

Lesions of Medial Prefrontal Cortex Disrupt the Acquisition But Not the Expression of Goal-Directed Learning

Sean B. Ostlund and Bernard W. Balleine

Department of Psychology and the Brain Research Institute, University of California, Los Angeles, Los Angeles, California 90095-1563

Several studies have established that pretraining lesions of the medial prefrontal cortex (mPFC) render instrumental actions insensitive to devaluation of the instrumental outcome and degradation of the action–outcome contingency. Nevertheless, it remains to be assessed whether the involvement of the mPFC in goal-directed action is limited to the acquisition or to the expression of the action–outcome association in performance. The current series of experiments investigated this issue by comparing the effects of mPFC lesions made either before or after initial training using sensitivity to outcome devaluation as an assay of goal-directed performance. Whereas pretraining lesions left performance insensitive to outcome devaluation, posttraining lesions spared this effect. To determine whether the effect of mPFC lesions on outcome devaluation was the result of a more fundamental deficit in response selection, experiment 2 assessed the impact of pretraining and posttraining lesions on the ability of the instrumental outcome to selectively reinstate the performance of its associated action after a period of extinction. Although both lesions attenuated the magnitude of instrumental reinstatement generally, they left intact the ability of the instrumental outcome to influence response selection. Experiment 3 investigated the relationship between the outcome-selective devaluation and reinstatement effects and found evidence that these effects are both behaviorally and neurally dissociable at the level of the mPFC. These results indicate that the mPFC is selectively involved in the acquisition, but not the permanent storage or expression, of action–outcome associations in instrumental conditioning.

Key words: reward; prefrontal cortex; response selection; instrumental conditioning; devaluation; priming

Introduction

There is considerable behavioral evidence that, in instrumental conditioning, rats encode the relationship between their actions and the specific goal or outcome of those actions. For example, after being trained on two distinct actions with unique outcomes, the devaluation of one of the two outcomes results in the selective reduction in the performance of its associated action relative to the other action (Colwill and Rescorla, 1985; Balleine and Dickinson, 1998). At a neural level, converging lines of evidence have implicated the medial prefrontal cortex (mPFC) in goal-directed instrumental actions (Dalley et al., 2004; Matsumoto and Tanaka, 2004). Electrophysiological studies using primates (Matsumoto et al., 2003) and rats (Mulder et al., 2003) have found neural activity in the mPFC related to specific action–outcome associations. Furthermore, several studies have shown that the rat mPFC is necessary for the normal organization of goal-directed action; pretraining neurotoxic lesions of the prelimbic area (PL) have been found to abolish the sensitivity of performance to outcome devaluation and selective degradation of the action–outcome contingency (Balleine and Dickinson, 1998; Corbit and Balleine, 2003; Killcross and Coutureau, 2003).

Lesions of the mPFC do not, however, block the acquisition of

instrumental actions altogether and leave intact several indicators of outcome-mediated response selection, including the selective facilitatory effects of noncontingent rewards or reward-related stimuli on performance (Corbit and Balleine, 2003). On the basis of this preserved function, it has been proposed that, in the absence of the mPFC, instrumental performance is primarily controlled by extraneous stimuli (cf. Balleine and Dickinson, 1998; Corbit and Balleine, 2003).

Because the specific role of the mPFC in action–outcome learning has, however, only been assessed after pretraining manipulations, it remains unknown whether it is selectively involved in encoding, storing, or implementing these associations. The first aim of the current study was, therefore, to compare the effects of pretraining and posttraining mPFC lesions on instrumental conditioning using outcome devaluation as a test of the integrity of action–outcome learning. Furthermore, we developed and used a direct test of outcome–action associations based on an outcome-specific reinstatement protocol in which, after training on two action–outcome relationships, choice performance was examined when one or other instrumental outcome was delivered noncontingently after a period of extinction. Together with outcome devaluation, this test enabled us to assess directly the effects of pretraining and posttraining lesions on action initiation based on both action→outcome and outcome→action associations.

Finally, we extended this assessment in experiment 3 to examine the role of the mPFC in the control of action initiation based on anticipated outcome value. Here, we assessed the effect of

Received May 13, 2005; revised July 14, 2005; accepted July 15, 2005.

This work was supported by National Institute of Mental Health Grant 56446.

Correspondence should be addressed to Sean Ostlund, Department of Psychology, University of California, Los Angeles, Box 951563, Los Angeles, CA 90095-1563. E-mail: sostlund@ucla.edu.

DOI:10.1523/JNEUROSCI.1921-05.2005

Copyright © 2005 Society for Neuroscience 0270-6474/05/257763-08\$15.00/0

devaluing the noncontingent outcome on outcome-mediated reinstatement in sham and in pretraining mPFC-lesioned rats. We predicted that, although a devalued outcome should not reinstate performance of an action delivering that outcome in sham rats, if performance in pretraining mPFC-lesioned rats is elicited only by the stimulus properties of the instrumental outcome, then they should show significant, selective reinstatement even when the reinstating outcome has been devalued.

Materials and Methods

Subjects and apparatus

Female Long-Evans rats, weighing between 225 and 250 g at the beginning of the experiment, were used as subjects. The rats were housed in pairs in transparent plastic tubs located in a temperature- and humidity-controlled vivarium. Throughout behavioral training and testing, rats were maintained at ~85% of their free-feeding body weight by restricting their food intake to between 10 and 12 g of their maintenance diet per day. This daily food allotment was reduced by half on outcome devaluation test days, when rats were provided with 1 h of *ad libitum* access to one of the training outcomes.

The behavioral procedures were performed in 16 identical Med Associates (East Fairfield, VT) operant chambers enclosed in sound- and light-attenuating shells. Each chamber was equipped with a recessed food magazine, located at the base of one end wall, through which 20% sucrose solution (0.1 ml) and food pellets (45 mg; Bio-Serv, Frenchtown, NJ) could be delivered using a syringe pump and pellet dispenser, respectively. An infrared photobeam crossed the magazine opening, allowing for the detection of head entries. Each chamber also contained a pair of retractable levers that were located to the left or right of the food magazine. A houselight (3 W, 24 V) located on the end wall opposite the magazine provided constant illumination, and an electric fan fixed in the shell enclosure provided background noise (~70 dB) throughout training and testing. A set of three microcomputers running the Med-PC program (Med Associates) controlled all experimental events and recorded lever presses and magazine entries.

