Neuron
Article

Behavioral Shifts and Action Valuation in the Anterior Cingulate Cortex

René Quilodran,[1,2] Marie Rothé,[1,2] and Emmanuel Procyk[1,2,*]
[1]Inserm, U846, Stem Cell and Brain Research Institute, 69500 Bron, France
[2]Université de Lyon, Lyon 1, UMR-S 846, 69003 Lyon, France
*Correspondence: procyk@lyon.inserm.fr
DOI 10.1016/j.neuron.2007.11.031

SUMMARY

Rapid optimization of behavior requires decisions about when to explore and when to exploit discovered resources. The mechanisms that lead to fast adaptations and their interaction with action valuation are a central issue. We show here that the anterior cingulate cortex (ACC) encodes multiple feedbacks devoted to exploration and its immediate termination. In a task that alternates exploration and exploitation periods, the ACC monitored negative and positive outcomes relevant for different adaptations. In particular, it produced signals specific of the first reward, i.e., the end of exploration. Those signals disappeared in exploitation periods but immediately transferred to the initiation of trials—a transfer comparable to learning phenomena observed for dopaminergic neurons. Importantly, these were also observed for high gamma oscillations of local field potentials shown to correlate with brain imaging signal. Thus, mechanisms of action valuation and monitoring of events/actions are combined for rapid behavioral regulation.

INTRODUCTION

Exploring is of primary importance for adapting to challenging situations, constructing novel internal maps, and developing the value of each choice in the context of new environments. Defining when exploration should be terminated is crucial as well. This refers to the ability of efficient adaptive systems to optimize performance and regulate the shift between exploring for rewards and exploiting known resources. The speed of adaptation is a key issue (Kawato and Samejima, 2007). Solutions emerge with models that alternate between flexible or controlled behavioral adaptations and poorly flexible long-term habituations (Daw et al., 2005a, 2005b; Doya, 2002; Sutton and Barto, 1998). One main difference between the two controls on behavior is the strong or weak impact of a particular outcome on the evaluation of the action that produced it (Daw et al., 2005b). An outcome can trigger immediate adaptation, as observed in many conditional protocols, or be used as just a piece of evidence weighted against recent reward history (Kennerley et al., 2006). Information issued from outcomes can either be used as positive and negative assessments of action value and/or as triggers to change or repeat a response. Where in the brain and how information about outcomes participates in fast action valuation and in shifting between behavioral modes is a central question.

Aston-Jones and Cohen recently proposed that a set of frontal areas (namely the anterior cingulate cortex [ACC] and the orbitofrontal cortex [OFC]) might have a key role in producing signals used to control the balance between exploratory and exploitative behavior (Aston-Jones and Cohen, 2005). The proposition is based on the hypothesis that OFC and ACC are critical players for processing rewards and costs, respectively. However, the role of these structures is still in debate. We focus on ACC functions, but see Rushworth and colleagues for a recent comparative review (Rushworth et al., 2007). The ACC has been reproducibly reported to activate during adaptive behavior and to shift activity between exploratory and exploitative behaviors (Procyk et al., 2000; Walton et al., 2004). The debates over ACC functions have often focused on its involvement and specificity in detection and evaluation of errors and/or conflict in processing. Holroyd and Coles proposed that ACC monitors behavioral errors based on negative prediction errors. The negative evaluation of performance would take place either at the level of motor response or at the level of external feedback indicating errors (Amiez et al., 2005; Holroyd and Coles, 2002). Some neurons in ACC code for a unidirectional error signal, a signal that could be dependent on dopaminergic afferences. This refers to the role of the dopaminergic system in reinforcement learning mechanisms by which prediction errors are used as teaching signals for neural plasticity (Schultz, 2006). Although numerous works have supported a role for ACC in error monitoring, other authors emphasize reports of ACC activation independent of error commission (Botvinick et al., 2004). Indeed, recent work showed mid-frontal event-related potentials possibly produced by the ACC and related to correct performance or to reward gain (Gehring and Willoughby, 2002; Vidal et al., 2003). Cohen and colleagues propose that ACC reacts whenever the current task induces conflict in processing for response selection, or even for feedback detection (Botvinick et al., 2004). The conflict-monitoring theory suggests that the expectancy violation due to the detection of an unexpected feedback can be conceived as a conflict. This formulation is in some aspects similar to the one invoking prediction errors. More recently, Rushworth and coworkers concluded that the ACC is critically involved in detecting both positive and negative outcomes for the purpose of action valuation, in other words, in a function that encompasses feedback detection and value adjustment (Rushworth

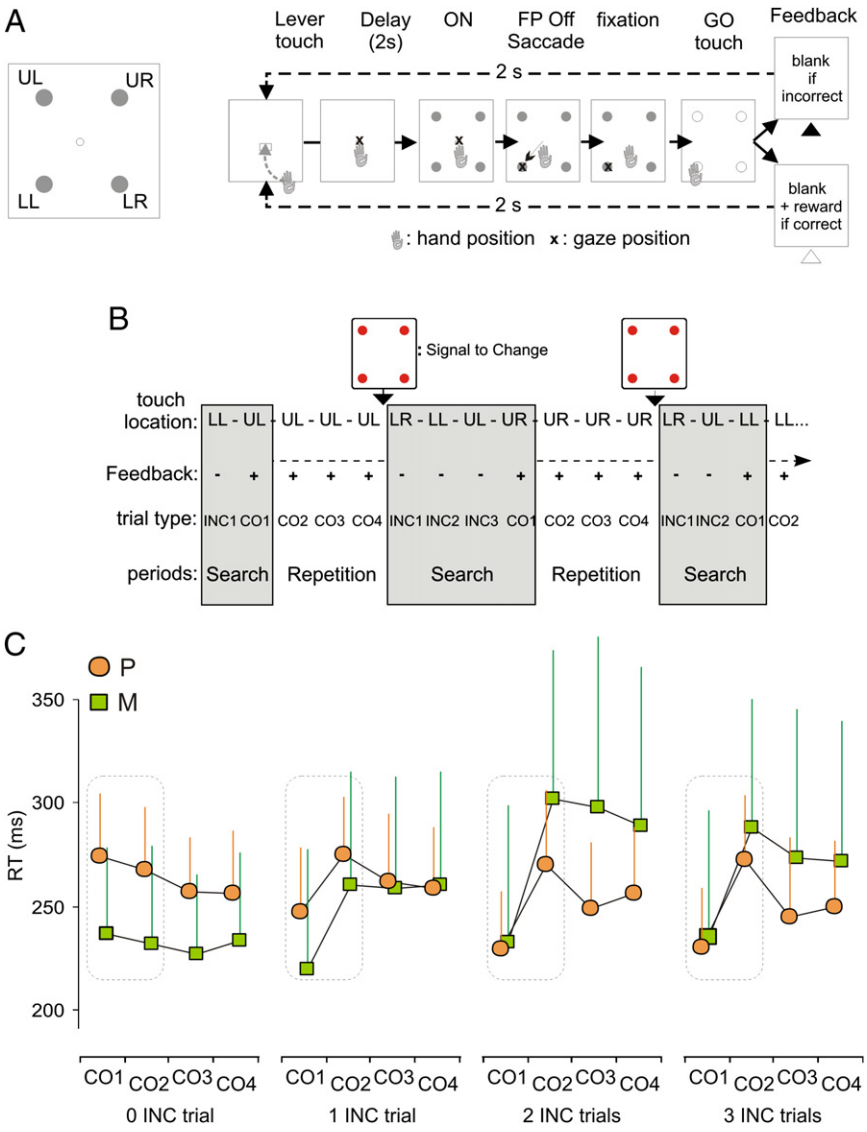

Figure 1. Behavioral Task and Performance

The animal had to search by trial and error for the correct target.

(A) A trial starts with a touch on the lever and onset of central fixation spot. After a delay period with eye fixation on the central spot, all four targets switched on, and the animal made a saccade toward and touched one of them. All targets switched off, and the feedback was given (no reward: negative, black arrowhead; reward: positive, white arrowhead).

(B) Example: In the first problem, the monkey discovered the solution (UL) in two trials (search period). After discovery, the animal was allowed to repeat the response (repetition period) (–, no reward; +, reward; INC, incorrect; CO, correct).

(C) Average reaction times (RT) for target touches (for monkeys M and P; 28 sessions each). RT for CO1, CO2, CO3, and CO4 are shown for problems with different search length (i.e., number of INC trials before the first correct CO1). Note differences for CO1 versus CO2 (paired t test on individual data; after 0 INC $p < 0.05$; after 1, 2, and 3 INC $p < 0.001$) and CO2 versus CO3 (after 0 INC $p < 0.01$; after 1 INC $p < 0.05$; after 2 and 3 INC $p < 0.001$), but no significant differences for CO3 versus CO4 in both animals. Both animals show an effect of search length on the difference in RT from search to repetition [CO1-CO2; ANOVA; monkey M: $F(3, 108) = 13.241$, $p < 0.00001$; monkey P: $F(3, 108) = 10.693$, $p < 0.00001$].

