European Journal of Neuroscience, Vol. 19, pp. 181–189, 2004

Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning

Henry H. Yin, Barbara J. Knowlton and Bernard W. Balleine
Department of Psychology, and Brain Research Institute, University of California, Los Angeles, 1285 Franz Hall, Los Angeles, CA 90095–1563, USA

Keywords: basal ganglia, dorsomedial striatum, instrumental conditioning, interval schedules, procedural learning, striatum

Abstract

Habits are controlled by antecedent stimuli rather than by goal expectancy. Interval schedules of feedback have been shown to generate habits, as revealed by the insensitivity of behaviour acquired under this schedule to outcome devaluation treatments. Two experiments were conducted to assess the role of the dorsolateral striatum in habit learning. In Experiment 1, sham operated controls and rats with dorsolateral striatum lesions were trained to press a lever for sucrose under interval schedules. After training, the sucrose was devalued by inducing taste aversion to it using lithium chloride, whereas saline injections were given to the controls. Only rats given the devaluation treatment reduced their consumption of sucrose and this reduction was similar in both the sham and the lesioned groups. All rats were then returned to the instrumental chamber for an extinction test, in which the lever was extended but no sucrose was delivered. In contrast to sham operated controls, rats with dorsolateral striatum lesions refrained from pressing the lever if the outcome was devalued. To assess the specificity of the role of dorsolateral striatum in this effect a second experiment was conducted in which a group with lesions of dorsomedial striatum was added. In relation now to both the sham and the dorsomedial lesioned groups, only rats with lesions of dorsolateral striatum significantly reduced responding after outcome devaluation. In conclusion, this study provides direct evidence that the dorsolateral striatum is necessary for habit formation. Furthermore, it suggests that, when the habit system is disrupted, control over instrumental performance reverts to the system controlling the performance of goal-directed instrumental actions.

Introduction

Recent studies of instrumental conditioning have shown that two different processes control the acquisition and performance of actions, such as lever pressing, which lead to rewards (Dickinson *et al*., 1995; Balleine & Dickinson, 1998). During acquisition, performance is largely goal-directed, controlled by the action–outcome contingency. Hence post-training devaluation of the outcome reduces the tendency of rats to perform an action leading to it (Adams, 1981; Adams & Dickinson, 1981; Balleine, 2001). With continued training, however, control over performance shifts from its consequences to its antecedents, i.e. responding becomes habitual or stimulus-driven, no longer sensitive to outcome devaluation treatments. This shift in control of instrumental performance from a goal-directed to a habit process has been argued to reflect the increasing involvement of stimulus–response (or S–R) associations as training proceeds (Dickinson, 1989; Dickinson *et al*., 1995a).

This effect of extended training is not due simply to rewarded or unrewarded repetition of an action. Rather, it appears to reflect a reduction, over time, in the perceived correlation between rate of performance and rate of outcome delivery (Baum, 1973; Dickinson, 1989, 1994). Consistent with this claim, instrumental performance can rapidly become habitual when the contingency between response and outcome delivery is weakened using interval schedules of reinforcement. Dickinson *et al.* (1983) compared ratio and interval schedules directly, and found that after post-training devaluation of the reward, only those trained on ratio schedules reduced responding compared with non-devalued controls; rats trained on interval schedules, by contrast, performed similarly whether the sucrose had been devalued or not.

Results such as these have been used to support the argument that goal-directed and habitual instrumental actions are controlled by distinct learning processes; the former involving the formation of action–outcome (R–O) associations that mediate flexible, deliberated actions and the latter involving the integration of sensory and motor representations in the classic S–R/reinforcement architecture mediating inflexible and automatic habitual behaviour (Dickinson, 1985, 1994). Furthermore, on the basis of the distinction between the content of these associative learning processes, it has been claimed that they may also involve distinct memory systems, e.g. the relations critical for deliberated action have been argued to be encoded in declarative memory whereas the critical S–R associations mediating habitual actions are encoded in procedural memory (Dickinson, 1980, 1989; Knowlton *et al*., 1996).

In rats, procedural memory is often thought to depend on the integrity of the dorsal striatum, particularly the more lateral region of this area. The extensive connections of the dorsolateral striatum (DLS) with the sensorimotor neocortex presents a neural circuitry well suited for S–R learning (Packard & McGaugh, 1996; Graybiel, 1998; Devan *et al*., 1999). Indeed, previous studies have provided strong, if

Correspondence: H. Yin, as above.
E-mail: hyin@ucla.edu

Received 7 August 2003, revised 16 October 2003, accepted 20 October 2003

doi:10.1046/j.1460-9568.2003.03095.x

somewhat indirect, evidence that the DLS is involved in habit formation (Devan *et al*., 1999; Jog *et al*., 1999; Packard & Knowlton, 2002).

These studies suggest that the transition in behavioural control, from a goal-directed to an habitual, S–R, process, requires plasticity in the DLS. Damage to this region should therefore be predicted to increase sensitivity to outcome devaluation even when rats are trained on interval schedules. We tested this hypothesis in Experiment 1 by assessing the effect of pre-training excitotoxic lesions of DLS on sensitivity to outcome devaluation induced after lever pressing had been trained under interval schedules. In Experiment 2, we replicated and extended Experiment 1 by including a group with lesions of the dorsomedial striatum (DMS) and by analysing lever press acquisition in more detail.

Materials and methods

This study was approved by the animal research committee at UCLA.

Experiment 1

Subjects and apparatus

In Experiment 1, 39 experimentally naïve male Long–Evans rats weighing between 400 and 480 g were used. The rats were housed singly and were handled daily for 1 week prior to surgery. Training and testing took place in 16 Medical Associates (East Fairfield, VT, USA) operant chambers housed within sound- and light-resistant shells. Each chamber was equipped with a pump fitted with a syringe that delivered 0.1 mL of a 20% sucrose solution into a recessed magazine in the chamber when activated, and two retractable levers on either side of the food magazine. A 3-W, 24-V house-light, mounted on the top-centre of the wall opposite the magazine, provided illumination. Microcomputers equipped with the MED-PC program (Medical Associates) controlled the equipment and recorded the lever-presses.

