



Published in final edited form as:

*J Exp Psychol Anim Behav Process.* 2013 January ; 39(1): 2–13. doi:10.1037/a0030941.

## Hierarchical and binary associations compete for behavioral control during instrumental biconditional discrimination

Laura A. Bradfield and Bernard W. Balleine

Brain & Mind Research Institute, University of Sydney

### Abstract

In two experiments we investigated the role of hierarchical S-(R-O) associations, as opposed to associative alternatives, in solving biconditional discrimination problems in rats. Using lesions of posterior dorsomedial striatum, known to attenuate R-O associative learning, and lesions of the dorsolateral striatum, that attenuate S-R learning, we found that whereas the lesions affecting R-O learning abolished biconditional discrimination, lesions of dorsolateral striatum did not (Experiment 1). Furthermore, in Experiment 2, we found, using a more challenging discrimination protocol, that dorsolateral striatal lesions actually enhanced biconditional discrimination learning. These results provide evidence that hierarchical S-(R-O) associations influence instrumental discrimination learning and compete with S-R associations for control of performance.

Experiments investigating instrumental conditioning typically train animals to perform an action to earn a specific outcome in the presence of particular stimuli. Both theory and evidence accumulating over many decades suggest that, as a consequence of this training, the animals encode various binary associations, whether between representations of the stimulus and response (S–R associations), the response and outcome (R–O associations), or the stimulus and outcome (S–O associations) (Bolles, 1972; Colwill & Rescorla, 1986; Dickinson, 1994). In addition to these binary associations, however, it has been suggested that instrumental training can also encourage the formation of hierarchical associations. For example, Colwill and Rescorla (1990) presented evidence that hierarchical associations between a discriminative stimulus and specific response-outcome associations are formed as a consequence of instrumental biconditional discrimination training. In Experiment 2 they used a design in which rats learned to perform two responses (R1 and R2) to earn two distinct outcomes (O1 and O2) during two discriminative stimuli (S1 and S2) such that, when S1 was presented, the rats earned O1 for performing R1 and O2 for performing R2 whereas, when S2 was presented these response-outcome contingencies were reversed (i.e. R1 earned O2 and R2 earned O1). Thus four hierarchical associations should have formed: S1-(R1-O1), S1-(R2-O2), S2-(R1-O2), and S2-(R2-O1). To establish evidence of hierarchical associations one of the two outcomes (O1) was subsequently devalued by pairing its consumption with lithium chloride after which the rats were allowed to choose between R1 and R2 in an extinction test in the presence of S1 and S2. If, as a consequence of training, the animals only encode the binary associations between the S's, R's and O's,

then no difference in the performance of R1 and R2 should be predicted because both R1 and R2 were equally associated with both O1 and O2 and both S1 and S2. If, however, rats are able to encode hierarchical associations, such that the specific R-O associations are controlled by the discriminative stimuli, then differences in the performance of R1 and R2 should emerge with the direction of the difference depending on stimulus presentation. Colwill and Rescorla's results were consistent with this latter prediction; that is, when S1 was presented R1 was reduced relative to R2 (S1:  $R1 < R2$ ), whereas, when S2 was presented R2 was reduced relative to R1 (S2:  $R1 > R2$ ).

Although simple binary associations cannot account for this result, there are possibilities other than the S-(R-O) hierarchical association favoured by Colwill & Rescorla that can. The simplest alternatives are perhaps other types of hierarchical association such as those illustrated in Figure 1. For example, an (S-O)-R structure, which provides that the performance of a response can be controlled by a specific S-O association, can explain biconditional discrimination performance if it is accepted that devaluing a specific O reduces this control. Likewise, an (S-R)-O structure can resolve the discrimination if the influence of a specific S-R association on performance is controlled by the value of the outcome with which it is associated. Another alternative is provided by Rescorla's (1991) suggestion that the rats in his experiments may have combined binary associations into ternary ones; i.e. having formed both an S-R and an R-O association from the S, R and O elements, the rats might then have formed a ternary S-R-O association (see Figure 1). This account can explain discrimination given that any one element can retrieve the other elements of the association and if performance is ultimately determined by the value of the outcome component.

Of these various explanations, only the (S-O)-R alternative has received any experimental attention, perhaps as a consequence of expectancy-based two-process theory (Trapold, 1970) and its popularity as an explanation of the differential outcomes effect in discrimination learning (Trapold & Overmeir, 1972). Rescorla (1992 - Experiment 3) devised a means of testing predictions from the (S-O)-R and S-(R-O) accounts by constructing a discrimination situation in which two antecedent S-O associations, S1-O1 and S2-O2, controlled two responses, R1 and R2, each earning the outcome that was not signaled by the stimulus; i.e.  $R1 \rightarrow O2$  and  $R2 \rightarrow O1$ . Thus, two hierarchical associations should have formed: (S1-O1)- $R1 \rightarrow O2$  whereas (S2-O2)- $R2 \rightarrow O1$ . If specific (S-O)-R associations control performance then devaluing O1 should reduce R1 relative to R2. Rescorla found, however, that devaluing O1 reduced R2 relative to R1 suggesting that R-O rather than S-O associations exerted greater control over performance and, therefore, that (S-O)-R hierarchical associations likely play little if any role in this form of discrimination.

To date there has not been any direct evidence to decide between the S-(R-O), the (S-R)-O and the S-R-O ternary accounts. One way in which these latter two structures might be thought to differ from the former is in the contribution of S-R associations to task solution. On both the (S-R)-O and the ternary account, of course, S-R associations are a necessary component of the underlying associative structure. As a consequence, a diminished capacity to form S-R associations might be expected to produce a deficit in performance. If, on the other hand, S-(R-O) associations governed performance, then a reduced capacity to form S-

R associations would be expected to leave performance intact. Indeed, any tendency to form S–R associations might be expected to interfere with accurate performance as a result of competition between S-R and S-(R-O) associations. That is, when the rat is faced with a choice between, say, R1 and R2 during S1 presentations, S-R associations would produce a tendency to increase the performance of both actions (i.e. both R1 and R2), which would compete with the hierarchical S1-(R1-O1) and S1-(R2-O2) associations that promote discriminative control (i.e. R1 during S1 and R2 during S2). Thus, from this perspective, a diminished ability to form S-R associations should either have no effect or, by removing a source of interference, improve performance on this task.

In this context, recent evidence suggesting that either lesions or pharmacological inactivation of the dorsolateral striatum (DLS) attenuates the ability of rats to form stimulus-response associations is of considerable interest (Yin, Knowlton, & Balleine, 2004, 2006). One area in which S–R associations are thought to play a central role is in the control of habitual instrumental performance after overtraining (Dickinson, 1985). Although undertrained actions are sensitive to outcome devaluation and to degradation of the response-outcome contingency, overtrained instrumental actions are not and persist when either the value of the outcome or the causal status of the instrumental action is changed (Dezfouli & Balleine, 2012). Furthermore, in a recent study we found evidence that correctly choosing between actions based on discriminative cues, in a task in which a simple S<sup>D</sup>-R association provided an optimal solution, was strongly attenuated by inactivation of the DLS (Balleine, Liljeholm, & Ostlund, 2009; see also McDonald & White, 1993).

Based on these findings, different predictions can be made about the likely effect of DLS lesions during a biconditional discrimination task. If accurate performance on this task requires the formation of S-R associations, then lesions of the DLS might be anticipated to attenuate performance on a biconditional task. If, however, solving the biconditional task requires the formation of hierarchical S-(R-O) associations then DLS lesions should not produce a deficit in performance. Indeed, by removing a potential source of interference, lesioning the DLS might well be anticipated to facilitate task performance relative to sham controls.

In contrast to these effects of DLS manipulation, lesions of an adjacent medial region of dorsal striatum, the dorsomedial striatum (DMS), do not affect these measures of S-R learning but, rather, appear to influence the ability of animals to form response-outcome associations (Balleine & Doherty, 2010). Thus, in undertrained animals, lesions or inactivation of the posterior DMS (pDMS) abolish sensitivity to both outcome devaluation and contingency degradation (Shiflett, Brown, & Balleine, 2010; Yin, Ostlund, Knowlton, & Balleine, 2005). Given that both the ternary and hierarchical accounts of biconditional discrimination performance rely upon animals being able to learn R-O associations, both of these accounts predict that lesions of the pDMS will attenuate performance.