Vertical lines indicate average error of the mean.

et al., 2004; Walton et al., 2004). A recent experiment in monkeys described positive and negative reward prediction error signals in the ACC during learning (Matsumoto et al., 2007).

However, experiments in humans have revealed, using various cognitive tasks, tonic and phasic ACC activations that are difficult to reconcile with a pure reinforcement learning account of ACC function. For instance, ACC is activated in dual tasks and self-selection of actions (Kondo et al., 2004; Walton et al., 2004). Several authors have proposed that ACC might form part of a dedicated network involve in task maintenance or task control that is in the ensemble of processes involved in regulating task performance (Braver et al., 2003; Dosenbach et al., 2006 ; Johnston et al., 2007; Roelofs et al., 2006).

The question thus remains of whether ACC is purely involved in processing negative and/or positive signals, in which circumstances, and if it produces signals for adapting behavioral strategy and/or for evaluating actions through reinforcement learning mechanisms.

Here we assess the characteristics of ACC feedback-related activity during a trial-and-error protocol that involved different types of behavioral adaptations. The task alternates exploration and exploitation periods, with sharp transitions between the two. We show that ACC activity reports and discriminates different types of feedbacks, negative and positive, that relate to diverse adaptations. A shift of activity from positive feedback in exploration to trial initiations in exploitation suggests that these activities participate in action valuation.

RESULTS

We studied ACC unit activities and local field potential (LFP) oscillations during a task that alternates exploration (trial-and-error) and exploitation (repetition) periods and that manipulates outcome expectation and outcome valence. Two monkeys had to search by trial and error which of four simultaneously presented targets was associated with a reward (Figure 1A). In each trial, the animal had to choose a target by fixating and then touching it. Targets switched off 600 ms after the touch. A reward (positive feedback) was delivered if the correct target was chosen. No reward was given in case of an incorrect choice (negative feedback). Each block of trials (or problem) contained a search period (exploration) during which the animal was

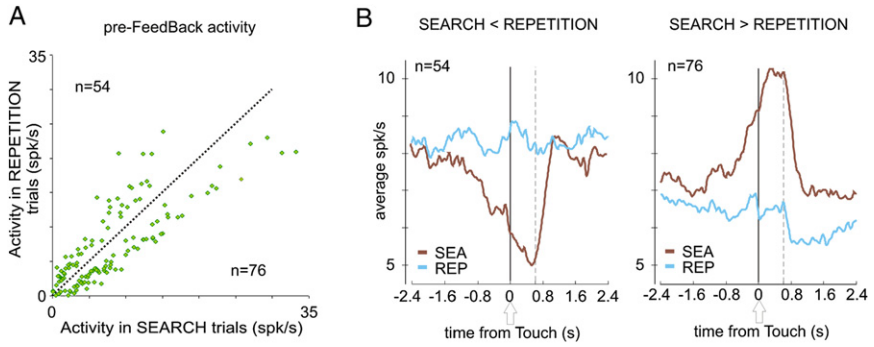

Figure 2. Pre-Feedback Population Activity

Population data for 130 neurons showing statistically significant differences in pre-feedback activity (0 ms to 400 ms after the touch; ANOVA $p < 0.01$) between SEARCH (INC and CO1) and REPETITION (COR) trials.

(A) The average epoch activity in the two periods plotted for each neuron shows two populations with higher or lower activity in search than in repetition periods.

(B) The population histograms for the two populations show the average search versus repetition effect. The dashed gray line indicates feedback onset.

searching for the rewarded target and, after its discovery, a repetition period (exploitation) during which the correct response was repeated at least three times (Figure 1B). A visual signal (signal to change [SC]) at the end of the repetition period indicated the beginning of a new problem. The animal started a trial by touching a target (this target is subsequently named "lever"). The touch induced the onset of a fixation point that marked the initiation of the trial. Subjects were required to fixate the fixation point until target onset and fixate the target once selected by eye (see Experimental Procedures). Any break in fixation requirements resulted in trial cessation (break fixation error—negative feedback). During recordings, monkeys performed optimal searches, i.e., did not repeat incorrect trials (INC), and optimal repetitions. Changes in reaction times (RT) between search and repetition revealed a behavioral shift after the first reward (CO1) (Figure 1C; see Figure S1 online for detailed analyses) (Procyk and Goldman-Rakic, 2006; Procyk et al., 2000). Interestingly, the behavioral shift effect (difference between CO1 and CO2 trials) increased with the length of search periods (see figure legend). This originated from reduced RT in successive search trials for monkey P and from increased RT in CO2 for monkey M (ANOVA, $p < 0.05$). Analyses of break fixations in repetition revealed that monkeys evaluated break of fixations as such and not as selection errors and suggested adaptations specific to these execution errors (see notes on performance in Experimental Procedures). Overall, behavioral data show that monkeys efficiently used each type of feedback.

ACC Feedback-Related Unit Activity

Neural activity was recorded in the dorsal bank of the cingulate sulcus within or anterior to the rostral cingulate motor area (Figure S2). Analyses of 546 ACC unit activities (349 and 197 in monkeys M and P) recorded for sufficient numbers of trials revealed a majority (n = 280; 51%) of neuronal activity with significant activation following feedback onset. Feedbacks were preceded by different levels of activity in search and repetition. 130 neurons out of 546 showed significant differences between search (INC and CO1) versus repetition (COR) trials (Figure 2). 64% of these cells also showed feedback-related activity. Individual examples are shown in Supplemental Data. 58% of cells had higher activity in search trials. Pre-feedback changes occurred at different times before or after the touch on targets (Figure S3). Although some individual cells had increased event-related activity in repetition, on average the cell population showing higher activity in repetition than search revealed a global reduction of neuronal activity during search in anticipation of feedbacks (Figure 2, center).

Computing the grand average over the 280 feedback-related activities clearly showed that both negative and positive feedbacks in search, but not positive feedback in repetition, induced increased ACC activity (Figure 3A). Note that the average activity is sustained after errors, but not after the first reward (CO1). We analyzed individual feedback-related data by separating four types of trials—incorrect (INC), first reward (CO1), reward in repetition (COR), and break of fixation (BKF)—by evaluating feedback preferences at the time of feedback and by grouping cells in different populations. The grouping procedure was performed on 234 cells (The analysis had two steps, one to detect significant activation for each feedback and one to group cells by their preference for feedbacks; this procedure was also compared to automatic hierarchical clustering Figure S6 and notes. Both methods gave similar results; the following proportions are issued from the first method—see Experimental Procedures, Figures S4 and S5, and Supplementary Notes for details and discussions.) The different populations were recorded in overlapping regions. The assessment of feedback preferences revealed a majority of ACC activities related to INC, CO1, and/or BKF feedbacks, indicating a bias toward processing feedbacks of search periods (exploration) or BKF feedbacks (Figure 3B). Major groups of ACC neurons were as follows. (1) Those that discriminated all feedbacks in search from positive feedback in repetition, regardless of whether outcomes in search were positive or negative. INC/CO1 neurons (7% in the two monkeys taken together) had increased activity for all INC feedbacks and the first reward (feedbacks in search) (Figure 3C, left; Figure 4A). Note that on average these activities started before feedback onset, which might relate to the anticipatory components of monitoring. (2) Conversely, COR neurons (6%) reacted only to rewards in the repetition period. These two groups probably concern a subpopulation of the activities reported to vary between search and repetition (Procyk et al., 2000). (3) Classically described error-related activity following incorrect choices (INC) represented 22% of the feedback-related population (Figure 3C, middle), i.e., 9.5% of total recorded neurons, as previously reported (Amiez et al., 2005; Ito et al., 2003). Of all feedback-related activities, 8% increased after BKF and INC and were classified as INC/BKF (Figure 3B). (4) 34% of feedback-related neurons were said to be BKF related (Figures 3B and 4B). (5) Most remarkably, a critical population reacted significantly only to the first reward (CO1) (17%). The same reward given in repetition (COR) elicited

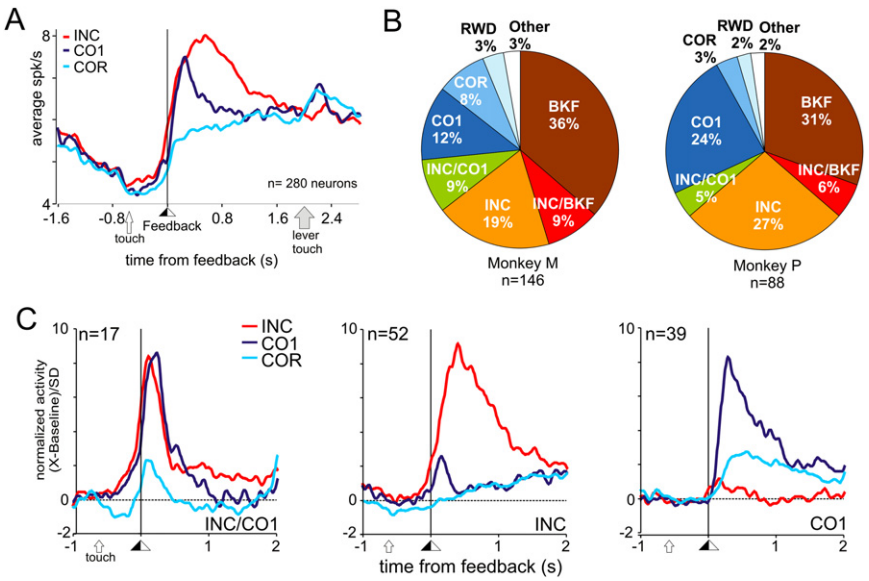

Figure 3. Populations of Feedback-Related Activity

(A) The grand average unit activity for 280 feedback-related units shows average activation for CO1 and INC. Time of target touch and time of lever touch are indicated on the x axes.

