European Journal of Neuroscience, Vol. 22, pp. 505–512, 2005

Blockade of NMDA receptors in the dorsomedial striatum prevents action–outcome learning in instrumental conditioning

Henry H. Yin, 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: APV, basal ganglia, goal-directed action, learning

Abstract

Although there is consensus that instrumental conditioning depends on the encoding of action–outcome associations, it is not known where this learning process is localized in the brain. Recent research suggests that the posterior dorsomedial striatum (pDMS) may be the critical locus of these associations. We tested this hypothesis by examining the contribution of *N*-methyl-D-aspartate receptors (NMDARs) in the pDMS to action–outcome learning. Rats with bilateral cannulae in the pDMS were first trained to perform two actions (left and right lever presses), for sucrose solution. After the pre-training phase, they were given an infusion of the NMDA antagonist 2-amino-5-phosphonopentanoic acid (APV, 1 mg/mL) or artificial cerebral spinal fluid (ACSF) before a 30-min session in which pressing one lever delivered food pellets and pressing the other delivered fruit punch. Learning during this session was tested the next day by sating the animals on either the pellets or fruit punch before assessing their performance on the two levers in extinction. The ACSF group selectively reduced responding on the lever that, in training, had earned the now devalued outcome, whereas the APV group did not. Experiment 2 replicated the effect of APV during the critical training session but found no effect of APV given after acquisition and before test. Furthermore, Experiment 3 showed that the effect of APV on instrumental learning was restricted to the pDMS; infusion into the dorsolateral striatum did not prevent learning. These experiments provide the first direct evidence that, in instrumental conditioning, NMDARs in the dorsomedial striatum are involved in encoding action–outcome associations.

Introduction

Animals adapt to changing environments by acquiring new actions to obtain desired goals. Extensive evidence across species suggests that this capacity depends on encoding the relationship between actions and their specific consequences or outcomes (Adams & Dickinson, 1981; Colwill & Rescorla, 1986; Corbit *et al*., 2001). Despite this advance in our understanding of the associative structure of instrumental learning, little direct evidence has emerged to suggest where in the brain action–outcome associations are encoded. The dorsomedial striatum, a critical component in the associative cortico-basal ganglia circuit, appears to be a plausible candidate structure for several reasons. First, it receives inputs from association cortices such as the prelimbic region of the prefrontal cortex as well as premotor areas such as the medial agranular cortex involved in the action monitoring and programming implicated in executive processes (Passingham *et al*., 1988; McGeorge & Faull, 1989; Nauta, 1989; Corbit & Balleine, 2003; Reep *et al*., 2003). The posterior part of the DMS (pDMS) also receives inputs from the basolateral amygdala (Kelley *et al*., 1982), a structure that, according to recent evidence, mediates the assignment of incentive value to the consequences of instrumental actions (Balleine *et al*., 2003; Everitt *et al*., 2003; Wang *et al*., 2005). Projections from the pDMS are therefore in a position to influence downstream motor control networks in the brainstem as well as the motor thalamocortical re-entrant network (Nauta, 1989). Finally, in a recent series of experiments (Yin *et al*., 2005), we found direct evidence that both pre- and post-training cell-body lesions of the pDMS as well as local inactivation of this area reduced the sensitivity of rats' performance both to non-contingent outcomes and to post-training reduction in outcome value, thus rendering their instrumental performance stimulus-bound and habitual.

The pDMS is therefore well positioned to combine the necessary inputs to form associations between actions and outcomes and provide the relevant directives to the motor system to guide instrumental performance. This hypothesis contrasts with other recent claims that the ventral (Kelley *et al*., 1997) or the posterolateral striatum (Andrzejewski *et al*., 2004) mediate learning critical to the acquisition of goal-directed actions. These studies only assessed changes in instrumental performance and did not directly assess changes in the content of learning. In the current study we used well-established behavioural assays that unambiguously distinguish action–outcome learning from other types of learning to assess the role of the pDMS in the formation of action–outcome associations. Given the evidence that NMDA receptor (NMDAR) activation is involved in long-term plasticity such as long-term potentiation in the dorsal striatum (Calabresi *et al*., 1992; Lovinger *et al*., 2003), we hypothesized that action–outcome association requires activation of NMDARs in the pDMS. This hypothesis was tested in rats that, after a period of pre-training (see Fig. 1 for the design), were given a bilateral infusion of either 2-amino-5-phosphonopentanoic acid (APV), a selective NMDAR antagonist, or artificial cerebral spinal fluid (ACSF) prior to a single learning session in which they were trained to press two levers for distinct outcomes. The next day we tested what the rats

