

Special issue: Original article

Dorsal anterior cingulate cortex integrates reinforcement history to guide voluntary behavior

Clay B. Holroyd^{*a*,*} and Michael G.H. Coles^{*b*}

^aUniversity of Victoria, Victoria, Canada ^bDonders Centre for Cognitive Neuroimaging, Nijmegen, The Netherlands

ARTICLE INFO

Article history: Received 31 January 2007 Reviewed 26 April 2007 Revised 30 May 2007 Accepted 29 August 2007 Published online 23 December 2007

Keywords: Anterior cingulate cortex Reinforcement Error-related negativity Prediction error Decision making

ABSTRACT

Two competing types of theory have been proposed about the function of dorsal anterior cingulate cortex (dACC): evaluative theories hold that dACC monitors ongoing behavior to detect errors or conflict, whereas response selection theories hold that dACC is directly involved in the decision making process. In particular, one response selection theory proposes that dACC utilizes reward prediction error signals carried by the midbrain dopamine system to decide which of several competing motor control systems should be given control over the motor system (Holroyd and Coles, 2002). The theory further proposes that the impact of these dopamine signals on dACC determines the amplitude of a component of the event-related brain potential called the error-related negativity (ERN). In the present study, we applied this theory to a decision making problem that requires participants to select between two response options in which an erroneous choice is not clearly defined. Rather, the reward received for a particular response evolves in relation to the individual's previous behavior. We adapted a computational model associated with the theory to simulate human performance and the ERN in the task, and tested the predictions of the model against empirical ERP data. Our results indicate that ERN amplitude reflects the subjective value attributed by each participant to their response options as derived from their recent reward history. This finding is consistent with the position that dACC integrates the recent history of reinforcements to guide voluntary choice behavior, as opposed to evaluating behaviors per se.

© 2007 Elsevier Masson Srl. All rights reserved.

1. Introduction

The ability to navigate and adapt to uncertain environments is mediated partly by dorsal anterior cingulate cortex (dACC), a region of medial frontal cortex believed to contribute to cognitive control (Ridderinkhof et al., 2004). At least two classes of theory have been proposed about the function of dACC. The first, motivated largely by event-related brain potential (ERP) and hemodynamic neuroimaging data, holds that dACC monitors ongoing performance to detect errors (Coles et al., 1998; Dehaene et al., 1994) or conflict (Botvinick et al., 2001, 2004; Yeung et al., 2004). Such *evaluative theories* propose that dACC is not directly involved in response selection but rather determines the success of ongoing behavior. In contrast, *response selection theories*, motivated largely by neurophysiological and neuroanatomical data, propose that dACC is directly involved

^{*} Corresponding author. Department of Psychology, University of Victoria, P.O. Box 3050 STN CSC, Victoria, Canada BC V8W 3P5. E-mail address: holroyd@uvic.ca (C.B. Holroyd).

^{0010-9452/\$ –} see front matter \odot 2007 Elsevier Masson Srl. All rights reserved. doi:10.1016/j.cortex.2007.08.013

in the decision making process (Holroyd and Coles, 2002; Holroyd et al., 2004b; Paus, 2001; Rushworth et al., 2004; Walton et al., 2004). For example, we have suggested that dACC may serve as a "motor control filter" that decides which of several competing motor control systems should be given responsibility over motor behavior (Holroyd and Coles, 2002). According to this view, dACC receives reward prediction error signals from the midbrain dopamine system called temporal difference errors (Schultz, 2002), which it utilizes for the flexible adaptation of behavior according to principles of reinforcement learning (Montague et al., 2004; Sutton and Barto, 1998). Note that in one respect, these two types of theories are antithetical to each other: whereas evaluative theories hold that performance monitoring occurs within dACC, response selection theories hold that this function must occur elsewhere in the brain, such as in the basal ganglia (Holroyd and Yeung, 2003; Holroyd et al., 2005).

We have further proposed that the impact of the temporal difference error signals on dACC determines the amplitude of an ERP component called the error-related negativity (ERN; Holroyd and Coles, 2002). The ERN appears to exist in two varieties: a response ERN and a feedback ERN (Holroyd et al., 2004b). The response ERN is seen when human participants commit errors in speeded response time tasks, peaking about 80 msec following the onset of the incorrect response (Falkenstein et al., 1990; Gehring et al., 1993). The feedback ERN is elicited when a feedback stimulus indicates to the participant that a task goal was not achieved, such as a previous response being incorrect (Holroyd and Coles, 2002; Miltner et al., 1997; Nieuwenhuis et al., 2004a). Both the response ERN (Dehaene et al., 1994; Debener et al., 2005; Holroyd et al., 1998) and the feedback ERN (Gehring and Willoughby, 2002a; Holroyd et al., 2004b; Miltner et al., 1997) appear to be generated in dACC, although the source of the feedback ERN is somewhat controversial (Nieuwenhuis et al., 2005). According to this "reinforcement learning theory of the ERN" (RL-ERN theory), the amplitude of the ERN should reflect the temporal difference error properties associated with the midbrain dopamine signal (Holroyd and Coles, 2002). This prediction has been confirmed in several experiments (e.g., Baker and Holroyd, 2006; Donkers and van Boxtel, 2005; Gibson et al., 2006; Holroyd and Coles, 2002; Holroyd et al., 2003, 2005; Potts et al., 2006; Holroyd and Krigolson, 2007; see also Holroyd et al., 2004c; Mars et al., 2005).

Although the RL-ERN theory is a response selection theory of dACC, it is sometimes perceived to be an evaluative theory (e.g., Walton et al., 2004). For example, Kennerley et al. (2006) have understood the theory to mean that dACC function involves "monitoring whether a single action achieved its expected outcome or in signaling the need for adaptive behavior", which they distinguish from "integrating reinforcement information over time to guide voluntary choice behavior". This misunderstanding may stem from the ambiguous meaning of the word "error", which has been interpreted differently by different investigators. For instance, Gehring and Willoughby (2002a, 2002b) have argued that the feedback ERN, and by extension dACC, is sensitive to reward value rather than to errors per se, but we have shown that the system that produces the feedback ERN can flexibly adapt to the current task context, however so defined (Holroyd et al., 2004a). Thus, whether dACC functions to maximize reward or to avoid errors

depends on how the participant perceives the goal of the task (Nieuwenhuis et al., 2004b; see also Holroyd et al., 2002); if the goal is to maximize reward, then a failure to do so constitutes an error by definition. Interestingly, the system that produces the ERN appears to categorize outcomes in a binary manner: as events that either do, or do not, indicate that the task goal has been achieved (Hajcak et al., 2006; Holroyd et al., 2006; Toyomaki and Murohashi, 2005; Yeung and Sanfey, 2004).