Surgery and histology

One group was given lesions of the DLS and the other given sham lesions. 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: 0.7 mm anterior to bregma and 3.6 mm lateral to midline; and 5 mm below the skull surface. For the lesion group ($n = 18$), 0.6 μL of *N*-methyl-D-aspartate (NMDA; 0.12 M; Sigma, St Louis, MO, USA) was infused into the DLS on each side over 3 min; for the sham group ($n = 21$), no infusion was given. Five 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 10% formalin solution for at least 48 h and then transferred to a 25% sucrose–formalin solution for at least 72 h before 55-μm coronal sections were cut throughout the anterior striatum. The slices were stained with thionin and examined. Determination of lesion extent was made blind to behavioural performance.

Lever press training

One week after surgery, rats were placed on a food-deprivation schedule, receiving 10–15 g of their maintenance diet daily sufficient to reduce and to maintain them at 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 cage.

Rats were first given three, 30-min, magazine training sessions in which the sucrose solution was delivered on a random time 60-s schedule with the lever removed. This experiment was conducted in two replications. For 21 rats (ten lesioned and 11 sham), lever press training consisted of three 30-min sessions in which the sucrose was delivered on continuous reinforcement, followed by two sessions each in which it was delivered on a random interval 15-s (RI-15), an RI-30 and then an RI-60 schedule, with one session per day. The other 18 rats (eight lesioned and ten sham) were run together as a single replication. They began training with the sucrose delivered on an FI-20 schedule instead of continuous reinforcement. These FI-20 sessions (60 min maximum, one session per day) ended as soon as a total of 100 sucrose rewards had been earned. All rats earned the reinforcers in four sessions and then received two sessions of training on the RI-30 and on the RI-60 schedules of reinforcement. For all rats in both replications each training session began with the illumination of the house light and insertion of the lever, and ended with the retraction of the levers and turning off of the house light.

Outcome devaluation by conditioned taste aversion

The day after the conclusion of training the sucrose reward was devalued using conditioned taste aversion. For each of three daily sessions the rats were allowed to drink the sucrose from a glass bottle attached to their home cages for 30 min. Immediately thereafter, 12 lesioned and 11 sham rats received an intraperitoneal injection of lithium chloride (0.15 M LiCl, 20 mL/kg), while six lesioned and ten sham rats received saline injections (0.15 M, 20 mL/kg).

Extinction test

The day after the final conditioning session, all the rats received a 5-min extinction test on the levers. The test 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.

Experiment 2

Subjects

The subjects were 60 experimentally naïve female Long–Evans rats weighing between 260 and 310 g at the start of the experiment. Rats were housed in the manner described for Experiment 1 and were trained and tested in the same operant chambers.

Surgery and histology

The procedures for surgery and histology were identical to those described for Experiment 1, except that a DMS group has been added, with the following coordinates: 0.7 mm anterior to bregma; 2 mm lateral to midline; and 5 mm ventral to the skull surface.

Procedure

The training procedures were similar to those used in the second replication described for Experiment 1. After magazine training, all rats were trained to lever press for sucrose on an FI-20 schedule of reinforcement in order to allow an assessment of changes in the rate of responding produced by each reward delivery during acquisition. These FI-20 sessions were maximally 90 min in duration, once per day and ended as soon as 100 sucrose rewards had been earned. All rats earned the reinforcers in three sessions. These training sessions were followed by two sessions in which the sucrose was delivered on an RI-30 schedule and two sessions in which it was delivered on an RI-60 schedule.

The procedures used in devaluation and the final extinction test were exactly as described for Experiment 1.

Results

NMDA lesions of the striatum (Experiments 1 & 2)

Figure 1 provides a schematic representation of the extent of damage to the striatum caused by NMDA infusions. Inspection of the stained tissue did not reveal damage outside the dorsal striatum, except in rats that were excluded as follows. Four rats with DLS lesions were excluded in Experiment 1. In all four cases, there was significant damage to the overlying cortex above the DLS. Five rats were excluded in Experiment 2. Of these, three were in the DLS group and two were in the DMS group. The reason for exclusion was also damage to the cortex above the lesion sites. Final group sizes for Experiment 1 were: DLS devalued ($n = 8$), DLS valued ($n = 6$), sham devalued ($n = 11$), sham valued ($n = 10$). Final groups sizes for Experiment 2 were: DLS devalued ($n = 9$), DLS valued ($n = 10$), DMS devalued ($n = 11$), DMS valued ($n = 11$), sham devalued ($n = 7$), sham valued ($n = 7$).

Figure 2a and b provides representative photomicrographs of both lesions and sham controls. In all lesioned animals, cell loss and gliosis were seen in the area beneath the lateral corpus callosum. The extensive tissue shrinkage resulted in a visible widening of the lateral ventricles.

Experiment 1

Our initial analysis of the results of Experiment 1 included replication as a factor but, because there was no significant main effect, or any interaction involving this factor (all F-values < 1), the data were collapsed across replication for presentation and analysis.

Training on the levers proceeded smoothly. In the sessions in which all animals received the same procedures, i.e. when the sucrose was delivered on the RI-30 and RI-60 schedules of reinforcement, performance was comparable across replications. Figure 4a shows the response rates on the last day of training on the RI-60 schedule. At the end of training, rats in both groups were allocated to LiCl and saline groups for the devaluation phase.

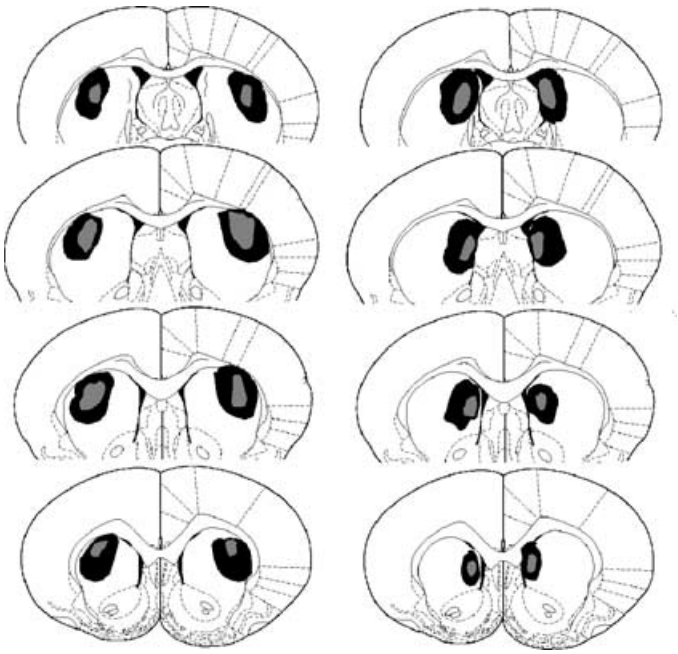

FIG. 1. Schematic representation of excitotoxic lesions of the DLS (left panel) and DMS (right panel). Shaded areas represent the maximum (dark grey) and minimum (light grey) extent of the lesions for rats included in the behavioural analyses. Coronal sections are illustrated for −0.26, 0.48, 1.0 and 1.6 mm (top to bottom, respectively) in the anteroposterior plane relative to bregma (adapted from Paxinos & Watson, 1998).