(B) Main types of feedback-related activity for each monkey. Note that most activities are related to uncertain feedbacks (break of fixation [BKF], INC, and CO1). Very few were specific to reward per se (RWD) or reward in repetition (COR). "Other" includes responses to CO1 and BKF, COR and BKF…(see Figure S5). The repartitions of types are not different in the two monkeys (χ^2 = 7.6164, df = 7, p value = 0.3676).

(C) Population histograms for three main categories aligned on INC, CO1, and COR.

no change in activity for these neurons (Figure 3C, right; Figures 4C and 7). Thus, this activity appears with the discovery of the correct target and the shift between search and repetition. The activity related to CO1 could be a reflection of positive prediction errors as describe for mesencephalic neurons. In such a case, CO1 responses should vary according to the probability to be rewarded, i.e., to the rank of CO1 trials in search periods. In other words, reward discovery in the first trial of search periods and reward discovery at the end of search periods (after targets have been eliminated by trial and error) should elicit different neural responses. A large majority (30 out of 31 neurons tested; 97%) of CO1 cells showed no effect of rank of the CO1 trial in search periods and thus did not vary according to prediction errors (ANOVA and Kruskall-Wallis, $p < 0.05$) (Figures 5A and 5B). However, 4 out of 17 INC/CO1 neurons (23%) showed an effect. 3 of these 4 neurons showed a progressive decrease of activity with the rank of CO1 trials (see an example in Figure 5C).

Transfer of Feedback-Related Activity to Trial Initiation

We further observed that 13.5% of feedback-related neurons (n = 38/280) significantly increased their activity *after* the lever touch following a rewarded choice (>3 SD from baseline before lever touch—see Experimental Procedures). The effect was also visible on grand average histograms (Figure 3A). We named this activity lever related, but note that the lever touch is immediately followed by fixation point onset that indicates trial initiation. Figure 5 shows three example units. Two of these units showed increased activity after the first correct feedback (panels A and B). This activity disappeared for COR trials, while activity at the initiation of the following trial increased. Interestingly, when the

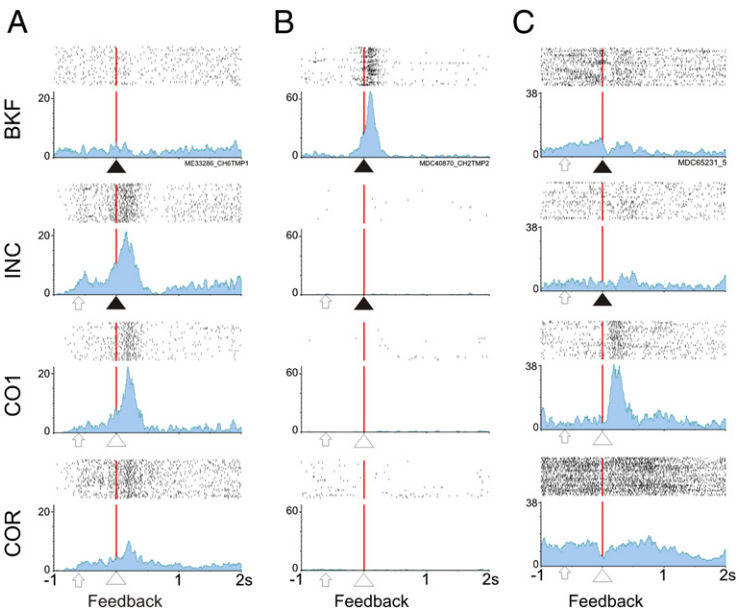

Figure 4. Examples of Feedback-Related Single Units

Three examples coding for INC/CO1 (A), BKF (B), and CO1 (C). Histograms are aligned on feedback with rasters on top. Activity is shown for four different feedbacks indicated on the left of the figure. Arrowheads indicate negative (black) or positive (reward, white) feedbacks.

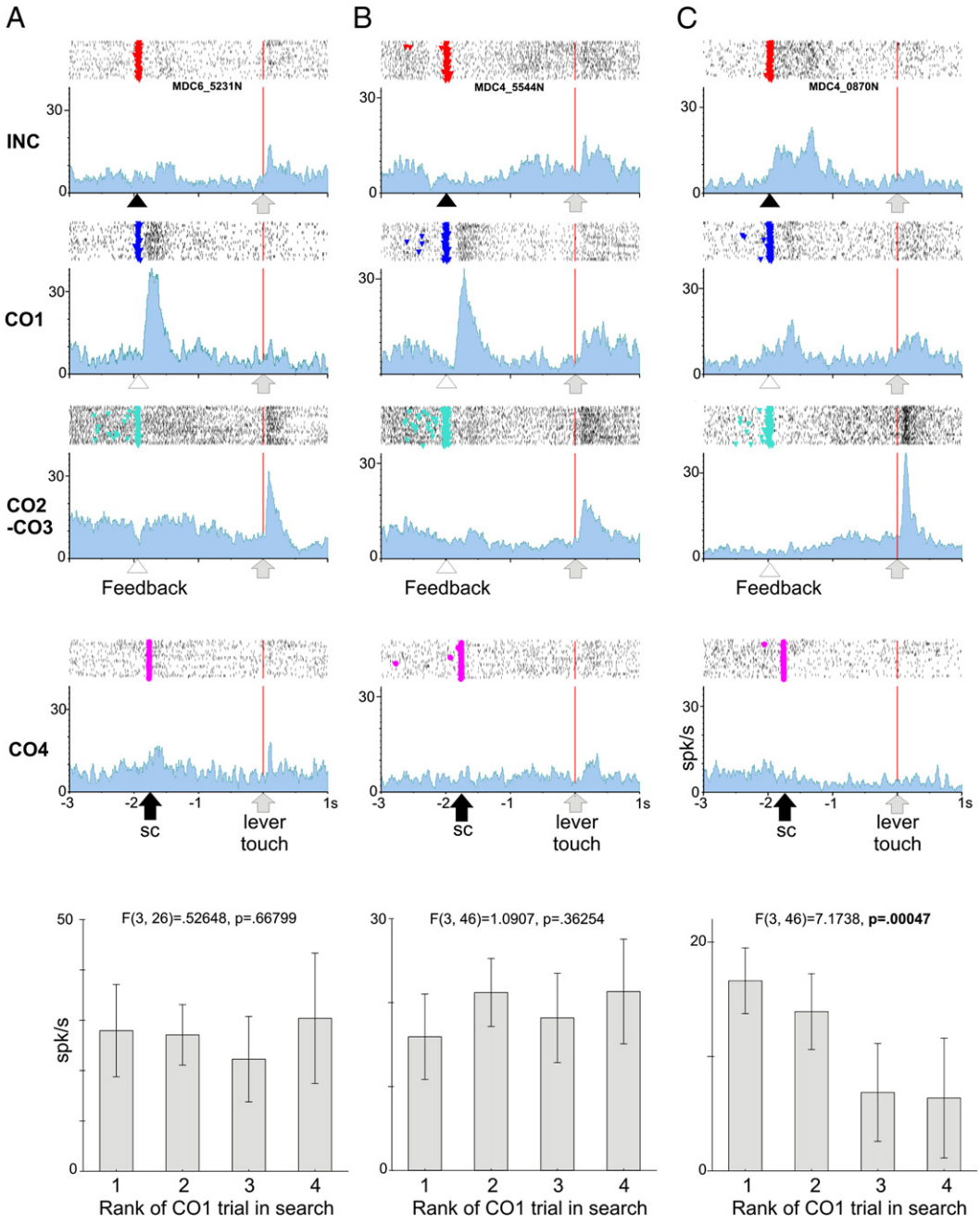

Figure 5. Three Single-Unit Examples of Reinforcement-Related Signal Transfer

The examples illustrate the transfer of feedback-related activity to lever-onset from the end of search (CO1) to repetition (CO2-CO3). Each column represents data for one single unit. Activity is aligned on lever touch. The three first rows present peristimulus histograms for INC, CO1, 2nd, and 3rd correct in repetition (CO2, CO3). In the fourth row, the activity is aligned on the lever that follows the signal to change response (SC; upward black arrow), i.e., the first lever touch of a problem. Note that feedback-related activity disappears in repetition and that lever-related activity increases. However, the lever-related activity disappears as soon as the monkey enters a new search following the SC. In the last raw, histograms of average activity measured for the four ranks of CO1 trials in search periods. Rank 1 corresponds to discovery in the first trial in search. Ranks 2, 3, and 4 correspond to discoveries after 1, 2, and 3 incorrect trials, respectively. Statistical values are the results of one-way ANOVAs. Error bars indicate 95% confidence intervals.

