nature neuroscience

Optimal decision making and the anterior cingulate cortex

Steven W Kennerley[1,3], Mark E Walton[1], Timothy E J Behrens[1,2], Mark J Buckley[1] & Matthew F S Rushworth[1,2]

Learning the value of options in an uncertain environment is central to optimal decision making. The anterior cingulate cortex (ACC) has been implicated in using reinforcement information to control behavior. Here we demonstrate that the ACC's critical role in reinforcement-guided behavior is neither in detecting nor in correcting errors, but in guiding voluntary choices based on the history of actions and outcomes. ACC lesions did not impair the performance of monkeys (*Macaca mulatta*) immediately after errors, but made them unable to sustain rewarded responses in a reinforcement-guided choice task and to integrate risk and payoff in a dynamic foraging task. These data suggest that the ACC is essential for learning the value of actions.

Fundamental to decision making is the ability to use past experience to select the best course of action from among competing alternatives. Although choices may be guided by cues or instructions, in many situations voluntary behavior requires the selection of actions based on the expected value of the available options[1–4]. Learning theories argue that the expected value of a given choice is derived from the recent history of outcomes of that choice[5,6]. Whereas dopamine is widely believed to contribute to the development of action-outcome predictions[6–9], we present data here suggesting that a cortical region, the anterior cingulate cortex (ACC), has an essential role in both learning and using extended action-outcome histories to optimize voluntary choice behavior.

The ACC has been described as an interface between motivation, cognition and action, and has been implicated in using reinforcement information to control behavior[10–16]. Activity within the ACC sulcus (ACC_S) has frequently been recorded on error trials[17–22]. In particular, it has been proposed that the ACC_S, more than simply representing when the consequences of actions diverge from predicted goals, might signal the overall context in which errors occur and may therefore facilitate corrective choices[18,22,23]. However, whereas ACC inactivation or lesions have been associated with some impairment of error correction[18,24–26], the degree of disruption is inconsistent and it remains unclear whether an emphasis on error processing provides a full account of ACC function. Notably, for an animal such as a macaque monkey foraging in an uncertain environment, actions are rarely categorically correct or erroneous, and both positive and negative outcomes are important sources of information that can be used to develop a prediction about the value of available options. Moreover, although it has received less attention, there is evidence to suggest that the ACC_S may be vital for using rewards to guide choice behavior[11,13,15,27]. Thus, although it is possible that ACC activity is modulated when errors occur because it is registering errors as immediate consequences of actions, the critical function of the ACC may instead be in the construction of an extended choice-outcome history (using both rewarded and nonrewarded outcomes) to guide future decisions.

To investigate whether the ACC_S is better characterized as detecting/correcting errors or as mediating adaptive decision making, we measured the effects of selective ACC_S lesions (**Fig. 1a**) in three of nine rhesus monkeys (*Macaca mulatta*) trained on tasks assessing error- and reward-guided action selection (**Fig. 1b,c**). The lesions included the region in macaque ACC in which neurons with error- and reward-related activity have previously been reported[13,17–19] and which is thought to be homologous to the human ACC region in which error- and conflict-related activity have been reported[10–12,20,23]. The results demonstrate that the critical function of the ACC_S is not simply in mediating adaptive behavior following errors but rather in integrating reinforcement information over time to guide voluntary choice behavior. ACC_S lesions did not impair monkeys' performance immediately after errors, but impaired their ability to sustain rewarded responses in a reward-guided choice task (experiment 1) and to choose optimally in a dynamic foraging task (experiment 2). It was not the case that ACC_S lesions impaired all types of cost-benefit decisions, as such lesions did not impair performance on a task assessing work-based decision making (experiment 3). The ACC_S may therefore be a critical component of a neural circuit that uses past action-reward history to learn the value of actions to guide choice behavior.

RESULTS

Experiment 1

Monkeys chose between lifting or turning a joystick for food reward (**Fig. 1b,c**). In experiment 1, only one of the two actions was rewarded at any time ('correct' response). Monkeys had to sustain a response for 25 rewarded movements, after which the action-outcome contingencies reversed, requiring a switch to the other response for

[1]Department of Experimental Psychology, South Parks Road, Oxford OX1 3UD, UK. [2]Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB), John Radcliffe Hospital, Headington, Oxford OX3 9DU, UK. [3]Present address: Helen Wills Neuroscience Institute, University of California Berkeley, 132 Barker Hall, MC #3190, Berkeley, California 94720-3190, USA. Correspondence should be addressed to S.W.K. (skennerley@berkeley.edu).

Received 7 February; accepted 23 May; published online 18 June 2006; doi:10.1038/nn1724

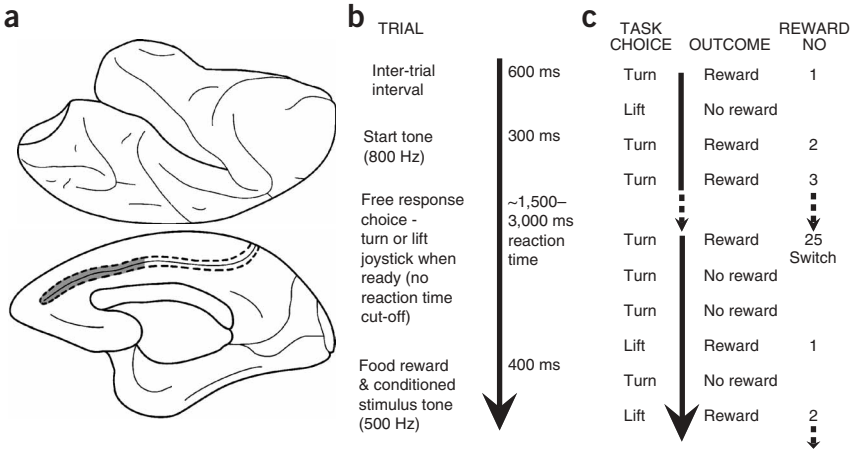

Figure 1 Diagram of the macaque brain and overview of experiments 1 and 2. (**a**) Intended ACC_S lesion (gray) on medial surface (bottom). (**b**) Schematic of a typical trial in experiments 1 and 2. On each trial, a start tone indicated that monkeys could choose one of two actions. If the chosen action was correct, a reward was delivered along with an auditory conditioned stimulus. (**c**) Schematic of action-outcome contingencies in experiment 1. Only one response was rewarded during a block of trials, though which of the two actions was correct switched every 25 rewarded trials. When an expected reward no longer followed a previously rewarded response (for example, turn movement), the monkey had to switch to the alternative response (for example, lift movement) to receive a reward. In experiment 2, rewards were assigned with independent probabilities to each response on each trial.

25 further rewards in order to obtain food at an optimal rate. Thus, the only information guiding choices was the expectation of reward based on the reinforcement of previous choices.

