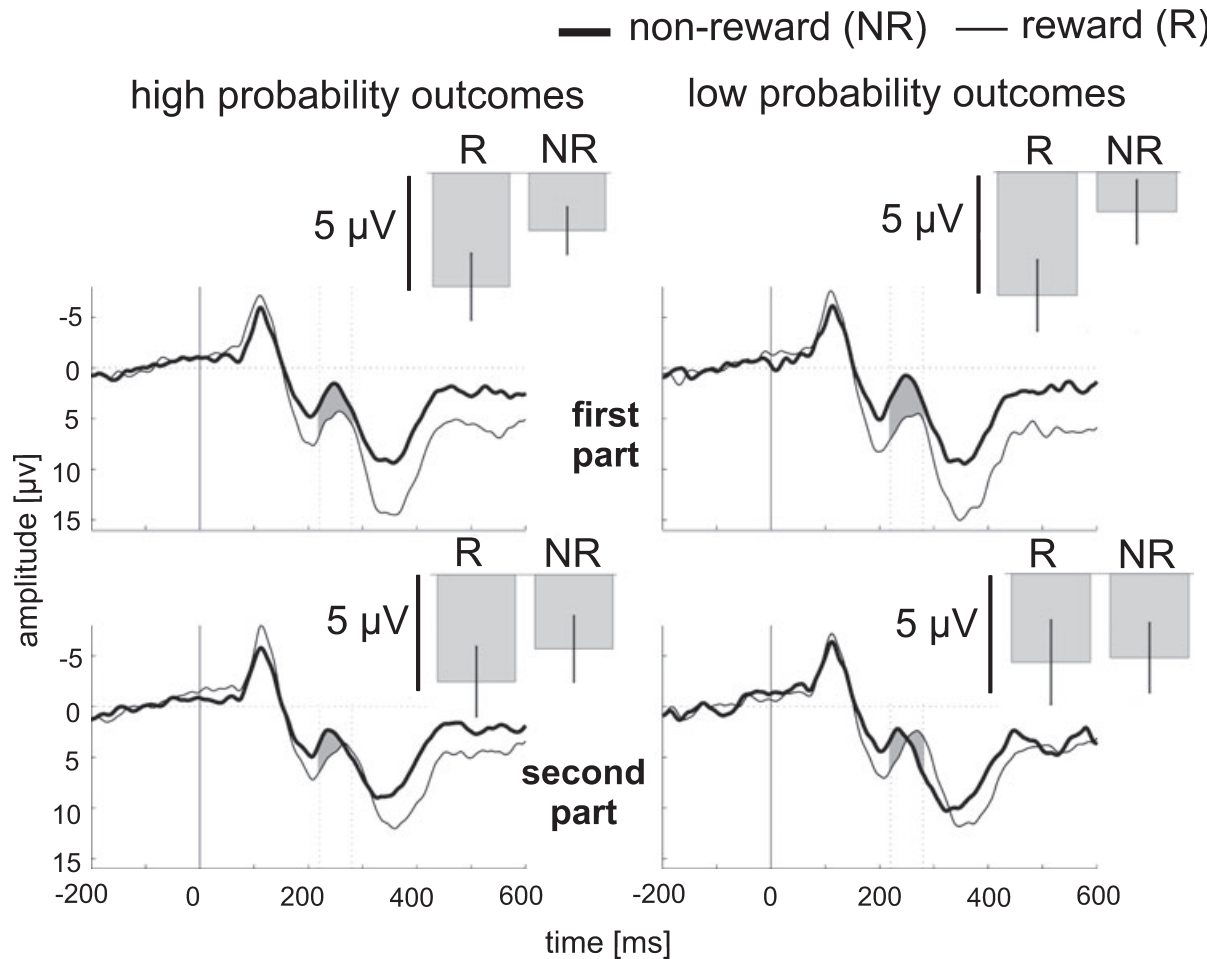


FIG. 7. Frontal ERPs for the first and second part of the experiment in the non-learners (pooled data from electrode positions FC3, FCz and FC4). On the left, ERPs for highly probable rewarding and non-rewarding outcomes are shown. The right side shows ERPs for improbable outcomes. The time window for the FRN is shaded in grey. Bars above the ERP traces show average amplitudes for rewarded (R) and non-rewarded (NR) trials in the analysed time window.

prediction. Consistent with this hypothesis, larger FRN amplitudes were observed in response to unexpected negative outcomes (Holroyd *et al.*, 2003, 2004a; Yasuda *et al.*, 2004). It is, however, as yet unclear whether FRN amplitude is modulated by the magnitude of the reward prediction error, as implicated by the RL theory.