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

Received 9 December 2004, revised 1 March 2005, accepted 16 March 2005

doi:10.1111/j.1460-9568.2005.04219.x

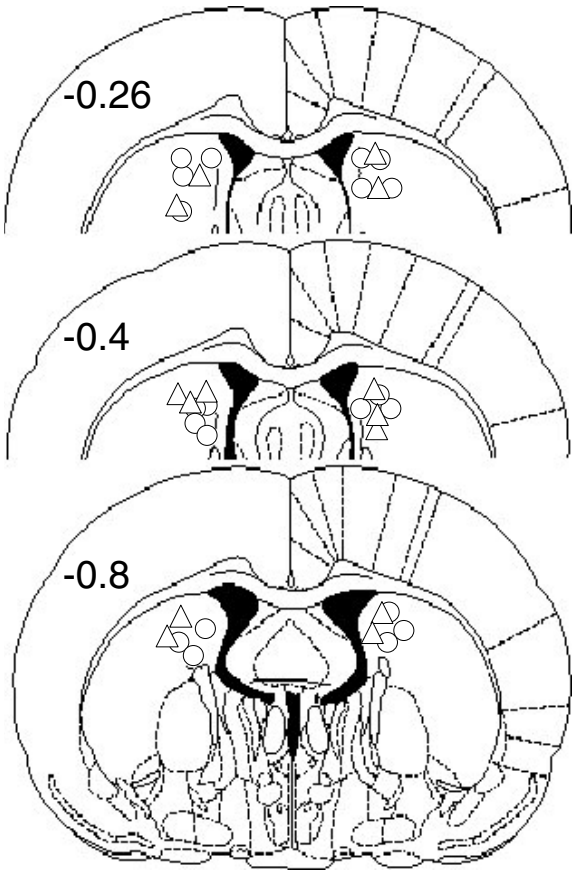

FIG. 1. Experiment 1. Schematic representation of the cannulae placement presented separately for the group infused with ACSF (circles) and with APV (triangles) during the training phase. The drawings of the coronal sections are taken −0.26, −0.4 and −0.8 mm relative to bregma (Paxinos & Watson, 1998).

had learned during this session using an outcome devaluation protocol. For this test rats were first allowed to consume one of the two outcomes for 1 h before a choice extinction test was given on the two levers.

We predicted that rats in the ACSF group would reduce their performance of the action that, in training, had delivered the outcome on which they were sated before the extinction test relative to the other action. If NMDARs in the pDMS are involved in the formation of action–outcome associations, then APV infusion during that session should block learning and render the rats' instrumental choice performance insensitive to the selective effects of outcome devaluation. Furthermore, on the above analysis this prediction should hold only for APV given before the critical training session and when APV is infused into the pDMS. To test these predictions, Experiment 2 compared the effects of APV infused immediately before the critical training session with the effects of APV infused 1 h after that session; and Experiment 3 compared the effect of infusing APV into the pDMS with its infusion into adjacent dorsolateral striatum (DLS).

Materials and methods

Experiment 1: The effects of NMDAR blockade on instrumental learning

Subjects and apparatus

Twenty-seven male Long–Evans rats weighing between 280 and 330 g were housed singly and handled daily for two days prior to surgery. The UCLA Animal Research Committee approved the study. Training and testing took place in 24 Medical Associates (East Fairfield, VT) operant chambers housed within sound- and light-resistant shells. Each chamber was equipped with a pump fitted with a syringe that could deliver 0.1 mL of either a 20% sucrose solution or of peach punch into a recessed magazine in the chamber. Each chamber was also equipped with a pellet dispenser that delivered one 45-mg pellet (Bio-Serve, Frenchtown, NJ) when activated. The chambers contained two retractable levers, that could be inserted to the left and right of the magazine. A 3 W, 24 V houselight mounted on the top-centre of the wall opposite the magazine provided illumination. Microcomputers 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. 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 to bregma, 2.6 mm lateral to the midline, and 4.5 mm ventral to the skull surface with reference to the atlas of Paxinos & Watson (1998). 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 striatum. The slices were stained with thionin and examined using a light microscope for cannulae placement.

Instrumental pre-training

Two days after surgery, rats were placed on a food deprivation schedule to reduce their weight to approximately 80% of their free-feeding weight. Once training began (10 days after surgery), the rats were fed each day after the training sessions, with free access to water while in their home cages. The pre-training phase began with a 30-min magazine training session in which the sucrose solution was delivered on a random time 60-s schedule with the levers retracted. Lever-press training began the next day. On each day, all rats were given two sessions, one on each lever in counterbalanced order. The levers were located on either side of the food magazine, at equal distances from it, and all reinforcers were delivered to the magazine. Each session began with the illumination of the house light and insertion of the lever and ended after 30 reinforcers had been earned with the retraction of the levers and turning off of the houselight. There was at least a 1-h break between the two sessions on each day. On the first day of training, the outcomes were delivered on a continuous reinforcement schedule. There followed two days on which the outcomes were delivered on a random ratio-5 schedule (RR-5) and two days on RR-10.

Instrumental acquisition

On the last day of lever-press training, in addition to their home chow, two new outcomes, i.e. food pellets (45 mg, Bio-serv, New Jersey) and peach punch (Tampico, Illinois) were given to the animals in their home cages to eliminate any neophobia to these novel rewards. The next day, all rats were given a single 30-min training session, after microinfusion of either APV or ACSF into the pDMS. Prior to this training session, the dummy cannulae were removed and injection cannulae (26 gauge; Plastics one, VA) were lowered into the guide cannulae, with an extension of 0.5 mm. The injection cannulae were connected by polyethylene tubing to 10-μL Hamilton syringes mounted on an infusion pump (Harvard). Following previous reports

(e.g. Ossowska & Wolfarth, 1995) APV (Sigma, Missouri; 1 μg per μL dissolved in ACSF, Sigma) was delivered at a rate of 0.25 μL per min for 2 min, with a total volume of 0.5 μL per side. The same volume of ACSF was used for the control infusions. One minute after infusion, the injection cannulae were removed and the dummy cannulae replaced. During the training session, one of the levers was inserted for 5 min, followed by its retraction and the insertion of the other lever counterbalanced across animals. For six rats in the ACSF group and eight rats in the APV group, the left lever earned pellets and right lever earned punch whereas the remainder in each group were given the opposite action–outcome pairings. Both outcomes were delivered on a RR-10 schedule of reinforcement.

