

Published in final edited form as:

Eur J Neurosci. 2013 March ; 37(6): 1012–1021. doi:10.1111/ejn.12106.

Striatum-dependent habits are insensitive to both increases and decreases in reinforcer value in mice

Jennifer J. Quinn^{1,*}, Christopher Pittenger², Anni S. Lee², Jamie L. Pierson¹, and Jane R. Taylor²

¹Department of Psychology, Miami University, 90 N. Patterson Ave., Oxford, OH 45056

²Division of Molecular Psychiatry, Yale University, CMHC Ribicoff Research Facilities, 34 Park Street, New Haven, CT 06508

Abstract

The mouse has emerged as an advantageous species for studying the brain circuitry that underlies complex behavior and for modeling neuropsychiatric disease. The transition from flexible, goal-directed actions to inflexible, habitual responses is argued to be a valid and reliable behavioral model for studying a core aspect of corticostriatal systems that is implicated in certain forms of psychopathology. This transition is thought to correspond to a progression of behavioral control from associative to sensorimotor cortico-basal ganglia networks. Habits form following extensive training and are characterized by reduced sensitivity of instrumental responding to reinforcer revaluation; few studies have examined this form of behavioral control in mice. Here we examine the involvement of the dorsolateral and dorsomedial striatum in this transition in the C57BL/6 inbred mouse strain. We provide evidence that damage to the dorsolateral striatum disrupts habitual responding – that is, it preserves sensitivity to changes in outcome value following either outcome *devaluation* or, for the first time in mice, outcome *inflation*. Together, these data show that instrumental responding in lesioned mice tracks the current value of a reinforcer and provide evidence that neuroanatomical mechanisms underlying habit learning in rats are preserved in mouse. This will allow for genetic and molecular dissection of neural factors involved in decision-making and mechanisms of aberrant habit formation.

Keywords

instrumental; learning; goal-directed; habitual; devaluation

Introduction

In an instrumental conditioning setting, an organism must perform a behavior in order to receive a reinforcer. In such a setting, a number of independent associations are formed between the behavior, the reinforcer, and various environmental stimuli (e.g., Colwill & Rescorla, 1988; Rescorla, 1992). During the initial acquisition of an instrumental action, such as a mouse nose-poking into a recessed aperture, performance is controlled by an expectation of the consequences of that action – that is, an action-outcome (A-O) association is formed. Consequently, a spontaneous change in performance occurs when the value of the outcome is altered following instrumental training – after outcome devaluation or inflation (Adams & Dickinson, 1981; Balleine & Dickinson, 1998). Following extensive training, however, such control by outcome value is diminished and performance shifts to depend

*Correspondence: Dr. Jennifer J. Quinn, Department of Psychology, 90 N. Patterson Avenue, Oxford, OH 45056, Fax: 513.529.2420, quinnjj@muohio.edu.

upon an association between antecedent environmental stimuli and the response – that is, a stimulus-response (S-R) association is formed. Accordingly, as training progresses, instrumental performance becomes increasingly insensitive to outcome devaluation (Balleine & Dickinson, 1998; Holland, 2004; Yin & Knowlton, 2006; Tricomi *et al.*, 2009). This criterion provides an objective means to discriminate between goal-directed actions and stimulus-driven habits.

This balance between goal-directed and habitual behavior depends upon corticostriatal systems (Yin & Knowlton, 2006; Hitchcott *et al.*, 2007; Wickens *et al.*, 2007; Grahm *et al.*, 2008; Hilário & Costa, 2008; Kubota *et al.*, 2009; Pennartz *et al.*, 2009; Balleine & O’Doherty, 2010; Lovinger, 2010; Corbit *et al.*, 2012; Lingawi & Balleine, 2012). Recent analyses suggest functional anatomical segregation within these cortical and striatal regions, with a sensorimotor-dorsolateral network supporting habitual, stimulus-response behaviors and a prefrontal-dorsomedial network mediating flexible, action-outcome behavior (Yin & Knowlton, 2006; Hitchcott *et al.*, 2007; Ashby *et al.*, 2010; Balleine & O’Doherty, 2010). This functional segregation of striatal subregions is supported by behavioral (e.g., Ragozzino *et al.*, 2002; Featherstone & McDonald, 2004; Faure *et al.*, 2005; Yin *et al.*, 2009) and electrophysiological (Carelli *et al.*, 1997; Jog *et al.*, 1999; Barnes *et al.*, 2005; Tang *et al.*, 2007) observations, as well as neural imaging (e.g., Tricomi *et al.*, 2009). Yin and colleagues have addressed this segregation directly using instrumental conditioning procedures in rats. Specifically, they have shown that in rats trained to express an instrumental habit, lesions or inactivations of the dorsolateral striatum restore sensitivity to reinforcer devaluation (Yin *et al.*, 2004) or the action-outcome contingency (Yin *et al.*, 2006), respectively. Conversely, manipulations of the dorsomedial striatum alter the acquisition of goal-directed responding (Yin *et al.*, 2005a; Yin *et al.*, 2005b).

We examined a role for the dorsal striatum in instrumental conditioning in mice. It is important to establish that the basic neuroanatomical mechanisms of habit learning in mice are similar to those in rats, since the molecular technologies suitable for future genetic dissection of habit learning mechanisms are predominantly available in mice (Hilário & Costa, 2008). We use insensitivity to reinforcer revaluation (both devaluation and inflation) to assess habitual responding, probing for the first time the involvement of the striatum in the resistance of habits to outcome inflation.

Materials and Methods

Experimental design

Four experiments were performed to assess the contributions of the dorsal striatum to habitual instrumental responding in mice. In experiments 1 and 2, large pre-training dorsal striatum lesions were used to assess its contribution to habitual responding as measured by sensitivity to either post-training reinforcer devaluation (experiment 1) or post-training reinforcer inflation (experiment 2). Experiments 3 and 4 assessed whether selective damage to the dorsolateral striatum is sufficient to disrupt habitual responding using reinforcer devaluation.

Subjects

C57Bl/6 mice (6–8 weeks of age) were purchased from Jackson Laboratories (Bar Harbor, ME) and housed in groups of 2–3 per cage. Mice were allowed at least one week of acclimation to the animal facility prior to surgery; vivarium rooms were temperature- and humidity-controlled with a 12:12hr light:dark cycle. During acclimation, mice had *ad libitum* access to food and water. All surgical and behavioral procedures were conducted during the light portion of the cycle and were approved by the Yale University IACUC.

Surgery

Mice were anesthetized with an intraperitoneal (i.p.) injection of tribromoethanol (275 mg/kg in saline with 2.5% 2-methyl-2-butanol). The animal's head was shaved, cleaned with 70% ethanol and Betadine, and mounted into a standard stereotaxic instrument. Sterile lubricant was generously applied to the eyes. The scalp was incised and the skin retracted. Bregma and lambda were leveled in the horizontal plane. Bilateral burr holes were drilled through the skull according to the following coordinates, measured from bregma: Large dorsal striatal surgeries in experiment 1 [AP +0.74 mm; ML \pm 2.3 mm; DV -3.5 mm], large dorsal striatal surgeries in experiment 2 [AP +0.74 mm; ML \pm 2.2 mm; DV -3.0 mm], dorsolateral surgeries in experiment 3 [AP +0.5 mm; ML \pm 2.8 mm; DV -3.0 mm], dorsomedial surgeries in experiment 4 [AP +0.74 mm; ML \pm 2.3 mm; DV -3.5 mm]. A 0.5 μ l Hamilton syringe was lowered to the target coordinates and N-methyl-D-aspartate (NMDA; 20 μ g/ μ l; Sigma, St. Louis, MO), dissolved in phosphate buffered saline (PBS), was infused. Sham animals received infusions of PBS alone. For the large dorsal striatum surgeries in experiments 1 and 2, the NMDA or vehicle was infused at a volume of 0.2 μ l per infusion (manually across 1 min). For the more restricted dorsolateral and dorsomedial surgeries in experiments 3 and 4 NMDA or vehicle was infused at a volume of 0.1 μ l per infusion (manually across 1 min). The syringe remained in place for an additional four minutes following each infusion to allow for diffusion of the drug. Following the final infusion, the incision was sutured and a topical antibiotic ointment applied. Mice were allowed to recover from surgery on a heat pad. Upon awakening, they were returned to their homecages and allowed to recover for at least 14 days prior to the start of behavioral testing. All mice were used in a water maze experiment prior to the start of the present behavioral experiments (Lee *et al.*, 2008). Experimental groups were counterbalanced to eliminate any contribution of animals' experience in this watermaze task to the current experiments.

