European Journal of Neuroscience, Vol. 22, pp. 513–523, 2005 © Federation of European Neuroscience Societies

The role of the dorsomedial striatum in instrumental conditioning

Henry H. Yin, Sean B. Ostlund, Barbara J. Knowlton and Bernard W. Balleine
Department of Psychology and Brain Research Institute, University of California, Los Angeles, Box 951563, Los Angeles, CA 90095–1563, USA

Keywords: basal ganglia, contingency, goal-directed action, learning, reward

Abstract

Considerable evidence suggests that, in instrumental conditioning, rats learn the relationship between actions and their specific consequences or outcomes. The present study examined the role of the dorsomedial striatum (DMS) in this type of learning after excitotoxic lesions and reversible, muscimol-induced inactivation. In three experiments, rats were first trained to press two levers for distinct outcomes, and then tested after training using a variety of behavioural assays that have been established to detect action-outcome learning. In Experiment 1, pre-training lesions of the posterior DMS abolished the sensitivity of rats' instrumental performance to both outcome devaluation and contingency degradation when tested in extinction, whereas lesions of the anterior DMS had no effect. In Experiment 2, both pre-training and post-training lesions of the posterior DMS were equally effective in reducing the sensitivity of performance both to devaluation and degradation treatments. In Experiment 3, the infusion of muscimol into the posterior DMS selectively abolished sensitivity of performance to devaluation and contingency degradation without impairing the ability of rats to discriminate either the instrumental actions performed or the identity of the earned outcomes. Taken together, these results suggest that the posterior region of the DMS is a crucial neural substrate for the acquisition and expression of action–outcome associations in instrumental conditioning.

Introduction

Studies of instrumental conditioning have established that, in rats, actions that are instrumental in gaining access to rewards, such as lever pressing for food, can be controlled by two distinct associative processes. During initial acquisition, actions appear to be goal-directed and mediated by the encoding of an association between the action and its specific consequences or outcome. Thus instrumental performance has been found to be sensitive both to non-contingent reward delivery and to post-training changes in outcome value (Adams & Dickinson, 1981; Colwill & Rescorla, 1986; Dickinson & Balleine, 1994, 2002; Balleine, 2001). After a period of training, however, control over performance has been found to shift to a stimulus-response process and, as a consequence, actions become stimulus-bound or habitual, and no longer sensitive to changes in either the instrumental contingency or reward value (Dickinson & Balleine, 1993, 1995; Dickinson *et al*., 1995; Balleine & Dickinson, 1998).

Many studies have suggested that the associative and the sensorimotor regions of the dorsal striatum may likewise play distinct roles in instrumental learning, with the former involved in the formation of action–outcome associations and the latter in stimulus–response (S–R) associations (Graybiel, 1998; Hikosaka *et al*., 2000). For example, electrophysiological studies measuring neural activity in the associative striatum or caudate nucleus in primates, the homologue of the dorsomedial striatum (DMS) in rats, have reported that activity correlated with the performance of skilled movements can be modulated by the expectancy of reward (Kawagoe *et al*., 1998; Hassani *et al*., 2001). In contrast, recordings from the sensorimotor striatum (i.e. roughly the dorsolateral striatum, DLS, in rats or putamen in primates), have generally failed to find neural activity specifically correlated with outcome expectancy (Kimura, 1990, 1992; Kimura *et al*., 1992; Jaeger *et al*., 1993; White & Rebec, 1993; Carelli *et al*., 1997; Jog *et al*., 1999).

In rats, several studies have suggested that habitual instrumental performance is mediated by the DLS (McDonald & White, 1993; Packard & McGaugh, 1996; reviewed in Packard & Knowlton, 2002). Indeed, using sensitivity to outcome devaluation as an index, we recently reported that lesions of the DLS disrupt S–R learning whilst leaving action–outcome learning intact (Yin *et al*., 2004). In contrast, in the same study we found that lesions of the DMS did not affect habit learning. No study, however, has directly assessed the role of the DMS in the performance of goal-directed instrumental actions using behavioural assays of sufficient power to elucidate the precise role of this region.

In the present study we assessed the role of the dorsal striatum in goal-directed action by evaluating the effects of selective pre-training (Experiment 1) and pre- and post-training (Experiment 2) excitotoxic lesions as well as reversible inactivation (Experiment 3) of the DMS, on the acquisition and expression of action–outcome learning in rats. In these experiments action–outcome learning was indexed by the rats' sensitivity to degradation of the instrumental contingency, outcome devaluation, outcome-mediated reinstatement, and performance on a heterogeneous chain of actions. Anatomical studies have shown distinct connectivity for the medial, associative striatum compared to the lateral, sensorimotor striatum (Cheatwood *et al*., 2003; Reep *et al*., 2003) but there appears to be considerable heterogeneity within the rat

Correspondence: Dr Bernard Balleine, as above.
E-mail: balleine@psych.ucla.edu

Received 10 January 2005, revised 1 March 2005, accepted 16 March 2005

doi:10.1111/j.1460-9568.2005.04218.x

dorsal striatum along the anterior–posterior axis as well (Kelley *et al*., 1982; Nauta, 1989). Therefore, in this study we also assessed the relative contributions of the anterior dorsomedial striatum (aDMS) and posterior dorsomedial striatum (pDMS) by comparing the effects of pre- and post-training lesions of these structures in Experiments 1 and 2.

Materials and methods

Experiment 1: Effect of anterior and posterior DMS lesions on outcome devaluation and contingency degradation

Subjects and apparatus

Twenty-six experimentally naïve male Long–Evans rats weighing between 420 and 530 g were used. The UCLA Animal Research Committee approved the study. The rats were housed singly and handled daily for one week prior to surgery. Training and testing took place in 16 Medical Associates (East Fairfield, VT) operant chambers housed within sound- and light-resistant walls. Each chamber was equipped with two retractable levers on either side of the food magazine, a pellet dispenser that delivered 45 mg Noyes pellets (formula A/I), and a pump with a syringe that delivered 0.1 mL of 20% sucrose solution into a recessed magazine in the chamber. A 3 W, 24 V house-light, mounted on the top-centre of the wall opposite the magazine, provided illumination. Computers equipped with the MED-PC program (Medical Associates, VT) controlled the equipment and recorded the lever presses.

Surgery and histology

Rats were anaesthetized with sodium pentobarbital (Nembutal; 50 mg/kg), treated with atropine (0.1 mg), and placed in a stereotaxic instrument (UCLA Animal Research Committee). Small holes were drilled into the skull bilaterally, and 28 gauge cannulae were lowered into the brain at the following coordinates: aDMS ($n = 8$), 1 mm anterior, 1.7 mm lateral to bregma, and 5 mm below skull surface; pDMS ($n = 10$), 0.4 mm posterior, 2.6 mm lateral to bregma, and 4.5 mm below skull surface. The sham group ($n = 8$) had holes drilled at the aDMS coordinates but received no NMDA infusion. For all lesions, the same concentration of NMDA (0.12 M; Sigma, St. Louis, MO) was used, and the volume infused was 0.4 μL per side over 3 min. Three minutes after the infusion, the cannulae were removed. At the end of the experiment, the rats were killed using a lethal barbiturate overdose and perfused transcardially with 0.9% saline followed by 10% formaldehyde solution. The brains were stored in a 25% sucrose-formalin solution for at least three days before 50-μm-coronal sections were cut throughout the anterior striatum. The slices were stained with thionin and examined with a microscope for areas of cell loss as well as shrinkage in the striatum and surrounding regions.

Training

Five days after surgery, rats were placed on a food deprivation schedule, receiving 10–15 g of their maintenance diet daily to reduce their weight to approximately 80% of their free-feeding weight. Once training began, they were fed each day after the training sessions, and had free access to water while in their home cages.

Ten days after surgery, rats were given two 30-min magazine training sessions in which the sucrose solution and the pellets were delivered on a random time 60-s schedule (once a minute on average) with the levers removed. The next day lever-press training began. On each day of lever-press training, all rats were given two 30-min sessions, one on each lever. Each training session began with the illumination of the house light and insertion of the lever, and ended with the retraction of the lever and turning off of the house light. For half of the rats in each group, the left lever earned pellets and the right lever sucrose. The other half received the opposite lever-outcome pairings. There was at least a one-hour break between the two training sessions with the order of sucrose and pellet sessions alternated each day. Progressively leaner schedules of reinforcement were used: continuous reinforcement (CRF) for 2 days then random ratio-5 for 2 days (RR-5; each response was rewarded at a probability of 0.2 on average), RR-10 for 2 days and finally RR-20 for 2 days.

