

# Event-Related Potential Activity in the Basal Ganglia Differentiates Rewards from Nonrewards: Temporospacial Principal Components Analysis and Source Localization of the Feedback Negativity

Dan Foti,<sup>1\*</sup> Anna Weinberg,<sup>1</sup> Joseph Dien,<sup>2</sup> and Greg Hajcak<sup>1</sup>

<sup>1</sup>Department of Psychology, Stony Brook University, Stony Brook, New York

<sup>2</sup>Center for Advanced Study of Language, University of Maryland, College Park, Maryland

---

**Abstract:** Event-related potential studies of reward processing have consistently identified the feedback negativity (FN), an early neural response that differentiates feedback indicating unfavorable versus favorable outcomes. Several important questions remain, however, about the nature of this response. In this study, the FN was recorded in response to monetary gains and losses during a laboratory gambling task, and temporospacial principal components analysis was used to separate the FN from overlapping responses. The FN was identified as a positive deflection at frontocentral recording sites that was enhanced for rewards compared with nonrewards. Furthermore, source localization techniques identified the striatum as a likely neural generator. These data indicate that this apparent FN reflects increased striatal activation in response to favorable outcomes that is reduced or absent for unfavorable outcomes, thereby providing unique information about the timing and nature of basal ganglia activity related to reward processing. *Hum Brain Mapp* 32:2207–2216, 2011. © 2011 Wiley Periodicals, Inc.

**Key words:** event-related potentials; visual evoked potentials; corpus striatum; putamen; reinforcement

---

## INTRODUCTION

Ongoing feedback from the environment informs decisions and shapes behavior by signaling the consequences of our actions. To better understand the processes through which environmental feedback is evaluated, researchers have examined event-related potentials (ERPs), scalp-recorded measures of neural activity with precise temporal resolution. For more than a decade, these studies have consistently observed an apparent negative deflection in the

ERP that differentiates negative outcomes, such as errors and monetary loss, from positive outcomes. Referred to as the feedback error-related negativity or feedback negativity (FN), this response peaks ~250–300 ms after feedback presentation and is maximal at frontocentral recording sites [Holroyd and Coles, 2002; Miltner et al., 1997].

When examined in the context of laboratory gambling tasks where individuals may win or lose money on each trial, several studies have demonstrated that the FN is sensitive to the valence but not the magnitude of outcomes, the FN distinguishes between monetary gains and losses, but is equivalent for larger compared to smaller losses [Hajcak et al., 2006; Sato et al., 2005; Yeung and Sanfey, 2004]. The valence of an outcome, however, is context dependent: “neutral” feedback indicating that money has neither been won nor lost may actually be favorable or unfavorable, depending on the range of other possible outcomes. Indeed, there is evidence that the FN tracks the relative valence of an outcome within the immediate context,

---

\*Correspondence to: Dan Foti, Department of Psychology, Stony Brook University, Stony Brook, NY 11794-2500.

E-mail: daniel.foti@stonybrook.edu

Received for publication 2 March 2010; Revised 21 June 2010; Accepted 7 September 2010

DOI: 10.1002/hbm.21182

Published online 8 February 2011 in Wiley Online Library (wileyonlinelibrary.com).

such that the amplitude of the FN elicited by neutral feedback depends on whether the alternative would have been to win or lose money on that trial [Holroyd et al., 2006; Holroyd et al., 2004]. Likewise, feedback is also evaluated relative to one's expectations immediately before the outcome, and this is reflected in the FN such that outcomes that violate reward predictions elicit a larger response [Hajcak et al., 2007; Holroyd et al., 2003]. Together, these lines of research indicate that the FN reflects a process in which outcomes are evaluated as either better or worse than expected.

Although there is a growing literature using the FN to examine the neural processing of environmental feedback, some fundamental questions remain about the nature and functional significance of this ERP component. For example, it remains unclear whether variation in the FN reflects activity related to negative feedback, positive feedback, or both [Holroyd, 2004]. The FN is often interpreted as a negative-going ERP component that is enhanced for feedback indicating errors and nonrewards, presumably reflecting a neural process that tracks the occurrence of unfavorable outcomes. However, this issue is obscured by the fact that many studies have used a difference wave approach, in which the FN is quantified as the numerical difference between negative and positive feedback [e.g., Dunning and Hajcak, 2007; Foti and Hajcak, 2009; Hajcak et al., 2007; Holroyd et al., 2008; Miltner et al., 1997]. Although this approach isolates variation in the ERP associated with feedback valence, it cannot attribute that variation to a specific outcome. The issue is compounded by the fact that ERP amplitudes are not inherently meaningful—an apparent decrease in amplitude of one component could instead be due to the onset of an overlapping component with opposite polarity (Luck, 2005). In this way, the FN may instead reflect a positive-going ERP response to favorable outcomes which is then reduced for unfavorable outcomes. Indeed, [Holroyd et al. 2008] have reported data that are consistent with the possibility that apparent variation in the FN is primarily due to activity related to positive feedback, although they could not conclude whether this response on positive feedback trials was a positive deflection in the ERP or the attenuation of a negative deflection.

In a similar vein, an observed change in FN amplitude could reflect variation in the P300, an ERP response that overlaps in time with the FN and is also sensitive to expectation violations [Courchesne et al., 1977; Duncan-Johnson and Donchin, 1977; Hajcak et al., 2005; Hajcak et al., 2007; Johnson and Donchin, 1980]. In several studies examining feedback processing, the P300 has been shown to be larger for positive compared to negative feedback [Hajcak et al., 2005; Hajcak et al., 2007; Holroyd et al., 2006; Holroyd et al., 2004], data which suggests that, similar to the FN, the P300 may be sensitive to outcome valence. Accordingly, the use of difference-wave and peak measures may confound variation in these two ERP responses across feedback type. By contrast, other studies have found the P300 is sensitive to outcome magnitude but insensitive to

valence [Goyer et al., 2008; Sato et al., 2005; Yeung and Sanfey, 2004], suggesting that the P300 may reflect a distinct aspect of feedback processing that is not captured by the FN. The heterogeneity of these findings underscores the importance of quantifying the FN and the P300 in such a way as to minimize the impact of component overlap.