animals were informed of the termination of the repetition period (signal to change [SC]) and thus that a new correct target was to be found, the activity after the lever touch disappeared. These neurons (n = 38) had higher lever-related responses after CO1 or COR rewards than after errors. They were selected without initial consideration for feedback-related activity. Remarkably, the average activity of these neurons revealed minor increases for incorrect feedbacks, a maximum peak for CO1 feedbacks, and no

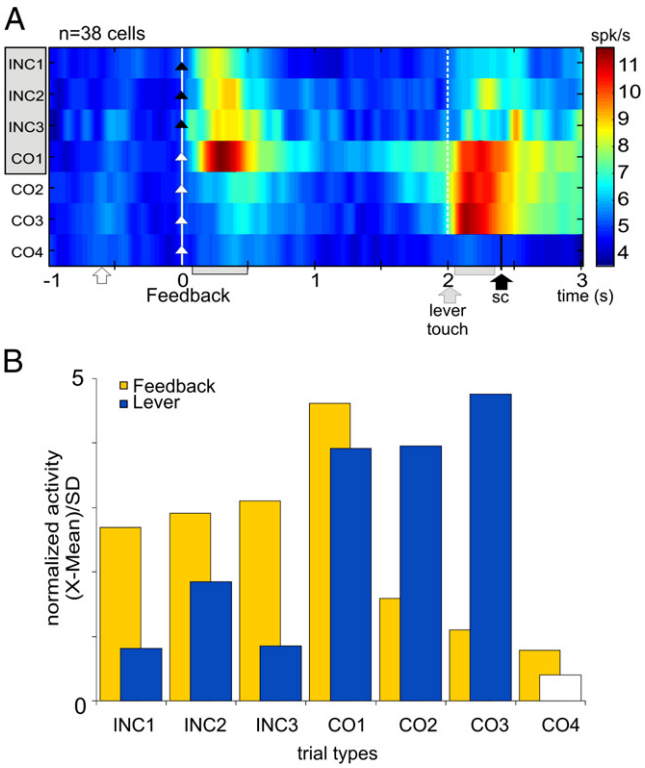

Figure 6. Signal Transfer in Population Activity

(A) Population activity for the different types of trials in an average problem for neurons showing significant activity at the lever touch (color indicates mean firing rate). Trial types are presented chronologically from top to bottom. The feedback-related response observed in search (search = gray area behind trial type names) is then transferred toward the lever touch after the first reward, after which no more feedback-related response is detectable. The signal to change in CO4 is indicated: black arrow, sc; lever touch in CO4 appears after 1.4 s and is not visible in the figure.

(B) Averaged standardized unit activity measured at the feedback and lever touch epochs (gray boxes on the abscissa in [A]). Note the shift after the first reward, CO1.

feedback-related activity during repetition (Figures 6A and 6B). 24/38 activities (18 in monkey M, 6 in monkey P) were significantly increased after INC and/or CO1 feedbacks. In summary, the population was reactive to the first reward, after which it reacted to the initiation of the next surely rewarded trial and not to the reward itself. Activity at the lever disappeared again when the change in reward contingencies were signaled.

High Gamma Oscillations Related to Feedbacks

Feedback-related ACC activity was also found in local field potentials. For 50% of recording sites (n = 55/110; monkey M, 34; monkey P, 21), the LFPs revealed significant increased power in high gamma bands after feedbacks (see example of time-frequency diagrams in Figure 7). These oscillations from 60 to 120 Hz, and often higher, are very similar to those described using intracranial or epipial electrodes in humans (Edwards et al., 2005; Mainy et al., 2006). We focused on variations related to feedback onset for INC, CO1, and COR, although changes were also found after BKF. The increased gamma was significantly higher during search—INC and/or CO1 feedbacks—than repetition (COR) in about 73% of sites (n = 40/55; monkey M, 26; monkey P, 14). This includes 20% (n = 11; monkey M, 5; monkey P, 6) of sites for which gamma increase was present only for CO1 feedbacks, and 36% (n = 20; monkey M, 12; monkey P, 8) for INC. An increase in gamma bands was observed during the repetition period or during both search and repetition periods in 27% of sites (n = 15; monkey M, 8; monkey P, 7). Regarding recording locations, no clustering was observed among the different feedback-related activity. Interestingly, feedback-related high gamma oscillations and simultaneously recorded unit activity were not always sharing the exact same functional properties (Figure 7). Yet, when feedback-related unit activity was recorded, feedback-related gamma was simultaneously observed in 58% of cases. When non-feedback-related activity was recorded, feedback-related gamma was found in only 32% of cases (one-sided proportions comparison test: $p < 0.007$). In summary, high gamma oscillations revealed a high incidence of activity related to the processing of both negative *and* positive feedbacks, when those feedbacks are relevant for behavioral adaptation. Remarkably, the transfer of neural responses from feedbacks to lever touch also appeared at the level of gamma oscillations (Figures 8A and 8B). Whereas we observed feedback-related gamma increases during search periods (INC and CO1), the gamma power increased and peaked in relation to lever touch after the first reward and during repetition. As for unit activity, ACC gamma activity related to outcome processing transferred to the initiation of trials once the positive value of a choice was established.

DISCUSSION

Recordings of unit activities and LFPs lead to four main findings on the role of ACC in behavioral adaptation.

First, the same ACC region processes and discriminates different negative and positive events. Remarkably, and in addition to previous reports of ACC responses to task-related rewards or free rewards (Amiez et al., 2005; Ito et al., 2003; Niki and Watanabe, 1979), we show a high incidence of specific responses to the first reward in a problem. Thus, ACC responses to positive feedback depend on its context, which, in the PS task, varies between uncertain (CO1) and highly expected (COR). It is possible that the bias of the literature toward a role for ACC in negative feedback processing derives in part from experimental biases, i.e., nonequivalence of the relevance of positive and negative outcomes. It is remarkable that most CO1 neuronal activities did not vary with prediction errors, although variations according to prediction errors are found for other types of activity. The CO1 signal thus appears as categorical information and not as a scalar measure of discrepancy between outcome and expectation as observed with other protocols (Matsumoto et al., 2007). One possibility then is that CO1 activity signals the end of the exploratory period and the shift toward a repetition mode of behavior.

Second, ACC unit activity reflects important feedback-related mechanisms that are also observable in LFP gamma oscillations. High gamma oscillations are a good marker for studying structure-function specificities in humans, and they seem to correlate strongly with functional magnetic resonance imaging bold signal

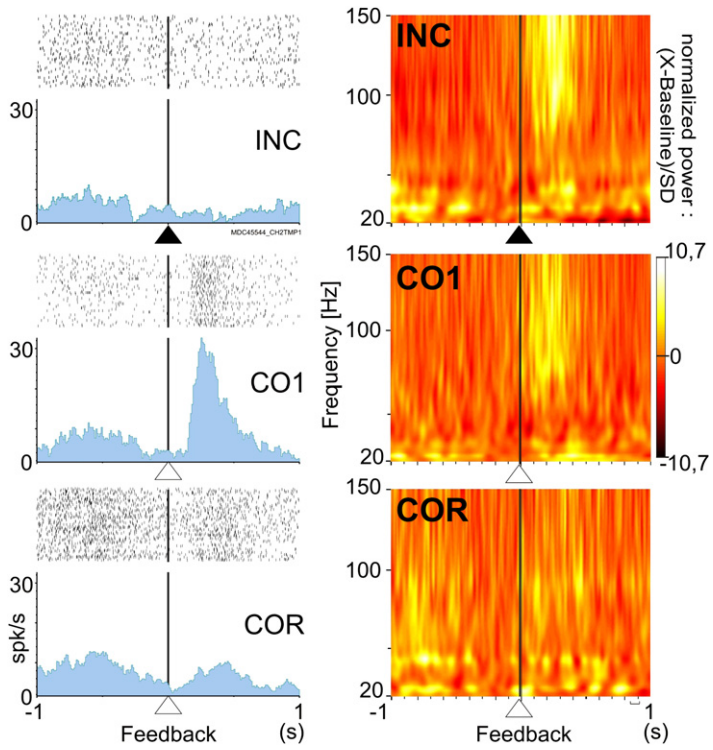

Figure 7. Simultaneous Feedback-Related LFPs and Single-Unit Recordings

Average firing rate of a single neuron (left) and time-frequency composition of LFPs (right) recorded simultaneously by the same electrode. Time-frequency diagrams show the normalized spectral content of LFP from −1 s to +1 s around feedback onset, expressed in number of sigma of the baseline (−300 to −100 ms). The unit is activated for the first reward only, whereas the LFPs show an increased power in the high gamma frequency band (60–120 Hz) for both incorrect and first correct trials.