Outcome devaluation

The day after the critical acquisition session we assessed what the rats had learned during the previous day using an outcome devaluation procedure. All rats were give the opportunity to consume either the pellets or the peach punch for 1 h in their home cages with seven rats in each lever–outcome assignment given access to the pellets in a bowl, whereas the remaining rats in each group were given access to the peach punch in a drinking tube. All rats consumed at least 10 mL of the peach punch or 5 g of the pellets during the pre-feeding session. Immediately after the pre-feeding session, the rats were returned to the operant chambers and given a 5-min choice extinction test on both levers. The test began with the illumination of the houselight and insertion of the levers, and ended with the retraction of the levers and the offset of the houselight. Neither the pellet nor the punch outcomes were delivered during this test.

Experiment 2: Comparing the effects of APV infusion before and after training

Twenty female Long–Evans rats were used in Experiment 2. It should be noted that the use of female vs. male rats was unlikely to influence the results of the present study; neither in our previous work nor in the literature is there any evidence of sex differences in the effects of reinforcer devaluation. The aim of Experiment 2 was to control for the general effects of APV infusion into the pDMS by comparing the effects of an infusion given before with one given 1 h after the critical instrumental acquisition session. Exactly the same procedures as those described in Experiment 1 were used in Experiment 2 except citrus-orange punch was used instead of peach-orange punch. Twelve rats received the infusion of APV before the critical training session as in Experiment 1. The remaining eight rats received the same infusion of APV except that it was given 1 h after training.

Experiment 3: Comparing the effects of APV infusion into pDMS and the DLS

Sixteen female Long–Evans rats were used in Experiment 3. Eight of the rats were implanted bilaterally with cannulae aimed at the pDMS, as described previously, whereas the others were implanted bilaterally with cannulae aimed at the DLS. For the DLS implants, the following coordinates were used: 0.7 mm anterior to bregma; 3.6 mm lateral to midline; and 5.0 mm below skull surface. The procedures for pre-training, the critical instrumental acquisition session and the devaluation test were the same as those described for Experiment 2 except that both groups were given an infusion of APV into either the pDMS or the DLS before the critical training session.

Results

Experiment 1: The effects of NMDAR blockade on instrumental learning

Schematic illustrations of cannulae placement are shown in Fig. 1. All rats learned to press both levers for sucrose during pre-training. Moreover, during the critical instrumental acquisition session, rats that received APV were found to respond on the two levers at a similar rate to those given the ACSF infusion. During this session, rats given the APV infusion responded at 6.44 presses per min on the lever that delivered the to-be-devalued outcome and 6.45 presses on the other lever, whereas rats given the ACSF infusion responded at 9.67 and 8.34 lever presses per min, respectively. A two-way ANOVA conducted using group (APV vs. ACSF) as a between-subjects factor and lever as a within-subject factor revealed no main effects of group or of lever and no interaction between these factors (largest $F_{1,25} = 1.6$, $P > 0.2$; $M_{SE} = 87.46$).

Although the APV did not significantly impair lever pressing performance during this instrumental acquisition session, the devaluation test conducted drug-free the next day revealed that there had been a significant effect of APV on what rats had learned during that session. As shown in Fig. 2, the critical data from the extinction test revealed a difference between the two groups: choice performance on the two levers in the rats infused during training with ACSF was

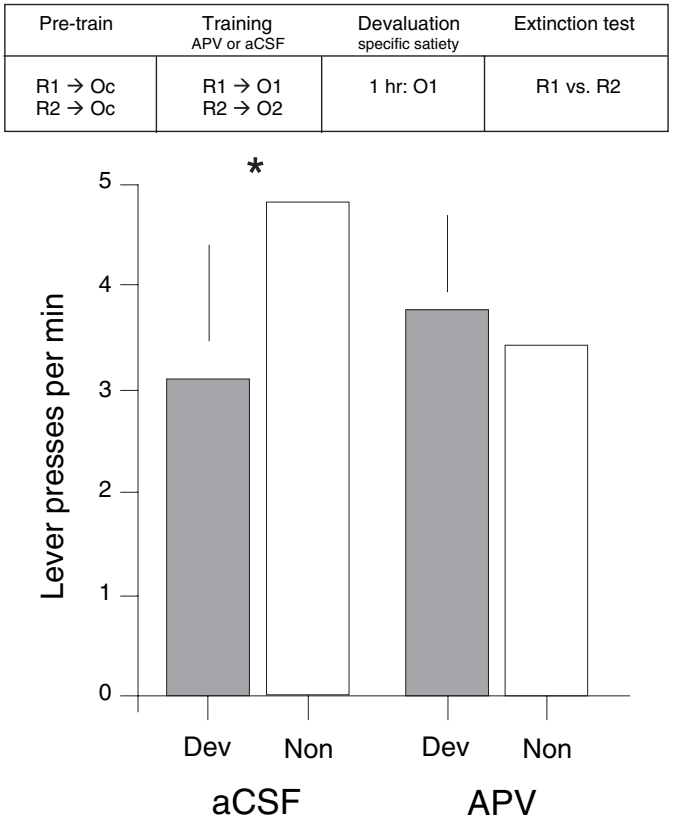

FIG. 2. Experiment 1. Experimental design and the results of the extinction test. R1 and R2. left and right lever presses; Oc, common outcome, 20% sucrose; O1 and O2, food pellet or peach punch outcomes. The graph presents the mean lever presses per min on both actions during the 5-min choice extinction test presented separately with rats infused with ACSF and with APV during the training phase and for the devalued action (Dev) that, in training, earned the devalued outcome and for the non-devalued action (Non) that, in training, earned the non-devalued outcome. Bars represent ± 1 standard error of the difference of the means. *$P < 0.05$.

sensitive to the selective satiety devaluation treatment whereas the choice performance of rats infused during training with APV appeared to be insensitive to the selective effects of outcome devaluation. Analysis of variance (ANOVA) found neither a main effect of group or of devaluation (largest $F_{1,25} = 2.35$, $P > 0.1$) but did find a significant interaction between these factors ($F_{1,25} = 5.86$, $P < 0.05$). Simple effects analyses conducted on this interaction showed that, whereas a significant devaluation effect emerged in the ACSF group, i.e. this group significantly reduced responding on the lever that, in training, had delivered the now devalued outcome ($F_{1,25} = 5.08$, $P < 0.05$), the APV group failed to show this effect responding similarly on the two levers ($F < 1$).