In this regard, our understanding of the FN may be enhanced through the application of a data reduction technique such as principal components analysis (PCA). There is a long history of using PCA as an empirically driven method of decomposing a raw ERP waveform across time into its underlying components [Donchin and Heffley, 1979], and with the recent proliferation of high-density ERP recording montages, PCA has proven to be useful in reducing the spatial dimensions of ERP datasets as well. For example, a combination of temporal and spatial PCA has been used to effectively separate the P300 from the Novelty P3 and Slow Wave components [Dien et al., 2003; Spencer et al., 2001], all of which are positive-going ERP responses at posterior sites that overlap in time.

In this study, we were interested in applying temporospatial PCA to ERP data recorded during a simple laboratory gambling task, with the goal of fully distinguishing the FN from overlapping responses. Unlike the conventional windowed difference wave approach, this approach allows us to more readily examine the latent structure of neural activity elicited by rewards versus nonrewards. A second goal of this study was to identify likely neural generators of ERP components that differentiate rewards from nonrewards. A recent report using a simulated dataset concluded that applying source localization techniques to PCA factors yields significantly more accurate results compared with localizing the ERP waveform directly [Dien, 2010b]. By applying both PCA and source localization techniques to the FN, we hope to better isolate ERP signals sensitive to reward versus nonreward and to identify the likely neural generator of this activity.

## METHODS

### Participants

Twenty-two undergraduate students (10 female, 12 male) participated in this study. The mean age of the sample was 18.36 (standard deviation = 1.05). This sample was drawn from a larger study that examined sad versus neutral mood inductions on ERP responses [Foti and Hajcak, 2010]; the current analyses were based only on subjects assigned to the neutral condition—induction details are described later under Procedure. No participants discontinued their participation in the experiment once the procedures had begun. All participants received course credit and \$5.00 (winnings from the gambling task) for their participation. Informed consent was obtained from participants before each experiment. This study was formally approved by the Stony Brook University Institutional Review Board.

### Task and Materials

The gambling task was administered using Presentation software (Neurobehavioral Systems, Inc., Albany, CA) to control the presentation and timing of all stimuli. On each trial, participants were shown a graphic displaying two doors (occupying 6° of the visual field vertically and 8° horizontally) and were told to choose which door they wanted to open. Participants were told to press the left mouse button to choose the left door or the right mouse button to choose the right door. Following each choice, a feedback stimulus appeared on the screen informing the participants whether they won or lost money on that trial. A green “↑” indicated a correct guess and a gain of \$0.50, whereas a red “↓” indicated an incorrect guess and a loss of \$0.25 (each occupying 3° of the visual field vertically and 1° horizontally). A fixation mark (+) was presented before the onset of each stimulus. At the end of each trial, participants were presented with the instruction “Click for the next round.” The task consisted of 40 trials total, with positive feedback given on exactly 20 trials (i.e., 50%). Feedback was presented in a random order for each participant. The order and timing of all stimuli was as follows: (i) the graphic of two doors was presented until a response was made, (ii) a fixation mark was presented for 1000 ms, (iii) a feedback arrow was presented for 2000 ms, (iv) a fixation mark was presented for 1500 ms, and (v) Click for the next round was presented until a response was made.

### Procedure

Following a brief description of the experiment, electroencephalographic (EEG) sensors were attached. Before the gambling task, all participants viewed two 5-minute film clips and performed two computer tasks unrelated to this study while listening to a neutrally valenced song [Foti and Hajcak, 2010]. To familiarize participants with the gambling task, they were given a practice block containing five trials. Participants then performed the main task, with the running total of money earned to that point presented after the first 20 trials. Finally, participants performed another unrelated computer task, were paid their winnings (\$5.00), and watched an amusing film clip. This experiment was conducted as part of a larger study examining the influence of mood on a range of ERP components, and the amusing film clip was used with all participants to serve as a positive mood induction at the conclusion of the laboratory session.

### Psychophysiological Recording, Data Reduction, and Analysis

The continuous EEG was recorded using a custom cap (Cortech Solutions, Wilmington, NC) and the ActiveTwo BioSemi system (BioSemi, Amsterdam, Netherlands). The signal was preamplified at the electrode with a gain of 16×

the EEG was digitized at 64-bit resolution with a sampling rate of 512 Hz using a low-pass fifth order sinc filter with a half-power cutoff of 102.4 Hz. Recordings were taken from 64 scalp electrodes based on the 10/20 system, as well as two electrodes placed on the left and right mastoids. The electro-oculogram was recorded from four facial electrodes: two 1 cm above and below the left eye, one 1 cm to the left of the left eye, and one 1 cm to the right of the right eye. Each electrode was measured online with respect to a common mode sense electrode that formed a monopolar channel. Off-line analysis was performed using Brain Vision Analyzer software (Brain Products, Munich, Germany). All data were re-referenced to the average of all scalp electrodes and band-pass filtered with cutoffs of 0.1 and 30 Hz. The EEG was segmented for each trial, beginning 200 ms before feedback onset and continuous for 1000 ms following feedback onset. Each trial was corrected for blinks and eye movements using the method developed by [Gratton et al. 1983].

Specific channels were rejected in each trial using a semi-automated procedure, with physiological artifacts identified by the following criteria: a step of >50  $\mu$ V between sample points, a difference of 300  $\mu$ V within a trial, and a maximum difference of <0.5  $\mu$ V within 100-ms intervals. Additional physiological artifacts were identified using visual inspection. Across channels and subjects, the mean number of usable trials was 18.64 (95% CI = 17.50–19.77) and 18.64 (95% CI = 17.30–19.97) for nonrewards and rewards, respectively. Of the 22 total participants, 12 had artifact-free data at frontocentral recorded sites (i.e., 20 usable reward trials and 20 usable nonreward trials), nine participants had between 15 and 19 usable trials for both nonrewards and rewards, and one participant had nine usable nonreward trials and six usable reward trials. Although the stability of the FN as a function of number of trials has not been empirically examined to date, there is data on the error-related negativity indicating that six trials are sufficient to achieve a stable measurement with a good signal-to-noise ratio for that ERP response [Olvet and Hajcak, 2009]; therefore, we elected to retain all 22 participants for further analysis.

