Psychophysiology, 50 (2013), 931–939. Wiley Periodicals, Inc. Printed in the USA.

DOI: 10.1111/psyp.12073

Learning-induced modulations of the stimulus-preceding negativity

JOAQUÍN MORÍS,[a,b] DAVID LUQUE,[a,b] AND ANTONI RODRÍGUEZ-FORNELLS[c,d,e]

[a]Department of Basic Psychology, University of Málaga, Málaga, Spain
[b]Institute of Biomedical Research of Málaga (IBIMA), Málaga, Spain
[c]Cognition and Brain Plasticity Group (Bellvitge Biomedical Research Institute), IDIBELL, L'Hospitalet de Llobregat, University of Barcelona, Barcelona, Spain
[d]Department of Basic Psychology, University of Barcelona, L'Hospitalet de Llobregat, Barcelona, Spain
[e]Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

Abstract

The neural basis of feedback expectation, which is crucial in learning theory, has only been minimally studied. Stimulus-preceding negativity (SPN), an ERP component that appears prior to the presentation of feedback, has been proposed as being related to feedback expectation. The present study showed, for the first time, amplitude modulations of the SPN component during learning acquisition in a trial-by-trial associative learning task. The results indicate that SPN could be a plausible electrophysiological index of the cognitive processes engaged while expecting the appearance of relevant feedback during reinforcement learning.

Descriptors: Associative theory, Expectation, Learning, Reinforcement, Stimulus-preceding negativity

In reinforcement learning, an informative stimulus or feedback indicates the consequence of a previous response and allows for the adjustment of future responses to maximize rewards and minimize punishments. The cognitive and neurophysiological models of reinforcement learning assume that this process takes place using an error computation analysis in which the difference between the expectation of an outcome and the actual outcome is computed (Holroyd & Coles, 2002; Rescorla & Wagner, 1972; Schultz, 2002). The neural bases of feedback and error detection and correction processing have received increasing attention during the last decade. For example, Schultz and colleagues (among others) have shown that the mesencephalic dopamine system is a good candidate for the neural implementation of error computation in the brain (Waelti, Dickinson, & Schultz, 2001). Holroyd and Coles (2002) later proposed their influential hypothesis, which suggests that feedback-related negativity (FRN), a frontocentral event-related potential (ERP) negativity that appears approximately 250 ms after the presentation of negative or worse than expected feedback, might index this error computation process.

The anterior cingulate cortex (ACC) appears to be one of the possible neural sources involved in the elicitation of FRN (Müller, Möller, Rodríguez-Fornells, & Münte, 2005). Holroyd and Coles proposed that FRN might be elicited because of an error signal conveyed by midbrain dopamine neurons to the ACC (for a review, see Càmara, Rodríguez-Fornells, & Münte, 2009). Among the evidence that supports the error computation account of FRN is the fact that FRN changes across learning, with larger FRN amplitudes at the beginning of learning than after extended training (Cohen, Elger, & Ranganath, 2007; Eppinger, Kray, Mock, & Mecklinger, 2008; Holroyd & Coles, 2002; Luque, López, Marco-Pallares, Càmara, & Rodríguez-Fornells, 2012; Müller et al., 2005; Walsh & Anderson, 2011). These results are in line with the predictions from the Holroyd and Coles model. This model establishes that participants acquire the actual response-outcome mapping during learning and, from that initial moment, are able to monitor their behavior using this information even in the absence of an external tutor (feedback).

Surprisingly, and to the best of our knowledge, no previous ERP studies have been conducted to investigate how expectation evolves during learning. The most likely ERP index of expectation during reinforcement learning is the stimulus-preceding negativity (SPN) (for a review, see Brunia, Hackley, van Boxtel, Kotani, & Ohgami, 2011). The SPN is a slow negative component with frontal distribution that progressively increases in amplitude prior to the presentation of informative feedback or motivationally relevant stimuli (e.g., monetary rewards, performance feedback, evocative photos, or painful affective salient stimulus; Brown, Seymour, Boyle, El-Deredy, & Jones, 2008; Donkers, Nieuwenhuis, & van Boxtel, 2005; Fuentemilla et al., 2013; Kotani, Hiraku, Suda, & Aihara, 2001; Kotani et al., 2003; Masaki, Takeuchi, Gehring, Takasawa, & Yamazaki, 2006; Masaki, Yamazaki, & Hackley, 2010; Ohgami, Kotani, Hiraku, Aihara, & Ishii, 2004; Ohgami et al., 2006; van Boxtel & Böcker, 2004). Although the amplitude of the SPN appears to be larger in the right prefrontal and precentral regions (Böcker, Baas, Kenemans, & Verbaten, 2001; Kotani et al., 2001; Brown et al., 2008; Ohgami et al., 2004), different topographies have been identified (Brunia et al., 2011; van Boxtel & Böcker, 2004). The neural sources of SPN have been attributed to the anterior insular cortex among other regions (Brunia, De Jong, van

This research was supported by grants P11-SEJ-7898 and P.08-SEJ-03586 SEJ from Junta de Andalucía and 2007-63691/PSIC and PSI2008-03901 from the Spanish Ministry of Science and Innovation.

Address correspondence to: Dr. Joaquín Morís, Departamento de Psicología Básica, Facultad de Psicología, Universidad de Málaga, Campus de Teatinos, s/n (29071), Málaga, Spain. E-mail: jmoris@uma.es

den Berg-Lenssen, & Paans, 2000). Brunia et al. (2011) have recently proposed that the SPN provides an index of neocortical processes underlying expectation of the outcome. Consistent with this proposal, functional neuroimaging studies in humans have shown that the anterior insular cortex is activated in the generation of prefeedback expectation, in comparison with a no-feedback condition (Kotani et al., 2009; Tsukamoto et al., 2006). Also, SPN has been related to the dopaminergic neurotransmitter system. This component is severely affected in patients in whom the dopaminergic system is severely compromised, such as patients with Parkinson's disease (Mattox, Valle-Inclán, & Hackley, 2006). Previous studies have shown that SPN can reflect not only that an outcome is expected, but also its motivational aspects (e.g., Donkers et al., 2005), and the implication of the participants in producing the outcome (Masaki et al., 2010), but no experiment has been reported that studies specifically how learning can lead to the creation of this expectation and if SPN is modulated accordingly.

A common aspect between the SPN and FRN components is that both have been associated with reward processing (feedback expectancy vs. evaluation of feedback) and dopaminergic midbrain regulatory mechanisms. However, few experiments have studied both components simultaneously (however, see Donkers et al., 2005; Masaki et al., 2006; 2010). Considering the proposals of Holroyd and Coles (2002) and Brunia et al. (2011) together, the SPN might index the outcome expectation in the prefeedback period, and FRN might index the difference between this expectation and the actual outcome after the presentation of the feedback. A possible prediction of this hypothesis is that prefeedback SPN should change throughout learning as in the case of the FRN.