Outcome devaluation

After the last day of training, all rats were given free exposure to one of the outcomes for 1 h in the home cage. Half of the rats in each lever-outcome assignment received 20 g of pellets in a bowl, and the remaining rats received sucrose (30 mL in a drinking bottle). Immediately after the pre-feeding session, a 5-min extinction test was given. The test began with the illumination of the house light and insertion of both levers at the same time, and ended with the retraction of the levers and the offset of the house light. The number of presses on each lever was recorded.

Contingency degradation

After the devaluation tests, the rats received 2 days of retraining on RR-20 schedules, followed by degradation training in which, for each rat, one instrumental outcome, either the pellets or the sucrose, was delivered non-contiguously such that its probability of delivery in each second of the training session was equally likely if the rats responded appropriately or not (for further details see Balleine & Dickinson, 1998; Balleine *et al*., 2003). For half of the rats, the response–pellet contingency was degraded, and for the other half the response–sucrose contingency was degraded. Two 20-min sessions were given each day, one on each lever with a break of at least 1 h between sessions. The order of the sessions was alternated. After 4 days of training, the rats received a 5-min choice extinction test on the two levers as the primary test of the effects of contingency degradation training. Both contingent and non-contingent rewards were omitted in this test in order to ensure responding was based on learning that had occurred during the training sessions and not new learning during the test session.

Experiment 2: Effect of pre- and post-training lesions of aDMS and pDMS on outcome devaluation and contingency degradation

Experiment 2 was designed to compare the effects of aDMS and pDMS lesions made either before or after instrumental training on the same tests as described in Experiment 1. To achieve this we replicated the pre-training lesion treatments in Experiment 1 and extended these findings by adding two further groups given post-training lesions of the aDMS and pDMS. As the aDMS and sham groups did not differ in Experiment 1, the effects of the various lesions were assessed against the former group.

Subjects and apparatus

The subjects were 30 experimentally naïve male Long–Evans rats, group-housed as in Experiment 1. Of these, six rats received aDMS lesions before training, seven rats received pDMS lesions before training, eight rats received aDMS lesions after training, and nine rats received pDMS lesions after training.

Training and devaluation

All rats were given instrumental training exactly as described in Experiment 1. Rats in the post-training groups then were given lesions of either the aDMS or pDMS as described above. After recovery, all rats were given an outcome devaluation test exactly as described in Experiment 1.

Contingency degradation

Four days after the devaluation test, all rats were retrained on a RR-20 schedule for four days, followed by five days of degradation training and a 5-min extinction test on the sixth day, using procedures identical to those used in Experiment 1.

Experiment 3: The effect of muscimol infusion into the pDMS on instrumental performance

Subjects and apparatus

Sixteen naïve female Long–Evans rats weighing between 240 and 300 g were used in Experiment 3. They were housed singly and handled daily for two days prior to surgery.

Surgery and histology

Small holes were drilled into the skull bilaterally, and 28 gauge cannulae were lowered into the brain at the following coordinates: 0.4 mm posterior and 2.6 mm lateral to bregma, and 4.5 mm below skull surface and held in place using dental acrylic. Otherwise the surgical procedures were the same as described above.

Training and devaluation

All rats were given instrumental training as described in Experiment 1 using progressively leaner ratio schedules of reinforcement, i.e. CRF, RR-5, RR-10, and RR-20. One day after the last day of RR-20 training, all rats were give free exposure to one of the outcomes for 1 h in their home cages. Half of the rats in each lever–outcome assignment received 20 g of pellets in a bowl, and the remaining rats received 30 mL of sucrose in a drinking bottle. Immediately after the pre-feeding session, the rats received infusions of either muscimol ($n = 8$) or artificial cerebral spinal fluid (ACSF, $n = 8$). The dummy cannulae were removed and injection cannulae (26 gauge; Plastics One, VA) were lowered into the guide cannulae extending 0.5 mm below the guide tip. The injection cannulae were connected by polyethylene tubing to 10-μL Hamilton syringes mounted on an infusion pump (Harvard Apparatus, USA). Muscimol (Sigma, USA; 1 μg per μL dissolved in ACSF; Sigma, USA) was delivered at a rate of 0.25 μL per min for 1 min, with a total volume of 0.25 μL per side. The same volume of ACSF was used as the control infusion. One minute after infusion, the injection cannulae were removed and the dummy cannulae replaced.

Five minutes after the infusion, a 10-min extinction test was given. The test began with the illumination of the house light and insertion of both levers and ended with the retraction of the levers and the offset of the house light. The number of presses on each lever was recorded. Immediately after the extinction test a 20-min rewarded test was given, in which actions on the two levers earned their respective outcomes as during training, but on a progressive ratio schedule with the ratio required to earn each reward increasing from CRF, to RR-5, RR-10 and finally to RR-20 for the remainder of the session.

Contingency degradation

All rats were retrained on a RR-20 schedule for one day, followed by three days of degradation training (see above). During this training three rats lost their cannulae assembly and were dropped from this phase leaving 13 rats for this experiment. All rats were given three days of contingency degradation training after which they received a 5-min choice extinction test immediately after an infusion of either muscimol ($n = 8$) or ACSF ($n = 5$).

Outcome-specific reinstatement

The next day, without additional training, all rats received a reinstatement session after infusions of either muscimol ($n = 7$) or ACSF ($n = 6$) counterbalanced with respect to the treatment on the previous day. For the first 20 min, responding was extinguished on both levers concurrently, i.e. the levers were extended but responses on them were not reinforced. After this period, a single delivery of one of the two instrumental outcomes was given into the food magazine and the effect of this outcome delivery on performance assessed for the next 2 min.

Action discrimination using a heterogeneous chain

Finally, three days after the reinstatement test, 12 rats were trained on a heterogeneous chain of instrumental actions for a food pellet reward (one further rat was dropped due to a loosened cannulae assembly). The instrumental chain consists of two lever-press responses: pressing on the left lever (distal response) followed by pressing on the right lever (proximal response), which earned the pellet. To train this chain, the rats first received one session of lever-press training on the right lever on a continuous reinforcement schedule, which ended after 30 pellets had been earned. The next day all rats were trained on the left lever such that pressing the left lever caused the insertion of the right lever, the pressing of which earned the pellet. They were then shifted to a RR-4 schedule on both levers, so that on average, four left lever responses were required to cause the appearance of the right lever and, on average, four further responses on the right, or proximal, lever earned the pellet.

After this second session performance on the chain was assessed in a test session conducted after an infusion of either muscimol ($n = 6$) or ACSF ($n = 6$) in which both levers were continuously present and with the pellet outcome delivered for responses on the left lever $\rightarrow$ right lever chain on an RR-4 schedule of reinforcement. During this session two conditional probabilities were examined: the probability of pressing the right lever in each second after the left lever had been pressed, i.e. $P(\text{LL} \rightarrow \text{RL})$, and the probability of pressing the left lever in each second after the right lever had been pressed, i.e. $P(\text{RL} \rightarrow \text{LL})$. This session ended after 30 reinforcers had been earned or after 20-min.

Results

Experiment 1: Dissociating the effects of anterior and posterior DMS lesions on outcome devaluation and contingency degradation

Figure 1 (right-hand panels) provides a schematic representation of the extent of damage to the striatum caused by NMDA infusions (Experiments 1 and 2). Although no volumetric analyses were made, the lesion placement was assessed by reconstructing the damaged area on standard stereotaxic atlas templates (Paxinos & Watson, 1998). Inspection of the stained tissue did not reveal damage outside of the dorsal striatum. Although the lesions were small, clear cell loss,

