

The Role of the Anterior Cingulate Cortex in Choices based on Reward Value and Reward Contingency

Yogita Chudasama¹, Teresa E. Daniels², Daniel P. Gorrin², Sarah E.V. Rhodes², Peter H. Rudebeck² and Elisabeth A. Murray²

¹Department of Psychology, McGill University, Montreal, Canada and ²Section on the Neurobiology of Learning and Memory, Laboratory of Neuropsychology, National Institute of Mental Health, National Institutes of Health, Bethesda, MD, USA

Address correspondence to Y. Chudasama, Department of Psychology, McGill University, 1205 Dr Penfield Avenue, Montreal, QC, Canada H3A 1B1. Email: yogita.chudasama@mcgill.ca

Although several studies have emphasized the role of the anterior cingulate cortex (ACC) in associating actions with reward value, its role in guiding choices on the basis of changes in reward value has not been assessed. Accordingly, we compared rhesus monkeys with ACC lesions and controls on object- and action-based reinforcer devaluation tasks. Monkeys were required to associate an object or an action with one of two reward outcomes, and we assessed the monkey's shift in choices of objects or actions after changes in the value of 1 outcome. No group differences emerged on either task. For comparison, we tested the same monkeys on their ability to make choices guided by reward contingency in object- and action-based reversal learning tasks. Monkeys with ACC lesions were impaired in using rewarded trials to sustain the selection of the correct object during object reversal learning. They were also impaired in using errors to guide choices in action reversal learning. These data indicate that the role of the ACC is not restricted to linking specific actions with reward outcomes, as previously reported. Instead, the data suggest a more general role for the ACC in using information about reward and nonreward to sustain effective choice behavior.

Keywords: action-reversal, monkey, object-reversal, reward value

Introduction

A wealth of evidence implicates the primate medial frontal cortex, especially the anterior cingulate cortex (ACC), in reward-guided choice behavior (for a review, see Rushworth et al. 2004). Neurophysiological studies in nonhuman primates have suggested that the ACC is important for encoding action–outcome associations (Shima and Tanji 1998; Matsumoto et al. 2003; Amiez et al. 2006; Quilodran et al. 2008; Kennerley et al. 2009; Luk and Wallis 2009; Hayden and Platt 2010). For example, the activity of neurons in the ACC is related to the actions that will be made to obtain a reward (Shima and Tanji 1998; Matsumoto et al. 2003), as well as the amount or type of reward associated with that action (Amiez et al. 2006; Kennerley et al. 2009; Luk and Wallis 2009; Hayden and Platt 2010). In addition, dynamic changes in the firing rates of ACC neurons have been reported following actions that are rewarded or unrewarded during the learning of action–outcome associations (Matsumoto et al. 2007; Quilodran et al. 2008).

Neuropsychological studies have also implicated the ACC in reward-guided action selection. Monkeys with lesions of the medial frontal cortex that include the ACC are impaired in their ability to select actions instructed by presentation of different food rewards with which those actions are associated (Hadland et al. 2003). In addition, monkeys with lesions

of the cortex lining the banks of the anterior portion of the cingulate sulcus are impaired in using reinforcement history (i.e. positive feedback) to maintain a rewarded or correct action (Kennerley et al. 2006). In nearly all of these studies, the ACC was implicated specifically in guiding the choices of actions based on a reward value that was held constant throughout testing. It is not known whether the ACC is important for guiding choice behavior when the biological value of the rewards change.

In monkeys, the ability to update the value of outcomes (i.e. reward) in the face of changing biological needs is assessed using satiety-specific devaluation of a food reward. Previous work has shown that one part of the macaque prefrontal cortex, the orbital prefrontal cortex (PFO), is important for the ability to update the value of food rewards associated with objects (Baxter et al. 2000; Izquierdo et al. 2004; Machado and Bachevalier 2007). This and related data (described above) have led to the idea that the ACC and PFO may play selective roles in mediating action–outcome and object–outcome processes, respectively (Murray and Izquierdo 2007; Ostlund and Balleine 2007; Rushworth et al. 2007; Rudebeck et al. 2008; Camille et al. 2011). Although this hypothesis of selective roles for the ACC and PFO has held up so far, there are still some gaps in our knowledge, and the precise contribution of the ACC in reward-guided choice behavior is still unclear. First, no study to date has directly examined the role of primate ACC in guiding object choices when the value of the reward is changed, as has been done for the PFO. Second, no study has directly compared the role of the ACC in learning about the action choices based on the reward value with those based on changes in reward contingency, as measured by reversal learning. Accordingly, we investigated the role of ACC in mediating these 2 different influences—changes in the reward value and changes in reward contingency—on both object and action choice behavior.

To investigate whether the ACC is essential for using the updated biological value of rewards to guide choices, we tested monkeys with ACC lesions on 2 tasks: object-based and action-based reinforcer devaluation (RD; Experiments 1A and 1B, respectively). In the action-based task, monkeys were trained to perform 2 different actions (e.g. a turn or pull action with a joystick), each associated with one of the 2 reward outcomes (e.g. peanut or banana pellet). In intact animals, the devaluation of one of the reward outcomes would be expected to yield a selective reduction in the performance of its associated action relative to the other action. For comparison, we tested the same groups of monkeys on their ability to make choices guided by reward contingency in object- and action-based reversal learning tasks (Experiments

2A and 2B, respectively). All behavioral testing was carried out postoperatively. Notably, we intended to assess the effects of ACC lesions on acquisition of the object- and action-based reversal-learning tasks, as opposed to performance of well-learned reversals (e.g. Kennerly et al. 2006).

If the ACC is essential for processing or using information about the updated value, then monkeys with ACC lesions should be impaired on one or both of the RD tasks. In addition, if the ACC is essential for action choices but not object choices, as the evidence suggests, then monkeys with ACC lesions should be impaired on both types of action tasks but not the object-based tasks. A role for ACC beyond the action tasks would suggest a more general role for this region in representing or using reward information to sustain effective choice behavior.

Materials and Methods

Subjects

A total of 8 adult, male rhesus monkeys (*Macaca mulatta*), ranging in weight from 4.75 to 9.3 kg at the start of behavioral training, were used for this study. They were housed individually or in pairs in temperature controlled rooms (76–80 °F) under diurnal conditions (12 h light/dark cycle). All monkeys were fed a controlled diet of primate chow (catalogue number 5038, PMI Feeds Inc., St Louis, MO, United States of America) supplemented with fresh fruit or vegetables. Water was available ad libitum. Four monkeys received bilateral ablations of the ACC, and 4 monkeys served as unoperated controls. In addition, the same monkeys received various tests of emotional responsiveness, which were performed after object devaluation and before action reversal. The order and timing of testing for all monkeys in the experiment was identical to ensure that all animals shared the same training history. The experiments described below were carried out in the following order: Experiment 1A, 2A, 2B and 1B. Thus, the action-reversal task (Experiment 2B) was conducted prior to the action RD task (Experiment 1B). All procedures were approved by the NIMH Animal Care and Use Committee.

Surgery

All lesions were performed under aseptic conditions. Anesthesia was induced with ketamine hydrochloride (10 mg/kg i.m.) and maintained with isoflurane gas (1–3%, to effect). During surgery, monkeys received isotonic fluids via an intravenous drip, and heart and respiration rates, body temperature, blood pressure, and expired CO₂ were monitored throughout the procedure.

For the lesion, a midline incision was made and the skin and galea were retracted to expose the cranium. A bone flap (~4 cm²) was first taken over the dorsal cranium. The dura mater was then cut along the lateral edge of the bone opening and reflected toward the midline. With the aid of an operating microscope, the boundaries of the lesion were identified. Then, using a combination of suction and electrocautery, ACC was removed by subpial aspiration through a fine-gauge metal sucker, insulated except at the tip. The boundaries of the intended lesion extended from the rostral tip of the cingulate sulcus, rostrally, to an imaginary coronal plane through the spur of the arcuate sulcus, caudally. The ventral limit of the lesion was an imaginary line just above the dorsal aspect of the corpus callosum. The dorsal limit was the lip of the dorsal bank of the cingulate sulcus. Thus, the lesion included the cortex of the cingulate gyrus and the dorsal and ventral banks of the cingulate sulcus. The region approximates cytoarchitectonic areas 24a, 24b, 24c and the immediately adjacent ventral medial part of area 9 (Vogt 1993; Carmichael and Price 1994). Notably, the territory of the intended lesion includes the rostral cingulate motor area (Picard and Strick 1996).

At the completion of surgery, the wound was closed in anatomical layers with Vicryl sutures. Pre- and postoperatively, monkeys received dexamethasone sodium phosphate (0.4 mg/kg i.m.) and an antibiotic

(Cefazolin, 15 mg/kg i.m., or Di-Trim, 0.1 mL/kg, 24% w/v solution i.m.) to reduce inflammation and prevent infection, respectively. For 3 days after surgery, the monkeys also received a treatment regimen consisting of Ketoprofen (10–15 mg/kg, i.m.), acetaminophen (40 mg/kg i.m.), or Banamine (flunixin meglumine, 5 mg/kg i.m.). Monkeys received ibuprofen (100 mg) for 5 additional days.

Lesion Assessment

Lesions were assessed using postoperative T₁-weighted magnetic resonance (MR) imaging scans (1.5 Tesla magnet; fast-spoiled gradient; echo time, 5.8; repetition time, 13.1; flip angle 30; number of excitations, 8; 256 square matrix; field of view, 100 mm; 1 mm slices) which were performed an average of 3 months after surgery. Representative MR images from case ACC 2 are shown in Figure 1. In addition, to further document the location and extent of the ACC lesion, both coronal and sagittal views of postoperative T₁-weighted MR images for every operated monkey are provided in Supplementary Figure S1.

For the purpose of estimating the size of the lesion, MR scan slices for each monkey in the ACC group were matched to the drawings of coronal sections of a standard rhesus monkey brain at 1 mm intervals. The extent of the lesion visible in the MR images was then plotted onto the standard sections. Using a digitizer tablet (Wacom, Vancouver, WA, United States of America), we measured the volume of the spared ACC in each operated subject and compared it with the volume of the intended lesion plotted on the same standard sections. The volume of the lesion was then expressed as a percent of the total volume of the structure. For illustrative purposes, we also plotted the lesions onto standard Nissl-stained sections from a rhesus monkey (Fig. 1).

The lesions were essentially as intended. We estimated that the operated monkeys sustained on average 85.5% damage to the total volume of the ACC (see Table 1). Sparing, where it occurred, tended to be in the caudal portion of the intended removal. As expected, there was extensive overlap in the placement of the lesions in the 4 subjects (Fig. 1) and little inadvertent damage to adjacent structures.