(Edwards et al., 2005; Mainy et al., 2006; Niessing et al., 2005). LFP data can thus explain recent reports of increased bold signal in human ACC for both correct and incorrect performance in a learning paradigm (Walton et al., 2004) and can also clarify the observed absence of feedback-related ACC activation when measures for negative and positive feedbacks are directly contrasted (van Veen et al., 2004). We predict, from LFP recordings, that using the PS task in humans will reveal feedback-related increased bold signals in the same ACC area for both incorrect and first correct trials. However, in light of the present data we conclude that ACC encodes and discriminates both positive and negative outcomes.

Studying the relationship between unit activity and LFP is crucial. For instance, one key aspect in the debates over the role of ACC has been the discrepancy between human and monkey experiments, the later giving little, if any, support to the conflict-monitoring hypothesis (Botvinick et al., 2004; Ito et al., 2003; Nakamura et al., 2005; Rushworth et al., 2004). One can argue that discrepancies come from the use of different recording techniques that give access to distinct aspects of ACC, one related to its output (spike), the other to its input (Bold). LFPs reflect mainly synaptic potentials (unless local multiple spikes are highly correlated)—that is, the input of a given cortical area as well as its local intracortical processing—and seem to be a better predictor of Bold increases than single spike firing (Buzsaki, 2006; Logothetis, 2003; Logothetis and Wandell, 2004). Indeed, it has been proposed that although in many cases spikes, LFPs, and Bold can correlate with each other, in some possible instances Bold/LFPs and spike firing can vary independently (in case of modulation, habituation…) (Logothetis, 2003). Therefore, and although clear logical counter-arguments have already been developed (Nakamura et al., 2005), the LFPs/unit comparison needs to be discussed in the context of debates over ACC function and for comparisons between human and nonhuman experiments.

According to the conflict-monitoring theory, conflict would arise when an unexpected event conflicts with an expected one, and hereby would be detected at the level of performance feedback (van Veen et al., 2004). Our data show that overall, at unit and LFP levels, the ACC is responding for positive and negative feedbacks when those are relevant for adaptation. Overall, the majority of recordings show signals (spike or LFPs) that discriminate between positive and negative feedbacks. This suggests a role for ACC in processing the valence of relevant feedbacks. It is hard to fit conflict detection and specific feedback-related signals. Moreover, conflict-detection theory hardly explains the shifts in activity from feedback to trial initiation. We thus propose that explaining ACC feedback-related activity in terms of reinforcement learning mechanisms is more parsimonious.

Third, the transfer of reinforcement-related information from feedback to trial initiation reveals that ACC participates in fast learning mechanisms. It adjoins previous reports of ACC activity modulated by reward prediction (Amiez et al., 2005; Matsumoto et al., 2007). The transfer is comparable to the reinforcement learning effect observed for dopaminergic neurons whose activity related to unexpected rewards transfers to conditioned stimuli predicting future rewards (Schultz, 2000). The major differences in our protocol are that ACC activity transfer relates to trial initiation and that it occurs within a few hundred milliseconds. Indeed, fast valuation of action is an important requirement for organisms to correctly exploit resource discovery or learn from dangerous or painful incidents. Satoh and colleagues found that mesencephalic activity can relate to trial initiation and be modulated across trial and error similarly to our lever-related activity (Satoh et al., 2003). This supports close relationships between mesencephalic and ACC neurons. However, shifts in neural activity toward conditioned stimuli have also been observed in the locus coeruleus that might have important roles in fast adaptation (Bouret and Sara, 2005). In any case, we show here the first evidence of a link between ACC feedback-related neural activity and subsequent neurophysiological change in behavioral valuation. It suggests that in our protocol the neural response to the uncertain reward concerns its behavioral relevance.

ACC lesions impair the integration of reward history, and ACC unit activity reflects reward-action associations, average expected values, and negative reward prediction errors (Amiez

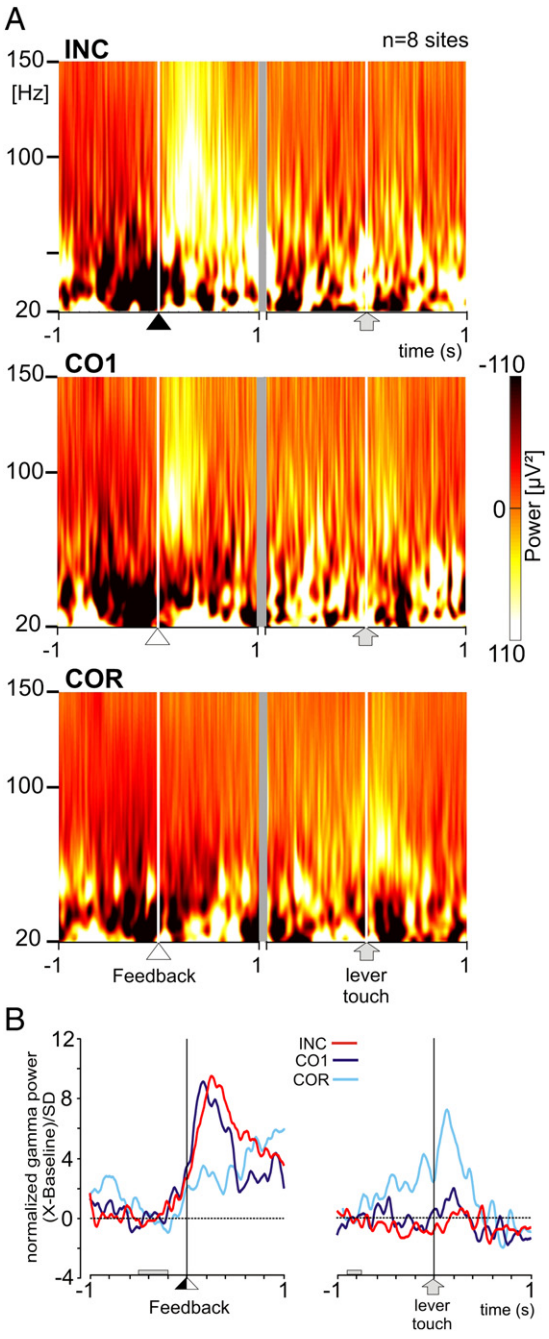

Figure 8. Average LFPs and Signal Transfer

(A) Time-frequency diagrams showing the averaged signals from eight sites with INC/CO1 gamma activity. Frequencies are shown from 20 to 150 Hz for clarity. Note the disappearance of feedback-related activity for COR and lever-related gamma activity appearing after CO1 and maximum after COR.

(B) Averaged gamma signal between 60 and 120 Hz for the same eight sites. Note the change in feedback-related activity and the appearance of lever-related activity from search to repetition trials. Profiles are expressed in standard deviations compared to respective baselines (gray boxed on abscissa). Profiles are produced separately for feedback and lever (see Experimental Procedures).

et al., 2005, 2006; Kennerley et al., 2006; Matsumoto et al., 2003). In addition, activity of ACC neurons reflects reduction in rewards and action shifting (Shima and Tanji, 1998). In this context, the present data clearly support a role for ACC feedback-related activity in updating action values (Kennerley et al., 2006; Rushworth et al., 2004). Recordings of mesencephalic dopaminergic neurons during the trial-and-error task revealed reward prediction error-related signals that could influence ACC feedback-related activity (Matsumoto et al., 2007; Satoh et al., 2003). The present findings suggest, however, that ACC computes various types of reward-related signals that might be more categorical than prediction errors. Models of the relationship between ACC and phasic dopaminergic signals (Holroyd and Coles, 2002) must thus take into account much more than detection of behavioral errors. Our finding on the varieties of feedbacks encoded in ACC, about learning effects, and other reports of ubiquitous modulations related to reward prediction errors suggest that, beyond direct dopamine-ACC relations, feedback detection and evaluation emerge from interactions within larger networks (Haruno and Kawato, 2006; Ullsperger and von Cramon, 2006). In particular, studying the top-down influences of ACC and the differential impacts of dopamine on different cortical targets (pyramidal and inhibitory neurons) and on ventral striatum will give essential information (Seamans and Yang, 2004; Tanaka, 2006). Such research might explain the relationships between the bidirectional coding in dopaminergic neurons (activation and inhibition for positive and negative errors, respectively) and the segregated, feedback-specific, unidirectional coding in ACC. Overall, the present data suggest important interactions but rule out the idea of ACC feedback-related signals as being simple reflectors of dopaminergic inputs. Simultaneous cortical and mesencephalic recordings will be needed to further study the interactions.