ACC_S lesions caused a marked decrement in overall performance in experiment 1A (**Fig. 2a**); unlike the control (CON) group, the ACC_S group made significantly more errors post operatively ($F_{1,7} = 8.75$, $P = 0.021$). Yet, rather than the deficit being one of detecting and/or correcting errors, the impairment reported here seems to have been caused by problems in sustaining rewarded behavior (**Fig. 2b**); the ACC_S group made significantly fewer correct responses on trials immediately following rewards (group × session: $F_{1,7} = 13.05$, $P = 0.009$) but were not impaired on the first trial after an error (group × session: $F_{1,7} = 0.53$, $P = 0.49$). Similarly, when we compared performance on the ten trials before an imposed response switch with that on the ten trials after a switch (**Fig. 2c**), we found that postoperative performance of the ACC_S group was significantly impaired (group × session: $F_{1,7} = 6.70$, $P = 0.036$), yet the ACC_S group showed a similar degree of impaired performance on the trials before and after a switch trial (group × session × switch: $F_{1,7} = 0.046\ 1$, $P = 0.84$). Moreover, the ACC_S-lesioned group did not differ from the control group in the number of trials they took to switch responses after a change in reinforcement assignment (**Supplementary Fig. 1** online). The ACC_S-lesioned group also did not differ from the control group in their inter-response times (**Supplementary Fig. 2** online). These results confirm that the deficit seen after ACC_S lesions was not simply related to changing to the new response after an error.

To examine the degree to which ACC_S lesions affect the use of positive reinforcement information to sustain correct performance, we performed a more detailed analysis ('EC' analysis, **Fig. 3a**,**b**) in which we examined the impact of each correct ('C') response on overall performance subsequent to each error ('E'). Whereas controls reached peak performance after 2–3 successive rewarded trials after an error ($EC_2 + 1$ or $EC_3 + 1$), the ACC_S group was impaired in their ability to use positive reinforcement to work out which action to make ($F_{1,7} = 12.46$, $P = 0.01$). Even after maintaining the same movement for more than eight consecutively rewarded trials, ACC_S-lesioned monkeys were still likely to revert back to the incorrect, unrewarded movement; the ACC_S group was significantly less likely to sustain the same movement for eight consecutively rewarded trials ($EC_8 + 1$) compared to the CON group ($t_7 = -5.0$, $P = 0.002$).

This inability to sustain rewarded action selection, yet with preserved error correction, remained evident in two further experiments. In experiment 1B, a salient sensory event, the brief extinction of the room light whenever an incorrect action was chosen, acted as an

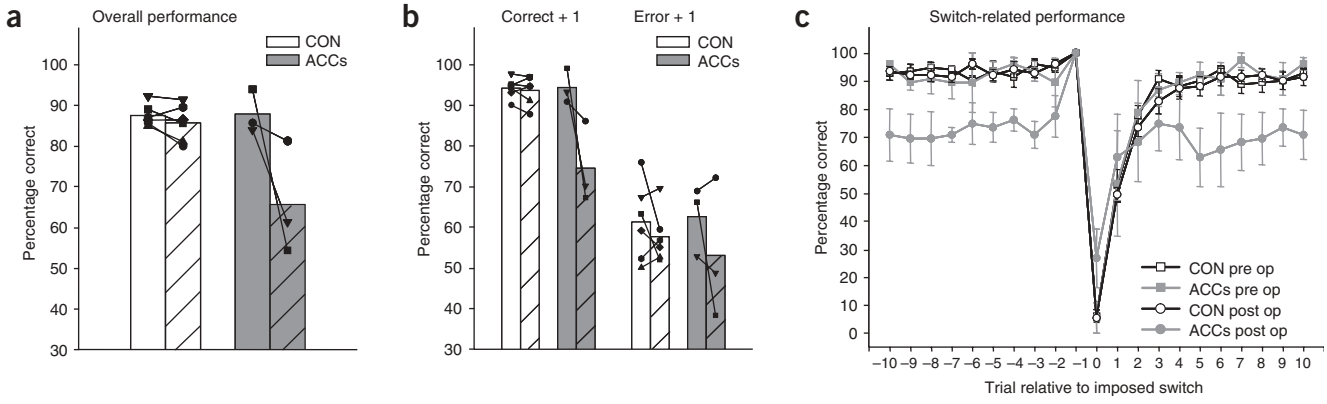

Figure 2 Performance in experiment 1A. (**a**) Overall performance of control (CON) and ACC_S-lesioned groups. Unfilled bars, preoperative performance. Hatched bars, postoperative performance. (**b**) Performance on the trial immediately following a rewarded ('Correct + 1') and nonrewarded ('Error + 1') response. Bars are labeled as in **a**. (**c**) Percent correct (± s.e.m.) on each of the 10 trials before and after an imposed switch (trial 0). The experimental design entailed that switches ('0' on *x*-axis) were only imposed after a correct response ('–1'). Notably, despite overall high levels of performance (as seen in **a**) and relatively good switching (as seen here), the monkeys rarely corrected an error on the very next trial (**b**, right).