Reinforcer Devaluation

Monkeys were tested on 2 versions of the RD task: object based and action based. In the object-based RD task (Experiment 1A), monkeys were required to associate *objects* with particular food rewards, and were then tested for their ability to adjust their choices of objects when the value of one of the foods was manipulated. The action-based RD task (Experiment 1B) was similar to the object-based version except that monkeys were required to associate *actions* with particular food rewards, and then were assessed for their ability to adjust their actions when a food value was changed. Unlike the object-based RD task, the action-based RD task did not require a choice between the 2 actions; rather, it involved a decision to perform an action or not.

Experiment 1A: Object-Based RD

Apparatus

Monkeys were tested in a Wisconsin General Test Apparatus (WGTA), which consists of a large monkey compartment that holds a monkey cage and a smaller test compartment that contains the test tray. The test tray measured 19.2 cm (width) × 72.7 cm (length) × 1.9 cm (height) and contained 2 food wells (diameter, 3.8 cm; depth, 0.6 cm) located 29 cm apart, center to center, on the midline of the tray. During test sessions, the test compartment was illuminated with two 60-W incandescent light bulbs and the monkey compartment was unlit. An opaque screen separated the monkey

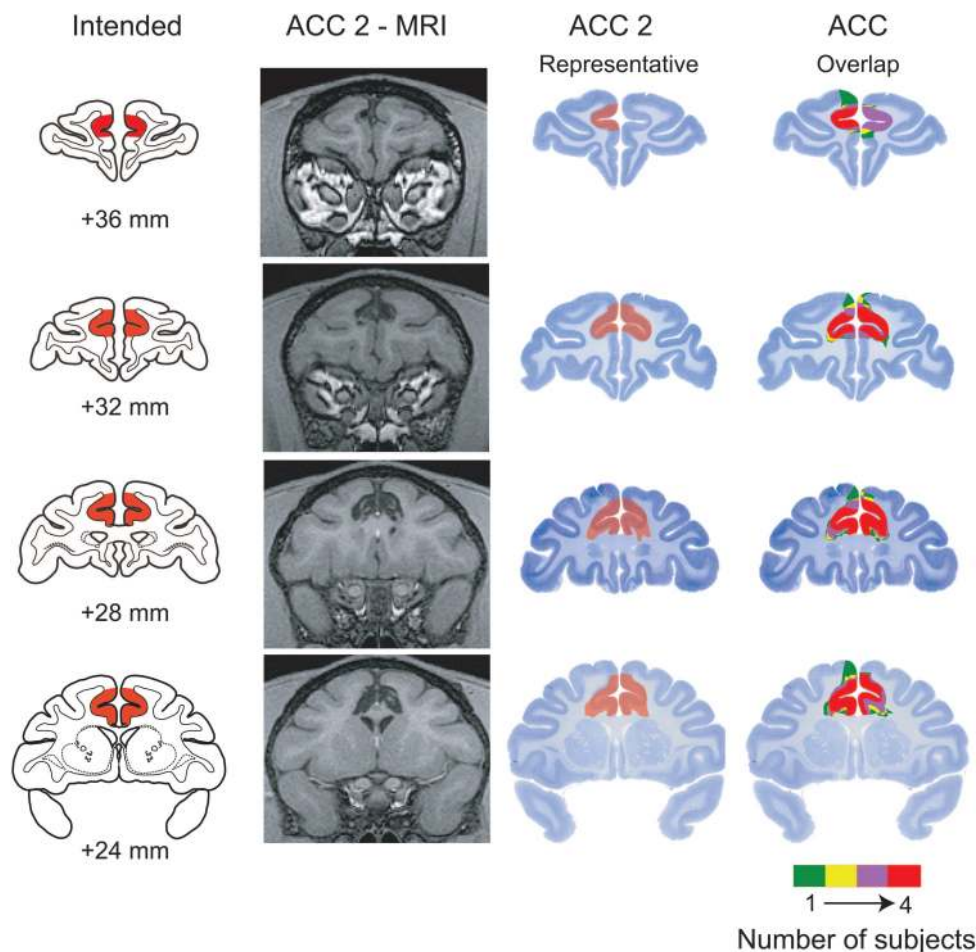


Figure 1. Location and extent of anterior cingulate cortex (ACC) lesion. The first column (far left) shows the intended lesion (shaded in red) drawn on coronal sections from a standard rhesus monkey brain. The second column provides postoperative T_1 -weighted MR images from case ACC 2 at matching levels. The third column shows the lesion in case ACC 2 plotted on Nissl-stained sections from a macaque brain, again at levels matching those in the first column. The fourth column provides an overlap of the lesions for the 4 subjects, plotted on top of each other. Colors show degree of overlap, as indicated in the legend. Numerals indicate distance in millimeters (mm) from the interaural plane.

Table 1
Estimated percent damage to the anterior cingulate cortex

Monkey	Estimated percent damage by volume		
	Left	Right	Mean
ACC 1	75.4	68.5	72.0
ACC 2	79.5	80.1	79.8
ACC 3	90.4	97.3	93.8
ACC 4	97.7	94.7	96.2

Mean, average of the values for the left and right hemispheres; Left, left hemisphere; Right, right hemisphere.

compartment from the test compartment. A separate screen, located between the experimenter and the test compartment, was fitted with a 1-way viewing screen. This screen allowed the experimenter to view the monkey's responses during the trial without being seen by the monkey. Testing was carried out in a darkened room.

We used 120 "junk" objects that varied in color, shape, and size. Several dark gray matboard plaques (7.6 cm on each side), and 3 junk objects were dedicated to pretraining. Each monkey was assigned 2 different foods (food 1 and food 2)

that were roughly equally palatable as determined from food preference tests (see Behavioral Procedure). The 2 foods were selected from the following 6: banana-flavored pellets (P.J. Noyes, Inc., Lancaster, NH, United States of America), half-peanuts, raisins, sweetened dried cranberries (Craisins, Ocean Spray, Lakeville-Middleboro, MA, United States of America), "fruit snacks" (Giant Food, Inc., Landover, MD, United States of America) or chocolate M&Ms (Mars Candies, Hackettstown, NJ, United States of America).

Behavioral Procedure

Pretraining

All behavioral testing took place after surgery. Before formal behavioral training, monkeys were introduced to the WGTA and allowed to take food freely from the test tray. Monkeys were then trained by successive approximation to displace plaques overlying the food wells to obtain a half peanut hidden underneath. The procedure was then repeated with 3 objects dedicated to this phase. Each monkey was required to complete a single session consisting of 10 plaque and 40 object trials. Each item was presented, singly, overlying a baited food well.

Food Preference

Monkeys were assessed for their preferences for 6 different foods. On each trial, the monkey was presented with 2 different foods, but was allowed to choose and eat only one of the foods. Trials were separated by 10 s. Fifteen individual food pairings appeared twice in each session. Hence, each food was encountered 10 times per 30-trial session. Monkeys were tested for a total of 15 days. The total number of choices of each food across the last 5 days of testing, when food preferences had stabilized, was tabulated. For each monkey, 2 foods that were approximately equally preferred were designated as *food 1* and *food 2*.

60-Pair Discrimination Learning

Monkeys learned to discriminate 60 trial-unique pairs of objects. For each pair, 1 object was baited with food (positive object) and the other was never baited (negative object). Half of the positive objects were assigned to be baited with food 1, and the other half were assigned to be baited with food 2. On each trial, the monkey was presented with a single pair of objects and was allowed to displace only 1 item. If the monkey displaced the positive object, it was allowed to retrieve the food hidden underneath it. If the monkey displaced the negative object, no reward was provided. A single session comprised 60 trials, 1 per pair, each separated by 20 s. The presentation order of the object pairs and the food reward assignments remained constant across sessions; the left–right position of the positive object followed a pseudorandom order. Criterion was set at a mean of 90% correct responses over 5 consecutive sessions (i.e. a minimum of 270 correct responses in 300 trials).

Reinforcer Devaluation: Test 1

Monkeys' choices of objects were assessed in 4 critical test sessions. In these sessions, only the positive (baited) objects were used and they were paired to produce a total of 30 pairs, each comprised of a *food-1* object and a *food-2* object. Each pair appeared once per session; each member of the pair covered the same food with which it had been associated in the discrimination learning phase. During critical test sessions, monkeys were allowed to displace one of the objects in each pair and to obtain the food reward. Objects were paired anew for each session. Two of the critical sessions were preceded by a selective satiation procedure, described below, intended to devalue one of the 2 foods. The other 2 critical test sessions were not preceded by selective satiation and provided baseline measures. At least 2 days of rest followed a session preceded by a selective satiation procedure. In addition, following the 2 days of rest, the monkeys were given 1 regular training session with the original set of 60 object pairs presented for choice as during original learning. This session was given to ensure that there was no long-lasting effect of selective satiation. On the rare occasion that a monkey did not score 54 correct responses or more, another session was given. Critical sessions were administered in the following order for each monkey: baseline session 1; session preceded by selective satiation with *food 1*; baseline session 2; session preceded by selective satiation with *food 2*.

The effect of RD was quantified as a "difference score," which was the change in choices of object type (*food-1* and *food-2* associated objects) in the sessions preceded by selective satiation relative to the mean of the baseline sessions.

The final difference score for each monkey was derived by summing the difference score from each of the 2 critical sessions preceded by selective satiation. The greater the shift in responses *away* from objects overlying the sated food, the higher the difference score.

Reinforcer Devaluation: Test 2

Approximately 1 month after RD test 1, the monkeys were retrained on the 60 pairs to the same criterion as before. After monkeys had reattained criterion, the RD procedure was repeated in the same manner as before.

Selective Satiation

A food box (8 × 10 × 7.5 cm) containing a known quantity of food 1 or food 2 was attached to the monkey's home cage. Each monkey was given 15 min, unobserved, to eat as much as it wanted. At the end of the 15 min, the food box was checked to see whether the monkey had eaten all of the food. If the box was empty, it was refilled. Thirty minutes after the food box had first been attached to the home cage, an experimenter started to observe the monkey's behavior. The selective satiation procedure was deemed to be complete when the monkey refrained from retrieving food from the box for 5 min. The monkey was then taken to the WGTA within 10 min and the test session conducted.

Results

60-Pair Discrimination Learning and Relearning

As shown in Figure 2A, the 2 groups of monkeys readily acquired the 60 discrimination problems. A repeated-measures ANOVA revealed a main effect of session, reflecting monkeys' increased accuracy across sessions [$F_{9, 54} = 26.62$; $P < 0.001$]. There was no main effect of group [$F_{1, 6} = 0.94$; $P > 0.05$] or interaction [$F_{9, 54} = 0.49$; $P > 0.05$], showing that the 2 groups did not differ in their rate of learning. Consistent with this finding, the 2 groups did not differ in the number of errors [mean (\pm SEM): CON 140.5 (± 25.6), ACC 198 (± 42.86); $t_{(6)} = 1.32$; $P > 0.05$], trials [mean (\pm SEM): CON 324.5 (± 62.3), ACC 507 (± 135.46); $t_{(6)} = 1.41$; $P > 0.05$], or sessions required to attain criterion [mean (\pm SEM): CON 7.75 (± 1.44), ACC 11.75 (± 2.95); $t_{(6)} = 1.40$; $P > 0.05$]. In addition, roughly 1 month later, there were no group differences in errors, trials, or sessions required to relearn the 60 pairs prior to RD Test 2 [all $P > 0.05$].