Experiment 2: Comparing the effects of APV infusion before and after training

A schematic illustration of the cannulae placements for the rats used in Experiment 2 is shown in Fig. 3. All rats learned to press both levers for sucrose during the five sessions of pre-training. As in Experiment 1, during the critical training session rats that received APV before the session pressed at a similar rate as those that were given the infusion after the session; rats given the APV infusion before training responded at 7.98 presses per min on the lever that delivered the to-be-devalued outcome and 7.74 presses on the other lever. Rats that were to receive the APV infusion after this training session responded at 10.2 and 5.83 lever presses per min, respectively. A two-way

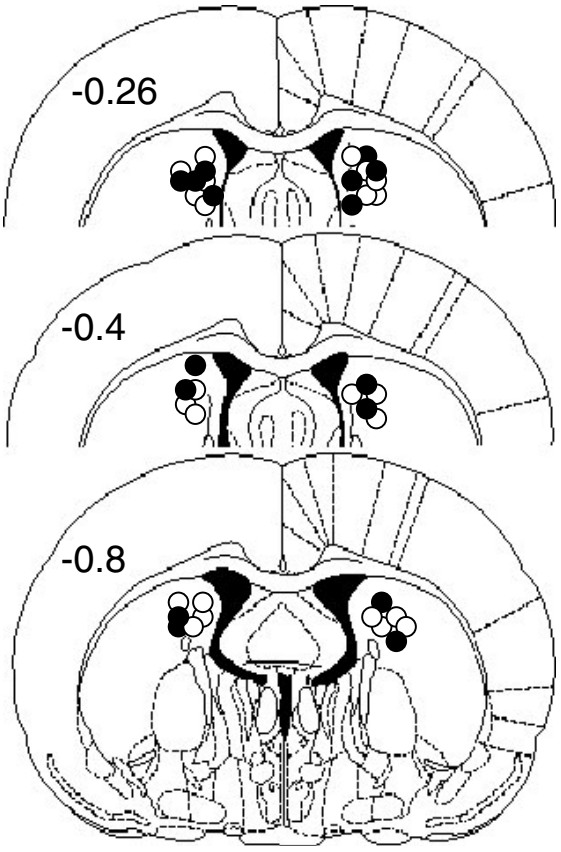

FIG. 3. Experiments 2. Schematic representation of the cannulae placement presented separately for the group infused with APV either before (white circles) or after (black circles) the training phase. The drawings of the coronal sections are taken −0.26, −0.4 and −0.8 mm relative to bregma (Paxinos & Watson, 1998).

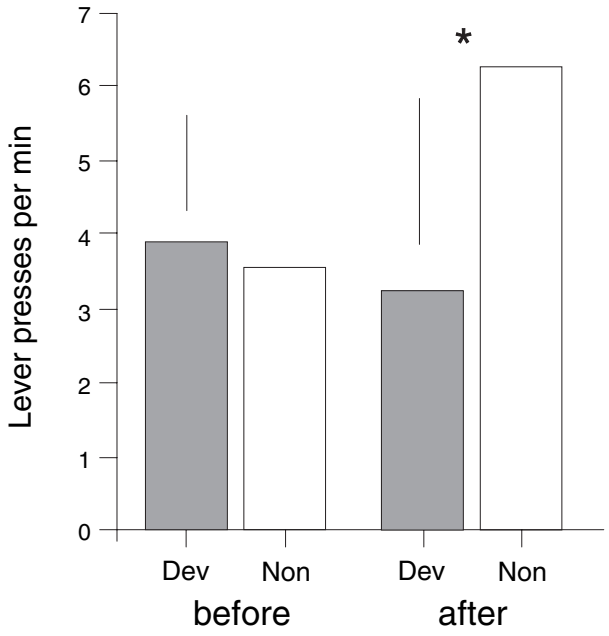

FIG. 4. Experiment 2. The graph presents the mean lever presses per min on both actions during the 5-min choice extinction test presented separately with rats infused with APV before training and those infused after training and for the devalued action (Dev) that, in training, earned the devalued outcome and for the non-devalued action (Non) that, in training, earned the non-devalued outcome. Bars represent ±1 standard error of the difference of the means; *$P < 0.05$.

ANOVA conducted on this training data using group (before vs. after) as a between-subjects factor and devaluation (separating performance on the to be devalued lever vs. from that on the other lever) as a within-subject factor revealed no main effects of either group ($F < 1$) or of lever ($F_{1,18} = 2.95$ $P > 0.05$), and no interaction between these factors ($F_{1,18} = 2.38$, $P > 0.05$; $M_{SE} = 41.50$).

