NeuroImage 63 (2012) 1078–1090

Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Modulation of feedback related activity in the rostral anterior cingulate cortex during trial and error exploration

Céline Amiez [b,1], Jérôme Sallet [a,*,1], Emmanuel Procyk [a], Michael Petrides [b]

[a] INSERM U846, Stem Cell and Brain Research Institute, 18 Avenue du Doyen Lepine, 69000 Bron, France
[b] Montreal Neurological Institute, Cognitive Neuroscience Unit, McGill University, 3801 University Street, Montreal, Quebec, Canada H3A 2B4

ARTICLE INFO

Article history:
Accepted 18 June 2012
Available online 23 June 2012

Keywords:
Anterior cingulate cortex
Reward prediction error
Trial and error exploration
Exploitation
fMRI

ABSTRACT

The present functional magnetic resonance imaging (fMRI) experiment aims at clarifying the role of the rostral anterior cingulate cortex (rACC) in the evaluation of feedback in a deterministic environment. We tested, in particular, the response of the rACC to the detection of different types of feedback, and to varying levels of outcome expectancy. We used a problem-solving task in which subjects had to discover, in successive trials, which one of the four presented stimuli was associated with a positive feedback, the other ones being associated with error feedback. In this task, two periods alternated: 1. an exploratory period in which error feedback indicated to adapt the following response appropriately (and continue to explore), and first positive feedback indicated to change strategy (i.e. to shift from explorative to exploitative behavior), and 2. an exploitative period in which subjects had to repeat the correct choice. The rACC is recruited in the exploratory period during the analysis of both error and first correct positive feedback. In addition, the rACC activity was modulated by positive reward prediction error values (i.e. the difference between obtained and expected feedback). Altogether, these results reveal the critical role of the rACC in the evaluation of salient feedback for learning optimal strategies.

© 2012 Elsevier Inc. All rights reserved.

Introduction

In non-routine situations, the detection and evaluation of behavioral feedback are of critical importance to adjust subsequent choices efficiently. In deterministic environments, whereas the commission of an error indicates the need to modify subsequent behavioral choices, the occurrence of a first positive feedback indicates the need to stop exploring the environment and to start exploiting the available resources. There is a growing body of evidence from fMRI and electroencephalographic (EEG) experiments in human subjects suggesting that the feedback-related activity observed in the anterior cingulate cortex (ACC) is not restricted to error processing, as initially suggested, but appears to result from other computations related to outcome properties, such as the analysis of the volatility or uncertainty of outcome estimates (Behrens et al., 2007; Cohen et al., 2007; Oliveira et al., 2007; Walton et al., 2004), the interpretation of feedback as success or failure (Jung et al., 2010), or the estimation of the cost in term of effort in choice selection (Croxson et al., 2009). Similarly, both electrophysiological (Amiez et al., 2006; Hayden et al., 2011a, 2011b; Ito et al., 2003; Kennerley and Wallis, 2009; Matsumoto et al., 2007; Quilodran et al., 2008; Sallet et al., 2007; Seo and Lee, 2009; Shima and Tanji, 1998) and lesion studies (Amiez et al., 2006; Hadland et al., 2003; Kennerley et al., 2006; Rudebeck et al., 2008; Shima and Tanji, 1998) in monkeys have suggested that the ACC evaluates current feedback that is pertinent to behavioral adjustment.

Specifically, two hypotheses can be distinguished concerning the role of the ACC in the evaluation of feedback. The first proposes that the ACC evaluates any feedback (and not only negative) and only in the context where it is salient for learning optimal strategies, e.g. in non-routine situations (Amiez et al., 2012; Behrens et al., 2007; Holroyd et al., 2009; Quilodran et al., 2008; Rushworth and Behrens, 2008; Rushworth et al., 2011). The second assumption is that the ACC is sensitive to both feedback value and expectancy (Amiez et al., 2005; Jessup et al., 2010; Matsumoto et al., 2007; Sallet et al., 2007). The latter is mainly based on the hypothesis that ACC feedback-related activity (in particular for negative feedback) depends mainly on direct dopaminergic inputs from the mesencephalon (Holroyd and Coles, 2002). Indeed, activity of both ACC and dopaminergic neurons has been shown to correlate with reward prediction error (RPE), i.e. the difference between the obtained and the expected reward (Amiez et al., 2005; Kennerley and Walton, 2011; Matsumoto et al., 2007; Sallet et al., 2007). Note however that the respective strength of coding for negative and positive prediction error is in question (Cohen et al., 2011; Kennerley and Walton, 2011).

The aim of the present functional magnetic resonance imaging (fMRI) study was to assess these two propositions. For this purpose,

* Corresponding author at: Decision and Action Laboratory, Department of Experimental Psychology, University of Oxford, South Parks Road, OX1 3UD, UK.
E-mail address: jerome.sallet@psy.ox.ac.uk (J. Sallet).
[1] The first two authors have contributed equally to this work.

1053-8119/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2012.06.023

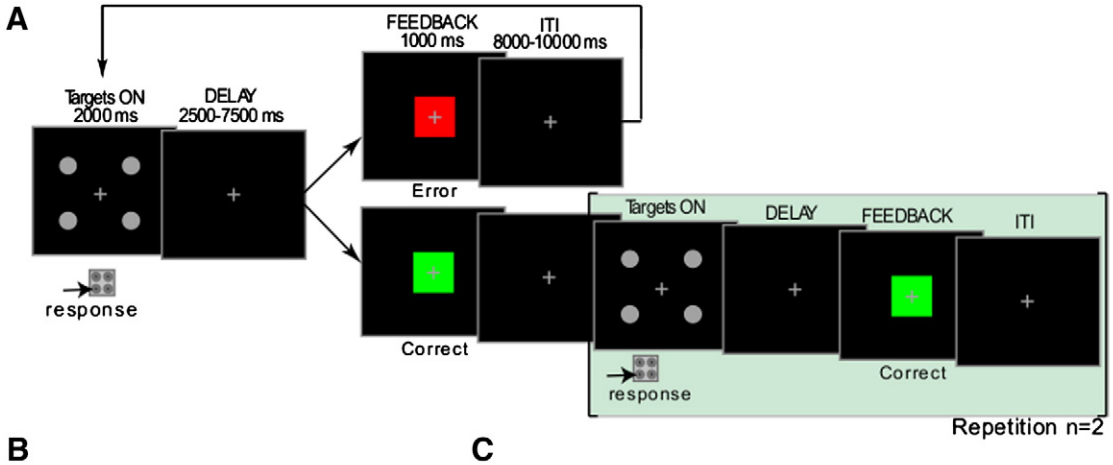
we adapted a problem-solving task developed for electrophysiological recording in monkeys (Quilodran et al., 2008; Rothe et al., 2011; Vezoli and Procyk, 2009). Subjects had to discover which one of the four presented stimuli was associated with correct feedback (i.e. green square), the other ones being associated with error feedback (i.e. red square). Once the stimulus associated with correct feedback was discovered (explorative period), the subjects had to repeat the same choice on 2 additional trials (exploitative period). In the explorative period of this task, therefore, both negative and positive feedback informed the subjects of how to adjust their behavior on the subsequent trial. However, whereas error feedback indicated the need to switch response on the next trial and keep searching for the rewarded target, the first positive feedback indicated the need to switch strategy (i.e. to halt exploration and start exploiting) but to keep the response identical on the following trial. In the exploitative period of the task, positive feedback indicated maintenance of both the response and the ongoing strategy (i.e. exploit).

If the first hypothesis is correct, the same region of ACC should be more active at the occurrence of both negative and positive feedback in the explorative period of the task than at the occurrence of positive feedback in the exploitative period of the task. If the second is correct, the ACC feedback-related activity during the exploration period should vary with reward prediction error.

Material and methods

Subjects

Data from 11 right-handed normal human subjects (6 females and 5 males, mean age = 27.9 years ± 3.6 SD) were recorded and analyzed



B

Problems	Trials 1	2	3	4	5	6
Type I	C1	R	R			
Type II	I1	C2	R	R		
Type III	I1	I2	C3	R	R	
Type IV	I1	I2	I3	C4	R	R

C

Trials	I3	I2	I1	R	C4	C3	C2	C1
RPE values	-0.5	-0.33	-0.25	0	0	0.5	0.66	0.75

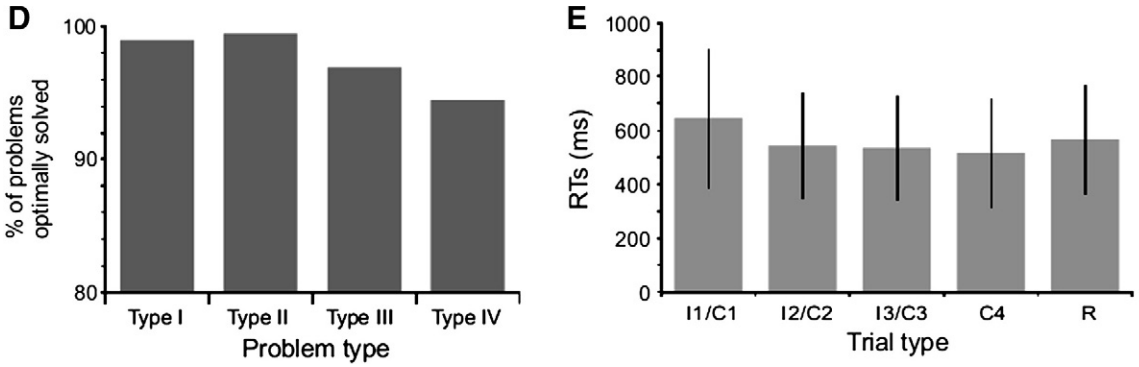


Fig. 1. Behavioral tasks and performance. A. Experimental design: The problem solving task involving the search by trial and error of the one of four stimuli that is associated with positive feedback. Each trial started with a 2000 ms target presentation. During this period, the subjects had to choose one stimulus using one of four computer mouse buttons. If a stimulus was selected within this 2000 ms period, the stimuli were switched off and a white central fixation cross appeared for a 2500 to 7500 ms (with an average of 3500 ms) period. This delay preceded the appearance of the feedback stimulus, a green square (positive) or a red square (negative), which was displayed on the screen for 1000 ms. The offset of the feedback stimulus was followed by a 8000 ms to 10000 ms inter-trial interval (ITI). After an incorrect choice, subjects continued to search for the correct stimulus while keeping in memory the previous erroneous choices. Discovering the correct stimulus (i.e. which was indicated by the green square feedback) ended the explorative period. The subjects then repeated the correct choice twice (i.e. exploitative period, see shaded area). At the end of the second exploitative trial, a blue ellipse appeared on the center of the screen for 500 ms. This ellipse indicated that a new Problem would start. B. Problem types and trials. Four types of problem (i.e. types I, II, III, and IV) and 8 types of trial (I1, I2, I3, C1, C2, C3, C4, R) occurred 3 times in each functional run. C. Reward Prediction Error (RPE) values associated with each trial type. D. Subjects' mean success rate. The percent of correct searches was 99% in Type I, 99.5% in Type II, 97% in Type III, and 94.4% in Type IV. Note that only correct searches (i.e. searches in which the subjects had not repeated previously erroneous choices) were entered in the fMRI data analysis. e. Subjects' mean reaction time ± SD. The subjects' mean reaction time in the explorative period was 581.89 ± 230.09 ms and 568.55 ± 202.09 ms in the exploitative period ($t = -1.789$, $P < 0.07$, unpaired *t*-test).

in this experiment (see Supplemental material). Informed, written consent was obtained from all the participants according to the institutional guidelines established by the Ethics Committee of the Montreal Neurological Institute and Hospital.