Instrumental Conditioning Chambers

Fourteen similar mouse instrumental chambers (15 cm deep \times 17 cm wide \times 12 cm high), each housed within a melamine sound-attenuating box, were used for these experiments (Med-Associates; Georgia, VT). The side walls of each chamber were made of modular stainless steel panels and the front door, ceiling and back wall were made of clear Plexiglas. The grid floor consisted of 19 stainless steel rods placed in parallel 0.75 cm apart (center-to-center). Each chamber was equipped with a 28V house light located at the top of the middle panel on the left side wall, three adjacent nosepoke apertures located at the bottom of the left side wall, and a food magazine located at the bottom of the middle panel on the right side wall. Each nosepoke aperture was equipped with a light and photobeam sensor. The food magazine, which delivered a single 20 mg pellet (Bioserv, Frenchtown, NJ) when activated, was also equipped with a photobeam sensor. A fan was located on the inside of each sound attenuating box to provide constant background noise and ventilation.

Instrumental Acquisition

At the start of the behavioral experiments, all mice were placed on a food restriction schedule that allowed them 1.5 hours free access to homecage food per day for the duration of the experiment. After five days on the food restriction schedule, instrumental acquisition commenced and consisted of: 1 day magazine training, 2 or 3 days fixed interval 20 sec training (FI20), 2 or 3 days random interval 30 sec training (RI30) and 2 or 5 days random interval 60 sec training (RI60). Each daily training session was 45 min in duration.

Magazine Training—At the start of the magazine training session, the houselight and fan were turned on and remained on until the end of the session. One minute into the magazine training session, a food pellet was delivered into the food magazine. Once the mouse

retrieved the pellet (as measured by a photobeam break inside the food magazine), the clock was reset for one minute, after which another food pellet was delivered. This continued for the entire 45 minute session.

Fixed Interval Training—At the start of each fixed interval (FI) training session, the houselight, fan and lights inside of the three nosepoke apertures were turned on; they remained on throughout the session. Either the left or right nosepoke aperture was designated as the “active” nosepoke for each animal, while the other two nosepokes were designated the “inactive” nosepokes. The active nosepoke assignment (left or right) was balanced across animals in each condition. On the FI20 schedule, the first active nosepoke after 20 seconds had elapsed resulted in the delivery of a food pellet in the magazine. Upon delivery, the 20 sec clock was reset. This continued for the entire 45 min session. Nosepokes to the inactive apertures were never reinforced.

Random Interval Training—During RI training, reinforcement followed the first active nosepoke after a random interval, with an average length of 30 sec (RI30) or 60 sec (RI60), had elapsed. Following each nosepoke, the next random interval was generated automatically and the clock was reset. This continued for the entire 45 min session. Again, inactive nosepokes were never reinforced.

Reinforcer Devaluation by Conditioned Taste Aversion (CTA)

Twenty-four hours following the final instrumental acquisition session in experiments 1, 3 and 4, individual mice were placed into a novel polypropylene tub (28cm × 18cm × 12cm; Allentown Caging, Allentown, NJ) with approximately 7 grams of food pellets (identical to the reinforcer used during instrumental acquisition) distributed on the floor of the tub. The mice were left undisturbed for 15 min, during which they freely consumed food pellets. Following the 15 min, mice were removed and injected with either 0.9% saline (“valued” mice) or 0.15M lithium chloride (“devalued” mice; 40ml/kg i.p.). Mice were returned to their homecage and later allowed their daily 1.5hr of homecage food access beginning approximately 4 hours following the injection. Approximately 4 hours following homecage feeding, “valued” mice received an injection of 0.15M lithium chloride (40ml/kg i.p.) and “devalued” mice received an equal volume injection of 0.9% saline. Thus, all mice had equal exposure to lithium chloride, but “devalued” mice had LiCl paired with access to food pellets while “valued” animals had unpaired presentations. This was performed once daily across two consecutive days. Consumption was measured by weighing the food pellets before and after the consumption period and using the difference as the amount consumed.

Reinforcer Inflation by Extended Food Restriction

The day after the final instrumental acquisition session in experiment 2, mice were assigned to “inflation” or “no inflation” conditions. On this day, mice in the inflation condition did not receive their daily 1.5 hr access to homecage food. Mice in the no inflation condition were fed normally for the 1.5 hr period.

Instrumental Testing

Extinction test—Two days following the last CTA session, or the day after reinforcer inflation, testing was conducted in the conditioning chamber used for instrumental acquisition. The test session was identical to interval training except that no reinforcement was delivered. A 5 min test session length was chosen since no observable extinction occurs over this short period. Active nosepokes during this 5 minute trial were normalized to active nosepokes during the first 5 min of the final acquisition day.

Consumption test—Immediately following the instrumental test, animals received a final 15 min consumption test to assess whether the instrumental extinction test alters responding for the valued or devalued food pellet reinforcer.

CTA transfer test—In experiment 3, mice underwent a final 15 min test session three days after the extinction test. During this test in the instrumental conditioning chamber, all of the nosepoke apertures were blocked and 10 reinforcer pellets were freely available in the food magazine. The number of pellets consumed during this session was assessed for each mouse. This test was conducted in order to assess whether the effects of dorsolateral lesions in this experiment resulted from differential transfer of the conditioned taste aversion from the context in which it took place to the instrumental conditioning context.

Histology

Lesions were documented using both Nissl staining (Cresyl Violet) and immunohistochemistry for GFAP and NeuN, as described previously (Lee et al., 2008); sample lesions documented by immunostaining are shown in Supporting Information online (Supplementary Figure 1). For Nissl staining, fresh frozen brains were sliced on a cryostat at 40 μ m and stained using standard techniques. For immunohistochemical documentation of lesions, brains were rapidly dissected and fixed overnight in 4% paraformaldehyde/PBS at 4°. After fixation, brains were equilibrated with 30% sucrose and sliced on a microtome at 40 μ m. Slices were stored in cryoprotectant solution (30% glycerine, 30% ethylene glycol, 0.2x PBS) at 4°. Floating sections were washed 3 \times 10min with PBS, blocked with PBS/0.3% Triton/2% goat serum (Sigma) for one hour with gentle shaking, and then immunostained overnight for GFAP (Sigma rabbit polyclonal anti-GFAP IgG, G9269, 1:500) and NeuN (Chemicon International mouse monoclonal anti-NeuN IgG, MAB377; 1:1000) in PBS/0.3% Triton. The following day slices were rinsed twice in PBS/0.3% Triton and twice in PBS, stained for one hour with secondary antibodies (FITC Goat anti-rabbit IgG 1:300; rhodamine goat anti-mouse IgG; 1:300) in PBS/0.3% Triton/2% goat serum, washed again 3 \times 10min in PBS/0.3% Triton, and mounted on glass slides. GFAP and NeuN immunoreactivity were visualized on an upright Nikon fluorescent microscope.

Statistical Analysis

For each dependent measure, an initial 2 \times 2 ANOVA (or RM-ANOVA) was used to compare the two independent factors: surgery (sham vs. lesion) and devaluation/inflation (valued vs. devalued/non-inflated vs. inflated). Following a significant interaction or main effect in these overall analyses, Fisher's PLSD tests were used to further evaluate individual group differences.

Results

Lesions of the dorsal striatum preserve sensitivity to post-training reinforcer devaluation (Experiment 1)

Experiment 1 addressed whether large pre-training lesions of the dorsal striatum (including damage to both lateral and medial subregions) altered sensitivity to post-training reinforcer devaluation by conditioned taste aversion in mice.

Histology—Figure 1a provides a schematic representation of the extent of damage to the striatum caused by NMDA infusions in Experiment 1. Analyses revealed significant damage to the dorsal striatum, with the lesion centering in the lateral portion of the striatum, in all mice. Maximal lesions involved some damage to the overlying cortex. There were no cases of thalamic damage or damage to the posterior tail of the striatum. One mouse was excluded from all analyses following examination of the stained tissue revealing damage to the ventral

striatum unilaterally. Group sizes were: sham/valued 8; sham/devalued 9; lesion/valued 7; lesion/devalued 7.

Acquisition—Across the eleven days of acquisition, the number of active nose pokes increased [$F(10,270) = 68.69, p < .0001$] with no effect of lesion condition or the to-be-valued/devalued condition (Supplementary Figure 2a). The number of inactive nose pokes was tenfold lower but also increased over acquisition days [$F(10,270) = 4.22, p < .0001$] with no effect of lesion condition or the to-be-valued/devalued condition (Supplementary Figure 2b). The total number of reinforcers earned increased over acquisition days [$F(10,270) = 37.90, p < .0001$] with no effect of lesion condition or the to-be-valued/devalued condition (Supplementary Figure 2c).

