Published in final edited form as:

Eur J Neurosci. 2010 April ; 31(7): 1312–1321. doi:10.1111/j.1460-9568.2010.07153.x.

Posterior dorsomedial striatum is critical for both selective instrumental and Pavlovian reward learning

Laura H. Corbit¹ and Patricia H. Janak^{1,2}

¹ Ernest Gallo Clinic and Research Center, University of California at San Francisco, 5858 Horton Street Suite 200, Emeryville, CA 94698, USA

² Wheeler Center for the Neurobiology of Addiction and Department of Neurology, University of California at San Francisco, San Francisco, CA, USA

Abstract

The dorsal striatum (DS) has been implicated in instrumental learning but its role in the acquisition of stimulus-driven behaviour is not clear. To explore the contribution of the DS to both response-outcome (R-O) and stimulus-outcome (S-O) associative learning, we pharmacologically inactivated subregions (dorsolateral, anterior dorsomedial and posterior dorsomedial) of the DS during acquisition sessions in which subjects acquired two unique, novel R-O pairs or two unique, novel S-O pairs. To test whether specific R-O or S-O associations were learned under inactivation, rats were tested following selective-satiety devaluation of one outcome under drug-free conditions. In the instrumental task, control rats and rats with dorsolateral striatum (DLS) inactivation during learning responded less on the lever that had earned the devalued outcome than on the alternative lever at test, indicating that the DLS is not critical for the formation of R-O associations. In contrast, rats with inactivation of the medial DS (DMS) (either anterior or posterior) during learning responded indiscriminately, suggesting failure to acquire the novel R-O associations. In the Pavlovian task, both controls and rats with anterior DMS inactivation during learning responded less in the presence of the stimulus predicting the devalued outcome, whereas rats with DLS or posterior DMS inactivation during learning responded equally to the stimuli, indicating that they had not acquired the novel S-O associations. These data confirm that the DLS and anterior region DMS mediate different aspects of reward-related learning, and suggest that the posterior DMS may mediate a function common to both forms of learning (R-O and S-O). Finally, we demonstrate that both S-O and R-O associations are required for selective Pavlovian-instrumental transfer.

Keywords

basal ganglia; caudate putamen; devaluation; instrumental conditioning; Pavlovian-instrumental transfer; rat

Introduction

The dorsolateral striatum (DLS) has been strongly implicated in habit learning and the formation of stimulus-response (S-R) associations presumed to underlie this form of learning (for review, see Yin & Knowlton, 2006). However, previous findings also implicate the dorsal striatum (DS) in stimulus-reward learning (Aosaki *et al.*, 1994) and stimulus-guided performance in a variety of tasks (Hassani *et al.*, 2001; Cromwell & Schultz, 2003;

Corbit & Janak, 2007b) that are unlikely to be controlled solely by S-R associations, suggesting that the DS may play a more broad role in mediating the influence of predictive stimuli on performance than is captured by the S-R framework.

We have recently reported that reversible inactivation of the DLS substantially reduces the influence of conditioned stimuli on the performance of independently-acquired instrumental actions using the Pavlovian-instrumental transfer (PIT) paradigm (Corbit & Janak, 2007b). In contrast, inactivation of the medial DS (DMS) eliminated the selectivity of the PIT while leaving the general excitatory impact of the stimuli intact. The PIT task has three phases: rats receive Pavlovian training wherein discrete stimuli are paired with reward delivery, instrumental training where the performance of different responses (such as lever presses) earns unique rewards and, finally, the impact of the stimuli on responding is assessed. Presentation of reward-predictive stimuli typically elevates the performance of responses associated with the same reward (e.g. Kruse *et al.*, 1983). Importantly, the influence of stimuli on the performance of the independently-acquired response is first assessed at the time of testing and, as such, there is no opportunity for the prior establishment of classic S-R associations between the discrete Pavlovian stimulus and the instrumental lever-press response (although retrieved stimulus properties of the outcome acting through an outcome-response (O-R) mechanism are likely to be involved) (Balleine & Ostlund, 2007; de Wit *et al.*, 2009; Balleine & O'Doherty, 2010).

Although the demonstration that inactivation of the DLS greatly attenuates the impact of stimuli on responding again implicates the DLS in the control of stimuli influences on responding (Corbit & Janak, 2007b), the specific role of the DLS remains ambiguous. For selective transfer effects to emerge, subjects must form and utilize specific response-outcome (R-O) and stimulus-outcome (S-O) associations as well as evaluate and integrate congruent information carried in each. Deficits in any of these processes would result in deficits in PIT. Hence, although both the DLS and DMS are necessary for selective transfer, the specific function of each structure is difficult to ascertain from the deficits in PIT alone.

The aim of the current study was to characterize the role of the DLS and DMS in the acquisition of the associative components required for the ultimate demonstration of selective PIT effects. Although we found a deficit following inactivation of a relatively anterior DMS (aDMS), previous work has shown that the posterior DMS (pDMS) is importantly involved in R-O learning (Yin *et al.*, 2005). Further, the role of these subregions in specific S-O learning has not been tested. Therefore, we addressed the role of the DLS, aDMS and pDMS in the acquisition of specific R-O and S-O associations using reversible inactivation techniques. Based on previous findings (Corbit & Janak, 2007b), we predicted that the DMS may be required for selective R-O learning, whereas the DLS may be critical for stimulus-related learning.

Materials and methods

Subjects and apparatus

Forty-six naive male Long-Evans rats (Harlan, Indianapolis, IN, USA) weighing approximately 350 g were singly housed and had free access to water in the home cage. Feeding was restricted to maintain weights at approximately 90% of free-feeding weights. All procedures were approved by the Institutional Animal Care and Use Committee of the Ernest Gallo Clinic and Research Center following guidelines instituted by the National Institutes of Health. Training and testing took place in operant chambers (Med Associates, East Fairfield, VT, USA) as described previously (Corbit & Janak, 2007a).

Baseline Instrumental Task training

Rats were initially trained to respond on two levers to earn a 20% polycose solution. In an attempt to establish similar responding on the two levers, one lever was inserted at a time for 3 min and was then retracted and the other lever inserted for 3 min. This sequence was repeated five times to comprise the 30-min session. The identity of the first lever was alternated across days. For the first two sessions, responding was reinforced on a continuous reinforcement schedule and thereafter rats received 2 days each of random-ratio (RR)2, RR3 and RR4 schedules of reinforcement prior to surgery.

Specific response-outcome training under inactivation

Following recovery from surgery, stable responding for polycose was re-established (two sessions; RR4) prior to the inactivation sessions. At 10 min prior to the session for the specific R-O training, rats received an infusion of baclofen/muscimol to inactivate either the DLS, aDMS or pDMS, or an equal volume of saline. During these sessions, responding on one lever resulted in the delivery of 0.1 mL of a 20% sucrose solution and responding on the other lever resulted in the delivery of a single 45-mg food pellet (grain-based dustless precision pellets; no. F0165, BioServ, Frenchtown, NJ, USA). As in the initial training, access to the two levers was alternated in 3-min intervals and responding was rewarded according to independent RR4 schedules. There were three of these training sessions.