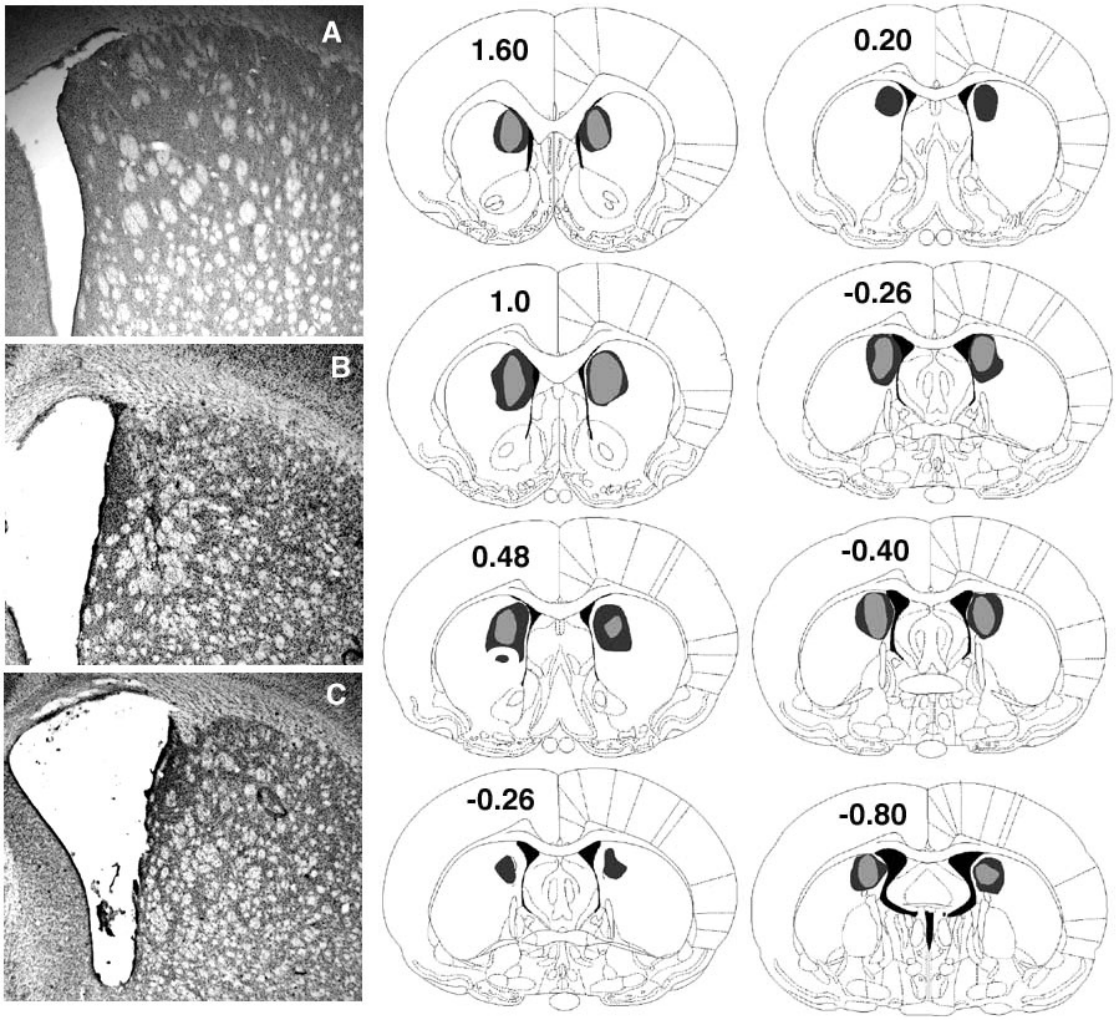

FIG. 1. Experiments 1 and 2. (A–C) Photomicrographs of representative lesions: (A) sham; (B) aDMS; (C) pDMS. Center and right-hand columns, schematic illustration of the NMDA lesions of the aDMS (centre), and the pDMS (right) in coronal sections (Paxinos & Watson, 1998); grey areas represent the smallest extent of damage, and black areas the largest extent of damage. Numbers indicate distance from bregma in mm.

gliosis, as well as shrinkage were seen in the targeted striatal areas just beneath the corpus callosum. All lesions were restricted to the medial striatum, within 3 mm to either side of midline. Representative photomicrographs of these lesions are shown in Fig. 1A–C. As can be clearly seen in this figure, the most notable feature of lesions of the DMS is a visible widening of the ventricles compared to sham operated controls, reflecting shrinkage due to cell loss in the target region.

Training and outcome devaluation

Lesions of the pDMS reduced instrumental performance. Although all of the rats learned to press the levers and increased their performance with increasing ratio requirements, their response rates were quite low compared to those in the aDMS and sham groups. An ANOVA conducted on the average performance on the two levers during the final training session revealed a significant main effect of lesion ($F_{2,23} = 19.1$, $P < 0.001$). *Posthoc* tests (Fisher's PLSD) revealed that performance was reduced in the pDMS group relative to both the aDMS ($P < 0.001$) and sham groups ($P < 0.001$), but the latter two groups did not differ ($P > 0.05$). The means from the final session of training are presented in Fig. 2.

The results from the devaluation test are shown in the right-hand panels of Fig. 2. Clearly, lever pressing in the sham group and the aDMS group was selectively reduced by the devaluation treatment. This sensitivity, however, was completely abolished in the pDMS group and the specific-satiety treatment appeared to have no differential effect on performance. A repeated-measures ANOVA using devaluation as a within-subject factor and lesion as a between-subjects factor showed a significant main effect of lesion ($F_{2,21} = 4.2$, $P < 0.05$), of devaluation ($F_{1,21} = 7.3$, $P < 0.01$), and an interaction between these factors ($F_{2,21} = 5.0$, $P < 0.05$). Simple main effects analyses revealed a significant devaluation effect in both the sham ($F_{1,\ 7} = 7.3$, $P < 0.05$) and the aDMS groups ($F_{1,\ 5} = 13.6$, $P < 0.05$), but no effect in the pDMS group ($F < 1$).

Contingency degradation

Performance on the final day of contingency degradation training and during the critical extinction test is presented in Fig. 3. The pattern of results from the extinction test conducted after contingency degradation training was very similar to that observed in the devaluation test. There was a main effect of degradation ($F_{2,21} = 8.5$, $P < 0.01$) but no main effect of lesion and, largely due to the three-group design, the

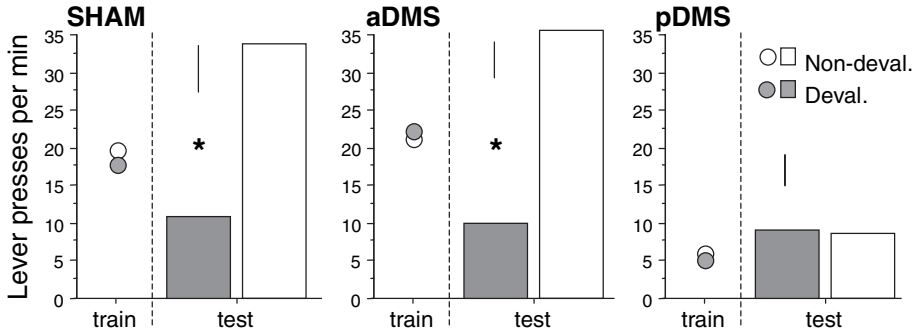

FIG. 2. Experiment 1. Responses rates during the 5-min devaluation test. Response rates on the actions that, in training, delivered the now devalued and non-devalued outcomes are shown both for the last day of training, before devaluation, on the left, and during the extinction test immediately after devaluation treatment, on the right, for the sham, aDMS and pDMS groups. ★ indicates $P < 0.05$. Each vertical line represents 1 SED (standard error of the difference of the means, a measure of within-subject variance).

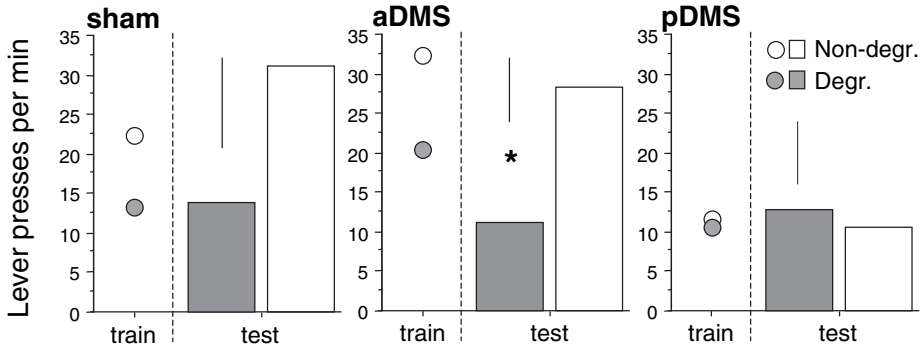

FIG. 3. Experiment 1. Response rates during the extinction test after contingency degradation training, shown separately for the action that had the action–outcome contingency degraded and that for which the contingency was not degraded. Left panel, last day of contingency degradation training; right panel, during the critical extinction test. ★ indicates $P < 0.05$. Each vertical line represents 1 SED (standard error of the difference of the means, a measure of within-subject variance).

lesion–degradation interaction was only marginally significant ($F_{2,21} = 2.67$, $P = 0.09$). Comparisons conducted across the within-subjects variable in each group revealed, nevertheless, that whereas responding on the degraded lever was significantly reduced compared to that on the control lever in both the sham ($F_{1,7} = 5.13$, $p = 0.058$) and the aDMS ($F_{1,5} = 13.9$, $P < 0.05$) groups, this comparison failed to reach significance in the pDMS group ($F_{1,9} = 2.56$, $P > 0.1$). Thus, rats with pDMS lesions also appeared to be relatively insensitive to selective degradation of the instrumental contingency.