Outcome devaluation

The data from the devaluation phase are presented in Fig. 3. Clearly, taste aversion learning was not affected by the lesion, and LiCl injection produced a robust aversion in both groups. By contrast, all rats receiving saline injections continued to consume the sucrose across sessions. A mixed ANOVA conducted using group, devaluation and session as factors revealed main effects of devaluation ($F_{1,31} = 34.5$, $P < 0.05$) and of session ($F_{2,31} = 7.94$, $P < 0.05$), as well as an interaction between these factors ($F_{2,31} = 14.1$, $P < 0.05$), indicating that the devaluation effect became stronger over sessions. There was, however, no main effect of the lesion or any significant interactions involving this factor (largest $F_{1,31} = 1.52$, $P > 0.05$).

Extinction test

The results from the extinction test in Experiment 1 are presented in Fig. 4b as a percentage of the mean response rate on the last day of training. Inspection of this figure suggests that, although the devaluation of the sucrose was effective in sham controls, it did not transfer to the test; lever pressing appeared to be similar in both sham groups whether the sucrose had been paired with an injection of LiCl or with saline. A two-way ANOVA, with lesion type (DLS, sham) and devaluation treatment (devalued, non-devalued) as factors, revealed a significant interaction ($F_{1,31} = 4.4$, $P < 0.05$). When devalued and valued conditions were compared within each lesion type, a simple main effect of devaluation emerged in the DLS group ($F_{1,31} = 4.7$, $P < 0.05$), but not in the sham group ($F < 1$).

Experiment 2

In Experiment 2 we compared the rate of acquisition of lever pressing in all of the rats in each of the lesion conditions on an FI-20 schedule of reinforcement. This procedure enabled us to examine the impact of each reward delivery on subsequent performance both in terms of responding in the subsequent 20-s interval and also in terms of the inter-reinforcement interval (IRI). These data are presented in Fig. 5 averaged across bins of ten reinforcers for the first 100 reinforced actions. It appears from this figure that the groups acquired lever pressing at comparable rates but to quite different asymptotes, i.e. that the impact of reward was similar on learning, but different on performance.

A repeated-measures ANOVA was conducted on the responses per reinforcer data (Fig. 5a), with bins and lesion type as factors. It revealed significant main effects of lesion type ($P < 0.01$), of bin ($P < 0.01$) and, importantly, a significant interaction between these factors ($F_{18,468} = 15$, $P < 0.01$). Fisher's protected least significant difference test (PLSD) was used to compare the groups at different points during training. Significant differences emerged during the course of training: starting in bin 5, the DMS group made more responses per reinforcer than the sham controls, and this pattern persisted until the end of FI-20 training ($P < 0.01$ for all comparisons). Starting in bin 3 the DLS group made fewer responses per reinforcer than the sham controls, and this difference also persisted ($P < 0.01$ for all but bins 7 and 8).

Outcome devaluation

The data from the devaluation phase are presented in Fig. 6. Taste aversion learning was not affected by either the dorsolateral or the dorsomedial lesions. All rats that received LiCl and sucrose pairings reduced consumption of sucrose. A repeated-measures ANOVA on

sucrose consumption showed a main effect of devaluation ($F_{1,49} = 69.4$, $P < 0.001$), of session ($F_{2,49} = 40.1$, $P < 0.001$) and an interaction between these factors ($F_{2,49} = 57.2$, $P < 0.001$), but no main effect of lesion type nor any interaction involving this factor (all F-values < 1).

Extinction test

The mean response rates for each of the groups from the critical extinction test are presented in Fig. 7b as a percentage of the rate of lever pressing on the last day of training. As in Experiment 1, devaluation did not affect responding in the sham group, but a robust devaluation effect emerged in the DLS group. This effect appeared to be selective to damage to the DLS: rats with DMS lesions did not show a clear effect of devaluation. A two-way ANOVA revealed a significant interaction between lesion type and devaluation treatment ($F_{2,49} = 3.5$, $P < 0.05$). Simple main effects analyses conducted on the significant interaction revealed that neither the sham controls ($F < 1$) nor rats with