Test for response-outcome learning using devaluation

Rats had free access to either pellets or sucrose for 1 h in the home cage prior to a two-lever choice test conducted in extinction. The test session was 15 min long. There were no infusions prior to the test sessions. The animals received 1 day of retraining on the selective R-O associations under inactivation prior to a second test. This test was identical to the first except that the rats were pre-fed the other outcome.

Baseline Pavlovian Task training

Two auditory stimuli (white noise and 5-Hz clicker) served as conditioned stimuli and were adjusted to 80 dB in the presence of background noise of 60 dB provided by a ventilation fan. Both stimuli were initially paired with delivery of a 20% polycose solution (with 0.9% NaCl; w/v). Six 2-min presentations of each stimulus were given in each session in random order interspersed with inter-trial intervals averaging 5 min in duration. During each stimulus, 0.1 mL of the outcome (polycose) was delivered on a random time 30-s schedule to separate recessed magazines (e.g. white noise-left and 5-Hz clicker-right). Rats received six sessions each of 75 min in length. The number of magazine entries during each stimulus and in a 2-min pre-stimulus interval was measured.

Specific stimulus-outcome training under inactivation

At 10 min prior to the sessions for the specific S-O training, rats received an infusion of baclofen/muscimol to inactivate either the DLS, aDMS or pDMS, or an equal volume of saline (control group). During these sessions, one stimulus predicted the delivery of 0.1 mL of a 20% sucrose solution and the other stimulus predicted the delivery of a single 45-mg food pellet (BioServ). As in training, the stimuli were 2 min in duration and rewards were delivered according to a random time 30-s schedule. The two rewards were delivered to separate but adjacent recessed magazines. These sessions had eight trials and lasted 50 min. The sessions were shortened to limit changes in the efficacy or potential spread of the baclofen/muscimol. There were three of these training sessions.

Test for stimulus-outcome learning using devaluation

Rats had free access to either pellets or sucrose for 1 h in the home cage prior to a stimulus test conducted in extinction. The test was 30 min long and each stimulus was presented twice and entries into the two magazines were recorded. There were no infusions prior to the test sessions. The rats received one retraining session with the specific S-O pairs under inactivation prior to a second test. This test was identical to the first except that rats were pre-fed the other outcome.

Pavlovian-instrumental transfer test

To allow a full return to a state of hunger following the final devaluation test, a minimum of 2 days was allowed prior to two PIT tests conducted in extinction. During each test, one lever was available and each stimulus was presented twice interspersed with intervals of no stimulus ($\emptyset$). The 22-min test contained eight 2-min bins (two white noise trials and two clicker trials alternated with four $\emptyset$ trials in the following order: white noise, 5-Hz clicker, 5-Hz clicker and white noise). Each stimulus bin was separated from the subsequent baseline ($\emptyset$) bin by 1 min.

Consumption test of specific satiety

To assess the ability of the pDMS rats to distinguish between the two outcomes, they were given a consumption version of the specific-satiety devaluation test. One of the two outcomes was devalued in the same manner as in the tests described above but now the dependent measure used was consumption of either that same outcome or a different outcome. Animals received two tests. In each test, following an infusion of baclofen/muscimol, they received 45 min of *ad-libitum* access to one of the foods, either pellets or sucrose, in the home cage. The pre-fed food was removed and then the animals were presented with the test food, either the sucrose or pellets, for an additional 15 min and their consumption in that test period was recorded. In the first test, half of the animals received the same food as that which they had just had access to and, for the remaining animals, the different food was presented. In the second test, the animals were pre-fed the same food as in the first test; however, the food item presented in the test period was the opposite to that of the first test. As a result, animals that had received the same food during the pre-feeding and testing in the first test now received the different food during the second consumption test and vice versa.

Surgery

Stereotaxic surgery was conducted under isoflurane anaesthesia (2–5% in oxygen for induction; 1–2% for maintenance) to implant 26-gauge guide cannulas (Plastics One, Roanoke, VA, USA) targeted at either the DLS (AP, +1.2 mm; ML, $\pm$ 3.4 mm; DV, –1.0 mm, coordinates relative to bregma, and dura for DV), aDMS (AP, +1.2 mm; ML, $\pm$ 1.5 mm; DV, –1.4 mm) or pDMS (AP, –0.3 mm; ML, $\pm$ 2.6 mm; DV, –1.2 mm). Final group sizes were as follows: DLS: control, N = 6, inactivation, N = 9; aDMS: control, N = 6, inactivation, N = 10; and pDMS: control, N = 5, inactivation, N = 9. The tips of the guide cannulas were positioned 3 mm dorsal to the intended infusion site and anchored with machine screws and dental acrylic. Rats were treated with buprenorphine (0.05 mg/kg, s.c.) prior to surgery and had children's Tylenol (cherry flavour; 20 mg/100 mL; McNeil-PPC Inc., Fort Washington, PA, USA) available in their drinking water for 1 week following surgery to minimize pain.

Infusions

The rats were gradually habituated to dummy removal and the handling required for microinfusion procedures during the recovery and retraining period. They were given mock

infusions prior to the final baseline training session wherein the dummy cannulas were removed, short injectors (1 mm beyond the guide rather than the 3-mm projection used for the actual infusions) were inserted into the guides and the rats were held in the lap of the experimenter for 3 min to simulate the infusion procedures that would commence the following day. Prior to the novel R-O and S-O training sessions, rats in the DLS, aDMS and pDMS inactivation groups received infusions of a combination of the GABA_B receptor agonist, baclofen, and the GABA_A receptor agonist, muscimol (1.0/0.1 mM, Sigma, St Louis, MO, USA), and control rats for each placement received saline vehicle via an infusion cannula (33 gauge; Plastics One; 0.3 μ L/min; total volume of 0.3 μ L delivered per hemisphere). Infusions took place 10 min prior to the session and were conducted over 1 min, and the cannulas were left in place for an additional 2 min to allow for diffusion.

Histology

The animals were overdosed on sodium pentobarbital and perfused transcardially. Coronal sections (50 μ m) of formalin-fixed tissue were sliced, mounted and stained with thionin to allow verification of infusion placement and assessment of any extraneous damage.

Data analysis

Data were analysed using repeated-measures ANOVA and simple effects and Tukey *post-hoc* analyses were used to further assess significant main effects and interactions. Preliminary ANOVAS indicated no effect of either stimulus (click vs. noise) or lever (left vs. right) for both Pavlovian and instrumental training (F -values < 1); therefore the data were collapsed across those factors.

Results

Histology

One subject was excluded from the DLS group; all others were included in the behavioural analyses. Figure 1 displays the placement of the cannula tips within the DLS, aDMS or pDMS for these subjects.