Behavioral protocol

The present fMRI study was designed to examine changes in cerebral activity related to the processing of various feedback signals generated

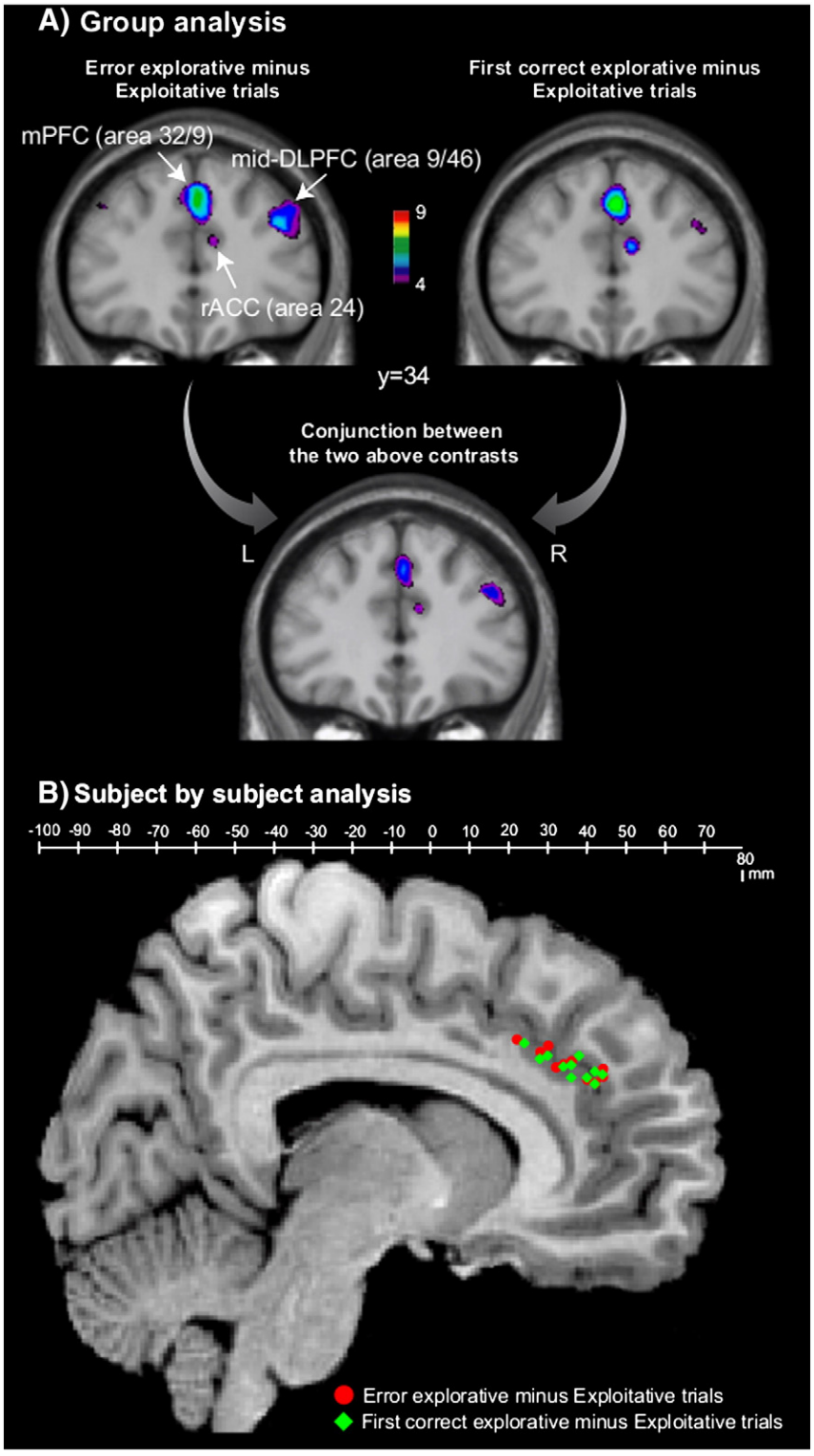


Fig. 2. Activity during the presentation of the feedback within the rACC. A. Average activity increases in the rACC from all 11 subjects in the comparisons "all error explorative minus exploitative trials" and "all first correct explorative minus exploitative trials", and in the conjunction analysis of these two contrasts. The t statistical map of activity has been superimposed on the average T1 anatomical acquisition of the 11 subjects transformed into standard stereotaxic space. The Y value indicates the antero-posterior level of the stereotaxic coordinates expressed in mm within the MNI stereotaxic proportional system. The color scale indicates the t-value range. B. Schematic representation of the location of the activity increases obtained in each individual subject in the rACC in both the contrasts "all error explorative minus exploitative trials" (red circles) and "all first correct explorative minus exploitative trials" (green rhombus). These activity increases are displayed on a typical single subject's brain rendered in standard stereotaxic space. Abbreviations: R and L, right and left hemisphere; mPFC, medial Prefrontal Cortex; DLPFC, Dorso-Lateral Prefrontal Cortex.

during self-organized behavior and to examine how RPE values associated with feedback signals are encoded. In the problem-solving task, the subject was presented with four identical stimuli (four disks) and had to discover by trial and error in successive trials which one of these stimuli was associated with the positive feedback, the other ones being associated with an error feedback. Two periods could be distinguished in this task: an explorative and an exploitative period. The combination of an explorative and an exploitative period constituted a "problem". In each trial of the explorative period, four identical stimuli appeared simultaneously at the four locations (see Fig. 1A) on the screen for a 2 s duration. During these 2 s, the subjects had to select one of these stimuli by pressing the corresponding mouse button. Note that the four mouse buttons matched the spatial position of the targets (see Fig. 1A). After a delay varying from 2.5 to 7.5 s (average = 3.5 s), a feedback stimulus appeared for a 1 s duration. If the stimulus choice was incorrect (i.e. the correct location was selected), the feedback stimulus was a red square (see Fig. 1A) and, after an inter-trial interval pseudo-randomly varying between 8 and 10 s, the same 4 identical stimuli were presented again and the subject had to select another location. Testing proceeded in this manner until the subject found the correct stimulus, the correct choice being indicated by the positive feedback stimulus, i.e. a green square. After the first correct choice (associated with the presentation of the first positive feedback), the explorative period terminated and the exploitative period started. The exploitative period consisted of two trials during which the subject had to keep selecting the correct stimulus. Then, the occurrence of a blue ellipse for 500 ms informed the subject that the exploitative period was over and that a new problem would start after an inter-trial interval lasting 8 to 10 s. The same 4 stimuli were used across the experiment. Subjects were instructed not to use verbal strategies during the performance of the task. This instruction was given in an attempt to minimize variability in the strategies employed by the subjects to solve the problems.

The problems were defined by sequences of trials that were themselves labeled according to the obtained feedback at the end of each trial. Trials were classified as follows: I1, I2, and I3 referred to the subject choosing incorrectly on the first, second, and third trial of the explorative period, respectively; C1, C2, C3, and C4 referred to the subject choosing correctly on the first, second, third, and fourth trial of the explorative period, respectively; R referred to the exploitative period.

Based on the sequence of trials, we categorized different problem types. They were defined by the following combination : Problems of Type I contained C1 and 2 R trials; Problems of Type II were composed of I1, C2, and 2 R trials; Problems of Type III were composed of I1, I2, C3, and 2 R trials; finally, Problems of Type IV contained I1, I2, I3, C4, and 2 R trials (Fig. 1B). Problems of Type 1, 2, 3, and 4, are therefore composed of 3, 4, 5, and 6 trials, respectively.

In order to obtain identical numbers of each problem type, the length of the explorative period (i.e. the number of trials required to discover the stimulus associated with the correct feedback) was randomly determined by the computer but the subject was not informed of this bias.

For each type of incorrect trial (I1, I2, I3) and correct trial (C1, C2, C3, C4) during the explorative period, negative and positive reward prediction errors (RPE) were calculated as follow:

$$RPE = [r_{obt}] - \left[r_{exp} \cdot p_{cor}\right]$$

where r_{obt} is the obtained feedback, r_{exp} is the expected feedback (= 1), and p_{cor} is the probability of selecting the correct stimulus.

The formula was based on the following assumptions: 1. the subject is always aiming for a positive feedback, 2. the value of the obtained feedback is "0" or "1" for negative and positive feedback, respectively, 3. the probabilities of choosing the correct stimulus in the different trials

Table 1
Brain regions showing increased activity during the presentation of the feedback in the contrast "all error explorative trials minus exploitative trials" and "all first correct explorative trials minus exploitative trials". All t-statistics were significant at $P<0.05$ (see Material and methods for the explanations concerning the determination of significant t-statistics thresholds). Note that the activity in the fronto-insula extended to the inferior frontal junction (IFJ: (x, y, z) 52, 10, 12, t-value = 4.58 in the contrast "all error explorative trials minus exploitative trials" and 51, 9, 13, t-value = 5.01, in the contrast "all first correct explorative trials minus exploitative trials"). Note that the frontopolar activity was located in area 10 at the intersection between the vertical branch of the intermediate frontal sulcus, the lateral and the intermediate frontomarginal sulci (Petrides and Pandya, 2004).