Surgical procedures

Rats were provided *ad libitum* access to their maintenance chow on the day before and on the 5 d that followed surgery, regardless of group assignment. At the time of surgery, rats were anesthetized with pentobarbital (Nembutal; 50 mg/kg) and were administered atropine (0.1 mg) before being placed in a stereotaxic frame (Stoelting, Wood Dale, IL). An incision was made into the scalp to expose the skull surface, and the incisor bar was adjusted to place bregma and lambda in the same horizontal plane. For all rats, two small holes were drilled into the skull above the target structure. Excitotoxic lesions were made by infusing 0.4 μ l of NMDA (20 μ g/ μ l in PBS) over 4 min into the mPFC of each hemisphere (all coordinates relative to bregma; anteroposterior, +3.3; mediolateral, \pm 0.7; dorsoventral, -3.5) using a 1 μ l Hamilton syringe. The needle was left in place for an additional 4 min to allow for diffusion of the drug. Sham lesions were made using the same procedures except that the needle was not lowered and no drug was infused. A recovery period of 10 d was provided between surgery and behavioral testing. Rats were handled daily and returned to the food deprivation schedule during the last 5 d of this period.

Histology

After behavioral testing, the rats received a lethal overdose of sodium pentobarbital and were perfused transcardially with 0.9% saline followed by 10% buffered formalin solution. The brains were then extracted and postfixed in a 25% sucrose-formalin solution. After several days, the brains were frozen, and 50 μ m coronal sections of the prefrontal cortex were collected on glass slides. The sections were stained with thionin and examined with a microscope to assess the placement and extent of neuronal damage through comparison with sham control sections and the stereotaxic atlas by Paxinos and Watson (1998).

Behavioral analysis

Behavioral data were analyzed as the mean number of lever presses per minute. However, to assess the effect of the noncontingent outcome

delivery on response selection, the results of reinstatement testing were subjected to additional analysis as the mean percentage of total responses (i.e., [responses on the reinstated lever / (responses on the reinstated lever + responses on the other lever)] \times 100). Specifically, we compared rats' choice performance during the reinstatement phase (i.e., after outcome delivery) to their baseline choice performance during the extinction phase of the test. This measure is particularly sensitive to changes in the distribution of responses across actions (i.e., in choice performance) independently of individual variability in response magnitude.

Experiment 1: effect of posttraining mPFC lesions on instrumental conditioning

Behavioral training. Rats were handled on each of the 5 d that preceded behavioral training. Over the next 2 d, rats received daily magazine training sessions, during each of which 15 pellet and 15 sucrose presentations were made on independent random time 60 s schedules. On each of the next 11 d, the two instrumental responses (left and right lever press) were rewarded with unique outcomes (either pellets or sucrose) in separate daily training sessions, such that only one response and one outcome were available in any given session. For one-half of the subjects in each group, left lever presses earned pellets and right lever presses earned sucrose solution, whereas the remaining subjects were trained on the opposite action–outcome contingencies. Each training session lasted for 30 min. The two daily sessions were separated by at least 30 min, and session order was alternated over days. Rats were continuously reinforced for lever pressing during the first 2 d of training. The reinforcement schedule was then gradually shifted in 3 d blocks to random ratio-5 (RR-5; probability of reward for each response is 0.2 on average), RR-10, and finally RR-20.

Surgery. Rats were assigned to surgery groups in a quasirandom manner; initial random group assignments were adjusted using baseline instrumental performance to control for response biases. Two days after the end of instrumental training, rats received either sham ($n = 8$) or excitotoxic ($n = 10$) lesions of the mPFC using the procedures described above.

Retraining. After the recovery period, rats were given 3 d of retraining on each response. The procedures used for retraining were identical to those used during the last 3 d of instrumental training. We predicted that if the expression of previously acquired goal-directed learning depends on the mPFC, then posttraining lesions of this structure should disrupt the reacquisition of lever pressing during this retraining period.

Reinstatement testing. Rats received two sessions of reinstatement testing, one session with each outcome, throughout which both responses were continuously available but were not rewarded. Each test session consisted of three separate phases: (1) a 20 min extinction phase, used to suppress the rate of responding; (2) an outcome delivery phase, during which a single noncontingent outcome (either the food pellet or sucrose solution) was delivered; and (3) a 3 min reinstatement phase, used to assess the effects of the outcome delivery on subsequent instrumental performance. A response on either lever during the extinction phase delayed the delivery of the reinstating outcome by 15 s to minimize the risk of an accidental pairing between it and any residual (i.e., nonextinguished) lever pressing. The reinstatement phase was initiated by the first magazine entry made after the outcome delivery. One-half of the rats in each group received a food pellet during the first reinstatement test and sucrose during the second reinstatement test, whereas the remaining rats received the opposite arrangement. Test order was counterbalanced with training contingency. One session of retraining (RR-20) with each response was provided on the day between the two reinstatement tests.

Outcome devaluation testing. After the second reinstatement test, all of the rats received a day of retraining (RR-20) on both actions before outcome devaluation testing. Rats received two sessions of outcome devaluation testing, one session with each outcome. Before each test session, one of the training outcomes was selectively devalued using a specific satiety procedure (cf. Balleine and Dickinson, 1998). Specifically, rats received 1 h of unrestricted access to either the pellets or the sucrose solution in their home cage immediately before they were returned to the experimental chamber for a 5 min choice extinction test. During the test session, both levers were inserted into the box, but no rewards were

delivered. Test order was reversed during devaluation testing relative to reinstatement testing. For example, if the order was pellets then sucrose across reinstatement tests, then a sucrose-then-pellets order was used during devaluation testing. This procedure ensured that one-half of the rats in each group were prefed on pellets before the first devaluation test and sucrose before the second devaluation test, whereas the remaining half received the opposite arrangement. One session of retraining (RR-20) with each response was provided on the day between outcome devaluation tests.

Experiment 2: effect of pretraining versus posttraining mPFC lesions on reinstatement and outcome devaluation

Pretraining surgery. One-half of the rats were assigned randomly to pretraining surgery groups. Of these rats, 12 received excitotoxic mPFC lesions and five received sham lesions using the procedures described above. The remaining 17 rats were assigned to posttraining lesion groups (see below). All rats, regardless of group, received the same handling and feeding treatment during this phase of the experiment.

Behavioral training. Magazine and instrumental training were conducted using exactly the same procedures described in experiment 1.

Posttraining surgery. Of those rats assigned to posttraining surgery groups, 12 received excitotoxic mPFC lesions and five received sham lesions using the procedures described above generating three final groups: group pre ($n = 12$) and group post ($n = 12$), which were given lesions of the mPFC before and after training, respectively, and group sham ($n = 10$), half of which was given sham surgery before training and half after training. Posttraining group assignments were again made in a quasirandom manner; initial random group assignments were adjusted using baseline instrumental performance to control for any response biases. All rats, regardless of group, received the same handling and feeding treatment during this phase of the experiment.

Outcome devaluation testing. After the recovery period, outcome devaluation testing was conducted using the procedures described in experiment 1.