The aim of the current study was to test these predictions. Experiment 1 compared the effect of DLS and pDMS lesions on a biconditional discrimination task similar to that employed by Colwill and Rescorla (1990 – Experiment 2). In Experiment 2 we conducted a similar

assessment except that we trained the rats on a more difficult version of the task than that employed in Experiment 1.

## Experiment 1

The purpose of Experiment 1 was to assess the effects of DLS lesions on an instrumental biconditional discrimination task. The behavioral design for this experiment was based on that used by Colwill and Rescorla (1990 – Experiment 2) and is shown in Table 1. During each stage rats received two discriminative stimuli, a clicker and noise (S1 and S2, counterbalanced), during which the left and right levers (R1 and R2) were made available. Rats were trained with one of the two outcomes during Stage 1 (O1: either pellets or sucrose counterbalanced) and with the alternate outcome in Stage 2 (O2: either sucrose or pellets). The different discriminative stimuli signalled different R-O relations; i.e., S1: (R1→O1) and S2: (R2 → O1) during Stage 1, and S1: (R2 → O2) and S2: (R1 → O2) during Stage 2. These discriminations were combined during Stage 3 in a manner consistent with Stages 1 and 2 such that S1: R1 → O1, R2 → O2 and S2: R1 → O2, R2 → O1. After this training but prior to test, one of the two outcomes was devalued by prefeeding it to satiety. Rats were then given a choice extinction test during which each S<sup>D</sup> was presented and pressing on both levers was recorded.

Sham rats that learned the associations necessary to complete the biconditional discrimination task should, on test, choose the lever associated with the nondevalued outcome, the identity of which was dependent on which S was presented. For example, if O1 was devalued, then rats should perform R2 > R1 in the presence of S1, but R1 > R2 in the presence of S2. Likewise, if O2 were devalued then rats should perform R1 > R2 in the presence of S1 but R2 > R1 in the presence of S2. If rats solve this task using a hierarchical S-(R-O) association then a similar effect should emerge in a group given lesions of the DLS. In contrast, if, as a component of the associative structure, rats need to form S-R associations to solve the task then lesions of the DLS should be expected to produce a deficit in performance relative to sham controls. Hence, in this test, our initial predictions were assessed by contrasting performance in the DLS group with the sham controls.

Secondly, we compared the effects of sham and DLS lesions with those produced by lesions of the pDMS on this task. As suggested above, lesions of pDMS were expected to produce a deficit in the capacity of animals to solve this task using the R-O association. On this basis we predicted that pDMS lesions would produce a deficit in performance. This will either differentiate it from both the sham and DLS groups or from the sham group alone. If the former then we anticipate that a contrast comparing performance in the DLS and sham groups against the pDMS group should generate a significant interaction; a finding that would provide clear evidence for the hierarchical S-(R-O) account.

## Method

**Subjects and Apparatus**—Thirty male Long-Evans rats, weighing between 300 and 400g at the beginning of the experiment, were used as subjects. The rats were housed in pairs in transparent yellow-tinted plastic tubs (.5m<sup>3</sup>) located in a temperature- and humidity-controlled vivarium. Throughout behavioral training and testing, rats were maintained at

~85% of their free-feeding body weight by restricting their food intake to 10g (each) of their maintenance diet per day. This daily food allotment was reduced by half on outcome devaluation test days, when rats were provided with 1 h of *ad libitum* access to one of the training outcomes. Rats were provided tap water *ad libitum* while in their home cage.

The behavioral procedures were performed in 16 identical Med Associates (East Fairfield, VT, USA) operant chambers enclosed in sound- and light-attenuating shells. Each chamber was equipped with a recessed food magazine, located at the base of one end wall, through which 20% sucrose solution (0.1 ml) and food pellets (45 mg; Bio-Serv, Frenchtown, NJ, USA) could be delivered using a syringe pump and pellet dispenser, respectively. Each chamber also contained a pair of retractable levers that were located to the left or right of the food magazine. A houselight (3 W, 24 V) located on the end wall opposite the magazine provided constant illumination, and an electric fan fixed in the shell enclosure provided background noise (~70 dB) throughout training and testing. Two microcomputers running the Med-PC program (Med Associates) controlled all experimental events and recorded lever presses and magazine entries. The boxes also contained a white-noise generator, a sonalert that delivered a 3 kHz tone, and a solenoid that, when activated, delivered a 5 Hz clicker stimulus. All stimuli were adjusted to 80 dB in the presence of background noise of 60 dB provided by a ventilation fan. Outcome devaluation procedures took place in transparent plastic tubs that were smaller, but otherwise identical to the cages in which rats were housed.

**Surgery**—Rats were provided *ad libitum* access to their maintenance chow prior to, and on the 4 days that followed, surgery. Approximately 10 min before surgery rats were injected intraperitoneally (i.p.) with 1.3 ml/kg of the anaesthetic ketamine (CenVet Australia Pty Ltd, Sydney, Australia) at a concentration of 100 mg/mL and with the muscle relaxant xylazine (0.3mL/kg; Rompun; Bayer, Sydney, Australia) at a concentration of 20 mg/mL. Each rat was then placed in a stereotaxic frame (Stoelting, Wood Dale, IL, USA). An incision was made into the scalp to expose the skull surface, and the incisor bar was adjusted to place bregma and lambda in the same horizontal plane. For all rats, two small holes were drilled into the skull above the target structure(s). Excitotoxic lesions were made by infusing 0.6 $\mu$ l (pDMS lesions: all coordinates in millimetres relative to bregma: anteroposterior, -0.4, mediolateral,  $\pm$  2.2, dorsoventral, -4.5) or 0.5 $\mu$ l (DLS lesions: anteroposterior, + 0.7, mediolateral,  $\pm$  3.4, dorsoventral, -5) of NMDA (10mg/mL) in sterilized 0.1M phosphate buffered saline (PBS) pH 7.2 over 6 min (pDMS lesions) or 5 min (DLS lesions). The needle was left in place for 4 min prior to removal to allow for diffusion. Sham-operated rats underwent the same procedures (half with Sham pDMS and half with Sham DLS insertions) but no neurotoxin was infused. Immediately after surgery, rats were injected intramuscularly (i.m.) with 0.2 mL of a 300mg/mL solution of procaine penicillin, and subcutaneously (s.c.) with 5 mg/kg carprofen. Rats were given 7 days to recover from surgery before experimentation commenced. Rats were weighed and handled daily and subject to food deprivation in the last 3 days of this period.

**Histology**—At the conclusion of the experiments rats were deeply anaesthetised with sodium pentobarbital (100mg/kg i.p.) and perfused transcardially with ~ 400mL of 4%

paraformaldehyde in 0.1M phosphate buffer (PB), pH 7.4. Brains were postfixed for 1 h in the same fixative and placed in PBS pH 7.2 20% sucrose solution overnight. Brains were blocked using a matrix aligned to the atlas of Paxinos and Watson (1998) and 40  $\mu$ m coronal sections were cut using a cryostat. Every fourth section was collected on a slide and stained with cresyl violet. Slides were examined for placement and extent of the lesion; the latter was assessed by microscopically examining sections for areas of marked cell loss as well as general shrinkage of a region relative to controls.

### Procedure

**Magazine Training:** On days 1 and 2 all rats were placed in operant chambers for 20 min. In each session the house light was illuminated at the start of the session and turned off when the session was terminated. During magazine training, 20 pellet and 20 sucrose outcomes were delivered into the magazine on independent random time (RT) 60 s schedules. No levers were extended during magazine training.

**Lever training:** On day 3 the rats were trained to lever press on a continuous reinforcement schedule. Both levers were paired with both outcomes within the same session. Specifically, the left lever was extended for two 10 min periods; during the first period the lever earned O1 and during the second O2. The right lever was then extended for two 10 min periods, earning O1 in the first and O2 in the second period. Each of the 10 min periods was separated by 2 min during which both levers were retracted and the houselight switched off.