The results from the choice extinction test conducted after outcome devaluation are presented in Fig. 4. First it should be noted that we were able to replicate the effect observed in Experiment 1. An infusion of APV given prior to instrumental training appeared to block action–outcome learning. In contrast, the same was not true of performance in the rats given the APV infusion 1 h after training. Performance in this group was sensitive to the selective satiety devaluation treatment; pressing was reduced on the lever that, in training, had delivered the now devalued outcome relative to the other action. The statistical analysis found neither a main effect of group ($F_{1,18} = 2.28$, $P > 0.05$) nor of devaluation ($F_{1,18} = 3.92$, $P > 0.05$) but found a significant interaction between these factors ($F_{1,18} = 6.85$, $P < 0.05$). Simple effects analyses conducted on this interaction revealed that, whereas rats that received APV after training significantly reduced responding for the devalued outcome ($F_{1,18} = 5.88$, $P < 0.05$), those that received APV before training failed to do so ($F < 1$).

Experiment 3: Comparing the effects of APV infusion into pDMS and the DLS

A schematic illustration of the cannulae placements for the rats used in Experiment 3 is shown in Fig. 5. As in previous studies, all of the rats in Experiment 3 learned to press both levers for sucrose during the five sessions of pre-training. Moreover, during the critical training session, rats that received an infusion of APV into the pDMS responded on the two levers at a similar rate to those given the

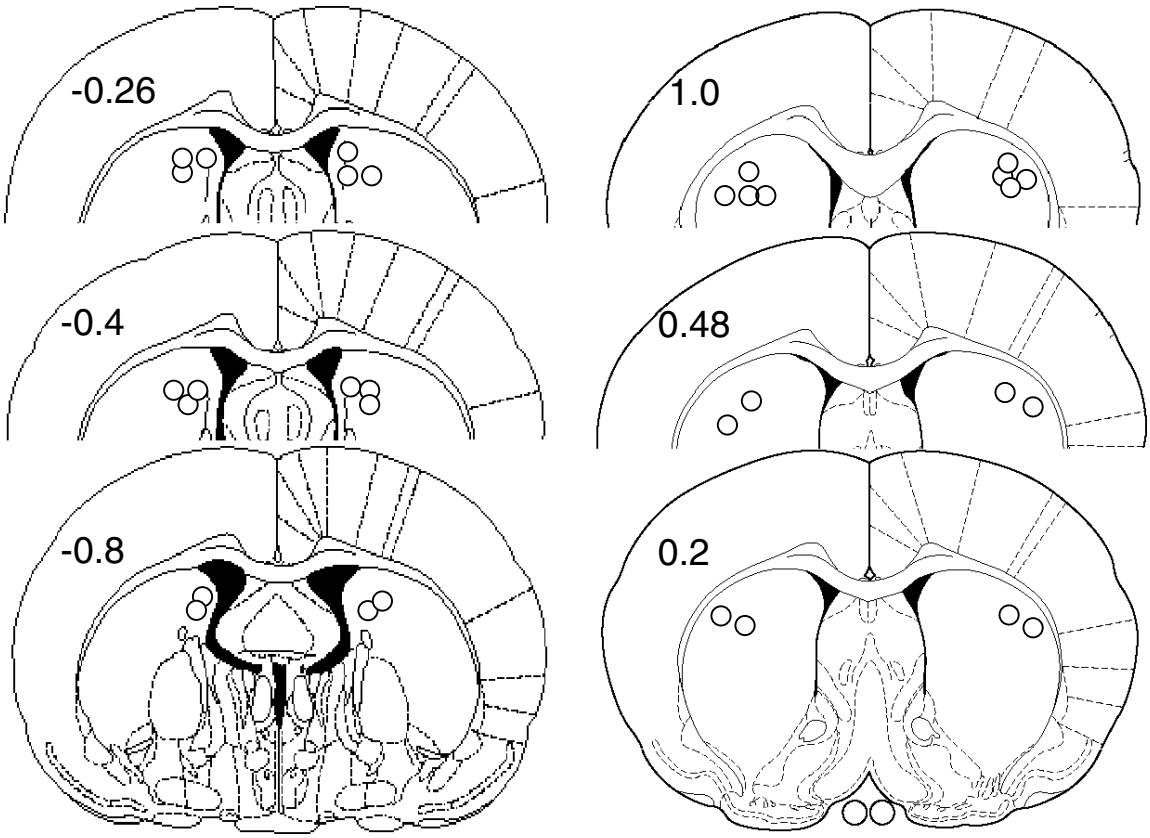

FIG. 5. Experiment 3. Schematic representation of the cannulae placement presented separately for the group infused with APV into the pDMS (left panels) and for the group infused with APV into the DLS (right panels). The drawings of the coronal sections are taken −0.26, −0.4 and −0.8 mm relative to bregma for the pDMS and 1 mm, 0.48 mm and 0.2 mm relative to bregma for the DLS (Paxinos & Watson, 1998).

infusion of APV into the DLS. During this session, rats given the APV infusion into the pDMS responded at 8.14 presses per min on the lever that delivered the to-be-devalued outcome and 7.73 presses on the other lever whereas rats given the APV infusion into the DLS responded at 4.90 and 4.12 lever presses per min, respectively. Although there was a numerical reduction in performance in the DLS group, a two-way ANOVA conducted using group (DLS vs. pDMS) as a between-subjects factor and devaluation (separating performance on the to-be-devalued lever from that on the other lever) as a within-subject factor, showed no significant effect of group, of lever or of the interaction between these factors (largest $F_{1,14} = 1.12$; $M_{SE} = 92.93$).