	Feedback: all error explorative minus exploitative trials				Feedback: all first correct explorative minus exploitative trials			
	MNI coordinates				MNI coordinates			
	x	y	Z	t-Statistic	x	y	z	t-Statistic
Right hemisphere								
Rostral ACC	9	34	22	4.06	8	34	21	5.41
MPFC (area 32/8)	3	32	41	6.01	2	33	40	7.10
Mid-DLPFC (area 9/46)	39	34	30	5.43	38	34	32	4.43
Mid-DLPFC (area 46)	42	50	10	6.01	44	52	10	7.36
Orbitofrontal cortex (area 11)	22	44	−14	7.12	22	46	−12	6.91
Frontopolar cortex (area 10)	29	60	−5	7.38	20	60	−11	6.84
Fronto-insula	30	22	−4	8.94	30	20	−5	9.00
Striatum	14	8	6	5.10	14	10	0	6.25
Retrosplenial cortex	1	−28	29	5.10	3	−28	30	6.19
Middle temporal cortex	61	−24	−8	5.23	66	−28	−5	6.39
Mid-inferior parietal lobule	42	−42	48	6.08	40	−46	44	6.19
Precuneus (area 7m)	11	−64	54	5.69	10	−73	45	6.39
Left hemisphere								
Mid-DLPFC (area 9/46)	−44	32	38	4.11	−42	34	36	4.09
Orbitofrontal cortex (area 11)	−21	46	−14	4.58	−21	46	−13	4.30
Frontopolar cortex (area 10)	−29	58	0	5.43	−30	56	0	4.36
Fronto-insula	−33	22	2	7.06	−30	24	0	7.50
Striatum	−16	6	2	3.67	−11	8	0	5.01
Mid-inferior parietal lobule	−36	−44	46	6.47	−38	−44	39	5.60
Cerebellum	−30	−60	−30	6.33	−30	−59	−30	6.05
Cerebellum	−8	−64	−51	4.97	−7	−60	−51	5.53
Cerebellum	−34	−66	−43	4.38	−42	−64	−42	5.66

Table 2
Brain regions showing increased activity in the conjunction analysis between the contrasts "all error explorative minus exploitative trials" and "all first correct explorative minus exploitative trials". All t-statistics were significant at $P<0.05$, except the mid-DLPFC and the frontopolar cortical increases in activity in the left hemisphere. However, given that the increases in activity in the opposite hemisphere were statistically significant, it is unlikely that these results are false positive.

Conjunction analysis between the contrasts "all error explorative minus exploitative trials" and "all first correct explorative minus exploitative trials"				
	MNI coordinates			
	x	y	z	t-Statistic
Right hemisphere				
Rostral ACC	8	34	22	4.69
MPFC (area 32/8)	2	32	40	5.57
Mid-DLPFC (area 9/46)	40	34	30	4.80
Mid-DLPFC (area 46)	40	48	10	5.57
Orbitofrontal cortex (area 11)	22	46	−12	5.51
Frontopolar cortex (area 10)	32	60	−4	4.98
Fronto-insula	30	20	−4	7.75
striatum	14	8	8	4.69
Retrosplenial cortex	0	−28	30	5.75
Middle temporal cortex	61	−22	−10	5.22
Mid-inferior parietal lobule	42	−44	46	5.22
Precuneus (area 7m)	8	−72	48	5.69
Left hemisphere				
Mid-DLPFC (area 9/46)	−44	32	38	3.03
Orbitofrontal cortex (area 11)	−22	46	−12	4.10
Frontopolar cortex (area 10)	−30	54	2	3.29
Fronto-insula	−30	24	0	6.63
Striatum	−14	10	2	3.86
Mid-inferior parietal lobule	−42	−48	54	6.10
Cerebellum	−30	−58	−30	5.22
Cerebellum	−8	−60	−48	4.74
Cerebellum	−38	−66	−44	4.63

of each explorative period were calculated based on the number of available alternatives, i.e. the more options available, the lower the probability of discovering the correct stimulus. For instance, in the first trial where 4 alternatives were available, the probability of choosing the correct stimulus was $P=1/4=0.25$; in the second trial where 3 alternatives were available, $P=1/3=0.33$; in the third trial where 2 alternatives were available, $P=1/2=0.5$, and in the fourth trial $P=1$. The RPE values related to each trial are represented in Fig. 1C.

One day prior to the scanning session, the subjects were trained on the task until they reached a level of 95% correct performance on all problem types. Problems were labeled as correct when no repetition of incorrect choices occurred in the explorative period and no incorrect choices occurred in the exploitative period. Only data obtained from problems solved correctly were used for fMRI analysis.

Scanning procedure and fMRI data analysis

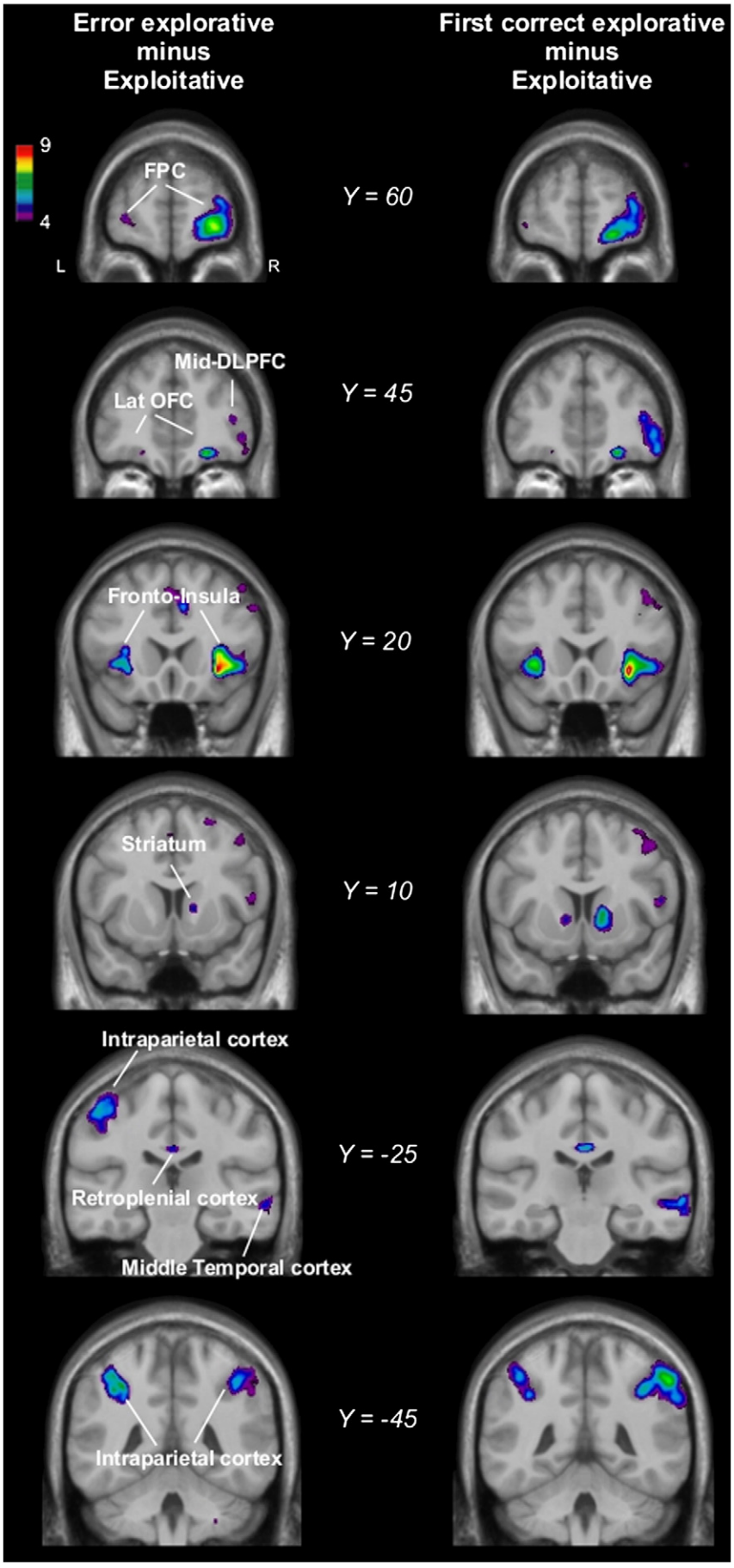
Each volunteer was scanned in a 1.5 T Siemens Sonata MRI Scanner (Siemens AG, Erlangen, Germany). After a high-resolution T1 anatomical scan (whole head, 1 mm³ isotropic resolution), five (in 2 subjects) or six (in 9 subjects) functional runs of 260 images each (38 oblique T2* gradient echoplanar images (EPI), voxel size = 3.4×3.4×3.4 mm, TR = 3.5 s, TE = 45 ms, flip angle = 90°) sensitive to the blood oxygenation level dependent (BOLD) signal were acquired. The FOV covered the whole brain. The first trial onset in each run was synchronized with the scanner acquisition via a trigger signal generated by the scanner. Behavioral and imaging data were acquired on all trials. In each functional run, the subjects performed 10 (in 2 subjects) or 12 problems (in 9 subjects), i.e. 3 problems of each type pseudo-randomly selected by the computer from a predetermined list of problems to be solved (see Fig. 1B). Since each problem was composed of a number of trials varying from 3 to 6 trials (see description of the problem type above), since each run was composed of 3 problems of each type, and since 9 participants performed 6 runs and 2 participants performed 5 runs, the total number of trials studies in the present paper is 3456 (i.e. [9 participants∗6 runs∗(3∗3+3∗4+3∗5+3∗6) trials]+[2 participants∗5 runs∗(3∗3+3∗4+3∗5+3∗6) trials]).

Images from all runs were first realigned with an AFNI image registration software using the third frame of the first run as a reference (Cox and Jesmanowicz, 1999), and then smoothed with an in-house MINC blurring software (mincblur) using a 6 mm full-width half-maximum (FWHM) isotropic Gaussian kernel. Subsequently, all images were non-linearly transformed into MNI stereotaxic space using in-house dedicated software (Collins et al., 1994). The non-linear transform was estimated on MRI data blurred with an 8 mm FWHM Gaussian kernel and a 3D lattice grid with 4 mm spacing between nodes. The data analysis was performed using fMRIstat (available at http://www.math.mcgill.ca/keith/fmristat) (Worsley et al., 2002). The statistical analysis of the fMRI data was based on a linear model with correlated errors. The design matrix of the linear model was first convolved with a hemodynamic response function modeled as a difference of two gamma functions timed to coincide with the acquisition of each slice (Friston et al., 1995). Temporal drift was removed by adding polynomial covariates in the frame times, up to degree three, to the design matrix (Worsley et al., 2002). Sufficient sampling across the hemodynamic response function was acquired in this experiment due to the complete desynchronization of the trial onset time from the acquisition exploitative time. This desynchronization was achieved by having a fixed exploitative time of acquisition (TR = 3.5 s) and a varied total trial duration. The correlation structure was modeled as an autoregressive process of degree one (Bullmore et al., 1996). At each voxel, the autocorrelation parameter was estimated from the least squares residuals using the Yule–Walker equations, after a bias correction for correlations induced by the linear model. The autocorrelation parameter was first regularized by spatial smoothing with a 15 mm FWHM Gaussian filter and then used to 'whiten' the data and the design matrix. The linear model was then re-estimated using least squares on the whitened data to produce estimates of effects and their standard errors (Worsley et al., 2002).