Reinforcer Devaluation: Tests 1 and 2

Figure 2B shows the mean difference scores for each group; higher scores reflect greater sensitivity to changes in the value of the food reward. As predicted, monkeys in the CON group tended to avoid choosing the objects overlying the devalued food. In both tests, the control monkeys showed the difference scores that were reliably higher than expected by chance [Test 1: $t_{(3)} = 7.24$; $P < 0.05$; Test 2: $t_{(3)} = 8.95$; $P < 0.05$]. The difference scores were further analyzed using an ANOVA with Test as the repeated-measures factor. As expected, there was a significant main effect of test, reflecting the tendency of monkeys to score higher on Test 2 relative to Test 1 [$F_{1, 6} = 6.95$; $P < 0.05$]. Despite the ACC lesion group exhibiting on average lower difference scores relative to controls, there was neither a significant main effect of group [$F_{1, 6} = 3.49$; $P > 0.05$] nor a significant group × test interaction [$F_{1, 6} = 1.53$;

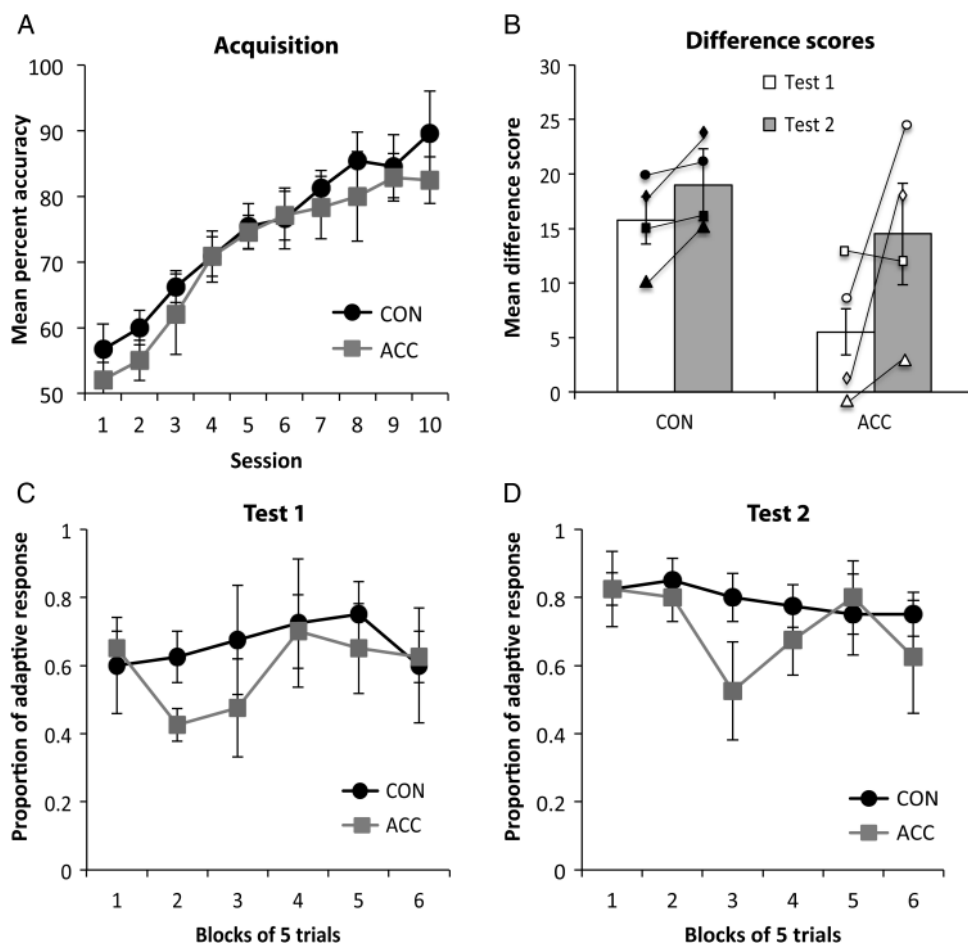


Figure 2. Performance on the object RD task. (A) Mean accuracy (\pm SEM) of controls (CON, filled circles) and anterior cingulate cortex (ACC, gray squares) during acquisition of the 60 object discrimination problems over 10 sessions; (B) The mean difference score (\pm SEM) of controls (CON) and anterior cingulate cortex (ACC) groups for Test 1 (white bars) and Test 2 (gray bars) of the object-based RD task. Symbols represent the scores of individual monkeys. Both groups of monkeys showed a normal enhancement of the devaluation effect in the second test relative to the first. CON group: filled circle, CON 1; filled square, CON 2; filled triangle, CON 3; filled diamond, CON 4. ACC group: open circle, ACC 1; open square, ACC 2; open triangle, ACC 3; open diamond, ACC 4; (C,D) trial-by-trial object choices on the critical test sessions after selective satiation for Test 1 and Test 2, respectively. CON, unoperated control monkeys; ACC, monkeys with lesions of the anterior cingulate cortex.

Table 2

Extent of lesion in each ACC monkey, and their respective scores for the object- and action-based devaluation tasks

Monkey	Lesion volume (%)	Object devaluation		Action devaluation	
		Difference score		Probability of response	
		Test 1	Test 2	Deval	Nondeval
ACC 1	72.0	9	25	0.56	0.67
ACC 2	79.8	13	12	0.58	0.67
ACC 3	93.8	-1	3	0.96	0.85
ACC 4	96.2	1	18	0.81	0.83
Mean	85.5	5.5	14.5	0.73	0.76
SEM	5.76	3.3	4.66	0.10	0.05

Deval, response associated with the devalued (sated) food; Nondeval, response associated with the nondevalued food.

$P > 0.05$], showing that the RD effects did not differ between the groups.

To further explore possible group differences, we also analyzed the proportion of adaptive responses on the critical test sessions. Figure 2C,D illustrates the proportion of adaptive responses after selective satiation for Test 1 and Test 2,

respectively. This measure, unlike the difference score, is independent of the baseline data. Each trial of the 30-trial session following selective satiation was scored systematically with 1 (one) or 0 (zero). A score of 1 indicates that the chosen object was associated with the higher valued food, whereas a score of 0 indicates that the chosen object was associated with the devalued food. Data for the 2 sessions (1 after devaluation of each food type) were averaged and then collapsed into 6 blocks of 5 trials each. The higher the score, the more adaptive the response. A repeated-measures comparison of the 6 five-trial blocks for each test revealed no main effect of group for Test 1 [$F_{1,6} = 0.268$; $P > 0.05$] or for Test 2 [$F_{1,6} = 0.492$; $P > 0.05$]. Thus, the proportion of adaptive responses, like the difference scores, show that the monkeys with ACC lesions were performing at a level indistinguishable from that of the controls on the object-based devaluation task.

Finally, given the variation in difference scores obtained by monkeys in the ACC group, we considered the extent of the lesion for each ACC monkey (% damage by volume) and their respective scores on the object-based devaluation task. As shown in Table 2, the monkey with the largest ACC lesion (ACC 4) received the second highest difference score on Test

2 and, in addition, showed the largest increase in score from Test 1 to Test 2. The monkey with the smallest lesion (ACC 1) also obtained a high difference score. Taken together, these findings argue against the possibility that the lesion size is systematically related to difference scores.

Food Consumption During Selective Satiation

During Test 1, the 2 groups ate equivalent amounts [means (\pm SEM): CON, 131.0 g (\pm 9.74); ACC, 133.0 g (\pm 11.0)]. During Test 2, monkeys with ACC lesions consumed less food on average than controls [means (\pm SEM): CON, 132.4 g (\pm 7.73); ACC, 114.6 g (\pm 9.92)]. However, a repeated-measures ANOVA confirmed that the amounts eaten during the satiation procedures did not differ significantly between groups [$F_{1,6} = 0.08$; $P > 0.05$], between tests [$F_{1,6} = 2.07$; $P > 0.05$], nor was there a group \times test interaction [$F_{1,6} = 4.63$; $P > 0.05$].

Experiment 1B: Action-Based RD

Apparatus

Monkeys were seated in a primate chair and tested in a custom-made sound-attenuating chamber (inside dimensions 108 \times 83 \times 51 cm; Crist Instrument Co., Hagerstown, MD, United States of America) that was dimly lit. A joystick (Measurement Systems, Inc., Norwalk, CT, United States of America) and a small speaker were mounted inside the chamber. The joystick was positioned 10 cm in front of the monkey and constrained to move in only 2 directions from a central origin. This configuration permitted only 2 movements: "pull" and "turn." The movements could only be performed using the right hand. A food cup located to the right of the joystick received food rewards, which were delivered from an automatic food dispenser mounted on the top of the chamber (BRS/LVE, Inc., Laurel, MD and Med Associates, St Albans, VT, United States of America). Two food types served as rewards: half peanuts and chocolate M&Ms (Mars Candies, Hackettstown, NJ, United States of America). In addition, banana-flavored pellets (P.J. Noyes, Inc., Lancaster, NH, United States of America) were used for pretraining purposes only. The task was controlled by CORTEX software (NIMH, <http://dally.nimh.nih.gov/>).

Behavioral Procedure

Pretraining

Each monkey was first shaped to move the joystick. Four monkeys, 2 from each group, learned to make the pull movement first followed by the turn movement. The remaining 4 monkeys were trained to make the turn movement first and then the pull movement. During this phase, the joystick was constrained such that only 1 movement could be made per session; the other movement was blocked. The start of each trial was signaled by a series of rapid clicks for 0.5 s. On each trial, to obtain a food reward, the monkey was required to move the joystick in the specified direction within 25 s. A successful action consisted of movement of the joystick from the center position to the fullest extent of either the pull or the turn movement. Completion of the turn or pull movement led to delivery of a single-food pellet that was concomitant with a tone and the initiation of 0.6 s intertrial interval (ITI). Additional movements made during the ITI were not

rewarded, and instead reset the ITI. A failure to make the specified movement within 25 s led to termination of the trial and initiation of the ITI. Monkeys were required to make 150 rewarded movements per day for 6 consecutive days after which they were trained to make the other movement to the same criterion.

Action–Outcome Learning

During action–outcome learning, each movement was paired with a different food (peanut or M&M), neither of which had been used for pretraining. The action–food pairings were counterbalanced across groups. We chose to train monkeys on a random ratio schedule for 2 reasons. First, it has been shown that behavior generated under such a schedule is more likely to remain goal directed, relative to other training schedules (Dickinson et al. 1983). Second, it made the response more robust to extinction (see below), which facilitated our evaluation of action–outcome associations uncontaminated by new learning (Ostlund and Balleine 2005).