Experiment 2: Effects of pre- and post-training lesions of aDMS and pDMS on outcome devaluation and contingency degradation

In this experiment we sought to replicate the effects of the pre-training lesions of the selective striatal areas found in Experiment 1, and to extend these findings by comparing them with post-training lesions of the same areas. Such a comparison is useful for two reasons. First, it helps to determine the locus of the effect; if only pre-training lesions are effective in abolishing sensitivity to devaluation, it would suggest that pDMS lesions affect the acquisition, but not the expression, of the learned behaviour; if, however, both pre- and post-training lesions are effective, it would suggest that the pDMS is involved in the expression as well as the acquisition of goal-directed behaviour. Second, lever-press rates were very low for rats receiving pre-training pDMS lesions, due to impaired acquisition. The post-training lesion group allowed us to evaluate the effects of pDMS lesions in rats that had shown normal response rates during acquisition.

Outcome devaluation

Figure 4 illustrates the results from the extinction test after the selective satiety devaluation treatment. There was a main effect of group ($F_{3,26} = 5.1$, $P < 0.01$), no main effect of devaluation ($F_{1,26} = 2.1$, $P > 0.1$), and no interaction between these factors ($F < 1$). The only group that showed a significant devaluation effect was the aDMS-pre-group ($F_{1,5} = 14.4$, $P < 0.01$). Interestingly, performance of rats in the aDMS-post-group showed attenuated sensitivity to devaluation with no significant difference in responding between the devalued and non-devalued conditions ($F_{1,7} = 1.2$, $P > 0.1$). However, given the numerical difference between the conditions, and the reliable effect of devaluation in the aDMS-pre-group, it appears that the aDMS does not play a critical role in goal-directed action. In contrast, both the pDMS-pre- and pDMS-post-groups failed to show a selective devaluation effect and both groups performed similarly on the devalued and the non-devalued actions (both Fs < 1).

Contingency degradation

The training and test data from contingency degradation are shown in Fig. 5. Before degradation training, all rats received 4 days of retraining on RR-20. Therefore, the post-training lesion groups must now be considered pre-training groups with respect to contingency degradation. Moreover, when the time of lesion (pre- or post-training)

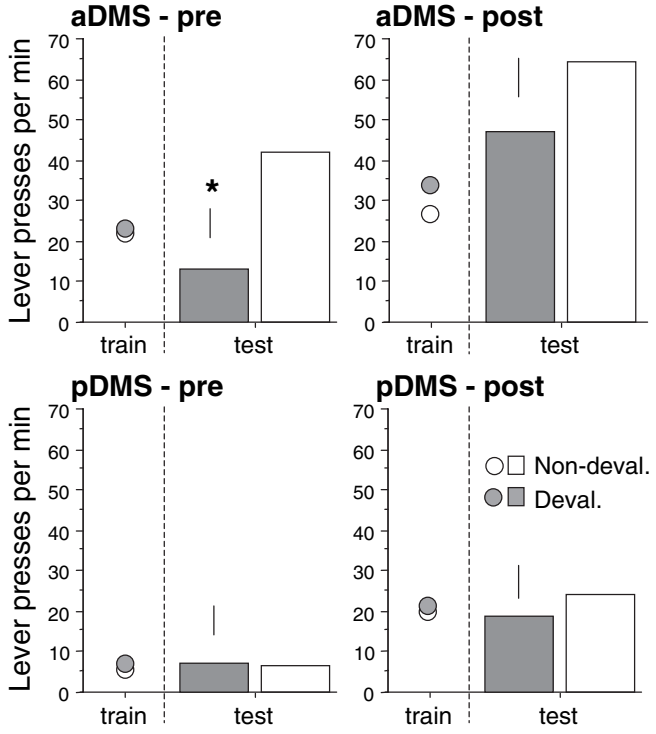

FIG. 4. Experiment 2. Responses rates during the 5-min devaluation test. Response rates on the actions that, in training, delivered the now devalued and non-devalued outcomes are shown both for the last day of training, before devaluation, on the left, and during the extinction test immediately after devaluation treatment, on the right, for the aDMS-pre and aDMS-post (top panels) and the pDMS-pre and pDMS-post (lower panels). ★ indicates $P < 0.05$. Each vertical line represents 1 SED (standard error of the difference of the means, a measure of within-subject variance).

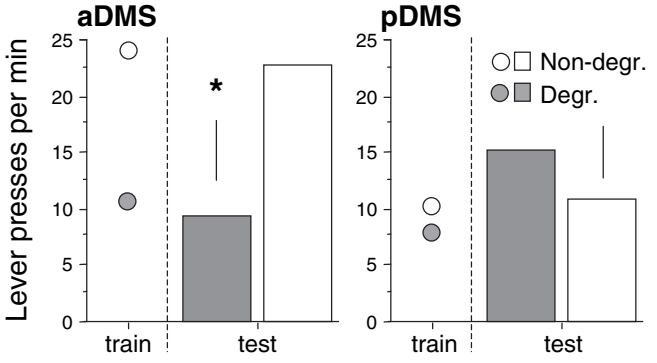

FIG. 5. Experiment 2. Response rates during the extinction test after contingency degradation training. Response rates are shown separately for the action that had the action–outcome contingency degraded and that for which the contingency was not degraded, both for the last day of contingency degradation training, on the left, and during the critical extinction test, on the right, for the aDMS and pDMS groups. ★ indicates $P < 0.05$. Each vertical line represents 1 SED (standard error of the difference of the means, a measure of within-subject variance).

was included as a factor, there was no interaction between this factor and degradation ($F < 1$). For these two reasons, data from pre-training and post-training lesioned groups were combined, and a two-way ANOVA was conducted using factors of lesion type, separating performance in the aDMS and pDMS lesioned rats, and of contingency. This analysis revealed no main effect of lesion type or of contingency, but there was a significant interaction between these factors ($F_{1,28} = 10.9$, $P < 0.01$), showing that the degradation training had different effects on the two lesion groups. Simple main effects analyses revealed a significant effect of degradation in the aDMS group ($F_{1,13} = 10.4$, $P < 0.01$), but not in the pDMS group ($F_{1,15} = 2.0$, $P > 0.05$).

Experiment 3: The effect of muscimol infusion into the pDMS on instrumental performance

Outcome devaluation

Figure 6 shows the cannulae placement for the rats in Experiment 3. All of the rats learned to press both levers during the training phase of this experiment. The results from the extinction test are shown in Fig. 7. Lever pressing in the ACSF group was selectively reduced by the devaluation treatment. This sensitivity, however, was abolished in the muscimol group. A repeated-measures ANOVA using devaluation and minute of testing (1–10) as within-subjects factors and drug treatment as a between-subjects factor showed that there was no main effect of drug ($F_{1,14} = 3.04$, $P > 0.1$), but there was a significant main effect of devaluation ($F_{1,14} = 6.16$, $P < 0.05$), and more importantly, an interaction between drug and devaluation ($F_{1,14} = 6.74$, $P < 0.05$). Simple main effects analyses conducted on this analysis revealed that, whereas the ACSF group selectively reduced performance on the lever earning the devalued outcome ($F_{1,7} = 7.04$, $P < 0.05$), the muscimol group did not ($F_{1,7} < 1$). There was also a marginally significant main effect of minute of testing ($F_{9,14} = 1.72$, $P = 0.09$), but no interaction between drug and minute, and no three–way interaction (Fs < 1).

On the rewarded test that followed, however, both groups were able to reduce lever pressing leading to the devalued outcome (Fig. 7). There was a main effect of devaluation treatment ($F_{1,14} = 8.44$, $P < 0.05$) but no main effect of group ($F < 1$), and no interaction between these two factors ($F < 1$). Moreover, there was no significant main effect of minute of testing and no interaction between devaluation and minute (Fs < 1).

Contingency degradation

Again, as illustrated in Fig. 8, the results from the degradation test were similar to those from the devaluation test. ANOVA revealed a main effect of group ($F_{1,11} = 11.53$, $P < 0.01$), a main effect of degradation ($F_{1,11} = 13.13$, $P < 0.01$), and an interaction between these factors ($F_{1,11} = 16.88$, $P < 0.01$). Simple main effects analyses conducted on the significant interaction revealed that, whereas the ACSF group performed relatively fewer responses on the degraded action ($F_{1,4} = 10.3$, $P < 0.05$), responding in the muscimol group did not differ ($F < 1$).

Outcome-specific reinstatement

In this experiment, we examined the effect of pDMS inactivation on extinction of lever pressing, and on the reinstatement of lever pressing by a single presentation of the outcome. It has been demonstrated that, after a response has undergone extinction, a presentation of the outcome or of cues associated with it can selectively reinstate that response; after re-exposure to the outcome, rats will selectively increase responding on the lever originally earning that outcome, relative to the lever earning a different outcome (Ostlund & Balleine, 2003). If the pDMS is necessary for utilizing action–outcome information, then inactivation of this structure should render this representation inaccessible for the purpose of outcome-specific reinstatement.