The results from the choice extinction test of Experiment 3 are presented in Fig. 6. Again we were able to replicate Experiments 1 and 2 finding that APV infused into the pDMS during instrumental training blocked action–outcome learning during that session. As shown in Fig. 6, no outcome devaluation effect was observed in the pDMS group; indeed, if anything, at least numerically this group responded more on the devalued than the non-devalued action. In contrast, the group given the infusion of APV into the DLS during training showed evidence of outcome devaluation. Although their performance was generally lower than in previous experiments, this group nevertheless reduced their performance of the action that, in training, had delivered the devalued outcome relative to the other action. The statistical analysis found neither a main effect of group or of devaluation (Fs < 1) but revealed a significant interaction between these factors ($F_{1,14} = 8.12$, $P < 0.05$). Simple effects analyses again revealed a significant outcome devaluation effect in the group given the infusion of APV into the DLS during training ($F_{1,14} = 4.88$, $P < 0.05$) but found no difference in performance on the two actions in the pDMS group ($F_{1,14} = 3.29$, $P > 0.05$).

Discussion

This series of experiments provides the first direct evidence that, in instrumental conditioning, the encoding of specific action–outcome associations, which are critical in the acquisition of goal-directed actions, may be localized to a discrete brain region, the pDMS, and that NMDAR-dependent plasticity in this region may mediate this type of learning. In all three experiments presented here, blockade of NMDARs in the pDMS prevented action–outcome learning, as indexed by outcome devaluation, without significantly altering the rats' performance of instrumental actions.

The basic design used in the current experiments was central to our approach. Rats were first trained to perform two lever-press actions for a common outcome before they were given a single session of training during which the two actions were rewarded with unique outcomes. This learning session was designed to allow the rats to encode unique action–outcome associations, and the outcome devaluation test was used to assay whether the rats had acquired these discrete relations between their actions and outcomes. In this test, the ability of the rats to choose between the two training actions was assessed in extinction, i.e. in the absence of reward, after one of the two outcomes had been devalued by specific satiety. Experiment 1 provided evidence that rats given an infusion of ACSF into the pDMS encoded the action–outcome associations during the

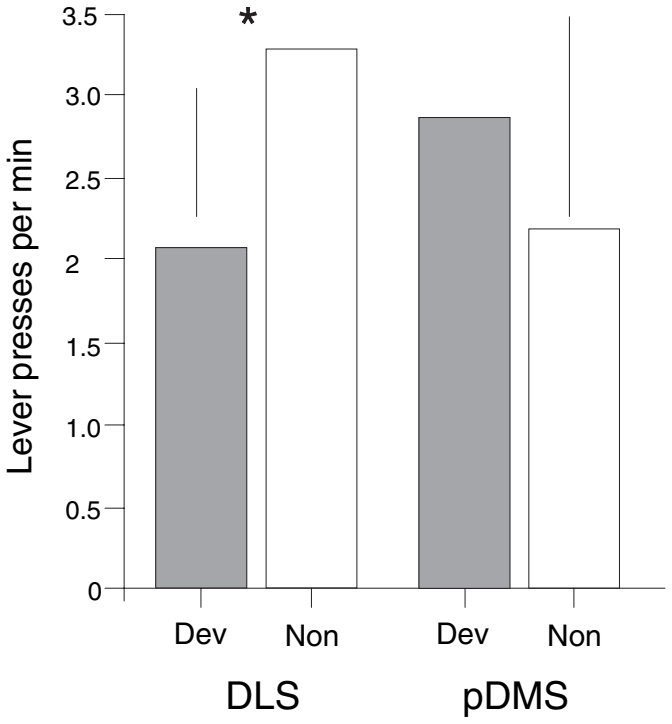

FIG. 6. Experiment 3. The graph presents the mean lever presses per min on both actions during the 5-min choice extinction test presented separately with rats infused with APV into the DLS and those infused with APV into the pDMS and for the devalued action (Dev) that, in training, earned the devalued outcome and for the non-devalued action (Non) that, in training, earned the non-devalued outcome. Bars represent ± 1 standard error of the difference of the means; $*P < 0.05$.

acquisition session; when the rats in this group were sated on one or other of the outcomes earned during acquisition, they subsequently performed fewer responses on the lever that had previously earned that outcome. In contrast, rats given the training session after an infusion of APV into the pDMS performed both actions at a similar rate during the test.

The results of Experiment 1 are consistent with the hypothesis that the pDMS is the critical site of the plasticity underlying instrumental learning. Experiments 2 and 3 provided further tests of this hypothesis by assessing the temporal and spatial specificity of the effects of APV on instrumental learning. These experiments also allowed us to address some alternative explanations. For example, it is possible that the infusion of APV into the pDMS modified performance on test not because it altered learning but because it had non-associative effects on subsequent performance. Nevertheless, whether characterized in terms of a general motor deficit, interference with retrieval or with some other general cognitive capacity, each of these accounts predicts that an infusion of APV given either before or after the initial session of instrumental training should produce a similar deficit in subsequent choice performance. Results from Experiment 2 established, however, that this was not the case. The choice performance of rats given an infusion of APV prior to instrumental training was again insensitive to outcome devaluation whereas choice performance in rats given the APV 1 h after the learning session remained sensitive to devaluation, indicating that action–outcome learning was intact in the latter group. This result therefore bolsters the claim that the effects of APV in Experiment 1 were specific to the learning processes engaged during instrumental training and were not due to an effect on performance.