**Stage 1 of Discrimination Training:** Stage 1 training occurred on days 4 to 14. During this stage rats were trained with O1 only (pellets or sucrose counterbalanced). During S1 (i.e. the tone or the noise, counterbalanced), responding on one lever earned O1 and responding on the other earned nothing (S1: R1  $\rightarrow$  O1, R2  $\rightarrow$   $\emptyset$ ). During S2 presentations R1 earned nothing whereas responding on R2 earned O1 (S1: R1  $\rightarrow$   $\emptyset$ , R2  $\rightarrow$  O1). During each session of this stage and each of the remaining stages of discrimination training (i.e. Stages 2, and 3), there were twelve 1 min presentations of S1 and twelve 1 min presentations of S2. Responding in the presence of these stimuli was reinforced on a RI-20 s schedule and the order of stimulus presentation was pseudo-random. The inter-trial-interval (ITI) was random about a mean of 45 s, and during this time the levers were retracted.

**Stage 2 of Discrimination Training:** Stage 2 training occurred on days 15 to 25. During this stage rats were trained with O2 only. During S1 presentations R1 earned nothing and R2 earned O2 (S1: R1  $\rightarrow$   $\emptyset$ , R2  $\rightarrow$  O2). During S2 presentations R1 earned O2 and on R2 earned nothing (S1: R1  $\rightarrow$  O2, R2  $\rightarrow$   $\emptyset$ ).

**Stage 3: Biconditional Discrimination Training:** Stage 3 training occurred on days 26 – 32. During this stage rats were trained with both O1 and O2 in a manner consistent with Stage 1 and Stage 2 training. That is, during S1 presentations R1 earned O1 and R2 earned O2 (S1: R1  $\rightarrow$  O1, R2  $\rightarrow$  O2). During S2 presentations R1 earned O2 and R2 earned O1 (S1: R1  $\rightarrow$  O2, R2  $\rightarrow$  O1). All other aspects of training were the same as Stage 1.



**Devaluation extinction tests:** Following the final day of biconditional discrimination training, rats were given access *ad libitum* to either the pellets (25g placed in a bowl in the devaluation cage) or the sucrose solution (100mL in a drinking bottle fixed to the top of the devaluation cage) for 1 hr. The aim of this prefeeding procedure was to satiate the animal specifically on the preferred outcome, thereby reducing its value relative to the non-preferred outcome (cf. Balleine and Dickinson, 1998). Rats were then placed in the operant chamber for an extinction test during which each stimulus was presented twice (i.e. 4 stimulus presentations in total). The levers were extended during stimulus presentations and retracted during the ITI, which was random about a mean of 45 s. Lever presses were recorded but no outcomes were delivered. The next day a second devaluation test was administered with the opposite outcome; i.e., if rats were previously preferred pellets they received sucrose, and if preferred sucrose they now received pellets. Rats were then placed into the operant chambers for a second extinction test conducted as for the first.

## Results and Discussion

**Histology**—Histological analysis determined the size and placement of lesions. Figure 2 shows a representation of the minimal (black) and the maximal (grey) extent of the lesions based on the stereotaxic atlas of the rat brain by Paxinos and Watson (1998). Figure 3 displays photomicrographs taken of representative lesions of the sham, pDMS and DLS groups. pDMS lesions targeted the most caudal extent of the pDMS and produced substantial cell loss and shrinkage in this region. DLS lesions targeted the mid anterior-posterior point of the dorsal striatum and produced substantial cell loss in this region. Four animals were excluded from the experiment because of incorrect lesion placement or size. Thus 26 animals were included in the analysis (Sham,  $n = 10$ , pDMS,  $n = 7$ , DLS,  $n = 9$ ).

**Stage 1 training**—The mean response rates for the various groups during acquisition are shown in the top left panel of Figure 4 collapsed across stimulus identity. Inspection of the figure suggests that rats in each group performed more of the reinforced than the nonreinforced response (as guided by the relevant stimuli) and that responding was similar across the groups. These observations were supported by the statistical analysis. For this and the following experiment, the decision-wise error rate ( $\alpha$ ) was set at .05 level for each orthogonal contrast using the procedure described by Hays (1972). The error rate ( $\alpha$ ) was similarly set at .05 for simple effects. There were no overall differences between groups in lever press responding during the  $S^D$  presentations (all  $F_s(1, 23) < 1, p > .05$ ). There was a main effect of lever,  $F(1, 23) = 128.36, p < .05$ , suggesting that rats responded more on the reinforced lever than the nonreinforced lever (as guided by each  $S^D$ ) averaged across group. This main effect did not interact with any of the group main effects (all  $F_s(1, 23) < 4, p > .05$ ). Simple effects analysis supported this as rats in the Sham group,  $F(1, 23) = 51.77, p < .05$ , pDMS group,  $F(1, 23) = 56.01, p < .05$  and DLS group,  $F(1, 23) = 24.11, p < .05$ , responded more on reinforced lever than the nonreinforced lever (as guided by each  $S^D$ ).

**Stage 2 training**—Responding during Stage 2 acquisition is shown in the top right panel of Figure 4 collapsed across stimulus identity. Inspection of the figure suggests that each group acquired lever-press responding during stimulus presentations and, again, that rats in each group responded similarly and performed more of the reinforced than the

nonreinforced response (as guided by the relevant stimuli). Statistical analyses revealed no overall differences between groups in lever press responding during the  $S^D$  presentations (all  $F$ s (1, 23) < 2,  $p$  > .05). There was a main effect of lever,  $F$  (1, 23) = 95.837,  $p$  < .05, suggesting that rats responded more on the reinforced lever than the nonreinforced lever averaged across group, that did not interact with any of the group main effects (all  $F$ s (1, 23) < 3,  $p$  > .05). Again, simple effects analysis supported this; rats in the Sham group,  $F$  (1, 23) = 38.55,  $p$  < .05, pDMS group,  $F$  (1, 23) = 41.97,  $p$  < .05 and DLS group,  $F$  (1, 23) = 17.95,  $p$  < .05, responded more on reinforced lever than the nonreinforced lever (as guided by each  $S^D$ ).

**Stage 3 training**—Rate of responding during stage 2 is shown in the bottom left panel of Figure 4 collapsed across stimulus identity. Inspection of the figure suggests that there were no differences between groups in lever pressing and, statistically, there were no overall differences between groups in lever press responding during the  $S^D$ 's (all  $F$ s < 1,  $p$  > .05).

**Extinction Test**—The data from the extinction tests are presented in the bottom right panel of Figure 4 collapsed across stimulus identity and across the two tests. From this figure it is apparent that, as predicted by the hierarchical account, rats in both the DLS and Sham groups responded more on the lever associated with the non-devalued outcome than that associated with the devalued outcome, as guided by each of the discriminative stimuli. In contrast, rats in the pDMS group did not appear to show this effect and responded similarly on both levers. These observations were supported by the statistical analysis. As described in the introduction, we tested two contrasts. First, Shams were compared to the DLS group. For this contrast, there was no main effect of group (Sham vs. DLS),  $F$  (1, 23) = 2.78,  $p$  > .05, suggesting that these two groups did not differ in overall responding. There was a main effect of lever,  $F$  (1, 23) = 6.3,  $p$  < .05 suggesting that averaged over group rats responded more on the nondevalued than the devalued lever. However, there was no group  $\times$  lever interaction,  $F$  (1, 23) = .03,  $p$  > .05, suggesting that both the Sham and DLS groups responded more on the nondevalued than the devalued levers.

The second orthogonal contrast compared performance in the sham and the DLS group against that in the pDMS group. This analysis found no main effect of Sham/DLS vs. pDMS,  $F$  (1, 23) = 3.21,  $p$  > .05, indicating that the Sham/DLS groups (averaged) and the pDMS group displayed, similar rates of overall responding. This main effect did, however interact with lever. That is, there was a DLS/Sham vs. pDMS Group  $\times$  Lever interaction,  $F$  (1, 23) = 4.622,  $p$  < .05, suggesting that only the Sham and DLS groups responded more on the devalued than the nondevalued lever, whereas the pDMS group responded similarly on both levers. This was supported by simple effects analysis which revealed that, although both the Sham group,  $F$  (1, 23) = 5.96,  $p$  < .05, and the DLS group,  $F$  (1, 23) = 6.62,  $p$  < .05 responded more on the nondevalued lever than the devalued lever, the pDMS group did not,  $F$  (1, 23) = 0.13,  $p$  > .05.