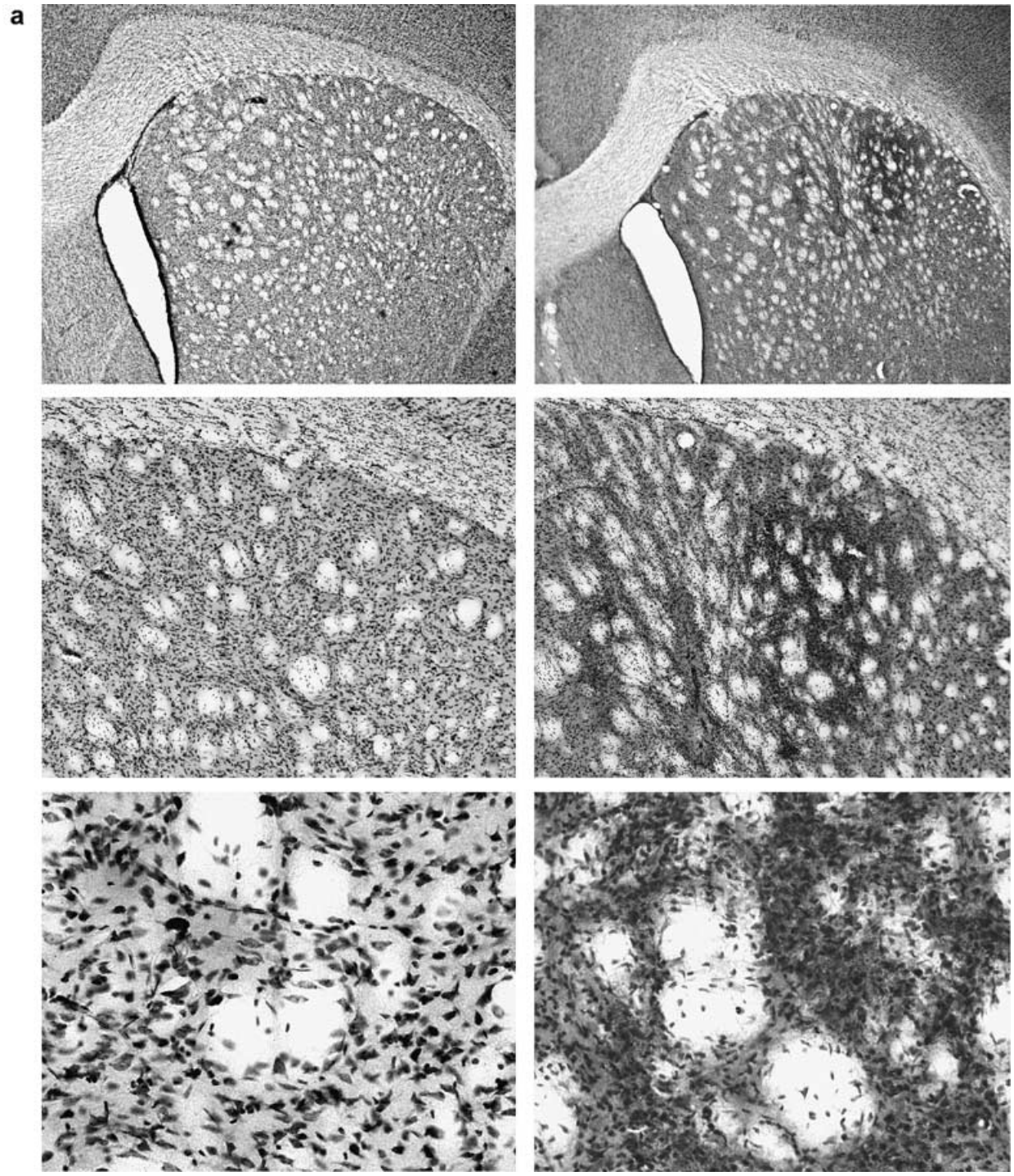

FIG. 2. (a) Photomicrographs of representative lesions of the DLS (right panels) and sham controls (left panels) shown, from top to bottom, in progressively higher levels of magnification. (b) Photomicrographs of representative lesions of the DMS (right panels) and sham control (left panels) shown, from top to bottom, in progressively higher levels of magnification.

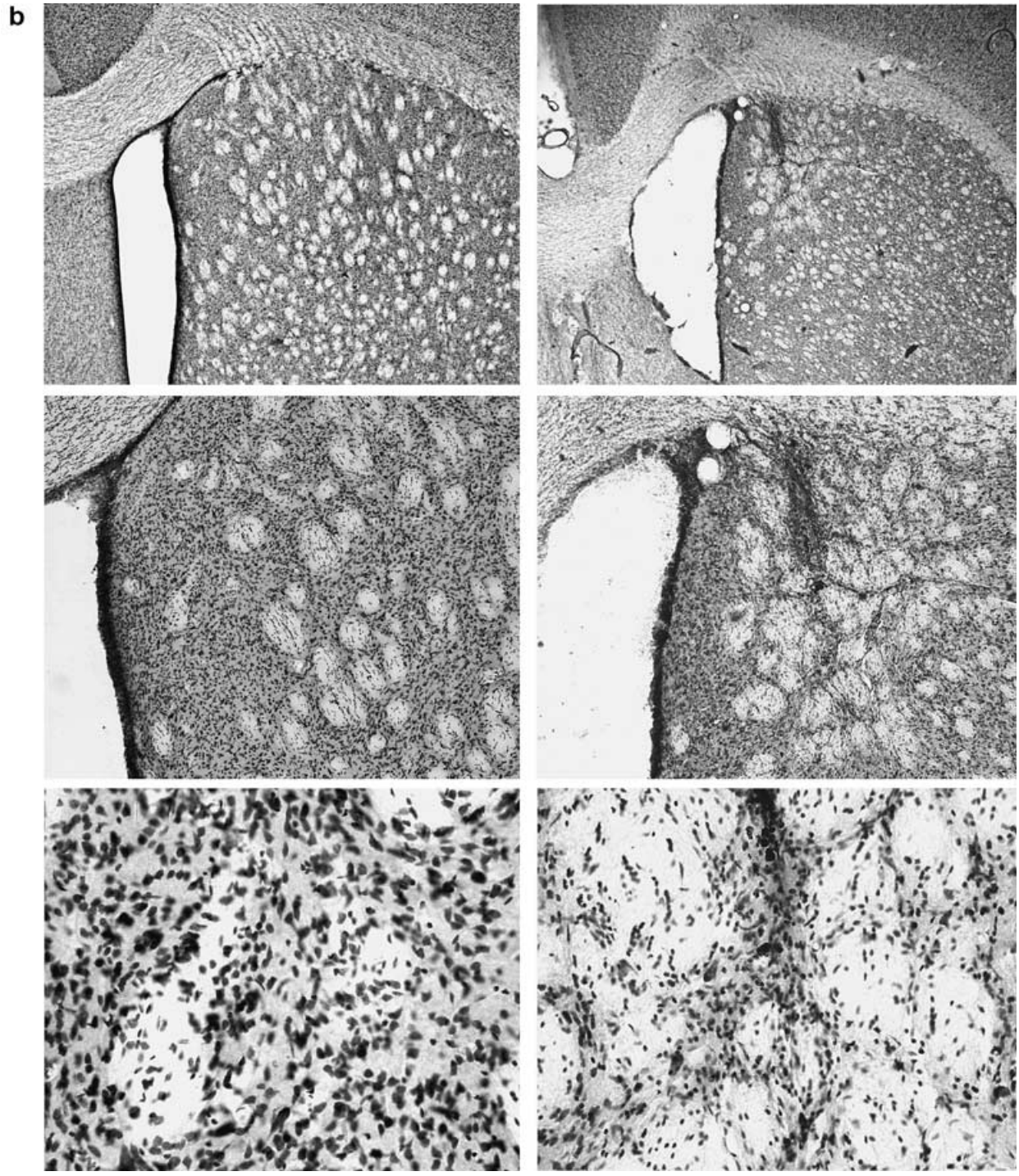

FIG. 2. continued

DMS lesions ($F_{2,49} = 1.9$, $P > 0.05$) were sensitive to devaluation treatment, whereas the DLS rats significantly reduced lever pressing in extinction after the devaluation treatment ($F_{2,49} = 8.7$, $P < 0.001$).