Instrumental learning and performance

As pointed out in the Introduction, previous attempts to establish the locus of instrumental learning have suggested instead that regions of the ventral (Kelley *et al*., 1997) and posterolateral striatum (Andrzejewski *et al*., 2004) are critical for the formation of action–outcome associations. These conclusions were drawn on the basis of the finding that APV infused into these regions produces a deficit in the performance of lever pressing during instrumental acquisition. There are, however, several problems with this analysis. First and foremost, it conflates changes in performance with changes in learning. Although deficits in performance can sometimes be indicative of failures of learning, they can also be produced by a wide variety of non-associative factors. Conversely, a failure to observe a deficit in performance is no guarantee that the learning supporting performance is actually normal. For these reasons, when analysing the role of a brain structure in learning it is important to use behavioural assays that directly assess the content of learning, though unfortunately this is seldom done. Furthermore, and particularly in instrumental conditioning where performance can be controlled by several learning processes, measuring performance alone cannot tell us whether action–outcome learning, or some other learning process that can also influence performance, has been affected by the experimental manipulation in question.

In the current series of experiments, we consistently found that, whereas APV infused into the pDMS had no apparent effect on performance, it had a significant effect on learning. As the results of Experiment 2 make clear, the design used in the current study has substantial advantages over previous approaches. It allows us to assess changes in learning in the absence of any direct influence of drug on performance. And, more importantly, this design allows us to assess directly the content of learning using a clear-cut assay for action–outcome encoding rather than having to infer changes in learning from changes in performance.

Moreover, there is in fact direct evidence that the ventral striatum (including nucleus accumbens core and shell) does not mediate the formation of action–outcome associations (e.g. Balleine & Killcross, 1994; Corbit *et al*., 2001; de Borchgrave *et al*., 2002). In addition, the results of Experiment 3 suggest that, if the posterolateral striatum is involved in instrumental performance its role may be limited to the performance of habitual rather than goal-directed actions (cf. Yin *et al*., 2004). In Experiment 3, we compared the effects of the infusion of APV into pDMS with an infusion into the dorsolateral striatum (DLS). Replicating Experiments 1 and 2, it was found that the infusion of APV into the pDMS abolished the sensitivity of the rats' performance to outcome devaluation. In marked contrast, however, the APV infusion into the DLS did not disrupt action–outcome learning and the performance of the rats in the DLS group remained sensitive to outcome devaluation in the extinction test.

S–R learning in the striatum

The dorsolateral striatum, as a component of the sensorimotor cortico-basal ganglia circuit, has been implicated in the formation of stimulus–response (S–R) associations critical for procedural learning and the performance of habitual rather than goal-directed actions (Robbins & Everitt, 2002; White & McDonald, 2002; Yin *et al*., 2004). When instrumental responses are controlled by S–R associations they are impervious to devaluation manipulations because the outcome itself is not part of the associative structure controlling responding. Consequently, changes in outcome value can have no direct influence on performance (Dickinson & Balleine, 1994). In a recent paper we

reported direct evidence in support of this claim, showing that lesions of the dorsolateral but not the dorsomedial striatum rendered actions sensitive to outcome devaluation that had become habitual in intact rats (Yin *et al*., 2004). Conversely, the failure to observe sensitivity to selective outcome devaluation in rats given APV infusions into the pDMS suggests that instrumental performance in these rats was controlled by S–R processes. Previous studies have indeed observed a general decrease in response rates of habitual actions after satiety induced by pre-feeding (e.g. Dickinson *et al*., 1995), indicating that the rate of habitual responding can be directly influenced by the general motivational state of the animal. Thus, the non-specific decrease in response rate after pre-feeding seen in the animals that had received APV in the pDMS is consistent with the idea that these rats were responding habitually.

S–R learning plays a central role in reinforcement learning models currently popular in theories of basal ganglia functioning (Sutton & Barto, 1998). In line with the actor-critic formulation of these models, the dorsal striatum has recently been identified as the actor and the ventral striatum as the critic (O'Doherty *et al*., 2004). Nevertheless, these models of adaptive behaviour, as currently formulated, cannot explain the acquisition and performance of deliberate, goal-directed actions (Dayan & Balleine, 2002). The acquisition of these actions involves the formation of action–outcome associations, a process that, in view of our results, requires NMDAR activation in a distinct striatal region, the pDMS.

The function of striatal plasticity in cortico-striatal circuits

Although NMDAR-mediated plasticity such as long-term potentiation at the corticostriatal synapse has long been identified in the dorsomedial striatum (e.g. Partridge *et al*., 2000), to our knowledge this study presents the first evidence for one possible functional significance of this type of plasticity in behaviour. It remains unknown at present how this plasticity functions within the larger circuit known to contribute to instrumental performance. The existence of parallel cortico-basal-ganglia re-entrant loops has been well documented (Alexander *et al*., 1986; Kelly & Strick, 2004). Like other striatal regions, the dorsomedial striatum can send outputs that can ultimately re-enter the thalmocortical network (Groenewegen, 2003). In addition, outputs from the pDMS are also expected to influence downstream networks in the brainstem involved in motor control and behavioural arousal (Nauta, 1989). When considering the functional significance of pDMS plasticity in behaviour, it is therefore critical to consider this structure in the larger context of the functional circuit to which it belongs, as a critical component of the associative cortico-basal ganglia circuit. Viewed from this perspective, action–outcome learning could take place in this region through the association of converging inputs from the cortex, particularly the prefrontal region, that are involved in the representation of actions and outcomes. NMDAR-dependent plasticity in the pDMS is a plausible candidate mechanism for this type of association, although this hypothesis remains to be tested.