Conditioned Taste Aversion (CTA)—Across the two days of CTA, pellet consumption dramatically decreased in devalued animals but remained stable in the valued groups [$F(1,27) = 24.78, p < .0001$] (Figure 2). This change in consumption across the two CTA days was independent of lesion condition.

Habit and Consumption Tests—Active nose poke responding during the non-reinforced test was converted to a percent of baseline responding during the first five minutes on the last acquisition day. There was a significant lesion X devaluation interaction [$F(1,27) = 6.80, p = .01$]. Planned comparisons using Fisher's PLSD ($p < .05$) showed that devalued animals had fewer active nose poke responses than valued animals in the lesion, but not sham, condition (Figure 3a); thus, the control animals' behavior was driven by a S-R habit, while the lesioned animals' responding remained dependent on reinforcer value. Despite the interaction observed in active nose poke responses, both lesion and sham devalued groups consumed fewer reinforcer pellets than the valued animals in a post-test free consumption session [$F(1,27) = 148.91, p < .0001$] (Figure 3b).

Lesions of the dorsal striatum preserve sensitivity to post-training reinforcer inflation (Experiment 2)

Experiment 2 addressed whether similar lesions altered sensitivity to post-training reinforcer inflation by extended food restriction in mice.

Histology—Figure 1b provides a schematic representation of the extent of damage to the striatum caused by NMDA infusions in Experiment 2. Lesions were more restricted to the dorsolateral striatum compared to those of Experiment 1. All mice showed significant damage to the overlying cortex. There were no cases of ventral striatal damage, thalamic damage, or damage to the posterior tail of the striatum. Group sizes were: sham/non-inflated 9; sham/inflated 9; lesion/non-inflated 8; lesion/inflated 6.

Acquisition—Across the six days of acquisition, the number of active nose pokes increased [$F(5,140) = 61.62, p < .0001$] with no effect of lesion condition or the to-be-inflated/non-inflated condition (Supplementary Figure 2d). The number of inactive nose pokes decreased over acquisition days [$F(5,140) = 4.04, p < .01$] with no effect of lesion condition or the to-be-inflated/non-inflated condition (Supplementary Figure 2e). The total number of reinforcers earned increased over acquisition days [$F(5,140) = 34.84, p < .0001$] with no effect of lesion condition or the to-be-inflated/non-inflated condition (Supplementary Figure 2f).

Extinction and Consumption Tests—Active nose poke responding during the non-reinforced test was converted to a percent of baseline responding during the first five minutes on the last acquisition day. Although the lesion X inflation interaction did not reach

significance ($p = 0.15$), the results of Experiment 1 warrant the following planned orthogonal contrasts. Sham inflated animals did not differ from sham non-inflated animals ($p > 0.1$). Further, sham animals (inflated and non-inflated) did not differ from lesion non-inflated animals ($p > 0.1$). Importantly, lesion inflated animals showed significantly higher active nosepoke responding compared with sham (inflated and non-inflated) plus lesion non-inflated animals ($p = 0.01$). These results confirm preserved sensitivity to increases in reinforcer value following dorsolateral striatum lesions (Figure 3c). Both lesion and sham inflated groups consumed more reinforcer pellets compared with the non-inflated animals [$F(1,28) = 8.57, p < .01$] during the consumption test (Figure 3d), confirming the efficacy of the inflation procedure.

Restricted dorsolateral, but not intermediate, striatum lesions preserve sensitivity to post-training devaluation-induced changes in reinforcer value (Experiments 3 & 4)

Experiments 3 and 4 addressed whether disruption of the most lateral region of the dorsal striatum is sufficient to alter sensitivity to post-training reinforcer devaluation by conditioned taste aversion in mice as seen in Experiment 1.

Histology—Figure 4 provides a schematic representation of the extent of damage to the striatum caused by NMDA infusions in Experiments 3 (4a) and 4 (4b). There were no cases of ventral striatal damage, thalamic damage, or damage to the posterior tail of the striatum in either experiment. In Experiment 3, lesions were restricted to the most lateral portion of the dorsal striatum and were highly uniform across animals. Significant damage was observed in the striatal matter along the edge of the corpus callosum, with no apparent damage to white matter (presumably because NMDA diffusion was constrained by the white matter). In Experiment 4, significant damage was observed in the dorsal striatum, though these lesions spared the most lateral portion. Group sizes for Experiment 3 were: sham/valued 10; sham/devalued 10; lesion/valued 9; lesion/devalued 9. Group sizes for Experiment 4 were: sham/valued 9; sham/devalued 9; lesion/valued 7; lesion/devalued 8.

Acquisition

Restricted dorsolateral lesions (Exp. 3): Across the six days of acquisition, the number of active nosepokes increased [$F(5,170) = 74.89, p < .0001$] with no effect of lesion condition or the to-be-valued/devalued condition (Supplementary Figure 3a). The number of inactive nosepokes was 5–10 fold lower but also increased over acquisition days [$F(5,170) = 7.57, p < .0001$] with no effect of lesion condition or the to-be-valued/devalued condition (Supplementary Figure 3b). The total number of reinforcers earned increased over acquisition days [$F(5,170) = 48.44, p < .0001$] with no effect of lesion condition or the to-be-valued/devalued condition (Supplementary Figure 3c).

Intermediate dorsal lesions (Exp. 4): Across the six days of acquisition, the number of active nosepokes increased [$F(5,145) = 81.48, p < .0001$] with no effect of lesion condition or the to-be-valued/devalued condition (Supplementary Figure 3d). The number of inactive nosepokes was 5–10 fold lower but also increased over acquisition days [$F(5,145) = 3.10, p < .05$] with no effect of lesion condition or the to-be-valued/devalued condition (Supplementary Figure 3e). The total number of reinforcers earned increased over acquisition days [$F(5,145) = 55.72, p < .0001$] with no effect of lesion condition or the to-be-valued/devalued condition (Supplementary Figure 3f).

Conditioned Taste Aversion (CTA)

Restricted dorsolateral lesions (Exp. 3): Across the two days of CTA, pellet consumption decreased in all animals [$F(1,34) = 101.64, p < .0001$]; however, it decreased significantly

more in the devalued animals as indicated by a day X reinforcer value interaction [$F(1,34) = 18.26, p = .0001$] (Figure 5a). This change in consumption across the two CTA days was independent of lesion condition.

Intermediate dorsal lesions (Exp. 4): Across the two days of CTA, pellet consumption decreased in all animals [$F(1,29) = 50.06, p < .0001$]; however, it decreased significantly more in the devalued animals, as indicated by a day X reinforcer value interaction [$F(1,29) = 5.74, p < .05$] (Figure 5b). This change in consumption across the two CTA days was independent of lesion condition.

Habit and Consumption Tests

Restricted dorsolateral lesions (Exp. 3): Active nosepoke responding during the non-reinforced test was converted to a percent of baseline responding during the first five minutes on the last acquisition day. There was a significant lesion X devaluation interaction [$F(1,34) = 4.47, p < .05$]. *A priori* planned comparisons using Fisher's PLSD ($p < .05$) showed that devalued animals had fewer active nosepoke responses than valued animals in the lesion, but not sham, condition (Figure 6a), indicating that the lesion enhanced the influence of outcome value on instrumental responding. Despite the interaction observed in active nosepoke responses, both lesion and sham devalued groups consumed fewer reinforcer pellets compared with the valued animals [$F(1,34) = 50.63, p < .0001$] (Figure 6b).

Intermediate dorsal lesions (Exp. 4): Active nosepoke responding during the non-reinforced test was converted to a percent of baseline responding during the first five minutes on the last acquisition day. There was no effect of either the surgery condition or the reinforcer value on responding (Figure 6c), indicating that both groups responded in a habitual manner. Despite this lack of differences in active nosepoke responding, both lesion and sham devalued groups consumed fewer reinforcer pellets compared with the valued animals [$F(1,29) = 77.03, p < .0001$] (Figure 6d).

CTA Transfer Test—One day following the extinction and consumption tests in Experiment 3, a reinforcer pellet consumption test was conducted in the instrumental conditioning chamber to confirm that the CTA transferred to the context in which all of the instrumental behavior is being assessed. Figure 6e shows that the CTA does indeed transfer to the instrumental chamber. Both lesion and sham devalued groups consumed fewer free reinforcer pellets in the instrumental chamber compared with the valued animals [$F(1,34) = 8.57, p < .01$].