Reinstatement testing. After outcome devaluation testing, reinstatement testing was conducted using the procedures described in the experiment 1. Moreover, the same outcome-related counterbalancing procedure was used to assign test order (e.g., if the order was pellets then sucrose during outcome devaluation testing, then the order sucrose then pellets was used during reinstatement testing).

Experiment 3: effect of pretraining mPFC lesions on the sensitivity of reinstatement to the devaluation of the reinstating outcome

Surgery. Rats were randomly assigned to either the mPFC ($n = 8$) or sham ($n = 8$) surgery group. Surgeries were performed before training in the manner described above.

Behavioral training. Magazine and instrumental training were conducted using exactly the same procedures described in experiment 1, except that each training session terminated after either 30 min had elapsed or 30 outcomes had been earned, whichever came first.

Devalued reinstatement testing. To assess the dependence of instrumental reinstatement on the incentive value of the noncontingent outcome, the rats were given two sessions of devalued reinstatement testing. These tests were conducted using the procedures described in experiment 1, except that the outcome to be delivered was devalued immediately before the test session using the selective satiety procedure described in experiment 1. One-half of the rats in each group were first tested with food pellets (i.e., they were sated on pellets before they received a pellet reinstatement test) and then tested with sucrose, whereas the remaining rats received the opposite test order. Outcome assignments were counterbalanced with regard to the training contingencies.

Results

Histology

Figure 1 shows the maximal and minimal areas of damage in mPFC-lesioned rats. Histological analysis revealed that, in general, the NMDA infusions resulted in substantial bilateral damage of the mPFC centered around the PL, but also extending to both the dorsal bank of the infralimbic area (IL) and the ventral bank

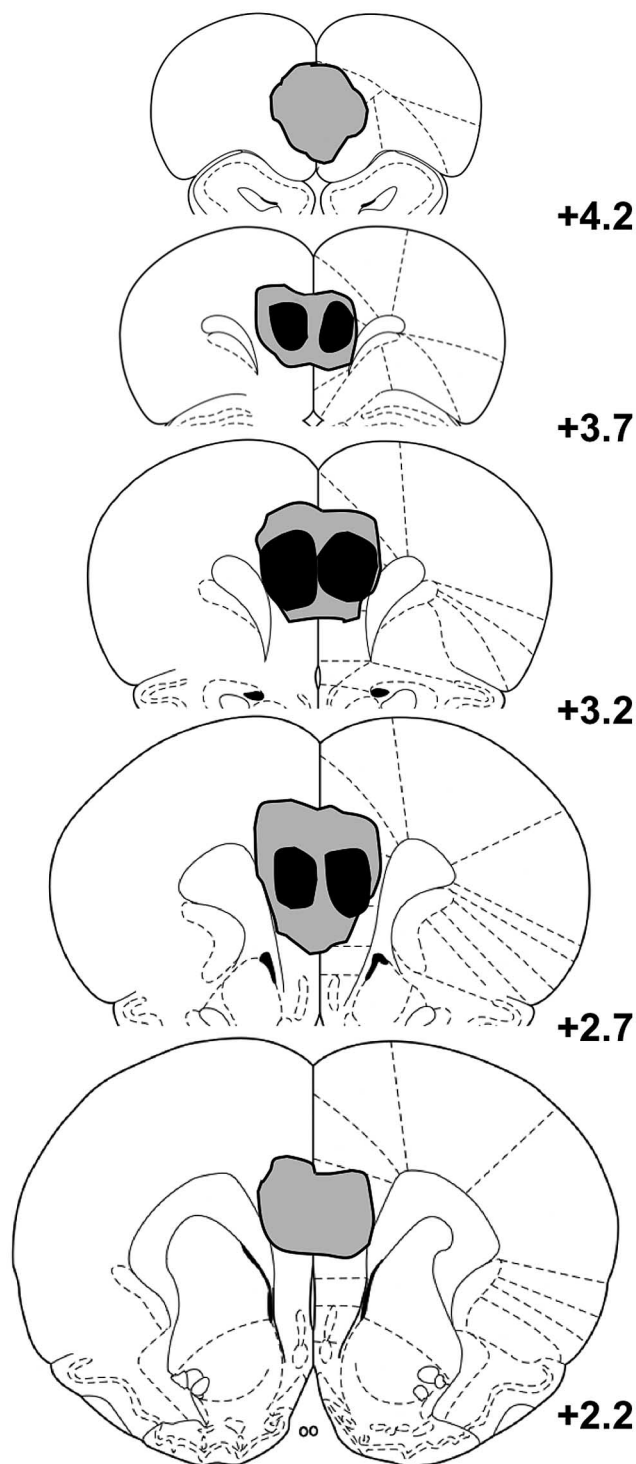


Figure 1. Schematic representation of minimum (black) and maximum (gray) extent of mPFC damage. Coronal sections are modified from Paxinos and Watson (1998). The number next to each section refers to its position in the anteroposterior plane relative to bregma (in millimeters).

of the anterior cingulate cortex in some rats, although this latter damage was not systematic. However, because it has been shown that selective IL lesions leave intact the sensitivity of instrumental performance to outcome devaluation (Killcross and Coutureau, 2003), it seems unlikely that the effects of mPFC lesions reported here were the consequence of the minimal and infrequent IL

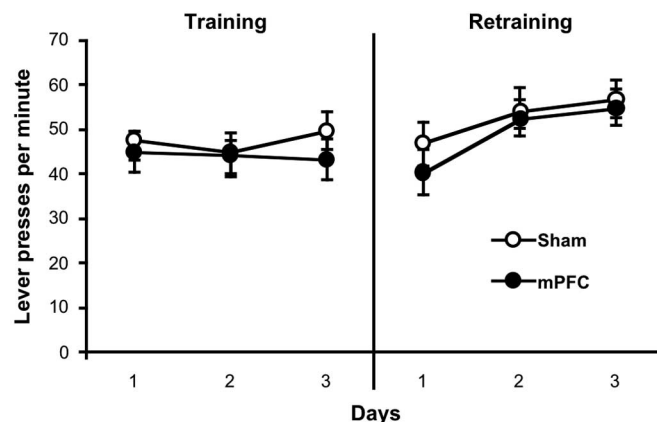


Figure 2. Reacquisition of instrumental performance. The mean number of lever presses per minute (± 1 SEM), averaged across daily sessions, are shown for the last 3 d of training (left) and 3 d of retraining after surgery (right). The data are plotted separately for the sham and mPFC group.

damage that we observed. The data from one lesioned rat from experiment 1, three lesioned rats from experiment 2 (two pretraining and one posttraining), and one lesioned rat from experiment 3 were excluded from the statistical analysis because they received only unilateral damage of the PL.