The results show that both the Sham and DLS groups successfully solved this instrumental biconditional discrimination task, responding more on the nondevalued lever than the devalued lever as guided by each stimulus. However, the pDMS group demonstrated



impaired responding relative to both the DLS and Sham groups and responded similarly on both levers regardless of which stimulus was presented.

These findings are not consistent with the use of S-R associations to solve this task; DLS lesioned rats demonstrated a similar profile of learning to the sham group. Rather, these results are consistent with the use of hierarchical S-(R-O) associations. The impairment in the pDMS group suggests that R-O associations were required to solve this task, as would be predicted if S-(R-O), or S-R-O associations were employed. However, the lack of impairment in the DLS group shows that S-R associations do not contribute positively to task performance, and this is uniquely consistent with the use of hierarchical S-(R-O) rather than the alternatives.

It was also anticipated that, if S-R and S-(R-O) associations compete, impairing the rats' capacity to form S-R associations using DLS lesions should have facilitated performance on this task. However, no such facilitation was observed and the DLS lesioned rats performed similarly to Shams. This raises the possibility that binary associations either did not form or did not influence performance, suggesting that hierarchical associations are the only influence on instrumental performance. On this position, any evidence for the influence of binary associations may simply reflect bias in the influence of one or other aspect of the S-(R-O) association on performance. However, another more plausible possibility is that, because of the complexity of the task and the small effect sizes typical of these types of tasks, it was not possible to observe an effect in the DLS group that was any larger than that displayed by the Shams (i.e. effect size was at ceiling). If this is true then perhaps a more difficult version of the current task would allow the facilitated performance in the DLS group to be revealed. This was assessed in Experiment 2.

## Experiment 2

This experiment examined whether facilitation of learning could be observed in DLS lesioned rats relative to Shams when a more difficult version of the instrumental biconditional discrimination task was employed. The design of this experiment is shown in Table 1. The design was similar to that of Experiment 1 except that Stages 1 and 2 of discrimination training were omitted. As such rats were required to learn each discrimination concurrently; specifically, rats received only the biconditional discrimination training stage employed in Stage 3 of Experiment 1 during which the relations S1: (R1 - O1), (R2 - O2) and S2: (R1 - O2), (R2 - O1) were trained. Test procedures were identical to those used in Experiment 1; prior to test, one of the outcomes was devalued by prefeeding it to satiety. Rats were then given a choice extinction test during which each S was presented and pressing on each lever recorded. Rats that had learned the biconditional discrimination task should choose the lever associated with the nondevalued outcome, the identity of which was dependent on which S was presented. For example, if O1 was devalued, then rats should perform  $R2 > R1$  in the presence of S1, but  $R1 > R2$  in the presence of S2. Likewise, if O2 were devalued then rats should perform  $R1 > R2$  in the presence of S1 but  $R2 > R1$  in the presence of S2.

It was expected that pDMS lesioned rats would show the deficit observed in Experiment 1 and that their biconditional discrimination performance on the levers would not differ during the devaluation test. Similarly (although in contrast to Experiment 1) it was expected that Sham lesioned rats would find this version of the instrumental biconditional discrimination task more challenging and that their differential performance in the devaluation test would be diminished, even relative to the levels observed in Experiment 1. Hence, we anticipated that, in a contrast of performance between sham and pDMS lesions groups, that there would be no essential difference in performance: both groups should lack the necessary associative structure to solve the discrimination. If, however, the associations used to solve this task are truly hierarchical, then removing interference from S-R associations by lesioning the DLS should enable DLS lesioned rats to have a relative advantage to the sham group in solving this more difficult task. As such, we predicted that a contrast comparing performance on the pDMS and sham groups against the DLS groups would reveal a significant interaction based on more substantial discrimination performance in the latter but not the former groups.

## Method

**Subjects and Apparatus**—Twenty-eight experimentally naïve male Long-Evans rats were the subjects. They were obtained from the same source and maintained under the same conditions as the rats in Experiment 1. All same apparatus was used and was as described for Experiment 1.

**Surgical Procedures and Histology**—All surgical procedures and histology were as described for Experiment 1.

### Procedure

**Magazine training:** The procedure for magazine training was identical to that described for Experiment 1.

**Lever Training:** The procedure for lever training was identical to that described for Experiment 1.

**Biconditional Discrimination Training:** Biconditional discrimination training took place over the next 11 days. This Stage was identical to Stage 3 in Experiment 1, with the exception that it was conducted over 11 rather than 6 days.

**Devaluation extinction tests:** The procedure for the devaluation extinction tests was identical to Experiment 1.

## Results and Discussion

**Histology**—Histological analysis determined the size and placement of lesions. Figure 5 shows a representation of the minimal (black) and the maximal (grey) extent of the lesions based on the stereotaxic atlas of the rat brain by Paxinos and Watson (1998). pDMS lesions targeted the most caudal extent of the pDMS and produced substantial cell loss and shrinkage in this region. DLS lesions targeted the mid anterior-posterior point of the dorsal striatum and produced substantial cell loss in this region. A total of 4 animals were excluded

from the experiment because of incorrect lesion placement or size. Thus 24 animals were included in the analysis (Sham,  $n = 8$ , pDMS,  $n = 8$ , DLS,  $n = 8$ ).

**Lever press acquisition**—The mean rate of responding during acquisition is shown in the left-hand panel of Figure 6 collapsed across stimulus identity. Inspection of the figure suggests that there were no differences between groups in the acquisition of lever pressing. These observations were supported by statistical analysis. There were no differences between groups in lever press responding during S presentation (all  $F_s(1, 21) < 1$ ,  $p > .05$ ).

**Extinction Test**—The mean rate of responding during the extinction test is presented in the right-hand panel of Figure 6 collapsed across stimulus identity and across the two tests. From this figure it is apparent that neither rats in the Sham nor the pDMS group responded according to the current value of the outcome as guided by each stimulus. That is, the Sham and pDMS groups responded similarly on the levers associated with the devalued and nondevalued outcomes. In contrast, rats in the DLS group responded more on the lever associated with the nondevalued outcome than the lever associated with the devalued outcome, as guided by each stimulus.

These observations were supported by the statistical analysis. As established in the introduction, we first compared responding in the Sham and pDMS groups. These groups did not differ from each other in overall responding,  $F(1, 21) = .39$ ,  $p > .05$  nor was there a main effect of lever,  $F(1, 21) = .38$ ,  $p > .05$ , suggesting that, averaged over group, rats did not differentially respond on the devalued and nondevalued levers. There was no group (Sham vs. pDMS)  $\times$  lever interaction,  $F(1, 21) = .44$ ,  $p > .05$  suggesting that rats in both groups responded equally on both levers. Next we compared these groups against the DLS lesioned animals to contrast performance in the Sham+pDMS vs. DLS groups. There was no main effect of Sham/pDMS vs. DLS,  $F(1, 21) = .21$ ,  $p > .05$ , indicating that the Sham/pDMS groups (averaged) and the DLS group displayed similar rates of overall responding. There was, however, a Sham/pDMS vs. DLS  $\times$  lever interaction,  $F(1, 21) = 5.89$ ,  $P < .05$ , suggesting that although the Sham and pDMS groups did not respond differentially on the devalued vs. nondevalued levers, the DLS group did. This is supported by simple effects analysis which showed that the DLS group responded more on the nondevalued lever than the devalued lever (as guided by the relevant stimuli),  $F(1, 21) = 5.46$ ,  $p < .05$ , whereas the Sham,  $F(1, 21) = 1.21$ ,  $p > .05$ , and pDMS groups,  $F(1, 21) = .87$ ,  $p > .05$ , did not.

The test performance in the Sham and pDMS groups suggested that they could not solve the biconditional discrimination task. By contrast, the DLS group preferentially chose the lever associated with the nondevalued outcome on test ( $R1 > R2$  and  $R2 > R1$ , dependant on the identity of S). These results are consistent with the prediction that the DLS group would demonstrate enhanced performance relative to the Sham and pDMS groups based on the suggestion that these lesions removed the S-R associations that would normally interfere with performance generated by the hierarchical S-(R-O) associations. Free from this interference, the DLS group were able to form these associations more readily than either the Sham group, that could form both R-O and S-R associations, or the pDMS group that could form S-R but not R-O associations. Thus, retaining the ability to form S-R

associations appears to impair the performance of both Sham and pDMS groups on this task whereas their absence appears to enhance performance.