In the present experiment, we show that dACC can integrate reinforcement information over time to guide voluntary choice behavior. To do so, we apply the RL-ERN theory to a decision making problem that requires participants to select between response options in which an erroneous choice is not clearly defined. In this task, developed by Herrnstein (1997a), the amount of reward that an individual receives depends on the history of his or her past responses. Thus, the reward received for a particular response is not stationary, but rather evolves in relation to the individual's previous behavior. Specifically, the reward is delivered as a function of the ratio of two button choices made by the participant over a set of previous responses (e.g., the previous 40 trials). Typically, the reward schedule is designed such that participants respond sub-optimally: unbeknownst to them, they would receive more rewards if they responded by pressing the two buttons at a different ratio. In general, participants tend to respond at the "matching point" of the reward function, where they obtain about an equal amount of reward from both button presses, rather than at other points of the reward function where they would obtain more reward overall, but where the rewards associated with the two responses are unequal.

We adapted a computational model associated with the RL-ERN theory (Holroyd and Coles, 2002; see also Holroyd et al., 2005; Nieuwenhuis et al., 2002) to simulate human performance, the response ERN and the feedback ERN in a Herrnstein matching task, and tested these predictions by collecting ERPs from participants engaged in the task (cf. Egelman et al., 1998). The free parameters of the model were determined by fitting the model's output to the empirically observed behavioral data, and the resulting simulated ERNs served as predictions for the empirical ERNs. We predicted that the amplitudes of the response ERN and the feedback ERN would track the value of the response functions associated with the participants' response ratios. That is, we predicted that the amplitudes of the response ERN and feedback ERN would reflect the subjective value attributed by each participant to their response options, as derived from their recent reward history. We expected that the ERN would not index the value itself, but would instead index the temporal difference of the value (Holroyd and Coles, 2002). In general, our experiment illustrates the role of dACC in integrating reinforcement history to guide voluntary behavior, as specified by the RL-ERN theory.

2. Methods

2.1. Reward functions

The task adopted in this study is derived from a class of twochoice decision making games developed by Herrnstein (1997a). On each trial, participants made one of two possible responses and were provided with feedback (the "reward"). The value of this feedback was determined by (1) the proportion of the previous 40 responses made with either hand, (2) the hand chosen on that trial, (3) a particular reward function designated for that block of trials, and (4) a small random element.¹ Fig. 1a and b illustrates the reward functions utilized in this experiment. In each case, the abscissa refers to the proportion of the previous 40 responses made with "hand 1". For example, a value of 1 indicates that all of the previous 40 responses have been made with hand 1 (and none with hand 2), whereas a value of 0 indicates that none of the previous 40 responses have been made with hand 1 (and all with hand 2). Note that responses made with hand 1 move the participant rightward along the figures, whereas responses made with hand 2 move the participant leftward along the figures. For both figures, the functions describe the amount of reward delivered for a response generated with either hand at a given proportion: solid lines correspond to hand 1, dashed lines to hand 2, and dotted lines to the average reward received at that proportion. In Function 1, note also that (a) participants receive the maximum average reward when zero of the 40 previous responses is made with hand 1 (i.e., all are made with hand 2); and (b) both responses elicit the same amount of reward when about 75% of the previous 40 responses were made with hand 1 (Fig. 1a). Furthermore, in Function 2, note that the maximum average reward occurs when the previous 40 responses are all made with hand 1 (Fig. 1b).

The experiment consisted of 40 blocks of 100 trials. Each block of trials was associated with a particular function, and each of the two functions occurred on 20 blocks of trials. Furthermore, the functions were counterbalanced across hands, such that hand 1 in both functions was mapped to the left hand on 10 blocks and to the right hand on 10 blocks. Lastly, at the start of each block the response history over the previous 40 (non-existent) trials was initialized by randomly selecting, without replacement, the number of hand 1 trials from one of 10 possible values (0, 4, 8, 13, 17, 22, 26, 31, 35, and 40). Thus, the 40 blocks are described by two functions, two hands, and 10 starting proportions; on each block a particular combination was selected at random, without replacement. The random element added to the reward on each trial was a randomly determined integer whose maximum absolute value was the minimum reward associated with that function, i.e., 10 in Function 1 and 15 in Function 2.

3. Experiment

3.1. Participants

Ten participants, ages 18–22, were paid \$5.00 per hour for their participation, plus a bonus which depended on performance. The experiment was performed in compliance with

institutional review board guidelines at the University of Illinois, Urbana-Champaign. Each session was completed in approximately 2 h.

3.2. Task

The participants sat comfortably about 1 m from a computer display in an electromagnetically shielded room. Participants were instructed that they would play a kind of "guessing game" in which their objective was to maximize their longterm reward using the scores they saw on the computer screen. On each trial (1.5 sec), participants were exposed to a warning cue (appearing as '+' on the computer screen) for 800 msec. During the first 600 msec of this period, participants were required to press one of two buttons with either the left or the right hand. At the end of the 800 msec interval, the warning cue was replaced by a feedback stimulus, which remained on the screen for 200 msec. The feedback stimulus could be of three types. If participants responded before the warning stimulus appeared, the word "EARLY" appeared, and no reward was recorded. If participants responded after the 600 msec limit had elapsed, the word "LATE" appeared, and no reward was recorded. If participants responded within that 600 msec period, then a numerical score was displayed, as determined from the reward function associated with that block of trials (see above). Participants were not informed of the nature of the reward distributions and had to infer the appropriate response strategies by trial-and-error. Stimuli measured about 2° across the visual field. Following each block of trials, the total amount of reward up to that point in the experiment was displayed in cents on the computer screen. This total reward was determined by summing the scores associated with Function 1 and dividing the result by 30,000, by summing the scores associated with Function 2 and dividing the result by 50,000, and by adding the results of the two calculations together. Dividing by these numbers ensured that the participants earned approximately \$6 in bonus money at the end of each session. Participants initiated each subsequent block of trials at their leisure by pressing a (third) button.

3.3. Data acquisition

A 128 channel system from Electrical Geodesics, Inc. (Eugene, Oregon; Tucker, 1993) was used with Netstation software for data acquisition and experimental control. Scalp impedances were less than 40 K. Electrodes were re-wetted during a 15 min break half way through the experiment. The sampling rate was 250 Hz, which for each 1.5 sec trial yielded 375 time points per channel.