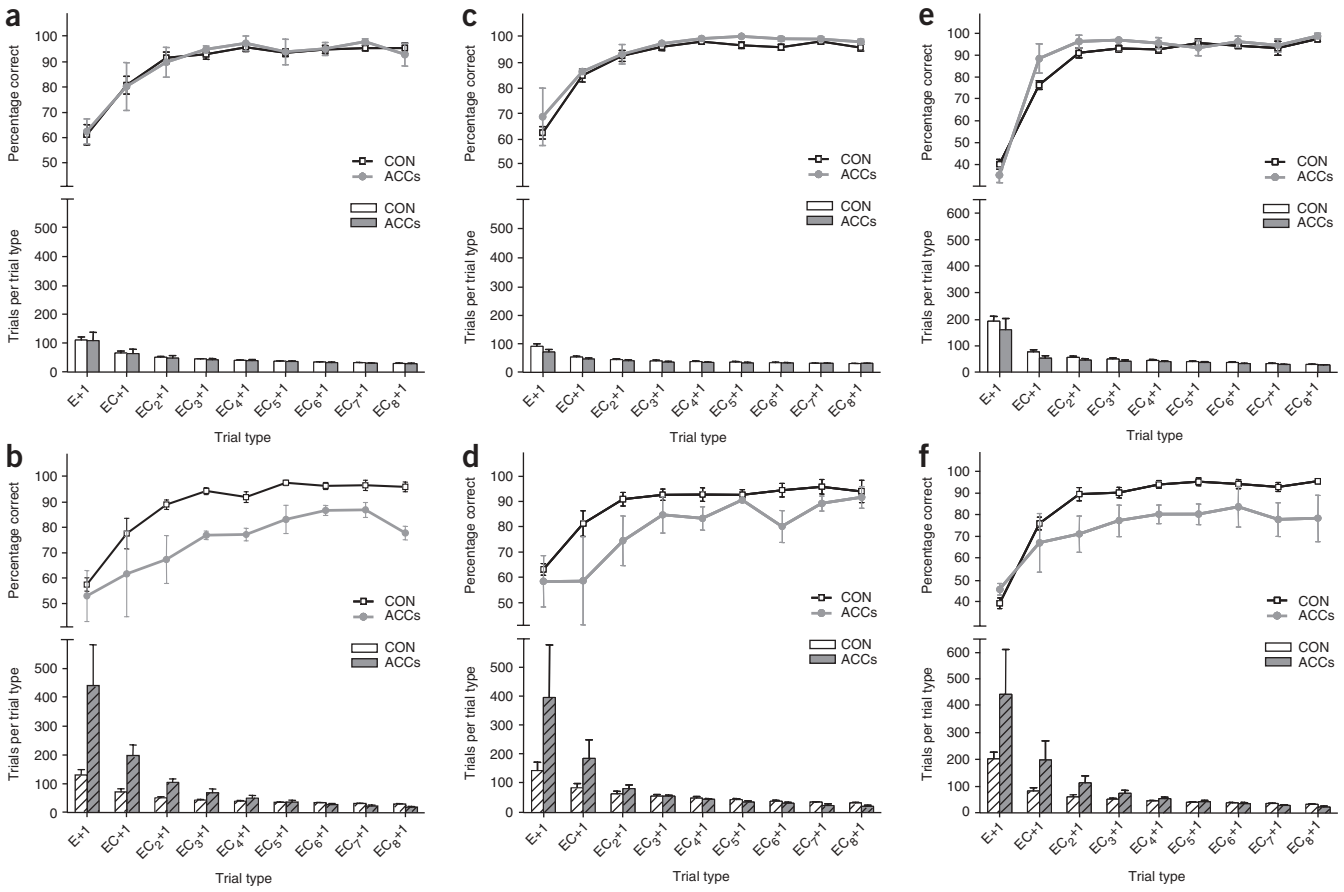

Figure 3 Performance for sustaining rewarded behavior following an error in experiment 1. (**a**–**f**) Preoperative (**a**,**c**,**e**) and postoperative (**b**,**d**,**f**) performance in experiments 1A, 1B and 1C, respectively. Each line graph shows the percentage of trials of each trial type that were correct (± s.e.m.). The trial types are plotted along the *x*-axis, starting with the trial immediately following an error ('E + 1'). The next data point corresponds to the trial after one error and one correct response ('EC + 1'). Similarly, 'EC_2 + 1' corresponds to the trial after one error and two correct responses, and so on. In each panel, moving from left to right corresponds to the monkey acquiring more instances of positive reinforcement after making the correct action, subsequent to an earlier error. Histograms indicate the number of instances of each trial type (± s.e.m.). Histogram bars are labeled as in **Figure 2a**.

additional error signal. In experiment 1C, a delay between action and food delivery, initially 500 ms long and then increasing by 500 ms for each subsequent error, indicated an incorrect response on the task-imposed switch trials. Regardless of the subtlety or the saliency of the error, we observed a similar pattern of results (**Figs. 3c–f**): in both experiments 1B and 1C, as in experiment 1A, the ACC_S-lesioned monkeys were not impaired on the trials that immediately followed errors (E + 1 condition, $P > 0.1$ in both cases); nonetheless, the EC analysis for experiments 1B and 1C showed that a history of rewarded action selection had less impact on the subsequent responses of ACC_S monkeys (experiment 1B, group × session: $F_{1,7} = 6.18$, $P = 0.042$; experiment 1C, group × session: $F_{1,7} = 6.35$, $P = 0.040$). In both tasks, the ACC_S group was significantly less likely to sustain the same movement for eight consecutively rewarded trials (EC_8 + 1) compared to the CON group (experiment 1B: $t_7 = -2.97$, $P = 0.021$; experiment 1C: $t_7 = -3.11$, $P = 0.017$). When pooling the data from experiments 1A and 1B, in which the categorization of errors was unambiguous, it was clear that the ACC_S-lesioned monkeys reacted in a similar way whether the errors occurred in the context of a task-imposed switch or at other times (**Supplementary Fig. 3** online). Compared to preoperative performance, postoperatively the ACC_S-lesioned monkeys reacted significantly worse following both switch errors ($F_{1,7} = 5.75$, $P = 0.048$) and nonswitch errors ($F_{1,7} = 12.26$, $P = 0.010$) as compared to controls.

The findings strengthen the conclusion drawn from experiment 1A that monkeys do not interpret the significance of errors and rewards in the same way as they interpret explicit sensory instructions to make one action or another. Instead, a single error is weighted against the recent reinforcement history. The impact of the reward history, however, was markedly diminished in the ACC_S monkeys. The tendency of ACC_S monkeys to be slightly quicker to update to the new response after changes in reinforcement assignment, a tendency evident in all three versions of experiment 1 (**Supplementary Fig. 1** online) might also be a reflection of the reduced impact of previous reinforcement history on choices.

As a complement of the EC analysis, we performed an 'EE' analysis to examine the impact of each additional error on overall performance after each initial error. We found no significant effect of the ACC_S lesion ($P > 0.1$). The analysis was not as powerful as the EC analysis, as the sample sizes were small, even after pooling data across experiments, because monkeys rarely made extended sequences of consecutive errors.

An alternative way to examine the influence of previous reinforcement history on subsequent choices is to use a multiple logistic regression analysis to obtain separate estimates of the weight of

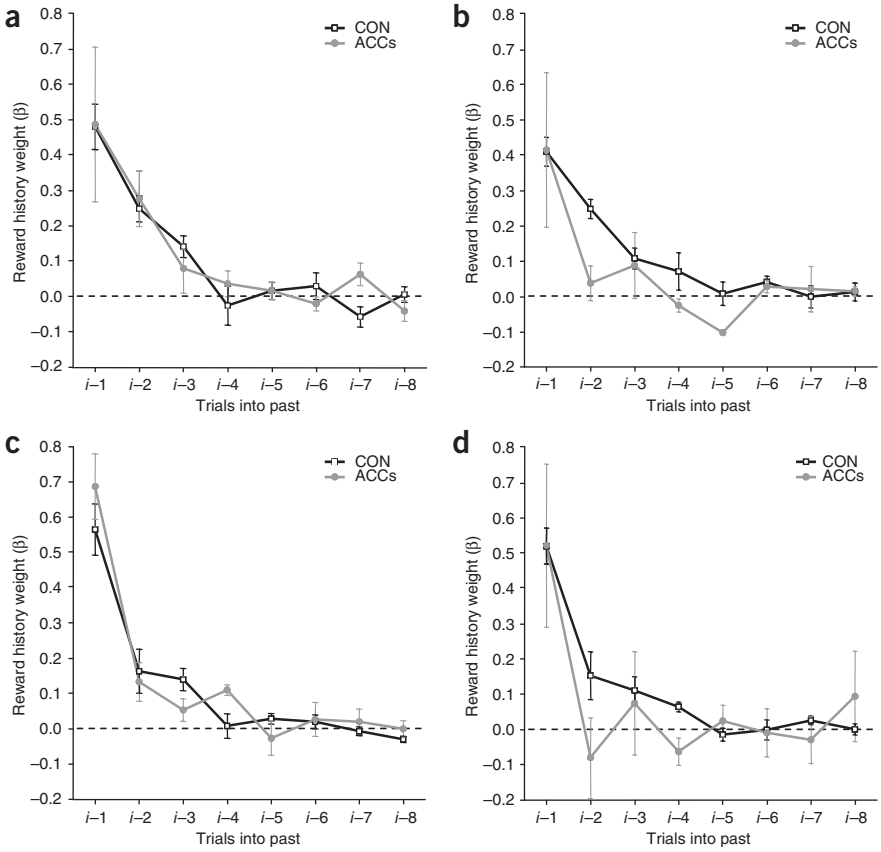

Figure 4 Estimates of the influence of previous reward history on current choice, in experiment 1. Each point represents a β value (± s.e.m.) derived from the multiple logistic regression of choice on the current trial (*i*) against the outcomes (rewarded or unrewarded) on the previous eight trials. Larger β values indicate greater influence of trial $i - x$ in the determination of the current choice (*i*). $i - 1$, previous trial; $i - 2$, two trials earlier, and so on. (**a**–**d**) Preoperative (**a**,**c**) and postoperative (**b**,**d**) performance in experiments 1A and 1B, respectively.

