Ventral striatum lesions enhance stimulus and response encoding in dorsal striatum

Amanda C. Burton[1,2], **Gregory B. Bissonette**[1], **Nina T Lichtenberg**[1], **Vadim Kashtelyan**[1], and **Matthew R. Roesch**[1,2]

[1]Department of Psychology, 1147 Biology-Psychology Building, University of Maryland, College Park, MD 20742

[2]Program in Neuroscience and Cognitive Science, 1147 Biology-Psychology Building, University of Maryland, College Park, MD 20742

Abstract

Background—The development of addiction is thought to reflect a transition from goal-directed to stimulus-response driven behavior, functions attributed to ventral (VS) and dorsal striatum (DS), respectively. In line with this theory, neuroadaptations that occur during prolonged drug use progress from VS to DS. Here, we ask if VS dysfunction alone, independent of drug use, can impact neural selectivity in DS.

Methods—To address this issue we recorded from single neurons in DS while rats performed an odor-guided choice task for differently valued rewards in rats with and without unilateral VS lesions. In a separate group of animals we used bilateral VS lesions to determine if VS was critical for performance on this task.

Results—We describe data showing that unilateral lesions of VS enhance neural representations in DS during performance of a task that is dependent on VS. Furthermore, we show VS is critical for reward-guided decision-making initially, but rats regain function after several days.

Conclusion—These results suggest that loss of VS function, independent of chronic drug use, can trigger stronger encoding in DS in a reward-guided decision-making task and that the transition from VS to DS governed behavior observed in addiction might be due, in part, to initial loss of VS function.

INTRODUCTION

Ventral and dorsal striatum perform critical roles in reward-guided decision-making and reinforcement learning, but it is still unclear how they interact. Together, with midbrain dopamine neurons, they form a circuit commonly referred to as the actor-critic model (1–

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

12). In this model, VS and dopamine neurons function to generate reward predictions and prediction errors, which modify action policies in DS so that desired outcomes can be obtained. This circuit is thought to be critical for drug-seeking and is affected by chronic drug use (13–17).

Many behaviors, including drug-seeking, are initially goal-directed, but eventually become stimulus driven or habitual with repetition. The transition away from goal-directed behavior toward stimulus driven habits is thought to depend upon a switch in control from VS to DS (18, 19), which is amplified by drugs of abuse (18, 20). Since many structural and functional alterations occur with extended drug use it is still unclear what might initiate this change. Importantly, VS appears to be one of the earliest brain regions to be affected by administration of drugs of abuse, suggesting that its disruption might be enough to initiate changes in downstream areas critical for stimulus driven behaviors. Here, we ask if loss of VS function alone, independent of drug use, might increase encoding in DS.

Consistent with this hypothesis, stimulus and response encoding in DS was enhanced after VS lesions during performance of a task that was dependent on VS. After several days of post surgery training, lesioned rats were able to make accurate reward-guided decisions, suggesting that enhanced encoding in DS might compensate for loss of VS function. These results demonstrate that disruption of decision-making with lesions to VS is enough to amplify signals in DS. This suggests that the main locus by which prolonged drug use transitions behavior away from goal-directed, to S-R driven, might reflect initial neuroadaptations in VS, and that this alone is enough to initiate changes in DS and enhance S-R learning.

METHODS

Subjects

Twenty-six male Long-Evans rats were obtained at 175–200g from Charles River Labs. Rats were tested at the University of Maryland, College Park in accordance with UM and NIH guidelines.

Surgical procedures

All surgical procedures were performed after training on the task described below. Ten rats had a drivable bundle of 10–25 μm diameter FeNiCr wires chronically implanted in the left or right hemisphere dorsal to DS (n = 10; 1 mm anterior to bregma, + or − 3.2 mm laterally, and 3.5 mm ventral to the brain surface)(21–23). VS lesions were made with a 2 μl Hamilton syringe, beveled edge facing the posterior direction, using 0.11M quinolinic acid, pH 7.4 in Dulbecco's PBS (Sigma). Quinolinic acid (0.3 μl) was delivered at 0.15 μl/min at coordinates: AP +1.9 ML ±1.9 DV −7.3. The remaining 4 rats served as controls, which received sham surgeries during which the Hamilton syringe loaded with saline was lowered to the same coordinates. In addition to rats that received electrodes, another group of rats only received bilateral sham (n = 6) or VS lesions (n = 8) to characterize behavior. Brains were removed and processed for histology using standard techniques at the end of the experiment (21).