## General Discussion

The experiments reported here provide evidence that a hierarchical S-(R-O) associative structure governs performance during an instrumental biconditional discrimination task and rules out control by a variety of alternative structures that rely on the formation of SR associations. In Experiment 1 rats with both Sham and DLS lesions successfully completed the task ( $R1 > R2$  and  $R2 > R1$  depending on the identity of S), but rats with pDMS lesions did not ( $R1 = R2$  regardless of which S was presented). This suggests a critical role for R-O but not S-R associations, consistent with the use of hierarchical S-(R-O), but not (S-R)-O or ternary S-R-O associations. When a more difficult version of the same task was employed in Experiment 2 DLS lesioned rats could ( $R1 > R2$  and  $R2 > R1$  depending on the identity of S) but Sham and pDMS lesioned rats could not ( $R1 = R2$  regardless of S) successfully complete the task. In other words, DLS lesions facilitated discrimination performance relative to Shams. This result provides further evidence against alternatives that require the formation of S-R associations to control performance. Had rats relied on this latter form of association then we should have expected to observe a deficit in the performance of DLS lesioned rats. In contrast, DLS rats performed better than Sham rats suggesting that they formed hierarchical associations free from interference created by concurrent S-R associations, whereas the performance of Sham rats suffered as a consequence of competition between these two types of association

Although competition between S-R and hierarchical alternatives provides a reasonable explanation of our data, it does not address why DLS lesions facilitated performance relative to Shams in Experiment 2 but not Experiment 1. One suggestion is that the differential discrimination training in Stage I and its reversal in Stage II of Experiment 1 discouraged or weakened any S-R associations formed by the Sham rats. If, as a consequence, S-R associations didn't contribute to performance in that group, then removing those associations by lesioning the DLS wouldn't be expected to have much of an effect. It should also be noted that the small effect sizes observed here, as in nearly every biconditional instrumental task reported (e.g. Colwill & Delameter, 1995; Colwill & Rescorla, 1990; Declercq & De Houwer, 2009; Rescorla, 1990), make it likely that any facilitation that may have arisen in the DLS group would have been difficult to observe, particularly if performance in both groups was already at ceiling. Indeed, Rescorla (1991) noted that these types of effects are usually small and incomplete, something that he attributed to the complexity of the task and the potential for generalisation. Current results are consistent with this, and further suggest that when the task is particularly complex, incomplete effects are also a result of competition between S-R and hierarchical S-(R-O) associations. Of course, it is possible that competition from other types of associations (e.g. S-O) could also contribute to the small effects sizes in these kind of tasks, and this is something this might warrant further experimental scrutiny.

Although the current results argue against explanations that rely upon intact S-R associations, such as hierarchical (S-R)-O or ternary S-R-O associations, their implications for configural learning theories has not yet been considered. Although Rescorla (1991)

presented a convincing argument against a configural account by manipulating various transfer effects, he did not tackle configural explanations of biconditional discrimination performance. Most configural theories have been developed to explain Pavlovian data, so to understand how they might explain biconditional instrumental discrimination performance it must be assumed that the outcome takes the place of the US. Following from that, animals must then construct configural units between stimuli and responses that enter into separate associations with the two outcomes. In other words, animals could solve this task by forming 4 separate configural associations: S1R1-O1, S1R2-O2, S2R1-O2, and S2R2-O1.

The effects of both pDMS and DLS lesions, however, argue against this kind of account, at least as envisioned by Pearce (1994). First, if DLS lesions suppress the S-R connections that underlie the formation of an SR configural unit, as they should on a configural explanation of their effects, then a configural account should predict impaired performance in the DLS group. However, current results found that lesioning the DLS facilitated performance. Second, a configural account also appears to predict that pDMS lesions will facilitate discrimination. In his explanation of feature negative discrimination learning (A+, AB-) Pearce states that, when presented, A excites its own configural unit (i.e. the configural unit for A) via its association with the US. However, he suggests it will also “fractionally arouse the AB configural unit” that has an inhibitory link with the US such that “trials with A will result in simultaneous excitatory and inhibitory links with the US” (pp 602). If in our study ‘R’ is assumed to replace ‘A’, then it will have simultaneous links with each US: O1 and O2 through the S1R1-O1 and S2R1-O2 associations. Each time R1 is presented, for example, it will excite the current configural unit (e.g. S1R1, as well as the configural unit for R), but will also fractionally arouse the competing SR configural unit containing R1 (e.g. S2R1), impairing discrimination. Therefore, as the pDMS group are unable to form R1-O1 and R2-O2 associations they should not suffer from this competition and discrimination should be facilitated for this group. This was not the observed result.

A separate possibility is that, rather than eradicating R-O associations, pDMS lesions simply cause a deficit in the animals’ ability to differentiate between the stimuli, responses, and/or outcomes used in the current tasks. However, pDMS lesioned rats were clearly able to discriminate in Stages I and II of Experiment 1 in which the outcome followed only one of the Rs during each S presentation. Alternatively, the deficit in performance for rats with pDMS lesions may have been due to an inability to encode the reduced value of the outcome on test. However, past results have shown that rats with pDMS lesions preferentially choose responses associated with the nondevalued outcome during a ‘rewarded’ test suggesting that they no longer desire the outcome (Yin et. al., 2005). Thus, the present results convincingly demonstrate that the deficit in performance caused by pDMS lesions was based on an inability to recall the R-O relation and could not form the S-(R-O) associations necessary to solve the discrimination.

Together, the results of Experiments 1 and 2 strongly argue in favour of a hierarchical S: (R-O) account and, when combined with other evidence (e.g. Rescorla, 1991), it might be concluded that hierarchical associations are ubiquitous throughout instrumental tasks that are too complex to be solved through the use of simple binary associations. There is, however, one result that requires some refinement of this hierarchical account: the case of

congruent/incongruent discriminations developed using the instrumental outcome as both a reward and a discriminandum. For example, in a congruent case, O1 signals that R1 will be followed by O1 and R2 by nothing (similarly O2 signals that R2 will be followed by O2 and R1 by nothing). Thus the rats should form two congruent hierarchical associations O1: R1-O1 and O2: R2-O2. In the incongruent case, however, O1 could signal that R1 will be followed by O2 and R2 by nothing, and O2 that R2 will be followed by O1 and R1 by nothing (O1: R1-O2, O2: R2-O1). A hierarchical account predicts that the congruent discrimination should be easy to solve because, for example, O1 lowers the threshold for R1-O1 responding, promoting R1 > R2 responding and similarly for O2. In the incongruent case, however, delivery of O1 (for example) signals the R1-O2 relation, promoting activation of O2, but activating O2 in this manner should also signal the R2-O1 relation because of the O2: R2-O1 association. Thus O1 presentations should encourage responding on both R1 and R2, making this discrimination more difficult to solve. In contrast to this prediction, it has been found that rats more readily solve the incongruent than the congruent discrimination (Dickinson & De Wit, 2003; Balleine & Ostlund, 2007) clearly arguing against the hierarchical account. Although it is unclear how this result can be directly reconciled with the current findings, at the very least it suggests that the instrumental outcome is represented differently when it serves as a discriminandum and when it acts to reward an instrumental action and that retrieval of the outcome as a reward does not activate its representation as a discriminandum and vice versa.

In conclusion, the current results demonstrate that animals show a bias towards the formation of hierarchical S-(R-O) associations over other types of associations in complex instrumental tasks. Using lesions to examine the relative contribution of S-R and R-O binary associations to biconditional discrimination learning, we have shown for the first time that the structure of the associations required to accurately perform such a task are exclusively hierarchical.