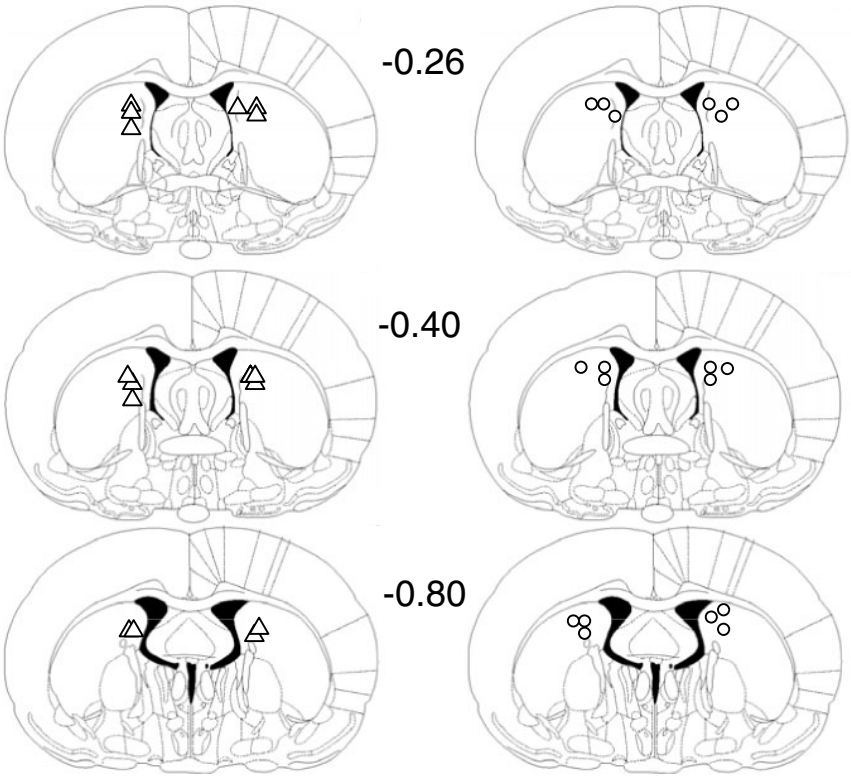

FIG. 6. Experiment 3. Schematic representation of the cannulae placement in coronal section (Paxinos & Watson, 1998). Triangles, ACSF group; circles, muscimol group. Numbers indicate distance from bregma in mm.

Figure 9 presents the critical data for the 20 min of extinction followed by the two-min test period after the single presentation of either the pellet or sucrose outcome. Muscimol infused into the pDMS had no effect on extinction performance. A two-way ANOVA conducted using group and bin as factors found no significant effect of bin ($F_{4,44} = 2.04$, $P > 0.05$), no effect of group, nor a group–bin interaction (largest $F = 1.23$). In contrast, inactivation of the pDMS had a clear effect on the ability of rats to retrieve action–outcome information during the reinstatement test. Whereas the vehicle group selectively increased performance on the lever that, in training, had delivered the reinstating outcome, the muscimol group did not show this pattern and, indeed, failed to show any reinstatement of instrumental performance. Using group (muscimol vs. vehicle), lever (reinstated vs. non-reinstated) and bin as factors, a mixed three-way ANOVA revealed no main effect of group ($F_{1,11} = 1.77$, $P > 0.05$), of lever ($F_{1,11} = 1.49$, $P > 0.05$), or of bin ($F_{7,77} = 1, 45$, $P > 0.05$), but a marginal interaction between group and lever ($F_{1,11} = 4.36$, $P = 0.06$). Simple effects analyses revealed that performance differed between groups on the reinstated lever ($F_{1,11} = 5.23$, $P < 0.05$) but did not differ on the control lever ($F < 1$). Furthermore, there was also a significant bin–group interaction ($F_{7,77} = 2.21$, $P < 0.05$), no significant lever–bin interaction ($F_{7,77} = 1.41$, $P > 0.05$) and no significant group × bin–lever interaction ($F_{7,77} = 1.06$).

Action discrimination on a chain schedule

The above deficits can be explained by an alternative account – that pDMS inactivation impairs the discrimination between actions. We tested this possibility using a heterogeneous chain of instrumental actions in which rats have to learn to press one lever, R1, and then another, R2, in order to earn reward, i.e. R1 → R2 → pellet. Clearly any inability to discriminate between actions would render performance insensitive to the imposition of this kind of chain; the probability of R1 → R2 would be similar to R2 → R1. If, however, rats successfully acquire this chain then the probability of performing R1 → R2 should be greater than R2 → R1.

As shown in the bottom panel of Fig. 10, the probability of R1 → R2 (calculated as actions per opportunity) although initially low, rapidly increased each second after performance of R1, peaking approximately 2-s after R1. The probability of R2 → R1 did not show this pattern; performance of R1 remained relatively unchanged throughout a 5-s time window after the performance of R2. In short, whereas R2 performance appeared to depend on R1, R1 performance appeared to be relatively independent of R2 in that same time window. In addition, the difference between R1 → R2 and R2 → R1 was similar in both vehicle and muscimol infused rats. These actions per opportunity data were analysed using a mixed ANOVA with group (i.e. muscimol vs. ACSF), order [i.e. P(P1 → P2) vs. P(R2 → R1)], and time as factors. This analysis found no main effect of group ($F < 1$), a significant main effect of order ($F_{1,10} = 66.9$, $P < 0.01$), and a significant main effect of time ($F_{1,10} = 3.36$, $P < 0.05$). There was, moreover, a significant interaction between time and order ($F_{1,10} = 5.13$, $P < 0.05$). Simple main effects analyses revealed a main effect of order for the vehicle group ($F_{5,1} = 40.12$, $P < 0.01$) as well as the muscimol group ($F_{5,1} = 27.27$, $P < 0.05$).

Discussion

Although it has long been argued that the striatum is involved in instrumental learning (Divac *et al.*, 1967; Konorski, 1967), the relevant early studies were conducted before differences within the striatum were well-understood, and without the benefit of contemporary behavioural analyses. According to older theories of instrumental conditioning, animals acquire actions solely through the formation of new S–R associations – an account that has informed the influential proposal that the striatum mediates habit or procedural learning, and that continues to exert influence on current models, e.g. of reinforce-

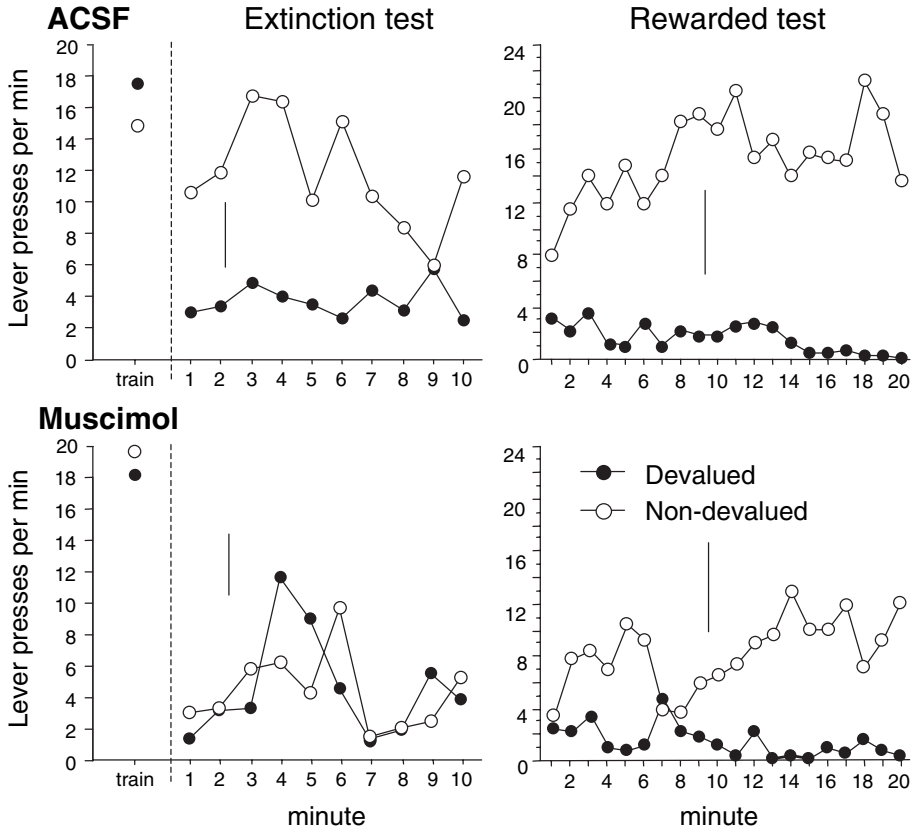

FIG. 7. Experiment 3. Response rates during the extinction and rewarded tests after selective satiety devaluation treatment in rats given an infusion of ACSF into the pDMS (top panels) or of muscimol into the pDMS (bottom panels). The far left panel shows the mean response rats during training. The central panel shows the response rates during the extinction test. The right panels show response rates during the rewarded test. The vertical lines represent ± 1 SED (standard error of the difference of the means, a measure of within-subject variance).

ment learning (Mishkin *et al*., 1984; Dayan & Balleine, 2002). Indeed, we have recently reported evidence in support of this latter claim: lesions of the DLS brought normally habitual actions under the control of the goal-directed system (Yin *et al*., 2004).