There is not a clear prediction about the direction of the hypothesized relation between the prefeedback SPN and the amount of learning, though. If in a reinforcement learning task the SPN is a direct correlate of feedback expectancy, it should increase its magnitude as learning progresses. Learning would lead to an improved performance, obtaining more positive feedbacks and to an increase in the expectancy of obtaining positive feedbacks, what in turn would be reflected in the SPN as an increase in its amplitude. This prediction would be consistent with the interpretation of the SPN as an anticipatory index underlying the expectation of the consequences of our actions (see Brunia et al., 2011). On the other hand, previous experiments have also shown that SPN can be correlated to the amount of information that a feedback provides. On the experiments described by Kotani et al. (2003), when participants completed trials that they knew would be followed by very informative feedbacks, the amplitude of SPN encountered showed a larger amplitude than in the case of less informative feedbacks. Similarly, a recent study showed that the amplitude of the SPN increased with very unexpected rewarding feedbacks (Fuentemilla et al., 2013). In the task used in the present study, at the beginning of learning, feedbacks might be very informative as they provide information about the relationships that are unknown to the participants. However, while learning progresses, feedback information decreases as participants might have already created a better representation of the associations involved in the task and they would have little need to update the knowledge acquired. According to this interpretation, SPN's amplitude would be expected to decrease as learning is progressively acquired, increasing as well its predictability. This pattern of results would be also consistent with results such as those of Chwilla and Brunia (1991) and of Foti and Hajcak, (2012), which found SPN of higher magnitude preceding true feedbacks that informed participants about their performance than preceding false feedbacks that did not provide any information. The aim of the present study is to explore these possibilities and to test if the magnitude of SPN is related to the amount of learning.

Method

The data analyzed in the present study are from the learning phase of an associative learning task reported in Luque et al. (2012). Additional information regarding participants, stimuli, task procedure, experimental design, and data recording are presented there, while these experimental details will be briefly summarized here.

Participants

Twenty-four right-handed young adults (seven males, mean age = 22, *SD* = 4.1, range = 19–38) participated in the experiment to completion. The participants did not report any health problems or previous neurological or psychiatric disorders when they were asked before the experiment. They received a financial compensation after completing the experiment.

Stimuli and Task

We used the allergy learning task in which participants learned in a trial-by-trial scheme which objects, among several, would cause an allergy to an imaginary patient. A total of 180 pictures of objects were used as potentially allergic stimuli. These objects belonged to one of the following six categories: fruits, foods (excluding fruits), clothes, office supplies, animals, and toys. The experiment was divided into 30 blocks, with 5 blocks per stimulus category. The experimental design was repeated across the different blocks, and a new set of stimuli was employed in each of these blocks.

Participants were asked to play the role of a fictitious allergist. At the beginning of each trial, they saw one object or a pair of objects presented as predictive information (cues). Supposedly the patient had been exposed to these objects. Using this information, they had to predict whether or not the mock patient in this situation would develop an allergy (outcomes). After responding, they saw if the patient had actually developed the allergy or not. Repeated experience with the cues and their associated outcomes allowed a progressive learning of the relations involved in that block. To study this progressive acquisition of learning and its electrophysiological correlates, the training trials of each independent block were divided into four bins. The initial bin, Bin 1, comprised the initial 25% of trials of each trial type of each block. Bin 2 contained the next 25%, and similarly for Bins 3 and 4. As training progressed in each block, participants were expected to have higher learning levels in each bin than in previous ones.

Each trial began with the presentation of a fixation point (a cross) for 1 s, followed by a cue combination in which one or two colored pictures were presented for 1.2 s. Subsequently, participants had to decide between two response options, "allergy" or "no allergy." To record these responses, participants had a mouse in their right hand and another in their left hand. For both hands, they pressed the left button of the mouse with the index finger. The outcome assignation (allergy or no allergy) to the participant's responses (right or left) was pseudorandom. On half of the blocks, the outcome allergy response was associated with the left hand response, on the other half with the right hand response. The order of the blocks was generated randomly for each participant. After pressing a button, the screen went blank for 1 s, and then feedback was presented for 1 s. Thus, the structure of a trial consisted of the sequential presentation of (a) a fixation point—1 s, (b) the cue(s)—

1.2 s, (c) the response screen (the duration was up to the participant's time for responding), (d) a blank screen—1 s, and (e) the feedback—1 s. Positive feedback was indicated by an iconic smiling face, and negative feedback was indicated by an iconic sad face.

Because the original study was focused on the effect of a cue-interaction manipulation over the FRN component, there were trials with only one predictive cue and others with two predictive cues (see, e.g., Dickinson & Burke, 1996). In single-cue trials, the cue appeared in the center of the screen. In compound-cue trials, the two cues appeared side by side in the center of the screen (they were separated from each other by 2 cm). In these trials, the position of each cue on the screen (right or left) was randomized.

As is usual in cue-interaction experiments, there were two phases: a training phase and a test phase (e.g., Dickinson & Burke, 1996; Luque, Flores, & Vadillo, 2013). Only the training phase data were relevant and analyzed for this study. It is important to note that fine-grained details about the cue-interaction design used in Luque et al. (2012) are not relevant for the aims of the present study. We used the information from Luque et al.'s training phase because participants learned cue-outcome associations in a trial-by-trial fashion, and these data are well suited to measure prefeedback expectation. Thus, the specific design is not relevant. In fact, as the reader will see below, for the present analyses the trials of all types of associations were collapsed.

During the learning phase, there were five different types of association to be learned in each block, and each trial type was associated with a given outcome. Three of these types of associations were associated with the outcome allergy (and, hence, the allergy responses increased with trial-by-trial learning). Among these three allergy associations, two were compound-cued, as they had two cues presented on each trial (e.g., to eat pineapple and cream [the cues] would cause allergy in the patient) while one was single-cued (e.g., to eat apple [the cue] would cause allergy in the patient). Each one of these allergy associations was presented 12 times within each block of learning. There were also two trial types that were associated with the outcome no-allergy. One of these trial types was one single-cued and was presented 12 times per block. The other was compound-cued and was presented 24 times per block. This asymmetry in the frequency of the associations was introduced because in the original study we needed a filler association to equal the number of allergy and no-allergy trials and also to equal the number of compound- and single-cued trials. These relations were probabilistic. Each trial type was associated with their corresponding outcome in 83% of the trials. Thus, the total number of learning trials in each block was 72, from which 60 were consistent cue(s) → outcome trials and 12 were "noise" trials, that is, inconsistent trials. The specific order of presentation of each trial into the block was pseudorandomized preventing more than two consecutive presentations of the same trial.

The roles of the cues were pseudorandomly assigned to the different stimuli with the constraints that all objects must be from the same category in each block and none of the stimuli could be repeated. The order of block presentation was counterbalanced across all participants. Because there were 30 blocks, and each block contained 72 learning trials, the whole experiment included a total of 2,160 learning trials.