Odor-guided delay/size choice task

Before surgery, all rats were trained on the odor-guided delay/size choice task. On each trial, nose poke into the odor port after house light illumination resulted in delivery of an odor cue to a hemicylinder located behind this opening (24, 25). One of three different odors (2-Octanol, Pentyl Acetate, or Carvone) was delivered to the port on each trial. One odor instructed the rat to go to the left to receive reward, a second odor instructed the rat to go to

the right to receive reward, and a third odor indicated that the rat could obtain reward at either well. Odors were presented in a pseudorandom sequence such that the free-choice odor was presented on 7/20 trials and the left/right odors were presented in equal proportions.

During the first day of training rats were first taught to simply nose poke into the odor port, then respond to the well for reward. On the second day, the free-choice odor was introduced and rats were free to respond to either well for reward. On each subsequent day, the number of forced-choice odors increased by 2 for each block of 20 trials. During this time we introduced blocks in which we manipulated the reward size and the length of the delay preceding reward. Once the rats were able to maintain accurate responding ($> 65\%$) on forced-choice trials through these manipulations, surgery was performed.

During recording, one well was randomly designated as short (500 ms) and the other long (1–7s) at the start of the session (Figure 1A: Block 1). In the second block of trials, these contingencies were switched (Figure 1A: Block 2). The length of the delay under long conditions abided by the following algorithm: the side designated as long started off as 1s and increased by 1s every time that side was chosen on a free-choice odor (up to a maximum of 7s). If the rat chose the side designated as long less than 8 out of the previous 10 free choice trials, the delay was reduced by 1s for each trial to a minimum of 3s. The reward delay for long forced-choice trials was yoked to the delay in free-choice trials during these blocks. In later blocks, we held the delay preceding reward delivery constant (500 ms) while manipulating the size of the expected reward (Figure 1A: Blocks 3 and 4). The reward was a 0.05 ml bolus of 10% sucrose solution. For big reward, an additional bolus was delivered 500 ms after the first bolus. Essentially there were four basic trial types (short, long, big, and small) by two directions (left and right) by two stimulus types (free- and forced-choice odor).

For behavior after bilateral lesions, only one manipulation varied each day. On each day, one well was randomly designated as high value (i.e. short delay or large reward depending on the day). The location of the high value outcomes switched every 60 correct trials. There were a total of 3 blocks each day. Delay manipulations occurred on days 1, 3, 5 and 7. Size manipulations occurred on days 2, 4, 6, and 8. All other contingencies were the same as during recording.

Single-unit recording

Procedures were the same as described previously (24, 26). Electrodes were advanced daily (40–80μm). Neural activity was recorded using four identical Plexon systems interfaced with odor discrimination chambers. Waveforms (>2.5:1 signal-to-noise) were extracted from active channels and recorded.

Neural analysis

Analysis epochs were computed by taking the total number of spikes and dividing by time. That analysis epoch was taken from 100 ms after odor onset to completion of the behavioral response. The activity of neurons during this epoch for which we examined differences between trial types at the single cell level only violated normality (Jarque-Bera; $p < 0.05$) in 6% of the 809 recorded neurons, which is fewer than expected from chance alone (chi-square; $p = 0.09$). A multi-factor ANOVA ($p < 0.05$) was performed for each neuron to determine if activity was modulated by odor type (free versus forced), response direction (left versus right), and expected outcome (short, long, big and small).

Neurons were also characterized by comparing firing rate during baseline to firing rate during the analysis epoch, averaged over all trial types (t test, $p < 0.05$). Baseline was the

average firing rate taken for 1s, starting 2s before odor onset. Chi-squares were performed to assess differences in the counts of neurons showing significant modulation across lesion and controls.