Stimulus-locked ERPs were averaged separately for nonrewards (i.e., monetary losses) and rewards (i.e., monetary gains), and the activity in the 200-ms window before feedback onset served as the baseline. The FN was quantified using temporospatial PCA, a technique which extracts linear combinations of data points that meet certain criteria that tend to distinguish between consistent patterns of electrocortical activity [Dien and Frishkoff, 2005]. This analysis was conducted using the ERP PCA Toolkit, version 1.3 [Dien, 2010a]. Following recently published sets of guidelines for applying PCA to ERP datasets [Dien, 2010b; Dien et al., 2005; Dien et al., 2007], a temporal PCA was performed on the data first to capture variance across time points. This PCA used all time points from each participant’s averaged ERP as variables, and it considered participants, trial types, and recording sites as observations. Promax rotation was used, and nine temporal factors were

extracted based on the resulting Scree plot [Cattell, 1966]. For each temporal factor, this analysis yielded factor scores for each combination of electrode, participant, and trial type, representing the amount of activity in the original data captured by that factor. The spatial distribution of these factor scores was then analyzed using spatial PCA. This PCA used all recording sites as variables, and it considered all participants, trial types, and temporal factor scores as observations. A separate spatial PCA was performed for each of the nine temporal factors. Infomax rotation was used, and based on the averaged Scree plot for all nine temporal factors, four spatial factors were extracted, yielding 36 unique factors combinations. The covariance matrix and Kaiser normalization were used for each PCA. The waveforms for each factor were reconstructed (i.e., converted to microvolts) by multiplying the factor pattern matrix with the standard deviations. Factors of interest were scored using the peak values on nonreward and reward trials. Statistical analysis was performed using SPSS (17.0; SPSS, Inc., Chicago, IL).

Source analysis techniques were used for those factors which significantly differentiated nonreward from reward trials. This analysis was conducted by specifying a pair of hemispheric dipoles (the second dipole mirroring position but not orientation) in Robert Oostenveld's FieldTrip (<http://fieldtrip.fcdonders.nl/start>) using a four-shell model. Mirrored dipoles were used because the neuroimaging literature indicates that even lateralized activity often involves both homologous hemispheric locations. The use of more than a pair of dipoles was avoided because of findings that localization accuracy suffers when trying to model simultaneous dipoles [Zhang et al., 1994], and indeed a major impetus for the use of the PCA was to avoid having to do so. Successfully meeting the generally accepted guideline for a good quality solution [residual variance (RV) no >10% RV] was taken as evidence that no more dipoles were needed for a given solution. Not meeting this criterion was taken as an indicator that robust source solutions would not be possible. Time windows were not specified because the spatial distribution of two-step PCA factors is identical across the entire time course (insofar as each temporospatial factor is characterized by a single set of temporal factor loadings and a single set of spatial factor loadings). The entire epoch was therefore selected for the fitting process. A grid scan first surveyed the head space for a rough estimate of the best starting position, thus minimizing the possibility of solving for a local minimum. An iterative algorithm was then used in which the program automatically shifted the position of the dipoles until it found a position of maximum fit using a maximum-likelihood estimation algorithm [Lutkenhoner, 1998].

Stability of the solution was then assessed with a jack-knife technique. First, the second (spatial PCA) solution was recomputed twenty-two times, each with one of the participants left out. The spatial factor best corresponding to that of the original (in terms of scalp topography) was

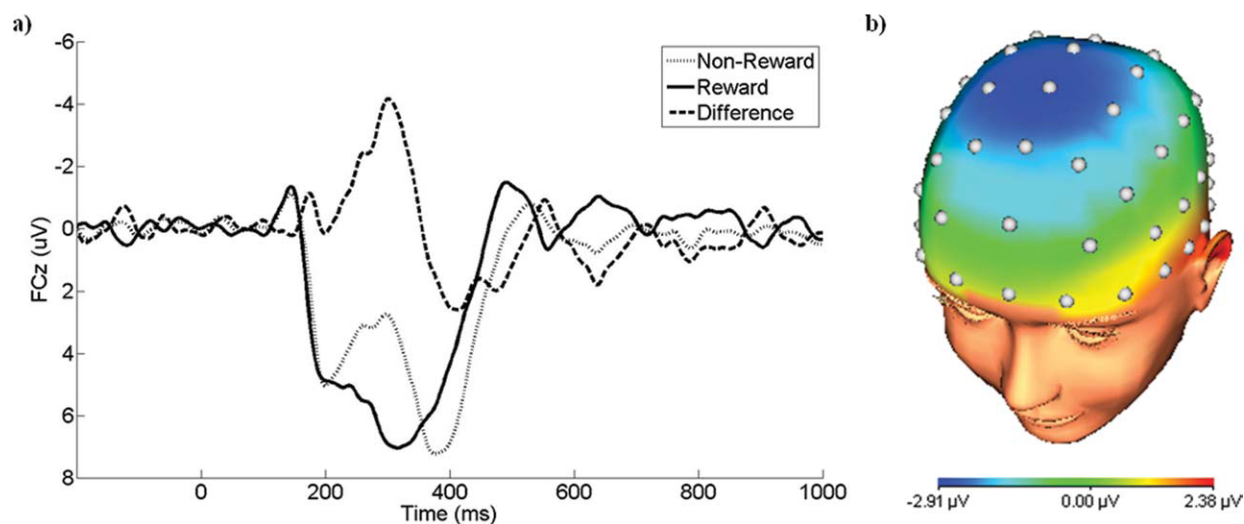
then identified for each jack-knife solution. The source analysis was then computed for each of the jack-knife solutions. In this manner, it is possible to estimate the extent to which the solution is affected by individual participants without falling prey to the instability of single-subject PCA and source analysis solutions. A graphical plot of the solutions can therefore provide a sense of the individual variability associated with a source solution. The jack-knife results were also utilized to perform an inferential test of which hemispheric dipole was stronger, using a published jack-knife test statistic [Miller et al., 1998]. The jack-knife procedure was implemented in the ERP PCA Toolkit version 2.17 [Dien, 2010a].

## RESULTS

Waveforms depicting the grand average ERPs (i.e., before PCA) are presented in Figure 1. Consistent with the FN literature, nonreward trials were associated with a relative negativity that peaked at ~300 ms following feedback onset and was maximal at frontocentral electrodes. This deflection in the waveform for nonreward trials was superimposed on the P300, which was maximal at ~400 ms.

Of the 36 total factor combinations yielded by the PCA, 12 accounted for at least 1% of the total variance in the data. One property of temporospatial PCA is that microvolt-scaled waveforms representing the portion of the data accounted for by each factor can be attained by multiplying the factor loadings, scores, and their standard deviations [Dien et al., 2003]. Following the suggestion of Dien et al. [Dien et al., 2005], visual inspection of the waveforms associated with these 12 factors was used to select those that most readily corresponded to ERP components relevant to the gambling paradigm used. Four factors were chosen for further statistical analysis, corresponding to the P200, FN, P300, and Slow Wave components (Table I). Factors not considered either did not closely correspond with a known ERP component, did not have a well-defined timing or scalp distribution, were too small to be of interest (e.g., a maximum deflection of <0.5  $\mu$ V), or reflected activity during the prebaseline period.