## Acknowledgements

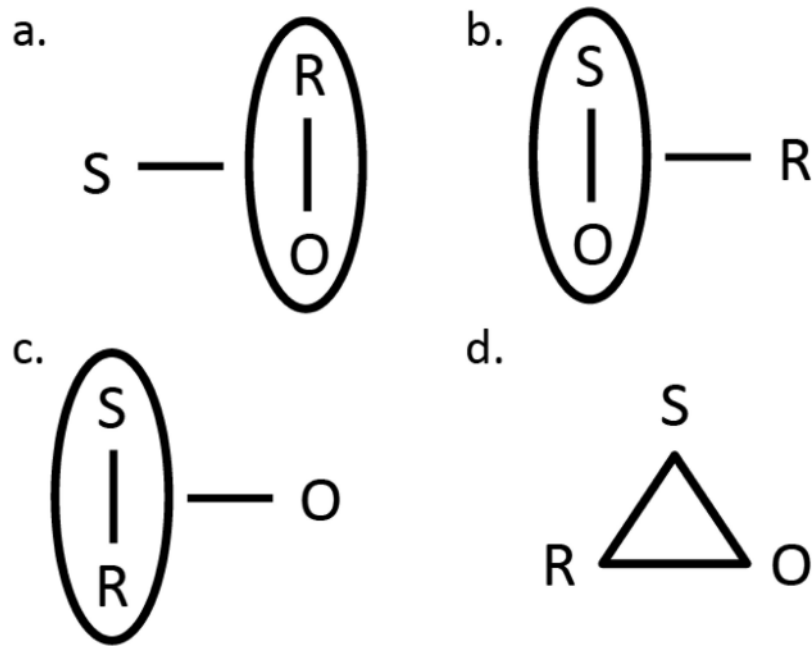
The research reported in this paper was supported by grants from the National Institute of Mental Health, #MH56446, the National Health and Medical Research Council, #633267 and a Laureate Fellowship from the Australian Research Council, FL0992409. We thank Lillian Ashmore for technical support.

## References

- Balleine B, Liljeholm M, Ostlund S. The integrative function of the basal ganglia in instrumental conditioning. *Behavioural Brain Research*. 2009; 199:43–52. doi: 10.1016/j.bbr.2008.10.034. [PubMed: 19027797]
- Balleine B, O'Doherty J. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology*. 2010; 35:48–69. doi: 10.1038/npp.2009.131. [PubMed: 19776734]
- Balleine BW, Ostlund SB. Still at the choice-point. *Action selection and initiation in instrumental conditioning*. *Annals of the New York Academy of Science*. 2007; 1104:147–171.
- Bolles RC. Reinforcement, expectancy, and learning. *Psychological Review*. 1972; 79:394–409. doi: 10.1037/h0033120.
- Colwill RM, Delamater BA. An associative analysis of instrumental biconditional discrimination learning. *Animal Learning and Behavior*. 1995; 23(2):218–233. doi: 10.3758/BF03199937.

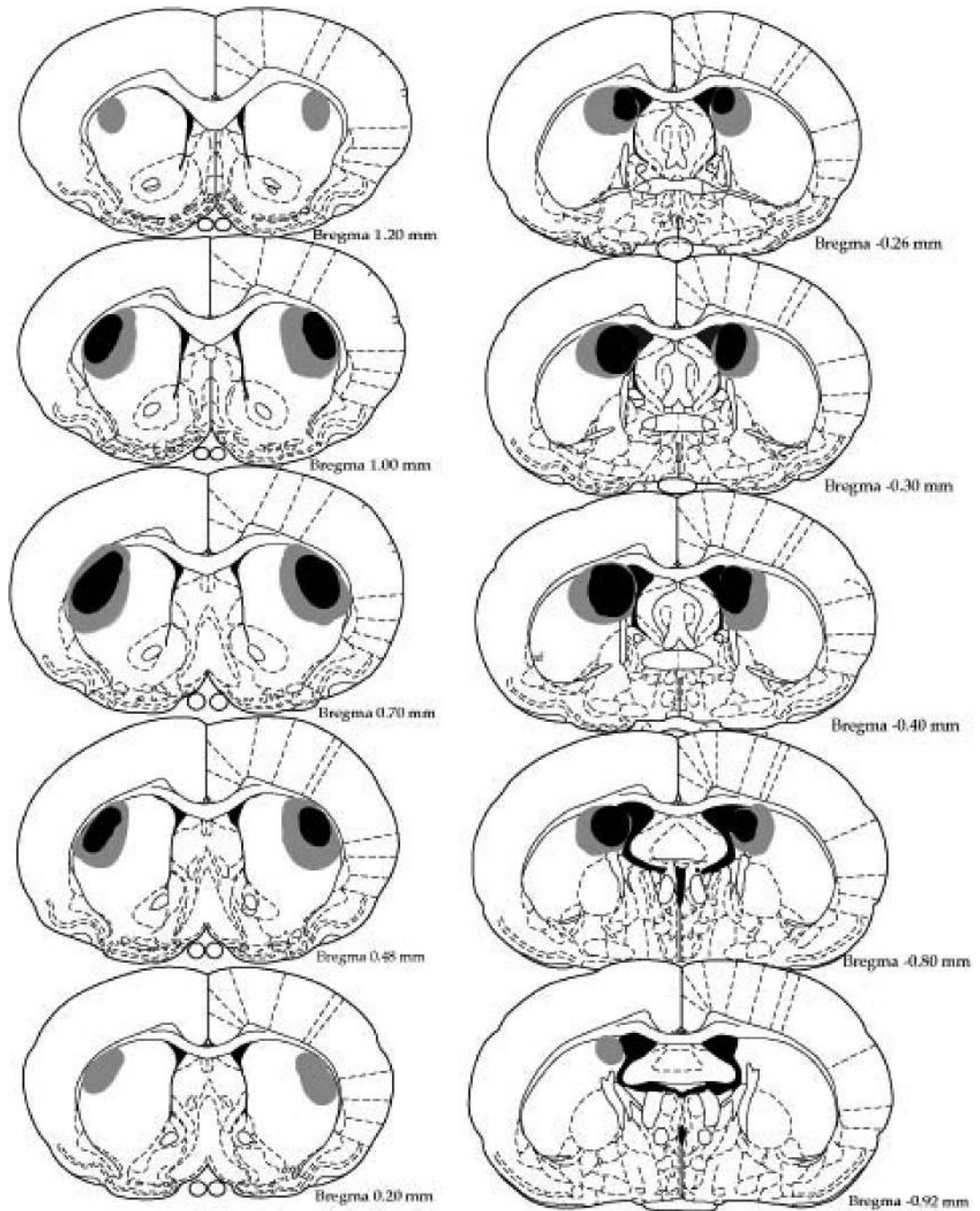


- Colwill, RM.; Rescorla, RA. Associative structures in instrumental learning.. In: bower, GH., editor. *The Psychology of Learning and Motivation*. Vol. 20. Academic Press; Orlando FL: 1986. p. 55-104.
- Colwill RM, Rescorla RA. Evidence for the hierarchical structure of instrumental learning. *Animal Learning and Behavior*. 1990; 18(1):71–82. doi: 10.3758/BF03205241.
- Declercq M, de Houwer J. Evidence for a hierarchical structure underlying avoidance behavior. *Journal of Experimental Psychology : Animal Behavior Processes*. 2009; 35(1):123–128. doi: 10.1037/a0012927. [PubMed: 19159168]
- Dezfouli A, Balleine BW. Habits, action sequences and reinforcement learning. *Eur J Neurosci*. 2012; 35(7):1036–1051. doi: 10.1111/j.1460-9568.2012.08050.x. [PubMed: 22487034]
- Dickinson A. Actions and habits: the development of behavioural autonomy. *Philosophical Transactions of the Royal Society of London*. 1985; B308:67–78. doi: 10.1098/rstb.1985.0010.
- Dickinson, A. Instrumental conditioning.. In: Mackintosh, NJ., editor. *Animal Cognition and Learning*. Academic Press; London: 1994. p. 4-79.
- Dickinson A, De Wit S. The interaction between discriminative stimuli and outcomes during instrumental learning. *Quarterly Journal of Experimental Psychology B- Comparative and Physiological Psychology*. 2003; 56(1):127–139. doi: 10.1080/02724990244000223. [PubMed: 12623543]
- McDonald RJ, White NM. A triple dissociation of memory systems: hippocampus, amygdala, and dorsal striatum. *Behavioral Neuroscience*. 1993; 107(3-22) doi: 10.1037/0735-7044.107.1.3.
- Paxinos, G.; Watson, C. 4th ed.. Academic Press; San Diego: 1998. *The rat brain in stereotaxic coordinates*.
- Pearce JM. Similarity and discrimination: a selective review and connectionist model. *Psychological Review*. 1994; 101(4):587–607. doi: 10.1037//0033-295X.101.4.587. [PubMed: 7984708]
- Rescorla RA. The role of information about the response outcome relation in instrumental discrimination learning. *Journal of Experimental Psychology: Animal Behavior Processes*. 1990; 16(3):262–270. doi: 10.1037/0097-7403.16.3.262.
- Rescorla RA. Associative relations in instrumental learning: The Eighteenth Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology*. 1991; 43:1–23.
- Rescorla RA. Response-outcome versus outcome-response associations in instrumental learning. *Animal Learning & Behavior*. 1992; 20:223–232. doi: 10.3758/BF03213376.
- Shiflett MW, Brown RA, Balleine BW. Acquisition and performance of goal-directed instrumental actions depends on ERK signaling in distinct regions of dorsal striatum in rats. *Journal of Neuroscience*. 2010; 30(8):2951–2959. doi: 10.1523/JNEUROSCI.1778-09.2010. [PubMed: 20181592]
- Trapold MA. Are expectancies based upon different positive reinforcing events discriminably different? *Learning and Motivation*. 1970; 1:129–140. doi: 10.1016/0023-9690(70)90079-2.
- Trapold, MA.; Overmier, JB. The second learning process in instrumental conditioning.. In: Black, AA.; Prokasy, WF., editors. *Classical Conditioning: II. Current research and theory*. Appleton-Century-Crofts; New York: 1972. p. 427-452.
- Yin HH, Knowlton BJ, Balleine BW. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur Journal of Neuroscience*. 2004; 19(1): 181–189. doi: 10.1111/j.1460-9568.2004.03095.x.
- Yin HH, Knowlton BJ, Balleine BW. Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. *Behavioral Brain Research*. 2006; 166:189–196. doi: 10.1016/j.bbr.2005.07.012.
- Yin HH, Ostlund SB, Knowlton BJ, Balleine BW. The role of the dorsomedial striatum in instrumental conditioning. *European Journal of Neuroscience*. 2005; 22(2):513–523. doi: 10.1111/j.1460-9568.2005.04218.x. [PubMed: 16045504]