In a second step, runs and subjects were combined using a mixed effects linear model for the effects (as data) with fixed effects standard deviations taken from the previous analysis. This was fitted using residual error maximum likelihood implemented by the expectation maximization algorithm. Experimental runs within a subject were combined with a fixed effects analysis that involved estimating the ratio of the random effects variance to the fixed effects variance, then smoothing this ratio with an infinitely large FWHM Gaussian filter, yielding a global ratio of one. The random effects analysis across subjects was performed by smoothing the variance ratio with a 15 mm FWHM Gaussian filter. The variance of the effect was then estimated by the smoothed ratio multiplied by the fixed effects variance. The amount of smoothing was chosen to achieve one hundred effective degrees of freedom (Worsley et al., 2002).

The resulting T statistic images were thresholded using the minimum given by a Bonferroni correction and random field theory to account for multiple comparisons. Significance was assessed on the

Fig. 3. Activity increases during the presentation of the feedback in the comparison "all error explorative minus exploitative trials" (left panels) and "all first correct explorative minus exploitative trials" (right panels) located in the prefrontal, the temporal, and the parietal cortex. The average activity increases are obtained from all 11 subjects. The t statistical map of activity has been superimposed on the average T1 anatomical acquisition of the 11 subjects transformed into standard stereotaxic space. The Y value corresponds to the antero-posterior level of the stereotaxic coordinates expressed in mm within the MNI stereotaxic proportional system. The color scale indicates the t-value range.


Error explorative minus Exploitative
First correct explorative minus Exploitative
9
4
L
R
FPC
Y = 60
Mid-DLPFC
Lat OFC
Y = 45
Fronto-Insula
Y = 20
Striatum
Y = 10
Intraparietal cortex
Retroplenial cortex
Middle Temporal cortex
Y = -25
Intraparietal cortex
Y = -45

basis of the spatial extent of consecutive voxels. Concerning the group analysis, a cluster volume extent >171.65 mm³ with a t-value >3 was significant (P<0.05), corrected for multiple comparisons using the method of Friston et al. (1995). In addition, for a single voxel in an exploratory explorative involving all peaks within an estimated gray matter of 600 cm³ covered by the slices, the threshold for reporting a peak as significant (P<0.05) was t=4.509 if the peaks were not predicted and t=4.036 if the peaks were predicted (Worsley et al., 1996). Concerning the individual subject analysis, we conducted a region of interest (ROI) analysis. It has been shown that the cingulate sulcus in the right hemisphere has a volume of 1273.67 mm^3 (Rametti et al., 2010). In this ROI analysis, a predicted cluster of voxels with a t-value >2.4 and with a volume extent >49.96 mm^3 in the right anterior cingulate sulcus, respectively, was significant (P<0.05), corrected for multiple comparisons using the method of Friston et al. (1995).

Our analysis was centered on the comparison of the BOLD signal related to the presentation of the feedback in: 1) error explorative trials versus correct exploitative trials (i.e. "all error explorative trials minus exploitative trials"), 2) first correct explorative trials versus correct exploitative trials (i.e. "all first correct explorative trials minus exploitative trials"), and 3) correct exploitative trials versus first correct explorative trials (i.e. "exploitative trials minus all first correct explorative trials"). The primary purpose of this experiment was to study the set of areas involved in evaluating feedback indicating the need to switch response (i.e. at the occurrence of an error feedback) or to switch strategy (i.e. at the occurrence of a first positive feedback) during exploration.

Mean percent of BOLD signal change was measured in each region of interest in both contrasts "all error explorative trials minus exploitative trials" and "all first correct explorative trials minus exploitative trials" as follows: first, the regions of interest were identified in each single subject on the basis of each one of the two above contrasts (i.e. "all error explorative trials minus exploitative trials" and "all first correct explorative trials minus exploitative trials"). Second, the percent of BOLD signal change was measured in the comparisons "I3 minus R" (corresponding to a RPE=−0.5), "I2 minus R" (RPE=−0.33), and "I1 minus R" (RPE=−0.25), within a gray matter volume of a 5 mm^3 radius centered on the peaks observed in each subject and in each region of interest revealed in the overall comparison "all error explorative trials minus exploitative trials". The percent of BOLD signal change was also measured using the same procedure in the comparisons "C4 minus R" (RPE=0), "C3 minus R" (RPE=0.5), "C2 minus R" (RPE=0.66), and "C1 minus R (RPE=0.75)" in each subject and in each region of interest revealed in the overall comparison "all first correct explorative trials minus exploitative trials". Finally, the percent of BOLD signal change obtained in the 11 subjects was averaged for each comparison.

Results

Behavioral results

The percent of optimally solved problems (i.e. problems in which no perseverative error during the explorative period and no response shift during the exploitative period were observed) decreased with the length of the explorative period ($F(3,788)=4.2879$, P<0.0052, one-way ANOVA) (Fig. 1D). This difference in performance was expected because of the higher cognitive demand in long searches.

The subject's reaction times (RTs) were analyzed as follows: I1 and C1 trials were pooled because both trials correspond to a 25% probability of choosing the correct stimulus on the current trial. Similarly, I2 and C2 trials (corresponding to a probability of 33%) and I3 and C3 trials (corresponding to a probability of 50%) were pooled. Finally, C4 was analyzed separately because it corresponds to a probability of 100% to select the correct stimulus. The subjects' results revealed that RTs in I1/C1 trials were slower than in the other trial types ($F(4,3456)=31.599$, P<0.0001, one-way ANOVA, Fig. 1E), suggesting that the RTs are increased when the probability of making an erroneous response is high. Such phenomenon has been shown in previous studies in monkeys (Procyk and Goldman-Rakic, 2006; Procyk and Joseph, 1996; Procyk et al., 2000).

Overall fMRI results

We first assessed whether the ACC was only involved in error processing or rather in the analysis of pertinent feedback leading to response adaptation (after the occurrence of an error feedback) and strategy adaptation (after the occurrence of the first positive feedback) during the explorative period of the task. Both the contrasts "all error explorative trials minus exploitative trials" and "all first correct explorative trials minus exploitative trials" revealed an increase in activity in the rostral ACC (rACC), respectively at MNI coordinates (x, y, z) 9, 34, 22 (t-value=4.06) and 8, 34, 21 (t-value=5.41), i.e. at a similar location (Fig. 2A, Table 1; paired *t*-test using y-coordinates: P=0.5117). Interestingly, both a group analysis and a subject by subject analysis confirmed that the rACC region involved in the processing of both error and first positive feedback in the explorative period overlapped perfectly. This region was located in the cingulate sulcus immediately dorsal to the genu of the corpus callosum (see Fig. 2B). A conjunction analysis (see Friston et al., 2005) between the contrasts "all error explorative trials minus exploitative trials" and "all first correct explorative trials minus exploitative trials" confirms that the rACC in the right hemisphere was commonly activated in both contrasts (at MNI coordinates (x, y, z) 8, 34, 22 (t-value=4.69)), together with area 32/8 (at MNI coordinates X=2, Y=32, Z=40, t-value=5.57), and area 9/46 (at MNI coordinates X=40, Y=34, Z=30, t-value=4.80) (see Table 2, Fig. 2A). To assess whether rACC activity was modulated by the valence of the feedback during explorative trials (error versus first correct feedback), we contrasted error explorative and first correct explorative trials. There was no effect within the rACC (t-value=−1.26 at MNI coordinates 8, 34, 22). Although this negative result should be interpreted with care due to the small sample size, it might reflect the multiplicity of information encoding in the rACC (see Discussion).

More precisely, the increased activity observed in the rACC in both the contrasts "all error explorative trials minus exploitative trials" and "all first correct explorative trials minus exploitative trials" was located in the ventral bank of the cingulate sulcus in 8 of the 11 subjects in the comparison "all error explorative trials minus exploitative trials" and in 9 of the 11 subjects in the comparison "all first correct explorative trials minus exploitative trials". The other increases in activity (3 subjects in the comparison "all error explorative trials minus exploitative trials" and 2 subjects in the comparison "all first correct explorative trials minus exploitative trials") were located in the dorsal bank of the cingulate sulcus. Finally, only one subject showed an increase in activity in the cingulate gyrus, and not in the sulcus, in the comparison "all first correct explorative trials minus exploitative trials".

On the medial wall of the frontal cortex, in the right hemisphere, the contrasts "all error explorative trials minus exploitative trials" and "all first correct explorative trials minus exploitative trials" also revealed a separate cluster of increased activity in area 32/8 (Table 1, Fig. 2A). In the prefrontal cortex, both contrasts revealed increased activity in the mid-dorsolateral prefrontal (area 9/46 bilaterally and area 46 in the right hemisphere), in the orbitofrontal cortex (area 11 bilaterally), in the frontopolar cortex (area 10 bilaterally), and in the fronto-insular cortex, bilaterally. Increased activity in both contrasts was also observed, bilaterally, in the striatum (Fig. S2) and in the mid-inferior parietal cortex. Furthermore, increased activity was observed in the retrosplenial cortex and the middle temporal cortex (area 21) in the right hemisphere (Table 1, Fig. 3). Overall, all activity increases were located at similar locations in both contrasts.

Table 3
Brain regions showing increased activity during the presentation of the feedback in the contrast "all correct exploitative trials minus all first correct explorative trials". All t-statistics were significant at $P<0.05$. Note that the frontopolar activity was located medially to the vertical branch of the intermediate sulcus, at the intersection between the lateral and the medial surface. This activity was located 2.5 cm dorsal, 2 cm medial, and 0.6 cm anterior to the one observed in the contrasts "all error explorative trials minus exploitative trials" or "all first correct explorative trials minus exploitative trials".

Feedback: all correct exploitative minus all first correct explorative trials				
	MNI coordinates			
	x	y	z	t-Statistic
Right hemisphere				
Ventromedial prefrontal cortex	6	39	−14	5.158
Parietal operculum	64	−11	13	5.615
	41	−20	18	5.224
Superior temporal gyrus	54	−29	10	5.810
Posterior insula	40	−26	6	4.897
Dorsal posterior cingulate cortex (area 23)	7	−17	48	5.484
Ventral posterior cingulate cortex (area 23)	3	−56	18	4.767
Primary motor cortex	2	−27	63	5.028
	62	0	6	5.484
Left hemisphere				
Frontopolar cortex	−10	64	25	5.484
Ventromedial prefrontal cortex	−3	38	−17	6.397
Orbitofrontal cortex (area 13)	−38	32	−13	5.289
Middle temporal gyrus	−60	0	−19	5.289
Superior temporal gyrus	−40	−58	28	5.550
Ventral posterior cingulate cortex (area 23)	−6	−55	14	7.637
Hippocampus	−24	−22	−13	5.419

Finally the contrast "exploitative trials minus all first correct explorative trials" revealed that numerous brain regions were more active during the exploitative period than in the explorative period (see Table 3). They include the frontopolar cortex, the ventromedial prefrontal cortex, the orbitofrontal cortex (area 13), the right dorsal posterior cingulate cortex, and the ventral posterior cingulate cortex (Fig. 4).