influence of each previous trial outcome in the reward history on the current trial[28,29]. The analysis generated a set of weights, or β values, reflecting the strength of influence of the outcome of the previous trial ($\beta_1 X_{i-1}$) up through the influence of the outcome of the trial performed eight trials earlier, in the following form (**Supplementary Methods** online): $\log(p(Y_i = 1))/\log(p(Y_i = 0)) = \beta_1 X_{i-1} + \beta_2 X_{i-2} + \beta_3 X_{i-3} + \ldots \beta_8 X_{i-8}$, where $p(Y_i = 0)$ and $p(Y_i = 1)$ represent the probability of reward allocation on the current trial for a lift or turn of the joystick respectively, and X can take the values –1 or 1 for previously rewarded lifts or turns. The β values were then plotted separately for the control and ACC_S groups, for the experiments in which errors could be unambiguously categorized and compared statistically (experiments 1A and 1B, **Fig. 4**). In experiment 1A (**Fig. 4a**), the influence of previous outcomes on the current choice gradually declined with increasing separation from the current trial, and outcomes more than five trials earlier had little influence on determining choice on the current trial. After the ACC_S lesion (**Fig. 4b**), the influence of past trials declined significantly more quickly with separation from the current trial (group × session: $F_{1,7} = 12.26$, $P = 0.010$). Past reward history had less of an effect on performance in experiment 1B when an explicit cue told monkeys to switch responses (**Fig. 4c**). Nevertheless, once again the ACC_S lesion caused the influence of previous outcomes to wane significantly more quickly (**Fig. 4d**, group × session: $F_{1,7} = 7.80$, $P = 0.027$).

Experiment 2

Experiment 2 tested monkeys on a 'dynamic foraging' or discrete-trial matching protocol[3,28], which better reproduces the conditions faced by an animal foraging in an uncertain environment. In this experiment, monkeys were again free to choose between two actions, but instead of having response outcomes that were either categorically correct or categorically incorrect, the two actions were rewarded according to unequally assigned probabilities (0.4:0.1, 0.5:0.2, 0.75:0.25 and 1:0). This meant that experiment 2 was not best performed using the 'win-stay, lose-switch' strategy that was optimal in experiment 1. As in other matching tasks[3,28,30], reward allocation occurred independently on each trial for each action (that is, lift or turn movement), and, once allocated to a particular action, reward remained available for that action until the monkey selected that action. To optimize foraging efficiency, it was therefore necessary to sample both action alternatives, integrating trial-by-trial reinforcement information over time to develop a sense of the utility of each action.

For each ratio of reward probabilities (0.4:0.1, 0.5:0.2, 0.75:0.25 and 1:0), we defined the optimum response allocation ratio (r_{opt}, **Supplementary Methods** and **Supplementary Figs. 4** and **5** online). At r_{opt}, monkeys received rewards at the maximum possible average rate, and the reward rate associated with each action was the same ('matching'). The ACC_S group was generally much slower than the control group to approach the optimum ratio threshold of action choices (**Fig. 5a** and **Supplementary Fig. 6** online; main effect of group: $F_{1,5} = 16.12$, $P = 0.01$). Notably, this was only observed in the three conditions in

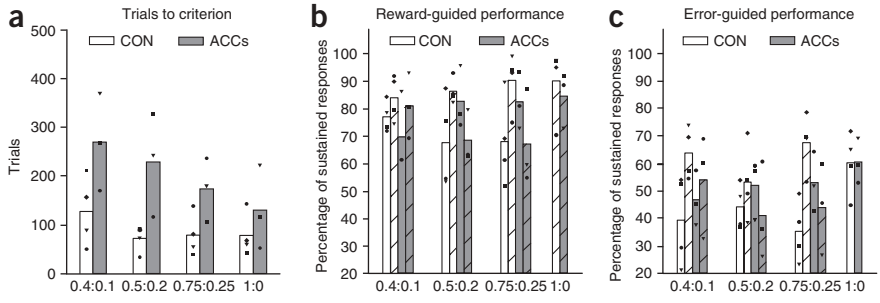

Figure 5 Postoperative performance in the matching task in experiment 2. (**a**) Number of trials required to exceed the optimum response-ratio threshold (**Supplementary Methods**). (**b**,**c**) Percentage of sustained low probability (*L*, unfilled bars) and high probability (*H*, hatched bars) responses (that is, successive *L* or successive *H* responses) following a rewarded (**b**) or a nonrewarded (**c**) response.

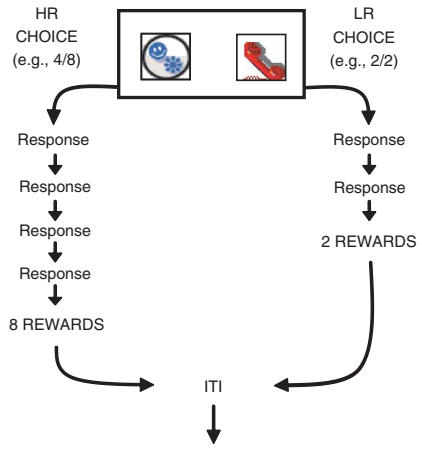

Figure 6 Schematic of a sample choice trial in experiment 3. Selection of one stimulus caused the other to disappear. The monkey then had to touch the chosen stimulus according to the stimulus' defined response cost in order to receive the number of food pellets associated with the stimulus.

which the outcome of each response was probabilistic, with the ACC_S monkeys taking significantly more trials to reach this criterion (for 0.4:0.1, 0.5:0.2 and 0.75:0.25, all $P < 0.05$, control versus ACC_S), but not when one action was deterministically rewarded (1:0, $P > 0.1$).

Again, the maladaptive behavior was not confined to trials following nonrewarded choices. ACC_S-lesioned monkeys were both less likely to sustain a rewarded action following the selection of the high probability (H) response compared to the low probability (L) response (**Fig. 5b**, group × response: $F_{1,5} = 22.52$, $P = 0.005$) and less likely to sustain an unrewarded action following an H response compared to an L response (**Fig. 5c**, group × response: $F_{1,5} = 10.32$, $P = 0.024$). This emphasizes the importance of the ACC_S for processing the behavioral value of rewarded and nonrewarded actions within the context of a dynamic environment.

Experiment 3

It was not the case that ACC_S lesions impaired all types of cost-benefit decisions. Six of the same monkeys were concurrently, and preoperatively, taught a decision-making task in which they chose between two visual stimuli that differed in the number of responses required before reward, the quantity of food that the monkey received or both (**Fig. 6**). Preoperatively, all monkeys consistently chose the option that either led to a greater reward or, in the case where both stimuli led to identical rewards, required less work. In contrast to experiments 1 and 2, performance in experiment 3 was unimpaired following surgery (**Fig. 7**), with all monkeys continuing to select the appropriate high reward or low work option.