The waveforms and scalp topography associated with Temporal Factor 3, Spatial Factor 1 (TF3/SF1), the PCA factor corresponding to the FN, are presented in Figure 2. As evident in the grand average ERP waveforms, this PCA factor represents a relative negativity for nonreward compared to reward trials that is maximal at frontocentral sites, and this difference across feedback types was statistically significant ( $t(21) = 6.40$ ,  $P < 0.001$ ). Additionally, the PCA waveform reveals that this variation in the ERP is driven by a positive response to reward trials that is absent on nonreward trials. Source localization of this difference between nonrewards and rewards identified the putamen as a likely neural generator, with Talairach coordinates of (23.0, 2.7, 6.3) and RV of 1.5%, indicating that



**Figure 1.**

a) The grand average waveform of the feedback negativity at FCz for nonreward and reward trials, as well as the difference. b) The scalp topography of the difference between nonreward and reward trials at a window spanning 250–350 ms. [Color figure can be viewed in the online issue, which is available at [wileyonlinelibrary.com](http://wileyonlinelibrary.com).]

rewards elicit increased activity in the putamen compared with nonrewards. Jack-knife analysis indicated this source solution to be highly stable across subsamples: of the 22 jack-knife solutions, 18 localized to the putamen (RV ranging from 1.4 to 2.6%). The remaining four localized to the globus pallidus (Talairach: 23.5, -13.3, 2.6; RV = 1.6%), lentiform nucleus (Talairach: 23.76, -1.4, -0.2; RV = 1.7%), claustrum (Talairach: -32.5, -10.2, 1.9; RV = 1.8%), and anterior commissure (Talairach: -30.8, -11.1, -2.5; RV = 1.8%). Across all of the jack-knife solutions, there was also evidence of lateralization, with consistently greater dipole strength in the right hemisphere ( $t(21) = 2.90, p < 0.01$ ).

Comparisons across feedback type revealed that the P200, P300, and Slow Wave components did not significantly differ between nonreward and reward trials (all  $ps > 0.05$ ), indicating that variation in the scalp-recorded ERP waveform across trial types was primarily due to the FN. Figures 3 and 4 display the overlaid waveforms for the four PCA factors that we selected for statistical analysis, illustrating visually how they combine to yield the

observed negative deflection to nonrewards in the original ERP waveform. In short, both nonrewards and rewards elicit equivalent P200 and Slow Wave responses at frontocentral recording sites, peaking at ~200 and 400 ms, respectively. Only rewards, however, elicit a third, intermediate positivity, peaking at ~300 ms, and the absence of this mid-range positivity on nonreward trials gives rise to the observed negativity in the scalp-recorded ERP.

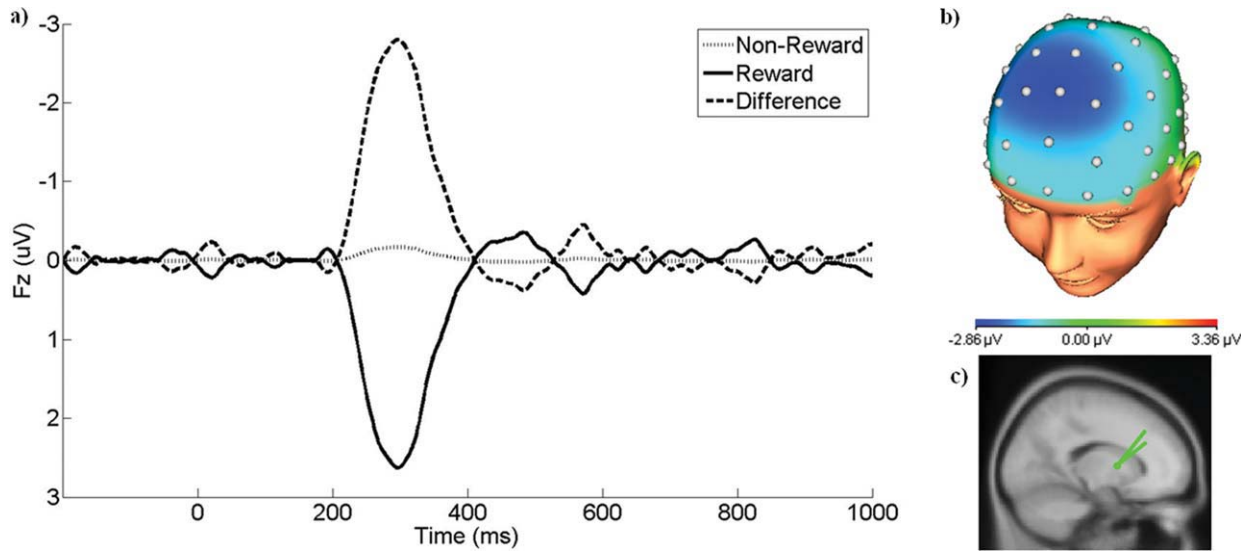
## DISCUSSION

In the current study, reward trials were associated with a positive-going deflection in the ERP at frontocentral sites that was maximal at 297 ms following feedback onset. On nonreward (i.e., monetary loss) trials, this component was reduced and was less positive. Previous ERP work on feedback processing has typically interpreted the FN as a negativity in response to unfavorable outcomes [e.g., Milner et al., 1997], whereas the current results indicate that variation in the ERP across feedback types may actually be