Discussion

This series of experiments demonstrates a role for the most lateral portion of dorsal striatum in the performance of habitual responses in mice. While control animals display an insensitivity to post-training decreases (Exp. 1) or increases (Exp. 2) in reinforcer value, large lesions of dorsal striatum that include both lateral and medial portions preserve this sensitivity to changes in reinforcer value. Importantly, we show further that limited lesions of the most lateral aspect of the dorsal striatum are sufficient to preserve sensitivity to changes in reinforcer value (Exp. 3) and, in fact, lesions of dorsal striatum that avoid this extreme lateral portion fail to preserve this sensitivity (Exp. 4). These data suggest a role for this most lateral region of dorsal striatum in habitual responding and reinforce the heterogeneity of function within the dorsolateral striatum.

Outcome devaluation is one means of assessing goal-directed responding (Yin & Knowlton, 2006; Balleine & O'Doherty, 2010; Lingawi & Balleine, 2012). Typically, a given action is reinforced by pairing that action with a desirable outcome. Following such training, the outcome is devalued either by pairing its consumption with nausea or by allowing the animal to freely consume the outcome until satiated. Humans and other animals performing on the basis of the action-outcome association will decrease subsequent performance for the now devalued outcome. However, performance of a stimulus-elicited response that is independent of the outcome value – that is, an S-R habit – will persist despite such outcome devaluation. Goal-directed actions and stimulus-response habits rely on distinct corticostriatal networks (e.g., Tricomi *et al.*, 2009; Yin *et al.*, 2004; Lingawi & Balleine, 2012).

Here we show that damage to the dorsolateral striatum restores sensitivity to changes in outcome value following outcome devaluation in mice. This recapitulates an earlier finding in rats (Yin *et al.*, 2004), confirming the conservation across species (at least in rodents) of this anatomical dissociation within the striatum. Our lesions were more restricted to the most lateral portion of the striatum than the lesions reported previously in rats, further constraining this anatomical correlate of S-R habit learning. We also show, for the first time, that striatal lesions restore sensitivity to changes in outcome value following reinforcer *inflation*. This indicates that the response changes in these lesioned animals truly mirror the current value of the reinforcer; the disruption of habitual responding is not specifically tied to the devaluation procedure.

A previous study in rats found posterior dorsomedial striatal lesions to disrupt instrumental acquisition (Yin *et al.*, 2005b). Our experiments do not recapitulate this finding. While this could result from a between-species difference, the more likely explanation is that the difference between these reports derives from a difference in the lesions: our intermediate lesions did not extend into the most medial and posterior portion of the striatum. This observation further emphasizes functional differentiation within the rodent striatum.

Conservation of this medial-lateral functional dissociation across rodent species is an important finding given the recent surge of interest in identifying the cellular and molecular mechanisms mediating the learning and performance of goal-directed actions and stimulus-response habits using genetically modified mice. Critical work in mouse models of habitual responding has enabled the detection of the roles of adenosine signaling (Yu *et al.*, 2009) and the cannabinoid system (Hilário *et al.*, 2007) in the expression and development of S-R habits. Additionally, our work in the four core genotypes mouse model has shown that genetic, but not hormonal, sex determines the rate of habit formation for both a natural reinforcer (Quinn *et al.*, 2007) and ethanol (Barker *et al.*, 2010). It is worth noting that two of these studies have shown significant devaluation effects on instrumental performance following pairings of the instrumental reinforcer (food pellet or alcohol) with LiCl in mice (Quinn *et al.*, 2007; Barker *et al.*, 2010).

A recent study by Lederle *et al.* (2011) suggests that instrumental performance in at least some mouse strains may be less sensitive to LiCl-induced devaluation effects. This suggestion relies upon a comparison of their data with another study conducted using satiety-induced devaluation, where instrumental performance was reduced in devalued mice (Hilário *et al.*, 2007). It is possible that the mice in the Lederle *et al.* (2011) study were performing a habitual response and performance would, therefore, be insensitive to reinforcer devaluation while the mice in the earlier study using satiety-induced devaluation were performing an action-outcome instrumental response. This is underscored by the fact that other studies, including the present data, have shown significant LiCl-induced

devaluation effects on instrumental performance in mice (Quinn *et al.*, 2007; Barker *et al.*, 2010).

The striatum is also involved in a variety of other behaviors. While the instrumental habit paradigm employed here is particularly powerful in that manipulations of reinforcer value allow the outcome independence of the S-R association to be rigorously established (e.g., Yin & Knowlton, 2006), the striatum, and the dorsolateral striatum in particular, has consistently been associated with relatively inflexible, cue-driven behavior across a variety of paradigms. For example, lesions or genetic perturbations specifically disrupt cue-driven navigation in mice (Pittenger *et al.*, 2006; Lee *et al.*, 2008) and rats (e.g., Packard & McGaugh, 1996). Specific dorsolateral striatal lesions and localized pharmacological manipulations, also impair cue-driven navigational learning (Devan *et al.*, 1999), as well as sequence learning (Yin, 2010), and aspects of cocaine-driven habitual responding (Belin & Everitt, 2008) in rats. However, few previous studies have explored specific manipulations of striatal subregions in mice (Yin *et al.*, 2009).

While stereotyped, habitual behaviors can be rather adaptive – allowing for precisely executed, rapid performance of frequent behaviors – their inflexibility can also become detrimental and can contribute to drug addiction (Everitt & Robbins, 2005) and to other neuropsychiatric conditions (e.g., Leckman & Riddle, 2000; Graybiel, 2008). Increasing our knowledge of the anatomical, neurochemical, and molecular mediators of habitual behavior may provide insight into the pathophysiological neural functions underlying intrusive, compulsive behaviors in psychiatric illnesses such as obsessive-compulsive disorder, Tourette syndrome, and addiction (Leckman & Riddle, 2000; Canales, 2005; Everitt & Robbins, 2005; Albin & Mink, 2006; Graybiel, 2008; Torregrossa *et al.*, 2008).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by PHS DA011717 (JRT), DA027844 (JRT), MH066172 (JRT), AA017776 (JRT), K08MH081190 (CP), the Interdisciplinary Research Consortium on Stress, Self-control and Addiction (UL1-DE19586 and the NIH Roadmap for Medical Research/Common Fund, AA017537), a NARSAD Young Investigator Award (CP), and the CT Department of Mental Health and Addiction Services, through its support of the Ribicoff Research Facilities at the Connecticut Mental Health Center (CP, JRT).