Given our present anatomical knowledge of the striatum, it is likely that subregions within the striatum, with their distinct anatomical connections, serve different behavioural functions. The present results establish clear evidence for this suggestion showing that the DMS, particularly the pDMS, is a critical locus for the acquisition and expression of the instrumental action–outcome association. These results provide evidence therefore that there may be a functional dissociation between adjacent regions of the dorsal striatum, between the dorsomedial region that mediates goal-directed actions and the dorsolateral region that mediates habitual, stimulus-driven actions.

The S-R habit system is thought to be characterized by lower response rates and slower, more gradual acquisition (Dickinson & Balleine, 1993). For example, response rates under interval schedules of reinforcement, which generate habitual responding insensitive to devaluation, are significantly lower than those under ratio schedules, which tend to generate goal-directed actions (Dickinson *et al*., 1983). On this view, the present results suggest that when the pDMS is disrupted, the neural circuit required for goal-directed actions is dysfunctional, requiring animals to rely on the habit system instead to acquire lever pressing. Accordingly, in Experiment 1, responding in these animals was characterized by slower acquisition and insensitivity to both outcome devaluation and to contingency degradation. Furthermore, in Experiment 2, we found that the role of the pDMS was not limited to acquisition; post-training lesions produced a striking deficit both on outcome devaluation and on contingency degradation, confirming that this structure is necessary for both the acquisition and expression of instrumental actions. Finally, Experiment 3 showed that inactivation of the pDMS abolished sensitivity to outcome devaluation, contingency degradation and selective outcome-induced reinstatement without affecting either performance in extinc-

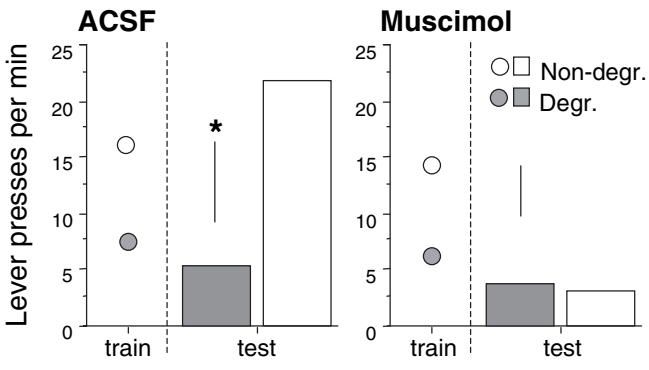

FIG. 8. Experiment 3. Response rates during the extinction test after contingency degradation training. Response rates are shown separately for the action that had the action–outcome contingency degraded and that for which the contingency was not degraded both for the last day of contingency degradation training, on the left, and during the critical extinction test, on the right, for rats given ACSF infusions (left panel) and rats given muscimol infusions (right panel) into the pDMS. ★ indicates $P < 0.05$. Each vertical line represents 1 SED (standard error of the difference of the means, a measure of within-subject variance).

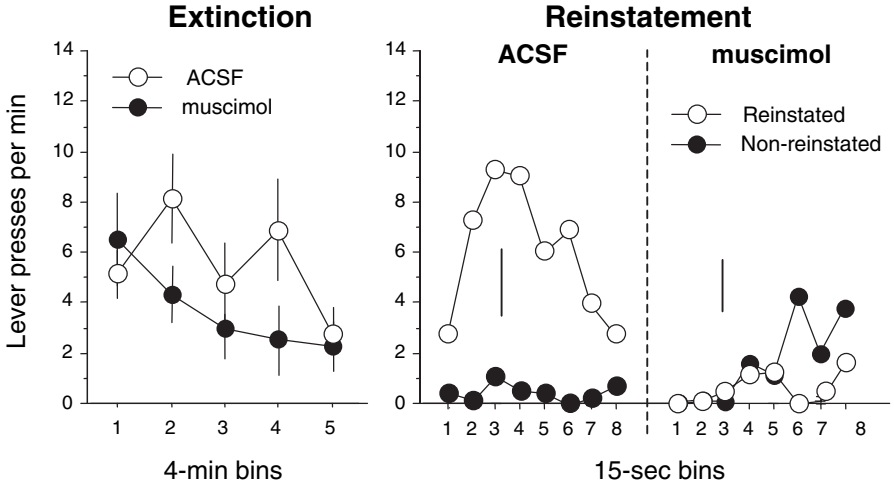

FIG. 9. Experiment 3. Results of the selective reinstatement test. The left panel shows response rates during the 20-min extinction phase in rats given ACSF and muscimol infusions into the pDMS (± 1 SEM). The right panel shows response rates in the 2 min immediately after the delivery of a single instrumental outcome (either a food pellet or 0.1 mL of 20% sucrose solution), in rats given infusions of ACSF or muscimol into the pDMS, for the reinstated action that previously earned the reinstating outcome and for the non-reinstated action that earned the other outcome. Error bars in the right panels represent ± 1 SED (standard error of the difference of the means, a measure of within-subject variance).

tion or the ability of rats to perform accurately on a chain of instrumental actions in order to gain access to reward.

The results of these experiments are therefore consistent with the claim that the pDMS is a critical structure in the acquisition and expression of instrumental learning. One implication of this argument is that neural plasticity underlying action–outcome learning in instrumental conditioning should involve the pDMS. In a companion paper we report the results of a series of experiments designed to test this prediction by examining the effects of infusing the NMDA antagonist APV into the pDMS during instrumental acquisition (Yin *et al*., 2005). Generally, this study found clear evidence that the infusion of APV rendered rats' instrumental performance insensitive to subsequent devaluation of the instrumental outcome but only if the APV was infused immediately prior to acquisition and only if the infusion was into the pDMS. APV infused into the dorsolateral striatum was without effect (Yin *et al*., 2005).

Alternative accounts of the present results

The effects observed in the current study were found in choice performance and, as has been well documented (see, for example, Colwill & Rescorla, 1986), reliable outcome devaluation and contingency degradation effects on choice performance do not require high rates of performance, but only consistent changes in the distribution of responses across actions. Therefore, the failure of rats successfully to alter their choice performance after lesions or inactivation of the pDMS shows that the deficit lies in the effects of these manipulations on their representational capacity rather than on their motor performance.

For example, reduced performance in the lesioned rats could be taken as evidence that, rather than a specific functional deficit, these rats suffered from a general motor deficit and were unable to respond at a sufficient level to show devaluation and degradation effects. But, by definition, a general motor impairment should be expected to reduce motor activity on all tasks equally. This is not what we have found either in the current study or in other experiments in which we have assessed the effects of lesions of dorsomedial striatum. First, it was shown in a previous study that pDMS lesions did not affect response latency in the performance of rats on a T-maze, even though choice behaviour based on the flexible use of place cues was impaired (Yin & Knowlton, 2004). Second, lesions of the dorsmedial striatum that overlapped with those of the current study were found to have no effect on lever-press performance when rats were trained on interval schedules of reinforcement (Yin *et al*., 2004). Although this training rendered the rats' performance habitual, the lesions clearly did not produce any evidence of a general motor impairment. Finally, in the current study, Experiment 3 was explicitly designed to test the possibility that the observed deficits in pDMS-lesioned rats could be attributed to a general motor deficit. Although muscimol infusions reduced responding during the extinction test after devaluation or degradation, this reduction in performance was not an inevitable consequence of pDMS inactivation. Prior to the reinstatement test, the level of performance on the two levers during the 20-min extinction phase did not differ between the muscimol and vehicle groups (see Fig. 9). Thus, the reduction of performance after devaluation and degradation treatments appears to have been caused by these manipulations specifically, and was not a general effect of pDMS inactivation.