Many previous studies did not find evidence of such a modulation (e.g. Holroyd *et al.*, 2003; Hajcak *et al.*, 2005, 2006; Cohen *et al.*, 2007). Only two recent reports indicated a significant FRN amplitude modulation by prediction error magnitude, based on modification of a time estimation task and modulation of expectancy by frequency of error feedback (Holroyd & Krigolson, 2007; see Miltner *et al.*, 1997) or a gambling task with the modulation of expectancy on a trial by trial basis (Hajcak *et al.*, 2007).

As Hajcak *et al.* (2007) have pointed out, the inconsistencies in the findings described above may, in part, be related to the lack of assessment of subjects' actual expectations in the earlier studies. The authors gave subjects the opportunity to choose one of four doors to win a prize. Reward expectations of different strengths were induced by varying the number of doors that actually hid a prize (one, two or three). In contrast to their own earlier study (Hajcak *et al.*, 2005), subjects' expectations were assessed by asking on every trial whether they expected to receive reward or not (Hajcak *et al.*, 2007), which led to a modulation of FRN amplitude depending on reward probability.

This modulation was, however, only observed when subjects expressed their expectation after they had chosen one of the four doors (Hajcak *et al.*, 2007; experiment 2). The authors concluded that only expectations that are closely linked to actions are capable of influencing neural processing in the ACC and therefore FRN amplitude. This action-related interpretation of the FRN is in line with the RL theory, which assumes that ACC-related neural processing is used to optimize choice behaviour in terms of maximizing positive performance outcomes. It is further corroborated by the finding that the FRN is much larger when active responding is required (Yeung *et al.*, 2005).

The present study aimed to further determine the conditions that lead to an expectation-relation modulation of FRN amplitude. In contrast to the study by Hajcak *et al.* (2007), subjects were not asked whether they expected to receive reward or not. Rather, subjects could choose between stimuli, which were associated with reward at different probabilities. The advantage of this procedure is that subjects' expectations can be inferred directly from their choice behaviour, providing an expectation measure that is closely linked to actions, as suggested by Hajcak *et al.* (2007).

In line with the findings of Hajcak *et al.* (2007), the present study provides empirical evidence for a relationship between the size of negative prediction errors and FRN amplitude. When subjects chose a