It is important to note that these effects were confined to the extinction test and were not present in the final training session (RI-60) – see Fig. 7a. A two-way ANOVA conducted on these data revealed no main effect of lesion (DLS or sham), of devaluation (to-be-devalued or to-be-valued) nor any interaction between these two factors (all F-values < 1).

Discussion

This study directly assessed the involvement of the dorsal striatum in habit learning using a post-training outcome devaluation procedure as a behavioural assay of habitual control of instrumental performance. It is well established that training rats to press a lever for a reward on an interval schedule of reinforcement rapidly renders performance of that action habitual, i.e. performance rapidly becomes insensitive to post-training changes in the value of the outcome. If this insensitivity is

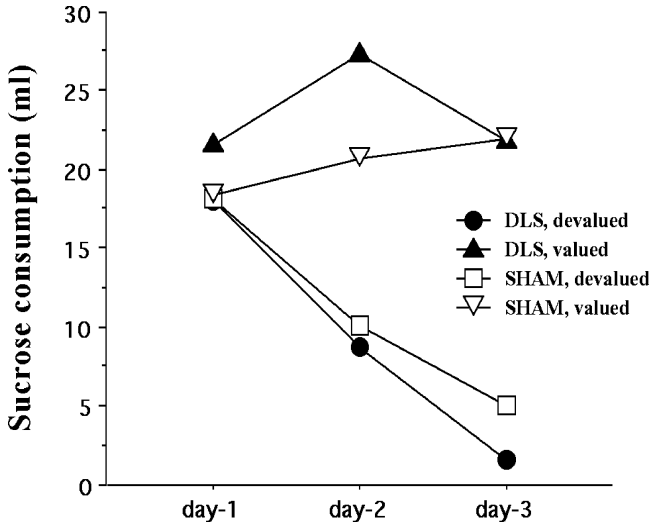

FIG. 3. Sucrose consumption over 3 days of taste aversion training for rats receiving LiCl injections (devalued) and saline injections (valued). DLS, dorsolateral striatum lesioned group; SHAM, sham operated control group.

induced by the development of an S–R habit involving the DLS, then lesions of this area should not only impair acquisition, but also increase the sensitivity of performance to outcome devaluation. To test this hypothesis, groups of lesioned and sham operated rats were trained to press a lever for a sucrose reward delivered on an interval schedule. The sucrose was then devalued for half of the rats by pairing its consumption with illness induced by an injection of lithium chloride before the effect of this treatment was assessed in an extinction test. Consistent with well-established findings in normal rats from behavioural experiments, although the sham controls stopped consuming the sucrose, when returned to the instrumental chamber with the lever extended, they performed similarly to rats for which the sucrose was not devalued. In striking contrast, instrumental performance in rats given lesions of the DLS was found to be highly sensitive to the current value of the training sucrose reward; rats with DLS lesions reduced their performance on the lever after the sucrose was devalued. This finding suggests that instrumental performance in the DLS rats was controlled by goal expectancy; when the value of that goal was reduced, their performance was also reduced.

To confirm this observation and assess its specificity, in Experiment 2 we replicated Experiment 1 and added a third group with NMDA lesions of the DMS. We replicated the result in the sham and dorsolateral groups in Experiment 1. Moreover, only the DLS was found to be required for habit learning, as rats with lesions of the DMS also responded habitually: they were as insensitive as shams to the effects of outcome devaluation.

One potential alternative interpretation of these results remains to be addressed. In these experiments devaluation training was conducted in the home cage and, as a consequence, in the absence of direct evidence that that the aversion to sucrose transferred effectively to the instrumental chamber, we cannot rule out the possibility that the insensitivity to devaluation in the sham controls was due to context specificity of the aversion. There are, however, a number of reasons to resist this interpretation of the data. First, although context specificity of taste aversion learning has been reported, it is a weak effect. There is no doubt that aversions can be conditioned to contextual cues, but typically post-conditioning changes in context appear only mildly to affect the rate of the extinction of the aversion over many repeated presentations of the aversive solution itself (Bonardi *et al*., 1990), and

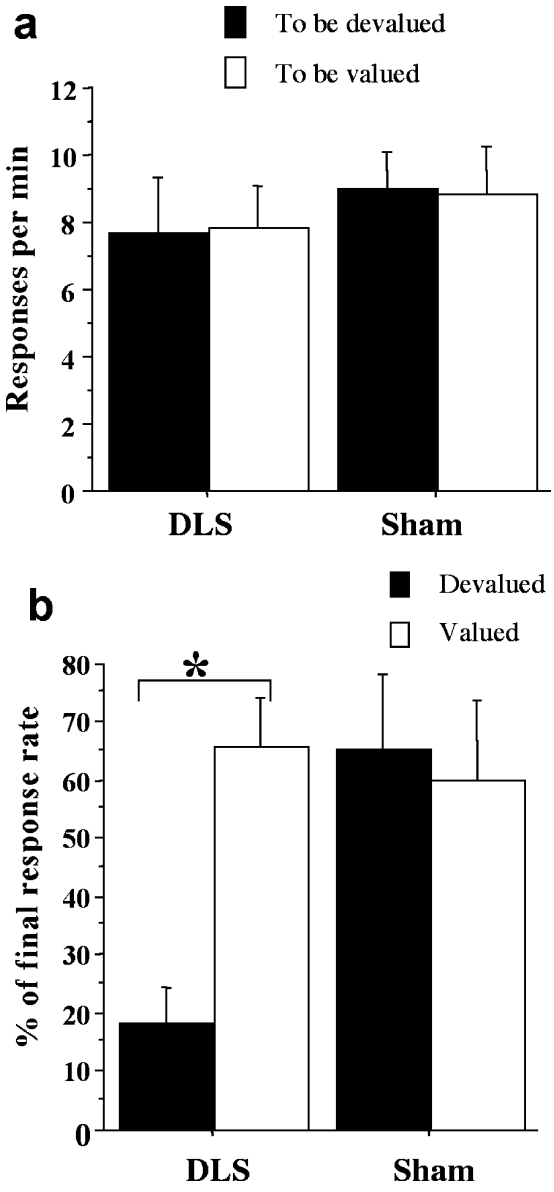

FIG. 4. Performance on a 5-min extinction test. (a) Response rates on the last day of training, on an RI-60 schedule. (b) Response rate during a 5-min extinction test after outcome devaluation as a percentage of the response rate on the last day of training. Error bars show one SEM; $^{*}P < 0.05$; DLS, dorsolateral strial lesioned group; Sham, sham operated controls.