As in pretraining, the start of each trial was signaled by a series of rapid clicks for 0.5 s and the joystick was constrained so that only 1 movement could be made per session. On each trial, the monkey was required to move the joystick in the specified direction within 25 s. A correct movement was rewarded with one of the foods delivered according to a probability, P . Monkeys were first trained on a schedule that delivered reward at $P = 0.75$ and then moved successively to a reward schedule of $P = 0.50$, $P = 0.35$ and finally, $P = 0.25$. The criterion for advancing to the next reward schedule was set at 20 rewards per session for 3 consecutive days. Reward delivery was concomitant with a tone and the initiation of a 2-s ITI. Failure to respond within 25 s led to termination of the trial and initiation of the ITI. A session ended after 20 rewarded responses had been completed, 30 min had passed or 5 consecutive trials were terminated, whichever came first. One turn and one pull action sessions were run per day. These sessions were run back-to-back, and the order in which they were run was alternated daily.

Action Devaluation: Extinction Test

Monkeys were assessed in 4 critical test sessions each preceded by a selective satiation procedure (see below). As was the case for the action–outcome learning stage, pairs of action–food test sessions were run back-to-back. The pull action was assessed in critical test sessions 1 and 4, and the turn action was assessed in critical test sessions 2 and 3. Each test session was conducted under extinction (i.e. responses were not rewarded) and comprised 12 trials. Each pair of test sessions was preceded by a selective satiation procedure (see below) and was followed by 2 days of rest. In addition, following 2 days of rest, the monkeys were retrained on the original action–outcome associations at $P = 0.25$ until they performed for 2 consecutive days at criterion. The effect of RD was quantified as probability of response (i.e. whether the monkey responded or not on a trial), and latency to respond across all 4 critical test sessions. If monkeys failed to respond within the 25-s trial limit, a latency of 25 s was scored.

Selective Satiation

The same selective satiation procedure used for the object-based RD experiment was used for the action-based devaluation experiment.

Note on Task Design

It would have been preferable to employ analogous versions of object- and action-based devaluations tasks. Indeed, we initially trained monkeys on a version of an action devaluation task that was formally analogous to the object devaluation task. When given a choice between actions, however, monkeys often developed a strong preference for 1 action. Despite considerable time and effort spent attempting to train monkeys to perform 2 different actions for different food rewards in a choice format, we were unable to do so. Ultimately, we decided to employ a different design, the one described here in which monkeys are allowed to perform only 1 action at a time. Although the monkeys did not get to choose between the 2 trained actions, they did get to choose whether to respond or not. If the actions are associated with particular food rewards, and if monkeys are truly satiated on 1 reward, then intact monkeys should in theory reduce their responding for the satiated food. We note that this is the approach often used in experiments with rodents (e.g. Ostlund and Balleine 2005).

Results

Action–Outcome Learning

Monkeys in both groups readily acquired the pull and turn movements for the 2 food outcomes. The number of sessions required to attain criterion on the $P=0.25$ reward schedule

did not differ significantly between the groups [$F_{1, 6}=0.00$; $P>0.05$; mean session (\pm SEM): Con = 2.5 (± 0.5); ACC = 2.5 (± 0.5)].

Action Devaluation: Extinction Test

After the selective satiation procedure, the performance of the 2 actions was assessed in extinction. Figure 3 shows the mean performance of the devalued and nondevalued actions for ACC and CON groups collapsed across the extinction tests. Data are plotted in blocks of 3 trials. A $2 \times 2 \times 4$ (group \times action \times block) repeated-measures ANOVA conducted on the latency data produced a main effect of block [$F_{3, 18}=11.10$; $P<0.01$] confirming that both groups took longer to produce an action with increasing numbers of trials, as expected under extinction conditions (Fig. 3A,B). There was also a main effect of action [$F_{3, 18}=7.17$; $P<0.05$]; monkeys displayed shorter latencies to perform actions that were associated with the higher valued food. In addition, an action \times block interaction [$F_{3, 18}=4.64$; $P<0.05$] confirmed that this devaluation effect increased in size as the session progressed. However, there were no main effects or interactions involving group [all $F<1$, $P>0.05$], indicating that the ACC lesion did not disrupt performance on this task.

A similar ANOVA conducted on the response probability data also produced a main effect of block [$F_{3, 18}=5.75$; $P<0.01$], indicating that the performance of the actions in

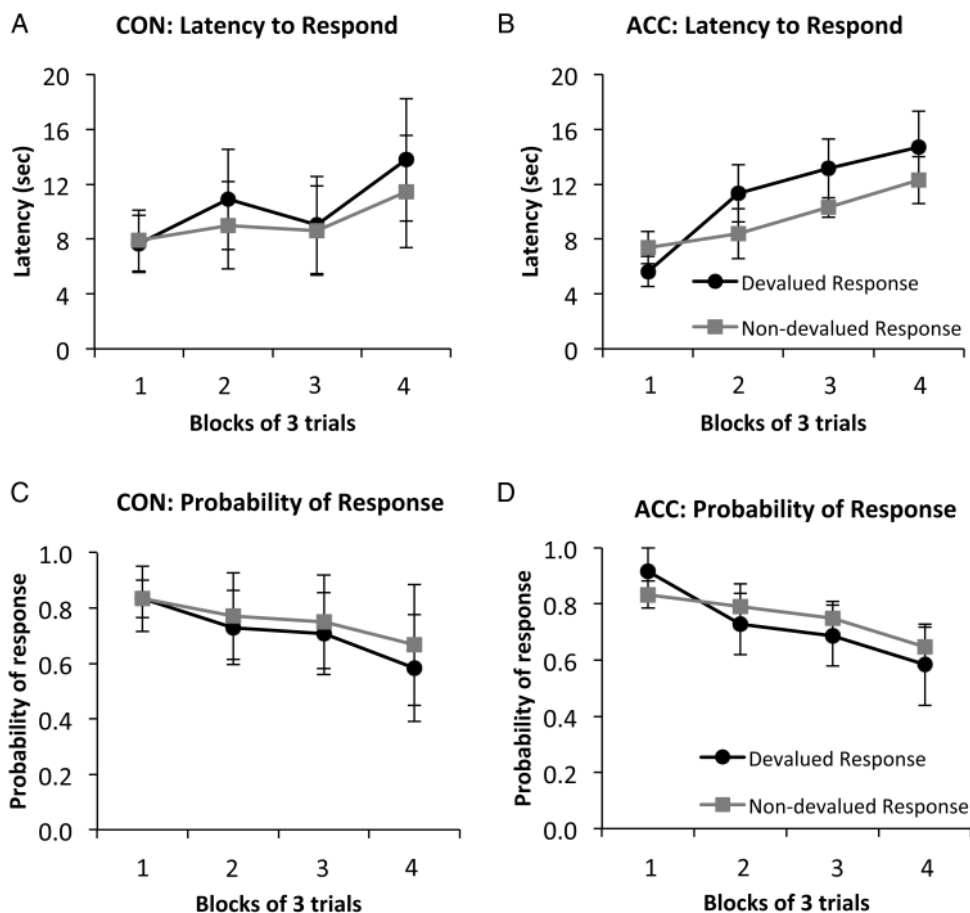


Figure 3. Mean latency (\pm SEM) to respond (A,B), and mean probability (\pm SEM) of response (C,D) collapsed across all 4 critical sessions in extinction in the action-based RD task. Following selective satiation, monkeys were evaluated for their performance of both the response associated with the devalued food (devalued response, black circles) and the response associated with the nondevalued food (nondevalued response, gray squares). CON, unoperated control monkeys; ACC, monkeys with lesions of the anterior cingulate cortex.

both groups decreased across the test session. Although there was a trend toward a main effect of action, which would indicate an effect of RD (see Fig. 3C,D), the groups did not differ significantly in terms of probability of responding [all $F < 1$, $P > 0.05$]. Inspection of the scores in Table 2 further confirms a lack of relationship between the lesion size and performance on the action-based devaluation task.

Experiment 2: Reversal Learning

For comparison with the effects of ACC lesions on the object- and action-based RD, the same monkeys with ACC lesions and their controls were tested on object reversal learning (Experiment 2A) and action reversal learning (Experiment 2B). These tasks required monkeys to associate one of the two *objects* or one of the two *actions* with a food reward, and then to respond flexibly to changes in reward contingency.

Experiment 2A: Object Reversal Learning

Apparatus

Monkeys were trained in the same WGTA used in Experiment 1A. Two objects, novel at the beginning of testing, were used throughout. A half-peanut served as the food reward.

Behavioral Procedure

Monkeys were trained on a single-object discrimination problem and its reversal. On the very first trial of initial learning, both objects were either baited or unbaited (counterbalanced within group). The object that was chosen was designated as the correct stimulus if it had been baited, or the incorrect stimulus if it had been unbaited. This procedure was used to prevent any systematic bias in scores that might have been caused by object preferences. On each trial thereafter, the 2 objects were presented, 1 baited and 1 unbaited, each overlying the 2 food wells. The monkey was allowed to displace one of the objects and, if correct (i.e., baited), to retrieve the food reward hidden underneath. The ITI was 10 s and the left and right positions of the correct object were determined pseudorandomly. A session comprised 30 trials. Criterion was set at 28 correct responses in 30 trials (93%) on 1 day, followed by at least 24 correct in 30 trials (80%) the next day. After monkeys had attained criterion, the stimulus-reward contingencies were reversed. That is, the object that was initially baited became the incorrect, unbaited object, and *vice versa*. The monkeys were then trained to the same criterion as in initial learning. This procedure was repeated until a series of 7 serial reversals had been completed.

Results

The number of errors scored during acquisition of the initial object discrimination and the subsequent 7 reversals is illustrated in Figure 4A. Both groups readily acquired the initial object discrimination [mean errors to criterion (\pm SEM): ACC 3.5 (\pm 1.3); CON 8.5 (\pm 3.2); $t_{(6)} = 1.41$, $P > 0.05$]. During object reversal learning, both groups made fewer errors with the increasing number of reversal [$F_{6, 36} = 8.39$; $P < 0.001$]. Although the ACC group showed a propensity to make more errors on specific reversals relative to the CON group, yielding a marginally significant group \times reversal interaction [$F_{6, 36} = 2.19$;

$P = 0.06$], the groups did not differ from each other [$F_{1, 6} = 2.23$; $P > 0.05$].

To further explore the trend for a significant interaction between group \times reversal, we examined the number of errors scored before and after the first correct choice following a reversal. During reversal, a preponderance of errors made immediately after a reversal but *before* a correct choice might indicate that a monkey is unable to use negative feedback (i.e. unrewarded choices) to switch their choice to the rewarded option. In contrast, a preponderance of errors made *after* a correct choice would signify an inability to use positive feedback (i.e. rewarded choices) to guide subsequent responses. To control for differences in the number of rewards earned across monkeys, only trials before monkeys had earned 52 rewards after each reversal were included in this analysis. This was the minimum number of rewards earned per reversal. Figure 4B (Total errors) shows the ACC lesion group scored more errors than controls. A 2×2 (group \times error type) repeated-measures ANOVA revealed a significant group \times error type interaction [$F_{1, 6} = 18.5$; $P < 0.05$]. Accordingly, we also conducted separate ANOVAs on the mean number of errors for each error type. This analysis confirmed that, relative to controls, monkeys with ACC lesions committed more errors *after* the first correct response [$F_{1, 6} = 11.2$; $P < 0.05$], suggesting a difficulty in sustaining the choice of the correct object. Errors made *before* the first correct response did not differ significantly between groups [$F_{1, 6} = 1.79$; $P > 0.05$].