Instrumental task

Baseline training—To ensure that rats would perform the task and thus have exposure to the novel R-O associations following inactivation, we first trained rats to press two separate levers for a common reward (polycose). The data for the final day of pre-training are shown in Fig. 2A. Following training, rats were implanted with cannulas and assigned to the control or treatment groups. Analysis of the performance of the three control groups indicated no differences and so these animals were combined to form a single control group. There were no differences between groups in pre-training ($F_{3,41} = 1.1$, $P > 0.05$).

Specific response-outcome training under inactivation—Prior to the following three sessions, rats received a bilateral 0.3- μ L infusion of a cocktail of the GABA_B agonist, baclofen (1.0 mM), and the GABA_A agonist, muscimol (0.1 mM). During these sessions, responding on one lever delivered pellets, whereas responding on the other lever delivered a sucrose solution. To examine whether rats in each group responded similarly following the infusions and thus had similar opportunity to learn the novel R-O associations, we examined the average number of lever presses during the inactivation sessions. Although there was an overall effect of group ($F_{3,41} = 2.89$, $P < 0.05$; Fig. 2B), *post-hoc* analyses indicated that none of the experimental groups differed from controls although the aDMS group did respond significantly more than the pDMS group (Tukey test). Rats responded more on the lever that earned pellets than the one that delivered sucrose ($F_{1,41} = 32.4$, $P < 0.05$);

however, this effect did not interact with group ($F_{3,41} = 1.5$, $P > 0.05$). Given this response bias, we were careful in subsequent tests to balance the order of testing (e.g. pellets devalued vs. sucrose devalued) within each group and both outcomes were tested in both the devalued and non-devalued condition. In order to rule out the possibility that the rats decreased responding across the multiple infusions, we examined performance across the infusion days and found that rats increased their responding across days ($F_{2,6} = 28.5$, $P < 0.01$). There was no interaction with group, indicating that each group increased responding across days ($F_{6,82} = 1.1$, $P > 0.05$).

Utilization of response-outcome associations: instrumental devaluation—The data from the inactivation sessions suggest that, regardless of group, all rats were exposed to and thus had the opportunity to learn the novel R-O associations. Nonetheless, performance may reflect learning occurring prior to the inactivation session, i.e. that responding leads to reward in general without specific reference to the novel outcomes introduced in the inactivation phase. To probe the content of the learning that occurred under inactivation, we altered the value of one of the new outcomes using specific-satiety devaluation and then allowed rats the opportunity to respond on the two levers in the absence of the outcomes. We predicted that, for the control group, devaluation of one outcome should selectively reduce performance of the response that previously earned the pre-fed outcome. To the extent that specific regions of the DS are critical for the acquisition of novel R-O associations, inactivation of these regions during outcome-specific training should render rats unable to selectively adjust responding during the test session following devaluation. The statistical analysis revealed no overall effect of group ($F_{3,41} = 1.95$, $P > 0.05$) but a significant effect of devaluation ($F_{1,41} = 16.35$, $P < 0.01$) and an interaction between these factors ($F_{3,41} = 3.56$, $P < 0.05$). As seen in Fig. 2C, control-treated rats responded less on the lever that had previously earned the currently-devalued outcome compared with the other lever ($F_{1,6} = 11.45$, $P < 0.01$). A similar pattern emerged for rats with inactivation of the DLS ($F_{1,8} = 9.33$, $P < 0.05$). In contrast, rats in both the aDMS and pDMS groups failed to selectively reduce responding on one lever compared with the other (aDMS: $F_{1,9} = 1.1$, $P > 0.05$; pDMS: $F_{1,8} = 0.65$, $P > 0.05$). These groups did show a relatively low level of responding on both levers, which is probably due to the method of devaluation. This is consistent with previous results (Yin *et al.*, 2005). Because we used selective satiety, rats that are usually trained and tested while hungry, as a result of the pre-feeding procedure, are considerably less hungry than in a typical training session. The decrease in hunger and resultant motivation could produce a decrease in responding overall. For animals that have selective associations related to the specific pre-fed outcome, the impact of this treatment is more selective and the animals preserve responding for the non-devalued outcome. In animals without these selective associations, the satiety treatment may act to decrease responding in a non-selective fashion. All groups consumed a similar amount of the freely available outcome during the pre-feeding ($F_{3,41} = 1.1$, $P > 0.05$) and the fact that this level of consumption was sufficient to produce a devaluation effect in the control group suggests that the satiety treatment itself was effective. The efficacy of the satiety treatment is further evaluated by a consumption test (data presented below). Notably, analysis of the magazine entry data from the test session indicated no significant group differences ($F_{3,41} = 1.76$, $P > 0.05$) and so the alterations in responding during the devaluation test cannot be explained by competition from the magazine entry response.

Pavlovian task

Baseline training—The same animals were subsequently trained in a Pavlovian task wherein presentation of one stimulus (e.g. clicker) indicated that reward would be delivered to a food magazine situated on the left side of one chamber wall, whereas presentation of a different stimulus (e.g. noise) indicated that reward would be delivered to a magazine

situated on the right side of that same chamber wall. As seen in Fig. 3A, all groups performed similarly during pre-training with polycose reward ($F_{3,41} = 0.54, P > 0.05$). Rats made more food magazine entries during stimulus presentations than during a pre-stimulus period of equal length ($F_{1,41} = 282.0, P < 0.01$) and made more entries to the magazine that was rewarded on a given trial (identified by stimulus identity) than to the alternate magazine ($F_{1,41} = 111.9, P < 0.01$), and did so predominantly during the stimulus presentations (stimulus $\times$ magazine interaction, $F_{1,41} = 173.3, P < 0.01$). Neither of these factors interacted with group (F -values < 1).

Specific stimulus-outcome training under inactivation

Initial analyses indicated that rats made a similar number of magazine entries during the three inactivation sessions ($F_{2,6} = 2.7, P > 0.05$) and therefore the data were averaged across infusion day for subsequent analyses. During the inactivation sessions all rats made more entries into the magazine during stimulus presentations than during baseline intervals and were selective in this behaviour, making more correct magazine entries, based on which stimulus was present (Fig. 3B) [main effect of stimulus (baseline vs. stimulus): $F_{1,41} = 294.6, P < 0.01$, magazine (rewarded vs. alternate): $F_{1,41} = 250.9, P < 0.01$ and stimulus $\times$ magazine interaction: $F_{1,41} = 248.2, P < 0.01$]. There was a main effect of group ($F_{3,41} = 3.0, P < 0.05$) but no interactions with this factor. *Post-hoc* analyses (Tukey test) indicated no group differences in baseline responding; however, during stimulus presentations for the rewarded magazine, rats in the aDMS group responded more than the pDMS group and more than all other groups for entries to the alternate magazine. Responding during the rewarded stimulus by the aDMS or pDMS groups did not significantly differ from that of the saline control group. Notably, all subjects consumed all rewards in all sessions.