References

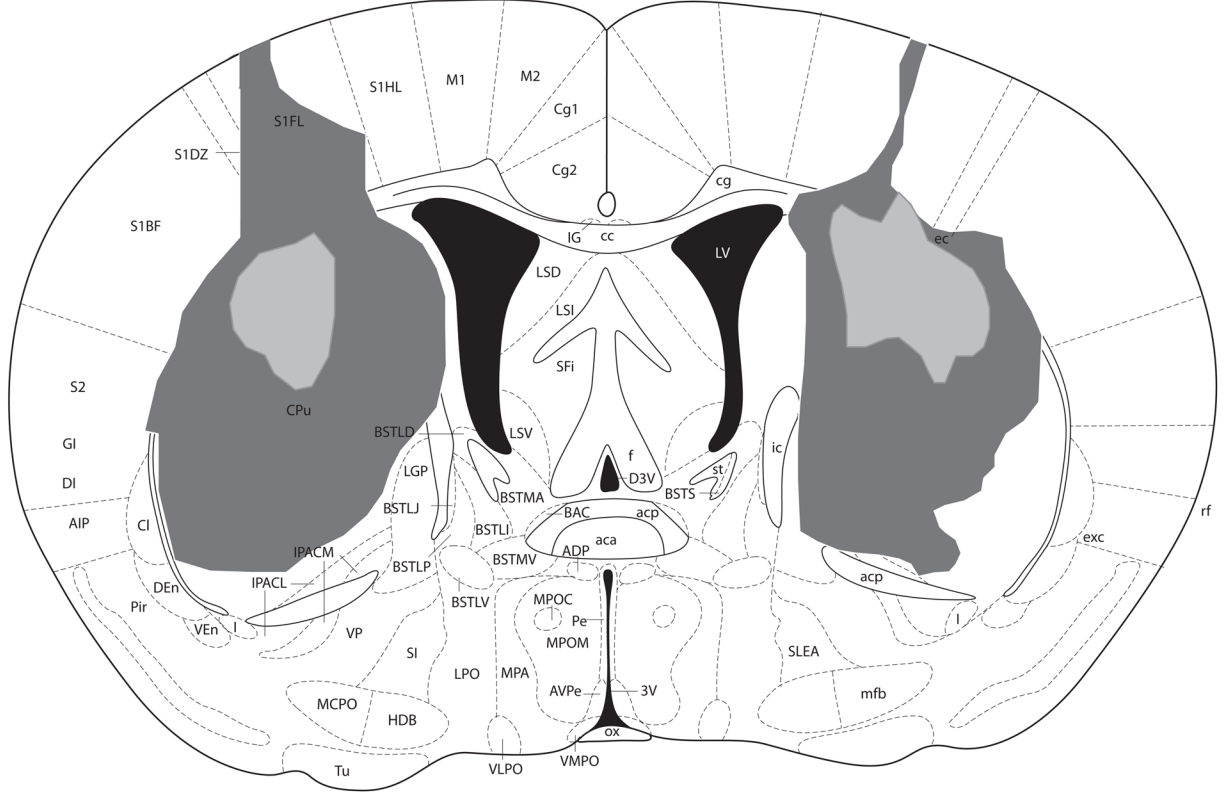
- Adams C, Dickinson A. Instrumental responding following reinforcer devaluation. *Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology*. 1981; 33:109–121.
- Albin RL, Mink JW. Recent advances in Tourette syndrome research. *Trends Neurosci*. 2006; 29:175–182. [PubMed: 16430974]
- Ashby FG, Turner BO, Horvitz JC. Cortical and basal ganglia contributions to habit learning and automaticity. *Trends Cogn Sci*. 2010; 14:208–215. [PubMed: 20207189]
- Balleine BW, Dickinson A. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology*. 1998; 37:407–419. [PubMed: 9704982]
- 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]
- Barker JM, Torregrossa MM, Arnold AP, Taylor JR. Dissociation of genetic and hormonal influences on sex differences in alcoholism-related behaviors. *J Neurosci*. 2010; 30:9140–9144. [PubMed: 20610747]

- Barnes TD, Kubota Y, Hu D, Jin DZ, Graybiel AM. Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature*. 2005; 437:1158–1161. [PubMed: 16237445]
- Belin D, Everitt BJ. Cocaine seeking habits depend upon dopamine-dependent serial connectivity linking the ventral with the dorsal striatum. *Neuron*. 2008; 57:432–441. [PubMed: 18255035]
- Canales JJ. Stimulant-induced adaptations in neostriatal matrix and striosome systems: transiting from instrumental responding to habitual behavior in drug addiction. *Neurobiol Learn Mem*. 2005; 83:93–103. [PubMed: 15721792]
- Carelli RM, Wolske M, West MO. Loss of lever press-related firing of rat striatal forelimb neurons after repeated sessions in a lever pressing task. *The Journal of neuroscience: the official journal of the Society for Neuroscience*. 1997; 17:1804–1814. [PubMed: 9030639]
- Colwill R, Rescorla R. The role of response-reinforcer associations increases throughout extended instrumental training. *Animal Learning & Behavior*. 1988; 16:105–111.
- Corbit LH, Nie H, Janak PH. Habitual alcohol seeking: time course and the contribution of subregions of the dorsal striatum. *Biological Psychiatry*. 2012; 72:389–395. [PubMed: 22440617]
- Devan BD, McDonald RJ, White NM. Effects of medial and lateral caudate-putamen lesions on place- and cue-guided behaviors in the water maze: relation to thigmotaxis. *Behav Brain Res*. 1999; 100:5–14. [PubMed: 10212049]
- Everitt BJ, Robbins TW. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat Neurosci*. 2005; 8:1481–1489. [PubMed: 16251991]
- Faure A, Haberland U, Conde F, El Massioui N. Lesion to the nigrostriatal dopamine system disrupts stimulus-response habit formation. *The Journal of neuroscience: the official journal of the Society for Neuroscience*. 2005; 25:2771–2780. [PubMed: 15772337]
- Featherstone RE, McDonald RJ. Dorsal striatum and stimulus-response learning: lesions of the dorsolateral, but not dorsomedial, striatum impair acquisition of a simple discrimination task. *Behavioural brain research*. 2004; 150:15–23. [PubMed: 15033275]
- Grahn JA, Parkinson JA, Owen AM. The cognitive functions of the caudate nucleus. *Prog Neurobiol*. 2008; 86:141–155. [PubMed: 18824075]
- Graybiel AM. Habits, rituals, and the evaluative brain. *Annu Rev Neurosci*. 2008; 31:359–387. [PubMed: 18558860]
- Hilário MR, Clouse E, Yin HH, Costa RM. Endocannabinoid signaling is critical for habit formation. *Front Integr Neurosci*. 2007; 1:6. [PubMed: 18958234]
- Hilário MR, Costa RM. High on habits. *Front Neurosci*. 2008; 2:208–217. [PubMed: 19225594]
- Hitchcott PK, Quinn JJ, Taylor JR. Bidirectional modulation of goal-directed actions by prefrontal cortical dopamine. *Cerebral cortex (New York, NY: 1991)*. 2007a; 17:2820–2827.
- Hitchcott PK, Quinn JJ, Taylor JR. Bidirectional modulation of goal-directed actions by prefrontal cortical dopamine. *Cereb Cortex*. 2007b; 17:2820–2827. [PubMed: 17322558]
- Holland PC. Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *J Exp Psychol Anim Behav Process*. 2004; 30:104–117. [PubMed: 15078120]
- Jog MS, Kubota Y, Connolly CI, Hillegaart V, Graybiel AM. Building neural representations of habits. *Science (New York, NY)*. 1999; 286:1745–1749.
- Kubota Y, Liu J, Hu D, DeCoteau WE, Eden UT, Smith AC, Graybiel AM. Stable encoding of task structure coexists with flexible coding of task events in sensorimotor striatum. *J Neurophysiol*. 2009; 102:2142–2160. [PubMed: 19625536]
- Leckman JF, Riddle MA. Tourette's syndrome: when habit-forming systems form habits of their own? *Neuron*. 2000; 28:349–354. [PubMed: 11144345]
- Lederle L, Weber S, Wright T, Feyder M, Brigman JL, Crombag HS, Saksida LM, Bussey TJ, Holmes A. Reward-related behavioral paradigms for addiction research in the mouse: performance of common inbred strains. *PLoS One*. 2011; 6:e15536. [PubMed: 21249214]
- Lee AS, Duman RS, Pittenger C. A double dissociation revealing bidirectional competition between striatum and hippocampus during learning. *Proc Natl Acad Sci U S A*. 2008; 105:17163–17168. [PubMed: 18955704]