Next, to explore the possibility that monkeys with ACC lesions were impaired in using feedback from either the rewarded or unrewarded trials, we conducted a series of trial-by-trial analyses to probe the way in which monkeys used rewarded and unrewarded trials to guide their subsequent choices (Kennerley et al. 2006; Rudebeck and Murray 2008). Specifically, we analyzed monkey's choices following rewarded (Correct + 1) or unrewarded (Error + 1) trials. As shown in Figure 5A, monkeys with ACC lesions exhibited a different pattern of behavior to controls, being slightly less likely to select the rewarded object following a correct response (Correct + 1) and more likely to select the rewarded object following an error (Error + 1). A 2×2 (group \times feedback) repeated-measures ANOVA revealed a significant effect of feedback as all monkeys were more likely to select the rewarded object following a previously rewarded choice as opposed to an unrewarded choice [$F_{1, 6} = 203.28$; $P < 0.05$]. There was also a significant group by feedback interaction [$F_{1, 6} = 11.29$; $P < 0.05$], but no effect of group [$F_{1, 6} = 3.11$; $P > 0.05$] indicating that rewarded and unrewarded trials were differentially influencing the choice behavior of monkeys with ACC lesions. Further analysis confirmed this difference (1-way ANOVA, correct + 1: [$F_{1, 6} = 11.63$; $P < 0.05$]; 1-way ANOVA, error + 1: [$F_{1, 6} = 6.52$; $P < 0.05$]).

Two additional analyses were conducted. The first analysis assessed each monkey's choices after different numbers of consecutive correct choices following an initial error ("EC" analysis). For example, monkeys might make 1 (EC), 2 [EC (2)], or 3 [EC(3)] correct choices after an error. Only trial types where there were at least 10 instances for each monkey were included. Based on this criterion, trial types EC + 1 – EC (9) + 1 were analyzed. The "EC" analysis revealed that the performance of the ACC lesion group was slightly below that of the controls following different numbers of consecutive correct trials following an error (Fig. 5B). A 2×9 (group \times trial type) repeated-measures ANOVA revealed a significant main

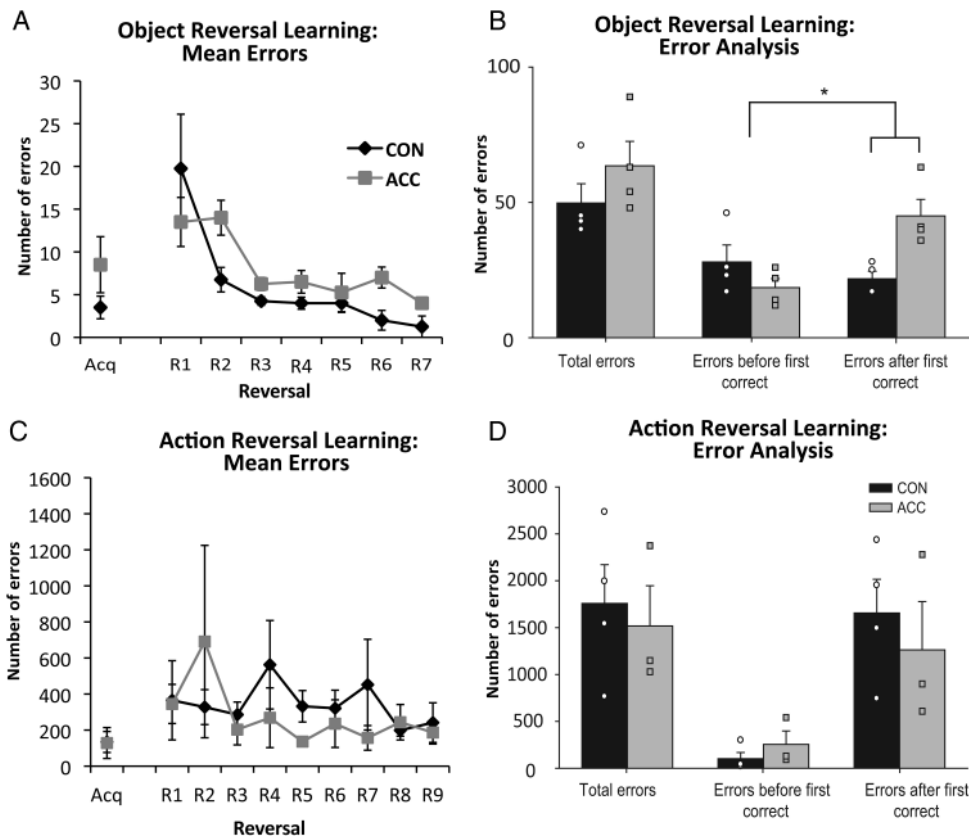


Figure 4. Performance on the object-based (A,B) and action-based (C,D) reversal learning tasks. (A,C) The mean number of errors (\pm SEM) committed during acquisition (Acq) of the initial discrimination and subsequent reversals. (B,D) The total number of errors made by CON and ACC groups during reversal, as well as the number of errors committed immediately after the reversal but before a correct response (errors before first correct) relative to errors made after a correct response (errors after first correct). Symbols represent the scores of individual monkeys. CON, unoperated control monkeys; ACC, monkeys with lesions of the anterior cingulate cortex.

effect of trial type [$F_{8,48} = 2.3$; $P < 0.05$], but no trial type by group interaction [$F_{8,48} = 1.04$; $P > 0.4$]. There was also a significant effect of group [$F_{1,6} = 7.84$; $P < 0.05$]. The second analysis determined each monkey's likelihood of selecting the correct object following consecutive unrewarded choices (errors) following an initial error ("EE" analysis). Only trial types with at least 10 instances for each monkey were included. Only trial types E+1 and EE+1 met this criterion. Figure 5C shows that monkeys with ACC lesions were more likely than controls to select the rewarded object following an error. A 2×2 (group \times trial type) repeated-measures ANOVA revealed a significant main effect of trial type [$F_{1,6} = 53.86$; $P < 0.001$], and of group [$F_{1,6} = 6.48$; $P < 0.05$], but no group by trial type interaction [$F_{1,6} = 0.22$; $P > 0.5$].

Experiment 2B: Action Reversal Learning

Apparatus

Monkeys were tested in the same apparatus used in Experiment 1B. Banana-flavored pellets (P.J. Noyes, Inc., Lancaster, NH, United States of America) served as the food reward throughout the experiment.

Behavioral Procedure

As was the case for the object reversal-learning task, the action reversal-learning task comprised an acquisition phase and a reversal phase. The 2 actions used in this experiment

were pull and turn, which were shaped during pretraining according to the protocol provided in Experiment 1B. It is important to note that the action reversal task was conducted prior to the action devaluation task (Experiment 1B). Consequently, the action reversal task was the first time any of the monkeys had experience performing different actions by manipulating a joystick.

During the acquisition phase, the joystick was unconstrained and the monkeys were free to select between the 2 actions. The trial onset was indicated by a series of rapid clicks presented for 0.5 s. The performance of the correct action led to the delivery of the food reward, a tone, and termination of the trial. The performance of the incorrect action terminated the trial. A new trial was initiated after a 0.6-s ITI. Failure to make the specified movement within 25 s led to termination of the trial and initiation of the ITI. The action that was rewarded in initial learning was counterbalanced within groups. A session ended after 150 rewarded trials had been completed, 30 min had passed, or 5 consecutive trials had been terminated due to response omission (i.e. failure to respond). Criterion was set at a minimum of 24 correct responses in the last 30 trials (80%), or 27 correct responses in 30 trials at any point within the session, for 2 consecutive days. After monkeys attained criterion, the action-reward contingencies were reversed. That is, the action that had initially led to reward was no longer rewarded, and *vice versa*. The criterion for each reversal was the same as for acquisition. A correction procedure was instituted if monkeys failed to select