Experiment 1: effects of posttraining mPFC lesions on instrumental conditioning

Instrumental training and retraining

The results of instrumental training are presented in Figure 2 (left) as the mean number of lever presses per minute averaged across levers during each of the last 3 d of training. An ANOVA performed on this data found no effect of day ($F < 1$), group ($F < 1$), nor a day-by-group interaction ($F_{(2,26)} = 1.04$; $p > 0.05$). After recovery from surgery, rats were given 3 d of retraining to determine whether mPFC lesions affected the retraining of previously acquired instrumental performance. These data are presented in Figure 2 (right), which clearly shows that both the sham and lesioned group maintained high levels of instrumental performance and increased their rate of lever pressing similarly over days of retraining. An ANOVA confirmed this conclusion, revealing a main effect of day ($F_{(2,26)} = 14.38$; $p < 0.0001$) but no effect of group ($F < 1$) nor a group-by-day interaction ($F < 1$). Thus, posttraining mPFC lesions did not impair the retraining of previously acquired lever pressing.

Reinstatement testing

The data from reinstatement testing are presented in Figure 3. As shown in the left panel, the nonreinforcement procedure was similarly effective at suppressing the instrumental performance of sham and lesioned rats. The extinction data were analyzed using an ANOVA, which revealed a main effect of block ($F_{(3,39)} = 94.91$; $p < 0.0001$) but no effect of group ($F < 1$), response ($F < 1$; to-be-reinstated vs other), nor any interactions between these variables (largest $F_{(3,39)} = 1.92$; $p > 0.14$).

Data from the reinstatement phase are presented as the percentage of the total responses made for the noncontingent, or reinstating, outcome (Fig. 3, right, choice performance). Clear and selective reinstatement was observed in both groups on this choice measure. Nevertheless, mPFC lesions appeared to reduce the overall magnitude of reinstatement; the overall response rate summed across both reinstated and the other action (± 1 SEM)

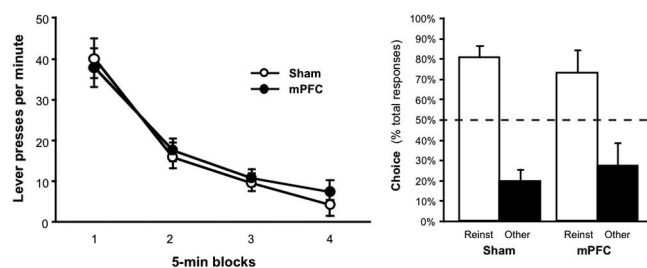


Figure 3. Extinction and selective reinstatement of instrumental performance. Left, The mean number of lever presses per minute (± 1 SEM) during consecutive 5 min blocks of extinction for the sham and mPFC groups. Right, The mean percentage of total responses (± 1 SEM) made on each lever after the noncontingent outcome delivery for the sham and mPFC group. The data are plotted according to whether the action was trained with the outcome delivered during reinstatement (Reinst) or other outcome (Other).

during these tests was 26.1 (± 6.8) for the sham group and 10.5 (± 3.0) for the mPFC group. As is clear in Figure 3, however, the relative magnitude of the effect of noncontingent outcome delivery on response selection was similar in the sham and mPFC groups. An ANOVA conducted on the mean responses per minute revealed a main effect of response ($F_{(1,13)} = 16.75$; $p < 0.01$) but no effect of group ($F_{(1,13)} = 3.92$; $p = 0.07$) and no response-by-group interaction ($F_{(1,13)} = 4.1$; $p = 0.06$). Although the data suggest that the sham group responded more on the reinstated action than the mPFC group, this effect did not reach significance ($F_{(1,13)} = 4.21$; $p = 0.06$). Moreover, both the sham group ($F_{(1,7)} = 12.64$; $p < 0.01$) and lesioned group ($F_{(1,6)} = 6.34$; $p < 0.05$) performed significantly more responses on the reinstated action than on the other action. This conclusion was also supported by the analysis of reinstatement choice performance (Fig. 3, right) (see Materials and Methods). Both groups increased their choice of the reinstated action after the noncontingent outcome delivery (i.e., during the reinstatement phase) relative to their baseline choice performance during the extinction phase, which was 46.3% (± 5.0) for the sham group and 47.3% (± 4.3) for the mPFC group. An ANOVA found a main effect of test phase ($F_{(1,13)} = 18.04$; $p < 0.01$) but no effect of group ($F < 1$) and no phase-by-group interaction ($F < 1$), indicating that both groups increased their choice of the action that had earned the reinstating outcome after it had been delivered.

Outcome devaluation testing

The results of outcome devaluation testing are presented in Figure 4 as the mean number of lever presses per minute, separately plotted for the action that had earned the devalued outcome during training and for the other action. Clearly, both groups displayed sensitivity to the reduction in outcome value, making fewer responses on the lever that, in training, had delivered the outcome with which they were satiated before testing than on the other lever. An ANOVA revealed a significant main effect of response ($F_{(1,13)} = 26.27$; $p < 0.001$) but no effect of group ($F < 1$) and no response-by-group interaction ($F_{(1,13)} = 2.64$; $p > 0.05$). Additional analysis indicated that both the sham group ($F_{(1,7)} = 21.98$; $p < 0.01$) and mPFC group ($F_{(1,6)} = 11.75$; $p < 0.05$) showed a reliable devaluation effect. These findings contrast with previous reports that pretraining mPFC lesions disrupt the sensitivity of instrumental performance to a reduction in outcome value (Balleine and Dickinson, 1998; Corbit and Balleine, 2003; Killcross and Coutureau, 2003). Instead, the current results suggest that mPFC lesions spare the outcome devaluation effect if they are made after initial training. In this experiment, however,

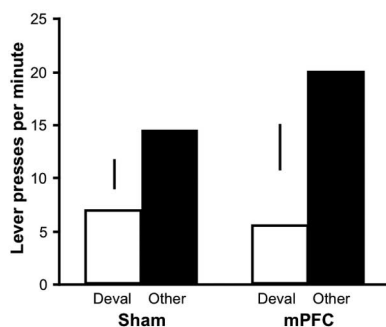


Figure 4. Sensitivity of instrumental performance to outcome devaluation. The mean number of lever presses per minute during the outcome devaluation test for the sham and mPFC group is shown. The data are plotted according to whether the action was trained with the outcome devalued at test (Deval) or the other outcome (Other). The vertical bars represent 1 SE of the difference between means across actions for each group.