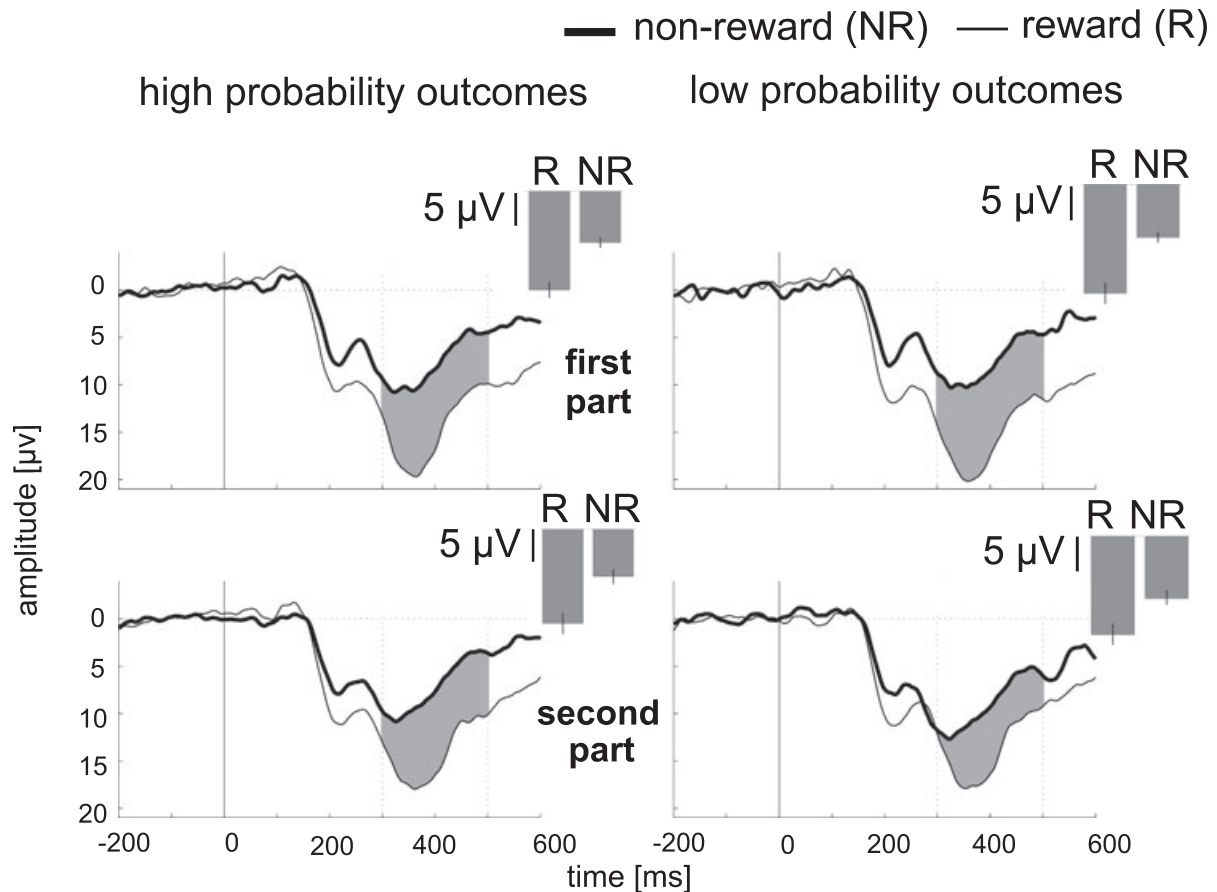


FIG. 8. Parietal ERPs for the first and second part of the experiment in the non-learners (pooled data from electrode positions P3, Pz and P4). On the left, ERPs for highly probable rewarding and non-rewarding outcomes are shown. The right side shows ERPs for low-probability outcomes. The time window for the P300 is shaded in grey. Bars above the ERP traces show average amplitudes for rewarded (R) and non-rewarded (NR) trials in the analysed time window.

stimulus with a reward probability of 1/3, the average – and thus expected – reward value was 1/3 of 5 cent. Although the reward expectation was thus not very high and non-reward was more or less expected, receiving non-reward was still accompanied by a small negative prediction error, which elicited a small, but significant FRN compared with expected reward (rewarding outcome following a 2/3-choice). Non-reward following a 2/3-choice led to a larger prediction error, which was reflected in a significantly larger FRN compared with unexpected reward (reward following a 1/3-choice).

Source analysis confirmed the well-documented finding of an association of FRN with ACC activity (Miltner *et al.*, 1997; Gehring & Willoughby, 2002). Extending previous findings, our data also suggest that the modulatory effect of negative prediction error magnitude on FRN amplitude is related to the ACC. For both smaller and larger negative prediction errors, the main source of the FRN difference between non-reward and reward was located in the ACC. However, for larger prediction errors the source covered a more extensive region, extending into the medial and lateral superior frontal cortex. Additionally, two sources in the temporal lobe emerged. Although source analysis results generally have to be interpreted with caution, the present findings are plausible, at least for the sources in the frontal lobe. Functional neuroimaging studies have provided evidence of pronounced ACC activity related to prediction errors (e.g. Holroyd *et al.*, 2004b). Several areas within the prefrontal cortex have also been linked to reward-related processing, and prediction errors in particular, with some studies showing preferential left hemisphere

involvement (e.g. Ramnani *et al.*, 2004; Dreher *et al.*, 2006). Interestingly, however, a region in the right dorsolateral prefrontal cortex, the location of which is very near to the prefrontal source observed in the present study, has been described to code prediction errors in the context of non-monetary feedback (Fletcher *et al.*, 2001; Turner *et al.*, 2004). In accordance with the present findings, activation in this region appears to decrease when outcomes become more predictable, and the largest responses are observed when predictions are violated (Fletcher *et al.*, 2001).