Behavioral analysis

Behavior during performance of the task was evaluated by computing percent choice of high and low value outcomes on free-choice trials, and percent correct and reaction time (odor offset to odor port exit) on forced-choice trials. Multi-factor ANOVAs were performed on these behavior measures to assess differences between lesions and controls. Factors included value (high versus low), manipulation (delay vs size), lesion (sham vs lesion), day (1–4) and block (1–3).

RESULTS

Rats were trained on a reward-guided decision-making task previously used to characterize encoding in several areas (Fig. 1A)(25, 27). Rats learned to nosepoke in a central odor port, wait for delivery of an odor (500 ms) and then respond to one of two fluid wells located to the left or right of the odor port. Odors signaled forced- and free-choice trials. On forced-choice trials, rats had to respond to the fluid well signaled by the odor (left or right) to receive reward. On free-choice trials rats were free to select either well. Over the course of four trial blocks we independently manipulated the length of the delay to (0.5 s or 1–7 s) or the size of (1 vs 2–3 boli) reward, making one fluid well better than the other. Essentially there were four predicted outcomes (short, long, big and small) by two directions (left and right) by two trial-types (free and forced odors).

After training on this task rats were split into two groups: in both groups we implanted recording electrodes in DS, in the lesion group (n = 7) rats received unilateral VS lesions in the same hemisphere as the DS electrode. Control rats received sham lesions (n = 4; see methods for further detail). Unilateral lesions were chosen to examine the impact of VS lesions on DS neural selectivity, with a minimal impact on behavior. Lesions and electrode positions are illustrated in Figure 1E–F.

Both control and lesion rats perceived differently delayed and sized rewards as having different values across all four trial blocks. There was a main effect of value with no interaction with lesion (ANOVA; $p < 0.05$). On free-choice trials, rats chose the well associated with large reward and short delay significantly more often than the well associated with small reward and long delay, respectively (Fig. 1B; t test; p's < 0.05). On forced-choice trials, rats were more accurate and faster on large reward and short delay trials, as compared to their respective counterparts, small reward and long delay (Fig. 1C–D; t test; p's < 0.05). Thus, performance on free- and forced-choice trials was modulated by the predicted outcomes in both size and delay trial blocks for both controls and lesions.

The only significant difference between the lesion and sham groups was that rats with VS lesions were significantly slower to decide when to leave the odor port (reaction time = port exit minus odor offset) compared to controls (Fig. 1D; control versus lesion; ANOVA; interaction effect; $p < 0.05$). Note, that this does not reflect a gross motor deficit because rats with lesions moved from the odor port to the fluid well (movement time) significantly faster than controls (t test; $p < 0.05$), suggesting that decisions took longer to process in rats with VS lesions, but the ability to move or act on that decision was unimpaired.

Stimulus and response encoding in DS was enhanced after VS lesions

We recorded activity from 457 and 352 DS neurons from lesion and control rats, respectively (Fig. 1E and F). As described previously, neural activity in DS was highly

associative, being modulated by all aspects of the task: trial-type (free vs forced odors), response direction (left vs right) and expected outcome (short, long, big and small) (27). This is illustrated by the activity of the single cell example in figure 2A during performance of size blocks. This neuron fired the strongest for forced-choice odors that predicted large reward and was directionally tuned in that firing was stronger for movement in the direction contralateral to the recording site (i.e. right).

The average firing over all recorded DS neurons from lesions and controls are shown in figure 2B–C. To make this plot, we sorted trials based on preferred trial-type according to which condition produced the maximal response. The remaining trials were categorized relative to the preferred trial type depending on whether the response and outcome were in the same or opposite direction, and were the same or opposite value for both manipulations (delay and size). For example, if a neuron fired maximally for big forced-choice trials to the right, then big reward and right became the preferred outcome and direction, small and left became non-preferred outcome and direction, and short and long became outcomes of the same and opposite value, respectively.