Utilization of stimulus-outcome associations: Pavlovian devaluation

The acquisition of the specific S-O associations was probed using outcome devaluation. Although all subjects responded during inactivation training sessions, and hence had the opportunity to learn the S-O associations, data from the devaluation test sessions suggest that some groups failed to acquire the novel associations (Fig. 3C). After one of the outcomes was devalued, control rats made fewer entries during presentation of the stimulus that predicted delivery of the devalued outcome than they did during presentations of the stimulus predicting the non-devalued outcome. Rats in the aDMS group responded less overall compared with the control group; nonetheless, they still showed less responding during the stimulus paired with the devalued outcome. In contrast, rats in the DLS and pDMS groups, in addition to demonstrating low levels of responding, failed to show any selective stimulus control over this behaviour. Statistical analyses confirm an effect of both devaluation ($F_{1,41} = 26.76, P < 0.01$) and group ($F_{3,41} = 5.57, P < 0.01$) as well as an interaction between these factors ($F_{3,41} = 8.55, P < 0.01$). Simple effects analyses demonstrate that both control ($F_{1,16} = 38.5, P < 0.01$) and aDMS ($F_{1,9} = 16.5, P < 0.01$) groups responded less in the presence of the stimulus paired with the devalued outcome, whereas the DLS ($F_{1,8} = 0.38, P > 0.05$) and pDMS ($F_{1,8} = 2.84, P > 0.05$) groups did not show selective responding during presentation of the two stimuli.

Pavlovian-instrumental transfer test

Next we tested in the same subjects whether selective R-O learning or selective S-O learning (or both) is necessary for the expression of selective PIT; selectivity is demonstrated when stimuli previously paired with a specific outcome enhance lever-press responding selectively on the lever that was previously reinforced by delivery of that same outcome. If selective R-O learning is required, then the subjects within the aDMS and pDMS inactivation groups that did not acquire selective R-O learning should fail to show selective PIT. If selective S-O learning is required, then the subjects within the DLS and pDMS inactivation groups that did

not acquire selective SO learning should fail to show selective PIT. If both selective R-O and S-O learning are required, all three groups should have a deficit. For this test, the effect of presentation of the two stimuli originally presented during the specific S-O training on lever-press responses on each of the two levers was measured; no outcomes were delivered during these tests. As seen in Fig. 4, upon stimulus presentation, control rats increased their response rates compared with periods when no stimuli were present. The magnitude of this increase was selective, in that it was greater when the stimulus predicted the same outcome (e.g. food pellets) as the available lever. In contrast, although all of the experimental groups showed some increase in responding when the stimuli were presented, in no case was it selective (Fig. 4). The general excitatory effect of the stimuli (including the increase during the 'different' stimulus for control rats) is anticipated based on the pre-training phases wherein both stimuli (and responses) were paired with a common outcome. Statistical analyses confirm an effect of group ($F_{3,41} = 4.0$, $P < 0.05$), stimulus ($F_{2,82} = 36.2$, $P < 0.01$) and an interaction between these factors ($F_{6,82} = 2.3$, $P < 0.05$). Simple effects analyses indicated an effect of stimulus in each group (all P -values < 0.01), confirming that stimulus presentation elevated responding from baseline; however, paired t-tests performed for each group indicated that only rats in the control group showed a greater increase in responding when the stimulus predicted the same outcome as the lever [$t(16) = 3.7$, $P = 0.002$ significant with the Bonferroni correction; P -values for all other groups > 0.05]. This finding confirms that both selective R-O and S-O associations are needed for the demonstration of selective PIT. Again, analysis of the magazine entry data during PIT testing indicates no differences between groups ($F_{3,41} = 0.89$, $P > 0.05$) and so response competition does not readily explain the observed differences in lever-press performance. One possible concern is that one of the two outcomes was devalued relatively more recently prior to PIT testing than the other. However, we do not believe that this should systematically bias the result of these tests for the following reasons. First, a minimum of 2 days were allowed between the last devaluation test and the PIT test and so the animals should have returned to a hungry state. Further, both outcomes underwent devaluation (although one outcome was devalued relatively more recently than the other) and both levers were tested for PIT and, as such, even if one outcome remained somewhat more devalued, this outcome was tested as both the 'same' and 'different', as was the relatively non-devalued outcome, and so any residual devaluation effect should not systematically bias the test results. Finally, the control animals that were treated identically to the experimental groups with respect to the devaluation manipulations showed a selective PIT effect.

Consumption test of specific satiety

Both the aDMS and DLS groups showed a selective devaluation effect in one of the other tests, providing evidence that these rats could discriminate between the different outcomes, thereby pointing to more selective deficits in either instrumental or Pavlovian learning, respectively. The pDMS group was the only group to show impairments in both the instrumental and Pavlovian tasks. An explanation for these effects could be that the pDMS is involved in a function common to both tasks, such as processing unique properties (taste, smell, specific nutrient content, etc.) of the different outcomes, a function required for selective performance in either task. To test the effects of pDMS inactivation on the ability to distinguish between the outcomes, rats were given a consumption version of the specific-satiety devaluation test. Following inactivation of the pDMS, one of the two outcomes was devalued in the same manner as in the tests described above but now the dependent measure used was consumption of either that same outcome or a different outcome. Normal performance on this test (i.e. decreased consumption of the same food item relative to a different food) requires that rats distinguish between the two food items. For the purpose of analysis, the amount of each type of food consumed (either grams of pellets or millilitres of sucrose) was converted to the equivalent number of outcomes (i.e. every 45 mg of pellets is

equivalent to one pellet outcome, whereas every 0.1 mL of sucrose solution is equivalent to one sucrose outcome). The results indicate that, after pre-feeding of one of the food outcomes, the rats ate more of a different food than of the one that they had just eaten. Importantly, the magnitude of this effect was similar between the control [mean devalued, 26.7 (SE, 8.6); mean non-devalued, 82.8 (SE, 15.8)] and pDMS [mean devalued, 42.3 (SE, 8.3); mean non-devalued, 82.9 (SE, 15.1)] groups, suggesting that the pDMS rats could indeed distinguish between the two outcomes and showed a selective devaluation effect when measured by consumption. The statistical analysis supported this claim, revealing no effect of group ($F_{1,10} = 0.4$, $P > 0.05$), a significant effect of devaluation ($F_{1,10} = 6.2$, $P < 0.05$) and no interaction between these factors ($F_{1,10} < 1$).

Discussion

Here we show that the DMS (either its anterior or posterior region) is critical for the formation of specific R-O associations, whereas the DLS and pDMS are required for the formation of selective S-O associations. These findings indicate that the DLS and aDMS mediate different aspects of reward-related learning and suggest that the pDMS supports both forms of learning.