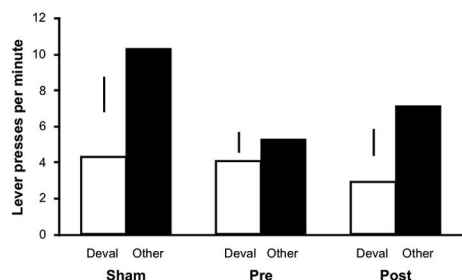


Figure 5. Sensitivity of instrumental performance to outcome devaluation. The mean number of lever presses per minute during the outcome devaluation test for groups sham, pre, and post is shown. The data are plotted according to whether the action was trained with the outcome devalued at test (Deval) or the other outcome (Other). The vertical bars represent 1 SE of the difference between means across actions for each group.

rats were given multiple days of instrumental retraining after surgery, making an interpretation of the results difficult. Experiment 2 had two aims: (1) to replicate the effects of posttraining mPFC lesions on outcome devaluation and reinstatement without providing additional training between surgery and testing and (2) to contrast directly the effects of these tests on performance in groups of pretraining- and posttraining-lesioned rats.

Experiment 2: effect of pretraining versus posttraining mPFC lesions on reinstatement and outcome devaluation

Outcome devaluation testing

Figure 5 shows the results of the outcome devaluation test in experiment 2. Although pretraining mPFC lesions disrupted the sensitivity of instrumental performance to outcome devaluation, lesions made after training failed to have this effect. An ANOVA revealed a significant main effect of response ($F_{(1,28)} = 32.30$; $p < 0.001$), a response-by-group interaction ($F_{(2,28)} = 4.17$; $p < 0.05$), but no overall effect of group ($F_{(2,28)} = 2.99$; $p = 0.07$). Additional analysis indicated that both group sham ($F_{(1,9)} = 17.45$; $p < 0.01$) and group post ($F_{(1,10)} = 22.23$; $p < 0.001$) performed fewer responses for the devalued outcome than for the other outcome, whereas group pre showed no effect of response ($F_{(1,9)} = 1.16$; $p > 0.05$). Therefore, the impact of mPFC damage on outcome devaluation performance appears to critically depend on the time of surgery, suggesting that the contribution of the mPFC to goal-directed action is limited to the acquisition, but not the storage or the expression, of action–outcome learning.

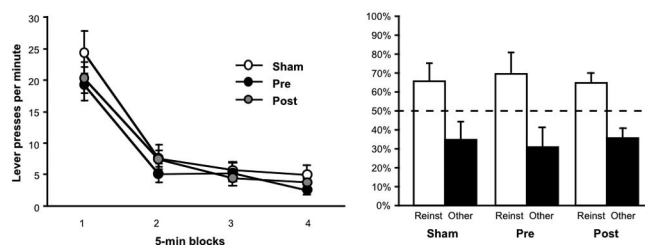


Figure 6. Extinction and selective reinstatement of instrumental performance. Left, The mean number of lever presses per minute (± 1 SEM) during consecutive 5 min blocks of extinction for the sham and mPFC groups. Right, The mean percentage of total responses (± 1 SEM) made on each lever after the noncontingent outcome delivery for groups sham, pre, and post. The data are plotted according to whether the action was trained with the outcome delivered during reinstatement (Reinst) or the other outcome (Other).

Reinstatement testing

The data from the extinction phase of reinstatement testing are presented in Figure 6 (left). The extinction procedure was equally effective in suppressing the instrumental performance of each group. An ANOVA performed on these data found a main effect of block ($F_{(3,84)} = 98.73$; $p < 0.0001$) but no effect of group ($F_{(2,28)} = 2.51$; $p > 0.05$), response ($F < 1$; reinstated vs other), nor any interaction between these variables ($F < 1$).

The results of the reinstatement phase are presented in the right panel of Figure 6. Lesions of mPFC, regardless of whether they were made before or after training, were found generally to attenuate the overall rate of performance in the reinstatement test. Average total response rates during this test (± 1 SEM) were the following: group sham, 22.5 (± 4.3); group pre, 11.3 (± 3.4); group post, 10.2 (± 2.3). Nevertheless, the lesions clearly left intact the selectivity of restatement on choice performance. An ANOVA conducted on the response per minute data revealed a main effect of group ($F_{(2,28)} = 4.10$; $p < 0.05$) and of response ($F_{(1,28)} = 14.45$; $p < 0.001$) but no group-by-response interaction ($F < 1$). *Post hoc* analysis of these data using Fisher's PLSD indicated that both group pre ($p < 0.05$) and group post ($p < 0.05$) made fewer total responses than group sham but did not differ from each other ($p > 0.05$). The preservation of selective reinstatement after mPFC damage is particularly clear in Figure 6 (right), which presents the test data as the mean percentage of total responses for the reinstating outcome (choice performance). All groups increased their choice of the reinstated action relative to their baseline choice performance during extinction, which was 50.5% (± 3.1) for group sham, 51.3% (± 5.0) for group pre, and 46.3% (± 2.9) for group post. An ANOVA found an effect of test phase ($F_{(1,28)} = 10.28$; $p < 0.01$) but no effect of group ($F < 1$) and no phase-by-group interaction ($F < 1$), indicating that the influence of noncontingent outcome delivery on response selection was similar across groups.

Although there is considerable evidence that outcome devaluation performance is mediated by the action→outcome relationship to which the rats were exposed during training (Colwill and Rescorla, 1986; Rescorla and Colwill, 1989), the associative structure underlying reinstatement performance appears, rather, to be mediated by the ability of the outcome to prime the response through an outcome→action association (for discussion, see Corbit and Balleine, 2003; Dickinson and de Wit, 2003). In experiment 3, we conducted a more direct investigation of this issue by assessing the sensitivity of instrumental reinstatement to devaluation of the reinstating outcome in both sham and pretraining mPFC-lesioned rats. If the outcome produces reinstatement by acting as a goal of the reinstated action, then, in sham

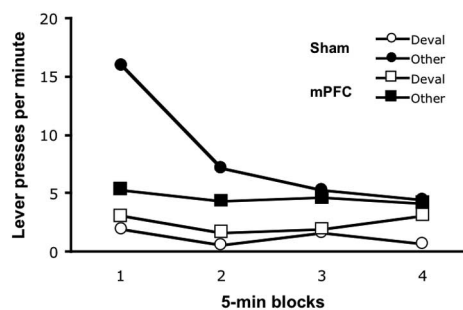


Figure 7. Extinction (baseline) during devalued reinstatement testing. The mean number of lever presses per minute during consecutive 5 min blocks of extinction for the sham and mPFC group. The data are plotted according to whether the action was trained with the outcome devalued at test (Deval) or the other outcome (Other). Note that the devalued outcome was also used to reinstate instrumental performance (see Fig. 8).

rats, this effect should depend on the current value of the goal; devaluing the outcome should reduce the ability of the outcome to reinstate performance of its associated action. If this is true, reinstatement in pretraining mPFC-lesioned rats should be expected to differ from shams and to depend solely on the discriminative properties of the instrumental outcome. Such a result would confirm that, whereas in the normal case, the mPFC functions to allow rats to respond flexibly with respect both to the sensory and emotional properties of goals, mPFC lesions renders them insensitive to the latter aspect and allows response selection only on the basis of the sensory properties of the goal.