**TABLE I. PCA factor combinations selected for statistical analysis**

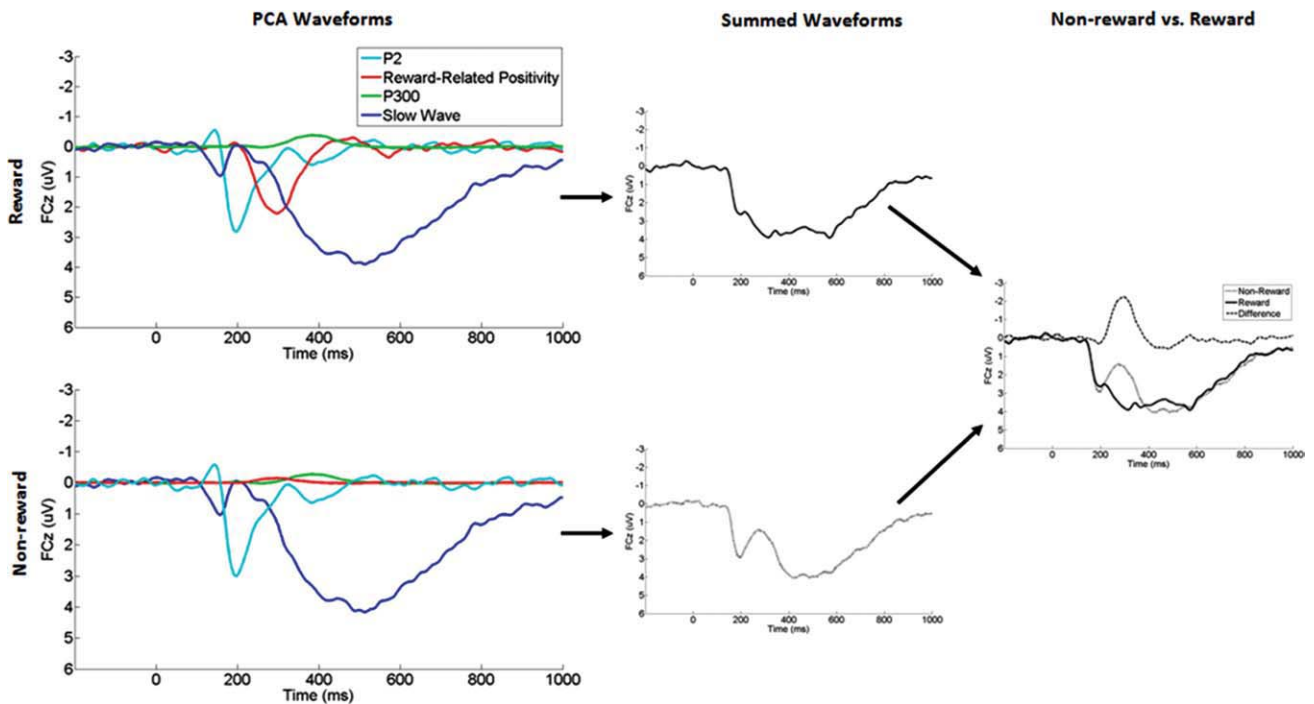
Corresponding ERP component	Temporospatial factor combination	Variance explained (%)	Temporal loading peak (ms)	Spatial distribution	Nonreward vs. reward, $t(21)$
P200	TF5/SF1	2.1	197	Frontal positivity	0.63
FN	TF3/SF1	3.4	297	Frontocentral positivity	6.40 <sup>a</sup>
P300	TF2/SF1	4.7	385	Parietal positivity	1.91
Slow Wave	TF1/SF2	11.2	514	Centroparietal positivity	0.85

<sup>a</sup> $P < 0.001$



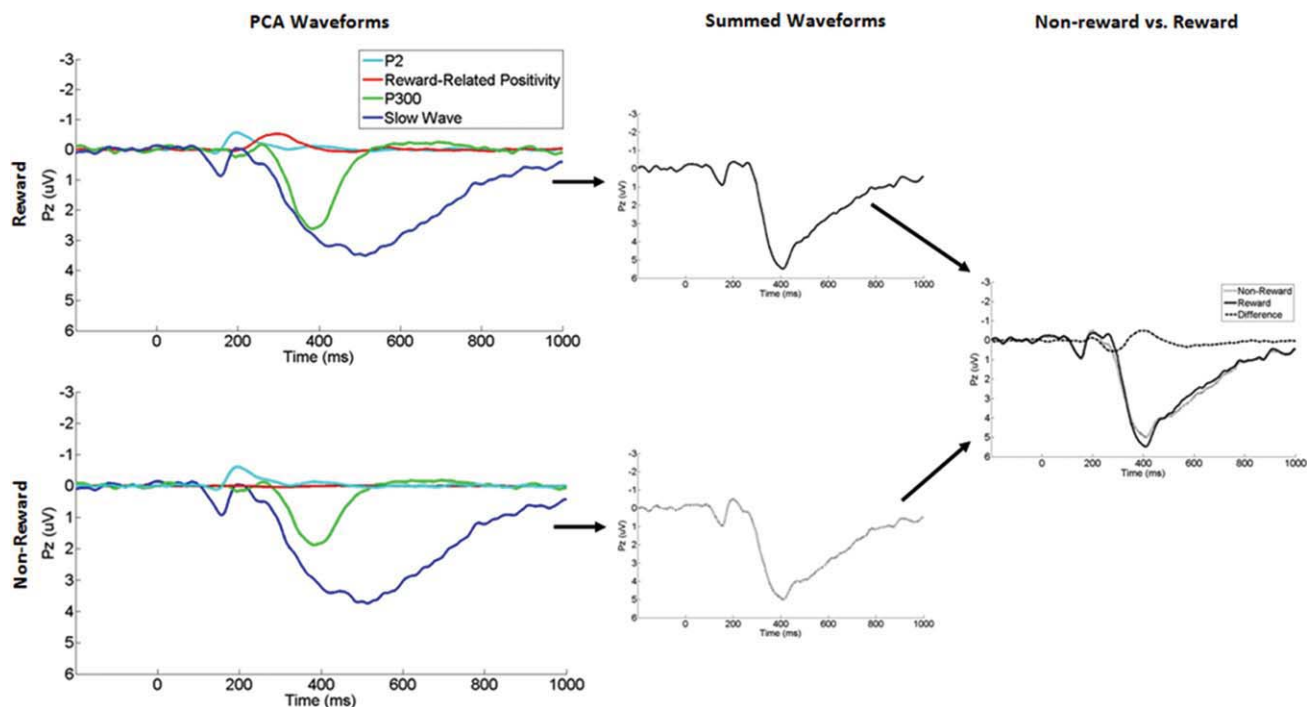
**Figure 2.**

a) The waveforms at Fz representing the portion of the ERP associated with TF3/SF1, the PCA factor corresponding to the feedback negativity. Waveforms are presented for nonreward and reward trials, as well as the difference. b) The scalp topography of the difference between nonreward and reward trials at 297 ms, where the temporal loading is maximal. c) The dipole source associated with TF3/SF1. [Color figure can be viewed in the online issue, which is available at [wileyonlinelibrary.com](http://wileyonlinelibrary.com).]