Data Acquisition

Electroencephalogram (EEG) data (250 Hz sampling rate, 0.01–50 Hz band-pass filtering, half-amplitude cutoffs) were recorded using 29 scalp electrodes placed at standard positions (10-20 system) (electrode positions: Fp1/2, F3/4, C3/4, P3/4, O1/2, F7/8, T3/4, T5/6, FC1/2, FC5/6, CP1/2, CP5/6, PO1/2, Fz, Cz, Pz). Data were rereferenced offline to linked mastoids. An electrode placed on the outer canthus of the right eye was used to record the horizontal electrooculogram (EOG), while for the vertical EOG another one was placed approximately 2 cm below the right eye and centered under the pupil. All data were screened for artifacts, and those trials in which the base-to-peak amplitude was higher than $\pm 100\,\mu V$ were removed from the analyses (11.2% of the trials). For the figure presented, the data were filtered using a 7 Hz low-pass Butterworth filter, with a roll-off of 12 dB/oct, as implemented in the ERPLAB toolbox V3.0 (Luck & López-Calderon, 2012).

Event-Related Potential Data Analysis

The training trials of all the blocks were divided into four bins for the ERP analysis. Each bin was composed of a quartile of the training trials. More specifically, the first 18 trials of every block were assigned to the first bin, Bin 1, the following 18 trials to Bin 2, and similarly for Bins 3 and 4. By doing this, 25% of the trials of each trial type of the block were assigned to each bin. All the trials of the training phase were assigned this way to a bin, including the filler association. For the first analysis, those trials within each bin with correct responses were selected. The mean number of correct response trials without artifacts per bin was, respectively, 316.43 (*SEM* = 9.35), 369.38 (*SEM* = 13.87), 387.05 (*SEM* = 15.35), and 409.48 (*SEM* = 15.57). Grand averages were calculated using the EEG signal of the trials in each bin in which participants provided the correct response. Following standard learning models (e.g., Rescorla & Wagner, 1972), participants should expect to obtain a positive feedback in these trials, and this expectation should be stronger as learning advances. EEG epochs were calculated using the feedback onset as reference. The interval between 1,200 to 1,000 ms prior to the feedback served as baseline. SPN occurs before the feedback, and its value is maximal right before it (Brunia et al., 2011); therefore, the mean voltages during three time windows (−600 to −400 ms, −400 to −200 ms, −200 to 0 ms) were calculated as SPN indices.

In a second analysis, the trials with correct and incorrect responses were compared. In this case, the mean voltages were calculated as described before, but including also those trials in which participants gave incorrect responses. It was expected that, as training advanced, the difference between expectancy in correct and incorrect trials should increase, as participants learned the relations between cues and responses. We tested if this same pattern took place in the SPN. The mean number of trials with incorrect responses and without artifacts for each bin was, respectively, 229.5 (*SEM* = 7.43), 180.77 (*SEM* = 8.8), 164.53 (*SEM* = 9.44), and 141.69 (*SEM* = 9.73). Greenhouse-Geisser epsilon correction was applied whenever the sphericity assumption was not met.

Results

Behavioral Data

There was a significant learning effect, and the percentage of trials with a correct response increased on each bin, $F(1.484,29.671) = 65.169$, $p < .001$, $\eta_p^2 = 0.765$, showing a significant linear trend, $F(1,20) = 88.218$, $p < .001$, $\eta_p^2 = 0.815$. The mean percentages (and *SEM*) for Bins 1 to 4 were, respectively, 59.2%

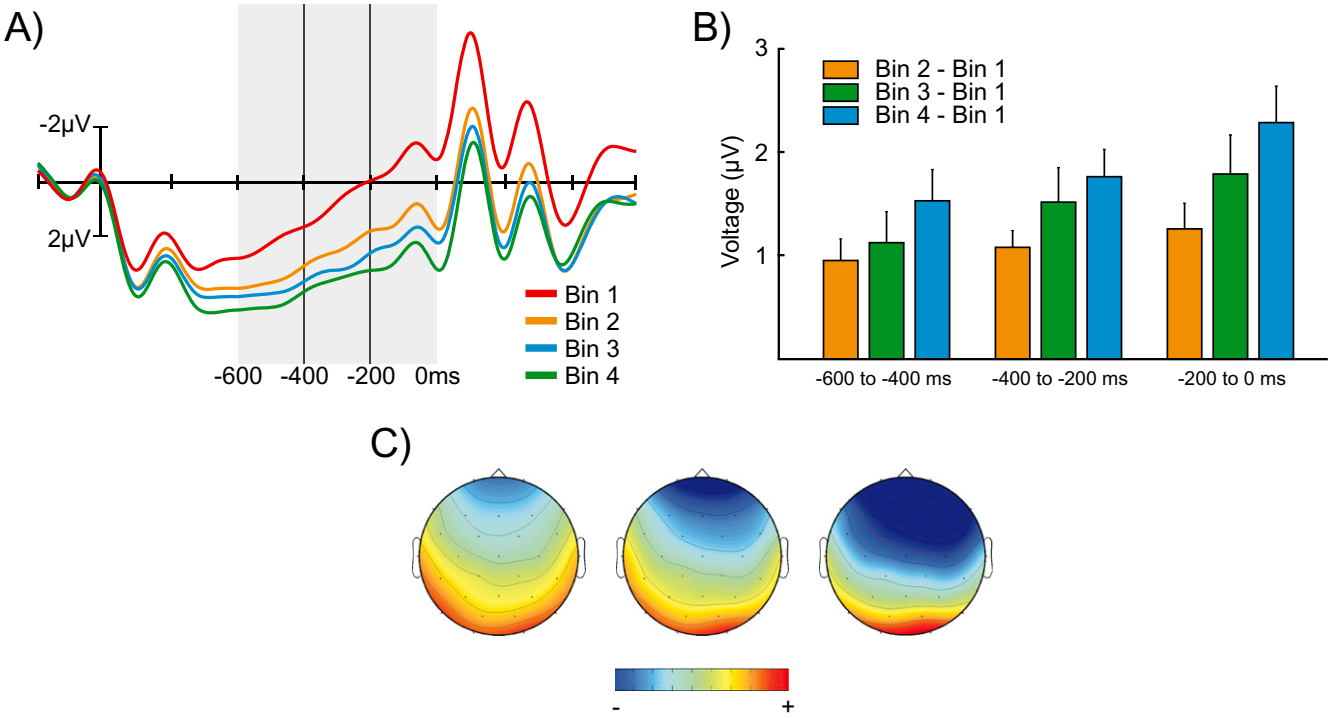

Figure 1. A: Response-locked grand average at the Fz location showing the SPN component and the changes in its amplitude during learning. Feedback was presented at 0 ms. A slow increasing negativity appeared before the presentation of this feedback. This negativity decreased as learning progressed. B: The changes in the SPN component amplitude throughout learning in the three time windows described before. The bars represent the difference between the mean voltages of Bin 1 and Bins 2, 3, and 4 in each of the three time windows for all the nine electrodes analyzed. C: The topographical maps of the difference waveforms between Bin 4 and Bin 1 for the same time windows are represented (the maximum and minimum values for each map were 0 and –3 μV, respectively).

(1.7), 68.8% (2.49), 72% (2.67), and 76.5% (2.9). Post hoc analyses showed that there were significant differences between all the comparisons ($p < .001$ for all tests).