3.4. Data analysis

The electroencephalogram data were segmented with Netstation and analyzed using in-house software. Ocular artifact was removed with the eye-movement correction procedure described by Gratton et al. (1983). Data were re-referenced off-line to link-mastoid electrodes (where channels 57 and 101 were taken as the left and right mastoids, respectively). Single trial ERP data were baseline corrected on the basis of

¹ This departs from the standard version of the task in which the reward schedule is entirely deterministic. Participants vary widely in their performance on the standard version (Herrnstein, 1997a), with some participants maximizing reward and others exhibiting matching behavior (Montague and Berns, 2002). We added a small random value to each reward to increase task difficulty, thereby ensuring that most participants would match. Note also that, unlike some matching tasks, this task did not implement a "changeover delay".



Fig. 1 – Reward functions and behavioral data. (a–f) Abscissa: proportion of the previous 40 responses made with hand 1. (a, b) Reward distributions: "Function 1" (a) and "Function 2" (b). Ordinate: amount of reward. Solid lines: average reward delivered for responses made with hand 1 at each proportion. Dashed lines: average reward delivered for responses made with hand 2 at each proportion. Dotted lines: expected amount of reward, averaged across hands, for responses made at each proportion. (c, d) Histograms of empirical behavioral data, pooled across subjects, for Function 1 (c) and Function 2 (d). (e, f) Histograms of simulated behavioral data, pooled across subjects, for Function 1 (e) and Function 2 (f). (c–f) Solid lines: number of responses made with hand 1 at each proportion. Dashed lines: number of responses made with hand 2 at each proportion. Dotted lines: total number of responses, pooled across hands, made at each proportion. Vertical lines in (a–f) indicate regions of interest for data analysis (see Section 2).

200 msec interval preceding stimulus onset, and were filtered with a passband of 1-20 Hz using the Interactive Data Language digital filter algorithm. For each participant, average ERPs were derived separately for each function by averaging all trials associated with that function across blocks and across hands. Separate "stimulus-locked" and "responselocked" averages were created by averaging the trials according to stimulus onset and response onset, respectively. Grand averages were derived by averaging these ERPs across participants. Furthermore, for each function, histograms of trials were created for responses made with either hand, as a function of the proportion of the previous 40 responses that were made with hand 1. For example, if on 10 trials for Function 1 a participant responded with hand 1 and the previous 40 responses were all made with hand 2, then the associated histogram (for hand 1, Function 1) would have a value of 10 at the bin corresponding to proportion zero. For both functions, regions of interest associated with the most trials were then determined. In Function 1, this area spanned where 60-85% of the previous 40 responses were generated with hand 1 (Fig. 1c); in Function 2, the area spanned where 0-25% of the previous 40 responses were made with hand 1 (Fig. 1d). For each participant, response-locked and feedback-locked ERPs were derived for both of these regions by averaging together the trials within that region according to hand. Thus, with 10 participants, two functions, two hands, two averaging methods (response, feedback), and 128 channels, this procedure yielded 10,240 ERPs.

We applied a spatial principal components analysis (PCA) to summarize activity that co-varied across electrodes (see Donchin and Heffley, 1978 for an introduction to PCA and its temporal application to the ERP). This procedure is commonly used to identify spatial factors that correspond to particular ERP components (e.g., Spencer et al., 2001) including the ERN (Holroyd et al., 1998; Krigolson and Holroyd, 2006). We restricted our analysis to time periods associated with the ERN, namely, between 0 and 100 msec following the response in the response-locked waveforms (Falkenstein et al., 1990; Gehring et al., 1993), and between 150 and 300 msec following presentation of the feedback stimulus in the feedback-locked waveforms (Holroyd and Coles, 2002; Miltner et al., 1997). Thus, the spatial PCA was conducted on the data associated with these periods across the 10,240 ERPs described above. This procedure produced 10 factors, each associated with 80 sets of factor scores corresponding to the two temporal periods of interest; the spatial factor scores can be thought of as "virtual ERPs" that describe the time course of the ERP component associated with that factor (Spencer et al., 2001). The ERN was identified with a factor that exhibited (1) factor loadings with a scalp distribution consistent with the ERN (frontalcentral), and (2) factor scores (virtual ERNs) that behaved like the ERPs recorded at frontal-central scalp areas.

As of yet there is no consensus about how best to measure ERN amplitude. According to the RL-ERN theory, the ERN is associated with the impact of negative temporal difference errors on dACC, but it is also possible that positive temporal difference errors on dACC would elicit a positivity in the time range of the ERN (Holroyd and Coles, 2002; Holroyd, 2004; Pakzad-Vaezi et al., 2006; see also Potts et al., 2006). Our approach has been to compare differences in ERN amplitude across conditions, for example, by comparing the difference associated with unexpected rewards and punishments with the difference associated with expected rewards and punishments (e.g., Holroyd et al., 2003; Holroyd, 2004; Holroyd and Krigolson, 2007). This procedure allows for our ERN measure to capture neural activity associated with both positive and negative temporal difference errors. In the present study, we were interested in isolating the relative value attributed by participants to each response hand. To do so, we first integrated the areas under the curves described by the virtual ERNs (the factors' scores associated with the ERN factor); this step is analogous to finding the average value of the ERPs recorded at frontal-central scalp sites within a specified time window. Then, we computed the difference between these measures for each hand in each condition (i.e., hand 1-hand 2), for both the response ERN and feedback ERN. The difference measures provide an index of the relative size and sign of the temporal difference errors associated with responses produced by either hand (see below): positive values indicate that the system evaluates the eliciting event as better for hand 1 than for hand 2, whereas negative values indicate that the system evaluates the eliciting event as better for hand 2 than for hand 1. Note that the PCA minimizes possible contamination of the ERN measure from other overlapping ERP components (Spencer et al., 2001).

3.5. Simulation

Our simulation combined aspects of our previous model of the ERN (Holroyd and Coles, 2002) with Egelman et al.'s (1998) temporal difference model of human behavior on the Herrnstein task. For statistical purposes, we ran the present model 10 times, simulating data for 10 participants. Each trial was divided into 60 time steps, with presentation of the warning cue at time step 10, response generation at time step 30, and feedback presentation at time step 60. Stimulus input and response output were sent to a "value" layer that identified states v^i , where i corresponds to the state and v to that of state's activation: when the warning cue was on, when a left response had been generated, and when a right response had been generated (Fig. 2). The activation of these units was set to 0 (minimum) when inactivated, and to 1 (maximum) when activated. Presentation of the warning cue activated the cue unit in the value layer, and this unit remained activated until a response was generated. Response generation activated the associated unit in the value layer, and this unit remained activated until the reward was delivered.²

At every time t, the value of the state of the system was determined as

$$\widehat{V}_t = \sum_{i=1}^3 z_t^i \upsilon_t^i$$

where \hat{V}_t corresponds to a weight associated with each unit. Thus, the value at any time t was simply the weight associated with state i. At each time step t, furthermore, a temporal difference error δ_t was computed as

² Note that the response unit remained activated even after the response had been generated overtly, and thus maintained a memory of the response event.