Experiment 3: effect of pretraining mPFC lesions on the sensitivity of reinstatement to the devaluation of the reinstating outcome

As shown in Figure 7, the groups displayed different patterns of responding during the extinction phase of the reinstatement test; whereas the sham group performed the action that led to the devalued outcome at a much lower rate than the other action, the mPFC group performed both actions at a low rate. An ANOVA revealed a main effect of response ($F_{(1,13)} = 39.72; p < 0.0001$) and block ($F_{(3,39)} = 9.23; p < 0.0001$) but no group effect ($F_{(1,13)} = 1.09; p < 0.05$). The ANOVA also found a significant group-by-block interaction ($F_{(3,39)} = 5.45; p < 0.01$), response-by-block interaction ($F_{(3,39)} = 8.92; p < 0.001$), group-by-response interaction ($F_{(1,13)} = 11.09; p < 0.01$), as well as a group-by-response-by-block interaction ($F_{(3,39)} = 7.71; p < 0.001$). Additional analysis revealed the source of the three-way interaction: whereas the groups responded for the devalued outcome at similarly low rates throughout extinction (group, $F_{(1,13)} = 2.48, p > 0.05$; block, $F_{(4,52)} = 2.69, p < 0.05$; group-by-block, $F < 1$), the sham group responded more for the nondevalued (other) outcome early, but not late, in the session (group, $F_{(1,13)} = 4.15, p = 0.06$; block, $F_{(4,52)} = 8.08, p < .0001$; group-by-block, $F_{(4,52)} = 6.09, p < 0.001$).

The results of the reinstatement test phase are presented in Figure 8. As in previous tests, the sham group responded at a generally higher rate in the reinstatement test than the mPFC lesioned rats [i.e., $8.3 (\pm 1.8)$ and $4.8 (\pm 1.0)$, respectively]. Nevertheless, whereas the lesioned group performed more responses on the lever that, in training, had delivered the reinstating outcome than on the other lever, the sham group showed, if anything, the opposite pattern of responding. This pattern is not surprising, however, given that the groups displayed different baseline rates of responding on the two actions during extinction. Thus, the effect of the devalued outcome delivery on response

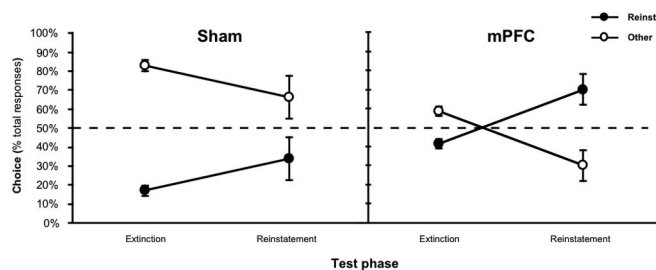


Figure 8. Impact of a devalued, noncontingently delivered outcome on response selection. The mean percentage of total responses made on each lever before (Extinction) and after (Reinstatement) the noncontingent outcome delivery for the sham group (left) and mPFC group (right). The data are plotted according to whether the action was trained with the outcome delivered during reinstatement (Reinst) or other outcome (Other). Note that the reinstating outcome had been devalued immediately before the test. Error bars represent SEM.

selection is better characterized by the shift in choice performance across test phases, as presented in Figure 8. As is clear from this figure, the sham group was, in general, less likely to choose the lever that, in training, delivered the devalued outcome than the mPFC group. Both groups, however, increased their choice of the devalued outcome after it had been noncontingently delivered. An ANOVA conducted on these data revealed a main effect of test phase ($F_{(1,13)} = 8.60; p < 0.05$) and group ($F_{(1,13)} = 18.51; p < 0.001$) but no phase-by-group interaction ($F < 1$). Thus, despite the fact that the reinstating outcome had been devalued before testing, it retained its capacity to guide response selection. This was true regardless of whether instrumental performance was sensitive to outcome devaluation, as in the case of the sham group, or not, as in the case of the mPFC group, indicating that the impact of a noncontingent outcome on response selection was mediated predominately by the sensory features of the outcome representation and not by its incentive properties.

Discussion

The current results indicate that the mPFC plays a stage-dependent role in instrumental conditioning. Consistent with previous reports (Balleine and Dickinson, 1998; Corbit and Balleine, 2003; Killcross and Coutureau, 2003), mPFC lesions made before initial training disrupted the sensitivity of instrumental performance to a reduction in expected outcome value. In contrast, lesions made after training had no effect on outcome devaluation. These results, together with the finding that posttraining lesions failed to disrupt the performance of previously acquired lever pressing, suggest that the mPFC is primarily involved in the acquisition and not expression of goal-directed, instrumental learning.

In contrast to its role in instrumental conditioning, the mPFC does not appear to be necessary for encoding the stimulus–outcome associations that underlie appetitive pavlovian conditioning (Corbit and Balleine, 2003). This pattern of results markedly contrasts with recent reports on the role of the orbitofrontal cortex (OFC) in appetitive learning (Holland and Gallagher, 2004). For instance, OFC lesions made either before or after pavlovian conditioning were shown to disrupt the normal sensitivity of conditioned approach behavior to unconditioned stimulus (US) devaluation (Gallagher et al., 1999; Pickens et al., 2005), suggesting that this structure is necessary for the storage and/or expression of stimulus-mediated outcome expectancies. Moreover, although additional research will be needed to determine what, if any, role the OFC plays in instrumental learning, the available

data tentatively suggests that it is selectively involved in pavlovian learning (Chudasama and Robbins, 2003).

Interestingly, although lesions of the mPFC made either before or after initial training reduced the overall magnitude of instrumental reinstatement, they left intact the selectivity of this effect, confirming a previous observation that mPFC lesions spared both outcome- and stimulus-mediated response priming (Corbit and Balleine, 2003) (experiments 2 and 3). Thus, the current results indicate that the influence of a noncontingent outcome presentation on response initiation is dissociable from its influence on response selection, with the mPFC playing a clear and potentially critical role in the former but not the latter. Support for this conclusion can be found in several recent reports on the reinstatement of drug-seeking behavior. For example, the mPFC has been implicated in drug-, cue-, stress-, and acute food deprivation-induced reinstatement of drug seeking (McFarland and Kalivas, 2001; Park et al., 2002; See, 2002; Capriles et al., 2003; Shalev et al. 2003), suggesting that it is a critical structure in the general motor output pathway mediating reinstatement of performance (McLaughlin and See, 2003; McFarland et al. 2004). Unfortunately, in most self-administration studies (but see Leri and Stewart, 2001), reinstatement is compared across responses that differ dramatically in their baseline rates of performance, making it difficult to determine whether the impact of noncontingent drug delivery on response selection remains intact after disruption of mPFC function. Such a finding would surely have important implications for the reinstatement model of drug relapse (Shaham et al., 2003).