The role of the FRN in learning

Comparing the pattern of FRN amplitudes before and after subjects had gained insight into reward probabilities on an individual basis permitted the conclusion that the observed modulations of FRN amplitude were induced by learning, as they were not observed in the pre-learning period or in non-learners. It should be noted that the three-way ANOVA including the factor period (i.e. before vs after learning) only yielded a near-significant three-way interaction in the learners. Follow-up ANOVAs did, however, indicate a clear pattern, with a significant interaction between outcome and probability in the post-learning but not in the pre-learning period. A possible reason for the lack of a significant three-way interaction may relate to the definition of pre- and post-learning periods. Shortly before subjects changed their behaviour accordingly, they might have already had some idea concerning the rule underlying reward frequencies, so that

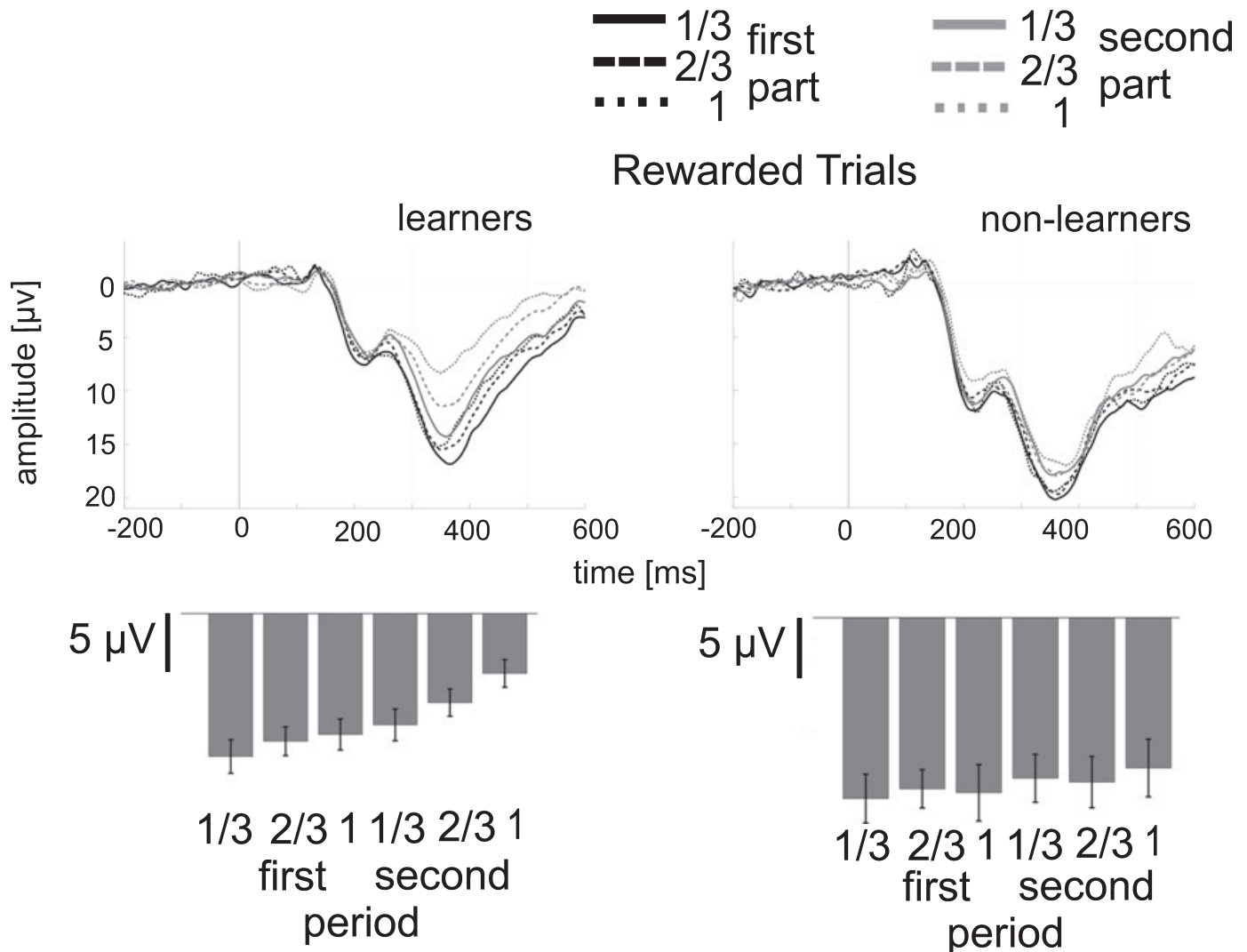


FIG. 9. Parietal ERPs for pre- and post-learning periods in learners and non-learners in response to reward following the choice of 1/3, 2/3 and 1-stimuli (pooled data from electrode positions P3, Pz and P4).

the learning onset trial as defined in the present study might not perfectly reflect the onset of insight.