Electrophysiological Data

Figure 1A presents the grand averages for the electrode Fz. In this figure, the initial trigger is the onset of the response of the participant, which is followed by the presentation of the feedback 1,000 ms after the response. During this foreperiod, a development of the SPN component is observed with drastic amplitude changes across the different learning bins.

A repeated measures analysis of variance (ANOVA) was carried out using the mean voltages of a subset of nine right, central, and left electrodes of frontal, central, and parietal positions (F3, Fz, F4, C3, Cz, C4, P4, Pz, and P4, see Figure 2). It had four factors: learning (4 levels, Bin 1 to Bin 4), time window (3 levels, –600 to –400 ms, –400 to –200 ms, and –200 to 0 ms), laterality (3 levels, left, central, right), and electrode (3 levels, frontal, central, and parietal). Post hoc *t* tests were carried out to disentangle the significant effects found.

There was a significant effect of learning, $F(1.498,32.966) = 28.298$, $p < .001$, $\eta_p^2 = 0.563$, and a significant linear trend, $F(1,23) = 38.577$, $p < .001$, $\eta_p^2 = 0.637$. There were significant differences between all the levels of the factor, ($p < .03$ in all the comparisons), and each one was more positive than the previous. The time window factor was also significant, $F(1.198,26.353) = 38.679$, $p < .001$, $\eta_p^2 = 0.637$, with significant differences between all levels of the factor, ($p < .001$ in all the comparisons).

There was a significant interaction between both factors, $F(3.492,76.831) = 8.616$, $p < .001$, $\eta_p^2 = 0.281$, that was best fitted with a linear–linear model, $F(1,23 = 23.044)$, $p < .001$, $\eta_p^2 = 0.512$. Figure 1B shows that the difference between the learning bins increased as the time window of analysis was closer to the feedback, which explains the interaction and trend between the two factors.

There were significant laterality differences, $F(1.634,35.939) = 22.704$, $p < .001$, $\eta_p^2 = 0.508$, between left and central and left and right ($p < .001$ for both) but not between right and central ($p = .092$) electrodes. Electrode factor showed significant differences, $F(1.197,26.33) = 16.211$, $p < .001$, $\eta_p^2 = 0.424$, with significant differences between the three levels of the factor ($p < .022$ in all the comparisons).

The Learning × Time Window × Electrode interaction was significant, $F(4.063,89.38 = 4.271)$, $p = .003$, $\eta_p^2 = 0.163$, and the Learning × Time Window × Laterality × Electrode approached significance, $F(7.828,172.223) = 1.905$, $p = .064$, $\eta_p^2 = 0.08$. Given that these two interactions are the most relevant, they were further explored. Additional analysis showed that the learning and Learning × Time Window effects were significant in all the single electrodes ($F > 2.924$, $p < .02$ for the nine electrodes). Also, the best model to explain the Learning × Time Window × Laterality × Electrode was a linear–linear–linear–linear model, $F(1,23) = 6.45$, $p = .019$, $\eta_p^2 = 0.227$. According to this model, the mean amplitude would decrease over bins, increase over time windows, and this

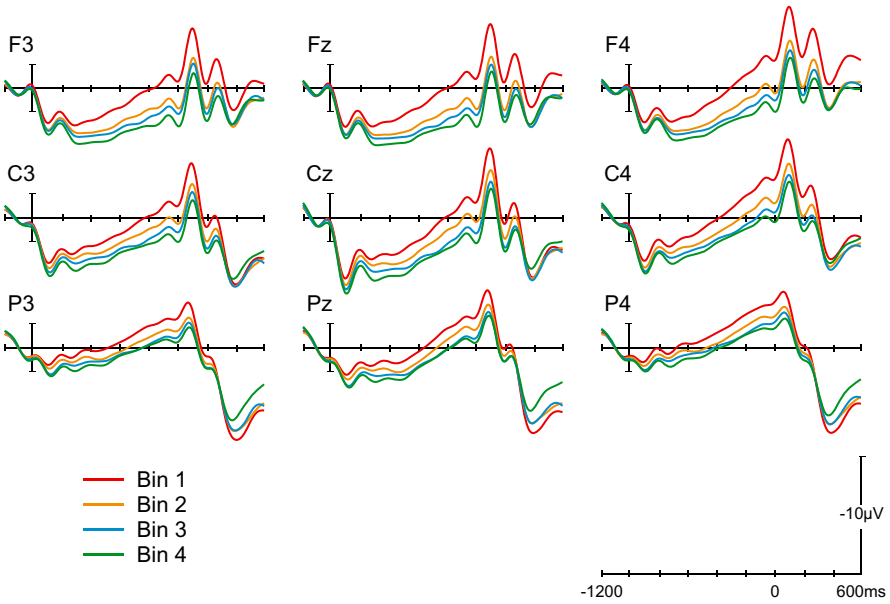

Figure 2. Grand-average waveforms showing the SPN modulations in amplitude during learning at frontal (F3, Fz, F4), central (C3, Cz, C4), and parietal locations (P4, Pz, P4) in the four learning bins. Feedback stimulus is presented at 0 ms.

effect would be maximum in anterior electrodes (decreasing in more posterior electrodes), and in right electrodes (being less pronounced in central and left electrodes). It is interesting to note that, although the more negative voltages in the time windows analyzed seem to be located in the parietal area (see Figure 2), the maximum learning effect, as described, was not seen in these electrodes. The topographies shown in Figure 1C support this analysis, as the maximum difference between Bin 1 and Bin 4 in the subset of electrodes analyzed was found close to F4.

To test the robustness of these effects, equivalent analyses were carried out using different baselines placed after the feedback or after the response but before the feedback. The results were the same, finding a significant learning bin effect with a linear trend in all of them (see Table 1 for a summary).

Table 1. *Additional Analyses with Different Baselines*

	Prefeedback		Postfeedback
	−850 to −750	−600 to −400	500 to 600
F	23.654	12.575	41.543
gl1, gl2	1.95, 42.92	2.131, 46.876	2.08, 45.78
p	< .001	< .001	< .001
η_p^2	0.518	0.364	0.654
Linear trend	$p < .001$	$p < .001$	$p < .001$

Note. Three additional repeated measures ANOVAs were calculated using different baselines (–850 to –750 and –600 to –400 ms previous to the presentation of feedback but after the response, and 500 to 600 ms after feedback presentation). In each of them, the mean voltages for the correct trials of each learning bin of 9 electrodes, 3 frontal (F3, Fz, and F4), 3 central (C3, Cz, and C4), and 3 parietal (P3, Pz, and P4) were used. The critical comparisons (effect of learning bin, and its linear trend) are reported. In the three cases, the learning bin factor was significant, and a linear trend was the model that best fitted the results, indicating that SPN amplitude decreased as learning progressed. The values of the second row indicate the freedom degrees of the factor learning bin (gl1) and of the error term (gl2).