DISCUSSION

Many accounts of ACC function have emphasized its involvement in error processing. Error-related activity can be recorded from individual or populations of cells in the region of the ACC_S (refs. 13,17,19–21). It has also been argued that the ACC_S does not simply detect errors, but rather signals a change in outcome that drives changes in behavior[18,31]. Despite the emphasis on the role of the ACC in error processing, it is important to note that many ACC_S cells are responsive to both rewards and errors[13,19,32].

The present experiments sought to quantify the role of the ACC_S in guiding voluntary behavior on the basis of both rewards and errors. In experiment 1, only one of the two actions was rewarded at any time (correct response) and monkeys had to sustain a response until the action-outcome contingencies reversed, at which point they were required to switch to the other response and sustain it for 25 further reward trials in order to obtain food at an optimal rate. Removal of the ACC_S caused significant impairments in all versions of experiment 1, but, despite the emphasis that has been placed on the monitoring and detection of errors by the ACC_S, the monkeys with ACC_S lesions did not perform significantly worse than control monkeys on the trials that followed errors (**Fig. 2b**,**c** and **Fig. 3b**,**d**,**f**). The number of trials it took both control and ACC_S monkeys to try an alternative response following an imposed switch in reinforcement varied in similar ways as the saliency of the error was changed (**Supplementary Fig. 1**). When the error became more salient, monkeys were slightly quicker to try an alternative response than when the error was less salient. Monkeys in both groups took approximately three trials to notice that the reinforcement was being increasingly delayed in experiment 1C. Thus ACC_S-lesioned monkeys seemed to be as quick to detect a decrement in reinforcement as the controls. The ACC_S-lesioned monkeys were, however, less likely to repeat a response that had been rewarded (**Fig. 2b** and **Fig. 3b**,**d**,**f**). Rather than an error detection deficit, the pattern of impairment suggested that the ACC_S area is vital for sustaining rewarded action selection.

However, inspection of the control monkeys' performance (**Fig. 3a**,**c**,**e**) revealed another notable finding: even after performing tens of thousands of trials on three different versions of the task during which failure to receive an expected reward always indicated that the other response was correct, controls never immediately corrected more than an average of 68% of their errors. This suggests that errors and rewards do not naturally operate like the explicit sensory cues that instruct actions in a conditional learning protocol[4,33]. Instead, to a foraging animal, a single negative outcome is a piece of evidence that a given choice may no longer be optimal, a fact that is weighed against the recent history of reinforcement.

Notably, this suggests that the ACC_S deficit might not simply be one of using errors to guide action, but rather a failure to integrate reinforcement history over time to develop a representation of the value for each option. To test this hypothesis, we calculated the influence of each previous outcome on subsequent choices (**Fig. 4**). Whereas in all

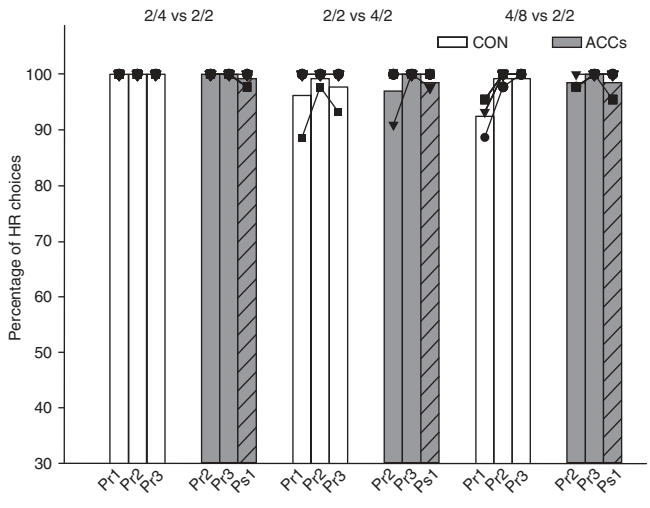

Figure 7 Percentage of high reward options selected for each of the three choice pairs across three preoperative (unfilled bars) and one postoperative (hatched bars) sessions. Numbers above bars represent the designation of the choice pair (work to be performed in terms of number of screen presses, and reward available for each stimulus).

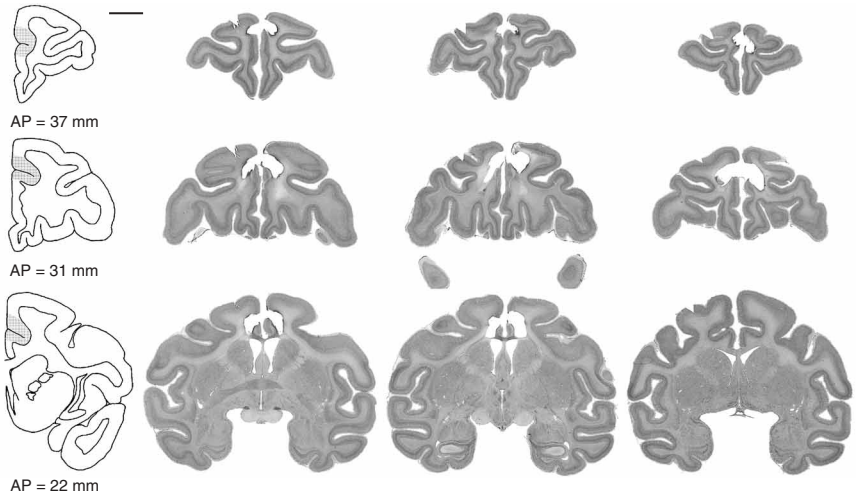

Figure 8 Coronal sections showing the cingulate lesion in all three monkeys that received surgery in experiments 1–3. Left, reference coronal slices redrawn from the atlas in ref. 50. Stippling indicates the intended lesion on a coronal cross-section. AP, approximate anterior-posterior position coordinates according to the atlas. At right, three coronal sections through the brains of monkeys ACC_S1, ACC_S2 and ACC_S3 (from left to right). In each case, the first section is taken at approximately the level of the emergence of the rostral sulcus, the second just rostral to the genu of the corpus callosum and the third section at the level of the bow of the arcuate sulcus. The lesion in monkey ACC_S3 did not extend as far caudally as it did in monkeys ACC_S1 and ACC_S2. Scale bar, 10 mm (in each photographed coronal section).

monkeys the influence of previous outcomes waned as more trials separated the outcome from the current choice, the influence of previous reward history was significantly lower in the ACC_S group.

The finding that both midbrain dopaminergic neurons and ACC neurons respond when errors occur has led to the proposal that these two regions may be parts of a reinforcement/error-driven learning system[22]. The ACC sulcus is the recipient of dopaminergic innervation from the midbrain, including the ventral tegmental area and the substantia nigra[34]. Activity of some midbrain dopamine neurons reflects the difference between the reward value of the current action and the 'value estimate' of the action, which is derived from the average magnitude of reward on the most recent trials[6]. Similarly, the EC and β-value analyses in experiment 1 revealed that the likelihood that the correct response would be selected following a rewarded trial was dependent on the previous reward history, with the control group typically reaching peak performance after receiving 2–3 consecutive rewards. In contrast, it seems that the ACC_S group was unable to properly use previous reward history to develop a value estimate for each action needed to perform optimally (**Fig. 4b**,**d**). One of the functions of the ACC_S may therefore be to build, modify and/or use action-outcome contingencies to develop value estimates for the available options, which can then be used to guide optimal choice behavior.