Finally, adaptation of values goes with adaptation of behavior. Our data show that ACC discriminates between different types of feedback, allowing appropriate behavioral adaptations (e.g., stay after CO1, shift in response after INC, increase control on fixation after a BKF—see Supplemental Notes about unspecific activity for CO1 and INC). Activity shows different dynamics after feedback, depending on the feedback valence. ACC also codes for visual signals, indicating a need to engage explorations (Amiez et al., 2005). Thus, the output of ACC is unlikely to be devoted to one specific type of adaptation (for instance, shifting response, or pure valuation in terms of reward). Recent data show that another area of the medial frontal cortex, the pre-SMA, has an important role in correctly shifting from automatic to controlled behaviors (Isoda and Hikosaka, 2007). The respective role and collaboration between the two areas are a major issue (Akkal et al., 2002; Rushworth et al., 2004). It is likely that the specificity of ACC relates in part to its position within the reward system and to the use of outcome information for action value adjustments and behavioral regulation or global changes in goal-directed policy.

Computational modeling suggests that ACC participates in the regulation of cognitive control notably exercised by lateral prefrontal cortex and that it receives and/or controls modulatory signals that mark transitions between routine and nonroutine behaviors (Aston-Jones and Cohen, 2005; Cohen et al., 2004;

Dehaene et al., 1998). As such, it is a key component of a so-called multiple-demand network or global workspace (Dehaene et al., 1998; Duncan, 2006). Recent works in humans suggest that the medial frontal cortex, including part of ACC, with the anterior insula/frontal operculum participate in maintenance of task sets or task control (Braver et al., 2003; Dosenbach et al., 2006). We previously showed that ACC expresses different processing states between search (exploration) and repetition (exploitation) (Procyk et al., 2000), and the present data reveal that these differences are covering different events of trials (delay, target onset, movement, and feedback expectation). ACC then produces specific signals in anticipation or at the onset of events, especially when task control is high and when actions are not yet valued, that is, during exploration. In similar conditions, lateral prefrontal cortex activity is modulated with analogous dynamics between search and repetition (Procyk and Goldman-Rakic, 2006). Our data are compatible with a role for ACC within a regulatory loop involving both noradrenergic and dopamine-based reinforcement learning mechanisms. ACC reward-discovery signal might serve as a trigger to shift from an exploratory to an exploitative mode, and conversely for error-related signals (Aston-Jones and Cohen, 2005; Dehaene et al., 1998). This interpretation can also be formulated in light of the recent demonstration that ACC feedback-related activity correlates with the volatility or uncertainty of the reward environment and with the subjects' learning rates (Behrens et al., 2007). Specific ACC states and dedicated processing of outcomes in uncertain environments would convey the tuning of metaparameters such as learning/adaptation rate (Behrens et al., 2007; Doya, 2002). Although we do not have direct evidence of the consequences of ACC feedback-related activity on lateral prefrontal cortex or on other structures including neuromodulatory systems, ACC itself is likely to be involved in anticipating and characterizing multiple relevant events to trigger appropriate reactions that in fine aim a behavioral valuation and possibly global behavioral regulation. Thus, the function we attribute to ACC activations is clearly not only to evaluate feedbacks but is also to participate in monitoring the different steps of the task at hand to optimize action adaptation and valuation. A dysfunction of these mechanisms represents the core feature of cognitive alterations observed in addiction and mental illness.

Conclusion

The ACC produces signals that discriminate between various behaviorally relevant positive and negative feedbacks, suggesting a role in triggering appropriate adaptations. Specific signals might be at the origin of transitions between different behavioral policies (exploration and exploitations). Transfer of reinforcement signals from time of reinforcement to trial initiation reveals mechanisms that pertained to reinforcement learning and possibly reflects direct interactions between ACC information processing and monoaminergic functions.

Our data reinforce the proposal that ACC is important for establishing action valuations. But they also emphasize a combined role in monitoring events/actions for behavioral regulation when task control is high, underlining the intimate link between control and action valuation.

EXPERIMENTAL PROCEDURES

Housing, surgical, electrophysiological, and histological procedures were carried out according to the European Community Council Directive (1986) (Ministère de l'Agriculture et de la Forêt, Commission nationale de l'expérimentation animale) and Direction Départementale des Services Vétérinaires (Lyon, France). Each animal was seated in a primate chair within arm's reach of a tangent touch-screen (Microtouch System) coupled to a TV monitor. An arm-projection window was opened in the front panel of the chair, allowing the monkey to touch the screen with one hand. A computer recorded the position and accuracy of each touch. It also controlled the presentation via the monitor of visual stimuli (color shapes), which served as light targets (CORTEX software, NIMH Laboratory of Neuropsychology, Bethesda, MD). Eye movements were monitored using an infrared system (Iscan Inc., USA). Four target items (disks of 5 mm in diameter) were used: upper left (UL), upper right (UR), lower right (LR), lower left (LL) (Figure 1A). A central white disk served as fixation point (FP). The lever was disposed just below the FP.

Task

Two male rhesus monkeys were trained in the problem solving task (PS) (Figures 1A and 1B). Monkeys had to find by trial and error which target, presented in a set of four, was rewarded. Each trial started by the onset of a starting target named "lever." The animal had to start a trial by touching the lever and holding his touch. The FP appeared, and the animal had to fixate it with his gaze. A delay period (2 s) followed, and ended by the simultaneous onset of the four targets. At the FP offset, the animal made a saccade toward a target, fixated it (0.5 s), and then touched it following the GO signal. All targets switched off at the touch, and a 0.6 s delay followed before the feedback was given. A reward (fruit juice) was delivered for choosing the correct target (positive feedback; white arrowhead). If a choice in one trial was incorrect (no reward, negative feedback; black arrowhead), the monkey could select another target in the following trial and so on until the solution was discovered (search period). Each touch was followed by an interval of 2 s.

The animal had to search for the correct target by trial and error. After discovery, the animal was allowed to repeat the response. In 90% of cases, after the third repetition, a red flashing signal (the four targets in red) indicated the start of a new problem, i.e., a search for a new correct target. In 10% of cases, the repetition lasted for 7 or 11 trials. A problem was composed of two periods: a "search" period that included all incorrect trials up to the first correct touch, and a "repetition" period wherein the animal was required to repeat the correct touch (Figure 1B). The different types of trials are indicated (−, no reward; +, reward; INC, incorrect; CO, correct).

Notes on Performance

The two monkeys worked on the PS task with optimal performances: average number of trials in search: 2.4 ± 0.15 trials in monkey M, 2.65 ± 0.23 in monkey P (optimal 2 or 2.5; see Procyk and Goldman-Rakic, 2006), in repetition: 3.14 ± 0.7 trials in monkey M, 3.4 ± 0.55 in monkey P. A break of fixation (BKF) in repetition (24% and 20% of repetition trials for monkeys M and P) did not induce particular change in response selection in the following trial, preserving good overall performance in repetition. This suggests that monkeys detected and evaluated BKF as such and not as a potential error in selecting a target. This is a behavioral parallel of the discrimination seen at the neural level. Based on previous publications (Amiez et al., 2005; Procyk and Goldman-Rakic, 2006), we hypothesized that reactions to BKF would vary depending on when the BKF occurred within trials (early when monkeys did not yet overtly selected a target, or later when monkeys were selecting the target by gaze). Effectively, analyses showed that, in repetition, monkeys were less likely to repeat a BKF or to make an INC trial after a BKF done at target selection time than after a BKF early in a trial (χ^2 on numbers of INC, BKF, COR, and Other after early (before 900 ms after trial start)and late (between 1800 and 2500 ms) BKF: Chi = 79.42, df = 4, p = 2.30911e−16, and Chi = 84.18, df = 4, p = 2.265524e−17 for monkeys M and P, respectively).

RT were significantly different for CO1 and CO2 (i.e., at the shift between search and repetition) for the two monkeys (t test, df = 27, monkey M: t = −5.03, p < 0.0001; P: t = −4.8, p < 0.0001; calculated for 28 days).