As described previously and consistent with the single cell example, population activity was highly selective, firing for a specific combination of odor, outcome and direction (27). Although activity was strongest for the preferred outcome and direction (by definition), there was little modulation by other trial types (Fig. 2B and C; blue versus other colors). This was true for both experimental and control groups. The only noticeable difference between population firing in sham and lesion rats is that cue-evoked activity appeared to be enhanced, and overall stronger, in rats with VS lesions (Fig 2B versus 2C). Indeed, the difference between the trial type that elicited the strongest firing and the trial-type designated as its opposite value and direction was significantly stronger in lesions compared to controls (t test; 2.9 versus 2.2, $p < 0.05$). Thus, at the population level, neural selectivity in DS was stronger in rats with VS lesions.

Consistent with a general increase in firing observed after lesions, the count of neurons that significantly increased and decreased firing during the decision period were significantly different between the two experimental groups (Fig. 2D). For this analysis, we compared average firing over all trial-types during the decision period (100 ms after odor onset to fluid well entry) to baseline (1 s prior to trial start; t test; $p < 0.05$). For controls, the number of neurons that increased and decreased firing were 116 (33%) and 144 (40%), respectively. This proportion was flipped for lesions, with 193 (42%) and 129 (28%) showing increased and decreased firing, respectively (Fig. 2D; chi-square; $p < 0.05$). The total number of responsive cells (increasing plus decreasing) were not significantly different between the two groups (control = 260(73%); lesion = 322(70%); chi-square; $p = 0.7$). Thus, the overall increase in firing observed at the population level reflects an altered frequency of increasing- and decreasing-type neurons in lesions compared to control rats.

To determine if selectivity at the single neuron level was different between lesions and controls we characterized firing by performing a ANOVA with stimulus (free- versus forced odors), response direction (contra vs ipsi), and predicted outcome (short, long, big, and small) as factors during the period starting 100 ms after odor onset and ending upon fluid well entry.

The results of this analysis for the entire population of responsive neurons are illustrated in figure 2E. The height of each bar represents the percent of neurons that showed a significant ($p < 0.05$) main or interaction effect of outcome (short, long, big, and small), response direction (contra or ipsilateral to the recording site), and/or stimulus-type (free- and forced-

choice odor). Each factor was then broken down by which trial-type produced the maximal firing.

As in our previous data set, an equal number of neurons from control rats fired maximally for each of the four predicted outcomes (Fig. 2E; black bars; short, long, big and small) (27). This effect was also present in rats with VS lesions (Fig. 2E; gray bars). The proportion of neurons selective for the four outcomes did not significantly differ between the two groups (Fig. 2e; short, long, big and small; chi-square; p's > 0.05).

Activity in DS was highly directional, firing more strongly for one direction over another, as illustrated by the single cell example and the population. Remarkably, this directional response selectivity was enhanced in rats with VS lesions, specifically for movements made in the contralateral direction relative to the recording site. The counts of neurons selective for contralateral movement in lesioned rats significantly outnumbered those selective for contralateral movement in controls (Fig. 2E; response direction; contralateral; black vs gray; chi-square; p < 0.05).

We have also shown that activity in DS is stimulus selective, with the majority of neurons showing maximal firing for free-choice odors (27). Consistent with those findings, counts of neurons that differentiated between free and forced odors were more than expected by chance alone (Fig. 2e; black bars; stimulus-type; chi-square; p < 0.05) and the frequency of neurons firing more strongly for free-choice odors were in the majority (Fig. 2e; black bars; stimulus-type; chi-square; p < 0.05). Like directional response tuning, stimulus selectivity was enhanced in rats with VS lesions, however, the elevated frequency of neurons showing increased firing in lesioned rats on free-choice trials only approached significance when examining the population as a whole (Fig. 2E; #: chi-square, p = 0.11).