To investigate this hypothesis—that the ACC_S deficit might not simply be one of detecting errors, but a failure to integrate reinforcement history over time in order to develop a representation of the value of each option—experiment 2 tested monkeys on a dynamic foraging or discrete-trial matching protocol[3,28,30]. Such protocols mimic the situations faced by a monkey foraging in an uncertain environment, where rewards may have only a probabilistic association with a response and where the rewards associated with one response may become more plentiful again if that response is not made for some time. In keeping with other matching tasks[3,28,30], a decision's value varied with the time since it was last made, because the expected reward was contingent on the rate of making that decision. To optimize foraging efficiency, it was therefore necessary to sample both action alternatives, integrating action outcome information (both reward and error) over time to develop a representation of the utility of each action.

When performing this task, the control monkeys quickly learned to make more of the response that had a greater probability of reward. The ACC_S monkeys, however, were significantly slower to approach an optimal level of performance when the rewards were delivered probabilistically, but were no worse than controls when one response was always consistently associated with reward and the other was never followed by a reward (**Fig. 5**). The results also confirmed that it was possible to observe altered error correction behavior after ACC_S disruption[18,26] but that the deficit only appeared once the context provided by the recent reward history was taken into account (**Fig. 5b**,**c**). Whereas the results of experiment 1 may suggest that ACC_S-lesioned monkeys could detect changes in action-outcome contingency (**Supplementary Fig. 1**), the results of both experiments 1 and 2 demonstrate that these monkeys do not use the history of previous reinforcement to guide subsequent choices. Experiments 1 and 2 together demonstrate that it does not matter whether the reinforcement history contains positive outcomes ('rewards') or negative outcomes ('errors'): in either case, the influence of action-outcome on subsequent choices is reduced after an ACC_S lesion. In experiment 2, compared to the control group, the ACC_S monkeys seemed unable to interpret the significance of rewarded and nonrewarded actions in the overall context of the dynamically changing reward probabilities associated with each action, causing them to be more likely in general to repeat the low (*L*) than the high probability (*H*) response after an error (**Fig. 5c**). It has been suggested that midbrain dopamine neurons, with which the ACC is interconnected, encode a context-dependent prediction error generated from the history of reinforcement[6,8]. Similarly, it has been argued that blood oxygenation signals in the ACC reflect not just the occurrence of errors but also their likelihood[23]. The present results suggest that the ACC response to both error and reward outcomes is influenced by the reinforcement context.

It could be argued that the ACC_S deficits in the reinforcement-guided choice tasks can be attributed to a general failure of working memory, which might compromise recall of the action performed on the previous trial. Although there clearly is a mnemonic component in remembering a history of past actions and outcomes, unlike lateral prefrontal lesions large ACC lesions that include the ACC_S do not impair working memory tasks such as delayed alternation or delayed response tasks[26,35]. Similarly, it is unlikely that the effect of the ACC_S lesion was simply to disconnect parts of lateral prefrontal cortex and the hippocampal formation through damage to the cingulum bundle[36], as complementary lesions to the anterior cingulate gyrus did not cause the same pattern of impairments (data not shown). Moreover, the finding that ACC_S lesions did not impair performance when actions were deterministically rewarded—the 1:0 condition of experiment 2, in which the requirement to sustain one action and inhibit the other was maximal—suggests that the ACC_S deficit cannot be explained simply as a failure of inhibition, as has been proposed for other prefrontal brain areas[37,38].

Although ACC_S lesions disrupted the monkeys' ability to use reinforcement history to guide their choices, it was not the case that

this lesion induced irrational decisions about reward in all circumstances. Monkeys with ACC_S lesions consistently selected optimally between visual stimuli that differed in their work requirements and/or reward quantity. Although activity in dorsal ACC changes in relation to reward size[27,39] and ACC_S cells track progress through a series of movements toward a reward[40], such information is also represented in other interconnected areas including the ventral striatum, orbitofrontal and prefrontal cortex, and adjacent cingulate gyrus[39,41,42].

The current results imply that even in the absence of the ACC_S, monkeys can maintain representations of reward magnitude, some aspects of response cost or the combined value of these factors, but not of reward probability. The assessment of the reward probability associated with an action depends, at least in part, on ACC_S integrity[1,27] even if, in the intact brain, reward probability information is widely distributed throughout the brain, for example in regions such as the parietal cortex[2]. It is unlikely that the ACC_S is solely responsible for assessing the reward probability associated with an action; it is likely to be the cortical component of a distributed circuit that also includes the caudate and midbrain dopaminergic system with which the ACC is strongly interconnected[6,34,43]. Neurons in the dorsolateral prefrontal cortex also encode information about the past history of the monkey's actions and the rewards that have been received[44].

Whereas the current results have been considered in relation to an error processing account of ACC function, it should be noted that an alternative account of ACC function has emphasized its role in detecting response conflict[12] or error likelihood[23] rather than error monitoring *per se*. Although single neuron recording and lesion studies have not confirmed the importance of the ACC_S in protocols where response conflict is induced by discrepant visual instructions[19,24–26,45], it is possible that the ACC_S is concerned with the resolution of response conflict on the basis of reward history (as in experiments 1 and 2, but not experiment 3).

The failure of the ACC_S-lesioned monkeys to continue making a response even when that response was being rewarded (in experiment 1, **Fig. 3**) was a notable aspect of the impairment. It has been suggested that the locus coeruleus–norepinephrine system is concerned with the transition from sustained responding when task contingencies are well understood, and can therefore be exploited, to a more explorative mode of behavior in which the monkey searches for alternative actions[46,47]. Moreover it has been suggested that the locus coeruleus activity patterns are, in turn, influenced by the ACC (ref. 46). The current results are consistent with such a hypothesis.

Conclusions

By examining error detection and correction within the context of a reversal task and a dynamic foraging task, the present results illustrate the vital role of the ACC_S in integrating reinforcement information over time rather than in monitoring whether a single action achieved its expected outcome or in signaling the need for adaptive behavior. During foraging and everyday decision making, there is often only a probabilistic chance, rather than a certainty, of success. In such situations, the lack of reward on a particular occasion may not necessarily signal the need to switch to an alternative course of action. Rather than simply detecting errors, the ACC_S may therefore be the cortical component of a distributed circuit for learning and maintaining the ongoing values of actions[6,27,43]. Unlike other premotor regions, the ACC_S has both connections to prefrontal and subcortical limbic structures and projections to the spinal cord[48,49], and it is also thought to receive reward-prediction error signals generated by midbrain dopamine neurons[22]. The ACC_S is thus in a prime anatomical position to compute the contextual value of each option based on multiple decision variables such as action and reward history, risk and expected payoff. Ultimately, such flexible coding of choice context and expected and obtained outcome, if integrated across time, would place the ACC in a prime position to determine which actions are worth making.

METHODS

Subjects. We used 9 male rhesus macaque monkeys, aged 4–6 years and weighing 4–10 kg. After preoperative testing, anterior cingulate sulcus (ACC_S) lesions were made in three monkeys (ACC_S1, ACC_S2, ACC_S3). The lesions included both banks of the ACC_S from its rostral inception to a caudal position level with the midpoint of the precentral dimple (**Figs. 1** and **8**; **Supplementary Methods**). The other six monkeys served as controls (CON) for experiments 1 and 3. Four of the monkeys also acted as controls in experiment 2. Surgical monkeys represented the full range of performance, and there was no preoperative difference between control and experimental groups. The studies were carried out under project and personal licenses from the British Home Office.