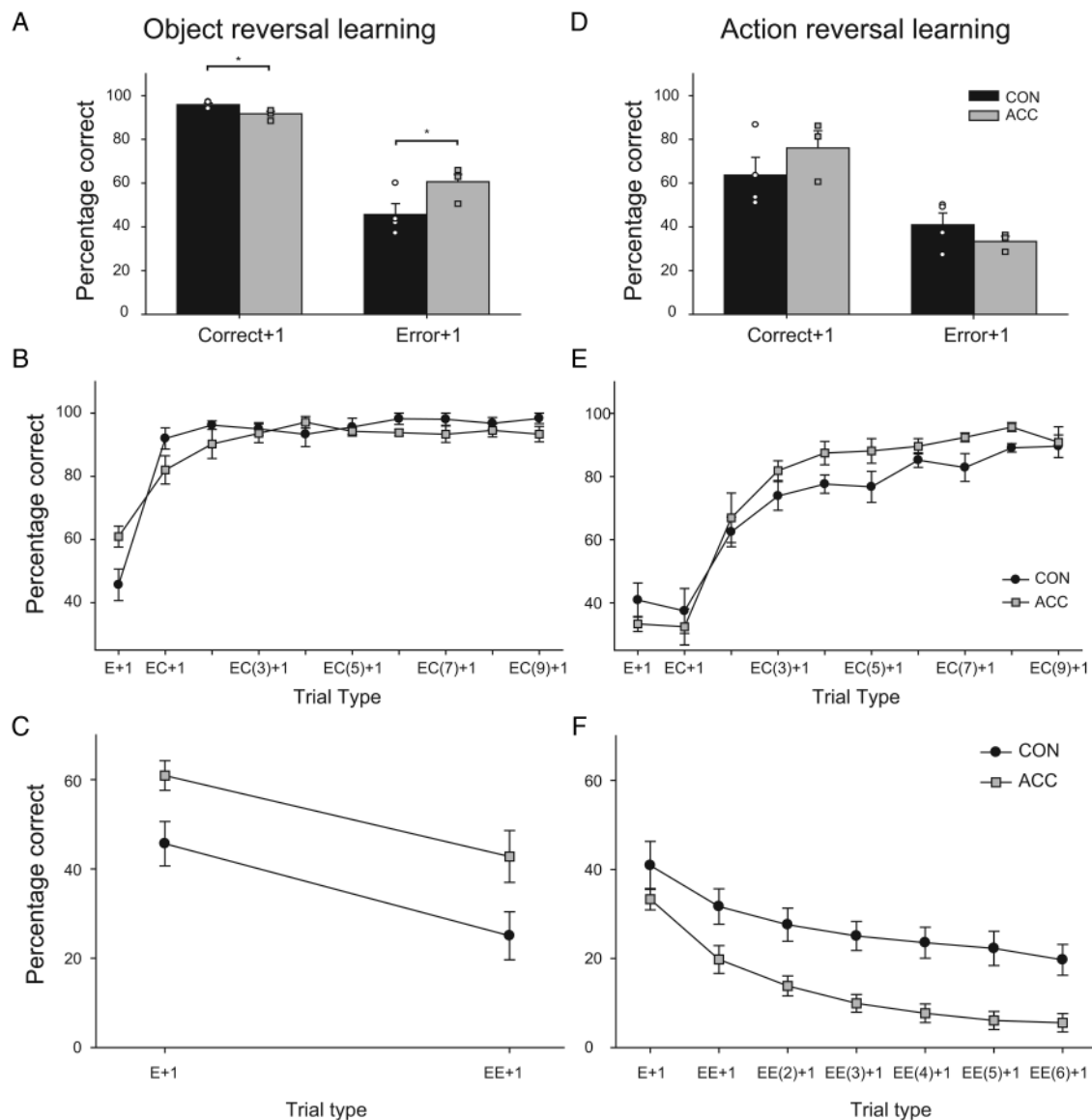


Figure 5. Top 2 panels show mean (\pm SEM) percent correct responses on trials immediately after a correct response (Correct + 1) or an incorrect response (Error + 1) for object reversal learning (A) and action reversal learning (D). Middle 2 panels show mean (\pm SEM) performance for sustaining the rewarded action after an error ("EC" analysis) for object reversal learning (B) and action reversal learning (E). The first trial type (E + 1) corresponds to the performance immediately after a single error. The second trial type (EC + 1) corresponds to performance immediately after a single error followed by a correct response. Similarly, the third trial type (EC(2) + 1) corresponds to performance immediately after a single error followed by 2 serially correct responses, and so on. Bottom 2 panels show results of "EE" analysis. The mean (\pm SEM) performance corresponds to switching to a correct response following one or a series of consecutive errors (E + 1, EE + 1, EE(2) + 1, etc.) for object reversal learning (C) and action reversal learning (D). CON, unoperated control monkeys; ACC, monkeys with bilateral lesions of the anterior cingulate cortex.

the new rewarded action during reversal learning for 4 consecutive days (i.e. never selected the correct action in 4 sessions). This ensured that the act of manipulating the joystick did not behaviorally extinguish. During the correction procedure, the joystick was constrained to the rewarded action for a single 150-trial session. The following day, monkeys were presented with an unconstrained joystick and were free to select between the 2 actions. Monkeys completed a total of 9 serial reversals.

Results

Both groups successfully acquired the initial action-reward association [mean sessions (\pm SEM): CON 2.75 (\pm 0.85); ACC 4.75 (\pm 2.59); $t_{(6)} = 0.73$; $P > 0.05$; mean total errors (\pm SEM):

CON 134 (\pm 58.16); ACC 128.25 (\pm 74.14); $t_{(6)} = 0.06$; $P > 0.05$]. When the action-reward contingencies were reversed, 1 monkey in the ACC group (ACC 1) failed to acquire the first reversal. The remaining 3 monkeys with ACC lesions required an average of 11 (\pm 2.30) sessions to acquire the first reversal. Monkey ACC 1 failed to reach criterion in 71 sessions. In addition, ACC 1 received 29 sessions of correction training (only 1 action available; the other blocked) interspersed with the regular sessions. Even with the aid of the correction training, ACC 1 made more errors during the first reversal relative to the other 3 monkeys [mean total error for reversal 1: ACC 1, 1808; ACC 2, 212; ACC 3, 560; ACC 4, 263]. Because this monkey could not complete the first reversal, even with the benefit of a large amount of correction training, we terminated the action-based reversal training for this monkey.

Accordingly, there are no reversal data for this monkey to analyze in the ANOVA examining the scores of the 2 groups (ACC lesion, controls); statistical analysis was conducted with an ACC group consisting of $n = 3$. Nevertheless, if the total errors and sessions scored by ACC 1 (not including the errors and sessions in correction training) are included in an analysis of the first reversal, the groups did not differ significantly from each other in terms of errors [$t_{(6)} = 0.79$; $P > 0.05$] or sessions [$t_{(6)} = 1.46$; $P > 0.05$].

Figure 4C shows the mean number of errors committed during acquisition of the action-reward association and during the subsequent 9 reversals. The groups did not differ in the number of errors scored during acquisition of the initial action-reward association [$t_{(5)} = 0.37$; $P > 0.05$]. A 2×9 (group \times reversal) repeated-measures ANOVA on errors to criterion revealed no effect of group [$F_{5,40} = 0.24$; $P > 0.05$], no effect of reversal [$F_{5,40} = 0.84$; $P > 0.05$], and no group \times reversal interaction [$F_{5,40} = 0.83$; $P > 0.05$].

For comparison with the object reversal task and also because previous studies have shown that, unlike controls, monkeys with ACC lesions are unable to sustain a rewarded action following a reversal (Kennerley et al. 2006), a trial-by-trial analysis was also conducted on the action reversal-learning data. To control for differences in the number of rewards earned between the different monkeys, we included only the trials before monkeys had earned 300 rewards after each reversal. This is the minimum number of rewards per reversal. As before, the total errors committed after a reversal were divided into those that occurred before the first correct (rewarded) choice following a reversal, and those that occurred after the first correct choice. This analysis revealed that monkeys in both CON and ACC groups performed similarly in that they made fewer errors immediately after a reversal but before a rewarded choice, as opposed to after a rewarded choice (Fig. 4D). A 2×2 (group \times error type) repeated-measures ANOVA showed an effect of error type [$F_{1,5} = 16.08$; $P < 0.05$], but not a group by error type interaction [$F_{1,5} = 0.73$; $P > 0.05$] nor an effect of group [$F_{1,5} = 0.158$; $P > 0.05$].

Despite this negative finding observed for the ACC lesion group in the action reversal task, we further examined the monkey's choices following single rewarded (Correct + 1) and unrewarded trials (Error + 1) as well as choices after multiple rewarded or unrewarded trials ("EC" and "EE" analyses). Following either rewarded or unrewarded trials, monkeys with ACC lesions and unoperated controls performed similarly (see Fig. 5D). Monkeys in both groups were more likely to select the rewarded action again following a previously rewarded action as opposed to following an unrewarded action. A 2×2 (group \times feedback type) repeated-measures ANOVA revealed a significant effect of feedback [$F_{1,5} = 16.155$; $P < 0.05$] but no group by feedback interaction [$F_{1,5} = 1.5$; $P > 0.05$] nor an effect of group [$F_{1,5} = 0.24$; $P > 0.05$]. For the "EC" analysis, trial types EC + 1 – EC(9) + 1 were included as these trial types had at least 30 instances per monkey. A 2×9 (group \times trial type) repeated-measures ANOVA failed to show any differences between the 2 groups (group-by-trial type interaction, [$F_{8,40} = 1.29$; $P > 0.05$]; effect of group, [$F_{1,5} = 1.54$; $P > 0.05$]). As shown in Figure 5E, however, both groups were more likely to select the rewarded option after longer sequences of choosing the rewarded action (effect of "EC" trial type, [$F_{8,40} = 63.62$; $P < 0.05$]). For the "EE" analysis trial types, E + 1

to EE(6) + 1 were included as these trial types had at least 30 instances per monkey. A 2×7 (group \times trial type) repeated-measures ANOVA confirmed a significant main effect of trial type (effect of "EE" trial type [$F_{6,30} = 73.9$; $P < 0.05$]); monkeys in both groups were less likely to select the correct action following increasing numbers of consecutive errors (Fig. 5F). Unlike the case for the EC analysis, however, the EE analysis revealed a main effect of group [$F_{1,5} = 8.28$; $P < 0.05$], due to the greater tendency of monkeys in the ACC group, relative to controls, to make incorrect choices after strings of errors. There was no group by trial type interaction ([$F_{6,30} = 0.28$; $P > 0.05$]). Thus, there was some evidence that the ACC group differed from controls in their use of error feedback in action-based reversal learning, but the direction of the effect was opposite to that observed for object-based reversal learning (compare and contrast Fig. 5C,F).

Discussion

The present study examined the role of the ACC in learning to select between objects and between actions when those choices were guided by changes in reward value or reward contingency. Monkeys with bilateral lesions of the ACC differed from controls in their ability to use reward information to guide choices during object reversal learning, and in their ability to switch their response after making strings of errors in action reversal learning. No group differences emerged in the object- and action-based devaluation tasks. These results have important implications for our understanding of the way in which goal-directed choice behavior is influenced by the ACC.

Choices Based on Object and Action Values

Updating the Value of Object–Outcome Associations

We found that when monkeys with ACC lesions were required to choose objects on the basis of the current, associated food value, they performed just as well as controls. Monkeys with ACC lesions showed an intact ability to update reward information (through selective satiation) and shifted their object choices after changes in the value of the associated food reward. Their behavior in these respects was indistinguishable from that of unoperated controls. Nevertheless, the ACC group scored lower than controls on average on Test 1 (Fig. 2B), a pattern not unlike that observed after lesions of medial PFO area 14 (Rudebeck and Murray 2011). Although speculative, 1 possibility is that the object-based RD task is supported by both Pavlovian and instrumental mechanisms, and that the ACC lesion disrupts 1 type of learning but not the other. To evaluate this idea, future studies should assess the effects of changes in outcome value using tasks that more specifically target Pavlovian and instrumental learning.

This lack of effect of ACC lesions is in marked contrast to the effect of PFO lesions, which yield significantly reduced RD effects when assessed with the same methods (Izquierdo et al. 2004; Machado and Bachevalier 2007; Rudebeck and Murray 2011; see also West et al. 2011). In addition, PFO must interact with the amygdala in order to update the expected value of reward outcomes (Baxter et al. 2000). Thus, although the PFO and ACC share access to information about reward via their interactions with the ventral striatum, amygdala, and the midbrain dopaminergic system (Barbas and De Olmos 1990;

Williams and Goldman-Rakic 1993; Ferry et al. 2000; Ghashghaei and Barbas 2002; Morecraft et al. 2007), it is only the PFO, via its interactions with the amygdala, that is essential for using updated value to guide object choices.

There are differences in the strength of motor, hippocampal, and sensory inputs to the ACC and PFO. For example, PFO, especially the lateral portion of PFO (including areas 13m, 13l, and 11l of Carmichael and Price 1994) that is essential for RD effects (Rudebeck and Murray 2011) receives inputs from almost all the sensory modalities. Although auditory sensory inputs are scarce, PFO receives multimodal inputs from the perirhinal cortex as well as inputs from visual, gustatory, olfactory, and somatic sensory cortex (Suzuki and Amaral 1994; Kondo et al. 2005; Saleem et al. 2008). In contrast, the ACC receives comparatively few projections from the foregoing sources and instead is more heavily interconnected with the parahippocampal cortex, entorhinal cortex, dorsal bank of the superior temporal sulcus, and hippocampal formation (Barbas and Blatt 1995; Barbas et al. 1999; Kondo et al. 2005; Saleem et al. 2008). The ACC is also more closely related to motor and premotor cortex than is PFO (Carmichael and Price 1995; Luppino et al. 2003; Morecraft et al. 2007). This difference in the pattern of connectivity of the PFO and ACC suggests that these regions operate within functionally distinct networks, and is consistent with a role for PFO but not ACC in guiding object choices based on the current biological value of reward.