**Figure 3.**

The waveforms of the four temporospatial factors at FCz, their algebraic sum, and the comparison between nonreward and reward trials. [Color figure can be viewed in the online issue, which is available at [wileyonlinelibrary.com](http://wileyonlinelibrary.com).]



**Figure 4.**

The waveforms of the four temporospatial factors at Pz, their algebraic sum, and the comparison between nonreward and reward trials. [Color figure can be viewed in the online issue, which is available at [wileyonlinelibrary.com](http://wileyonlinelibrary.com).]

driven by the modulation of a *positivity* to *favorable* outcomes. In fact, this builds upon the results of a recent study published by [Holroyd et al. 2008], in which they compared ERPs recorded during a time estimation tasks and an oddball task in order to differentiate responses elicited specifically by feedback from those elicited more generally by salient, task-relevant stimuli. They concluded that the FN elicited by negative feedback on the time estimation task was similar to the N200 that was elicited by target stimuli in the oddball task—indeed, there exists a considerable literature on the N200 as a general response to salient, task-relevant stimuli [for a review, see Folstein and Van Petten, 2008]. [Holroyd et al. 2008], therefore, asserted that the modulation of the FN by feedback type in the time estimation task was driven by positive feedback, either as a reduced N200 or as an opposite-going positive ERP component in the same range. They were unable to distinguish between these two possibilities, however, as they analyzed the difference between the ERP response to the target stimulus on the oddball task and the ERP responses to positive and negative feedback on the time estimation task. Here, we shed new light on this question by providing direct evidence supporting the possibility that the FN is driven by the modulation of a reward-related positivity. We used temporospatial PCA to decompose the ERP responses to rewards and nonrewards, rather than analyz-

ing the difference wave, and found that rewards were associated with an increased positivity compared with nonrewards. This positivity occurs in between two other positive deflections, the P200 and Slow Wave components, and the reduction of this reward-related positivity yields an apparent negativity in the scalp-recorded ERP on nonreward trials. In fact, we observed this same pattern in a prior application of PCA to ERP data within a separate sample [Foti and Hajcak, 2009], although we did not elaborate on the point in that report. Across these multiple studies, there appears to be converging evidence indicating a positive ERP follows rewards, which is reduced for feedback indicating nonreward.

Importantly, this study suggests that this differentiation of rewards from nonrewards may be partly due to activation in the putamen, a region of the striatum. Traditionally, subcortical regions such as the putamen have been thought to contribute little to scalp-recorded EEG signals due to the physical properties of spiny stellate neurons [Lorente de No, 1947], but this perspective has been challenged in recent years [Sander et al., 2010]. For example, [Rektor 2002] recorded the readiness potential in epileptic patients using both scalp and intracranial electrodes. By using implanted EEG electrodes, Rektor localized the scalp-recorded signal to restricted areas in the motor and somatosensory cortex, but these cortical generators alone

were insufficient to account for the magnitude and widespread scalp distribution of the readiness potential. Activation consistent with the readiness potential was also found in numerous subcortical regions, including the putamen, indicating that these regions may also contribute to the scalp-recorded signal. In fact, a recent simulation of single-cell neuronal activity indicated that spiny stellate cells produce a current dipole that is stronger than previously thought and on the same order of magnitude as pyramidal cells [Murakami and Okada, 2006]. Building on this result using a whole-brain anatomical model, Monte-Carlo simulations indicated that activity in subcortical regions contributes to the scalp-recorded EEG, and that signals from the putamen may be detected at the scalp with relatively few trials, even in the presence of random cortical activity [Attal et al., 2009].

In addition, localization of EEG activity in response to rewards versus nonrewards in the putamen is notable in light of strong evidence from single-neuron recording studies in animals [Schultz, 2002] and neuroimaging studies in humans [Delgado, 2007] implicating the striatum as a critical brain region involved in reward processing and guiding goal-directed behavior. For example, activity within several striatal regions, including the putamen, was found to differentiate rewards from nonrewards, but was insensitive to reward magnitude [Elliott et al., 2003] data which parallels findings on the FN [Hajcak et al., 2006; Sato et al., 2005; Yeung and Sanfey, 2004]. There is also evidence that, like the FN, the putamen is sensitive to violations of reward expectations [McClure et al., 2003; O'Doherty et al., 2003]. Specifically, [McClure et al. 2003] manipulated expectations by conditioning participants to expect the delivery reward at a set time within each trial and then interspersing several trials with a delayed reward. The delivery of an unexpected reward was associated with increased putamen activity compared to expected rewards; conversely, the absence of an expected reward was associated with decreased putamen activity compared with periods where no reward was expected. This pattern closely parallels the modulation of the FN by outcome expectancy [Hajcak et al., 2007; Holroyd et al., 2003], whereby the difference between nonreward and reward trials is larger for unpredicted than predicted outcomes. Although we localized the EEG signal to a pair of hemispheric dipoles, there was consistently stronger activation in the right dipole compared to the left, suggesting some lateralization of putamen activity in the current study. Neuroimaging studies, meanwhile, have reported reward-related putamen activity that is right-lateralized [Elliott et al., 2003], left-lateralized [McClure et al., 2003; O'Doherty et al., 2003; Pessiglione et al., 2006], and bilateral [Haruno and Kawato, 2006; Kirsch et al., 2003; Tobler et al., 2006]. By localizing EEG activity to the putamen, this study provides further insight into the FN. Linking this ERP signal to a substantial existing neuroimaging literature on reward processing, the FN seems to be generated in part by reward-related striatal activity that is increased in response to rewards compared to nonrewards.

The FN has previously been discussed in terms of reinforcement learning theory, by which feedback elicits phasic increases and decreases in midbrain dopamine signals when outcomes are better or worse than expected, respectively [Holroyd and Coles, 2002]. In two previous reports, the FN was localized to the anterior cingulate cortex [Gehring and Willoughby, 2002; Miltner et al., 1997]. Linking the anterior cingulate cortex to reinforcement learning models, Holroyd and Coles proposed that the FN reflects disinhibition of anterior cingulate cortex neurons from dopaminergic inputs from the basal ganglia [Holroyd and Coles, 2002]. Indeed, using neuroimaging techniques, activity in the anterior cingulate cortex has been related to feedback processing in some studies [Bush et al., 2002; Knutson et al., 2000] but not others [Elliott et al., 2003; Nieuwenhuis et al., 2005].