Finally, to examine further the response selection strategy used by rats with pretraining mPFC lesions, we assessed the sensitivity of their reinstatement performance to devaluation of the noncontingent outcome. Whereas the response selectivity of reinstatement performance was sensitive to outcome devaluation in shams, this was not true for the lesioned group, which displayed robust selective reinstatement for a devalued outcome. This finding is consistent with the claim that whereas the reinstatement performance of mPFC-lesioned rats was mediated solely by the discriminative stimulus properties of the outcome, presumably through an outcome→action association, the performance of sham rats was guided both by these discriminative properties and by the incentive properties of the expected outcome, the latter influencing performance through the action→outcome association (Colwill and Rescorla, 1986; Rescorla and Colwill, 1989).

Previous research suggests that instrumental performance can be supported by either of two distinct learning systems, each characterized by the unique associations that it supports (Dickinson, 1989; Balleine and Dickinson, 1998). The goal-directed system is thought to encode the instrumental relationships that exist between individual actions and their respective outcomes. According to this account, response selection and initiation is a product of both instrumental learning and the incentive value of the outcome anticipated as a consequence of that action. Alternatively, the habit system is thought to support the acquisition of instrumental performance by associating individual responses with stimuli that prevailed during training and, hence, primarily reflects the formation stimulus–response associations. Rewards are considered to play only a reinforcing function in habit learning (i.e., the instrumental outcome acts to strengthen stimulus–response associations but it does not become encoded as a goal).

The disruptive effects of pretraining mPFC lesions on sensitivity to devaluation and contingency degradation indicate that this structure is a critical component of the goal-directed learning

system that normally mediates instrumental performance. On its own, this finding is consistent with the notion that the mPFC is responsible for encoding action–outcome associations. However, if these associations are permanently stored in the mPFC, then posttraining lesions of this structure should (1) disrupt the maintenance of previously acquired instrumental performance and (2) leave any residual responding insensitive to outcome devaluation. Contrary to these predictions, we found that posttraining lesions had no detectable effect on the maintenance of lever pressing or the sensitivity of instrumental performance to a reduction in outcome value. Therefore, although the mPFC is clearly required for the acquisition of goal-directed actions, it does not appear to permanently store this information.

This finding is consistent with the theory, recently elaborated on by Miller and Cohen (2001), that, early in training, the prefrontal cortex is critical for keeping active the representations of individual actions and their outcomes, thereby providing information about specific action–outcome contingencies to other structures capable of storing this information in the long-term. Thus, after extended training, it is proposed that storage and expression of action–outcome memories becomes independent of the prefrontal cortex. The current results, therefore, favor the view that instrumental learning is supported by a distributed goal-directed learning system that involves the mPFC and other, presumably closely connected, brain structures. Several candidates have been identified by recent anatomical and behavioral research. The basolateral amygdala (BLA), for instance, shares rich reciprocal connections with the mPFC (Krettek and Price, 1977; Cassell and Wright, 1986). Furthermore, pretraining lesions of the BLA result in many of the same behavioral deficits as do those of the mPFC. For example, BLA lesions have been shown to disrupt both outcome devaluation and contingency degradation performance (Balleine et al., 2003; Corbit and Balleine, 2005). There is, however, some evidence that the involvement of the BLA in action–outcome learning is limited to the acquisition of reward representations. For instance, it has been shown recently that intra-BLA protein synthesis inhibition disrupts the consolidation (and reconsolidation) of incentive learning, the process whereby animals update changes in reward value (Wang et al., 2005). Moreover, BLA lesions made before appetitive pavlovian conditioning attenuate the impact of US devaluation on conditioned approach behavior (Hatfield et al., 1996; Blundell et al., 2003), suggesting that it may play a general role in attaching incentive value to stimuli (Baxter and Murray, 2002; Holland and Gallagher, 2004).

The mPFC, however, also projects to several discrete regions of the striatum. For example, both the dorsomedial striatum (DMS) and ventral striatum, particularly the core of the nucleus accumbens, receive afferents from the mPFC (McGeorge and Faull, 1989; Berendse et al. 1992). Furthermore, amygdalostratial projections from BLA show considerable overlap with corticostriatal projections from mPFC (McDonald, 1991), placing the DMS and accumbens core in a prime position to integrate information about reward value with action–outcome information from the mPFC. There is also strong evidence from behavioral studies that these structures contribute to goal-directed action. For example, pretraining lesions of the core have been shown to disrupt rats' sensitivity to outcome devaluation but do not impair contingency degradation learning, suggesting that this area is involved in modulating performance according to the incentive value of expected outcomes but is not necessary for action–outcome encoding (Corbit et al., 2001). The DMS, however, appears to play a more substantial role in goal-directed learning. For ex-

ample, both pretraining and posttraining DMS lesions, as well as muscimol-induced inactivation of this region, impair the sensitivity of rats to instrumental outcome devaluation and degradation of the instrumental contingency (Yin et al., 2005b), indicating that it, unlike the mPFC, plays a relatively long-lasting role in the expression of action–outcome learning. Furthermore, it has been shown recently that the infusion of APV, a selective NMDA receptor antagonist, into the DMS during instrumental learning blocks the acquisition of new action–outcome learning (Yin et al., 2005a). Interestingly, it appears that feedback from the DMS to the prefrontal cortex might be necessary for goal-directed instrumental learning because lesions of the mediodorsal thalamus, a likely hub for this feedback (Nauta, 1989), disrupt outcome devaluation performance and contingency degradation learning (Corbit et al., 2003). These findings, together with the results reported here, lead to the intriguing hypothesis that action–outcome learning depends on functional interaction between the mPFC and DMS during the early stages of training.