Response-outcome learning

When the DMS (either anterior or posterior) was inactivated while rats had the opportunity to learn novel R-O associations, the rats showed similar levels of lever-press performance and earned and consumed similar numbers of the novel rewards as control rats and so had exposure to, and thus opportunity to learn, the novel R-O associations. However, when the content of the learning was later assessed by devaluing one of the two outcomes and giving the rats the opportunity to respond on the two levers in extinction, rats with DMS inactivation during learning failed to respond selectively at test, indicating that they had not formed associations based on the specific or unique aspects of the R-O association or its components. In contrast, rats that had the DLS inactivated during the novel R-O training sessions, like control rats, showed a selective devaluation effect at the time of testing, indicating that this region is not critical for learning selective R-O associations. This is consistent with, and extends, previous findings by Yin *et al.* (2005). Interestingly, using permanent lesions of the aDMS and a longer training period, Yin *et al.* (2005) reported no deficit in R-O learning evaluated using similar devaluation procedures. Although differences in the training procedures may account for this discrepancy, the lesion technique is probably responsible for these differences, although this remains to be tested. For example, it is possible that, over the course of the relatively longer training period (8 vs. 3 days) of Yin *et al.* (2005), portions of the aDMS not affected by the lesion may have been able to compensate and support conditioning. It is also possible that our drug infusions affected a larger region than their excitotoxic lesions, which could account for why we found deficits that were not previously observed.

Stimulus-outcome learning

The impairment in R-O learning in the DMS groups could be explained by an inability to differentiate either the unique attributes of the two responses or the two outcomes. The data from our Pavlovian test argue against the latter, at least for the animals in the aDMS group. In this procedure, rats received local inactivation of a DS subregion prior to sessions wherein two stimuli predicted the delivery of two novel rewards. Again, all groups performed well during training and thus were exposed to the novel S-O contingencies. The content of the learning that occurred in these sessions was assessed by devaluing one of the two outcomes and examining the behavioural response to the two stimuli in extinction. During this test the control and aDMS groups made more entries to the magazine where the

currently non-devalued outcome had been delivered and did so during the appropriate stimulus, confirming that they learned the novel S-O associations and demonstrating that the aDMS is not required for animals to distinguish between alternate rewards or to update representations of their relative value. In contrast, rats in both the DLS and pDMS groups were impaired in the test sessions despite similar behaviour during training under inactivation.

Stimulus properties of the outcome: Outcome-Response learning

According to theories of goal-directed action, animals encode the relationship between a particular response and its consequences or outcome; action choice will depend not only on knowledge of this relationship but also on an evaluative process that examines the value of that outcome relative to alternatives (Adams & Dickinson, 1981; Dickinson & Balleine, 1994; Balleine & Dickinson, 1998). Nonetheless, although these associations may inform behavioural choice, both the outcome itself and stimuli that predict the outcome can influence response initiation and vigour. Contemporary theories of instrumental performance suggest that the outcome controls actions in two distinct ways: first, through the usual R-O association in which the response retrieves the outcome with which it has been associated as a goal and second, the stimulus properties of the outcome can select the response with which it has been associated (Balleine & Ostlund, 2007). On this account, anything that serves to retrieve the outcome, such as a previously paired stimulus, should consequently retrieve the response with which the outcome is associated (thus predicting selective transfer effects). Selection of that response, in turn, acts to retrieve the outcome, now as a goal. These two processes, although independent (Colwill & Rescorla, 1990; Corbit *et al.*, 2001; Corbit & Balleine, 2003a,b; Holland, 2004; Balleine & Ostlund, 2007), should normally work together to control response selection and initiation. Selective PIT may therefore involve integration of S-O and O-R associations and it is possible that our inactivation treatments during the R-O training sessions impaired O-R learning, which could subsequently impair selective PIT (Balleine & Ostlund, 2007). Although conditions can be arranged such that an outcome serves as a discrete cue for action selection (e.g. de Wit *et al.*, 2009), under free operant conditions the O-R association is probably formed incidentally as animals undergo R-O training. As such, it is difficult to isolate from the present experiments distinct roles of subregions of the DS in R-O vs. O-R learning. Although effects on the forward R-O learning can be determined from the devaluation tests, any effects specifically on O-R learning will be best understood in future studies using designs similar to those of de Wit *et al.* (2009).

Relation to Pavlovian-instrumental transfer

The above results suggest that the DLS is critical for the formation of specific S-O associations or for applying this information to direct selective responding, whereas the aDMS is involved in the formation of specific R-O associations. Further, and as expected, disruption of either form of learning interferes with the expression of selective PIT effects. This finding indicates that both selective R-O associations and selective S-O associations must be acquired by the aDMS and DLS, respectively, for selective transfer to occur and, further, that the expression of normal transfer thus depends upon both regions under normal conditions. In addition, inactivation of the pDMS during S-O and R-O learning, which resulted in an impairment of each, blocked selective PIT but in this case, because this was a within-subject design, we cannot define the deficit (R-O deficit, S-O deficit or general deficit in outcome representation) that led to the PIT impairment in this group. It is also possible, as noted above, that O-R learning was blocked by inactivation during the instrumental training sessions and the specific role of this association and the contribution of DS subregions to its control await further study. Although inactivation of the DLS, aDMS or pDMS during training blocked selective PIT, it is notable that a general excitatory effect of

the conditioned stimuli on lever-press responding in the PIT test was observed in these groups, as indicated by increases in responding on both levers during the presentation of both cues. This general excitatory effect could be the result of the baseline training sessions where responses and stimuli were paired with a common reward (polycose) (Corbit *et al.*, 2007). The association with reinforcement in general would be expected to elevate responding in response to both stimuli, whereas selective responding would require knowledge of the outcome-selective contingencies that were only encountered under inactivation. Taken with our previous findings that inactivation at the time of testing also disrupts PIT (Corbit & Janak, 2007b), our data suggest that the DMS and DLS contribute to both the formation and expression of selective R-O and S-O associations, respectively.

Outcome discrimination

The pDMS group was the only group to show impairments in both the instrumental and Pavlovian tasks. This may mean that the pDMS is independently recruited in both instrumental and Pavlovian learning processes or that the pDMS mediates a function that is common to both forms of learning. One possible explanation for the observed results could be that the pDMS is involved in representing the unique properties (taste, smell, specific nutrient content, etc.) of the different outcomes, a function that is required for selective performance in either of the tasks. To address this issue we conducted a consumption version of the devaluation task. The results from this test indicate that animals in the pDMS group were indeed able to selectively devalue one outcome, suggesting that they were able to differentiate the unique sensory features of the outcomes under inactivation. Hence, the results point to a deficit in forming associations (with either stimuli or responses) related to unique reinforcers or perhaps modifying performance based on associatively-activated representations of the outcomes rather than a deficit related to the basic discrimination between outcomes when they are encountered. Although instrumental and Pavlovian motivational influences can be dissociated (Colwill & Rescorla, 1990; Corbit *et al.*, 2001; Corbit & Balleine, 2003a; Holland, 2004; Balleine & Ostlund, 2007), both probably contribute to performance under many circumstances. As such there must be a cooperation or integration at some level and, given that the pDMS is critical for both functions, it is possible that this structure contributes to this coordination.