One primary strength of current study is the use of temporospatial PCA to isolate the FN from overlapping ERP responses. The processing of rewards involves a network of brain regions that includes not only the striatum but also the amygdala, medial prefrontal cortex, and orbitofrontal cortex [McClure et al., 2004], each of which may contribute to scalp-recorded EEG activity. Although the current results suggest that the FN may be generated in part by reward-related striatal activity, this is not to say that the striatum is the only brain region that contributes to the observed ERP response during reward processing. In fact, overlap between the FN and other ERP components, particularly the P300, is problematic for the use of source localization procedures [Nieuwenhuis et al., 2004]. Although the two components share certain properties, such as both being sensitive to violations of outcome expectancies [Hajcak et al., 2005; Hajcak et al., 2007], they also appear to capture distinct aspects of reward processing. This is reflected by their divergent scalp distributions as well as evidence that the P300, but not the FN, is sensitive to reward magnitude [Goyer et al., 2008; Sato et al., 2005; Yeung and Sanfey, 2004]. Further, the current data suggest that the P300 is insensitive to outcome valence. As such, applying source localization directly to an FN-P300 waveform may require the use of multiple dipoles to achieve adequate fit of the data, and it runs the risk of yielding a source that confounds these two responses. This situation is not unique to the FN and the P300, and it has been demonstrated that first applying PCA to ERP data may improve the accuracy of source localization when component overlap is a concern [Dien, 2010b; Dien et al., 2003]. Three additional PCA components consistent with the P200, P300, and Slow Wave were examined here, although none significantly differentiated rewards from nonrewards. Therefore, to the extent that we minimized the influence of overlapping responses on our measure of the FN, the current source localization results may represent a more accurate estimate of the neural generator of the FN. That being said, any application of source localization techniques cannot completely rule out alternative sources located along the axis of the dipole [Scherg, 1990].



For example, the source solution identified here is also consistent with a larger convex source that is located closer to the surface, including a generator within the anterior cingulate cortex. Alternatively, it is also possible that activity from both the anterior cingulate cortex and the striatum contribute to influence the scalp-recorded ERP waveform observed in response to feedback, and that the prominence of one source over another depends on the demands of a particular laboratory task, as well as the data analytic methods used. One way to more definitively rule out these other possibilities would be to apply EEG and functional magnetic resonance imaging techniques in the same sample and using the same paradigm. In this way, the scalp-recorded FN could be examined alongside hemodynamic measurements of reward-related activity in the putamen, anterior cingulate cortex, and other brain regions; this is a direction that we are currently pursuing.

In sum, this study extends the existing ERP literature on feedback processing by demonstrating that the FN might better be conceptualized as a frontocentral positivity that is enhanced for rewards compared with nonrewards, and that the relative increase in positivity for rewards compared with nonrewards is generated in part by the putamen. Temporospatial PCA was used to parse the ERP waveform into its underlying components, thereby separating the FN from the P300 and other overlapping responses and improving the accuracy of the source estimation. As with any attempt to identify the neural generator of an ERP response, however, these results must be interpreted with caution; even under ideal circumstances, source localization results are associated with an error on the order of 5–10 mm [Dien, 2010b]. If confirmed by future studies, these results indicate that the FN may be used as an ERP correlate of reward-related striatal activity. In this way, the FN may enhance our understanding of the precise role of the basal ganglia within the cascade of neurocognitive responses involved in the processing of rewards versus nonrewards.

## REFERENCES

- Attal Y, Bhattacharjee M, Yelnik J, Cottureau B, Lefèvre J, Okada Y, Bardinet E, Chupin M, Baillet S (2009): Modelling and detecting deep brain activity with MEG and EEG. *IRBM* 30:133–138.
- Bush G, Vogt BA, Holmes J, Dale AM, Greve D, Jenike MA, Rosen BR (2002): Dorsal anterior cingulate cortex: A role in reward-based decision making. *Proc Natl Acad Sci U S A* 99:523–528.
- Cattell RB (1966): The scree test for the number of factors. *Multivariate Behav Res* 1:245–276.
- Courchesne E, Hillyard SA, Courchesne RY (1977): P3 waves to the discrimination of targets in homogeneous and heterogeneous stimulus sequences. *Psychophysiology* 14:590–597.
- Delgado MR (2007): Reward-related responses in the human striatum. *Ann N Y Acad Sci* 1104:70–88.
- Dien J (2010a) The ERP PCA Toolkit: An open source program for advanced statistical analysis of event-related potential data. *J Neurosci Methods* 187:138–145.
- Dien J (2010b) Evaluating two-step PCA of ERP data with Geomin, Infomax, Oblimin, Promax, and Varimax rotations. *Psychophysiology* 47:170–183.
- Dien J, Frishkoff G (2005): Principal components analysis of event-related potential datasets. In: Handy TC, editor. *Event-Related Potentials: A Methods Handbook*. Cambridge, MA: The MIT Press. pp 189–208.
- Dien J, Beal DJ, Berg P (2005): Optimizing principal components analysis of event-related potentials: matrix type, factor loading weighting, extraction, and rotations. *Clin Neurophysiol* 116:1808–1825.
- Dien J, Khoe W, Mangun GR (2007): Evaluation of PCA and ICA of simulated ERPs: Promax vs. Infomax rotations. *Hum Brain Mapp* 28:742–763.
- Dien J, Spencer KM, Donchin E (2003): Localization of the event-related potential novelty response as defined by principal components analysis. *Brain Res Cogn Brain Res* 17:637–650.
- Donchin E, Heffley E (1979): Multivariate analysis of event-related potential data: A tutorial review. In D.Otto (Ed.), *Multidisciplinary perspectives in event-related potential research*. Washington, DC: U.S. Government Printing Office.
- Duncan-Johnson CC, Donchin E (1977): On quantifying surprise: the variation of event-related potentials with subjective probability. *Psychophysiology* 14:456–467.
- Dunning JP, Hajcak G (2007): Error-related negativities elicited by monetary loss and cues that predict loss. *Neuroreport* 18:1875–1878.
- Elliott R, Newman JL, Longe OA, Deakin JF (2003): Differential response patterns in the striatum and orbitofrontal cortex to financial reward in humans: a parametric functional magnetic resonance imaging study. *J Neurosci* 23:303–307.
- Folstein JR, Van Petten C (2008): Influence of cognitive control and mismatch on the N2 component of the ERP: A review. *Psychophysiology* 45:152–170.
- Foti D, Hajcak G (2009): Depression and reduced sensitivity to non-rewards versus rewards: Evidence from event-related potentials. *Biol Psychol* 81:1–8.
- Foti D, Hajcak G (2010): State sadness reduces neural sensitivity to non-rewards versus rewards. *Neuroreport* 21:143–147.
- Gehring WJ, Willoughby AR (2002): The medial frontal cortex and the rapid processing of monetary gains and losses. *Science* 295:2279–2282.
- Goyer JP, Woldorff MG, Huettel SA (2008): Rapid electrophysiological brain responses are influenced by both valence and magnitude of monetary rewards. *J Cogn Neurosci* 20:2058–2069.
- Gratton G, Coles MG, Donchin E (1983): A new method for off-line removal of ocular artifact. *Electroencephalogr Clin Neurophysiol* 55:468–484.
- Hajcak G, Holroyd CB, Moser JS, Simons RF (2005): Brain potentials associated with expected and unexpected good and bad outcomes. *Psychophysiology* 42:161–170.
- Hajcak G, Moser JS, Holroyd CB, Simons RF (2006): The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. *Biol Psychol* 71:148–154.
- Hajcak G, Moser JS, Holroyd CB, Simons RF (2007): It's worse than you thought: the feedback negativity and violations of reward prediction in gambling tasks. *Psychophysiology* 44:905–912.
- Haruno M, Kawato M (2006): Different neural correlates of reward expectation and reward expectation error in the putamen and caudate nucleus during stimulus-action-reward association learning. *J Neurophysiol* 95:948.