Updating the Value of Action–Outcome Associations

Our assessment of action-value associations, implemented with a joystick task, provided evidence that monkeys altered their actions in response to changes in the value of the expected outcome. Although RD—achieved via selective satiation—produced significant increases in latency to produce the action associated with a devalued food, it had no effect on the probability of completing that action. Still, our action-value task should have been sufficient to detect effects of the ACC lesion, had there been any, on linking actions with food value, albeit not in the context of choosing directly between 2 actions.

Our finding of the lack of effect of ACC lesions on a task that requires linking actions with the current biological value of specific rewards (i.e. action-based RD) is revealing. Experimental approaches using lesions, inactivations, and cell recordings emphasize a fundamental role for the ACC in linking actions with reward to guide voluntary choice behavior. This includes using past reinforcement experience for the purpose of adjusting subsequent actions (Matsumoto et al. 2003; Kennerley et al. 2006; Rudebeck et al. 2008). Consistent with this view, Hadland et al. (2003) reported that monkeys with lesions of ACC (as defined here), together with areas 32 and 25, were impaired in using the identity of specific rewards to guide their actions. Taken together, these results suggest that the ACC may be important for representing the contingency between actions and their outcomes as opposed to the value of those outcomes. This issue is discussed further in the Linking Actions with Reward Contingencies section.

Our negative result should be interpreted with caution, however, as it is difficult to rule out the possibility that monkeys in the ACC group sustained inadvertent damage to fibers in the cingulum bundle. In rodents, it has been found that cingulum transection can either exaggerate or mask the

effect of cingulate cortex lesions (Neave et al. 1996; Meunier and Destrade 1997). In addition, it remains to be seen how medial frontal regions other than the ACC contribute to representations of the action value. Preliminary data indicate that monkeys with lesions of prelimbic cortex (area 32), a region adjacent to ACC, are impaired in using reward value to shift their choice of action (Rhodes and Murray 2010). Future studies will need to examine the effects of selective, circumscribed lesions of ACC and other medial frontal cortical fields on action-value decision-making. Finally, although we initially intended to obtain a measure of action-based RD that involved a choice between 2 actions, this was not possible in the current study. As already noted, whereas our object-based RD task involved a choice between objects, our action-based RD task did not. Accordingly, the measures we employed to assess object-value and action-value associations are not directly comparable. Future studies should aim to develop more robust measures of action-value associations, preferably ones involving a choice between actions (e.g. Rhodes and Murray 2010).

Choices Based on Object and Action Reward Contingencies

Linking Objects with Reward Contingencies

Although monkeys with ACC lesions were not impaired on the object RD task, a specific impairment emerged in the ACC group when the reward contingencies were reversed in the object reversal-learning task (Experiment 2A). Whereas monkeys with ACC lesions readily reversed (or switched) their responses following an error, they made more errors following an initial correct response immediately after a reversal. This finding indicates that the monkeys with ACC lesions did not have a problem with reversing their responses, per se, but instead failed to benefit from correctly performed trials to the same extent as controls, and, therefore, failed to maintain a successful response strategy. This pattern of results matches precisely that reported by Kennerley et al. (2006) for the effects of ACC lesions on performance of well-learned action reversals.

Complementing our findings, it has recently been reported that the same monkeys with lesions of the ACC sulcus cortex studied by Kennerley et al. (2006) exhibited a subtle deficit in using reward information following an initial *error* to guide their subsequent choices during object reversal learning (Walton, Rudebeck et al. 2010). In addition, consistent with the impairment observed in object reversal learning, inactivating the ACC does not induce a general impairment in all aspects of stimulus selection but specifically disrupts reward-guided stimulus learning (Amiez et al. 2006). Thus, it appears that the ACC makes contributions to both object and action reward-guided choice behavior, especially during learning (see also Alexander and Brown 2011).

An inability to sustain the correct choice of object following a reversal has also been observed after damage to PFO (Rudebeck and Murray 2008; Rudebeck et al. 2008). Notably, the effects of PFO and ACC lesions on object reversal learning differ both qualitatively and quantitatively. Whereas monkeys with PFO lesions have difficulty in using positive (but not negative) feedback to guide subsequent choices (Rudebeck and Murray 2008), monkeys with ACC lesions show alterations in using both positive and negative feedback (Fig. 5A,C). In addition, the disruption in using positive feedback to guide

subsequent choices is more pronounced in monkeys with PFO lesions relative to those with ACC lesions.

Because monkeys with PFO lesions were found to be unimpaired on an action reversal task (Rudebeck et al. 2008), it has been proposed that PFO plays a specific role in integrating positive feedback to guide choices between objects rather than actions (Rushworth et al. 2007; Rudebeck and Murray 2008; Rudebeck et al. 2008; Walton, Behrens et al. 2010). In addition, as indicated in the Introduction, it has been hypothesized that the ACC plays a selective role in reward-guided action selection (Rushworth et al. 2007; Rudebeck et al. 2008). The latter idea appears to be incongruent, however, with the current finding that the ACC is necessary for effective reward guided *object* selection.

One possible explanation for this apparent discrepancy is that the performance of the object reversal task may involve not only object–outcome representations, but also action–outcome strategies. (For discussion of the processes contributing to performance on object discrimination tests, see Roberts 2006.) Thus, the possibility remains that ACC lesions compromise the appropriate action targeted toward the object, rather than selection of the object itself. According to this view, the impairments in sustaining the correct action in the context of an action reversal task (Kennerley et al. 2006) and of sustaining selection of the correct object in an object reversal task (current study) arise from one and the same impairment, namely, a failure in selecting goals for action. Such an impairment would fit within a simple hierarchical model of choice where the PFO is important for representing stimulus–outcome relationships which are then passed on to the more motor related ACC for voluntary choice (Rushworth and Behrens 2008; Cai and Padoa-Schioppa 2012). Still, as discussed earlier in the section Updating Value of Object–Outcome Associations, not all object-based choices are dependent on the ACC.

One possibility, not mutually exclusive with the foregoing, is that ACC is particularly important for guiding a choice when the reinforcement history of object and action choices needs to be taken into account. This idea is consistent with a role for ACC in both object and action reversal learning (Kennerley et al. 2006; current study), the performance of which benefits from knowledge of prior choice outcomes, and with the limited role for ACC in object and action devaluation tasks, which depend on the current value of the outcomes, as opposed to the history of choice outcomes. According to this view, the adverse effects of ACC lesions on object and action reward-guided choices would be instantiated by a diminished ability to represent or use knowledge of the history of outcomes of previous choices to guide the current one. This idea receives support from the finding that monkeys with ACC lesions are less influenced than controls by the history of rewarded and unrewarded trials (Kennerley et al. 2006) and more likely to switch responses after either correct or incorrect choices (present study).

Linking Actions with Reward Contingencies

We found direct evidence implicating the ACC in linking actions with reward contingency. Specifically, monkeys with ACC lesions were less likely to switch their response following strings of errors (see Fig. 5F), suggesting that they were impaired in using error feedback (i.e. nonreward) to guide their choice of action. This result appears to conflict with the

findings of Kennerley et al. (2006), who found that monkeys with lesions of cortex in the banks of the anterior cingulate sulcus were impaired in using positive feedback (i.e. reward) to guide their responses in an action reversal task. The monkeys studied by Kennerley et al. (2006) received lesions of the ACC *after* they had acquired the action–outcome contingencies and were highly experienced with reversals. Rather than re-examine the effect of lesions on the performance of well-learned reversals, we opted to examine the effect of ACC lesions on postoperative acquisition of action reversals. Accordingly, one possible explanation for the pattern of results is that the ACC is sensitive to changes in reward information following the performance of well-learned action reversals (as in Kennerley et al. 2006), whereas during initial learning of action reversals (as in the present study), the ACC is important for using error feedback to guide the choice of response. Consistent with this idea, neurophysiological studies have shown that many neurons in the ACC respond to both rewards and errors (Niki and Watanabe 1979; Ito et al. 2003; Matsumoto et al. 2003), and are sensitive to the context in which errors occur (Shima and Tanji 1998; Brown and Braver 2005). Thus, taking into consideration the differences in training protocol, the combined results from Kennerley et al. (2006) and the present study strongly suggest a role for the ACC in using information about both rewards and errors to guide behavior. This idea accords with recent electrophysiological data, which implicates the ACC not only in representing reward value (Kennerley and Wallis 2009; Kennerley et al. 2009) and linking specific actions with rewards (Shima and Tanji 1998; Procyk et al. 2000; Matsumoto et al. 2003), but also in more general the aspects of choice behavior regarding the success (or not) of choice behavior, including positive and negative outcomes of choices, outcomes not received, and unsigned reward prediction errors among other things (Matsumoto et al. 2007; Seo and Lee 2007; Quilodran et al. 2008; Hayden et al. 2009, 2011). Thus, ACC neurons provide a suite of signals that together promote efficient learning and adaptive behavior in both natural settings and in the laboratory.

Conclusions

In this study, we illustrate an important role for the ACC in integrating information about reward and errors when making decisions that require choices involving objects or actions, respectively. In object reversal learning, monkeys with ACC lesions were deficient in using reward information to guide their subsequent choice of object. In action reversal learning, monkeys with ACC lesions were impaired in their ability to switch their response after making consecutive errors. Taken together, these data indicate that the ACC promotes decision-making not only in the action–outcome domain, as previously reported (Rushworth et al. 2004, 2007), but also in the object–outcome domain. These data are consistent with the idea that the ACC represents various aspects of reward value and plays a general role in dictating how much influence individual outcomes should be given to guide adaptive behavior (Kennerley and Wallis 2009; Hayden et al. 2011; Kennerley and Walton 2011). We propose that the ACC is especially important for guidance of object and action choices whenever those choices are dependent on knowledge of the history of choice outcomes (reward and nonreward), as opposed to the current value of outcomes (Behrens et al.

2007). Whether the contribution of ACC to both object and action selection can be unified by consideration of a role for ACC in representation of higher order goals remains a topic for future investigation.

Supplementary Material

Supplementary material can be found at: <http://www.cercor.oxfordjournals.org/>

Funding

This work was supported by the Intramural Research Program of the National Institute of Mental Health at the National Institutes of Health.

Notes

We are grateful to Katherine S. Wright for help with brain lesion reconstructions and Ping-yu Chen for help assessing the volume of the lesions. *Conflict of Interest*: None declared.