Fig. 2 – Computational model. "Cue", "Left", and "Right": value layer units associated with activation states v1, v2 and v3, respectively, and weights z1, z2, and z3, respectively. "Feedback": trial score associated with the reward r. "TD": temporal difference unit that outputs the temporal difference error δ . "P": Softmax function unit that outputs the selected response. Note that the temporal difference error was used to update the value layer weights z1-z3.

$\delta_t = \widehat{V}_t + r_t - \widehat{V}_{t-1}$

where r_t is the magnitude of any feedback delivered to the system at that time step, as determined by the reward functions in Fig. 1a and b (see above; for convenience, the magnitudes of these rewards were scaled down by a factor of 10 for the simulation). The weights were initialized with small random values between .00 and .01, and learning occurred on each time step according to

$$\mathbf{z}_{t+1}^{\iota} = \mathbf{z}_{t}^{\iota} + \eta \delta_{t+1} \boldsymbol{v}_{t}^{\iota}$$

where η is the learning rate.

Response selection at time step 30 was determined according to a "softmax" probability function (Sutton and Barto, 1998):

 $P(\text{response } k \text{ selected}) = \frac{e^{z_{t}^{k}/\tau}}{\sum_{k=1}^{2} e^{z_{t}^{k}/\tau}}$

Here z_t^k corresponds to the weight associated with hand 1 or hand 2 at time t, and τ is a "temperature" parameter that controls the degree of exploration.

Values for η and τ were determined separately for both reward functions, using an optimization procedure that identified optimal values for both parameters. We used a reinforcement learning algorithm that identified values for η and τ through successive approximation. Each step of the optimization procedure consisted of choosing values for η and τ , running the simulation through a block of trials, and determining the discrepancy between the empirical and the simulated behavioral distributions (see Fig. 1). Possible values for η and τ initially ranged between 0 and 1, but the range was restricted to the most relevant areas as optimization progressed. For each function, the procedure was run for 48,000 blocks, yielding $\eta = .875$ and $\tau = .225$ for Function 1 and $\eta = .025$ and $\tau = .1375$ for Function 2.³

As with the empirical data, the simulated ERNs were measured in terms of differences, i.e.,

 $\text{ERN} = \delta_t(\text{hand 1}) - \delta_t(\text{hand 2})$

Note that the temporal difference error δ_t reflects the difference in value across successive states: positive temporal difference errors indicate that ongoing events are "better than expected", whereas negative temporal difference errors indicate that ongoing events are "worse than expected" (Sutton, 1988). Thus, for the response ERN, positive ERN values indicate that the system evaluates hand 1 responses as better than hand 2 responses, and conversely, negative values indicate that the system evaluates hand 2 responses as better than hand 1 responses. Likewise, for the feedback ERN, positive ERN values indicate that the feedback was relatively better than expected following hand 1 responses compared to hand 2 responses, and negative values indicate that the feedback was relatively better than expected following hand 2 responses compared to hand 1 responses. Note that the exploration parameter τ ensured that the simulation would occasionally produce responses associated with the lesser value.

4. Results

Fig. 1c and d illustrates histograms of the empirical behavioral data for Function 1 and Function 2, respectively, pooled across participants. For each function, participants' responses gravitated to parts of the reward distribution where the amount of reward obtained was less than the optimum. In Function 1, the vast majority of responses occurred near proportion = .75 (Fig. 1c; average proportion associated with the final 40 trials = 64.4% \pm 6.5%), where both response options elicit the same amount of reward (Fig. 1a). Note that, had participants responded exclusively with hand 2 in Function 1, they would have received the maximum reward possible. Indeed, the participants would have received more reward than they actually did had they responded at chance, producing half their responses with either hand. A small fraction of all the responses was in fact generated at the reward maximum; we observed this outcome to occur on blocks in which the participants' proportions were initialized with values close to zero.

³ This optimization procedure involved only behavioral data. Thus, the simulated ERNs in this study were not produced by discovering parameters that fit the simulated ERNs to the observed ERNs, but rather were an emergent property of the model itself. Note that any optimization procedure could have been used to find values for η and τ ; we chose a reinforcement learning approach for convenience.

In Function 2, nearly all responses were made with hand 2 (Fig. 1d; average proportion associated with the final 40 trials = $13.5\% \pm 13.9\%$), but participants would have maximized their reward had they responded exclusively with hand 1 (Fig. 1b). These observations are consistent with the results of previous studies (Egelman et al., 1998; Herrnstein, 1997a).

Fig. 1e and f depicts the simulated behavioral data. As can be seen, the behavior of the simulated participants was qualitatively similar to that of the real participants. In particular, nearly 75% of the responses were made with hand 1 in Function 1 (Fig. 1e), and nearly all responses were made with hand 2 in Function 2 (Fig. 1f). These results are not surprising given that the model parameters were optimized so as to fit the simulated behavioral data to the empirical behavioral data.

We restricted our ERP analysis to data associated with those parts of the reward functions where most responses were generated (see Section 2). Fig. 3 illustrates the ERPs associated with both functions, averaged over all trials and both hands, recorded at a frontal-central electrode location where the ERN reaches maximum amplitude (Falkenstein et al., 1990; Gehring et al., 1993; Miltner et al., 1997). The thin lines correspond to trials associated with Function 1 (solid) and Function 2 (dashed). Fig. 3a illustrates ERPs averaged according to the onset of the response at 0 msec ("response-locked" ERPs). Fig. 3b illustrates the development of the ERPs across the 1.5 sec trial interval, averaged according to the onset of the imperative stimulus ("stimulus-locked" ERPs). In this figure, the imperative and feedback stimuli are presented at 0 and 800 msec, respectively. In both figures, the "virtual ERNs" (i.e., the factor scores) are plotted over the ERPs during the ERN time windows (thick solid lines). The close correspondence of the virtual ERNs with the actual ERPs indicates that the PCA successfully identified ERN-related activity; for this reason we hereafter call the factor scores associated with this factor the ERN.