Experiment 1. Error- and reward-guided action selection. Monkeys were trained to lift or turn a joystick for 150 rewarded trials per day. Each testing session (pre- and postoperative) was composed of 5 consecutive testing days (750 rewarded trials). A correct trial yielded a reward food pellet (Noyes formula L/I, Research Diets) delivered to a food-well placed directly above the joystick. Only one of the two actions was ever rewarded at any one time (the 'correct' response). Monkeys were required to sustain a given response for 25 rewarded movements, at which point the action-outcome contingencies reversed, requiring the monkeys to switch to the other response and maintain choosing it for 25 rewards in order to continue to obtain food at an optimal rate. The intention was to quantify the degree to which any impairment after an ACC_S lesion was due to a failure to detect an error or a failure to integrate action outcomes (both rewarded and nonrewarded) before making the next choice.

Three distinct procedures of experiment 1 (experiments 1A–C) were used both before and after surgery (**Supplementary Methods**). In experiment 1A, an error was signaled by the simple omission of reward. In experiment 1B, errors were made salient by switching off the cubicle illumination for the duration of the subsequent intertrial interval (ITI). In experiment 1C, a delay between action and food delivery, initially 500 ms long and then increasing by 500 ms for each subsequent error, indicated an incorrect response on the task-imposed switch trials.

Experiment 2. Dynamic foraging probabilistic matching task. After postoperative testing on experiment 1C, monkeys were taught a dynamic probabilistic version of the joystick task. The protocol resembled a discrete trial version of a dynamic foraging task and was based on previous probability-matching protocols[3,28,30]. Monkeys either lifted or turned a joystick for 150 rewarded trials per day. Unlike experiment 1, however, rewards were not allocated to just one or the other movement with certainty; instead, on every trial the availability of reward after each movement was determined by two independent probabilistic algorithms. Notably, reward allocation occurred independently on each trial for each action, and, once allocated to a particular action, the reward remained available for that action until the monkey selected this action (**Supplementary Methods**).

The monkey's objective was to work out which response was the more profitable on any given testing day. We used 4 action-reward pair probability ratios: 0.4:0.1, 0.5:0.2, 0.75:0.25 and 1:0, where the numbers reflect the probability of reward for the high (*H*) and low (*L*) probability responses, respectively. The four reward ratios were counterbalanced for both the lift and turn movements, yielding eight conditions. For each action-reward ratio pair, we calculated the optimum ratio criterion of lift and turn responses to maximize expected reward, and then calculated the number of trials it took the monkeys to approach this criterion (**Supplementary Methods**).

In additional analyses, we considered whether monkeys changed to a new response or sustained the previous response after an unrewarded action. The data were plotted separately for trials in which the previous response was *H* or *L*. The nonappearance of reward on *L* and *H* trials should be treated differently by the monkeys because it may not be optimal to change away from the *H* response just because a given trial was unrewarded. To respond in this way, however, requires that the monkey remember the overall probabilistic value of

the response in addition to whether or not it had been followed by reward on the previous trial.

Experiment 3. Work-based cost-benefit decision making. Monkeys were trained to discriminate between different reward/effort stimuli. On choice trials, monkeys were presented with two different visual stimuli (6 cm × 6 cm) to the left and right of the center of the touch screen, at a distance of 20 cm from the front of the testing cage. On forced trials, only one stimulus was presented on one side of the screen, requiring monkeys to select the single displayed stimulus. Each stimulus differed in terms of its work quotient (the number of times it needed to be pressed before reward delivery) and/or its associated reward size. The association between each discrete stimulus and its work/reward ratio was kept constant throughout the study. Upon selecting one stimulus, a 350-Hz response tone was presented for 250 ms and both stimuli were extinguished; this was followed by the reappearance of the chosen stimulus 500 ms later at a random location on the screen. This would then be repeated (stimulus touched, response tone generated, stimulus extinguished, stimulus reappearance at new location) up to the number of times assigned for that stimulus until the monkeys made their final response, the selection of which resulted in the reward delivery along with the presentation of a second reward tone (500 Hz for 250 ms). The ITI was 6 s. During training and all subsequent testing, monkeys were always presented with the same pair of stimuli throughout the session. The stimulus with the greater reward size, or, if the reward sizes were equal, the stimulus requiring fewer responses, was termed the high-reward stimulus (HR); the other was the low-reward stimulus (LR).

Monkeys were tested with three choice pairs (2/4 versus 2/2; 2/2 versus 4/2; 4/8 versus 2/2; the first number indicates number of responses, the second the number of rewards for that particular stimulus) in three separate sessions, with approximately 90 d between sessions. Following surgery, the three ACC_S monkeys were retested with the same choice pairs.

Note: Supplementary information is available on the Nature Neuroscience website.

ACKNOWLEDGMENTS

We are grateful to D. Gaffan for advice and encouragement. Supported by the Medical Research Council. Additional support from the Clarendon Foundation (S.W.K.), the Wellcome Trust (M.E.W.) and the Royal Society (M.J.B and M.F.S.R.).