Reward prediction errors

Following this overall analysis, we examined, within all the brain regions described above, the relationships between the feedback-related activity and RPE values associated with the various trial types (i.e. trials I3, I2, I1, C4, C3, C2, and C1). Concerning positive RPE values (see green squares in Fig. 5), the average percent of BOLD signal change was linearly modulated in the right rACC as follows: the higher the positive RPE value, the higher the average percent of BOLD signal change ($F(3,40)=3.8362$, $P=0.0167$, one-way ANOVA, Table S1). Similar results were observed in the right hemisphere within the fronto-insular cortex ($F(3,36)=7.1438$, $P=0.0001$, one-way ANOVA; Fig. 5B and Table S1), the striatum (including the ventral caudate nucleus and putamen; $F(3,40)=3.1952$, $P=0.0336$, one-way ANOVA; Fig. 5C and Table S1), the retrosplenial cortex ($F(3,40)=5.4647$, $P=0.003$, one-way ANOVA; Fig. 6A and Table S1), and the mid-dorsolateral prefrontal cortical area 46 ($F(3,32)=2.9048$, $P=0.0498$, one-way ANOVA; Fig. 6B and Table S1). In the left hemisphere, note that the same pattern of modulation of the percent of BOLD signal was observed in the left fronto-insular cortex ($F(3,40)=2.8885$, $P=0.0473$, one-way ANOVA).

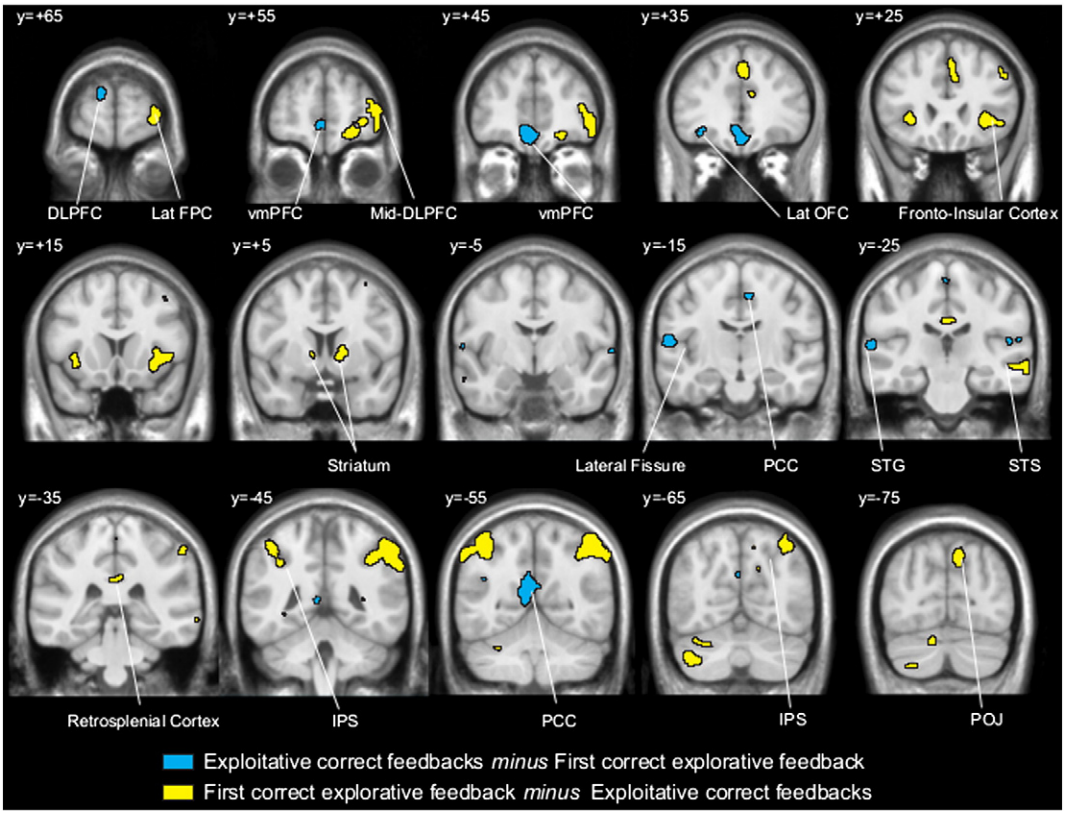
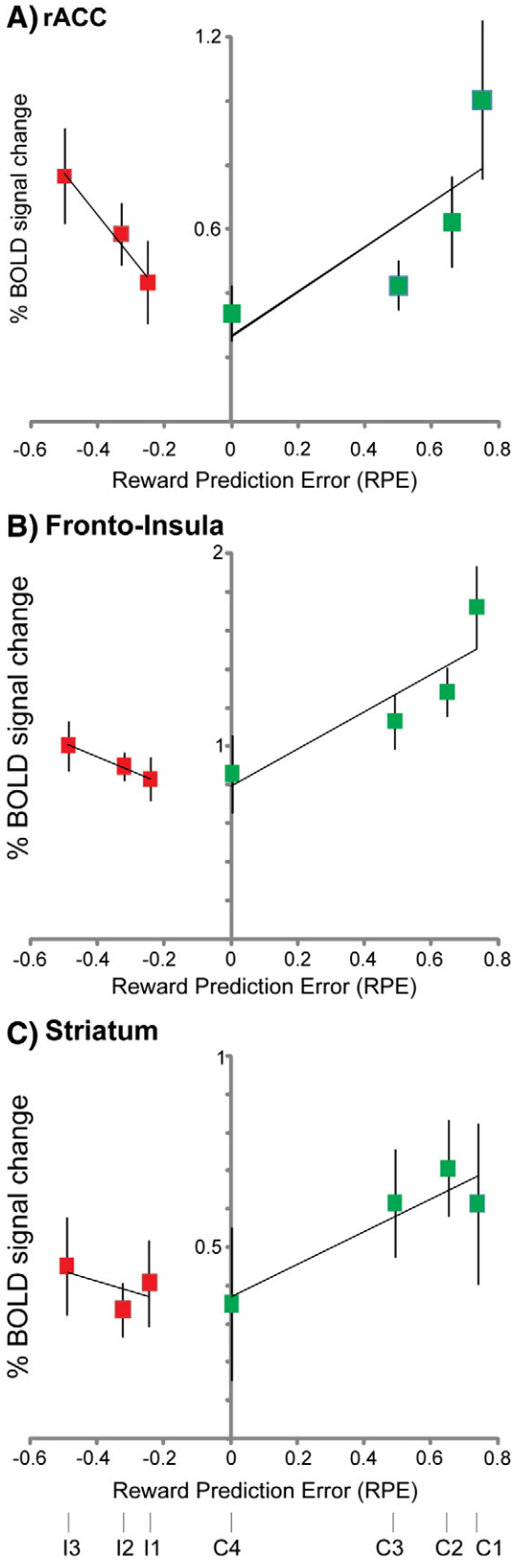


Fig. 4. Activity increases during the presentation of the feedback in the comparison "exploitative correct feedback minus all first correct explorative feedback" (blue clusters) and in the comparison "all first correct explorative feedback minus exploitative correct feedback" (yellow clusters). The average activity increases is obtained from all 11 subjects. The t statistical map of activity has been superimposed on the average T1 anatomical acquisition of the 11 subjects transformed into standard stereotaxic space. The Y value corresponds to the antero-posterior level of the stereotaxic coordinates expressed in mm within the MNI stereotaxic proportional system. All t-values were above the statistical threshold of $t=4.509$ (see Material and methods).

Concerning negative RPE (see red squares in Fig. 5), all the brain regions exhibiting a significant modulation of the percent of BOLD signal change at the occurrence of positive RPE display no significant modulation at the occurrence of negative RPE (see Supplemental data, Table S1). However, this result should be interpreted carefully given that the range of values used to assess sensitivity to negative RPE was shorter than the one used to assess sensitivity to positive RPE (i.e. whereas positive RPE values varied from 0 to 0.75, negative RPE values varied from −0.25 to −0.5), and given the small sample size used in the present study.



Interestingly, the pattern observed in the right retrosplenial cortex (Fig. 6A and Table S2), in the right mid-dorsolateral prefrontal cortex (Fig. 6B and Table S1), and in the left intraparietal cortex (Fig. 6C and Table S1) was very different from the one observed in the four other regions that showed a modulation of the percent of BOLD signal change by RPE values. The average percent of the BOLD signal change was highest in I1 and C1 trials (i.e. at the beginning of the explorative period) and lowest in I3 and C4 trials (i.e. at the end of the explorative period).

Discussion

The present study confirms some aspects of the two main hypotheses concerning the role of the rACC: 1) the rACC evaluates feedbacks in situations where they are salient for learning optimal strategies and 2) the rACC is sensitive to both feedback values and expectations. Our results demonstrate, first, that the evaluation of positive and negative feedback during trial and error exploration engages a common set of brain regions. Second, among those brain regions, the rACC, the striatum, and the fronto-insula performed both a categorical and a reinforcement-based evaluation of feedback during exploration.

The same rACC region analyzes error and first correct trials during exploration

The present data show that, within the rACC region, the same locus is involved in the analysis of both incorrect and first positive feedback during the exploration period of the task (Fig. 2). This finding shows that this region is involved in the analysis of feedback indicating the need either for response adaptation (i.e. to switch response and continue to explore at the occurrence of incorrect feedback) or for strategy adaptation (i.e. to stop exploring and start exploiting at the occurrence of a first positive feedback). This activation focus was located mainly within the cingulate sulcus. None of our subjects displayed a paracingulate sulcus in the right hemisphere (see Supplemental data), thus reducing rACC inter-subject variability (Palomero-Gallagher et al., 2008). The region of interest concerned mainly the 3rd cluster described by Beckmann et al. (Beckmann et al., 2009). This rACC subregion is likely to be the homologue of the anterior portion of the monkey cingulate sulcus (Dum and Strick, 1991).

Importantly, in recent electrophysiological studies in monkey using the same task as the one used in the present study, Quilodran et al. (2008) showed that the cortex within the dorsal bank of the cingulate sulcus processed both negative and positive feedback, specifically in the exploration period of the task, and Rothe et al. (2011) showed that this cingulate region interacted with the dorsolateral prefrontal cortex (DLPFC) only after such feedback indicating the need to update behavioral response or strategy. Interestingly, the present data also demonstrated increased activity within the mid-DLPFC (as in the

Fig. 5. Average (±s.e.m.) percent of BOLD signal change within the rACC (A), the frontal insula (B), and the striatum (C) in the right hemisphere from the 11 subjects in the comparisons "I3 minus R trials" (corresponding to a RPE = −0.5), "I2 minus R trials" (RPE = −0.33), "I1 minus R trials" (RPE = −0.25), "C4 minus R trials" (RPE = 0), "C3 minus R trials" (RPE = 0.5), "C2 minus R trials" (RPE = 0.66), and "C1 minus R trials" (RPE = 0.753). The mean percent of BOLD signal change varied significantly with the RPE value in the rACC ($F(6,70) = 2.5809$, $P < 0.02568$, one-way ANOVA), in the fronto-insula ($F(6,66) = 3.8634$, $P < 0.00230$, one-way ANOVA), and in the striatum ($F(6,70) = 3.5642$, $P < 0.00387$, one-way ANOVA). Red and green squares represent RPE values associated with negative and positive feedback respectively. Trial types corresponding to the different RPE values are indicated below the bottom figure. The diagonals represent linear regressions.