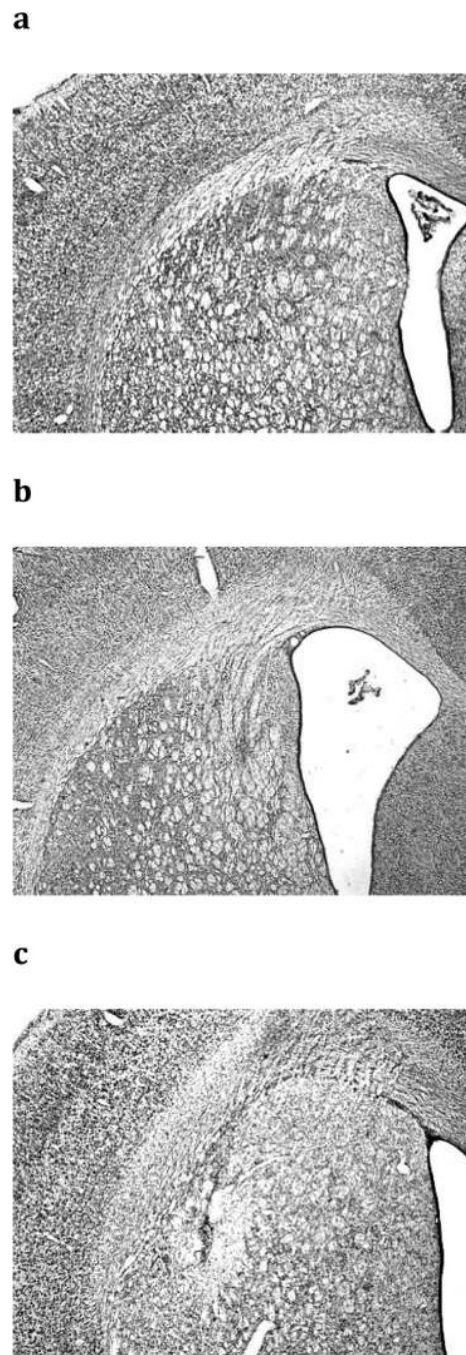


**Figure 1.**

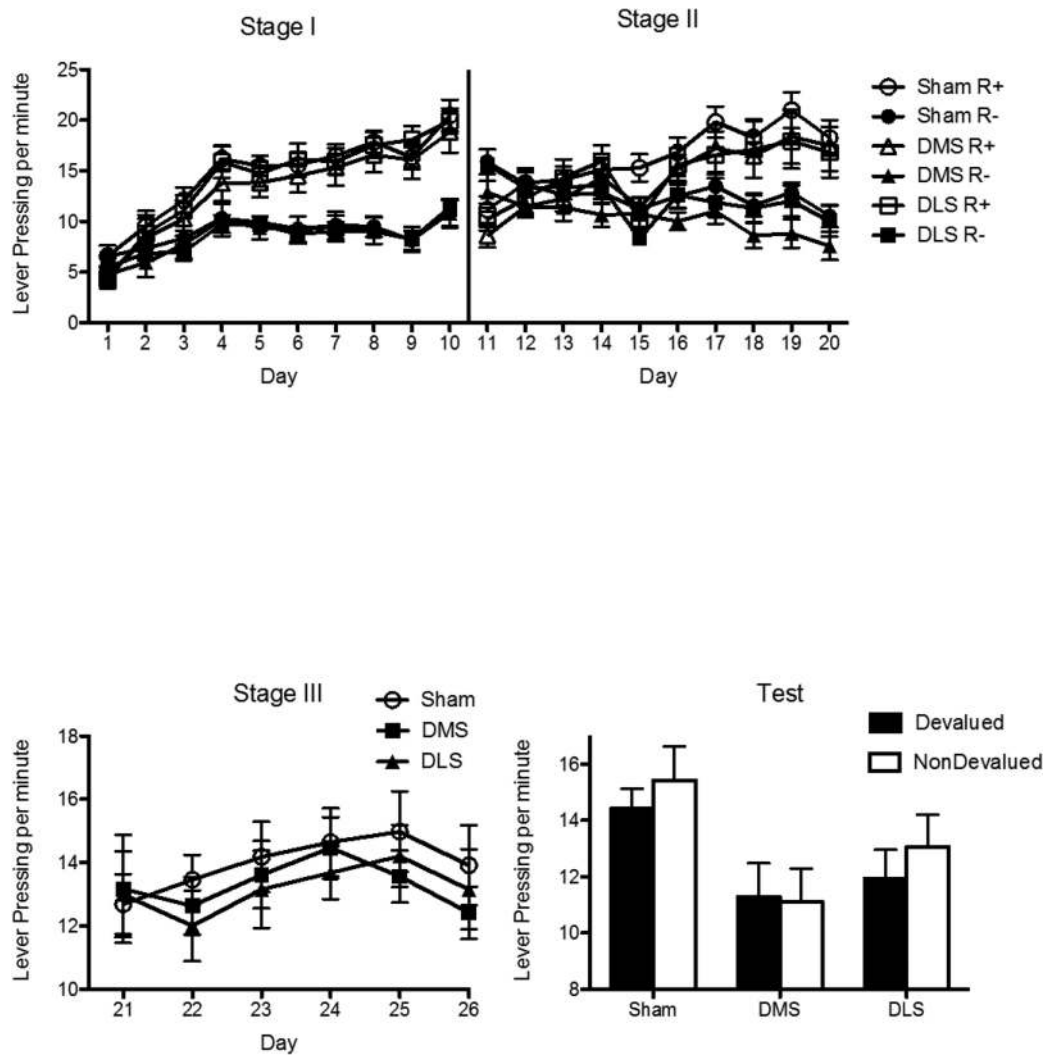
Alternative associative structures mediating biconditional discrimination. (a) The hierarchical structure favoured by Rescorla (1991) allowing S to control specific R-O associations. (b) The two-process alternative allowing the performance of R to be controlled by a specific S-O association and hence by the value of the O associated with S. (c) A hierarchical alternative with the performance based on a specific S-R association modulated by O and so controlled by the value of that specific O. (d) A non-hierarchical, ternary associative structure with the three terms together controlling R with such control reduced by the devaluation of O. S, stimulus, R, response, O, outcome.