Given that the SPN obtained was found following a motor response, the potential contribution of motor-related potentials to the SPN observed was analyzed. The lateral readiness potential (LRP; e.g., Hughes, & Yeung, 2011; Rodríguez-Fornells, Kurzbuch, & Münte, 2002) was calculated for the trials with correct responses of each bin. Given that in each bin exactly half of the trials had a correct response option associated with each hand, it was possible to calculate the LRP using the voltage at C4 and C3 electrodes. For each participant, the difference between the voltage of the central electrodes contralateral and ipsilateral to the hand with which the response was given on that trial (i.e., C4–C3 in the trials with left hand responses and C3–C4 for the trials with a right hand response) was calculated for each of the trials on each bin with a correct response. Then, these voltages were averaged for each bin. In the time window from −200 to 0 ms, where the SPN learning effect was maximum, there were no statistically significant differences between bins in the magnitude of the LRP, $F(2.366,47.313) = 1.535$, $p = .223$. Also, there were not significant correlations between the magnitude of the SPN and the LRP in any of the learning bins ($p > .185$ in all of them). Also, in the SPN analysis, in the time windows that were further from the response, the learning effect was greater than in the time windows closer to the response. This Time Window × Learning interaction and the absence of a significant LRP effect are contrary to what would be expected if the SPN effect found was caused by motor-related activity.

A second analysis was run to test the differences between trials with incorrect and correct responses. The topographical distribution of these differences is represented in Figure 3B. They were maximal in the frontal electrodes (F3, Fz, and F4). Because of this, a repeated measures ANOVA was run using the mean values of the frontal electrodes (F3, Fz, and F4) in the time window from −200 to 0 ms, with three factors, learning bin, response (correct or incorrect), and laterality (left, central or right). There was a significant response effect, $F(1,23) = 32,670$, $p < .001$, $\eta_p^2 = 0.598$. The amplitude of SPN in incorrect trials was more negative than in

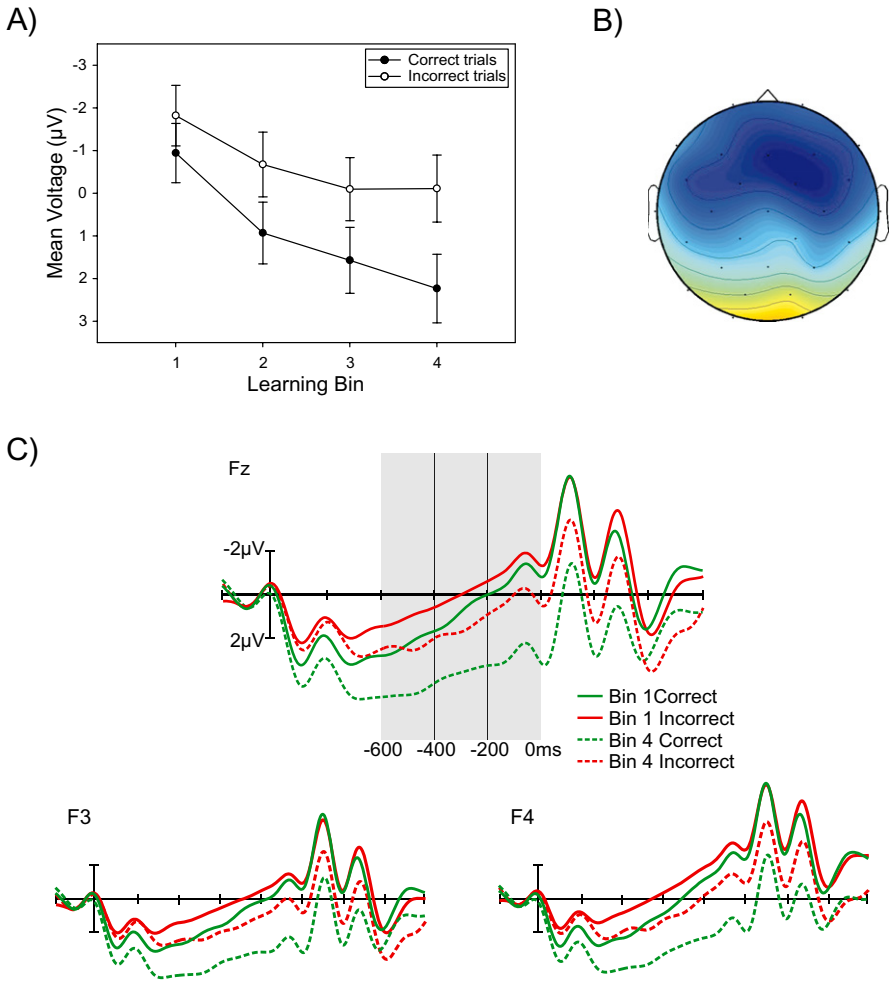

Figure 3. A: Mean differences in SPN amplitude (−200 to 0 ms time window) between correct and incorrect trials in each learning bin for frontal electrodes (F3, Fz, and F4), where the effect was maximum. B: Topographic distribution of the difference between correct and incorrect trials (Bin 1 minus Bin 4) in the time window from −200 to 0 ms (maximum value 1 μV, minimum value −1.6 μV). C: Comparison of grand-average waveforms for electrodes Fz, F3, and F4 of trials with correct and incorrect responses in Bins 1 and 4. It can be seen that incorrect trials have a more negative SPN than correct trials and that the difference in SPN between correct and incorrect trials is higher in Bin 4 than in Bin 1.

correct trials. This factor also interacted with the learning bin factor, $F(2.469,54.332) = 3.236$, $p = .037$, $\eta_p^2 = 0.128$. The three-way interaction Learning Bin × Response × Laterality did not reach statistical significance, $F(4.404,96.885) = 1.444$, $p = .221$, $\eta_p^2 = 0.062$. The Learning Bin × Response interaction showed a significant linear–linear trend, $F(1,23) = 6.240$, $p = .020$, $\eta_p^2 = 0.221$. As can be seen in Figure 3A, the difference between correct and incorrect trials increased as learning progressed (see also Figure 3C). This was due to the fact that the rate at which SPN mean voltage changes across learning bins was different in the case of correct and incorrect trials. In correct trials, SPN became more positive across learning bins than in incorrect trials. As can be seen in Figure 3A), this led to an increasing difference in mean voltage between the two conditions as learning progresses.

Finally, the results of Luque et al. (2012) showed that FRN changed across learning because the positivity observed in the positive feedback trials diminished across bins, while the negativity observed in the case of the negative feedbacks did not change (for similar results, see Eppinger et al., 2008). Given this, and the results of our previous analyses, we should expect a correlation between changes in the FRN and the SPN components. More specifically, across learning a negative correlation should exist between the amplitude of the SPN and the voltage in the FRN time window in the trials with correct responses followed by positive feedbacks. On the other hand, no correlation is expected between the SPN and the FRN in trials with incorrect responses and negative trials. Also, this correlation should be observed in a within-participant fashion, as for each participant the trials with more negative SPN should lead to more negative voltages in the FRN time window in the correct-positive trials.