Recordings

Monkeys were implanted with a head-restraining device, and an atlas-guided craniotomy was done to expose an aperture over the prefrontal cortex. A recording chamber was implanted with its center placed at stereotaxic anterior level +31. Neuronal activity was recorded using epoxy-coated tungsten electrodes (1–4 MOhm at 1 kHz; FHC Inc, USA). One to four microelectrodes were placed in stainless-steel guide tubes and independently advanced into the cortex through a set of micromotors (Alpha-Omega Engineering, Israel). Neuronal activity was sampled at 13 kHz resolution and LFP at 900 Hz. Recordings were referenced on the guide tubes in contact with the dura and containing the microelectrodes. Recording sites (see Figure S2) covered an area extending over about 6 mm (anterior to posterior), in the dorsal bank of the anterior cingulate sulcus, at stereotaxic antero-posterior levels superior to A+30, and at depths superior to 4.5 mm from cortical surface. Locations were confirmed by anatomical MRI and histology. This corresponds to a region recorded in previous reports and in which error-related activity has been observed (Amiez et al., 2006; Ito et al., 2003; Procyk et al., 2000). This part of the anterior cingulate cortex lies at the same anterior level as the SEF and includes part of and goes anterior to the rostral cingulate motor area (CMAr), as evaluated from previous publications (Ito et al., 2003; Shima et al., 1991).

Unit Activity

Single activity was identified using online spike sorting (MSD, AlphaOmega). The activity of single neurons was compared with respect to different events and outcomes resulting from different conditions by using averaged peristimulus histograms (PSTH) and trial-by-trial spike counts (NeuroExplorer, Nex Technology, USA, and MatLab—The MathWorks Inc.—home-made scripts). PSTHs had a binning of 0.01 s and were boxcar averaged. Neural activity was considered to be significantly different between conditions if it exceeded 5 standard deviations of the mean difference between trial types taken during the window −600/−200 ms preceding event alignment time and remained above this threshold for more than six 0.01 s bins. MatLab, R v2.5.0 (R Foundation for Statistical Computing) and Statistica (StatSoft Inc.) were used for analyses and graphics.

Neuronal activity was studied at the level of target touch and tested for having significant difference between trials in search periods (INC and CO1 combined) and trials in repetition periods (COR). Activity in each trial was measured in a post-touch epoch (+0 ms to +400 ms from touch time; the window was limited to +400 to avoid changes in activity related to feedbacks). A one-way ANOVA, Search versus Repetition, at $p < 0.01$ was used to select activity that discriminated between the two periods.

Neuronal unit activity was evaluated for being feedback related or lever related. For feedback-related measurements, the average activity aligned on the different feedback types (INC, CO1, COR, and BKF) were standardized to the mean and standard deviation of the baseline activity taken from −600 to −200 ms before feedback onset (Figure S4). Resulting data were expressed in number of sigma (SD) of the baseline. Data were analyzed with a first test (TEST1): an activity was said to be feedback related when, during the window +60 +800 it passed 5 SD (of the baseline) for more than six bins (6 × 0.01 ms). For Lever-related measurements, the average activity aligned on the first lever touch present after the different feedback types (INC, CO1, and COR) were standardized to the mean and standard deviation of the baseline activity taken from −300 to −20 ms before lever touch. Resulting data were expressed in number of sigma (SD) of this baseline. An activity was said to be lever related when the average activity measured in the window +50 +350 ms after lever touch passed 3 SD of the baseline.

Grouping, Clustering

Feedback-related activities were grouped in different types with the following strategy: if one activity was showed to be significant at the 5 SD test (TEST1) for only one feedback, then it was defined as being specific for this feedback; if one activity was significant at TEST1 for more than one feedback, we ran a post hoc analysis on differences between feedbacks (TEST2): the standardized average epoch activity was calculated for all feedbacks in the time window showing significant bins. If the standardized epoch averages passed 5 SD for several feedbacks and the difference between standardized averages was inferior to 3 SD, then the activity was defined as being related to the different feedbacks: for instance, a standardized epoch activity measured at 12.667, −0.6868, −3.5318, and 56.362 for INC, CO1, COR, and BKF, respectively, would be BKF, whereas a standardized epoch activity measured at 6.6393, 0.13089, −2.0113, and 9.5247 would be defined as being INC/BKF. Note that in the first case the activity is also significant for INC feedback; yet it is five times higher in standardized activity for BKF. To account for these differences and also for equivalence, we defined TEST2. The choice of a threshold at 3 SD is somewhat arbitrary, although we visually verified the test with PSTH. The raw statistical data for classification of CO1, INC, and INC/CO1 activity illustrates the implementation of the classification for neurons classified in each category after TEST2 (Figures S4 and S5). This classification incorporates specificity of activity and tendencies for activity to be more related to one particular feedback. Although it is not ideal, it gives clear ideas about tendencies within populations. Data obtained directly from TEST1 are presented in Figure S5.

We evaluated the grouping by feedback preference procedure by using an alternative method, hierarchical clustering. The details are given in Supplemental Notes and Figure S6. In short, both methods led to similar grouping of activity and feedback preferences.

Local Field Potentials

We analyzed the local field potentials (LFPs) at sites where unit activities have been recorded. Recordings at single recording sites were analyzed independently and selected for having correct recording quality (regarding electrical noise in particular). Although variations at different frequency bands were observed in time-frequency diagrams, we focused this study on the high gamma band, which showed unique relations to feedback onset.

LFPs were evaluated with the software package for electrophysiological analyses (ELAN-Pack) developed at the Inserm U821 laboratory (previously U280; Lyon, France; http://www.lyon.inserm.fr/821). LFPs were visualized and analyzed in the form of time-frequency graphics. Each set of data represents the activity recorded at one site (one depth for one track) in the ACC. The time-frequency graphics contained the averaged power of each frequency ranging from 2 Hz to 150 Hz in 2 Hz steps across all the trials for the entire data set. Data were analyzed in the time-frequency domain by convolution with complex Gaussian Morlet's wavelets with a ratio f/δ_f of 7. We applied a Blackman window (=100) to eliminate border effects. Trials were aligned on feedback onset (reward or no reward). Data were baseline subtracted. The average signal at all frequencies taken in the last 500 ms of trials was subtracted (common baseline) from the data because it seemed to be the more neutral period in terms of power variations for the different frequency bands.

The statistical evaluations of postfeedback increases in high gamma oscillations were performed as following. Analyses were made on time-frequency data computed for three types of feedbacks: INC, CO1, and COR. The acceptability threshold was defined at $p < 0.001$ for Kruskal and Wilcoxon tests. The analyses followed three systematic steps. (1) We applied for each type of feedback a Wilcoxon test in order to isolate significant changes in the high gamma frequency band (60–120 Hz). The level of gamma frequency power at the time and just after the feedback (from −100 to 400 ms) was compared to the averaged signal included between −300 and −100 ms before the feedback. A profile was generated to determine whether these changes corresponded to increases. (2) We then applied a Kruskal-Wallis test on the files where a statistical increase in gamma band had been found to determine the statistical differences in gamma power between the three events (INC, CO1, COR). To rank the signals related to each type of feedback from the higher to the lower (in order to categorize "gamma signals"), we compared the averaged power measured from −100 to 400 ms and from 60 to 120 Hz. (3) Classification: the files where a significant increase in the gamma band was observed were classified in different categories according to the type of feedback generating the signal and according to the rank of each signal in terms of power. We obtained five main categories of gamma related to:

(1) incorrect trials: 36.4% of the files with a gamma signal (INC; that means that the increase of power in the gamma band occurred only after the incorrect trials or that the level of power in the gamma band was statistically higher for this condition than for the other generating such a signal)
(2) first correct trials (CO1): 20%

(3) correct trials of the repetition period (COR): 21.8%
(4) trials of the search period (INC/CO1; where the power of the gamma signals following these two feedbacks were equal, if there was an increase of power in the gamma band after a COR, it was statistically less powerful): 16.4%
(5) the three types of feedbacks (ALL; in this case, the power of the signals were equal in the three conditions): 5.4%.

LFP Time-Frequency Diagrams and Profiles

The diagrams and profiles presented in Figure 8 were made independently for feedback onset and lever touch. The averaged raw power signal for the frequency band 60–120 Hz are expressed in standard deviations compared to a baseline taken from the averaged signal (from −500 to −200 ms before feedback and from −900 to −750 ms before the lever touch for feedback and lever-related signals, respectively). Because of occasional jitters in the time delay between feedback onset and lever touch and because of occasional breaks of fixation just after the lever touch, we aligned separately on feedback onset and lever touch and removed lever-related activity trials for which a break of fixation occurred just after lever touch.

Grand average unit activity (Figures 2 and 3A) and average gamma profiles (Figure 8) were smoothed with a loess fitting.