Note that when averaged across trials and across hands, the ERPs associated with Functions 1 and 2 behave in a comparable fashion (Fig. 3). However, the ERNs depart from one another when averaged according to hand, as illustrated in Fig. 4 (average number of trials per ERP: Function 1 - hand 1: 894, hand 2: 334; Function 2 - hand 1: 73, hand 2: 1079). The figure shows the response-locked (top row) and feedback-locked (bottom row) ERNs, for both hands and both functions, during the ERN time windows (negative is plotted up by convention). For the response-locked ERN, hand 1 (solid line) appears to be about the same magnitude as hand 2 (dashed line) in Function 1 (Fig. 4a), but hand 1 appears to be more negative than hand 2 in Function 2 (Fig. 4b). For the feedback-locked ERN, hand 1 appears to be more negative than hand 2 in Function 1 (Fig. 4c), and hand 1 appears to be more positive than hand 2 in Function 2 (Fig. 4d). The ERN measures, taken as the difference of these values (see Section 2), are plotted in Fig. 5 for both the empirical (Fig. 5a) and simulated (Fig. 5b) data. Recall that for the response ERN (solid lines in Fig. 5), positive ERN values indicate that the system evaluates hand 1 responses as better than hand 2 responses, and conversely, negative values indicate that the system evaluates hand 2 responses as better than hand 1 responses. Further, for the feedback ERN (dashed lines in Fig. 5), positive ERN values indicate that the feedback was relatively better than expected following hand 1



Fig. 3 – ERPs recorded frontal-centrally (channel 7 of the electrogeodesic dense electrode array; Tucker, 1993), averaged over all trials. (a) Response-locked ERPs; response onset occurs at .0 sec. (b) Stimulus-locked ERPs; stimulus onset occurs at .0 sec and feedback onset at .8 sec. Thin solid lines and thin dashed lines correspond to Function 1 and Function 2, respectively. Thick solid lines correspond to virtual ERNs (i.e., the spatial factor scores associated with the frontal-central spatial factor). Vertical lines centered around the virtual ERNs indicate time windows of interest (see Section 2). Negative is plotted up by convention.

responses compared to hand 2 responses, and negative values indicate that the feedback was relatively better than expected following hand 2 responses compared to hand 1 responses.

Insight into these results can be found by relating the response-locked ERNs to the behavioral data. For Function 1, the simulated response ERN difference measure was nearly zero (although hand 1 was slightly favored over hand 2; Fig. 5b: solid line, Function 1). This result predicts that participants should value both response options about equally in



Fig. 4 – ERN factor scores during the time windows of interest (see Section 2) for Function 1 (left column) and Function 2 (right column), associated with the response (top row) and feedback (bottom row). Solid and dashed lines correspond to choices made with hand 1 and hand 2, respectively. Values along the abscissa are relative to the time of the response (top row) and feedback onset (bottom row). Negative is plotted up by convention.

this region of the reward function. Indeed, Fig. 5a confirms that the response ERNs associated with each response option were nearly identical for Function 1 (Fig. 5a: solid line, Function 1; 95% confidence interval = -1.0, 2.8). Furthermore, a comparison of Fig. 1c with a indicates that most participants responded in Function 1 at a proportion where both choices obtained about equal amounts of reward. In contrast, for Function 2 the simulation predicts that the response ERN should be larger for hand 1 than for hand 2 (Fig. 5b: solid line, Function 2). These relationships were confirmed by the empirical data (Fig. 5a: solid line, Function 2; 95% confidence interval = -15.1, -3.2). These results suggest that for Function 2, hand 2 responses were perceived to be better by the participant than hand 1 responses. In fact, for all proportions in this function, hand 2 responses elicited more reward than hand 1 responses (Fig. 1b: dashed line > solid line for all proportions). Evidently for this reason participants responded almost exclusively with hand 2, which propelled them into the region of the reward function where minimal reward was actually obtained (Fig. 1d). Note that the response ERN difference

measure was more negative for Function 2 than for Function 1 (Fig. 5a, solid line; t = -3.3, p < .01), confirming that the relative value of hand 1 responses to hand 2 responses was greater (and more negative) for Function 2 than for Function 1. In other words, responses made with hand 2 were perceived to be better than response made with hand 1 for Function 2 but not for Function 1.

The feedback ERN difference measures were also consistent with the reinforcement learning framework. As with the response ERN measures, the feedback ERN measures for Function 1 were close to zero for both the simulated and empirical data (Fig. 5a and b: dashed lines, Function 1; 95% confidence interval = -11.8, 0.6). Evidently, this result obtained because the system expected and received equivalent amounts of reward from responses made with either hand, which at this response proportion (close to 75%, Fig. 1c) was in fact what occurred (Fig. 1a). In contrast, for Function 2 the simulation predicted that the feedback ERN should be more positive for hand 1 responses relative to hand 2 responses (Fig. 5b: dashed line, Function 2); this prediction was



Fig. 5 – Relative ERN values measured as ERN (hand 1) – ERN (hand 2). (a) Empirical data. (b) Simulated data. Solid lines: response-locked differences. Dashed lines: feedback-locked differences. Units for empirical data reflect integrated factor scores (see Section 2); units for simulated data reflect temporal difference error. Note that positive values indicate that the system evaluates the eliciting event as better for hand 1 than hand 2 (i.e., the ERN is smaller for hand 1 than hand 2), whereas negative values indicate that the system evaluates the eliciting event as better for hand 1 (i.e., the ERN is larger for hand 1 than hand 2).

confirmed by the empirical data (Fig. 5a; dashed line, Function 2; confidence interval = 3.1, 21.3). According to the theory, the outcomes of hand 1 responses were relatively better than expected compared to the outcomes of hand 2 responses. Perhaps surprisingly, this result holds even though the reward on hand 2 trials was larger than the reward on hand 1 trials for all proportions (Fig. 1b: dashed line > solid line for all proportions).

One explanation for this observation is as follows. Participants generally responded with hand 2 in Function 2 (Fig. 1d) and, as described above, the response ERNs indicate that hand 2 was favored over hand 1 (Fig. 5a, solid line, Function 2). Thus, participants tended to predict better outcomes following hand 2 responses than following hand 1 responses. If as proposed the feedback ERN reflects a reward prediction error signal, then its amplitude would have been modulated only by feedback that violated these predictions, namely feedback that was worse than expected following hand 2 responses and better than expected following hand 1 responses. This outcome obtained as participants continually favored hand 2 over hand 1, decreasing the proportion of hand 1 responses and yielding less reward overall. Therefore, following responses made with hand 2, the actual reward obtained was continually less than expected, eliciting negative prediction errors. In contrast, responses made with hand 1 increased the proportion of responses made with hand 1, leading to a commensurate increase in reward and positive prediction errors. Note that the feedback ERN difference measure was more positive for Function 2 than for Function 1 (Fig. 5a, dashed line; t = 2.6, p < .05), confirming that the relative value of the outcomes following hand 1 responses to the outcomes following hand 2 responses was greater for Function 2 than for Function 1. In other words, outcomes were perceived as being relatively better than expected

following hand 1 responses compared to hand 2 responses for Function 2 but not for Function 1.