To date, only a few studies have directly addressed the relationship between the FRN and learning performance. Early studies focused on interactions between the response-locked ERN and the FRN (Holroyd & Coles, 2002; Nieuwenhuis *et al.*, 2002). When responses and outcomes were matched in 100% of the trials, the amplitude of the FRN following negative feedback decreased and the amplitude of the ERN increased as soon as subjects learned the association between responses and outcomes. Negative feedback continued to elicit significant FRNs, when the response did not perfectly predict feedback valence (Holroyd & Coles, 2002; Nieuwenhuis *et al.*, 2002). Interestingly, Nieuwenhuis *et al.* (2002) observed lower FRN amplitudes as well as reduced learning performance in older compared with younger subjects, suggesting that the FRN reflects processes that are critical for learning from feedback. A more detailed analysis by Frank *et al.* (2005) showed that FRN amplitude does not correlate with overall learning performance, but with subjects' ability to learn from negative feedback. Negative learners, i.e. subjects who show a stronger tendency to avoid negative rather than seeking positive events, are also more sensitive to their own errors, as

revealed by larger ERN amplitudes in comparison to positive learners. This may be one reason why they learned better from their errors.

In the present study, modulations of FRN amplitude were induced by negative prediction errors of different magnitudes. It is not surprising that these modulations were only observed in the learners after they had gained insight into reward contingencies, as only learners were able to distinguish between stimuli associated with different reward magnitudes. The observed modulation thus would appear to be a by-product of learning. In our view, however, insight and adequate prediction error processing are prerequisites of further learning. Only the learners developed specific reward predictions, the violations of which were reflected in FRN amplitude modulations. But even after they had learned the rule, they kept using the feedback for: (i) checking whether their assumptions about reward probabilities matched the actual frequencies of reward and non-reward; and (ii) predicting the outcomes with even higher accuracy by looking for further rules. The non-learners, on the other hand, appear to have given up searching for a rule during the course of the experiment. This is reflected in altered processing of negative feedback in the second compared with the first part of the experiment: Only in the first part,

FRNs in response to negative feedback were significant in the non-learners.

Thus, while the FRN in the first stage does not seem to predict learning in the later stage, as both learners and non-learners showed similarly sized amplitudes, it is likely that only learners would have been able to use the feedback in the second stage for further behavioural adaptations. Although this idea was not specifically assessed in the present study, recent data by Cohen & Ranganath (2007) offer support for this hypothesis. They observed significantly larger FRN amplitudes on trials that were followed by a behavioural change (Cohen & Ranganath, 2007).

The P300 and positive outcomes

The P300 wave is typically elicited in the so-called oddball paradigm, when subjects actively attend to an infrequent stimulus in a series of frequent stimuli. Although other factors such as the interstimulus or intertarget intervals also affect P300 parameters (see Gonsalvez & Polich, 2002), P300 amplitude decreases with increasing frequency, i.e. the probability of the target stimulus (Donchin & Coles, 1988; Picton, 1992).

In the context of outcome processing, it has been suggested that the P300 codes reward magnitude independent of valence during reward-related processing (Yeung & Sanfey, 2004; Sato *et al.*, 2005). The results of the present study are, however, in line with the findings of Hajcak *et al.* (2007) who observed larger P300 amplitudes for rewarding compared with non-rewarding stimuli. This effect cannot be caused by the relative frequency of rewarding and non-rewarding stimuli, as it was observed in all conditions and experimental phases, i.e. also when learners received reward in more than 60% of the trials in the post-learning period. There is previous evidence of similar result patterns, although P300 findings were not specifically addressed. Visual inspection of ERPs in two studies by Holroyd and co-workers indicates that P300 amplitudes were enhanced when reward was unexpected (Holroyd *et al.*, 2003) or larger than expected (Holroyd *et al.*, 2004a).