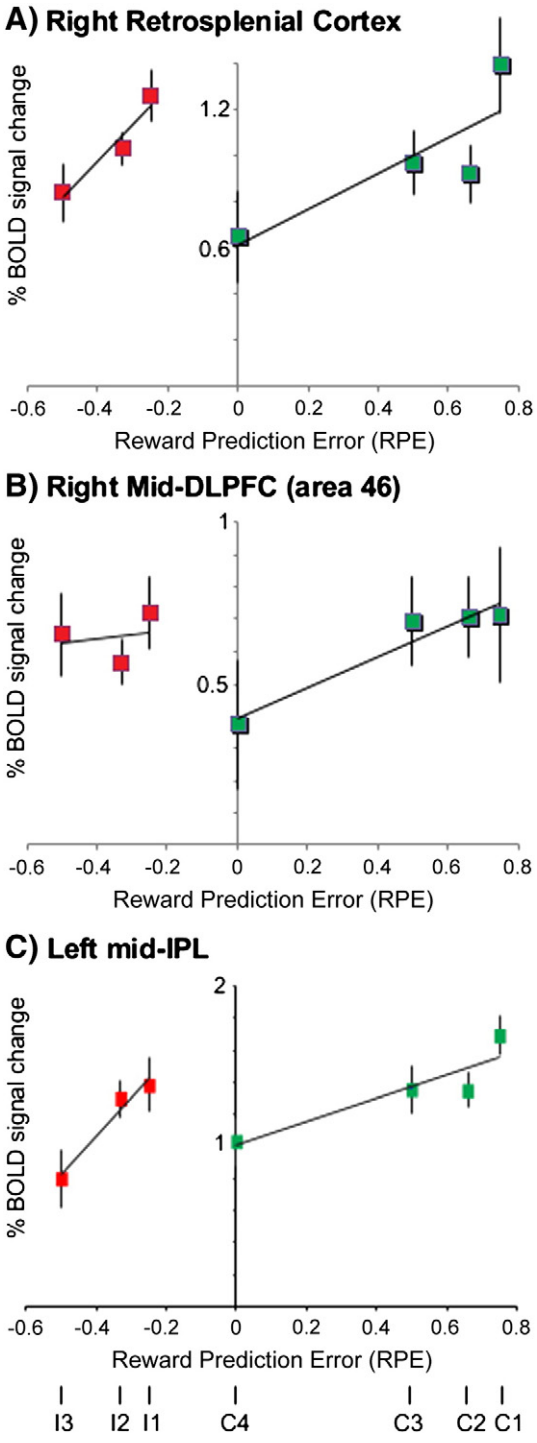


Fig. 6. Average (±s.e.m.) percent of BOLD signal change within the right retrosplenial cortex (A), the right mid-DLPFC area 46 (B), and the left intraparietal sulcal cortex (C) from the 11 subjects in the comparisons "I3 minus R trials" (corresponding to a RPE = −0.5), "I2 minus R trials" (RPE = −0.33), "I1 minus R trials" (RPE = −0.25), "C4 minus R trials" (RPE = 0), "C3 minus R trials" (RPE = 0.5), "C2 minus R trials" (RPE = 0.66), and "C1 minus CR trials" (RPE = 0.753). The mean BOLD signal change varied significantly with the positive RPE value in the retrosplenial cortex ($F(6,61) = 2.6765$, $P < 0.02265$, one-way ANOVA) and in the DLPFC ($F(3,32) = 2.9048$, $P = 0.0498$, one-way ANOVA). It varies for both negative and positive RPE in the IPL ($R^2 = 0.1105$, $F = 4.7213$, $P = 0.0361$ for positive RPE values; $R^2 = 0.1410$, $F = 5.0891$, $P = 0.0313$ for negative RPE values). Red and green squares represent RPE values associated with negative and positive feedback respectively. Trial types corresponding to the different RPE values are indicated below the bottom figure. The diagonals represent linear regressions.

rACC) at the occurrence of both incorrect and first positive feedback during the exploration period of the task. Based on the results of Rothe et al. (2011), we therefore suggest that the rACC and the mid-DLPFC may interact to process feedback indicating the need for response or strategy adjustment.

Our data, together with those of several electrophysiological studies in monkeys (Amiez et al., 2006; Hayden and Platt, 2010; Hayden et al., 2011a; Kennerley and Wallis, 2009; Quilodran et al., 2008; Sallet et al., 2007; Seo and Lee, 2009) suggest that the rACC is critical for the evaluation of feedback (positive and negative) when it is relevant for behavioral adaptation. Although the lack of significant difference when comparing activity related to the two different valences of feedback should be interpreted with care, overlapping activity in rACC for both negative and positive feedback was expected from monkey experiments showing that distinct populations of cells from the same rACC subregion encoded both positive and error feedback (Ito et al., 2003; Kennerley and Wallis, 2009; Quilodran et al., 2008; Sallet et al., 2007). These monkey experiments excluded a pure coding of saliency in ACC. Although saliency is a possible confound in the present experiment, we must recall that several neuroimaging studies in humans using different protocols have shown that rACC activation does not reflect a response to saliency, as opposed to activation observed in the ventrolateral prefrontal cortex, the anterior insula, or the supplemental motor area for instance (Amiez and Petrides, 2007; Litt et al., 2011). Whereas the dissociation between value and saliency effects can be established using single cell recordings, it is difficult to assess using fMRI (see discussion in Wallis and Rich, 2011), and our protocol does not clearly separate feedback evaluation from saliency. Future fMRI studies designed for efficient multi-voxel pattern analyses (Haxby, 2012; Norman et al., 2006) are required to evaluate whether relative activation patterns for positive and negative feedback in rACC can be disambiguated in the current protocol and to help eliminate saliency as a confound for reinforcement signals (Vickery et al., 2011).

Modulation of feedback-related activity by uncertainty during exploration

In order to assess whether the activity within brain regions involved in the analysis of both error and first correct feedback was modulated by expectations, we tested whether the BOLD signal in these regions was modulated by positive and negative reward prediction error (RPE) values. The results demonstrated that the brain regions involved could be classified into two groups: regions in which the percent of BOLD signal change was modulated by RPE and regions in which the BOLD signal was not modulated by the RPE.

First, the percent of BOLD signal change within rACC was modulated by positive RPE. This result is of critical importance because it supports a link between ACC function and reinforcement-learning and, possibly a role of afferent dopaminergic signals that encode RPE in modulating ACC activity (Cohen et al., 2007; Holroyd et al., 2003; Matsumoto et al., 2007; Quilodran et al., 2008; Rushworth and Behrens, 2008). Event-related potential studies have not been consistent in showing modulations of feedback-related potentials by RPE (Cohen et al., 2007; Hajcak et al., 2006; Holroyd et al., 2003). Our results are consistent with a recent EEG study showing that medial frontal feedback-related potentials are modulated during reward-based learning (Cohen et al., 2007), with a recent fMRI study showing that, during behavioral adaptation, ACC outcome-related activation varies with changing reward contingencies (Behrens et al., 2007), and with an fMRI study in which model-based analyses of BOLD signal using Q-Learning algorithms and instrumental learning showed a correlation between ACC activity and RPE (Haruno and Kawato, 2006a).

Second, the BOLD signal within rACC was not significantly modulated by negative RPE. A higher sensitivity to positive compared to negative events as in (Mirenowicz and Schultz, 1996) combined with the small range of values used for the negative RPE (from −0.25 to −0.5) and with the fact that the sample size used in the present study was small, might explain this result and therefore does not, at present, allow a

clear conclusion. However, we should note that, in monkey rACC, negative prediction error signals are less frequently reported than positive prediction error signals (Kennerley and Walton, 2011).

Notwithstanding these data, it must be emphasized that the role of DA in rACC for processing reward prediction error is far from clear. This is due to inconclusive evidence from human experiments, lack of clear and direct experimentation using invasive techniques in monkeys, and from a mismatch between the theoretical propositions and the dynamical properties of DA cortical release (Cohen et al., 2011; Lapish et al., 2007). Yet, monkey experiments have shown that DA can act on dissociated processes in prefrontal cortex via different receptors and at the scale of single trials in working memory tasks (Wang et al., 2004). The dual-state theory of prefrontal cortex dopamine function (Durstewitz and Seamans, 2008) proposed that DA action on prefrontal networks favors or restricts cognitive flexibility, a gating mechanism allowing appropriate updating of cognitive control (Montague et al., 2004). Indeed, rACC and lateral prefrontal activity shows clear changes between search and repetition trials (Procyk and Goldman-Rakic, 2006; Quilodran et al., 2008). Additional experiments are thus required to test whether and how DA release in rACC at the time of feedback participates in behavioral shift. The problem of RPE encoding, or activity modulation by RPE, within the rACC is also made more complex by the cortical dissociation of responses to positive and negative feedback (see above; Kennerley and Walton, 2011; Quilodran et al., 2008). For that reason, BOLD response, as measured here in rACC, violates the first rule of the axiomatic model of RPE (Caplin and Dean, 2008).

A modulation by positive RPE was also observed in the striatum and the fronto-insula in the right hemisphere. First, the increased activity in the dorsal striatum was located in the caudate nucleus and extended to the putamen in most subjects (Fig. 3; Supplemental material). These data confirm that the dorsal striatum participates in reward processing, and confirm its sensitivity to RPE signals (Amiez et al., 2012; Apicella et al., 2009, 2011; Brovelli et al., 2008; Haruno and Kawato, 2006b; Jessup and O'Doherty, 2011; Lau and Glimcher, 2007; Yamada et al., 2011; Zink et al., 2004). Interestingly, using a variant of the problem-solving task, Yamada et al. reported that different populations of striatal cells code for positive feedback, for error feedback, for probability, and for search-repeat effect (Yamada et al., 2011). Therefore, as the rACC, the dorsal striatum might perform a reinforcement-based evaluation of action outcomes as well as categorical-based evaluation. Second, rostral insula activity is known to be correlated with risk prediction error (Preuschoff et al., 2008). However, our experimental protocol does not permit us to disentangle RPE and risk prediction error (see Supplementary data) and it does not allow us to test alternatives hypotheses that have been proposed to explain activity of the insula. For instance, recent studies have suggested that the insula may encode the variance in the RPE (d'Acremont et al., 2009) or the salience of an outcome (Rutledge et al., 2010).