Differential connectivity of striatal subregions

The regional differences that we observed probably depend on the different cortical and subcortical connections of DS subregions (Alexander *et al.*, 1986; Parent & Hazrati, 1995; Voorn *et al.*, 2004; Haber *et al.*, 2006). Of particular interest are inputs from the basolateral complex of the amygdala (BLA). The BLA has been reported to project broadly throughout the striatum; however, the density of the BLA inputs appears to differ greatly depending on the subregion (Kelley *et al.*, 1982). Notably, the most anterior and lateral regions of the DS (largely including our DLS placement) do not receive robust input from the BLA although it has been reported to be functionally connected (via the substantia nigra) with the central nucleus of the amygdala (Hatfield *et al.*, 1996; El-Amamy & Holland, 2007). It is of interest that more posterior regions, including the region of pDMS targeted by us in the present study, appear to have more dense BLA projections. Notably, rats with BLA lesions are not impaired in the acquisition of simple Pavlovian or instrumental tasks, indicating that basic reward mechanisms are intact in these animals (Blundell *et al.*, 2001; Balleine *et al.*, 2003; Corbit & Balleine, 2005). Instead, the BLA may function to assign motivational significance to stimuli (and perhaps responses), particularly when animals must discriminate multiple rewards with reference to their sensory-specific properties (Corbit & Balleine, 2005; Balleine & Killcross, 2006; Johnson *et al.*, 2009), rather than to simply associate neutral events (Blundell *et al.*, 2003; Dwyer & Killcross, 2006). Impairment of this more specific function is strikingly similar to the pattern of deficits that we have found following pDMS

inactivation, suggesting a potential functional connection between these structures in the control of selective responding related to rewards with unique sensory features. Given the sparse BLA projections to the DLS, the BLA is unlikely to directly contribute to DLS-mediated S-O learning.

Other regions that may contribute to the differences in behaviour observed after inactivation of DS subregions could include cortical and thalamic inputs. For example, the DMS receives afferents from the pre-limbic cortex (Berendse *et al.*, 1992), a region that has been shown in the rat to be critical for the acquisition of goal-directed behaviours (Corbit & Balleine, 2003b; Ostlund & Balleine, 2005). The DMS also receives input from the pre-motor or medial agranular cortex involved in action monitoring and programming (Reep *et al.*, 2003). However, the DLS receives input from primary motor and somatosensory cortices (Ramanathan *et al.*, 2002; Alloway *et al.*, 2006) and is thus well situated to integrate sensory inputs with response generation. Indeed, this pattern of innervation has led to much work examining the role of the DLS in S-R learning (Packard & Knowlton, 2002; Featherstone & McDonald, 2004; Balleine *et al.*, 2009). Notably, the orbitofrontal cortex projects broadly through the striatum in a topographic manner (Schilman *et al.*, 2008). Hence, within the DLS, information regarding reward/reinforcing events from the orbitofrontal cortex (and perhaps midbrain dopaminergic inputs) may be integrated with sensory inputs. In this way, DLS function need not be limited to S-R learning in the classic sense but rather, as the current data suggest, the DLS may play a broader role in stimulus-reward learning and the control of stimulus influences on responding, discussed further below.

Relation to stimulus-response or habit learning

Much previous work has focused on the role of the DLS in S-R and habit learning (Packard & Knowlton, 2002; Yin & Knowlton, 2006). Within this framework, reward acts to strengthen an S-R association between environmental stimuli and the response performed and, as a result, the tendency to perform that response when those stimuli are encountered is increased (Thorndike, 1911; Hull, 1943). Critically, such an association does not rely on a representation of the outcome itself. It is interesting to note that, when the DLS was inactivated during the Pavlovian training sessions with novel outcomes, responding was not affected. This is intriguing because one might have expected performance during the inactivation sessions to be impaired if responding were supported by previous S-R associations acquired during the baseline training sessions (e.g. click-left magazine or noise-right magazine) that were dependent upon DLS function. However, impairment was observed only when rats were required to retrieve and/or update the current value of a particular outcome and its paired stimulus, a function not typically ascribed to the DLS. This suggests that, although basic Pavlovian reinforcement mechanisms may be intact, the DLS is critical for integrative or evaluative processes essential for specific stimuli, and their relation to unique rewards, to direct responding. This mirrors the deficits that we saw previously with inactivation of the DLS prior to a PIT test where simple instrumental responding was intact but the modulation of this responding by reward-predictive cues was lost (Corbit & Janak, 2007b). Hence, the DLS may integrate excitatory information carried by stimuli with the control of particular responses. This integration could occur through the special case of S-R learning where the retrieved stimulus properties of the outcome initiated by stimulus presentation (stimulus retrieves outcome) signal the associated response (outcome retrieves response) based on instrumental training and thus contribute to response initiation, although further experiments will be needed to test this specific hypothesis. The current results, together with our previous findings within the PIT paradigm, suggest that the DLS plays a more fundamental role in allowing excitatory stimuli to impact response selection and generation, which may include but is not limited to the S-R architecture.

Limitations and future directions

A possible concern in the current studies is that the same rats were used for the instrumental and Pavlovian tasks and that the order of these tests was not counterbalanced. This could provide an opportunity for learning in the first task to carry over to the second task. For example, for the aDMS group, despite impaired performance in the instrumental task (conducted first), it is possible that some learning occurred (e.g. discriminating the outcomes) that contributed to performance in the Pavlovian task (conducted second). There are several reasons why we do not believe that this explains the full pattern of results. First, rats in the DLS group showed normal performance in the instrumental task but a deficit in the Pavlovian task, indicating that ability to perform the Pavlovian task was not simply the result of learning occurring during the instrumental task. Second, devaluation by specific satiety is transient and reversible in that responding typically returns to baseline by the following day and as such it is not expected that experience with the devaluation treatment in the instrumental task would alter performance in the Pavlovian task. Further, these methods are similar to many previous permanent lesion studies where rats undergo repeated sequential testing on a variety of tasks. Nonetheless, future studies can reverse and/or balance the order of the tasks to rule out this possible confound.

Conclusions

In summary, in addition to previously established differences in anatomical connections and functional circuits, the current study demonstrates that the DS can be functionally divided based on the contribution to different types of associative learning. The dissociation of functions within the DS suggests control by functionally distinct parallel pathways and although integration, either through cooperation or competition, must ultimately occur to determine behavioural performance, the exact locus of this integration remains unknown (Balleine & O'Doherty, 2010). These findings highlight the importance of rostro-caudal anatomy, in addition to the well-noted mediolateral distinctions. Notably, similar rostro-caudal functional differences have been described in other striatal and limbic structures including the nucleus accumbens and ventral pallidum (Reynolds & Berridge, 2001, 2003; Smith & Berridge, 2005). Further, these data indicate that the role of the DS in stimulus-directed responding is more complex than simply mediating the formation and/or implementation of S-R learning and demonstrate that this region also contributes to aspects of goal-directed performance.

Acknowledgments

The research reported in this study was supported by NIH R01 AA018025 and by US Army Medical Research Acquisition Award W81XWH-07-1-0076. We thank Victoria Powell and Natalie Warrick for technical assistance.