Next, we asked if enhanced stimulus and response encoding was consistent across neurons that showed general increases and decreases in firing during the decision period. Remember that the ratio of increasing and decreasing neurons flipped after lesions, in that the number of increasing-type neurons were in the significant majority in the lesion, but not the control group (Fig. 2D). The population analysis performed earlier on all cells is broken down by increasing- and decreasing-type in Figure 3. Remarkably, response direction and stimulus encoding enhancement observed across the entire population was a result of specific changes in increasing- and decreasing-type neurons, respectively. For increasing-type cells the frequency of neurons that showed elevated firing for contralateral movement outnumbered those observed in controls (Fig. 3E; Response Direction; black versus gray; chi-square; p < 0.05). For decreasing cells (Fig. 3F), there was a significant increase (chi-square; p < 0.05) in the number cells that were modulated significantly by either free- or forced-choice odors (ANOVA; p < 0.05), with a trend toward a disproportionate increase on free-choice trials (Fig. 3F; #: chi-square; p = 0.057). We conclude that VS lesions induced increased firing during task performance and enhanced selectivity for stimuli and responses at the single unit and population level.

Reward-guided decision-making is disrupted after VS lesions

The results thus far suggest that VS is not critical for outcome encoding in DS and that associations related to direction and stimuli were enhanced after VS lesions. Although deficits in behavior have been observed after VS disruption, these results raise the important question of whether VS is critical for performance on our task (28–33). If VS is not critical, then we would not expect to see loss of outcome encoding in DS in the first place.

To address this issue, we trained a second group of animals in the same manner as the first, but instead of implanting electrodes we made bilateral lesions to VS (Fig. 4). We tested

them for eight days, alternating delay and size manipulations. Each day, we randomly selected which response direction would yield the high value reward (left or right). After 60 correct trials, contingencies were reversed for 60 trials and then reverted back to the original contingencies.

Figure 4 plots percent choice on free-choice trials (top row), and percent correct (middle row) and reaction time (bottom row) on forced-choice trials, averaged over the first two and last two test days for lesions and controls. A multi-factor ANOVA on number of free choices made to high and low valued outcomes produced significant (p < 0.05) interactions of lesion (control and lesion), block (1–3), value manipulation (delay vs size) and day (1–4). Post-hoc t tests revealed that controls (n = 5) chose short delay and large reward over long delay and small reward trials, respectively (Fig. 4A; control). This was significant in all three trial blocks for the delay manipulation and two out of the three blocks for the size manipulation. For lesions (n = 5), there was no difference between the selection of high and low value outcomes in any trial block during the first two days for either delay (Fig. 4A; lesion) or size (Fig. 4B; lesion) manipulation. Importantly, this was true during the first block of trials for both size and delay suggesting that lesioned rats genuinely had issues with selecting the more valuable option, instead of simply having difficultly reversing contingencies. These results add to the growing body of literature that has proposed different roles for VS during discounting (25, 31–33).

Although there was an initial impact of the lesion on free choice behavior, by the last two testing days for delay and size manipulations, both groups were significantly choosing the high over low value outcomes (t test; p < 0.05). The average over blocks during the last two days is shown in figure 4C. Unlike performance on the first two days of testing (Fig. 4A and B), lesioned rats significantly chose higher value reward in every block during the last two days (t test; p < 0.05).

Overall, lesioned animals were slower and less accurate on forced-choice trials (Fig 4D–I). The multi-factor ANOVA on percent correct and reaction time data produced main effects of lesion and interactions with day and value manipulation (ANOVA; p < 0.05). Post-hoc t tests reveal that percent correct scores were significantly better for both value manipulations across blocks in controls and lesions (Fig. 4D and E, t test; p < 0.05). Finally, in most cases the individual block comparisons between high and low value reaction times were not significantly different, but in all cases rats from both groups tended to be faster under high value conditions (Fig. 4G–I; high versus low).

We conclude that VS lesions do impair the ability to make reward-guided decisions on this task, but this impairment is transitory, possibly reflecting the increased selectivity for odor stimuli and response direction observed in DS after unilateral lesions.

DISCUSSION

The transition from goal-directed to S-R driven behavior is thought to depend on connectivity between VS and DS via midbrain DA neurons (1, 3–11, 34). Chronic drug use amplifies the strength of this transition, shifting the balance of encoding from VS to DS (6, 18–20, 35–37).