Finally, in sharp contrast to the above mentioned data, the percent of BOLD signal change in the right retrosplenial cortex (Fig. 6A), the right mid-DPFC (area 46; Fig. 6B), and the left intraparietal cortex (Fig. 6C) was higher at the occurrence of feedback presented at the beginning of the exploratory period. Instead of reflecting motivational salience or value, we suggest that those BOLD signal modulations reflect the decrease of working memory demands as the exploratory period progresses and, therefore, the number of options available decreases. This interpretation is supported by several studies carried out in both monkey and human subjects. For instance, the mid-DLPFC has been shown to be critically involved in the tracking (monitoring) of choices and other events in working memory in both monkeys (Petrides, 1991, 1995) and humans (Petrides et al., 1993, 2002). Recently, the intraparietal cortex has also been shown to be involved in the manipulation aspect of working memory processing (Champod and Petrides, 2007). The retrosplenial cortex, which is known to be involved in spatial navigation and learning (St-Laurent et al., 2009), is directly connected to the mid-DLPFC (Morris et al., 1999), and probably provides the means by which the mid-DLPFC can interact with the memory system so that choices and events in memory can be tracked (Petrides, 1996).

Finally, a difference in the current experimental protocol might explain differences between our results and those of previous published experiments. For instance, the binary property of the feedback (incorrect or correct) might explain why regions sensitive to reward values, for instance the frontopolar cortex, were not modulated (Boorman et al., 2009, 2011). The fact that we did not observe BOLD signal modulation with regard to RPE within the dorsomedial prefrontal cortex and the right intraparietal cortex may reflect their involvement in action update rather than in action valuation (Boorman et al., 2009; Isoda and Hikosaka, 2007).

Neural systems involved in the analysis of feedback during exploitation

While many studies focused on feedback processing in an exploratory-like environment (Behrens et al., 2007; Boorman et al., 2009, 2011; Brovelli et al., 2008; Daw et al., 2006; Walton et al., 2004; Yoshida and Ishii, 2006), rare are the experiments assessing the brain areas engaged during exploitative behavior (Landmann et al., 2007). Our data confirmed that the left rostral mid-DLPFC, the ventromedial prefrontal cortex (VMPFC), and area 23c of the posterior cingulate cortex (Vogt et al., 2006) are more engaged in the exploitative period. VMPFC activity may reflect the value of the chosen action compared to the alternative options (Boorman et al., 2009). However, it was shown that the higher the probability of the outcome, the higher the activity in the VMPFC (Daw et al., 2006). This might explain why feedback-related activity is stronger for correct feedback occurring in the exploitative period than for correct feedback occurring in the exploratory period. Posterior cingulate activity might reflect cognitive disengagement during the exploitative period of the task (Hayden et al., 2010).

Conclusion

The present study demonstrates the involvement of a set of brain regions involved in the evaluation of positive and negative feedback during exploration. Such evaluation is likely to carry information necessary for appropriate adaptation, such as maintaining exploration after errors or shifting toward exploitation following the occurrence of the first positive feedback. Among those brain structures, five regions displayed a percent of BOLD signal change that was modulated by the positive RPE. The pattern of BOLD signal modulation observed in the rACC, anterior insula, and the striatum in relation to positive RPE values illustrates the critical involvement of these regions in performing reinforcement-based evaluation of feedback in the exploratory period. By contrast, the pattern of BOLD signal modulation observed in the mid-DLPFC, the intraparietal cortex, and the retrosplenial cortex reflects the involvement of these structures in the working memory constraints inherent to the exploratory period.

Altogether these data suggest that feedback information content, uncertainty about feedback occurrence, and feedback valence are important factors, hereby joining the recent view that rACC activation reflects volatility or uncertainty of outcome estimates (Behrens et al., 2007; Cohen et al., 2007; Oliveira et al., 2007; Walton et al., 2004).

Acknowledgments

This work was supported by the Canadian Institutes of Health Research (CIHR) grant FRN 37753 to Michael Petrides. Jérôme Sallet was supported by the Bettencourt Schueller Foundation. Céline Amiez was supported by a CIHR fellowship and by Neurodis Foundation. Emmanuel Procyk was supported by ANR JCJC 0048. We thank

Louise Marcotte, Ronaldo Lopez and David Costa for their technical assistance. We thank Emily Rubin-Ferreira and MaryAnn P. Noonan for their comments on the manuscript and Erie D. Boorman and Nils Kolling for the helpful discussions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2012.06.023.

References

Amiez, C., Petrides, M., 2007. Selective involvement of the mid-dorsolateral prefrontal cortex in the coding of the serial order of visual stimuli in working memory. Proc. Natl. Acad. Sci. 104 (34), 13786–13791.

Amiez, C., Joseph, J.P., Procyk, E., 2005. Anterior cingulate error-related activity is modulated by predicted reward. Eur. J. Neurosci. 21, 3447–3452.

Amiez, C., Joseph, J.P., Procyk, E., 2006. Reward encoding in the monkey anterior cingulate cortex. Cereb. Cortex 16, 1040–1055.

Amiez, C., Hadj-Bouziane, F., Petrides, M., 2012. Response selection versus feedback analysis in conditional visuo-motor learning. Neuroimage 59, 3723–3735.

Apicella, P., Deffains, M., Ravel, S., Legallet, E., 2009. Tonically active neurons in the striatum differentiate between delivery and omission of expected reward in a probabilistic task context. Eur. J. Neurosci. 30, 515–526.

Apicella, P., Ravel, S., Deffains, M., Legallet, E., 2011. The role of striatal tonically active neurons in reward prediction error signaling during instrumental task performance. J. Neurosci. 31, 1507–1515.

Beckmann, M., Johansen-Berg, H., Rushworth, M.F., 2009. Connectivity-based parcellation of human cingulate cortex and its relation to functional specialization. J. Neurosci. 29, 1175–1190.

Behrens, T.E., Woolrich, M.W., Walton, M.E., Rushworth, M.F., 2007. Learning the value of information in an uncertain world. Nat. Neurosci. 10, 1214–1221.

Boorman, E.D., Behrens, T.E., Woolrich, M.W., Rushworth, M.F., 2009. How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. Neuron 62, 733–743.

Boorman, E.D., Behrens, T.E., Rushworth, M.F., 2011. Counterfactual choice and learning in a neural network centered on human lateral frontopolar cortex. PLoS Biol. 9, e1001093.

Brovelli, A., Laksiri, N., Nazarian, B., Meunier, M., Boussaoud, D., 2008. Understanding the neural computations of arbitrary visuomotor learning through fMRI and associative learning theory. Cereb. Cortex 18, 1485–1495.

Bullmore, E., Brammer, M., Williams, S.C., Rabe-Hesketh, S., Janot, N., David, A., Mellers, J., Howard, R., Sham, P., 1996. Statistical methods of estimation and inference for functional MR image analysis. Magn. Reson. Med. 35, 261–277.

Caplin, A., Dean, M., 2008. Axiomatic methods, dopamine and reward prediction error. Curr. Opin. Neurobiol. 18 (2), 197–202.

Champod, A.S., Petrides, M., 2007. Dissociable roles of the posterior parietal and the prefrontal cortex in manipulation and monitoring processes. Proc. Natl. Acad. Sci. U. S. A. 104, 14837–14842.

Cohen, M.X., Elger, C.E., Ranganath, C., 2007. Reward expectation modulates feedback-related negativity and EEG spectra. Neuroimage 35, 968–978.

Cohen, M.X., Wilmes, K., Vijver, I., 2011. Cortical electrophysiological network dynamics of feedback learning. Trends Cogn. Sci. 15 (12), 558–566.

Collins, D.L., Neelin, P., Peters, T.M., Evans, A.C., 1994. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. J. Comput. Assist. Tomogr. 18, 192–205.

Cox, R.W., Jesmanowicz, A., 1999. Real-time 3D image registration for functional MRI. Magn. Reson. Med. 42, 1014–1018.

Croxson, P.L., Walton, M.E., O'Reilly, J.X., Behrens, T.E., Rushworth, M.F., 2009. Effort-based cost–benefit valuation and the human brain. J. Neurosci. 29, 4531–4541.

d'Acremont, M., Lu, Z.L., Li, X., Van der Linden, M., Bechara, A., 2009. Neural correlates of risk prediction error during reinforcement learning in humans. Neuroimage 47, 1929–1939.

Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B., Dolan, R.J., 2006. Cortical substrates for exploratory decisions in humans. Nature 441, 876–879.

Dum, R.P., Strick, P.L., 1991. The origin of corticospinal projections from the premotor areas in the frontal lobe. J. Neurosci. 11, 667–689.

Durstewitz, D., Seamans, J.K., 2008. The dual-state theory of prefrontal cortex dopamine function with relevance to catechol-o-methyltransferase genotypes and schizophrenia. Biol. Psychiatry 64 (9), 739-479.

Friston, K.J., Holmes, A.P., Poline, J.B., Grasby, P.J., Williams, S.C., Frackowiak, R.S., Turner, R., 1995. Analysis of fMRI time-series revisited. Neuroimage 2, 45–53.

Friston, K.J., Penny, W.D., Glaser, D.E., 2005. Conjunction revisited. Neuroimage 25, 661–667.

Hadland, K.A., Rushworth, M.F., Gaffan, D., Passingham, R.E., 2003. The anterior cingulate and reward-guided selection of actions. J. Neurophysiol. 89, 1161–1164.

Hajcak, G., Moser, J.S., Holroyd, C.B., Simons, R.F., 2006. The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. Biol. Psychol. 71, 148–154.

Haruno, M., Kawato, M., 2006a. Heterarchical reinforcement-learning model for integration of multiple cortico-striatal loops: fMRI examination in stimulus-action-reward association learning. Neural Netw. 19, 1242–1254.

Haruno, M., Kawato, M., 2006b. 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, 948–959.

Haxby, J.V., 2012. Multivariate pattern analysis of fMRI: the early beginnings. Neuroimage 62 (2), 852–855.

Hayden, B.Y., Platt, M.L., 2010. Neurons in anterior cingulate cortex multiplex information about reward and action. J. Neurosci. 30, 3339–3346.

Hayden, B.Y., Smith, D.V., Platt, M.L., 2010. Cognitive control signals in posterior cingulate cortex. Front. Hum. Neurosci. 4, 223.

Hayden, B.Y., Pearson, J.M., Platt, M.L., 2011a. Neuronal basis of sequential foraging decisions in a patchy environment. Nat. Neurosci. 14, 933–939.

Hayden, B.Y., Heilbronner, S.R., Pearson, J.M., Platt, M.L., 2011b. Surprise signals in anterior cingulate cortex: neuronal encoding of unsigned reward prediction errors driving adjustment in behavior. J. Neurosci. 31, 4178–4187.