Supplemental Data

The Supplemental Data for this article can be found online at http://www.neuron.org/cgi/content/full/57/2/314/DC1/.

ACKNOWLEDGMENTS

We thank J.P. Joseph, H. Kennedy, O. Bertrand, J. Sallet, J. Vezoli, and P.-E. Aguera for help on data analyses, reading, and discussion of the manuscript; and K. Knoblauch for important helps with statistics. E.P. is funded by NRJ foundation, Région Rhône-Alpes, Agence National de la Recherche, Centre National de la Recherche Scientifique, and Institut National de la Santé et de la Recherche Médicale; R.Q. is funded by Facultad de Medicina Universidad de Valparaiso, projecto MECESUP UVA-106, and Fondation pour la Recherche Médicale; M.R. is founded by Ministère de l'Enseignement Supérieur et de la Recherche.

Received: April 10, 2007
Revised: October 17, 2007
Accepted: November 26, 2007
Published: January 23, 2008

REFERENCES

Akkal, D., Bioulac, B., Audin, J., and Burbaud, P. (2002). Comparison of neuronal activity in the rostral supplementary and cingulate motor areas during a task with cognitive and motor demands. Eur. J. Neurosci. *15*, 887–904.

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

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

Aston-Jones, G., and Cohen, J.D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. Annu. Rev. Neurosci. *28*, 403–450.

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

Botvinick, M.M., Cohen, J.D., and Carter, C.S. (2004). Conflict monitoring and anterior cingulate cortex: an update. Trends Cogn. Sci. *8*, 539–546.

Bouret, S., and Sara, S.J. (2005). Network reset: a simplified overarching theory of locus coeruleus noradrenaline function. Trends Neurosci. *28*, 574–582.

Braver, T.S., Reynolds, J.R., and Donaldson, D.I. (2003). Neural mechanisms of transient and sustained cognitive control during task switching. Neuron *39*, 713–726.

Buzsaki, G. (2006). Rhythms of the Brain (Oxford: Oxford University Press).

Cohen, J.D., Aston-Jones, G., and Gilzenrat, M.S. (2004). A systems-level perspective on attention and cognitive control. In Cognitive Neuroscience of Attention, M.I. Posner, ed. (New York: Guilford), pp. 71–90.

Daw, N.D., Niv, Y., and Dayan, P. (2005a). Actions, Policies, Values, and the Basal Ganglia. In Recent Breakthroughs in Basal Ganglia Research, E. Bezard, ed. (New York: Nova Science Publishers).

Daw, N.D., Niv, Y., and Dayan, P. (2005b). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat. Neurosci. *8*, 1704–1711.

Dehaene, S., Kerszberg, M., and Changeux, J.P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. Proc. Natl. Acad. Sci. USA *95*, 14529–14534.

Dosenbach, N.U., Visscher, K.M., Palmer, E.D., Miezin, F.M., Wenger, K.K., Kang, H.C., Burgund, E.D., Grimes, A.L., Schlaggar, B.L., and Petersen, S.E. (2006). A core system for the implementation of task sets. Neuron *50*, 799–812.

Doya, K. (2002). Metalearning and neuromodulation. Neural Netw. *15*, 495–506.

Duncan, J. (2006). EPS Mid-Career Award 2004: brain mechanisms of attention. Q. J. Exp. Psychol. (Colchester) *59*, 2–27.

Edwards, E., Soltani, M., Deouell, L.Y., Berger, M.S., and Knight, R.T. (2005). High gamma activity in response to deviant auditory stimuli recorded directly from human cortex. J. Neurophysiol. *94*, 4269–4280.

Gehring, W.J., and Willoughby, A.R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. Science *295*, 2279–2282.

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

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

Isoda, M., and 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., and Schall, J.D. (2003). Performance monitoring by the anterior cingulate cortex during saccade countermanding. Science *302*, 120–122.

Johnston, K., Levin, H.M., Koval, M.J., and Everling, S. (2007). Top-down control-signal dynamics in anterior cingulate and prefrontal cortex neurons following task switching. Neuron *53*, 453–462.

Kawato, M., and Samejima, K. (2007). Efficient reinforcement learning: computational theories, neuroscience and robotics. Curr. Opin. Neurobiol. *17*, 205–212.

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

Kondo, H., Osaka, N., and Osaka, M. (2004). Cooperation of the anterior cingulate cortex and dorsolateral prefrontal cortex for attention shifting. Neuroimage *23*, 670–679.

Logothetis, N.K. (2003). The underpinnings of the BOLD functional magnetic resonance imaging signal. J. Neurosci. *23*, 3963–3971.

Logothetis, N.K., and Wandell, B.A. (2004). Interpreting the BOLD signal. Annu. Rev. Physiol. *66*, 735–769.

Mainy, N., Kahane, P., Minotti, L., Hoffmann, D., Bertrand, O., and Lachaux, J.P. (2006). Neural correlates of consolidation in working memory. Hum. Brain Mapp. *28*, 183–193.

Matsumoto, K., Suzuki, W., and Tanaka, K. (2003). Neuronal correlates of goal-based motor selection in the prefrontal cortex. Science *301*, 229–232.

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

Nakamura, K., Roesch, M.R., and Olson, C.R. (2005). Neuronal activity in macaque SEF and ACC during performance of tasks involving conflict. J. Neurophysiol. *93*, 884–908.

Niessing, J., Ebisch, B., Schmidt, K.E., Niessing, M., Singer, W., and Galuske, R.A. (2005). Hemodynamic signals correlate tightly with synchronized gamma oscillations. Science *309*, 948–951.

Niki, H., and Watanabe, M. (1979). Prefrontal and cingulate unit activity during timing behavior in the monkey. Brain Res. *171*, 213–224.

Procyk, E., and Goldman-Rakic, P.S. (2006). Modulation of dorsolateral prefrontal delay activity during self-organized behavior. J. Neurosci. *26*, 11313–11323.

Procyk, E., Tanaka, Y.L., and Joseph, J.P. (2000). Anterior cingulate activity during routine and non-routine sequential behaviors in macaques. Nat. Neurosci. *3*, 502–508.

Roelofs, A., van Turennout, M., and Coles, M.G. (2006). Anterior cingulate cortex activity can be independent of response conflict in Stroop-like tasks. Proc. Natl. Acad. Sci. USA *103*, 13884–13889.

Rushworth, M.F., Walton, M.E., Kennerley, S.W., and Bannerman, D.M. (2004). Action sets and decisions in the medial frontal cortex. Trends Cogn. Sci. *8*, 410–417.

Rushworth, M.F., Behrens, T.E., Rudebeck, P.H., and Walton, M.E. (2007). Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. Trends Cogn. Sci. *11*, 168–176.

Satoh, T., Nakai, S., Sato, T., and Kimura, M. (2003). Correlated coding of motivation and outcome of decision by dopamine neurons. J. Neurosci. *23*, 9913–9923.

Schultz, W. (2000). Multiple reward signals in the brain. Nat. Rev. Neurosci. *1*, 199–207.

Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. Annu. Rev. Psychol. *57*, 87–115.

Seamans, J.K., and Yang, C.R. (2004). The principal features and mechanisms of dopamine modulation in the prefrontal cortex. Prog. Neurobiol. *74*, 1–58.

Shima, K., and Tanji, J. (1998). Role for cingulate motor area cells in voluntary movement selection based on reward. Science *282*, 1335–1338.

Shima, K., Aya, K., Mushiake, H., Inase, M., Aizawa, H., and Tanji, J. (1991). Two movement-related foci in the primate cingulate cortex observed in signal-triggered and self-paced forelimb movements. J. Neurophysiol. *65*, 188–202.

Sutton, R.S., and Barto, A.G. (1998). Reinforcement Learning: An Introduction (Cambridge, MA; London, England: MIT Press).

Tanaka, S. (2006). Dopaminergic control of working memory and its relevance to schizophrenia: a circuit dynamics perspective. Neuroscience *139*, 153–171.

Ullsperger, M., and von Cramon, D.Y. (2006). The role of intact frontostriatal circuits in error processing. J. Cogn. Neurosci. *18*, 651–664.

van Veen, V., Holroyd, C.B., Cohen, J.D., Stenger, V.A., and Carter, C.S. (2004). Errors without conflict: implications for performance monitoring theories of anterior cingulate cortex. Brain Cogn. *56*, 267–276.

Vidal, F., Burle, B., Bonnet, M., Grapperon, J., and Hasbroucq, T. (2003). Error negativity on correct trials: a reexamination of available data. Biol. Psychol. *64*, 265–282.

Walton, M.E., Devlin, J.T., and Rushworth, M.F. (2004). Interactions between decision making and performance monitoring within prefrontal cortex. Nat. Neurosci. *7*, 1259–1265.