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/natureneuroscience
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. McCoy, A.N. & Platt, M.L. Risk-sensitive neurons in macaque posterior cingulate cortex. *Nat. Neurosci.* **8**, 1220–1227 (2005).
2. Platt, M.L. & Glimcher, P.W. Neural correlates of decision variables in parietal cortex. *Nature* **400**, 233–238 (1999).
3. Sugrue, L.P., Corrado, G.S. & Newsome, W.T. Matching behavior and the representation of value in the parietal cortex. *Science* **304**, 1782–1787 (2004).
4. Murray, E.A., Bussey, T.J. & Wise, S.P. Role of prefrontal cortex in a network for arbitrary visuomotor mapping. *Exp. Brain Res.* **133**, 114–129 (2000).
5. Sutton, R.S. & Barto, A.G. *Reinforcement Learning* (MIT Press, Cambridge, Massachusetts, 1998).
6. Bayer, H.M. & Glimcher, P.W. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47**, 129–141 (2005).
7. Schultz, W. Getting formal with dopamine and reward. *Neuron* **36**, 241–263 (2002).
8. Nakahara, H., Itoh, H., Kawagoe, R., Takikawa, Y. & Hikosaka, O. Dopamine neurons can represent context-dependent prediction error. *Neuron* **41**, 269–280 (2004).
9. Satoh, T., Nakai, S., Sato, T. & Kimura, M. Correlated coding of motivation and outcome of decision by dopamine neurons. *J. Neurosci.* **23**, 9913–9923 (2003).
10. Paus, T. Primate anterior cingulate cortex: where motor control, drive and cognition interface. *Nat. Rev. Neurosci.* **2**, 417–424 (2001).
11. Rushworth, M.F., Walton, M.E., Kennerley, S.W. & Bannerman, D.M. Action sets and decisions in the medial frontal cortex. *Trends Cogn. Sci.* **8**, 410–417 (2004).
12. Botvinick, M.M., Braver, T.S., Barch, D.M., Carter, C.S. & Cohen, J.D. Conflict monitoring and cognitive control. *Psychol. Rev.* **108**, 624–652 (2001).
13. Matsumoto, K., Suzuki, W. & Tanaka, K. Neuronal correlates of goal-based motor selection in the prefrontal cortex. *Science* **301**, 229–232 (2003).
14. Walton, M.E., Devlin, J.T. & Rushworth, M.F. Interactions between decision making and performance monitoring within prefrontal cortex. *Nat. Neurosci.* **7**, 1259–1265 (2004).
15. Hadland, K.A., Rushworth, M.F., Gaffan, D. & Passingham, R.E. The anterior cingulate and reward-guided selection of actions. *J. Neurophysiol.* **89**, 1161–1164 (2003).
16. Morecraft, R.J. & Van Hoesen, G.W. Convergence of limbic input to the cingulate motor cortex in the rhesus monkey. *Brain Res. Bull.* **45**, 209–232 (1998).
17. Amiez, C., Joseph, J.P. & Procyk, E. Anterior cingulate error-related activity is modulated by predicted reward. *Eur. J. Neurosci.* **21**, 3447–3452 (2005).
18. Shima, K. & Tanji, J. Role for cingulate motor area cells in voluntary movement selection based on reward. *Science* **282**, 1335–1338 (1998).
19. Ito, S., Stuphorn, V., Brown, J.W. & Schall, J.D. Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science* **302**, 120–122 (2003).
20. Ullsperger, M. & von Cramon, D.Y. Error monitoring using external feedback: specific roles of the habenular complex, the reward system, and the cingulate motor area revealed by functional magnetic resonance imaging. *J. Neurosci.* **23**, 4308–4314 (2003).
21. Gehring, W.J., Goss, B., Coles, M.G., Meyer, D.E & Donchin, E. A neural system for error detection and compensation. *Psychol. Sci.* **4**, 385–390 (1993).
22. Holroyd, C.B. & Coles, M.G. The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol. Rev.* **109**, 679–709 (2002).
23. Brown, J.W. & Braver, T.S. Learned predictions of error likelihood in the anterior cingulate cortex. *Science* **307**, 1118–1121 (2005).
24. Fellows, L.K. & Farah, M.J. Is anterior cingulate cortex necessary for cognitive control? *Brain* **128**, 788–796 (2005).
25. Swick, D. & Turken, A.U. Dissociation between conflict detection and error monitoring in the human anterior cingulate cortex. *Proc. Natl Acad. Sci. USA* **99**, 16354–16359 (2002).
26. Rushworth, M.F., Hadland, K.A., Gaffan, D. & Passingham, R.E. The effect of cingulate cortex lesions on task switching and working memory. *J. Cogn. Neurosci.* **15**, 338–353 (2003).
27. Amiez, C., Joseph, J.P. & Procyk, E. Reward encoding in the monkey anterior cingulate cortex. *Cereb. Cortex,* published online 5 October 2005 (doi:10.1093/cercor/bhj046).
28. Lau, B. & Glimcher, P.W. Dynamic response-by-response models of matching behavior in rhesus monkeys. *J. Exp. Anal. Behav.* **84**, 555–579 (2005).
29. Lee, D., Conroy, M.L., McGreevy, B.P. & Barraclough, D.J. Reinforcement learning and decision making in monkeys during a competitive game. *Brain Res. Cogn. Brain Res.* **22**, 45–58 (2004).
30. Herrnstein, R.J. *The Matching Law: Papers in Psychology and Economics* (eds. Rachlin, H. & Laibson, D.I.) (Harvard Univ. Press, Cambridge, Massachusetts, 1997).
31. Bush, G. *et al.* Dorsal anterior cingulate cortex: a role in reward-based decision making. *Proc. Natl Acad. Sci. USA* **99**, 523–528 (2002).
32. Niki, H. & Watanabe, M. Prefrontal and cingulate unit activity during timing behavior in the monkey. *Brain Res.* **171**, 213–224 (1979).
33. Genovesio, A., Brasted, P.J., Mitz, A.R. & Wise, S.P. Prefrontal cortex activity related to abstract response strategies. *Neuron* **47**, 307–320 (2005).
34. Williams, S.M. & Goldman-Rakic, P.S. Widespread origin of the primate mesofrontal dopamine system. *Cereb. Cortex* **8**, 321–345 (1998).
35. Murray, E.A., Davidson, M., Gaffan, D., Olton, D.S. & Suomi, S. Effects of fornix transection and cingulate cortical ablation on spatial memory in rhesus monkeys. *Exp. Brain Res.* **74**, 173–186 (1989).
36. Morris, R., Pandya, D.N. & Petrides, M. Fiber system linking the mid-dorsolateral frontal cortex with the retrosplenial/presubicular region in the rhesus monkey. *J. Comp. Neurol.* **407**, 183–192 (1999).
37. Aron, A.R., Robbins, T.W. & Poldrack, R.A. Inhibition and the right inferior frontal cortex. *Trends Cogn. Sci.* **8**, 170–177 (2004).
38. Roberts, A.C. & Wallis, J.D. Inhibitory control and affective processing in the prefrontal cortex: neuropsychological studies in the common marmoset. *Cereb. Cortex* **10**, 252–262 (2000).
39. Knutson, B., Taylor, J., Kaufman, M., Peterson, R. & Glover, G. Distributed neural representation of expected value. *J. Neurosci.* **25**, 4806–4812 (2005).
40. Shidara, M. & Richmond, B.J. Anterior cingulate: single neuronal signals related to degree of reward expectancy. *Science* **296**, 1709–1711 (2002).
41. Wallis, J.D. & Miller, E.K. Neuronal activity in primate dorsolateral and orbital prefrontal cortex during performance of a reward preference task. *Eur. J. Neurosci.* **18**, 2069–2081 (2003).
42. Roesch, M.R. & Olson, C.R. Neuronal activity related to reward value and motivation in primate frontal cortex. *Science* **304**, 307–310 (2004).
43. Samejima, K., Ueda, Y., Doya, K. & Kimura, M. Representation of action-specific reward values in the striatum. *Science* **310**, 1337–1340 (2005).
44. Barraclough, D.J., Conroy, M.L. & Lee, D. Prefrontal cortex and decision making in a mixed-strategy game. *Nat. Neurosci.* **7**, 404–410 (2004).
45. Nakamura, K., Roesch, M.R. & Olson, C.R. Neuronal activity in macaque SEF and ACC during performance of tasks involving conflict. *J. Neurophysiol.* **93**, 884–908 (2005).
46. Aston-Jones, G. & Cohen, J.D. An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu. Rev. Neurosci.* **28**, 403–450 (2005).
47. Yu, A.J. & Dayan, P. Uncertainty, neuromodulation, and attention. *Neuron* **46**, 681–692 (2005).
48. He, S.Q., Dum, R.P. & Strick, P.L. Topographic organization of corticospinal projections from the frontal lobe: motor areas on the medial surface of the hemisphere. *J. Neurosci.* **15**, 3284–3306 (1995).
49. Bates, J.F. & Goldman-Rakic, P.S. Prefrontal connections of medial motor areas in the rhesus monkey. *J. Comp. Neurol.* **336**, 211–228 (1993).
50. Paxinos, G., Huang, X.F. & Toga, A.W. *The Rhesus Monkey Brain in Stereotaxic Coordinates* (Academic Press, San Diego, 1999).