Holroyd, C.B., Coles, M.G., 2002. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. Psychol. Rev. 109, 679–709.

Holroyd, C.B., Nieuwenhuis, S., Yeung, N., Cohen, J.D., 2003. Errors in reward prediction are reflected in the event-related brain potential. Neuroreport 14, 2481–2484.

Holroyd, C.B., Krigolson, O.E., Baker, R., Lee, S., Gibson, J., 2009. When is an error not a prediction error? An electrophysiological investigation. Cogn. Affect. Behav. Neurosci. 9, 59–70.

Isoda, M., Hikosaka, O., 2007. Switching from automatic to controlled action by monkey medial frontal cortex. Nat. Neurosci. 10, 240–248.

Ito, S., Stuphorn, V., Brown, J.W., Schall, J.D., 2003. Performance monitoring by the anterior cingulate cortex during saccade countermanding. Science 302, 120–122.

Jessup, R.K., O'Doherty, J.P., 2011. Human dorsal striatal activity during choice discriminates reinforcement learning behavior from the gambler's fallacy. J. Neurosci. 31, 6296–6304.

Jessup, R.K., Busemeyer, J.R., Brown, J.W., 2010. Error effects in anterior cingulate cortex reverse when error likelihood is high. J. Neurosci. 30, 3467–3472.

Jung, J., Jerbi, K., Ossandon, T., Ryvlin, P., Isnard, J., Bertrand, O., Guenot, M., Mauguiere, F., Lachaux, J.P., 2010. Brain responses to success and failure: direct recordings from human cerebral cortex. Hum. Brain Mapp. 31, 1217–1232.

Kennerley, S.W., Wallis, J.D., 2009. Evaluating choices by single neurons in the frontal lobe: outcome value encoded across multiple decision variables. Eur. J. Neurosci. 29, 2061–2073.

Kennerley, S.W., Walton, M.E., 2011. Decision making and reward in frontal cortex: complementary evidence from neurophysiological and neuropsychological studies. Behav. Neurosci. 125, 297–317.

Kennerley, S.W., Walton, M.E., Behrens, T.E., Buckley, M.J., Rushworth, M.F., 2006. Optimal decision making and the anterior cingulate cortex. Nat. Neurosci. 9, 940–947.

Landmann, C., Dehaene, S., Pappata, S., Jobert, A., Bottlaender, M., Roumenov, D., Le Bihan, D., 2007. Dynamics of prefrontal and cingulate activity during a reward-based logical deduction task. Cereb. Cortex 17, 749–759.

Lapish, C.C., Kroener, S., Durstewitz, D., Lavin, A., Seamans, J.K., 2007. The ability of the mesocortical dopamine system to operate in distinct temporal modes. Psychopharmacology (Berl) 191 (3), 609–625.

Lau, B., Glimcher, P.W., 2007. Action and outcome encoding in the primate caudate nucleus. J. Neurosci. 27, 14502–14514.

Litt, A., Plassmann, H., Shiv, B., Rangel, A., 2011. Dissociating valuation and saliency signals during decision-making. Cereb. Cortex 21, 95–102.

Matsumoto, M., Matsumoto, K., Abe, H., Tanaka, K., 2007. Medial prefrontal cell activity signaling prediction errors of action values. Nat. Neurosci. 10, 647–656.

Mirenowicz, J., Schultz, W., 1996. Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. Nature 379, 449–451.

Montague, P.R., Hyman, S.E., Cohen, J.D., 2004. Computational roles for dopamine in behavioural control. Nature 431 (7010), 760–767.

Morris, R., Pandya, D.N., Petrides, M., 1999. Fiber system linking the mid-dorsolateral frontal cortex with the retrosplenial/presubicular region in the rhesus monkey. J. Comp. Neurol. 407, 183–192.

Norman, K.A., Polyn, S.M., Detre, G.J., Haxby, J.V., 2006. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. Trends Cogn. Sci. 10 (9), 424–430.

Oliveira, F.T., McDonald, J.J., Goodman, D., 2007. Performance monitoring in the anterior cingulate is not all error related: expectancy deviation and the representation of action-outcome associations. J. Cogn. Neurosci. 19, 1994–2004.

Palomero-Gallagher, N., Mohlberg, H., Zilles, K., Vogt, B., 2008. Cytology and receptor architecture of human anterior cingulate cortex. J. Comp. Neurol. 508, 906–926.

Petrides, M., 1991. Monitoring of selections of visual stimuli and the primate frontal cortex. Proc. R. Soc. Lond. B Biol. Sci. 246, 293–298.

Petrides, M., 1995. Impairments on nonspatial self-ordered and externally ordered working memory tasks after lesions of the mid-dorsal part of the lateral frontal cortex in the monkey. J. Neurosci. 15, 359–375.

Petrides, M., 1996. Specialized systems for the processing of mnemonic information within the primate frontal cortex. Philos. Trans. R. Soc. Lond. B Biol. Sci. 351, 1455–1461 (discussion 1461–1452).

Petrides, M., Pandya, D.N., 2004. In: Paxinos, G., Mai, J.K. (Eds.), The Frontal Cortex. Elsevier Academic Press, San Diego, pp. 950–972.

Petrides, M., Alivisatos, B., Evans, A.C., Meyer, E., 1993. Dissociation of human mid-dorsolateral from posterior dorsolateral frontal cortex in memory processing. Proc. Natl. Acad. Sci. U. S. A. 90, 873–877.

Petrides, M., Alivisatos, B., Frey, S., 2002. Differential activation of the human orbital, mid-ventrolateral and mid-dorsolateral prefrontal cortex during the processing of visual stimuli. Proc. Natl. Acad. Sci. U. S. A. 99, 5649–5654.

Preuschoff, K., Quartz, S.R., Bossaerts, P., 2008. Human insula activation reflects risk prediction errors as well as risk. J. Neurosci. 28, 2745–2752.
Procyk, E., Goldman-Rakic, P.S., 2006. Modulation of dorsolateral prefrontal delay activity during self-organized behavior. J. Neurosci. 26 (44), 11313–11323.
Procyk, E., Joseph, J.P., 1996. Problem solving and logical reasoning in the macaque monkey. Behav. Brain Res. 82 (1), 67–78.
Procyk, E., Tanaka, Y.L., Joseph, J.P., 2000. Anterior cingulate activity during routine and non-routine sequential behaviors in macaques. Nat. Neurosci. 3 (5), 502–508.
Quilodran, R., Rothe, M., Procyk, E., 2008. Behavioral shifts and action valuation in the anterior cingulate cortex. Neuron 57, 314–325.
Rametti, G., Junqué, C., Bartrés-Faz, D., Zubiaurre-Elorza, L., Catalán, R., Penadés, R., Bargalló, N., Bernardo, M., 2010. Anterior cingulate and paracingulate sulci morphology in patients with schizophrenia. Schizophr. Res. 121 (1–3), 66–74.
Rothe, M., Quilodran, R., Sallet, J., Procyk, E., 2011. Coordination of high gamma activity in anterior cingulate and lateral prefrontal cortical areas during adaptation. J. Neurosci. 31, 11110–11117.
Rudebeck, P.H., Behrens, T.E., Kennerley, S.W., Baxter, M.G., Buckley, M.J., Walton, M.E., Rushworth, M.F., 2008. Frontal cortex subregions play distinct roles in choices between actions and stimuli. J. Neurosci. 28, 13775–13785.
Rushworth, M.F., Behrens, T.E., 2008. Choice, uncertainty and value in prefrontal and cingulate cortex. Nat. Neurosci. 11, 389–397.
Rushworth, M.F., Noonan, M.P., Boorman, E.D., Walton, M.E., Behrens, T.E., 2011. Frontal cortex and reward-guided learning and decision-making. Neuron 70, 1054–1069.
Rutledge, R.B., Dean, M., Caplin, A., Glimcher, P.W., 2010. Testing the reward prediction error hypothesis with an axiomatic model. J. Neurosci. 30, 13525–13536.
Sallet, J., Quilodran, R., Rothé, M., Vezoli, J., Joseph, J.P., Procyk, E., 2007. Expectations, gains, and losses in the anterior cingulate cortex. Cogn. Affect. Behav. Neurosci. 7, 327–336.
Seo, H., Lee, D., 2009. Behavioral and neural changes after gains and losses of conditioned reinforcers. J. Neurosci. 29, 3627–3641.
Shima, K., Tanji, J., 1998. Role for cingulate motor area cells in voluntary movement selection based on reward. Science 282, 1335–1338.
St-Laurent, M., Petrides, M., Sziklas, V., 2009. Does the cingulate cortex contribute to spatial conditional associative learning in the rat? Hippocampus 19, 612–622.
Vezoli, J., Procyk, E., 2009. Frontal feedback-related potentials in nonhuman primates: modulation during learning and under haloperidol. J. Neurosci. 29, 15675–15683.
Vickery, T.J., Chun, M.M., Lee, D., 2011. Ubiquity and specificity of reinforcement signals throughout the human brain. Neuron 72 (1), 166–177.
Vogt, B.A., Vogt, L., Laureys, S., 2006. Cytology and functionally correlated circuits of human posterior cingulate areas. Neuroimage 29, 452–466.
Wallis, J.D., Rich, E.L., 2011. Challenges of interpreting frontal neurons during value-based decision-making. Front. Neurosci., http://dx.doi.org/10.3389/fnins.2011.00124
Walton, M.E., Devlin, J.T., Rushworth, M.F., 2004. Interactions between decision making and performance monitoring within prefrontal cortex. Nat. Neurosci. 7, 1259–1265.
Wang, M., Vijayraghavan, S., Goldman-Rakic, P.S., 2004. Selective D2 receptor actions on the functional circuitry of working memory. Science 303 (5659), 853–856.
Worsley, K.J., Marrett, S., Neelin, P., Vandal, A.C., Friston, K.J., Evans, A.C., 1996. A unified statistical approach for determining significant signals in images of cerebral activation. Hum. Brain Mapp. 4, 58–73.
Worsley, K.J., Liao, C.H., Aston, J., Petre, V., Duncan, G.H., Morales, F., Evans, A.C., 2002. A general statistical analysis for fMRI data. Neuroimage 15, 1–15.
Yamada, H., Inokawa, H., Matsumoto, N., Ueda, Y., Kimura, M., 2011. Neuronal basis for evaluating selected action in the primate striatum. Eur. J. Neurosci. 34, 489–506.
Yoshida, W., Ishii, S., 2006. Resolution of uncertainty in prefrontal cortex. Neuron 50, 781–789.
Zink, C.F., Pagnoni, G., Martin-Skurski, M.E., Chappelow, J.C., Berns, G.S., 2004. Human striatal responses to monetary reward depend on saliency. Neuron 42, 509–517.