In any case, the P300 appears to be modulated by expectancy (Yeung & Sanfey, 2004; Hajcak *et al.*, 2005). In the learners of the present study, the P300 amplitude difference between rewarded and non-rewarded trials was significantly enhanced for less probable and thus unexpected outcomes, as reflected by a significant interaction between probability and outcome. This interaction can be interpreted in terms of larger surprise by unexpected outcomes. The absence of higher-order interactions suggests that unexpected outcomes elicited larger P300 amplitudes irrespective of the learning stage. This is somewhat surprising, as the term expectancy does only really apply to the post-learning period, after subjects had gained insight into reward probabilities. As already outlined in the discussion of the FRN results (see above), it cannot be excluded that subjects might have had some idea about the rule determining reward probability before they adapted their choice behaviour accordingly, so that the exact onset of learning is difficult to determine. This interpretation is also corroborated by our exploratory analysis based on rewarded trials only (see Results). In the pre-learning period, the P300 amplitude appears to be already modulated by expectation, as the P300 amplitude is larger following rewarded 1/3-choices compared with rewarded 2/3-choices or 1-choices. In the post-learning period, the modulatory effect appeared to be stronger, with a larger amplitude for 2/3- compared with 1-choices.

An interaction between the pre- and post-learning periods on the one hand and performance outcome on the other, which emerged in both learners and non-learners, might be related to an unspecific

adjustment of expectations, i.e. towards the end of the experiment subjects may not have been as surprised by positive outcomes as at the beginning.

Component overlap

Because of potential component overlap, the comparison of ERPs elicited by different experimental conditions is often difficult to interpret, because in many cases differences between conditions cannot unequivocally be ascribed to one particular component (Luck, 2005). For the FRN, several components have been described to play a potentially confounding role. In the time window of the FRN itself, i.e. approximately between 200 and 300 ms following a feedback stimulus, the positive component P2 (or P2a, see Potts *et al.*, 2006) may affect the FRN. Visual inspection of the ERPs in the present study indicates that a positive component appears to precede the FRN. This early difference in potentials following reward and non-reward is often considered as an early manifestation of the FRN being superimposed on a positive going potential (e.g. Yeung & Sanfey, 2004). Even if the positive peak is regarded as a separate potential, the modulatory effect of reward expectancy on ERP amplitude is clearly related to the FRN in our data, as the effect is driven by the FRN peak between 220 and 280 ms following feedback presentation.

Two other components, the N200, which might even reflect the same neural process as the FRN (see Holroyd, 2004), and the P300 may influence FRN amplitude, especially because both components appear to be mediated by factors that also affect the FRN, such as expectation. To minimize possible effects of component overlap, the analysis of the present study contrasted different outcomes in conditions that were matched for probability (i.e. expectancy in the post-learning period). Thus, the FRN or P300 amplitudes differences between the different outcome conditions cannot be accounted for by expectation-related modulations of other components.

Conclusions

Together with other recent reports, the present data provide further evidence of a critical role of the ACC in coding negative reward prediction errors, which are used to optimize behaviour on a given task, as was proposed by Holroyd & Coles (2002). While Hajcak *et al.* (2007) obtained a similar finding, the present study extends the previous results by showing that expectations, which are inferred from actions, are reflected in the amplitude of the FRN. It seems likely that the error signals are received from midbrain dopaminergic neurons, as their firing pattern shows parallels to the FRN amplitude effects of the present study, at least for negative prediction errors. Via connections with motor regions, the ACC is involved in movement generation (Dum & Strick, 1993; van Hoesen *et al.*, 1993; Diehl *et al.*, 2000) and uses the feedback signals to guide behaviour. Consistent with the present findings, the firing pattern of many ACC cells was found to be influenced by reward prediction in the monkey, and temporal inactivation of the ACC leaves monkeys unable to find the optimal performance strategy in a learning task (Amiez *et al.*, 2006).

The modulation of the FRN was only observed when subjects gained insight into reward contingencies, and could thus clearly be related to learning. However, the exact relationship between the FRN on the one hand and learning performance on the other remains to be clarified.

Positive feedback does not appear to affect early processing in the FRN time window, but it might be reflected in the P300, as the present study suggests in accordance with earlier reports. In contrast to the

FRN and negative outcomes, the P300 does, however, not seem to code the size of a reward prediction error.

Acknowledgement

We gratefully acknowledge support by the German Ministry for Education and Research (Bundesministerium für Bildung und Forschung, BMBF, Grant Nr. 01 GW0541).

Abbreviations

ACC, anterior cingulate cortex; BA, Brodmann area; EEG, electroencephalography; ERN, error-related negativity; ERP, event-related potentials; FRN, feedback-related negativity; LORETA, low-resolution brain electromagnetic tomography; RL, reinforcement learning.