References

- Balleine BW, Dickinson A (1998) Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37:407–419.
- Balleine BW, Killcross AS, Dickinson A (2003) The effect of lesions of the basolateral amygdala on instrumental conditioning. *J Neurosci* 23:666–675.
- Baxter MG, Murray EA (2002) The amygdala and reward. *Nat Rev Neurosci* 3:563–573.
- Berendse HW, Galis-de Graaf Y, Groenewegen HJ (1992) Topographical organization and relationship with ventral striatal compartments of prefrontal corticostriatal projections in the rat. *J Comp Neurol* 316:314–347.
- Blundell P, Hall G, Killcross S (2003) Preserved sensitivity to outcome value after lesions of the basolateral amygdala. *J Neurosci* 23:7702–7709.
- Capriles N, Rodaros D, Sorge RE, Stewart J (2003) A role for the prefrontal cortex in stress- and cocaine-induced reinstatement of cocaine seeking in rats. *Psychopharmacology* 168:66–74.
- Cassell MD, Wright DJ (1986) Topography of projections from the medial prefrontal cortex to the amygdala in the rat. *Brain Res Bull* 17:321–333.
- Chudasama Y, Robbins TW (2003) Dissociable contributions of the orbitofrontal and infralimbic cortex to pavlovian autoshaping and discrimination learning: further evidence for the functional heterogeneity of the rodent frontal cortex. *J Neurosci* 23:8871–8880.
- Colwill RM, Rescorla RA (1985) Postconditioning devaluation of a reinforcer affects instrumental responding. *J Exp Psychol Anim Behav Process* 11:120–132.
- Colwill RM, Rescorla RA (1986) Associative structures in instrumental conditioning. In: *The psychology of learning and memory*, Vol XX (Bower GH, ed), pp 55–104. New York: Academic.
- Corbit LH, Balleine BW (2003) The role of prelimbic cortex in instrumental conditioning. *Behav Brain Res* 146:145–157.
- Corbit LH, Balleine BW (2005) Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer. *J Neurosci* 25:962–970.
- Corbit LH, Muir JL, Balleine BW (2001) The role of the nucleus accumbens in instrumental conditioning: evidence of a functional dissociation between accumbens core and shell. *J Neurosci* 21:3251–3260.
- Corbit LH, Muir JL, Balleine BW (2003) Lesions of mediodorsal thalamus and anterior thalamic nuclei produce dissociable effects on instrumental conditioning in rats. *Eur J Neurosci* 18:1286–1294.
- Dalley JW, Cardinal RN, Robbins TW (2004) Prefrontal executive and cognitive functions in rodents: neural and neurochemical substrates. *Neurosci Biobehav Rev* 28:771–784.
- Dickinson A (1989) Expectancy theory in animal conditioning. In: *Contemporary learning theories* (Klein SB, Mowrer RR, eds), pp 279–308. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Dickinson A, de Wit S (2003) The interaction between discriminative stimuli and outcomes during instrumental learning. *Q J Exp Psychol B* 56:127–139.
- Gallagher M, McMahan RW, Schoenbaum G (1999) Orbitofrontal cortex and representation of incentive value in associative learning. *J Neurosci* 19:6610–6614.
- Hatfield T, Han JS, Conley M, Gallagher M, Holland P (1996) Neurotoxic lesions of basolateral, but not central, amygdala interfere with Pavlovian second-order conditioning and reinforcer devaluation effects. *J Neurosci* 16:5256–5265.
- Holland PC, Gallagher M (2004) Amygdala-frontal interactions and reward expectancy. *Curr Opin Neurobiol* 14:148–155.
- Killcross S, Coutureau E (2003) Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb Cortex* 13:400–408.
- Krettek JE, Price JL (1977) Projections from the amygdaloid complex to the cerebral cortex and thalamus in the rat and cat. *J Comp Neurol* 172:687–722.
- Leri F, Stewart J (2001) Drug-induced reinstatement to heroin and cocaine seeking: a rodent model of relapse in polydrug use. *Exp Clin Psychopharmacol* 9:297–306.
- Matsumoto K, Tanaka K (2004) The role of the medial prefrontal cortex in achieving goals. *Curr Opin Neurobiol* 14:178–185.
- Matsumoto K, Suzuki W, Tanaka K (2003) Neuronal correlates of goal-based motor selection in the prefrontal cortex. *Science* 301:229–232.
- McDonald AJ (1991) Organization of amygdaloid projections to the prefrontal cortex and associated striatum in the rat. *Neuroscience* 44:1–14.
- McFarland K, Kalivas PW (2001) The circuitry mediating cocaine-induced reinstatement of drug-seeking behavior. *J Neurosci* 21:8655–8663.
- McFarland K, Davidge SB, Lapish CC, Kalivas PW (2004) Limbic and motor circuitry underlying footshock-induced reinstatement of cocaine-seeking behavior. *J Neurosci* 24:1551–1560.
- McGeorge AJ, Faull RL (1989) The organization of the projection from the cerebral cortex to the striatum in the rat. *Neuroscience* 29:503–537.
- McLaughlin J, See RE (2003) Selective inactivation of the dorsomedial prefrontal cortex and the basolateral amygdala attenuates conditioned-cued reinstatement of extinguished cocaine-seeking behavior in rats. *Psychopharmacology* 168:57–65.
- Miller EK, Cohen JD (2001) An integrative theory of prefrontal cortex function. *Annu Rev Neurosci* 24:167–202.
- Mulder AB, Nordquist RE, Orgut O, Pennartz CM (2003) Learning-related changes in response patterns of prefrontal neurons during instrumental conditioning. *Behav Brain Res* 146:77–88.
- Nauta WJH (1989) Reciprocal links of the corpus striatum with the cerebral cortex and limbic system: a common substrate for movement and thought? In: *Neurology and psychiatry: a meeting of minds* (Mueller J, ed), pp 43–63. Basel: Karger.
- Park WK, Bari AA, Jey AR, Anderson SM, Speelman RD, Rowlett JK, Pierce RC (2002) Cocaine administered into the medial prefrontal cortex reinstates cocaine-seeking behavior by increasing AMPA receptor-mediated glutamate transmission in the nucleus accumbens. *J Neurosci* 22:2916–2925.
- Paxinos G, Watson C (1998) *The rat brain in stereotaxic coordinates*. San Diego: Academic.
- Pickens CL, Saddoris MP, Gallagher M, Holland PC (2005) Orbitofrontal lesions impair use of cue–outcome associations in a devaluation task. *Behav Neurosci* 119:317–322.
- Rescorla RA, Colwill RW (1989) Associations with anticipated and obtained outcomes in instrumental learning. *Anim Learn Behav* 17:291–303.
- See RE (2002) Neural substrates of conditioned-cued relapse to drug-seeking behavior. *Pharmacol Biochem Behav* 71:517–529.
- Shaham Y, Shalev U, Lu L, De Wit H, Stewart J (2003) The reinstatement model of drug relapse: history, methodology and major findings. *Psychopharmacology* 168:3–20.
- Shalev U, Roberts P, Shaham Y, Morales M (2003) Selective induction of c-Fos immunoreactivity in the prelimbic cortex during reinstatement of heroin seeking induced by acute food deprivation in rats. *Behav Brain Res* 145:79–88.
- Wang SH, Ostlund SB, Nader K, Balleine BW (2005) Consolidation and reconsolidation of incentive learning in the amygdala. *J Neurosci* 25:830–835.
- Yin HH, Knowlton BJ, Balleine (2005a) Blockade of NMDA receptors in the dorsomedial striatum prevents action–outcome learning in instrumental conditioning. *Eur J Neurosci* 22:502–512.
- Yin HH, Ostlund SB, Knowlton BJ, Balleine (2005b) The role of the dorsomedial striatum in instrumental conditioning. *Eur J Neurosci* 22:513–523.