Abbreviations

Ø	no stimulus
aDMS	anterior dorsomedial striatum
BLA	basolateral complex of the amygdala
DLS	dorsolateral striatum
DMS	dorsomedial striatum
DS	dorsal striatum
O-R	outcome-response

pDMS	posterior dorsomedial striatum
PIT	Pavlovian-instrumental transfer
R-O	response-outcome
RR	random-ratio
S-O	stimulus-outcome
S-R	stimulus-response

References

- Adams CD, Dickinson A. Instrumental responding following reinforcer devaluation. *Q J Exp Psychol* 1981;33B:109–121.
- Alexander GE, DeLong MR, Strick PL. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu Rev Neurosci* 1986;9:357–381. [PubMed: 3085570]
- Alloway KD, Lou L, Nwabueze-Ogbo F, Chakrabarti S. Topography of cortical projections to the dorsolateral neostriatum in rats: multiple overlapping sensorimotor pathways. *J Comp Neurol* 2006;499:33–48. [PubMed: 16958106]
- Aosaki T, Graybiel AM, Kimura M. Effect of the nigrostriatal dopamine system on acquired neural responses in the striatum of behaving monkeys. *Science* 1994;265:412–415. [PubMed: 8023166]
- Balleine BW, Dickinson A. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropsychopharmacology* 1998;37:407–419.
- Balleine BW, Killcross S. Parallel incentive processing: an integrated view of amygdala function. *Trends Neurosci* 2006;29:272–279. [PubMed: 16545468]
- Balleine BW, O'Doherty JP. Human and rodent homologues in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology* 2010;35:48–69. [PubMed: 19776734]
- Balleine BW, Ostlund SB. Still at the choice point: action selection and initiation in instrumental conditioning. *Ann N Y Acad Sci* 2007;1104:147–171. [PubMed: 17360797]
- Balleine BW, Killcross AS, Dickinson A. The effect of lesions of the basolateral amygdala on instrumental conditioning. *J Neurosci* 2003;23:666–675. [PubMed: 12533626]
- Balleine BW, Liljeholm M, Ostlund SB. The integrative function of the basal ganglia in instrumental conditioning. *Behav Brain Res* 2009;199:43–52. [PubMed: 19027797]
- Berendse HW, Galis-de Graaf Y, Groenewegen HJ. Topographical organization and relationship with ventral striatal compartments of prefrontal corticostriatal projections in the rat. *J Comp Neurol* 1992;316:314–347. [PubMed: 1577988]
- Blundell P, Hall G, Killcross S. Lesions of the basolateral amygdala disrupt selective aspects of reinforcer representation in rats. *J Neurosci* 2001;21:9018–9026. [PubMed: 11698612]
- Blundell P, Hall G, Killcross S. Preserved sensitivity to outcome value after lesions of the basolateral amygdala. *J Neurosci* 2003;23:7702–7709. [PubMed: 12930810]
- Colwill RM, Rescorla RA. Effect of reinforcer devaluation on discriminative control of instrumental behavior. *J Exp Psychol Anim Behav Process* 1990;16:40–47. [PubMed: 2303793]
- Corbit LH, Balleine BW. Instrumental and Pavlovian incentive processes have dissociable effects on components of a heterogeneous instrumental chain. *J Exp Psychol Anim Behav Process* 2003a; 29:99–106. [PubMed: 12735274]
- Corbit LH, Balleine BW. The role of prelimbic cortex in instrumental conditioning. *Behav Brain Res* 2003b;146:145–157. [PubMed: 14643467]
- Corbit LH, Balleine BW. Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of pavlovian-instrumental transfer. *J Neurosci* 2005;25:962–970. [PubMed: 15673677]
- Corbit LH, Janak PH. Ethanol-associated cues produce general pavlovian-instrumental transfer. *Alcohol Clin Exp Res* 2007a;31:766–774. [PubMed: 17378919]

- Corbit LH, Janak PH. Inactivation of the lateral but not medial dorsal striatum eliminates the excitatory impact of Pavlovian stimuli on instrumental responding. *J Neurosci* 2007b;27:13977–13981. [PubMed: 18094235]
- Corbit LH, Muir JL, Balleine BW. The role of the nucleus accumbens in instrumental conditioning: evidence of a functional dissociation between accumbens core and shell. *J Neuroscience* 2001;21:3251–3260.
- Corbit LH, Janak PH, Balleine BW. General and outcome-specific forms of Pavlovian-instrumental transfer: the effect of shifts in motivational state and inactivation of the ventral tegmental area. *Eur J Neurosci* 2007;26:3141–3149. [PubMed: 18005062]
- Cromwell HC, Schultz W. Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. *J Neurophysiol* 2003;89:2823–2838. [PubMed: 12611937]
- Dickinson A, Balleine BW. Motivational control of goal-directed action. *Anim Learn Behav* 1994;22:1–18.
- Dwyer DM, Killcross S. Lesions of the basolateral amygdala disrupt conditioning based on the retrieved representations of motivationally significant events. *J Neurosci* 2006;26:8305–8309. [PubMed: 16899725]
- El-Amamy H, Holland PC. Dissociable effects of disconnecting amygdala central nucleus from the ventral tegmental area or substantia nigra on learned orienting and incentive motivation. *European J Neuroscience* 2007;25:1557–1567.
- Featherstone RE, McDonald RJ. Dorsal striatum and stimulus-response learning: lesions of the dorsolateral, but not dorsomedial, striatum impair acquisition of a stimulus-response-based instrumental discrimination task, while sparing conditioned place preference learning. *Neuroscience* 2004;124:23–31. [PubMed: 14960336]
- Haber SN, Kim KS, Mailly P, Calzavara R. Reward-related cortical inputs define a large striatal region in primates that interface with associative cortical connections, providing a substrate for incentive-based learning. *J Neurosci* 2006;26:8368–8376. [PubMed: 16899732]
- Hassani OK, Cromwell HC, Schultz W. Influence of expectation of different rewards on behavior-related neuronal activity in the striatum. *J Neurophysiol* 2001;85:2477–2489. [PubMed: 11387394]
- Hatfield T, Han JS, Conley M, Gallagher M, Holland P. Neurotoxic lesions of basolateral, but not central, amygdala interfere with Pavlovian second-order conditioning and reinforcer devaluation effects. *J Neurosci* 1996;16:5256–5265. [PubMed: 8756453]
- Holland PC. Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *J Exp Psychol Anim Behav Process* 2004;30:104–117. [PubMed: 15078120]
- Hull, CL. *Principles of Behavior: An Introduction to Behavior Theory*. Appleton-Century-Crofts; New York: 1943.
- Johnson AW, Gallagher M, Holland PC. The basolateral amygdala is critical to the expression of pavlovian and instrumental outcome-specific reinforcer devaluation effects. *J Neurosci* 2009;29:696–704. [PubMed: 19158296]
- Kelley AE, Domesick VB, Nauta WJ. The amygdalostratial projection in the rat—an anatomical study by anterograde and retrograde tracing methods. *Neuroscience* 1982;7:615–630. [PubMed: 7070669]
- Kruse JM, Overmier JB, Kon WA, Rokke E. Pavlovian conditioned stimulus effects upon instrumental choice behavior. *Learn Motiv* 1983;14:165–181.
- Ostlund SB, Balleine BW. Lesions of medial prefrontal cortex disrupt the acquisition but not the expression of goal-directed learning. *J Neurosci* 2005;25:7763–7770. [PubMed: 16120777]
- Packard MG, Knowlton BJ. Learning and memory functions of the Basal Ganglia. *Annu Rev Neurosci* 2002;25:563–593. [PubMed: 12052921]
- Parent A, Hazrati LN. Functional anatomy of the basal ganglia. I The cortico-basal ganglia-thalamo-cortical loop. *Brain Res Brain Res Rev* 1995;20:91–127. [PubMed: 7711769]
- Paxinos, G.; Watson, C. *The Rat Brain in Stereotaxic Coordinates*. Academic Press; San Diego: 1998.
- Ramanathan S, Hanley JJ, Deniau JM, Bolam JP. Synaptic convergence of motor and somatosensory cortical afferents onto GABAergic interneurons in the rat striatum. *J Neurosci* 2002;22:8158–8169. [PubMed: 12223570]