References

- Amiez, C., Joseph, J.P. & Procyk, E. (2006) Reward encoding in the monkey anterior cingulate cortex. *Cereb. Cortex*, **16**, 1040–1055.
- Cohen, M.X., Elger, C.E. & Ranganath, C. (2007) Reward expectation modulates feedback-related negativity and EEG spectra. *Neuroimage*, **35**, 968–978.
- Cohen, M.X. & Ranganath, C. (2007) Reinforcement learning signals predict future decisions. *J. Neurosci.*, **27**, 371–378.
- Diehl, B., Dinner, D.S., Mohamed, A., Najm, I., Klem, G., LaPresto, E., Bingaman, W. & Luders, H.O. (2000) Evidence of cingulate motor representation in humans. *Neurology*, **55**, 725–728.
- Donchin, E. & Coles, M.G. (1988) Is the P300 component a manifestation of context updating? *Behav. Brain Sci.*, **11**, 355–372.
- Dreher, J.C., Kohn, P. & Berman, K.F. (2006) Neural coding of distinct statistical properties of reward information in humans. *Cereb. Cortex*, **16**, 561–573.
- Dum, R.P., Strick & P.L. (1993) Cingulate motor areas. In Vogt, B.A. & Gabriel, M. (Eds), *Neurobiology of Cingulate Cortex and Limbic Thalamus: A Comprehensive Review*. Birkhauser, Boston, pp. 414–441.
- Falkenstein, M., Hohnsbein, J., Hoormann, J. & Blanke, L. (1990) Effects of errors in choice reaction tasks on the ERP under focused and divided attention. In Brunia, C.H.M., Gaillard, A.W.K. & Kok, A. (Eds), *Psychophysiological Brain Research*. Tilburg University Press, Tilburg, The Netherlands, pp. 192–195.
- Fletcher, P.C., Anderson, J.M., Shanks, D.R., Honey, R., Carpenter, T.A., Donovan, T., Papadakis, N. & Bullmore, E.T. (2001) Responses of human frontal cortex to surprising events are predicted by formal associative learning theory. *Nat. Neurosci.*, **4**, 1043–1048.
- Frank, M.J., Worocho, B.S. & Curran, T. (2005) Error-related negativity predicts reinforcement learning and conflict biases. *Neuron*, **47**, 495–501.
- Gehring, W.J., Goss, B., Coles, M.G.H., Meyer, D.E. & Donchin, E. (1993) A neural system for error detection and compensation. *Psychol. Sci.*, **4**, 385–390.
- Gehring, W.J. & Willoughby, A.R. (2002) The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, **295**, 2279–2282.
- Gonsalvez, C.L. & Polich, J. (2002) P300 amplitude is determined by target-to-target interval. *Psychophysiology*, **39**, 388–396.
- Hajcak, G., Holroyd, C.B., Moser, J.S. & Simons, R.F. (2005) Brain potentials associated with expected and unexpected good and bad outcomes. *Psychophysiology*, **42**, 161–170.
- 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. *Biol. Psychol.*, **71**, 148–154.
- 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.
- van Hoesen, G.W., Morecraft, R.J. & Vogt, B.A. (1993) Connections of the monkey cingulate cortex. In Vogt, B.A. & Gabriel, M. (Eds), *Neurobiology of Cingulate Cortex and Limbic Thalamus: A Comprehensive Review*. Birkhauser, Boston, pp. 414–441.
- Holroyd, C.B. (2004) A note on the oddball N200 and the feedback ERN. In Ullsperger, M. & Falkenstein, M. (Eds), *Errors, Conflicts and the Brain: Current Opinions in Response Monitoring*. MPI for Cognitive Neuroscience, Leipzig.
- Holroyd, C.B. & Coles, M.G. (2002) The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.*, **109**, 679–709.
- Holroyd, C.B. & Krigolson, O.E. (2007) Reward prediction error signals associated with a modified time estimation task. *Psychophysiology*, **44**, 913–917.
- Holroyd, C.B., Larsen, J.T. & Cohen, J.D. (2004a) Context dependence of the event-related brain potential associated with reward and punishment. *Psychophysiology*, **41**, 245–253.
- Holroyd, C.B., Nieuwenhuis, S., Yeung, N., Nystrom, L., Mars, R.B., Coles, M.G. & Cohen, J.D. (2004b) Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nat. Neurosci.*, **7**, 497–498.