do not influence the size of the initial aversion. Furthermore, there is good evidence that outcome devaluation of instrumental outcomes using taste aversion procedures transfers well between potent discriminative cues (Colwill & Rescorla, 1990; Lopez *et al*., 1992). And, indeed, the devaluation procedures used in the current experiments replicated those reported in the Dickinson *et al*. (1983) study on which this series was based and in which perfectly good transfer of the aversion from the home cage to the operant chamber was observed. Perhaps more to the point, to our knowledge there is no evidence that either the DLS or the striatum in general is involved in the control, contextual or otherwise, of taste aversion learning or that lesions of the DLS impair context discrimination, as would be required in this case. In fact, to the contrary, a well-established finding from studies of maze learning is that rats with lesions of the DLS tend to be *more* responsive to contextual cues than are unlesioned controls (Packard & McGaugh,

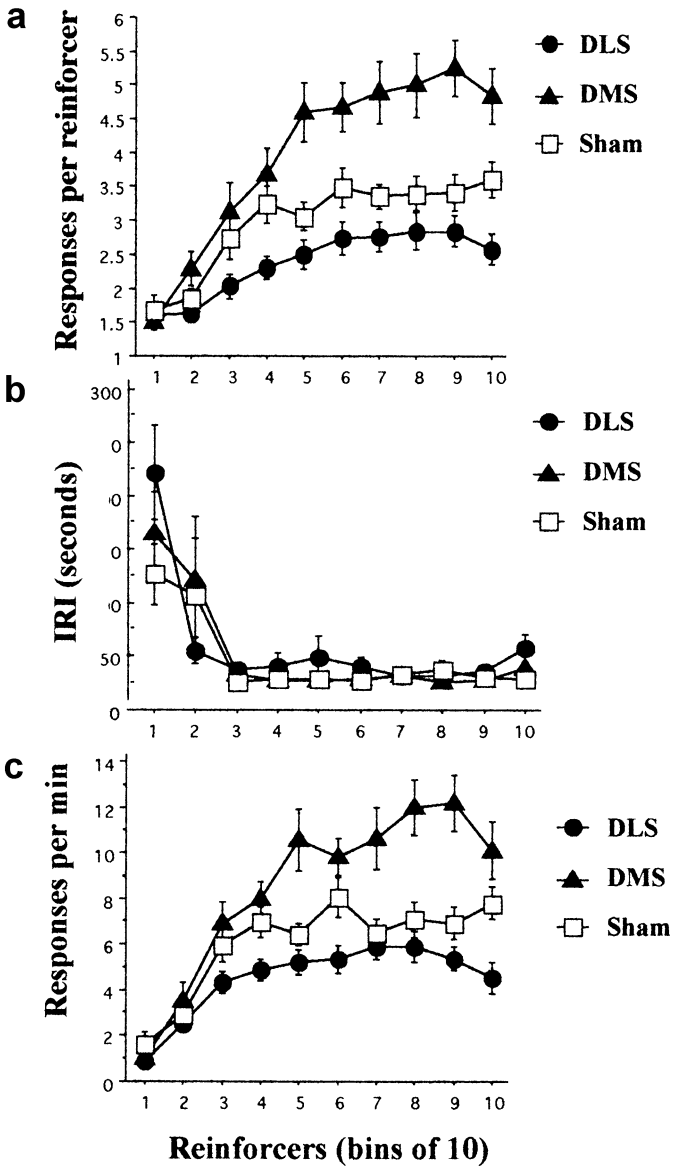

FIG. 5. Detailed analysis of early acquisition using a fixed interval 20-s (FI-20) schedule. (a) Responses per reinforcer on the *y*-axis, and the number of reinforcers on the *x*-axis (ten bins with ten reinforcers each). (b) Inter-reinforcer interval (IRI, in seconds). (c) Response rates (per min) during this period. DLS, dorsolateral striatal lesioned group; DMS, dorsomedial striatal lesioned group; Sham, sham operated controls. Error bars show one SEM.

1996). Thus, if anything, these data predict that DLS lesioned rats should have enhanced rather than impaired contextual discrimination and consequently should have been even less likely to show outcome devaluation than the sham group. In light of these considerations it is quite unlikely that the results of the current experiments reflect differential context control in the sham and DLS lesioned groups.

Dorsolateral striatum and habit learning

Although it is often argued that the dorsal striatum is involved in the acquisition of S–R habits, support for this claim to date has largely come from experiments in which the control of responding by an S–R process was inferred rather than directly demonstrated. Nor has it always been clear in previous work whether the dorsal striatum functions homogeneously as a whole, or if the DLS specifically is involved in habit learning (Packard & Knowlton, 2002).

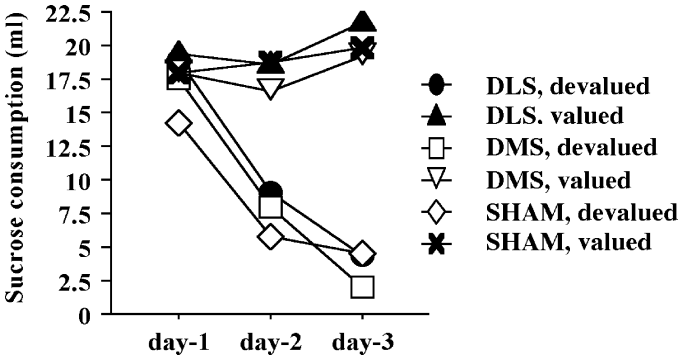

FIG. 6. Sucrose consumption over 3 days of taste aversion training for rats receiving LiCl injections (devalued) and saline injections (valued). DLS, dorsolateral striatal lesioned group; DMS, dorsomedial strital lesions; Sham, - sham operated control group.