- Reep RL, Cheatwood JL, Corwin JV. The associative striatum: organization of cortical projections to the dorsocentral striatum in rats. *J Comp Neurol* 2003;467:271–292. [PubMed: 14608594]
- Reynolds SM, Berridge KC. Fear and feeding in the nucleus accumbens shell: rostrocaudal segregation of GABA-elicited defensive behavior versus eating behavior. *J Neurosci* 2001;21:3261–3270. [PubMed: 11312311]
- Reynolds SM, Berridge KC. Glutamate motivational ensembles in nucleus accumbens: rostrocaudal shell gradients of fear and feeding. *Eur J Neurosci* 2003;17:2187–2200. [PubMed: 12786986]
- Schilman EA, Uylings HB, Galis-de Graaf Y, Joel D, Groenewegen HJ. The orbital cortex in rats topographically projects to central parts of the caudate-putamen complex. *Neurosci Lett* 2008;432:40–45. [PubMed: 18248891]
- Smith KS, Berridge KC. The ventral pallidum and hedonic reward: neurochemical maps of sucrose “liking” and food intake. *J Neurosci* 2005;25:8637–8649. [PubMed: 16177031]
- Thorndike, EL. *Animal Intelligence: Experimental Studies*. Macmillan; New York: 1911.
- Voorn P, Vanderschuren LJ, Groenewegen HJ, Robbins TW, Pennartz CM. Putting a spin on the dorsal-ventral divide of the striatum. *Trends Neurosci* 2004;27:468–474. [PubMed: 15271494]
- de Wit S, Ostlund SB, Balleine BW, Dickinson A. Resolution of conflict between goal-directed actions: outcome encoding and neural control processes. *J Exp Psychol Anim Behav Process* 2009;35:382–393. [PubMed: 19594283]
- Yin HH, Knowlton BJ. The role of the basal ganglia in habit formation. *Nat Rev Neurosci* 2006;7:464–476. [PubMed: 16715055]
- Yin HH, Knowlton BJ, Balleine BW. Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. *Eur J Neurosci* 2005;22:505–512. [PubMed: 16045503]

Fig. 1.

Reconstruction of microinfusion sites in the DS. The location of the cannula tips is represented by circles for the DLS group, squares for the aDMS group and triangles for the pDMS group. Numbers on the left indicate distance relative to bregma in mm. The microinfusion sites are superimposed on coronal sections (adapted from the rat brain atlas of Paxinos & Watson, 1998).

Fig. 2. Inactivation of the aDMS and pDMS impairs acquisition of novel R-O associations as revealed in a subsequent test of outcome devaluation. (A) Mean responding (+ SEM) on both levers combined during pre-training with the same outcome (polycose) earned by both levers. There were no differences among groups. (B) Mean responding per session (+ SEM) on both levers combined during specific R-O training during which rats were exposed to two unique R-O pairs following a pre-session microinfusion of saline (control) or baclofen/muscimol (aDMS, pDMS, and DLS groups). Data are averaged across three consecutive training/inactivation sessions. The aDMS group responded more than the pDMS group but neither group was significantly different from saline controls. (C) Mean responding (+ SEM)

on each of two levers during the 15-min devaluation test. Subjects were pre-fed with one of the two outcomes and the lever that previously delivered that outcome was termed 'devalued' and the lever that previously delivered the outcome that was not pre-fed was termed 'non-devalued'. Saline control rats and rats with prior inactivation of the DLS during specific R-O training show a significant decrease in responding on the devalued lever relative to the non-devalued lever. Rats with prior inactivation of the aDMS and pDMS during specific R-O training show no significant differences in responding on the two levers. Control, $N = 17$; aDMS, $N = 10$; pDMS, $N = 9$; DLS, $N = 9$. $*P < 0.05$, $**P < 0.01$.

Fig. 3. Inactivation of the DLS and pDMS impairs acquisition of novel S-O associations as revealed in a subsequent test of outcome devaluation. (A) Mean magazine (reward port) responding (+ SEM) during pre-training with the same outcome available at each magazine. 'Rewarded' refers to the magazine that delivered the reward during the current stimulus, whereas 'alternate' refers to the magazine that was rewarded during the other (currently not present) stimulus; 'pre' refers to the period prior to stimulus presentation; 'stimulus' refers to the period during the stimulus. There were no differences among groups. (B) Mean magazine responding per session (+ SEM) during specific SO training (average of three training days) during which rats were exposed to two unique S-O pairs following a pre-session

microinfusion of saline (control) or baclofen/muscimol (aDMS, pDMS and DLS). The aDMS group responded more than the pDMS group during stimulus presentation but neither group was significantly different from saline controls. The aDMS group made more entries to the alternate magazine than all other groups. (C) Mean magazine responding (+ SEM) during the stimuli or equivalent pre-stimulus periods in the devaluation test. Subjects were pre-fed with one of the two outcomes and the stimulus previously associated with that outcome was termed 'devalued' and the stimulus previously associated with the outcome that was not pre-fed was termed 'non-devalued'. Saline control rats and rats with prior inactivation of the aDMS during specific S-O training show a significant decrease in responding during the devalued stimulus relative to the non-devalued stimulus. Rats with prior inactivation of the DLS and pDMS during specific S-O training show no significant differences in responding during the two stimuli. Group N-values as in Fig. 1. * $P < 0.05$; ** $P < 0.01$.

Fig. 4.

Expression of PIT is impaired if subjects have failed to acquire either the selective R-O or selective S-O associations that underlie performance in this task. Mean lever-press responding (+ SEM) during test sessions during periods with no stimuli (baseline) and during presentation of the two stimuli used during the specific S-O training (no rewards were delivered). 'Same' refers to the stimulus paired with the same outcome as was previously earned by pressing the available lever; 'different' refers to the stimulus paired with the other outcome. All groups elevated responding from baseline during presentations of either stimulus. Only the control group showed a significantly greater enhancement of responding during presentation of the same vs. different stimulus. $**P < 0.01$.