In conclusion, whatever the functional importance of plasticity in the dorsomedial striatum turns out to be, our data serve to reinforce the more general claim that the striatum, like the cortex, is functionally heterogeneous (Devan *et al*., 1999; Robbins & Everitt, 2002; Palencia & Ragozzino, 2004). Analyses of learning processes in the dorsal striatum have previously focused on the function of this region in procedural learning involving the acquisition and performance of habitual actions mediated by S–R associations (Jog *et al*., 1999; Packard & Knowlton, 2002). There is now evidence sufficient to propose that two distinct learning processes take place within the dorsal striatum: one in the dorsolateral (or sensorimotor) striatum that mediates stimulus–response habits and a second in the dorsomedial (or associative) striatum that mediates goal-directed actions. These striatal regions appear to be critical components of distinct neural circuits that mediate distinct types of learning. Establishing the way in which these learning processes interact to control instrumental performance and the rules that control plasticity in these regions appears now to be a viable task for future research.

Acknowledgements

This research was supported by UCLA Dissertation Year Fellowship to HHY, NSF grant 9985417 to BJK, and NIMH grant MH56446 to BWB.

Abbreviations

APV, 2-amino-5-phosphonopentanoic acid; NMDARs, *N*-methyl-D-aspartate receptors; ACSF, artificial cerebral spinal fluid; DLS, dorsolateral striatum; DMS, dorsomedial striatum; pDMS, posterior dorsomedial striatum; 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.

Alexander, G.E., DeLong, M.R. & Strick, P.L. (1986) Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.*, **9**, 357–381.

Andrzejewski, M.E., Sadeghian, K. & Kelley, A.E. (2004) Central amygdalar and dorsal striatal NMDA receptor involvement in instrumental learning and spontaneous behavior. *Behav. Neurosci.*, **118**, 715–729.

Balleine, B.W. & Killcross, A.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.

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

Calabresi, P., Pisani, A., Mercuri, N.B. & Bernardi, G. (1992) Long-term potentiation in the striatum is unmasked by removing the voltage-dependent magnesium block of NMDA receptor channels. *Eur. J. Neurosci.*, **4**, 929–935.

Colwill, R.C. & Rescorla, R.A. (1986) Associative structures in instrumental learning. In G.H. Bower (Ed.) *The Psychology of Learning and Motivation*, Vol. 20, Academic Press, New York, pp. 55–104.

Corbit, L.H. & Balleine, B.W. (2003) The role of the prelimbic cortex in instrumental conditioning. *Behav. Brain Res.*, **146**, 145–157.

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.

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

Devan, B.D., McDonald, R.J. & White, N.M. (1999) Effects of medial and lateral caudate-putamen lesions on place- and cue-guided behaviors in the water maze: relation to thigmotaxis. *Behav. Brain Res.*, **100**, 5–14.

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

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

Everitt, B.J., Cardinal, R.N., Parkinson, J.A. & Robbins, T.W. (2003) Appetitive behavior: Impact of amygdala-dependent mechanisms of emotional learning. *Ann. N.Y. Acad. Sci.*, **985**, 233–250.

Groenewegen, H.J. (2003) The basal ganglia and motor control. *Neural Plast.*, **10**, 107–120.

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

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.

Kelley, A.E., Smith-Roe, S.L. & Holahan, M.R. (1997) Response-reinforcement learning is dependent on *N*-methyl-D-aspartate receptor activation in the nucleus accumbens core. *Proc. Natl Acad. Sci.*, **94**, 12174–12179.

Kelly, R.M. & Strick, P.L. (2004) Macro-architecture of basal ganglia loops with the cerebral cortex: use of rabies virus to reveal multisynaptic circuits. *Prog. Brain Res.*, **143**, 449–459.

Lovinger, D.M., Partridge, J.G. & Tang, K.C. (2003) Plastic control of striatal glutamatergic transmission by ensemble actions of several neurotransmitters and targets for drugs of abuse. *Ann. N.Y. Acad. Sci.*, **1003**, 226–240.

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.

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, Baselm, pp. 43–63.

O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K. & Dolan, R.J. (2004) Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science*, **304**, 452–454.

Ossowska, K. & Wolfarth, S. (1995) Stimulation of glutamate receptors in the intermediate/caudal striatum induces contralateral turning. *Eur. J. Pharmacol.*, **273**, 89–97.

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

Palencia, C.A. & Ragozzino, M.E. (2004) The influence of NMDA receptors in the dorsomedial striatum on response reversal learning. *Neurobiol. Learn. Mem.*, **82**, 81–89.

Partridge, J.G., Tang, K.C. & Lovinger, D.M. (2000) Regional and postnatal heterogeneity of activity-dependent long-term changes in synaptic efficacy in the dorsal striatum. *J. Neurophysiol.*, **84**, 1422–1429.

Passingham, R.E., Myers, C., Rawlins, N., Lightfoot, V. & Fearn, S. (1988) Premotor cortex in the rat. *Behav. Neurosci.*, **102**, 101–109.

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.

Robbins, T.W. & Everitt, B.J. (2002) Limbic-striatal memory systems and drug addiction. Neurobiol. *Learn. Mem.*, **78**, 625–636.

Sutton, R.S. & Barto, A.G. (1998) *Reinforcement Learning*. MIT Press, Cambridge, Mass.

Wang, S.-H., Ostlund, S.B., Nader, K. & Balleine, B.W. (2005) Consolidation and reconsolidation of incentive learning in the amygdala. *J. Neurosci.*, **25**, 830–835.

White, N.M. & McDonald, R.J. (2002) Multiple parallel memory systems in the brain of the rat. *Neurobiol. Learn. Mem.*, **77**, 125–184.

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., Ostlund, S.B., Knowlton, B.J. & Balleine, B.W. (2005) The role of the dorsomedial striatum in instrumental conditoning. *Eur. J. Neurosci.*, **22**, 513–523.