In order to test this prediction, we selected the trials with correct responses followed by positive feedbacks, correct-positive trials, and the trials with incorrect responses followed by negative feedbacks, incorrect-negative trials. For this analysis, the noise trials were excluded, given that they are small in number and they should not follow this same pattern, as surprise and infrequent feedbacks modulate the FRN (e.g., Hajcak, Holroyd, Moser, & Simons, 2005). For each trial, the value of the SPN was the voltage in the time window from −200 to 0 ms prefeedback in the Fz electrode using a baseline from −1,200 to −1,000 ms as in the previous

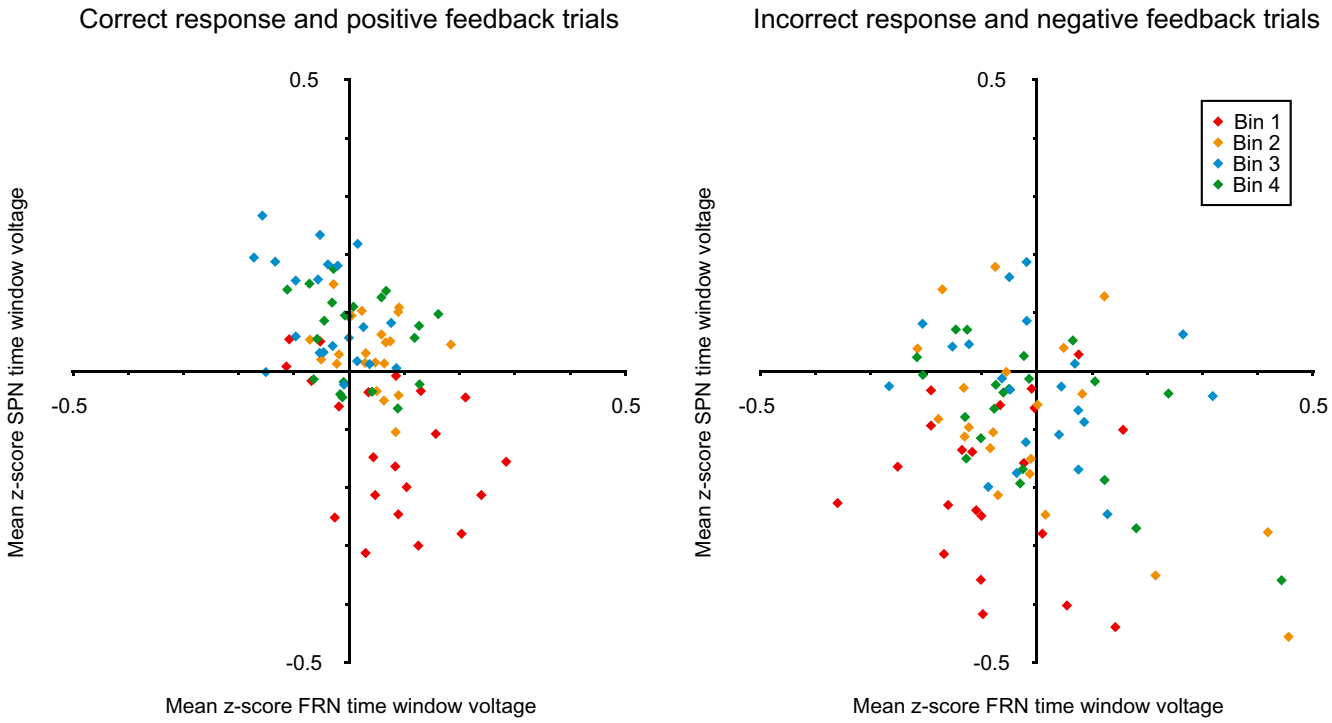

Figure 4. Correlation between the mean *z* scores of the voltages recorded during the SPN time window (−200 to 0 ms prefeedback, baseline from −1,200 to −1,000 ms prefeedback) and the FRN time window (250 to 350 ms postfeedback, baseline from −100 to 0 ms prefeedback) on each bin (Bins 1–4) for all participants. The right panel shows the scatter plot of the mean values of the correct-positive trials, in which participants made a correct response and obtained positive feedback. In these trials, a significant correlation was found between the mean scores of SPN and FRN. The left panel is the scatter plot of the incorrect-negative trials. For this type of trial, there was no significant correlation between FRN and SPN values across learning bins.

analysis. Following Luque et al. (2012), the FRN amplitude for each trial was the mean voltage in the 250 to 350 ms time window Fz electrode location, using a prefeedback baseline of −100 to 0 ms. SPN direct values for all of the trials pooled together were transformed to *z* scores. The same analysis was carried out for the FRN. Then, the mean *z* score for the SPN and the FRN components in correct-positive trials and in incorrect-negative trials was calculated for each bin in each participant (Figure 4).

Given that the prediction that is going to be tested is that the correlation between the SPN and the FRN in correct-positive trials should differ from the correlation in the incorrect-negative trials across learning, the difference between two correlations based on repeated measures has to be estimated. To do so, the best statistical option is a linear mixed model (e.g., Baayen, Davidson, & Bates, 2008; Gueorguieva & Krystal, 2004). The model tested if SPN scores could significantly predict FRN scores. It did so including two repeated measure factors, bin (Bins 1–4) and type of trial (correct-positive vs. incorrect-negative trials), and also a random factor, participants. The covariance matrix used was unstructured. The analysis showed that, as predicted, the crucial interaction between FRN and type of trial was statistically significant, $F(2,68.38) = 10.853$, $p < .001$. Further analysis showed that in the case of the correct-positive trials, there was a significant correlation between SPN and FRN scores, with a slope $\beta = -0.315$, $t(2,26.869) = 3.744$, $p < .001$, indicating a negative relation between the scores of these two variables. Meanwhile, in the case of the incorrect-negative trials, this test was not significant, $\beta = -0.047$, $t(2,73.122) = 0.44$, $p = .661$.

Discussion

The present study directly evaluated the hypothesis that the SPN could be considered an electrophysiological index of expectation during reinforcement learning. Our results showed a clear decrease in the SPN amplitude while learning progressed. Also, there was a difference in the SPN amplitude between trials with correct and incorrect responses. The former were more positive than the latter, and this difference increased with learning. This happened even though the interval from the response to the presentation of the feedback was shorter than the one used typically in SPN studies (e.g., Masaki et al., 2010). If a longer interval had been used, it might have allowed us to observe possibly clearer SPN modulations.

Previous studies supported the hypothesis that SPN reflects an expectation of a motivational outcome (Brunia et al., 2011). A very convincing result was obtained by Donkers et al. (2005) in which participants passively viewed a slot machine and experienced monetary gains and losses depending on the outcome of the machine. The slot machine consisted of three digits (from 0 to 9) that were shown sequentially on a computer screen. Participants were told that they would gain money (or lose, depending on the experimental condition) when the three digits were the same (e.g., a 9-9-9 play). The SPN analysis was focused on the time window between the presentation of the second and third digit. Consistent with the hypothesis of Brunia et al. (2011), they obtained a larger SPN in the trials in which the initial two digits were the same. In other words, a larger SPN occurred in the trials in which participants expected the third digit to determine whether they either gained or lost

money. Masaki et al. (2010) showed convergent evidence of the involvement of SPN in expectancy in a reward task. Two types of trials were included in that experiment. In the no-choice trials, participants had to press a button to know whether they gained money. The participants were told that the outcome in the no-choice trials did not depend on their behavior. In the choice trials, participants were informed that they could learn how to control the outcome with their responses (although the outcomes were at random). A larger SPN was observed in the choice trials compared to the no-choice trials. Neither of these previous experiments studied a possible effect of learning on the SPN. In contrast, we showed that the magnitude of the SPN is closely related to the learning processes in reinforcement learning tasks.