References

- Alexander WH, Brown JW. 2011. Medial prefrontal cortex as an action–outcome predictor. *Nat Neurosci*. 14:1338–1344.
- Amiez C, Joseph JP, Procyk E. 2006. Reward encoding in the monkey anterior cingulate cortex. *Cereb Cortex*. 16:1040–1055.
- Barbas H, Blatt GJ. 1995. Topographically specific hippocampal projections target functionally distinct prefrontal areas in the rhesus monkey. *Hippocampus*. 5:511–533.
- Barbas H, De Olmos J. 1990. Projections from the amygdala to basoventral and mediodorsal prefrontal regions in the rhesus monkey. *J Comp Neurol*. 300:549–571.
- Barbas H, Ghashghaei H, Dombrowski SM, Rempel-Clower NL. 1999. Medial prefrontal cortices are unified by common connections with superior temporal cortices and distinguished by input from memory-related areas in the rhesus monkey. *J Comp Neurol*. 410:343–367.
- Baxter MG, Parker A, Lindner CC, Izquierdo AD, Murray EA. 2000. Control of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex. *J Neurosci*. 20:4311–4319.
- Behrens TEJ, Woolrich MW, Walton ME, Rushworth MFS. 2007. Learning the value of information in an uncertain world. *Nat Neurosci*. 10:1214–1221.
- Brown JW, Braver TS. 2005. Learned predictions of error likelihood in the anterior cingulate cortex. *Science*. 307:1118–1121.
- Cai X, Padoa-Schioppa C. 2012. Neuronal encoding of subjective value in dorsal and ventral anterior cingulate cortex. *J Neurosci*. 32:3791–3808.
- Camille N, Tsuchida A, Fellows LK. 2011. Double dissociation of stimulus-value and action-value learning in humans with orbitofrontal or anterior cingulate cortex damage. *J Neurosci*. 31:15048–15052.
- Carmichael ST, Price JL. 1994. Architectonic subdivision of the orbital and medial prefrontal cortex in the macaque monkey. *J Comp Neurol*. 346:366–402.
- Carmichael ST, Price JL. 1995. Sensory and premotor connections of the orbital and medial prefrontal cortex of macaque monkeys. *J Comp Neurol*. 363:642–664.
- Dickinson A, Nicholas DJ, Adams CD. 1983. The effect of the instrumental training contingency on susceptibility to reinforcer devaluation. *Q J Exp Psycho Section B*. 35:35–51.
- Ferry AT, Ongür D, An X, Price JL. 2000. Prefrontal cortical projections to the striatum in macaque monkeys: evidence for an organization related to prefrontal networks. *J Comp Neurol*. 425:447–470.
- Ghashghaei HT, Barbas H. 2002. Pathways for emotion: interactions of prefrontal and anterior temporal pathways in the amygdala of the rhesus monkey. *Neuroscience*. 115:1261–1279.
- Hadland KA, Rushworth MFS, Gaffan D, Passingham RE. 2003. The anterior cingulate and reward-guided selection of actions. *J Neurophysiol*. 89:1161–1164.
- Hayden BY, Pearson JM, Platt ML. 2009. Fictive reward signals in the anterior cingulate cortex. *Science*. 324:948–950.
- Hayden BY, Pearson JM, Platt ML. 2011. Neuronal basis of sequential foraging decisions in a patchy environment. *Nat Neurosci*. 14:933–939.
- Hayden BY, Platt ML. 2010. Neurons in anterior cingulate cortex multiplex information about reward and action. *J Neurosci*. 30:3339–3346.
- Ito S, Stuphorn V, Brown JW, Schall JD. 2003. Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science*. 302:120–122.
- Izquierdo A, Suda RK, Murray EA. 2004. Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. *J Neurosci*. 24:7540–7548.
- Kennerley SW, Dahmubed AF, Lara AH, Wallis JD. 2009. Neurons in the frontal lobe encode the value of multiple decision variables. *J Cogn Neurosci*. 21:1162–1178.
- Kennerley SW, Wallis JD. 2009. Evaluating choices by single neurons in the frontal lobe: outcome value encoded across multiple decision variables. *Eur J Neurosci*. 29:2061–2073.
- Kennerley SW, Walton ME. 2011. Decision making and reward in frontal cortex: complementary evidence from neurophysiological and neuropsychological studies. *Behav Neurosci*. 125:297–317.
- Kennerley SW, Walton ME, Behrens TEJ, Buckley MJ, Rushworth MFS. 2006. Optimal decision making and the anterior cingulate cortex. *Nat Neurosci*. 9:940–947.
- Kondo H, Saleem KS, Price JL. 2005. Differential connections of the perirhinal and parahippocampal cortex with the orbital and medial prefrontal networks in macaque monkeys. *J Comp Neurol*. 493:479–509.
- Luk C-H, Wallis JD. 2009. Dynamic encoding of responses and outcomes by neurons in medial prefrontal cortex. *J Neurosci*. 29:7526–7539.
- Luppino G, Rozzi S, Calzavara R, Matelli M. 2003. Prefrontal and agranular cingulate projections to the dorsal premotor areas F2 and F7 in the macaque monkey. *Eur J Neurosci*. 17:559–578.
- Machado CJ, Bachevalier J. 2007. Measuring reward assessment in a semi-naturalistic context: the effects of selective amygdala, orbital frontal or hippocampal lesions. *Neuroscience*. 148:599–611.
- Matsumoto K, Suzuki W, Tanaka K. 2003. Neuronal correlates of goal-based motor selection in the prefrontal cortex. *Science*. 301:229–232.
- Matsumoto M, Matsumoto K, Abe H, Tanaka K. 2007. Medial prefrontal cell activity signaling prediction errors of action values. *Nat Neurosci*. 10:647–656.
- Meunier M, Destrède C. 1997. Effects of radiofrequency versus neurotoxic cingulate lesions on spatial reversal learning in mice. *Hippocampus*. 7:355–360.
- Morecraft RJ, McNeal DW, Stilwell-Morecraft KS, Gedney M, Ge J, Schroeder CM, van Hoesen GW. 2007. Amygdala interconnections with the cingulate motor cortex in the rhesus monkey. *J Comp Neurol*. 500:134–165.
- Murray EA, Izquierdo A. 2007. Orbitofrontal cortex and amygdala contributions to affect and action in primates. *Ann N Y Acad Sci*. 1121:273–296.
- Neave N, Nagle S, Sahgal A, Aggleton JP. 1996. The effects of discrete cingulum bundle lesions in the rat on the acquisition and performance of two tests of spatial working memory. *Behav Brain Res*. 80:75–85.
- Niki H, Watanabe M. 1979. Prefrontal and cingulate unit activity during timing behavior in the monkey. *Brain Res*. 171:213–224.
- Ostlund SB, Balleine BW. 2007. The contribution of orbitofrontal cortex to action selection. *Ann N Y Acad Sci*. 1121:174–192.
- Ostlund SB, Balleine BW. 2005. Lesions of medial prefrontal cortex disrupt the acquisition but not the expression of goal-directed learning. *J Neurosci*. 25:7763–7770.

- Picard N, Strick PL. 1996. Motor areas of the medial wall: a review of their location and functional activation. *Cereb Cortex*. 6:342–353.
- Procyk E, Tanaka YL, Joseph JP. 2000. Anterior cingulate activity during routine and non-routine sequential behaviors in macaques. *Nat Neurosci*. 3:502–508.
- Quilodran R, Rothé M, Procyk E. 2008. Behavioral shifts and action valuation in the anterior cingulate cortex. *Neuron*. 57:314–325.
- Rhodes SEV, Murray EA. 2010. Knowing what you know and how you get it: the involvement of prelimbic cortex and the amygdala in goal-directed behavior in rhesus monkeys. *Soc. Neurosci. Abstr.* #711.20.
- Roberts AC. 2006. Primate orbitofrontal cortex and adaptive behaviour. *Trends Cogn Sci*. 10:83–90.
- Rudebeck PH, Behrens TE, Kennerley SW, Baxter MG, Buckley MJ, Walton ME, Rushworth MFS. 2008. Frontal cortex subregions play distinct roles in choices between actions and stimuli. *J Neurosci*. 28:13775–13785.
- Rudebeck PH, Murray EA. 2008. Amygdala and orbitofrontal cortex lesions differentially influence choices during object reversal learning. *J Neurosci*. 28:8338–8343.
- Rudebeck PH, Murray EA. 2011. Dissociable effects of subtotal lesions within the macaque orbital prefrontal cortex on reward-guided behavior. *J Neurosci*. 31:10569–10578.
- Rushworth MFS, Behrens TEJ. 2008. Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat Neurosci*. 11:389–397.
- Rushworth MFS, Behrens TEJ, Rudebeck PH, Walton ME. 2007. Contrasting roles for cingulate and orbitofrontal cortex in decisions and social behaviour. *Trends Cogn Sci*. 11:168–176.
- Rushworth MFS, Walton ME, Kennerley SW, Bannerman DM. 2004. Action sets and decisions in the medial frontal cortex. *Trends Cogn Sci*. 8:410–417.
- Saleem KS, Kondo H, Price JL. 2008. Complementary circuits connecting the orbital and medial prefrontal networks with the temporal, insular, and opercular cortex in the macaque monkey. *J Comp Neurol*. 506:659–693.
- Seo H, Lee D. 2007. Temporal filtering of reward signals in the dorsal anterior cingulate cortex during a mixed-strategy game. *J Neurosci*. 27:8366–8377.
- Shima K, Tanji J. 1998. Role for cingulate motor area cells in voluntary movement selection based on reward. *Science*. 282:1335–1338.
- Suzuki WA, Amaral DG. 1994. Perirhinal and parahippocampal cortices of the macaque monkey: cortical afferents. *J Comp Neurol*. 350:497–533.
- Vogt BA. 1993. Structural organization of cingulate cortex: areas, neurons, and somatodendritic transmitter receptors. In: Vogt BA, Gabriel M, editors. *Neurobiology of Cingulate Cortex and Limbic Thalamus*. Boston: Birkhäuser. p. 19–70.
- Walton ME, Behrens TEJ, Buckley MJ, Rudebeck PH, Rushworth MFS. 2010. Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. *Neuron*. 65:927–939.
- Walton ME, Rudebeck PH, Behrens TEJ, Rushworth MFS. 2010. Cingulate and orbitofrontal contributions to valuing knowns and unknowns in a changeable world. In: Delgado MR, Phelps EA, Robbins TW, editors. *Decision Making, Affect, and Learning: Attention and Performance XXIII*. New York: Oxford University Press. p. 235–555.
- West EA, DesJardin JT, Gale K, Malkova L. 2011. Transient inactivation of orbitofrontal cortex blocks reinforcer devaluation in macaques. *J Neurosci*. 31:15128–15135.
- Williams SM, Goldman-Rakic PS. 1993. Characterization of the dopaminergic innervation of the primate frontal cortex using a dopamine-specific antibody. *Cereb Cortex*. 3:199–222.