- Lingawi NW, Balleine BW. Amygdala central nucleus interacts with dorsolateral striatum to regulate the acquisition of habits. *Journal of Neuroscience*. 2012; 32:1073–1081. [PubMed: 22262905]
- Lovinger DM. Neurotransmitter roles in synaptic modulation, plasticity and learning in the dorsal striatum. *Neuropharmacology*. 2010; 58:951–961. [PubMed: 20096294]
- Packard MG, McGaugh JL. Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. *Neurobiol Learn Mem*. 1996; 65:65–72. [PubMed: 8673408]
- Pennartz CM, Berke JD, Graybiel AM, Ito R, Lansink CS, van der Meer M, Redish AD, Smith KS, Voorn P. Corticostriatal Interactions during Learning, Memory Processing, and Decision Making. *J Neurosci*. 2009; 29:12831–12838. [PubMed: 19828796]
- Pittenger C, Fasano S, Mazzocchi-Jones D, Dunnett SB, Kandel ER, Brambilla R. Impaired bidirectional synaptic plasticity and procedural memory formation in striatum-specific cAMP response element-binding protein-deficient mice. *J Neurosci*. 2006; 26:2808–2813. [PubMed: 16525060]
- Quinn JJ, Hitchcott PK, Umeda EA, Arnold AP, Taylor JR. Sex chromosome complement regulates habit formation. *Nat Neurosci*. 2007; 10:1398–1400. [PubMed: 17952068]
- Ragozzino ME, Ragozzino KE, Mizumori SJ, Kesner RP. Role of the dorsomedial striatum in behavioral flexibility for response and visual cue discrimination learning. *Behavioral neuroscience*. 2002; 116:105–115. [PubMed: 11898801]
- Rescorla R. Response outcome versus outcome response associations in instrumental learning. *Animal Learning & Behavior*. 1992; 20:223–232.
- Tang C, Pawlak AP, Prokopenko V, West MO. Changes in activity of the striatum during formation of a motor habit. *The European journal of neuroscience*. 2007; 25:1212–1227. [PubMed: 17331217]
- Torreghrossa MM, Quinn JJ, Taylor JR. Impulsivity, compulsivity, and habit: the role of orbitofrontal cortex revisited. *Biol Psychiatry*. 2008; 63:253–255. [PubMed: 18194683]
- Tricomi E, Balleine BW, O'Doherty JP. A specific role for posterior dorsolateral striatum in human habit learning. *Eur J Neurosci*. 2009; 29:2225–2232. [PubMed: 19490086]
- Wickens JR, Horvitz JC, Costa RM, Killcross S. Dopaminergic mechanisms in actions and habits. *The Journal of neuroscience: the official journal of the Society for Neuroscience*. 2007; 27:8181–8183. [PubMed: 17670964]
- Yin HH. The sensorimotor striatum is necessary for serial order learning. *JNeurosci*. 2010; 30:14719–14723. [PubMed: 21048130]
- Yin HH, Knowlton BJ. The role of the basal ganglia in habit formation. *Nature reviews Neuroscience*. 2006; 7:464–476.
- Yin HH, Knowlton BJ, Balleine BW. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *The European journal of neuroscience*. 2004; 19:181–189. [PubMed: 14750976]
- Yin HH, Knowlton BJ, Balleine BW. Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. *The European journal of neuroscience*. 2005a; 22:505–512. [PubMed: 16045503]
- Yin HH, Knowlton BJ, Balleine BW. Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. *Behavioural brain research*. 2006; 166:189–196. [PubMed: 16153716]
- Yin HH, Mulcare SP, Hilário MR, Clouse E, Holloway T, Davis MI, Hansson AC, Lovinger DM, Costa RM. Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nat Neurosci*. 2009; 12:333–341. [PubMed: 19198605]
- Yin HH, Ostlund SB, Knowlton BJ, Balleine BW. The role of the dorsomedial striatum in instrumental conditioning. *The European journal of neuroscience*. 2005b; 22:513–523. [PubMed: 16045504]
- Yu C, Gupta J, Chen JF, Yin HH. Genetic deletion of A2A adenosine receptors in the striatum selectively impairs habit formation. *J Neurosci*. 2009; 29:15100–15103. [PubMed: 19955361]

a

Experiment 1



b

Experiment 2

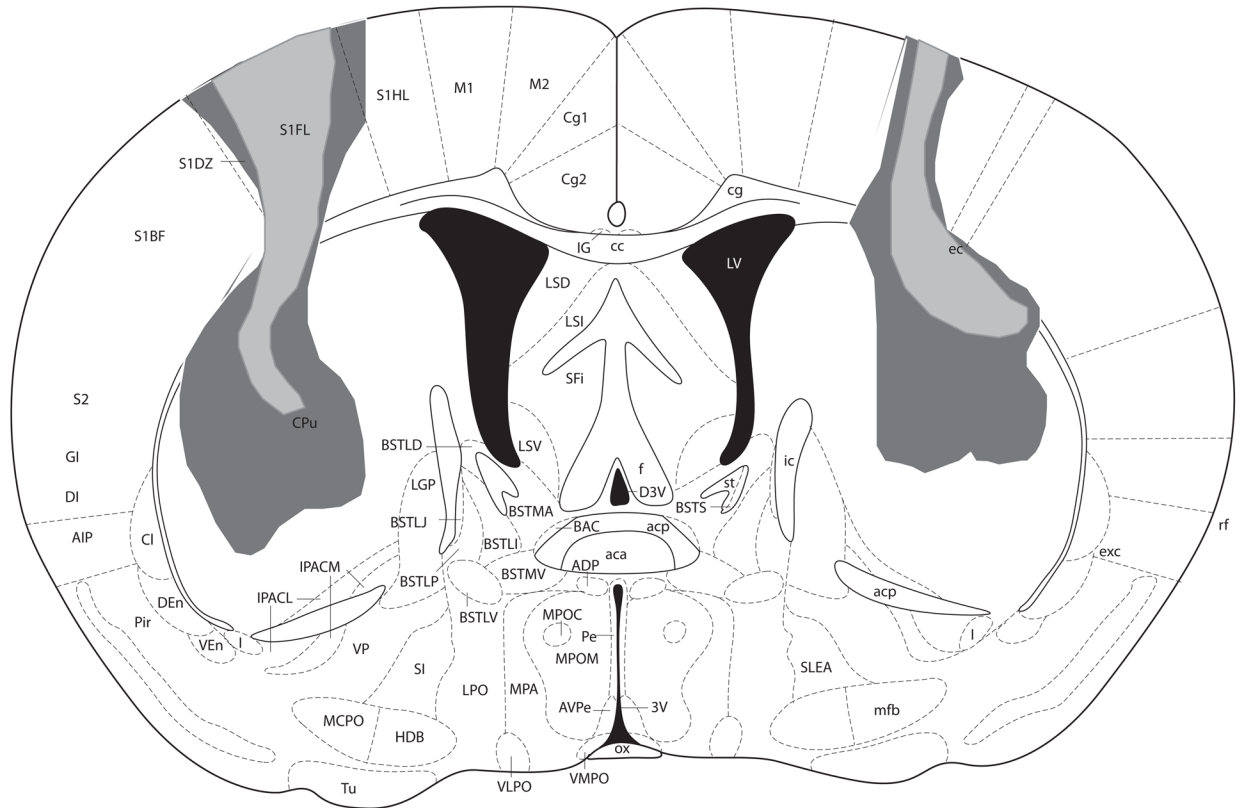


Figure 1. Schematic representation of NMDA-induced excitotoxic lesions of the dorsal striatum in Experiments 1 (a) and 2 (b). Shaded areas represent the minimum (light gray) and maximum (dark gray) extent of the lesions for mice included in all analyses (adapted from Paxinos & Franklin, 2001).