**Figure 2.** Minimal (black) and maximal (gray) extents of NMDA-induced excitotoxic DLS (left) and pDMS (right) in Experiment 1. DLS, dorsolateral striatum, pDMS posterior dorsomedial striatum.



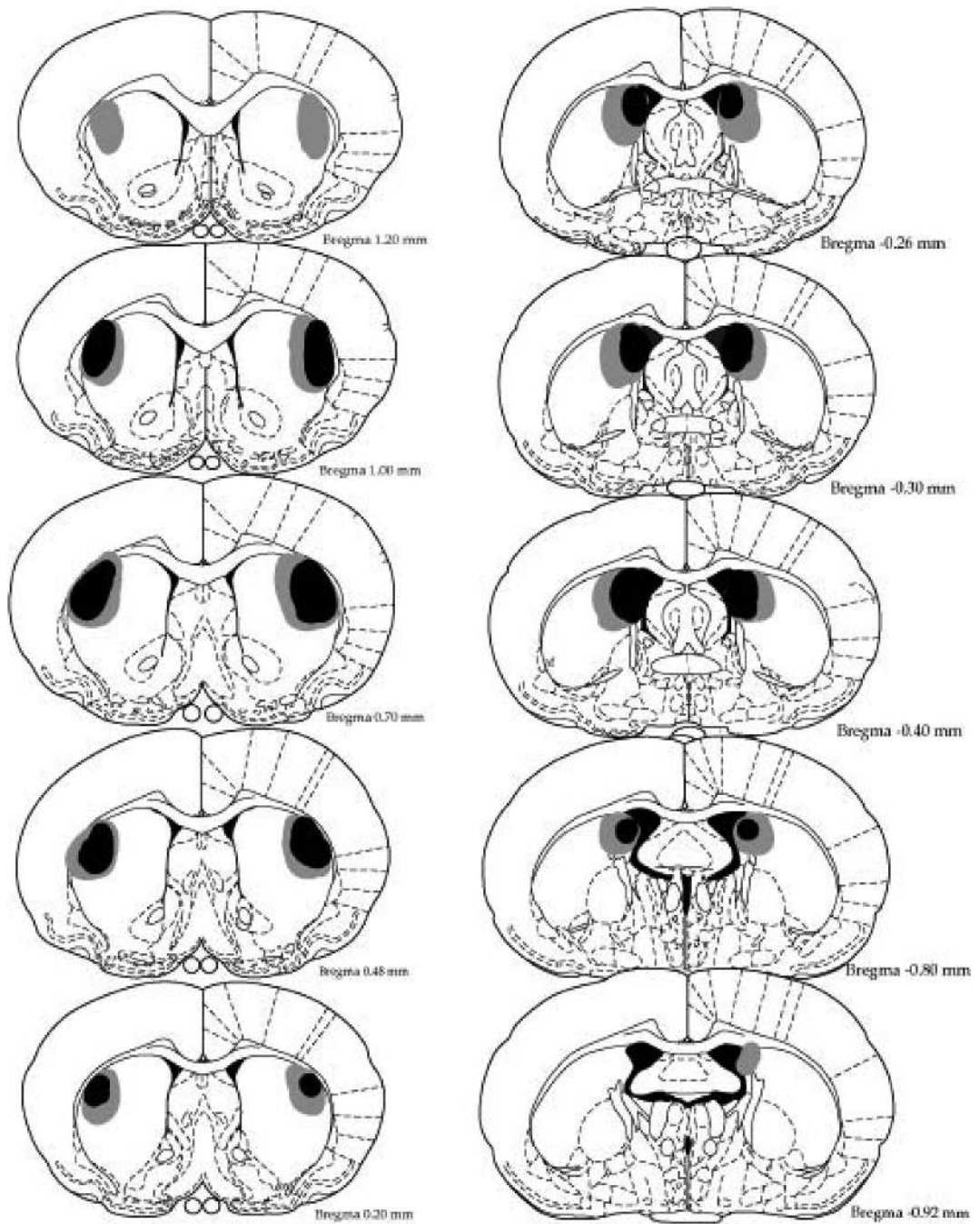
**Figure 3.**  
Photomicrographs of coronal sections showing a) Sham, b) DMS, and c) DLS lesions.  
pDMS, posterior dorsomedial striatum, DLS, dorsolateral striatum.



**Figure 4.**

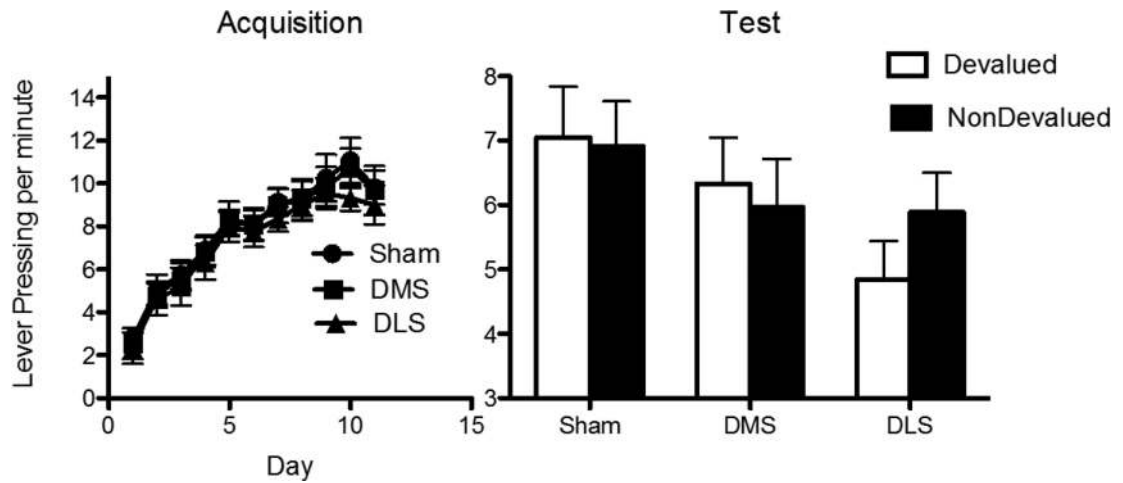
Mean rate of lever pressing ( $\pm 1$  SEM) during each stage of acquisition and on test during Experiment 1. The top two panels show that the Sham, pDMS and DLS groups all responded more on the reinforced than the nonreinforced lever, as guided appropriately by each stimulus, during Stages I and II of acquisition. The bottom left panel shows responding in the Sham, pDMS and DLS groups averaged over both levers during Stage III of acquisition (Both levers were reinforced at this stage). Groups did not differ from each other during acquisition. The bottom right panel shows responding on test. Both the Sham and DLS groups demonstrated evidence of having learned the instrumental biconditional discrimination task: NonDevalued > Devalued (as guided appropriately by each stimulus). By contrast, task performance was impaired for the pDMS group: NonDevalued = Devalued.





**Figure 5.** Minimal (black) and maximal (gray) extents of NMDA-induced excitotoxic DLS (left) and pDMS (right) in Experiment 2. DLS, dorsolateral striatum, pDMS, dorsolateral striatum.





**Figure 6.**

Mean rate of lever pressing ( $\pm 1$  SEM) during acquisition and test of Experiment 2. The left panel shows lever pressing during acquisition averaged over both levers. Groups did not differ on acquisition. The right panel shows lever press responding on test. The DLS group demonstrated evidence of having learned the instrumental biconditional discrimination task: NonDevalued > Devalued (as guided appropriately by each stimulus). Performance on this task was impaired for the Sham and pDMS groups: NonDevalued = Devalued.

**Table 1**

## Experiment 1

Stage I	Stage II	Stage III	Devalue	Test
S1: R1-O1, R2-	S1: R2- O2, R1-	S1: R1 - O1, R2 - O2	O1 or O2	S1: R1 vs. R2
S2: R2-O1, R1-	S2: R1-O2, R2-	S2: R1 - O2, R2 - O1		S2: R1 vs. R2

Design of Experiments 1 and 2. Experiment 1 incorporated each stage, Experiment 2 incorporated Stage III, Devalue, and Test stages only. R1 and R2 were the instrumental responses left lever press and right lever press, S1 and S2 were the discriminative stimuli tone and clicker, O1 and O2 were the two outcomes, sucrose and pellets.