Lastly, note that there appears to be a reciprocal relationship between the ERNs associated with the response and feedback: if in a given condition the response ERN was larger for responses made with one hand than for responses made with the other, then the opposite relationship tended to hold true for the feedback ERN (Fig. 5: solid and dashed lines are inversely related). For Function 2, hand 1 was more negative than hand 2 when associated with the response (solid line) but not when associated with the feedback (dashed line), for both the simulated and empirical data (t = 4.3, p < .005). For Function 1, hand 1 was more positive than hand 2 when associated with the response (solid line) but not when associated with the feedback (dashed line), for both the simulated and empirical data, although this difference did not reach significance (t = -1.7, p > .05).

5. Discussion

We applied the RL-ERN theory to predict the amplitude of the response and feedback ERNs in a decision making task in which an erroneous response was not clearly defined. Because the appropriateness of each choice was ambiguous, participants had to infer the optimal response strategy by integrating their response-reward histories over time. Nevertheless, as predicted by the RL-ERN theory, the relative amplitudes of the response ERN and the feedback ERN indicated sensitivity to the subjective values of participants' response options. In particular, when participants exhibited matching behavior in Function 1, receiving about the same amount of reward for either choice, the amplitudes of the response ERN and feedback ERN were about the same size following both button presses. In this case participants' behaviors settled on an equilibrium state in which both choices elicited equivalent amounts of reward, and there were no surprises. In contrast, in Function 2 participants responded nearly exclusively with the hand that minimized overall reward, even though they would have maximized their reward had they responded exclusively with the other hand. As suggested by their response ERNs, participants evidently evaluated the bad response option as being the better of the two and thus chose it over the alternative. Further, by choosing the worse option participants continually earned less reward than expected, except when they occasionally selected the better option, in which case they earned more reward than expected. Thus, the feedback ERN was relatively more positive following good choices than following bad choices. This result obtained even though the immediate reward following good choices was always smaller than the immediate reward following bad choices.

It is important to note that while the simulation captured the qualitative relationships between the relative amplitudes of the response ERN and feedback ERN, the fit was not perfect. In particular, the feedback ERN difference in Function 2 was smaller for the simulated data compared to the empirical data, even though the direction of the relationship (more positive for hand 1) was the same (Fig. 5: dashed lines, Function 2). Evidently the model identified the relative values of the two response options faster than the human participants did, and the size of the prediction errors fell more quickly. A comparable result was observed in a previous study, where it was suggested that an additional cognitive process in the human system might have continued to search for an appropriate response strategy in a circumstance where none was actually available (Holroyd and Coles, 2002). More research will be needed to investigate these discrepancies.

Given that the Herrnstein task does not involve objectively incorrect responses, these ERN results must be understood in a novel context. Classically, the response ERN occurs in speeded response time tasks immediately following errors (Falkenstein et al., 1990; Gehring et al., 1993), and recently it has also been observed in continuous motor control tasks (Krigolson and Holroyd, in 2007a, 2007b). Our results are consistent with previous suggestions that the response ERN may also reflect the value of a potential response during the decision making process (Hewig et al., 2007; Tucker et al., 1999). In this context it is relevant that the amplitude of the response ERN was modulated over a relatively wide time interval, from before the response to over 100 ms following it (Fig. 4b). This difference may have resulted from the fact that "errors" in this task were unlike the "fast guess" errors that elicit the typical ERN seen in speeded response time tasks, but rather occurred as participants sampled a response option that they perceived to be sub-optimal; in this case, the timing of the evaluative process might have occurred relatively independently of response generation. As well, note that the feedback ERNs reported here are different from the negative deflection that is typically seen in trial-and-error learning tasks (Miltner et al., 1997), as they appear to reflect variance in the amplitude of the positivity on correct trials (Fig. 4c and d), the P2a (Potts et al., 2006). Both positive and negative temporal difference errors can potentially impact the size of the ERN, the former decreasing and the latter

increasing its amplitude (Holroyd, 2004; Pakzad-Vaezi et al., 2006). For the present purpose, it is important to note that our difference measures preserved the relative values of the responses and outcomes without determining their absolute values.

Animals, including humans, live in an uncertain environment characterized by non-stationary reward contingencies. It is perhaps self-evident then that we evaluate individual rewards and punishments in the context of our reward histories, as the appropriateness of any individual action is difficult to determine on the basis of a single outcome. The mechanics of reinforcement learning have yet to be fully described, but prior responses and rewards clearly contribute to ongoing decision making (e.g., Catania, 1971; Lau and Glimcher, 2005; Sutton and Barto, 1998). Herrnstein proposed that animals keep track of the average yield associated with each available response option and shift to the option that currently provides the highest rate of return, a process that he called *melioration* (Herrnstein, 1997a, 1997b). Neurally, this decision making process is mediated by a distributed network of brain systems including orbitofrontal cortex, the basal ganglia and parietal cortex (Sanfey et al., 2006). As part of this network, the midbrain dopamine system may provide a "bottom up" learning signal that drives melioration (Egelman et al., 1998). Further, dACC may contribute to learning the values associated with actions, as the activity of this brain area depends on the history of previous response-outcome associations (Amiez et al., 2005; Gibson et al., 2006; Holroyd and Coles, 2002; Holroyd et al., 2004c; Kennerley et al., 2006; Mars et al., 2005; Matsumoto et al., 2007; Nieuwenhuis et al., 2002; Ullsperger and von Cramon, 2003; Volz et al., 2005). Our results demonstrate that the dACC is sensitive to the appropriateness of behavior even in tasks where response correctness is highly ambiguous and must be inferred by ongoing trial-and-error. These findings are consistent with the results of other recent ERN studies that have examined the neural mechanisms of decision making in complex learning tasks (Cohen and Ranganath, 2007; Frank et al., 2005; Hewig et al., 2007), and support our proposal that dACC uses such performance information to optimize the response selection process (Holroyd and Coles, 2002).