- Holroyd CB (2004): A note on the oddball N200 and the feedback ERN. In: Ullsberger M, Falkenstein M editors. *Errors, Conflicts, and the Brain: Current Opinions on Response Monitoring*. Leipzig: MPI of Cognitive Neuroscience.
- Holroyd CB, Coles MG (2002): The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychol Rev* 109:679–709.
- Holroyd CB, Hajcak G, Larsen JT (2006): The good, the bad and the neutral: electrophysiological responses to feedback stimuli. *Brain Res* 1105:93–101.
- Holroyd CB, Larsen JT, Cohen JD (2004): Context dependence of the event-related brain potential associated with reward and punishment. *Psychophysiology* 41:245–253.
- Holroyd CB, Nieuwenhuis S, Yeung N, Cohen JD (2003): Errors in reward prediction are reflected in the event-related brain potential. *Neuroreport* 14:2481–2484.
- Holroyd CB, Pakzad-Vaezi KL, Krigolson OE (2008): The feedback correct-related positivity: sensitivity of the event-related brain potential to unexpected positive feedback. *Psychophysiology* 45:688–697.
- Johnson R Jr, Donchin E (1980): P300 and stimulus categorization: two plus one is not so different from one plus one. *Psychophysiology* 17:167–178.
- Kirsch P, Schienle A, Stark R, Sammer G, Blecker C, Walter B, Ott U, Burkart J, Vaitl D (2003): Anticipation of reward in a nonaversive differential conditioning paradigm and the brain reward system: An event-related fMRI study. *Neuroimage* 20:1086–1095.
- Knutson B, Westdorp A, Kaiser E, Hommer D (2000): FMRI visualization of brain activity during a monetary incentive delay task. *Neuroimage* 12:20–27.
- Lorente de No R (1947): Action potential of the motoneurons of the hypoglossus nucleus. *J Cell Physiol* 29:207–287.
- Luck SJ (2005): *An introduction to the event-related potential technique*. Cambridge, MA: MIT Press.
- Lutkenhoner B (1998): Dipole source localization by means of maximum likelihood estimation I. Theory and simulations. *Electroencephalogr Clin Neurophysiol* 106:314–321.
- McClure SM, Berns GS, Montague PR (2003): Temporal prediction errors in a passive learning task activate human striatum. *Neuron* 38:339–346.
- McClure SM, York MK, Montague PR (2004): The neural substrates of reward processing in humans: the modern role of FMRI. *Neuroscientist* 10:260–268.
- Miller J, Patterson T, Ulrich R (1998): Jackknife-based method for measuring LRP onset latency differences. *Psychophysiology* 35:99–115.
- Miltner WH, Braun CH, Coles MGH (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.
- Murakami S, Okada Y (2006): Contributions of principal neocortical neurons to magnetoencephalography and electroencephalography signals. *J Physiol* 575(Pt 3):925–936.
- Nieuwenhuis S, Heslenfeld DJ, von Geusau NJ, Mars RB, Holroyd CB, Yeung N (2005): Activity in human reward-sensitive brain areas is strongly context dependent. *Neuroimage* 25:1302–1309.
- Nieuwenhuis S, Holroyd CB, Mol N, Coles MG (2004): Reinforcement-related brain potentials from medial frontal cortex: Origins and functional significance. *Neurosci Biobehav Rev* 28:441–448.
- O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ (2003): Temporal difference models and reward-related learning in the human brain. *Neuron* 38:329–337.
- Olvet DM, Hajcak G (2009): The stability of error-related brain activity with increasing trials. *Psychophysiology* 46:957–961.
- Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD (2006): Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* 442:1042–1045.
- Rektor I (2002): Scalp-recorded Bereitschaftspotential is the result of the activity of cortical and subcortical generators—A hypothesis. *Clin Neurophysiol* 113:1998–2005.
- Sander TH, Knosche TR, Schlogl A, Kohl F, Wolters CH, Hauelsen J, Trahms L (2010): Recent advances in modeling and analysis of bioelectric and biomagnetic sources. *Biomed Tech (Berl)* 55:65–76.
- 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.
- Scherg M (1990): Fundamentals of dipole source analysis. In: Grandori F, Romani GL, editors. *Auditory Evoked Magnetic Fields and Potentials*. New York: Karger. pp 1–30.
- Schultz W (2002): Getting formal with dopamine and reward. *Neuron* 36:241–263.
- Spencer KM, Dien J, Donchin E (2001): Spatiotemporal analysis of the late ERP responses to deviant stimuli. *Psychophysiology* 38:343–358.
- Tobler PN, O'Doherty JP, Dolan RJ, Schultz W (2006): Human neural learning depends on reward prediction errors in the blocking paradigm. *J Neurophysiol* 95:301.
- Yeung N, Sanfey AG (2004): Independent coding of reward magnitude and valence in the human brain. *J Neurosci* 24:6258–6264.
- Zhang Z, Jewett DL, Goodwill G (1994): Insidious errors in dipole parameters due to shell model misspecification using multiple time-points. *Brain Topogr* 6:283–298.