The DLS receives massive excitatory projections from the sensorimotor cortices, and is able to influence various motor systems via its projections to the brainstem and to the motor thalamocortical pathways (Graybiel, 1998). As with the cortical areas from where these projections arise, this part of the striatum also appears to be somatotopically

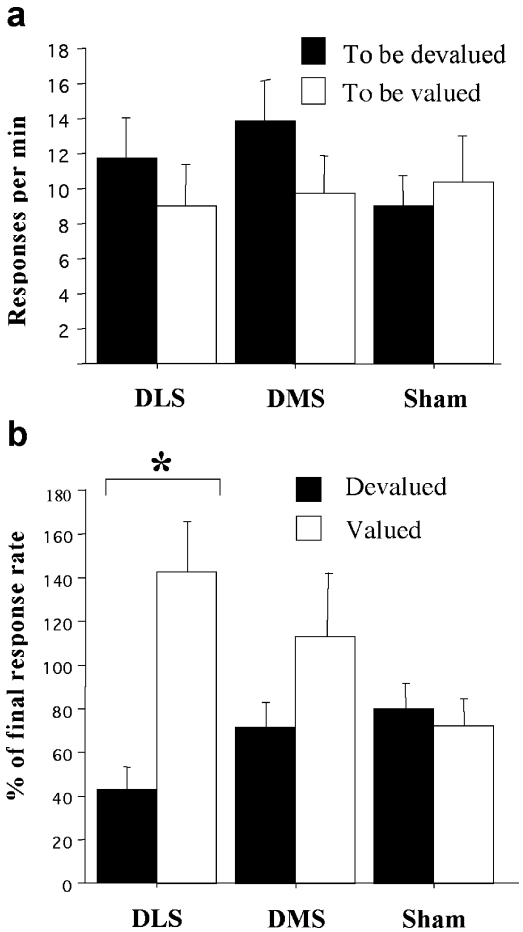

FIG. 7. Performance during a 5-min extinction test with no reward delivery. (a) Final response rates on the last day of training (RI-60). (b) Extinction performance as a percentage of the final response rates during the last day of training. DLS, dorsolateral striatal lesioned group; DMS, dorsomedial striatal lesioned group; Sham, sham operated controls. Error bars show one SEM; $^{*}P < 0.05$.

organized (Graybiel, 1990; West *et al*., 1990). The behavioural evidence generally accords with the anatomy. For example, maze studies using win–stay or response learning tasks have generally found significant impairments following lesions of the DLS (McDonald & White, 1993; Packard & McGaugh, 1996). However, in none of these maze studies has there been any attempt to establish whether performance was habitual by independent manipulation of the outcome.

By contrast, by using outcome devaluation as a behavioural assay to assess the relative control of instrumental performance by goal-directed and habitual processes, the current study has provided direct evidence for the involvement of the DLS in habit learning. The results of this study are therefore both strongly consistent with and significantly bolster the support for previous theoretical claims.

Neural substrates of actions and habits

In a detailed analysis of the early acquisition data (Fig. 5), rats with DMS lesions responded at a significantly higher rate relative to controls, whereas rats with DLS lesions responded less than controls. It is important to note, however, that this difference was only found early in training, i.e. under the FI-20 schedule, and was not found after additional training under variable interval schedules (see Figs 5a and 7a). Although previous reports have shown various dissociations between the medial and lateral regions of the dorsal striatum, the precise role of the DMS is still unknown. Some authors have claimed that the DMS may be involved in declarative types of learning such as place learning, via its connections with the medial temporal lobe (Devan & White, 1999; Devan *et al*., 1999). Our results suggest that it may suppress activity in the habit system, although further work is clearly needed to elucidate the functional role of the DMS, particularly its involvement in goal-directed actions.

There is now considerable evidence for the existence of a distributed neural network that supports the performance of non-habitual, goal-directed actions (Balleine & Dickinson, 1998, 2000; Corbit *et al*., 2001, 2003; Balleine *et al*., 2003). One property that components of this network have in common is goal expectancy, and recent neurophysiological work has found evidence of anticipatory, goal-related activity in neurons of various regions in the frontal lobe, in the caudate nucleus, which roughly corresponds to DMS in the rat, and in the nucleus accumbens (Apicella *et al*., 1991; Leon & Shadlen, 1999; Hikosaka & Watanabe, 2000; Hollerman *et al*., 2000; de Borchgrave *et al*., 2002; Lauwereyns *et al*., 2002a,b). Recent evidence, however, suggests that the accumbens may not be involved, strictly speaking, in instrumental incentive processes but rather is critical for the modulation of instrumental performance by Pavlovian cues (Balleine & Killcross, 1994; Parkinson *et al*., 2000; Corbit *et al*., 2001; Hall *et al*., 2001; de Borchgrave *et al*., 2002). By contrast, recordings from the DLS have failed to find neural activity 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), consistent with the view that the DLS mediates responding driven by antecedent stimuli.

According to contemporary theories of instrumental learning, normal acquisition and performance of instrumental actions requires the interaction of both S–R habit and a goal-directed, R–O, processes (Dickinson & Balleine, 1994; Dickinson *et al*., 1995). Although lever pressing in rats with DLS lesions was mildly impaired in the present study, they were clearly capable of acquiring and maintaining responding, even though independent evidence from devaluation suggests that their responding was not mediated by S–R learning. This finding therefore implies that S–R and R–O processes can operate independently to control instrumental performance, and that R–O processes can exert control over behaviour that is usually mediated by S–R learning. Such a conclusion is also supported by a recent study of the cortical substrates underlying instrumental conditioning, which found that, after infralimbic cortex lesions, sensitivity to devaluation was enhanced (Killcross & Coutureau, 2003). However, as the training procedures used in that particular study are quite different from ours (for instance, they used a within-subjects design with a mixture of ratio and interval schedules), it is difficult to compare these findings directly. Nevertheless, although the relationship between the infralimbic cortex and the DLS remains to be elucidated by further work, the study by Killcross & Coutureau (2003) supports this apparent distinction between a goal-directed action system sensitive to devaluation and the habit system insensitive to devaluation, and also suggests that rats can switch to the use of the action system when the S–R habit system is disrupted. As the switch between habitual and more flexible modes of behaviour is likely to be tightly regulated, the study of the neural mechanisms underlying this shift in control of performance will be central to our understanding of voluntary behaviour generally, and of certain disorders, particularly drug addiction, that appear to impair the regulation of these processes (Everitt & Wolf, 2002).