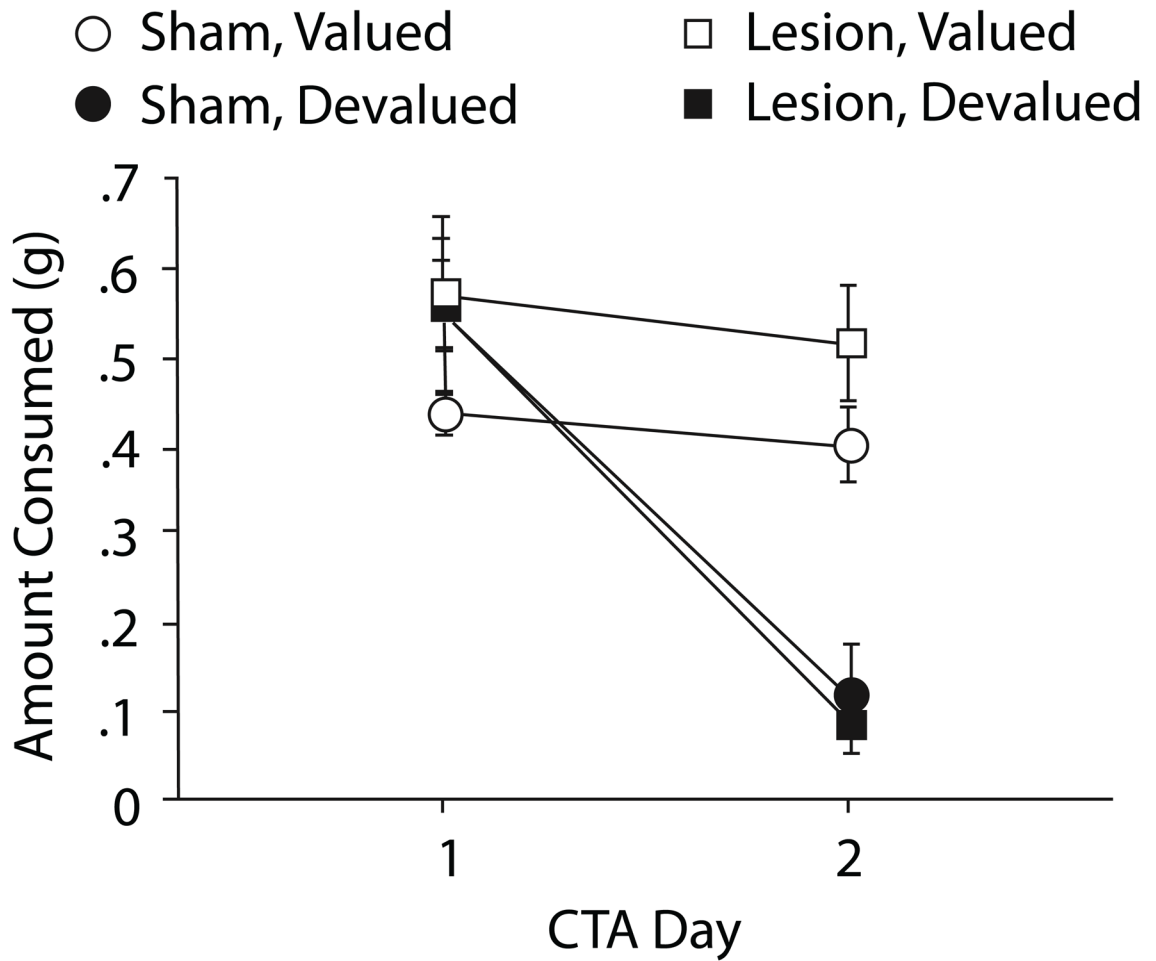


Figure 2. Conditioned taste aversion (CTA) acquisition for Experiment 1. On the first CTA day, mice equally consumed the food pellet reinforcer when allowed 15min free access. Devalued mice subsequently decreased consumption of the reinforcer demonstrating a conditioned taste aversion.

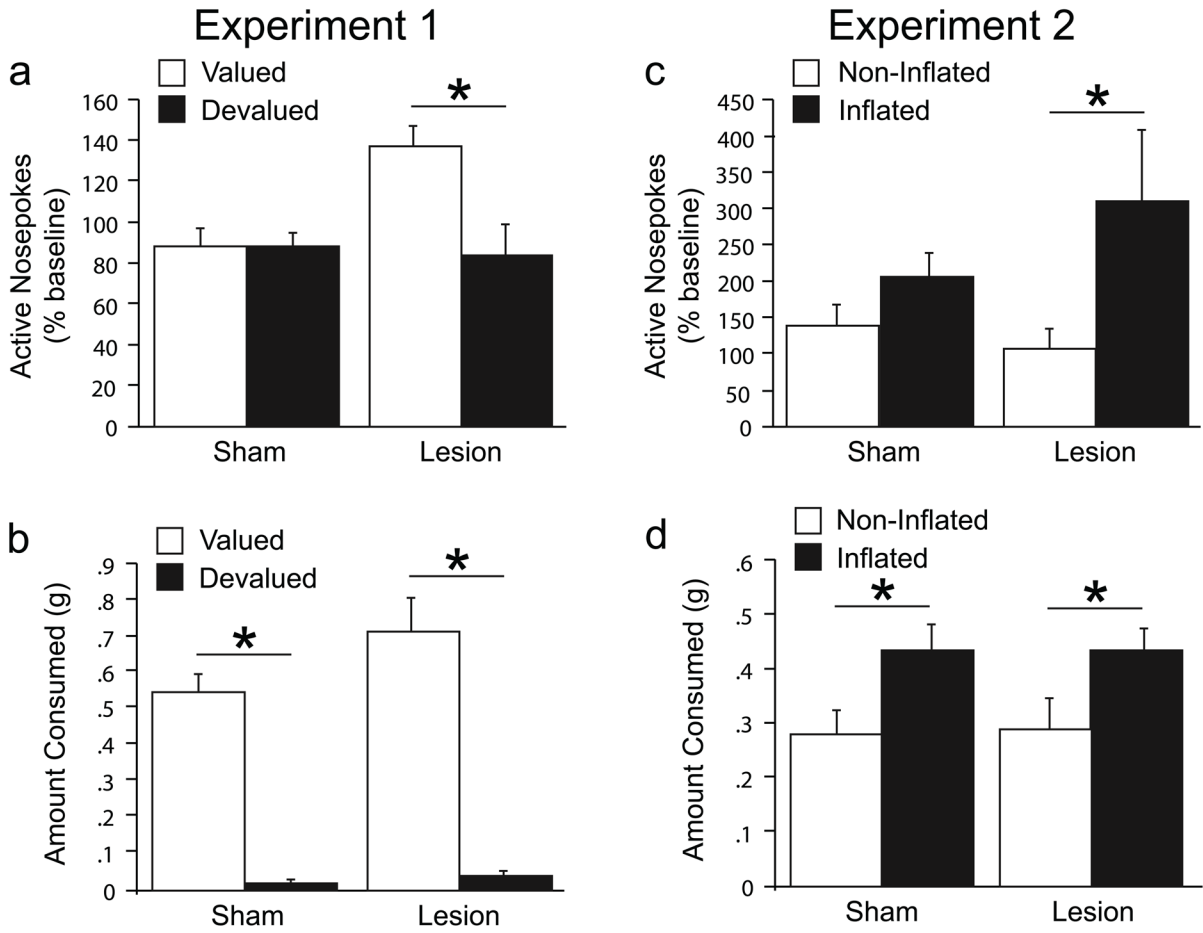
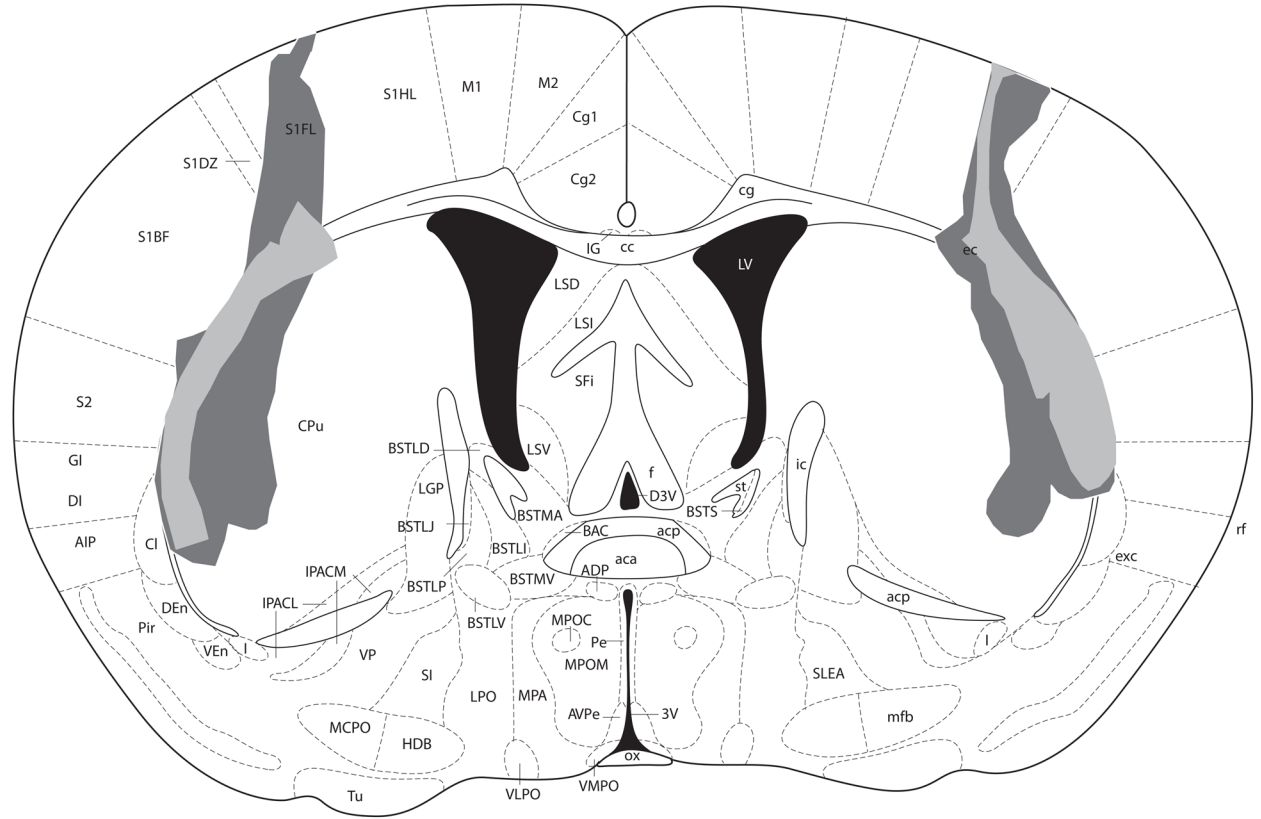