In the present study, SPN decreased in magnitude across learning. In the task used, learning should lead to better performance and an increased expectation of positive feedbacks. As described before, SPN has been interpreted as an anticipatory index underlying the expectation of the consequences of our actions. If SPN were directly related to this expectation, it should increase with learning. However, the opposite result was obtained. The fact that trials with incorrect responses have SPN values that are more negative than those trials with correct responses would not be predicted by this interpretation either. In probabilistic tasks, like the one used in the present study, humans as well as nonhuman animals tend to allocate their responses in a way proportional to the probability of reinforcement of each response option, following the matching law (Herrnstein, 1961; Savastano & Fantino, 1994). Although participants explore from time to time the response that is less reinforced, which in the present study would be the incorrect responses, they should expect a lower probability of obtaining positive feedbacks if that response is made. If in our task the SPN might directly reflect expectation of the outcome, the magnitude of the SPN observed should be smaller in this type of trial (participants might be anticipating a negative feedback). However, this was not the case. Nonetheless, given the experimental evidence available about the SPN, our results do not exclude the possibility that SPN can be modulated by expectation creation per se, only that in this case higher magnitudes of SPN were not associated with higher certainty and better performance due to learning.

Other explanations can offer more consistent and parsimonious explanations of the results obtained. One presented already in the introduction would be based on the expectation of information. This process is strongly related to expectation formation during learning, but it reflects different aspects of knowledge acquisition. In the initial trials, participants knew very little about the relations between cues and responses and, because of this, each feedback at this stage would provide much information. After additional training, feedbacks would become less informative as participants have already learned more about these relations. If the modulation of the SPN observed were due to these changes, we should expect a decrease in the magnitude of SPN as learning advances, the result obtained in our experiment. Similarly, participants chose less often the incorrect response option, and the probability of reinforcement of this option was lower. Because of this, they should have learned less about this response option than about the correct response. Also, as learning advances, the proportion of correct responses is higher, which would lead to an increased asymmetry in how much is known about these two situations. This would cause differences in how informative feedbacks are after correct and incorrect responses, and their respective SPN.

Our results, together with the previous literature, highlight the multifactorial nature of SPN and the fact that it is modulated by several different manipulations and factors. Although changes in the information provided by the feedback can explain well the results obtained, there are other possible explanations that we cannot discard. For example, the lower probability of reinforcement added to the reduced number of responses of that type made would lead to a situation of low certainty about the possibility of a positive feedback. Fuentemilla et al. (2013) have recently showed an increase in the amplitude of the SPN in situations in which their appearance was very unlikely compared to other outcomes more probable and equally desirable. Also, it is possible, given the previous studies of the SPN component (e.g., Kotani et al., 2001; Ohgami et al., 2006), that the modulations of the SPN shown could also reflect affective and motivational aspects of the expectation of the outcome acquired through learning. As training advances, participants might learn not only what kind of outcome will probably follow a response, but also about the emotional impact of its presentation. This possibility could be explored in future experiments that study how changes caused by learning in informational value and emotional aspects interact and affect SPN.

The SPN changed across learning, a result also found in the case of the FRN component. In the case of the FRN, previous studies have found that positive feedbacks show more negative voltages as learning progresses, while negative feedbacks produce a consistent response that does not change with more training (e.g., Luque et al., 2012). In this experiment, SPN values were negatively correlated with the values of the FRN in the trials with correct responses and positive feedbacks. This correlation was not significant in the case of the trials with incorrect responses and negative feedbacks. This finding supports the idea that the SPN and the FRN components are both correlates of different phases of the same learning system, feedback expectation, and processing.

The topographical distribution of the differences in SPN between correct and incorrect trials observed in the present experiment was similar to SPN distributions often observed in previous studies (van Boxtel & Böcker, 2004) (see Figure 3B). However, the scalp distribution of the difference waveform between the initial and final learning phases (Bin 1 vs. Bin 4) had its maximum amplitude at frontal and frontopolar sites (Figure 1C). Although this scalp distribution differs slightly from previous SPN studies in which the maximum amplitude was observed at the F4 electrode locations (van Boxtel & Böcker, 2004), it is similar to the one observed in Mattox et al. (2006). Mattox et al. (2006) measured SPN during two predictive learning tasks, both similar to the task used in the present study. For instance, in the authors' "weather prediction" task, participants had to predict the weather (sunny or cloudy) by pressing two different buttons. Participants had to base their predictions on the presentation of predictive cues (geometric designs). Mattox et al. (2006) did not study changes in SPN due to learning, but differences between participants with Parkinson's disease and healthy controls. Put together, the scalp distribution observed in both the present study and in Mattox et al. (2006) could reflect a differential engagement of some dopaminergic circuits when SPN is recorded during reinforcement learning.

Future experiments are required to carefully investigate the specific role of SPN as a possible correlate of information expectation during learning, the SPN subcomponents during learning, and the integration of the anterior insula in the brain network responsible for reinforcement learning. In sum, the SPN appears to offer a reliable component to measure online the cognitive processes that take place while waiting for a forthcoming feedback, which might be crucial for successful learning.

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390–412. doi: 10.1016/j.jml.2007.12.005

Böcker, K. B. E., Baas, J. M. P., Kenemans, J. L., & Verbaten, M. N. (2001). Stimulus-preceding negativity induced by fear: A manifestation of affective anticipation. *International Journal of Psychophysiology*, *43*, 77–90. doi: 10.1016/S0167-8760(01)00180-5

Brown, C. A., Seymour, B., Boyle, Y., El-Deredy, W., & Jones, A. K. P. (2008). Modulation of pain ratings by expectation and uncertainty: Behavioral characteristics and anticipatory neural correlates. *Pain*, *135*, 240–250. doi: 10.1016/j.pain.2007.05.022

Brunia, C. H., de Jong, B. M., van den Berg-Lenssen, M. M., & Paans, A. M. (2000). Visual feedback about time estimation is related to a right hemisphere activation measured by PET. *Experimental Brain Research*, *130*, 328–337. doi: 10.1007/s002219900293

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

Càmara, E., Rodríguez-Fornells, A., & Münte, T. F. (2009). Functional connectivity of reward processing in the brain. *Frontiers in Human Neuroscience*, *2*, 1–14. doi: 10.3389/neuro.09.019.2008

Chwilla, D. J., & Brunia, C. H. M. (1991). Event-related potentials to different feedback stimuli. *Psychophysiology*, *28*, 123–132. doi: 10.1111/j.1469-8986.1991.tb00400.x

Cohen, M., Elger, C. E., & Ranganath, C. (2007). Reward expectation modulates feedback related negativity and EEG spectra. *NeuroImage*, *35*, 968–978. doi: 10.1016/j.neuroimage.2006.11.056

Dickinson, A., & Burke, J. (1996). Within-compound associations mediate the retrospective revaluation of causality judgements. *Quarterly Journal of Experimental Psychology*, *49B*, 60–80.