Here we show that VS lesions alone, independent of drug use, can enhance stimulus and response selectivity in DS. It is important to note that we are not suggesting that VS lesions mimic structural changes that occur in addiction. However, this procedure does allow us to eliminate value signals generated by VS that we know are diminished after chronic cocaine use (36). The advantage of this study is that we were able to examine changes in DS

selectivity independent from other changes that might occur during acute or chronic drug use (e.g., receptor availability; disruption of other areas).

One interpretation of our data is that encoding in DS is heightened to compensate for the loss of VS function, as has been described for VS after DS lesions (38). During performance of our task, normal animals likely base decisions on a mixture of outcome expectancies and S-R contingencies. Without VS, the rats likely depend more heavily on S-R encoding in DS during decision-making, which allows behavior to recover over several days.

The mechanism by which VS alters encoding in DS after lesions and chronic drug use might reflect abnormal DA signals (6, 10, 11). Indeed, with extended drug self administration there is elevated DA efflux in DS (39–41), which might lead to an excessive stamping in of associations between stimuli and responses. However, recent work shows that increases in DA release in DS during extended periods of self-administration is dependent on VS (42). That is, lesions to VS abolish DA release to cues that predict cocaine. This suggests that lesions of VS do not increase cue-evoked DA release in DS, at least in the context of drug self administration. Whether or not this is true during unexpected reward delivery in a non-drug setting is unknown. Nevertheless, reduced DA signals arriving to DS might evoke compensatory mechanisms that elevate processing of stimuli and responses in brain areas that signal this information.

Regardless of the underlying mechanism, it is clear that VS lesions alone increase stimulus and response encoding in DS. This suggests that during the development of addiction, the transition from VS to DS governed behavior might be due, in part, to initial loss of VS function (43).

Acknowledgments

References

1. van der Meer MA, Redish AD. Ventral striatum: a critical look at models of learning and evaluation. Curr Opin Neurobiol. 2011; 21:387–392. [PubMed: 21420853]

2. Padoa-Schioppa C. Neurobiology of economic choice: a good-based model. Annu Rev Neurosci. 2011; 34:333–359. [PubMed: 21456961]

3. Barto A. Adaptive critics and the basal ganglia. Models of information processing the the basal ganglia. 1995

4. Niv Y, Schoenbaum G. Dialogues on prediction errors. Trends Cogn Sci. 2008; 12:265–272. [PubMed: 18567531]

5. Sutton, RaBAG. Reinforcement Learning: An introduction. Cambridge M: MIT press; 1998.

6. Takahashi Y, Schoenbaum G, Niv Y. Silencing the critics: understanding the effects of cocaine sensitization on dorsolateral and ventral striatum in the context of an actor/critic model. Front Neurosci. 2008; 2:86–99. [PubMed: 18982111]

7. Houk J, Adams JL, Barto AG. A model of how the basal ganglia generate and use neural signals that predict reinforcement. Models of information processing the basal Ganglia. 1995

8. Joel D, Niv Y, Ruppin E. Actor-critic models of the basal ganglia: new anatomical and computational perspectives. Neural Netw. 2002; 15:535–547. [PubMed: 12371510]

9. Redish AD. Addiction as a computational process gone awry. Science. 2004; 306:1944–1947. [PubMed: 15591205]

10. Haber SN, Fudge JL, McFarland NR. Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. J Neurosci. 2000; 20:2369–2382. [PubMed: 10704511]

11. Ikemoto S. Dopamine reward circuitry: two projection systems from the ventral midbrain to the nucleus accumbens-olfactory tubercle complex. Brain Res Rev. 2007; 56:27–78. [PubMed: 17574681]

12. O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ. Dissociable roles of ventral and dorsal striatum in instrumental conditioning. Science. 2004; 304:452–454. [PubMed: 15087550]

13. 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]

14. Koob GF, Volkow ND. Neurocircuitry of addiction. Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology. 2010; 35:217–238. [PubMed: 19710631]

15. Everitt BJ, Robbins TW. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. Nature neuroscience. 2005; 8:1481–1489.