Nor can the pattern of results in the current studies be explained by reference to a deficit in the ability of the rats to discriminate either the sensory properties of the outcomes used or to assess their reward value. In the rewarded test (Fig. 7) rats in the muscimol group selectively reduced performance on the lever yielding the devalued outcome, showing that, indeed, they could discriminate which outcome was which and attach a distinct value to each outcome. It could of course be argued that this evidence of successful discrimination comes after a period of extinction and so at a time when the effects of muscimol were wearing off. There are two problems with this argument. First, the rats in the muscimol group began to discriminate between the two actions (something they could not do in extinction) quite quickly after the outcomes started to be delivered in the reward test and certainly within 5 min. Second, in the reinstatement test, the infusions of muscimol were clearly effective in blocking reinstatement even after 20 min of extinction, i.e. at a time when the rats were reliably discriminating during the rewarded test.

Finally, we were able to show that pDMS inactivation did not significantly impair acquisition of a heterogeneous chain of instrumental actions (Fig. 10). Not only was the muscimol group able to acquire the chain contingency, it also showed similar

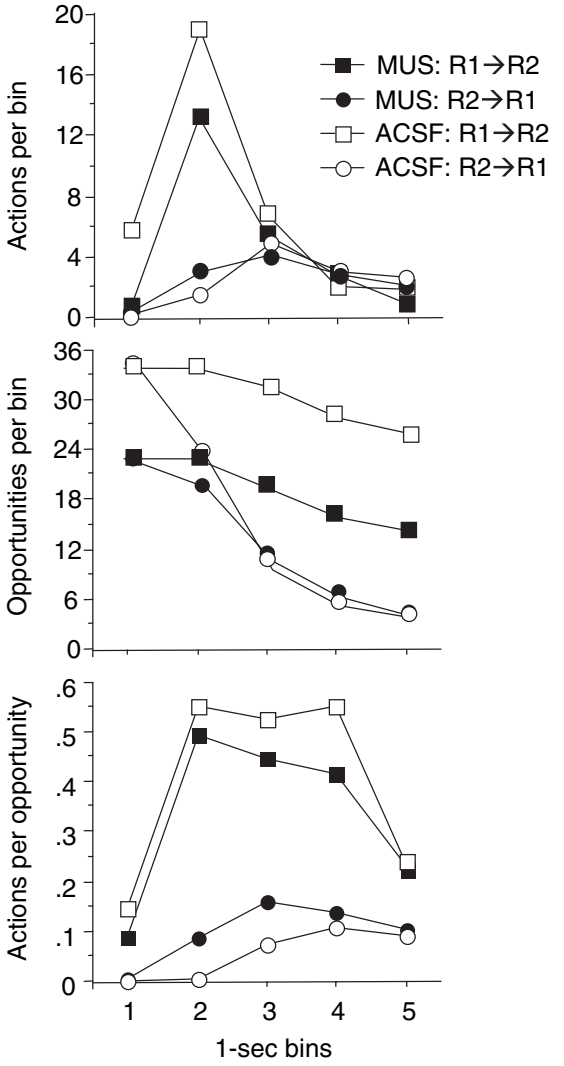

FIG. 10. Experiment 3. Action discrimination using a heterogeneous chain of lever-press responses: R1 → R2 → reward. The figure shows the number of responses per opportunity on one action in the chain in the 5-s period after the performance of the other action in the chain. R1 → R2, probability of second response after the first response in the chain; R2 → R1, probability of first response after the second response in the chain. Performance is shown separately for rats tested after infusions of ACSF and after infusions of muscimol into the pDMS.

response rates to the ACSF group on the distal, reward-seeking component of the chain (Corbit & Balleine, 2003). This result, then, also rules out the possibility that the effects of muscimol infusion into the pDMS on outcome devaluation and contingency degradation were due to an impaired ability to discriminate between the two actions.

The neural substrates of instrumental learning

Previous studies have shown that lesions of the basolateral amygdala (BLA), the core of the nucleus accumbens (NAC), the mediodorsal thalamus (MD), and the prelimbic (PL) region of the medial prefrontal cortex also produced deficits in instrumental conditioning (Balleine & Dickinson, 1998; Corbit *et al*., 2001; Balleine *et al*., 2003; Corbit *et al*., 2003). Important differences between these earlier findings and the present results should, however, be noted. For example, after BLA lesions, rats were not able to use outcome value as a means of discriminating between actions during a rewarded test, nor could they use outcomes as discriminative stimuli. This pattern suggests that BLA lesions impaired the ability of rats to encode the value of instrumental outcomes (Balleine *et al*., 2003; Blundell *et al*., 2001). Lesions of the NAC also impaired acquisition and reduced sensitivity to outcome devaluation, but not contingency degradation (Corbit *et al*., 2001), again suggesting that these lesions affect the motivational system by selectively disrupting the effect of the instrumental incentive process on performance. Thus the NAC, as a component of the limbic cortico-basal ganglia network, appears to mediate the ability of the incentive value of rewards and of cues associated with reward to affect instrumental performance, but does not play a direct role in action–outcome learning *per se* (cf. Balleine & Killcross, 1994; Parkinson *et al*., 2000; Corbit *et al*., 2001; Cardinal *et al*., 2002; de Borchgrave *et al*., 2002).

In contrast, pre-training lesions of either the MD or the PL produce deficits in instrumental conditioning that are similar to those of the pDMS, i.e. insensitivity to outcome devaluation and contingency degradation (Corbit & Balleine, 2003; Corbit *et al*., 2003; Killcross & Coutureau, 2003). This is not surprising. The PL is a major source of the excitatory corticostriatal inputs to the DMS (McGeorge & Faull, 1989), and the MD, also heavily connected with the DMS, may serve as an important locus at which output from the pDMS eventually reenters the thalamocortical network (Nauta, 1989). Nevertheless, recent work has revealed a significant functional difference between the PL and the pDMS. In contrast to the pDMS, both post-training lesions and inactivation of the PL have no effect on sensitivity to outcome devaluation (Ostlund & Balleine, 2003), i.e. the performance of goal-directed actions, once acquired, appears no longer to require the PL, whereas the pDMS is involved in both the acquisition and further deployment of instrumental learning. Although a more detailed analysis of the respective roles of the PL and of the DMS in instrumental learning is needed to clarify the functioning of this corticostriatal circuit, our results point to its crucial role in the acquisition of goal-directed actions.

Another interesting feature of our results is the finding that there may be significant differences between anterior and posterior regions in the DMS in instrumental learning. The lesion results indicate that the posterior region is more important for the acquisition of action–outcome learning than the more anterior region. Nauta (1989), on the basis of purely anatomical analysis, also pointed out a possible anterior-posterior dissociation, based on the extent of inputs from the limbic system. As the pDMS receives a greater input from the BLA (McGeorge & Faull, 1989), which also plays an important role in encoding the value of instrumental outcomes, the BLA–pDMS projection could be a critical pathway via which the value of the expected outcome interacts with knowledge of action–outcome contingency to guide instrumental behaviour. This possibility remains to be tested.

Finally, our results underscore the importance of analysing the behavioural functions of neural circuits in terms of the underlying associative processes that support performance. Although the performance of a lever-press action may appear to be quite similar in two rats, this does not mean that it is in fact controlled by the same underlying psychological processes or the same neural circuits, and tests must be conducted to establish whether the lever pressing in question is controlled by antecedent stimuli (stimulus-driven habits) or by the encoded action–outcome association (goal-directed actions). That these fundamentally distinct processes can be revealed by contemporary behavioural assays has far-reaching implications for the study of cerebral functioning.

Acknowledgements

This research was supported by a NSF graduate fellowship to HHY, NSF grant #9985417 to BJK and NIMH grant #56446 to BWB.

Abbreviations

ACSF, artificial cerebral spinal fluid; aDMS, anterior dorsomedial striatum; BLA, basolateral amygdala; CRF, continuous reinforcement; DLS, dorsolateral striatum; DMS, dorsomedial striatum; MD, mediodorsal thalamus; NAC, nucleus accumbens; pDMS, posterior dorsomedial striatum; PL, prelimbic region; RR, random ratio; S–R, stimulus–response.

References

Adams, C.D. & Dickinson, A. (1981) Instrumental responding following reinforcer devaluation. *Q. J. Exp. Psychol.*, **33B**, 109–122.

Balleine, B.W. (2001) Incentive processes in instrumental conditioning. In Klein, R.M.S. (Ed.), *Handbook of Contemporary Learning Theories.* LEA, Hillsdale, NJ, pp. 307–366.

Balleine, B.W. & Dickinson, A. (1998) Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology*, **37**, 407–419.

Balleine, B. & Killcross, S. (1994) Effects of ibotenic acid lesions of the nucleus accumbens on instrumental action. *Behav. Brain Res.*, **65**, 181–193.