Acknowledgments

This research was supported in part by National Institutes of Mental Health predoctoral fellowship no. MH11530, postdoctoral fellowship no. MH63550, and grant no. MH41445; and by Natural Sciences and Engineering Research Council Discovery Grant No. RGPIN 312409-05.

REFERENCES

- Amiez C, Joseph J-P, and Procyk E. Anterior cingulate error related activity is modulated by predicted reward. European Journal of Neuroscience, 21: 3447–3452, 2005.
- Baker TE and Holroyd CB. Which way do I go? Neural activation in response to feedback processing and decision making in a virtual T-maze task. Psychophysiology, 43: S22, 2006.

- Botvinick MM, Braver TS, Barch DM, Carter CS, and Cohen JD. Conflict monitoring and cognitive control. Psychological Review, 108: 624–652, 2001.
- Botvinick MM, Cohen JD, and Carter CS. Conflict monitoring and anterior cingulate cortex: an update. *Trends in Cognitive Sciences*, 8: 539–546, 2004.
- Catania AC. Reinforcement schedules: the role of responses preceding the one that produces the reinforcer. *Journal of the Experimental Analysis of Behavior*, 15: 271–287, 1971.
- Cohen MX and Ranganath C. Reinforcement learning signals predict future decisions. *Journal of Neuroscience*, 27: 371–378, 2007.
- Coles MGH, Scheffers MK, and Holroyd C. Berger's dream? The error-related negativity and modern cognitive psychophysiology. In Witte H, Zwiener U, Schack B, and D'ring A (Eds), Quantitative and Topological EEG and MEG Analysis. Jena-Erlangen: Druckhaus Mayer Verlag, 1998: 96–102.
- Debener S, Ullsperger M, Siegel M, Fiehler K, von Cramon DY, and Engel AK. Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. The Journal of Neuroscience, 25: 11730–11737, 2005.
- Dehaene S, Posner MI, and Tucker DM. Localization of a neural system for error detection and compensation. Psychological Science, 5: 303–305, 1994.
- Donchin E and Heffley III EF. Multivariate analysis of eventrelated potential data: A tutorial review. In Otto D (Ed), Multidisciplinary Perspectives in Event-related Potential Research. Washington, DC: US Government Printing Office, 1978: 555–572. EPA-600/9-77-043.
- Donkers FCL and van Boxtel GJM. Mediofrontal negativities to averted gains and losses in the slot-machine task: a further investigation. Journal of Psychophysiology, 19: 256–262, 2005.
- Egelman DM, Person C, and Montague PR. A computational role for dopamine delivery in human decision-making. *Journal of Cognitive Neuroscience*, 10: 623–630, 1998.
- Falkenstein M, Hohnsbein J, Hoormann J, and Blanke L. Effects of errors in choice reaction tasks on the ERP under focused and divided attention. In Brunia C, Gaillard A, and Kok A (Eds), Psychophysiological Brain Research. Tilburg, The Netherlands: Tilburg University Press, 1990: 192–195.
- Frank MJ, Woroch BS, and Curran T. Error-related negativity predicts reinforcement learning and conflict biases. *Neuron*, 47: 495–501, 2005.
- Gehring WJ, Goss B, Coles MGH, Meyer DE, and Donchin E. A neural system for error detection and compensation. *Psychological Science*, 4: 385–390, 1993.
- Gehring WJ and Willoughby AR. The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, 295: 2279–2282, 2002a.
- Gehring WJ and Willoughby AR. Response. Science, 296: 1611, 2002b.
- Gibson J, Krigolson OE, and Holroyd CB. Sensitivity of the feedback error-related negativity to reward probability. Psychophysiology, 43: S41–S42, 2006.
- Gratton G, Coles MGH, and Donchin E. A new method for off-line removal of ocular artifact. Electroencephalography and Clinical Neurophysiology, 55: 468–484, 1983.
- Hajcak G, Moser J, Holroyd CB, and Simons RF. The feedbackrelated negativity reflects the binary evaluation of good versus bad outcomes. *Biological Psychology*, 71: 148–154, 2006.
- Herrnstein RJ. Experiments on stable suboptimality in individual behavior. In Herrnstein RJ (Ed), The Matching Law. New York: Russell Sage Foundation, 1997a: 266–273.
- Herrnstein RJ. Melioration: a theory of distributed choice. In Herrnstein RJ (Ed), The Matching Law. New York: Russell Sage Foundation, 1997b: 274–292.