- Holroyd, C.B., Nieuwenhuis, S., Yeung, N. & Cohen, J.D. (2003) Errors in reward prediction are reflected in the event-related brain potential. *Neuroreport*, **14**, 2481–2484.
- Ito, S., Stuphorn, V., Brown, J.W. & Schall, J.D. (2003) Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science*, **302**, 120–122.
- Knutson, B., Westdorp, A., Kaiser, E. & Hommer, D. (2000) fMRI visualization of brain activity during a monetary incentive delay task. *Neuroimage*, **12**, 20–27.
- Lee, T.W., Girolami, M. & Sejnowski, T.J. (1999) Independent component analysis using an extended infomax algorithm for mixed subgaussian and supergaussian sources. *Neural Comput.*, **11**, 417–441.
- Luck, S.J. (2005) *An Introduction to the Event-Related Potential Technique*. The MIT Press, Cambridge, MA.
- Miltner, W.H.R., Braun, C.H. & Coles, M.G.H. (1997) Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a generic neural system for error detection. *J. Cogn. Neurosci.*, **9**, 788–798.
- Nieuwenhuis, S., Ridderinkhof, K.R., Talsma, D., Coles, M.G., Holroyd, C.B., Kok, A. & van der Molen, M.W. (2002) A computational account of altered error processing in older age: dopamine and the error-related negativity. *Cogn. Affect. Behav. Neurosci.*, **2**, 19–36.
- Pascual-Marqui, R.D., Michel, C.M. & Lehmann, D. (1994) Low resolution electromagnetic tomography: a new method for localizing electrical activity in the brain. *Int. J. Psychophysiol.*, **18**, 49–65.
- Picton, T.W. (1992) The P300 wave of the human event-related potential. *J. Clin. Neurophysiol.*, **9**, 456–479.
- Potts, G.F., Martin, L.E., Burton, P. & Montague, P.R. (2006) When things are better or worse than expected: the medial frontal cortex and the allocation of processing resources. *J. Cogn. Neurosci.*, **18**, 1112–1119.
- Ramnani, N., Elliott, R., Athwal, B.S. & Passingham, R.E. (2004) Prediction error for free monetary reward in the human prefrontal cortex. *Neuroimage*, **23**, 777–786.
- Sato, A., Yasuda, A., Ohira, H., Miyawaki, K., Nishikawa, M., Kumano, H. & Kuboki, T. (2005) Effects of value and reward magnitude on feedback negativity and P300. *Neuroreport*, **16**, 407–411.
- Schultz, W. (1999) The reward signal of midbrain dopamine neurons. *News Physiol. Sci.*, **14**, 249–255.
- Schultz, W. (2001) Reward signaling by dopamine neurons. *Neuroscientist*, **7**, 293–302.
- Shima, K. & Tanji, J. (1998) Role for cingulate motor area cells in voluntary movement selection based on reward. *Science*, **282**, 1335–1338.
- Talairach, J. & Tournoux, P. (1988) *Co-planar Stereotaxic Atlas of the Human Brain 3-Dimensional Proportion System: an Approach to Medical Cerebral Imaging*. G. Thieme, Stuttgart.
- Turner, D.C., Aitken, M.R., Shanks, D.R., Sahakian, B.J., Robbins, T.W., Schwarzbauer, C. & Fletcher, P.C. (2004) The role of the lateral frontal cortex in causal associative learning: exploring preventative and super-learning. *Cereb. Cortex*, **14**, 872–880.
- Weisberg, J., van Turennout, M. & Martin, A. (2007) A neural system for learning about object function. *Cereb. Cortex*, **17**, 513–521.
- Wills, A.J., Lavric, A., Croft, G.S. & Hodgson, T.L. (2007) Predictive learning, prediction errors, and attention: evidence from event-related potentials and eye tracking. *J. Cogn. Neurosci.*, **19**, 843–854.
- Yasuda, A., Sato, A., Miyawaki, K., Kumano, H. & Kuboki, T. (2004) Error-related negativity reflects detection of negative reward prediction error. *Neuroreport*, **15**, 2561–2565.
- Yeung, N., Holroyd, C.B. & Cohen, J.D. (2005) ERP correlates of feedback and reward processing in the presence and absence of response choice. *Cereb. Cortex*, **15**, 535–544.
- Yeung, N. & Sanfey, A.G. (2004) Independent coding of reward magnitude and valence in the human brain. *J. Neurosci.*, **24**, 6258–6264.