Balleine, B.W., Killcross, A.S. & Dickinson, A. (2003) The effect of lesions of the basolateral amygdala on instrumental conditioning. *J. Neurosci.*, **23**, 666–675.

Blundell, P., Hall, G. & Killcross, S. (2001) Lesions of the basolateral amygdala disrupt selective aspects of reinforcer representation in rats. *J. Neurosci.*, **21**, 9018–9026.

de Borchgrave, R., Rawlins, J.N., Dickinson, A. & Balleine, B.W. (2002) Effects of cytotoxic nucleus accumbens lesions on instrumental conditioning in rats. *Exp. Brain Res.*, **144**, 50–68.

Cardinal, R.N., Parkinson, J.A., Hall, J. & Everitt, B.J. (2002) Emotion and motivation: The role of the amygdala, ventral striatum, and prefrontal cortex. *Neurosci. Biobehav. Rev.*, **26**, 321–352.

Carelli, R.M., Wolske, M. & West, M.O. (1997) Loss of lever press-related firing of rat striatal forelimb neurons after repeated sessions in a lever pressing task. *J. Neurosci.*, **17**, 1804–1814.

Cheatwood, J.L., Reep, R.L. & Corwin, J.V. (2003) The associative striatum: cortical and thalamic projections to the dorsocentral striatum in rats. *Brain Res.*, **968**, 1–14.

Colwill, R.M. & Rescorla, R.A. (1986) Associative structures in instrumental learning. In Bower, G. (Ed.), *The Psychology of Learning and Motivation.* Academic Press, New York, pp. 55–104.

Corbit, L.H. & Balleine, B.W. (2003) Instrumental and Pavlovian incentive processes have dissociable effects on components of a heterogeneous instrumental chain. *J. Exp. Psychol. Anim. Behav. Process.*, **29**, 99–106.

Corbit, L.H., Muir, J.L. & Balleine, B.W. (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, L.H., Muir, J.L. & Balleine, B.W. (2003) Lesions of mediodorsal thalamus and anterior thalamic nuclei produce dissociable effects on instrumental conditioning in rats. *Eur. J. Neurosci.*, **18**, 1286–1294.

Dayan, P. & Balleine, B.W. (2002) Reward, motivation, and reinforcement learning. *Neuron*, **36**, 285–298.

Dickinson, A. & Balleine, B. (1993) Actions and responses: The dual psychology of behaviour. In Eilan, N. & McCarthy, R.A. (Eds), *Spatial Representation: Problems in Philosophy and Psychology.* Blackwell Publishers Inc., Malden, MA, US, pp. 277–293.

Dickinson, A. & Balleine, B.W. (1994) Motivational control of goal-directed action. *Anim. Learning Behav.*, **22**, 1–18.

Dickinson, A. & Balleine, B.W. (1995) Motivational control of instrumental action. *Curr. Directions Psychol. Sci.*, **4**, 162–167.

Dickinson, A. & Balleine, B.W. (2002) The role of learning in the operation of motivational systems. In Gallistel, C.R. (Ed.), *Learning, Motivation and Emotion*, Vol. 3, Steven's Handbook of Experimental Psychology, 3rd Edn. John Wiley & Sons, New York, pp. 497–533.

Dickinson, A., Balleine, B.W., Watt, A., Gonzales, F. & Boakes, R.A. (1995) Overtraining and the motivational control of instrumental action. *Anim. Learning Behav.*, **22**, 197–206.

Dickinson, A., Nicholas, D.J. & Adams, C.D. (1983) The effect of the instrumental training contingency on susceptibility to reinforcer devaluation. *Q. J. Exp. Psychology: Comp. Physiol. Psychol.*, **35**, 35–51.

Divac, I., Rosvold, H.E. & Szwarcbart, M.K. (1967) Behavioral effects of selective ablation of the caudate nucleus. *J. Comp. Physiol. Psychol.*, **63**, 184–190.

Graybiel, A.M. (1998) The basal ganglia and chunking of action repertoires. *Neurobiol. Learn. Mem.*, **70**, 119–136.

Hassani, O.K., Cromwell, H.C. & Schultz, W. (2001) Influence of expectation of different rewards on behavior-related neuronal activity in the striatum. *J. Neurophysiol.*, **85**, 2477–2489.

Hikosaka, O., Takikawa, Y. & Kawagoe, R. (2000) Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiol. Rev.*, **80**, 953–978.

Jaeger, D., Gilman, S. & Aldridge, J.W. (1993) Primate basal ganglia activity in a precued reaching task: preparation for movement. *Exp. Brain Res.*, **95**, 51–64.

Jog, M.S., Kubota, Y., Connolly, C.I., Hillegaart, V. & Graybiel, A.M. (1999) Building neural representations of habits. *Science*, **286**, 1745–1749.

Kawagoe, R., Takikawa, Y. & Hikosaka, O. (1998) Expectation of reward modulates cognitive signals in the basal ganglia. *Nature Neurosci.*, **1**, 411–416.

Kelley, A.E., Domesick, V.B. & Nauta, W.J. (1982) The amygdalostriatal projection in the rat – an anatomical study by anterograde and retrograde tracing methods. *Neuroscience*, **7**, 615–630.

Killcross, S. & Coutureau, E. (2003) Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb. Cortex*, **13**, 400–408.

Kimura, M. (1990) Behaviorally contingent property of movement-related activity of the primate putamen. *J. Neurophysiol.*, **63**, 1277–1296.

Kimura, M. (1992) Behavioral modulation of sensory responses of primate putamen neurons. *Brain Res.*, **578**, 204–214.

Kimura, M., Aosaki, T., Hu, Y., Ishida, A. & Watanabe, K. (1992) Activity of primate putamen neurons is selective to the mode of voluntary movement: visually guided, self-initiated or memory-guided. *Exp. Brain Res.*, **89**, 473–477.

Konorski, J. (1967) *Integrative Activity of the Brain*. University of Chicago Press, Chicago.

McDonald, R.J. & White, N.M. (1993) A triple dissociation of memory systems: hippocampus, amygdala, and dorsal striatum. *Behav. Neurosci.*, **107**, 3–22.

McGeorge, A.J. & Faull, R.L. (1989) The organization of the projection from the cerebral cortex to the striatum in the rat. *Neuroscience*, **29**, 503–537.

Mishkin, M., Malamut, B. & Bachevalier, J. (1984) Memories and habits: Two neural systems. In Lynch, G., McGaugh, J.L. & Weinberger, N. (Eds), *Neurobiology of Learning and Memory*, Guilford Press, New York, pp. 65–77.

Nauta, W.J.H. (1989) Reciprocal links of the corpus striatum with the cerebral cortex and limbic system: A common substrate for movement and thought? In Mueller, J. (Ed), *Neurology and Psychiatry: a Meeting of Minds.* Karger, Basel, pp. 43–63.

Ostlund, S.B. & Balleine, B.W. (2003) Medial prefrontal cortex lesions disrupt the reinstatement of instrumental behavior. *Soc. Neurosci. Abstr.* [Annual Meeting of the Society for Neuroscience, 2003, New Orleans].

Packard, M.G. & Knowlton, B.J. (2002) Learning and memory functions of the Basal Ganglia. *Annu. Rev. Neurosci.*, **25**, 563–593.

Packard, M.G. & McGaugh, J.L. (1996) Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. *Neurobiol. Learn. Mem.*, **65**, 65–72.

Parkinson, J.A., Cardinal, R.N. & Everitt, B.J. (2000) Limbic cortical-ventral striatal systems underlying appetitive conditioning. *Prog. Brain Res.*, **126**, 263–285.

Paxinos, G. & Watson, C. (1998) *The Rat Brain in Sterotaxic Coordinates*. Academic Press, San Diego.

Reep, R.L., Cheatwood, J.L. & Corwin, J.V. (2003) The associative striatum: organization of cortical projections to the dorsocentral striatum in rats. *J. Comp. Neurol.*, **467**, 271–292.

White, I.M. & Rebec, G.V. (1993) Responses of rat striatal neurons during performance of a lever-release version of the conditioned avoidance response task. *Brain Res.*, **616**, 71–82.

Yin, H.H. & Knowlton, B.J. (2004) Contributions of striatal subregions to place and response learning. *Learn. Mem.*, **11**, 459–463.

Yin, H.H., Knowlton, B.J. & Balleine, B.W. (2004) Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur. J. Neurosci.*, **19**, 181–189.

Yin, H.H., Knowlton, B.J. & Balleine, B.W. (2005) Blockade of NMDA receptors in the dorsomedial striatum abolishes action-outcome learning in instrumental conditioning. *Eur. J. Neurosci.*, **22**, 505–512.