Figure 3. Test session data for Experiments 1 and 2. Active nosepoke responding during the 5min instrumental test (a, c). Free access reinforcer consumption in the CTA context (Experiment 1) or novel context (Experiment 2) immediately following the instrumental test (b, d).

a

Experiment 3



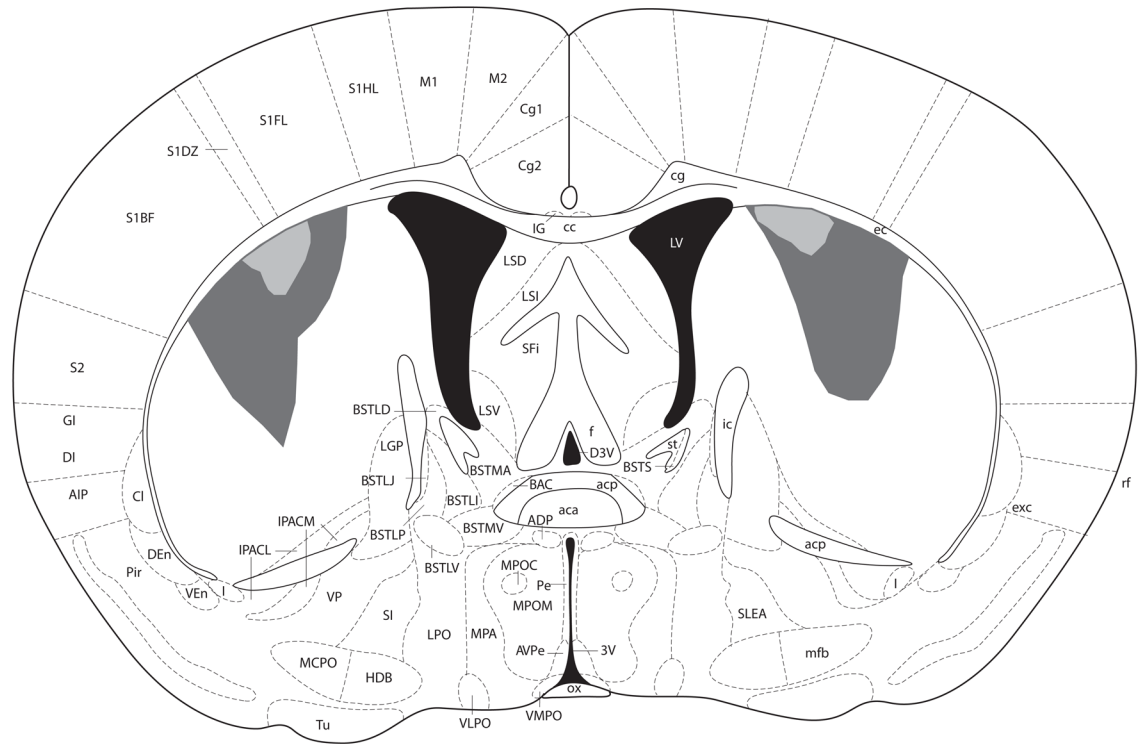
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

b

Experiment 4

**Figure 4.**

Schematic representation of NMDA-induced excitotoxic lesions of the lateral (a) and intermediate (b) dorsal striatum in Experiments 3 and 4, respectively. Shaded areas represent the minimum (light gray) and maximum (dark gray) extent of the lesions for mice included in all analyses (adapted from Paxinos & Franklin, 2001).

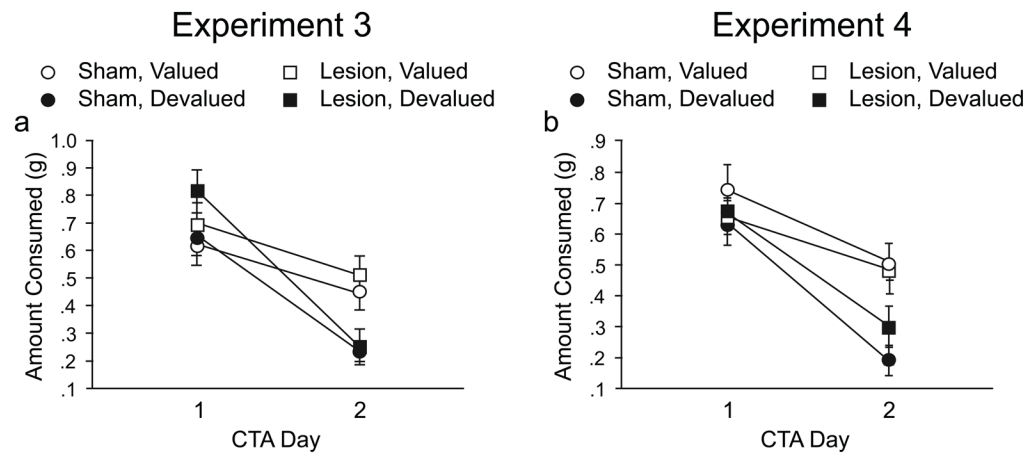


Figure 5. Conditioned taste aversion (CTA) acquisition for Experiments 3 (a) and 4 (b). On the first CTA day in both experiments, mice equally consumed the food pellet reinforcer when allowed 15min free access. Devalued mice subsequently decreased consumption of the reinforcer demonstrating a conditioned taste aversion.

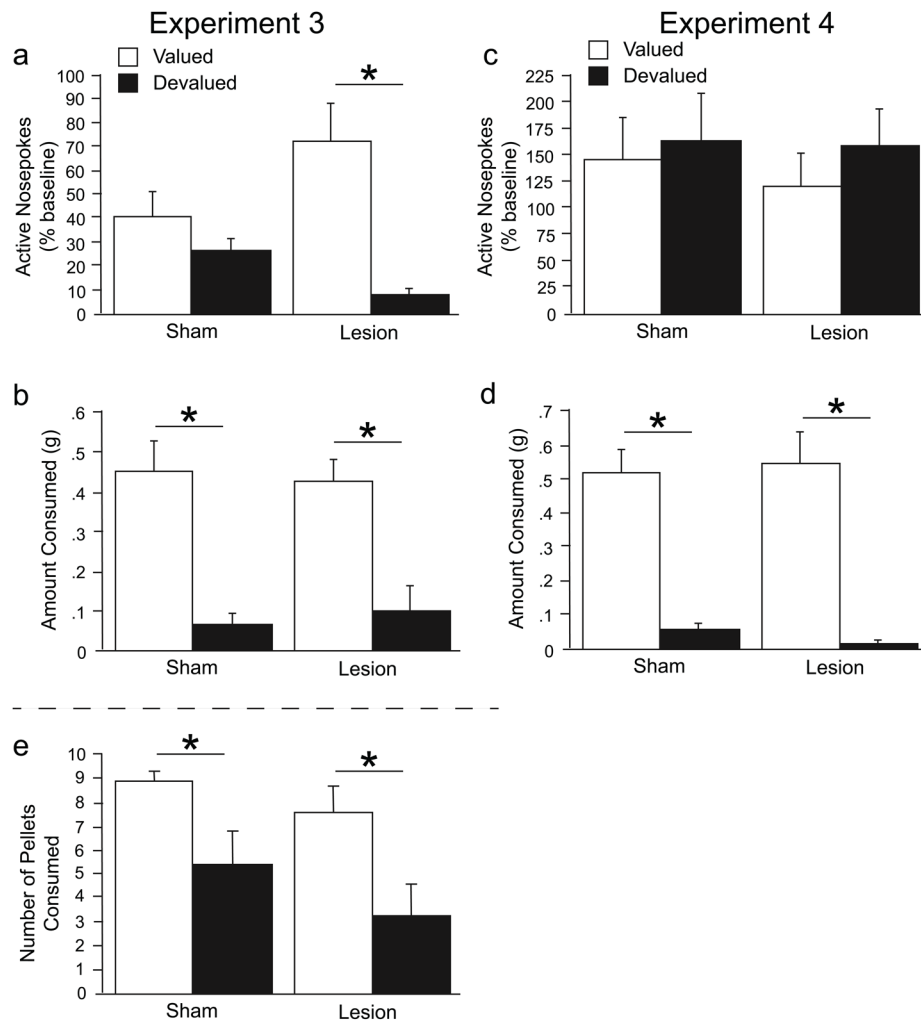


Figure 6. Test session data for Experiments 3 and 4. Active nosepoke responding during the 5min instrumental test (a, c). Free access reinforcer consumption in the CTA context immediately following the instrumental test (b, d). Free reinforcer consumption in the instrumental conditioning chamber (CTA transfer test) in Experiment 3 (e).