- Hewig J, Trippe R, Hecht H, Coles MGH, Holroyd CB, and Miltner WHR. An electrophysiological analysis of decisionmaking in blackjack. Cerebral Cortex, 17: 865–877, 2007.
- Holroyd CB. A note on the Oddball N200 and the feedback ERN. In Ullsperger M, and Falkenstein M (Eds), Errors, Conflicts, and the Brain: Current Opinions on Performance Monitoring. Leipzig: MPI of Cognitive Neuroscience, 2004: 211–218.
- Holroyd CB and Coles MGH. The neural basis of human error processing: reinforcement learning, dopamine, and the errorrelated negativity. *Psychological Review*, 109: 679–709, 2002.
- Holroyd CB and Krigolson OE. Reward prediction error signals associated with a modified time estimation task. *Psychophysiology*, 44: 913–917, 2007.
- Holroyd CB and Yeung N. Alcohol and error processing. Trends in Neurosciences, 26: 402–404, 2003.
- Holroyd CB, Dien J, and Coles MGH. Error-related scalp potentials elicited by hand and foot movements: evidence for an outputindependent error-processing system in humans. *Neuroscience Letters*, 242: 65–68, 1998.
- Holroyd CB, Coles MGH, and Nieuwenhuis S. Medial prefrontal cortex and error potentials. *Science*, 296: 1610–1611, 2002.
- Holroyd CB, Nieuwenhuis S, Yeung N, and Cohen JD. Errors in reward prediction are reflected in the event-related brain potential. *Neuroreport*, 14: 2481–2484, 2003.
- Holroyd CB, Larsen JT, and Cohen JD. Context dependence of the event-related brain potential associated with reward and punishment. *Psychophysiology*, 41: 245–253, 2004a.
- Holroyd CB, Nieuwenhuis S, Mars R, and Coles MGH. Anterior cingulate cortex, selection for action, and error processing. In Posner M (Ed), Cognitive Neuroscience of Attention. New York: Guilford Publishing, Inc., 2004b: 219–231.
- Holroyd CB, Nieuwenhuis S, Yeung N, Nystrom L, Mars R, Coles MGH, and Cohen JD. Dorsal anterior cingulate cortex shows fMRI response to internal and external error signals. *Nature Neuroscience*, 7: 497–498, 2004c.
- Holroyd CB, Yeung N, Coles MGH, and Cohen JD. A mechanism for error detection in speeded response time tasks. *Journal of Experimental Psychology: General*, 134: 163–191, 2005.
- Holroyd CB, Hajcak G, and Larsen JT. The good, the bad and the neutral: electrophysiological responses to feedback stimuli. *Brain Research*, 1105: 93–101, 2006.
- Kennerley SW, Walton ME, Behrens TEJ, Buckley MJ, and Rushworth MFS. Optimal decision making and the anterior cingulate cortex. Nature Neuroscience, 9: 940–947, 2006.
- Krigolson OE and Holroyd CB. Evidence for hierarchical error processing in the human brain. Neuroscience, 137: 13–17, 2006.
- Krigolson OE and Holroyd CB. Hierarchical error processing: different errors, different systems. Brain Research, 1155: 70–80, 2007a.
- Krigolson OE and Holroyd CB. Predictive information and error processing: the role of medial-frontal cortex during motor control. Psychophysiology, 44: 586–595, 2007b.
- Lau B and Glimcher PW. Dynamic response-by-response models of matching behavior in rhesus monkeys. *Journal of the Experimental Analysis of Behavior*, 84: 555–579, 2005.
- Mars RB, Coles MGH, Grol MJ, Holroyd CB, Nieuwenhuis S, Hulstijn W, and Toni I. Neural dynamics of error processing in medial frontal cortex. *Neuroimage*, 28: 1007–1013, 2005.
- Matsumoto M, Matsumoto K, Abe H, and Tanaka K. Medial prefrontal cell activity signaling prediction errors of action values. *Nature Neuroscience*, 10: 647–656, 2007.
- Miltner WHR, Braun CH, and Coles MGH. Event-related brain potentials following incorrect feedback in a time-estimation task: evidence for a "generic" neural system for error detection. *Journal of Cognitive Neuroscience*, 9: 788–798, 1997.
- Montague PR and Berns GS. Neural economics and the biological substrates of valuation. *Neuron*, 36: 265–284, 2002.
- Montague PR, Hyman SE, and Cohen JD. Computational roles for dopamine in behavioural control. Nature, 431: 760–767, 2004.

- Nieuwenhuis S, Ridderinkhof KR, Talsma D, Coles MGH, Holroyd CB, Kok A, and Van der Molen MW. A computational account of altered error processing in older age: dopamine and the error-related negativity. Cognitive, Affective, and Behavioral Neuroscience, 2: 19–36, 2002.
- Nieuwenhuis S, Holroyd CB, Mol N, and Coles MGH. Reinforcement-related brain potentials from medial frontal cortex: origins and functional significance. *Neuroscience and Biobehavioral Reviews*, 28: 441–448, 2004a.
- Nieuwenhuis S, Yeung N, Holroyd CB, Schurger A, and Cohen JD. Sensitivity of electrophysiological activity from medial frontal cortex to utilitarian and performance feedback. *Cerebral Cortex*, 14: 741–747, 2004b.
- Nieuwenhuis S, Slagter HA, von Geusau NJA, Heslenfeld D, and Holroyd CB. Knowing good from bad: differential activation of human cortical areas by positive and negative outcomes. *European Journal of Neuroscience*, 21: 3161–3168, 2005.
- Pakzad-Vaezi K, Krigolson OE, and Holroyd CB. Underlying unity: the relationship between the N200 and the feedback ERN. Psychophysiology, 43: S75, 2006.
- Paus T. Primate anterior cingulate cortex: where motor control, drive, and cognition interface. Nature Reviews Neuroscience, 2: 417–424, 2001.
- Potts GF, Martin LE, Burton P, and Montague PR. When things are better or worse than expected: the medial frontal cortex and the allocation of processing resources. *Journal of Cognitive Neuroscience*, 18: 1112–1119, 2006.
- Ridderinkhof KR, Ullsperger M, Crone EA, and Nieuwenhuis S. The role of the medial frontal cortex in cognitive control. *Science*, 306: 443–447, 2004.
- Rushworth MFS, Walton ME, Kennerley SW, and Bannerman DM. Action sets and decision in the medial frontal cortex. Trends in Cognitive Sciences, 8: 410–417, 2004.
- Sanfey AG, Lowenstein G, McClure SM, and Cohen JD. Neuroeconomics: cross-currents in research on decision making. Trends in Cognitive Sciences, 10: 108–116, 2006.

- Schultz W. Getting formal with dopamine and reward. Neuron, 36: 241–263, 2002.
- Spencer KM, Dien J, and Donchin E. Spatiotemporal analysis of the late ERP responses to deviant stimuli. Psychophysiology, 38: 343–358, 2001.
- Sutton RS. Learning to predict by the methods of temporal differences. Machine Learning, 3: 9–44, 1988.
- Sutton RS and Barto AG. Reinforcement learning: an introduction. Cambridge, MA: MIT Press, 1998.
- Toyomaki A and Murohashi H. Discrepancy between feedback negativity and subjective evaluation in gambling. *Neuroreport*, 16: 1865–1868, 2005.
- Tucker DM. Spatial sampling of head electrical fields: the geodesic sensor net. Electroencephalography and Clinical Neurophysiology, 87: 154–163, 1993.
- Tucker DM, Hartry-Speiser A, McDougal L, Luu P, and deGrandpre D. Mood and spatial memory: emotion and right hemisphere contributions to spatial cognition. Biological Psychology, 50: 103–125, 1999.
- Ullsperger M and von Cramon DY. Error monitoring using external feedback: specific roles of the habenular complex, the reward system, and the cingulate motor area revealed by functional resonance imaging. The Journal of Neuroscience, 23: 4308–4314, 2003.
- Volz KG, Schubotz RI, and von Cramon DY. Frontomedian activation depends on both feedback validity and valence: fMRI evidence for contextual feedback evaluation. *Neuroimage*, 27: 564–571, 2005.
- Walton ME, Devlin JT, and Rushworth MFS. Interactions between decision making and performance monitoring within prefrontal cortex. *Nature Neuroscience*, 7: 1259–1265, 2004.
- Yeung N, Botvinick M, and Cohen JD. The neural basis of error detection: conflict monitoring and the error-related negativity. Psychological Review, 111: 931–959, 2004.
- Yeung N and Sanfey AG. Independent coding of reward magnitude and valence in the human brain. *Journal of Neuroscience*, 24: 6258–6264, 2004.