Acknowledgements

We would like to thank Aaron Blaisdell, Michael Fanselow, Joaquin Fuster and Frank Krasne for comments on earlier drafts of this paper, and Denize Yanez for technical assistance. This research was supported by an NSF graduate fellowship to H.H.Y., NSF grant 9985417 to B.J.K. and NIMH grant MH 56446 to B.W.B.

Abbreviations

DLS, dorsolateral striatum; DMS, dorsomedial striatum; R–O, action–outcome; S–R, stimulus–response.

References

Adams, C.D. (1981) Variations in the sensitivity of instrumental responding to reinforcer devalaution. *Q. J. Exp. Psychol.*, **34B**, 77–98.

Adams, C.D. & Dickinson, A. (1981) Instrumental responding following reinforcer devaluation. *Q. J. Exp. Psychol.*, **33B**, 109–121.

Apicella, P., Ljungberg, T., Scarnati, E. & Schultz, W. (1991) Responses to reward in monkey dorsal and ventral striatum. *Exp. Brain Res.*, **85**, 491–500.

Balleine, B.W. (2001) Incentive processes in instrumental conditioning. In Klein, S.B. (Ed.), *Handbook of Contemporary Learning Theories*. Lawrence Erlbaum Associates, Inc., Publishers, Mahwah, 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.W. & Dickinson, A. (2000) The effect of lesions of the insular cortex on instrumental conditioning: evidence for a role in incentive memory. *J. Neurosci.*, **20**, 8954–8964.

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.

Baum, W.M. (1973) The Correlation-Based Law of Effect. *J. Exp. Anal. Behav.*, **20**, 137–153.

Bonardi, C., Honey, R.C. & Hall, G. (1990) Context specificity of conditioning in flavor-aversion learing: extinction and blocking tests. *Anim. Learning Behav.*, **18**, 229–237.

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.

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.

Colwill, R.M. & Rescorla, R.A. (1990) Effect of reinforcer devaluation on discriminative control of instrumental behavior. *J. Exp. Psychol. Anim. Behav. Process*, **16**, 40–47.

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.

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.

Devan, B.D. & White, N.M. (1999) Parallel information processing in the dorsal striatum: relation to hippocampal function. *J. Neurosci.*, **19**, 2789–2798.

Dickinson, A. (1980) *Contemporary Animal Learning Theory.* Cambridge University Press, Cambridge.

Dickinson, A. (1985) Actions and habits: the development of behavioural autonomy. *Phil. Trans. R. Soc. Lond.*, **B308**, 67–78.

Dickinson, A. (1989) Expectancy theory in animal conditioning. In Klein, S.B. & Mowrer, R.R. (Eds), *Contemporary Learning Theories: Pavlovian Conditioning and the Status of Traditional Learning Theories*. Lawrence Erlbaum Associates, Hillsdale, NJ: pp. 279–308

Dickinson, A. (1994) Instrumental conditioning. In Mackintosh, N.J. (Ed.), *Animal Cognition and Learning.* Academic Press, London, pp. 4–79.

Dickinson, A. & Balleine, B. (1994) Motivational control of goal-directed action. *Anim. Learning 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. 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. Psychol.*, **35B**, 35–51.

Everitt, B.J. & Wolf, M.E. (2002) Psychomotor stimulant addiction: a neural systems perspective. *J. Neurosci.*, **22**, 3312–3320.

Graybiel, A.M. (1990) Neurotransmitters and neuromodulators in the basal ganglia. *Trends Neurosci.*, **13**, 244–254.

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

Hall, J., Parkinson, J.A., Connor, T.M., Dickinson, A. & Everitt, B.J. (2001) Involvement of the central nucleus of the amygdala and nucleus accumbens core in mediating Pavlovian influences on instrumental behaviour. *Eur. J. Neurosci.*, **13**, 1984–1992.

Hikosaka, K. & Watanabe, M. (2000) Delay activity of orbital and lateral prefrontal neurons of the monkey varying with different rewards. *Cereb. Cortex*, **10**, 263–271.

Hollerman, J.R., Tremblay, L. & Schultz, W. (2000) Involvement of basal ganglia and orbitofrontal cortex in goal-directed behavior. *Prog. Brain Res.*, **126**, 193–215.

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.

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.

Knowlton, B.J., Mangels, J.A. & Squire, L.R. (1996) A neostriatal habit learning system in humans [see comments]. *Science*, **273**, 1399–1402.

Lauwereyns, J., Takikawa, Y., Kawagoe, R., Kobayashi, S., Koizumi, M., Coe, B., Sakagami, M. & Hikosaka, O. (2002a) Feature-based anticipation of cues that predict reward in monkey caudate nucleus. *Neuron*, **33**, 463–473.

Lauwereyns, J., Watanabe, K., Coe, B. & Hikosaka, O. (2002b) A neural correlate of response bias in monkey caudate nucleus. *Nature*, **418**, 413–417.

Leon, M.I. & Shadlen, M.N. (1999) Effect of expected reward magnitude on the response of neurons in the dorsolateral prefrontal cortex of the macaque. *Neuron*, **24**, 415–425.

Lopez, M., Balleine, B. & Dickinson, A. (1992) Incentive learning following reinforcer devaluation is not conditional upon the motivational state during re-exposure. *Q. J. Exp. Psychol. B*, **45**, 265–284.

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

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 Stereotaxic Coordinates*, Academic Press, San Diego.

West, M.O., Carelli, R.M., Pomerantz, M., Cohen, S.M., Gardner, J.P., Chapin, J.K. & Woodward, D.J. (1990) A region in the dorsolateral striatum of the rat exhibiting single-unit correlations with specific locomotor limb movements. *J. Neurophysiol.*, **64**, 1233–1246.

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.