Donkers, F. C., Nieuwenhuis, S., & van Boxtel, G. J. (2005). Mediofrontal negativities in the absence of responding. *Cognitive Brain Research*, *25*, 777–787. doi: 10.1016/j.cogbrainres.2005.09.007

Eppinger, B., Kray, J., Mock, B., & Mecklinger, A. (2008). Better or worse than expected? Aging, learning, and the ERN. *Neuropsychologia*, *46*, 521–539. doi: 10.1016/j.neuropsychologia.2007.09.001

Foti, D., & Hajcak, G. (2012). Genetic variation in dopamine moderates neural response during reward anticipation and delivery: Evidence from event-related potentials. *Psychophysiology*, *49*, 617–626. doi: 10.1111/j.1469-8986.2011.01343.x

Fuentemilla, L., Cucurell, D., Marco-Pallarés, J., Guitart-Masip, M., Morís, J., & Rodríguez-Fornells, A. (2013). Electrophysiological correlates of anticipating improbable but desired events. *NeuroImage*, *78*, 135–144. doi: 10.1016/j.neuroimage.2013.03.062

Gueorguieva, R., & Krystal, J. H. (2004). Move over ANOVA: Progress in analyzing repeated-measures data and its reflection in papers published in the Archives of General Psychiatry. *Archives of General Psychiatry*, *61*, 310–317. doi: 10.1001/archpsyc.61.3.310

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. doi: 10.1111/j.1469-8986.2005.00278.x

Herrnstein, R. J. (1961). Relative and absolute strength of response as a function of frequency of reinforcement. *Journal of the Experimental Analysis of Behavior*, *4*, 267–272.

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*, 679–709. doi: 10.1037//0033-295X.109.4.679

Hughes, G., & Yeung, N. (2011). Dissociable correlates of response conflict and error awareness in error-related brain activity. *Neuropsychologia*, *49*, 405–415. doi: 10.1016/j.neuropsychologia.2010.11.036

Kotani, Y., Hiraku, S., Suda, K., & Aihara, Y. (2001). Effect of positive and negative emotion on stimulus-preceding negativity prior to feedback stimuli. *Psychophysiology*, *38*, 873–878. doi: 10.1111/1469-8986.3860873

Kotani, Y., Kishida, S., Hiraku, S., Suda, K., Ishii, M., & Aihara, Y. (2003). Effects of information and reward on stimulus-preceding negativity prior to feedback stimuli. *Psychophysiology*, *40*, 818–826. doi: 10.1111/1469-8986.00082

Kotani, Y., Ohgami, Y., Kuramoto, Y., Tsukamoto, T., Inoue, Y., & Aihara, Y. (2009). The role of the right anterior insular cortex in the right hemisphere preponderance of stimulus-preceding negativity (SPN): An fMRI study. *Neuroscience Letters*, *450*, 75–79. doi: 10.1016/j.neulet.2008.11.032

Luck, S., & López-Calderon, J. (2012) ERPLAB V3.0. Retrieved from http://erpinfo.org/erplab

Luque, D., Flores, A., & Vadillo, M. A. (2013). Revisiting the role of within-compound associations in cue-interaction phenomena. *Learning & Behavior*, *41*, 61–76. doi: 10.3758/s13420-012-0085-3

Luque, D., López, F. J., Marco-Pallares, J., Càmara, E., & Rodríguez-Fornells, A. (2012). Feedback-related brain potential activity complies with basic assumptions of associative learning theory. *Journal of Cognitive Neuroscience*, *24*, 794–808. doi: 10.1162/jocn_a_00145

Masaki, H., Takeuchi, S., Gehring, W. J., Takasawa, N., & Yamazaki, K. (2006). Affective motivational influences on feedback-related ERPs in a gambling task. *Brain Research*, *1105*, 110–121. doi: 10.1016/j.brainres.2006.01.022

Masaki, H., Yamazaki, K., & Hackley, S. H. (2010). Stimulus-preceding negativity is modulated by action-outcome contingency. *NeuroReport*, *21*, 277–281. doi: 10.1097/WNR.0b013e3283360bc3

Mattox, S. T., Valle-Inclán, F., & Hackley, S. A. (2006). Psychophysiological evidence for impaired reward anticipation in Parkinson's disease. *Clinical Neurophysiology*, *117*, 2144–2153. doi: 10.1016/j.clinph.2006.05.026

Müller, S. V., Möller, J., Rodríguez-Fornells, A., & Münte, T. F. (2005). Brain potentials related to internal and external information used for performance monitoring. *Clinical Neurophysiology*, *116*, 63–74. doi: 10.1016/j.clinph.2004.07.009

Ohgami, Y., Kotani, Y., Hiraku, S., Aihara, Y., & Ishii, M. (2004). Effects of reward and stimulus modality on stimulus-preceding negativity. *Psychophysiology*, *41*, 729–738. doi: 10.1111/j.1469-8986.2004.00203.x

Ohgami, Y., Kotani, Y., Tsukamoto, T., Omura, K., Inoue, Y., Aihara, Y., & Nakayama, M. (2006). Effects of monetary reward and punishment on stimulus-preceding negativity. *Psychophysiology*, *43*, 227–236. doi: 10.1111/j.1469-8986.2006.00396.x

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York, NY: Appleton.

Rodríguez-Fornells, A., Kurzbuch, A. R., & Münte, T. F. (2002). Time course of error detection and correction in humans: Neurophysiological evidence. *Journal of Neuroscience*, *22*, 9990–9996.

Savastano, H. I., & Fantino, E. (1994). Human choice in concurrent ratio-interval schedules of reinforcement. *Journal of the Experimental Analysis of Behavior*, *61*, 453–463. doi: 10.1901/jeab.1994.61-453

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron*, *36*, 241–263. doi: 10.1016/S0896-6273(02)00967-4

Tsukamoto, T., Kotani, Y., Ohgami, Y., Omura, K., Inoue, Y., & Aihara, Y. (2006). Activation of insular cortex and subcortical regions related to feedback stimuli in a time estimation task: An fMRI study. *Neuroscience Letters*, *399*, 39–44. doi: 10.1016/j.neulet.2006.01.061

van Boxtel, G. J., & Böcker, K. B. E. (2004). Cortical measures of anticipation. *Journal of Psychophysiology*, *18*, 61–76. doi: 10.1027/0269-8803.18.2–3.61

Waelti, P., Dickinson, A., & Schultz, W. (2001). Dopamine responses comply with basic assumptions of formal learning theory. *Nature*, *412*, 43–48. doi: 10.1038/35083500

Walsh, M. M., & Anderson, J. R. (2011). Modulation of the feedback-related negativity by instruction and experience. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 19048–19053. doi: 10.1073/pnas.1117189108

(Received June 13, 2012; Accepted April 12, 2013)