16. Everitt BJ, Robbins TW. From the ventral to the dorsal striatum: Devolving views of their roles in drug addiction. Neuroscience and biobehavioral reviews. 2013

17. Hyman SE, Malenka RC, Nestler EJ. Neural mechanisms of addiction: the role of reward-related learning and memory. Annual review of neuroscience. 2006; 29:565–598.

18. 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]

19. Vanderschuren LJ, Di Ciano P, Everitt BJ. Involvement of the dorsal striatum in cue-controlled cocaine seeking. J Neurosci. 2005; 25:8665–8670. [PubMed: 16177034]

20. Everitt BJ, Dickinson A, Robbins TW. The neuropsychological basis of addictive behaviour. Brain Res Brain Res Rev. 2001; 36:129–138. [PubMed: 11690609]

21. Bryden DW, Burton AC, Kashtelyan V, Barnett BR, Roesch MR. Response inhibition signals and miscoding of direction in dorsomedial striatum. Frontiers in integrative neuroscience. 2012; 6:69. [PubMed: 22973206]

22. McDannald MA, Lucantonio F, Burke KA, Niv Y, Schoenbaum G. Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning. J Neurosci. 2011; 31:2700–2705. [PubMed: 21325538]

23. Singh T, McDannald MA, Haney RZ, Cerri DH, Schoenbaum G. Nucleus Accumbens Core and Shell are Necessary for Reinforcer Devaluation Effects on Pavlovian Conditioned Responding. Front Integr Neurosci. 2010; 4:126. [PubMed: 21088698]

24. Bryden DW, Johnson EE, Diao X, Roesch MR. Impact of expected value on neural activity in rat substantia nigra pars reticulata. Eur J Neurosci. 2011; 33:2308–2317. [PubMed: 21645133]

25. Roesch MR, Bryden DW. Impact of size and delay on neural activity in the rat limbic corticostriatal system. Front Neurosci. 2011; 5:130. [PubMed: 22363252]

26. Bryden DW, Johnson EE, Tobia SC, Kashtelyan V, Roesch MR. Attention for learning signals in anterior cingulate cortex. J Neurosci. 2011; 31:18266–18274. [PubMed: 22171031]

27. Stalnaker TA, Calhoon GG, Ogawa M, Roesch MR, Schoenbaum G. Neural correlates of stimulus-response and response-outcome associations in dorsolateral versus dorsomedial striatum. Front Integr Neurosci. 2010; 4:12. [PubMed: 20508747]

28. Giertler C, Bohn I, Hauber W. Transient inactivation of the rat nucleus accumbens does not impair guidance of instrumental behaviour by stimuli predicting reward magnitude. Behav Pharmacol. 2004; 15:55–63. [PubMed: 15075627]

29. Giertler C, Bohn I, Hauber W. The rat nucleus accumbens is involved in guiding of instrumental responses by stimuli predicting reward magnitude. Eur J Neurosci. 2003; 18:1993–1996. [PubMed: 14622231]

30. Brown VJ, Bowman EM. Discriminative cues indicating reward magnitude continue to determine reaction time of rats following lesions of the nucleus accumbens. Eur J Neurosci. 1995; 7:2479–2485. [PubMed: 8845953]

31. Floresco SB, St Onge JR, Ghods-Sharifi S, Winstanley CA. Cortico-limbic-striatal circuits subserving different forms of cost-benefit decision making. Cogn Affect Behav Neurosci. 2008; 8:375–389. [PubMed: 19033236]

32. Cardinal RN, Pennicott DR, Sugathapala CL, Robbins TW, Everitt BJ. Impulsive choice induced in rats by lesions of the nucleus accumbens core. Science. 2001; 292:2499–2501. [PubMed: 11375482]

33. Acheson A, Farrar AM, Patak M, Hausknecht KA, Kieres AK, Choi S, et al. Nucleus accumbens lesions decrease sensitivity to rapid changes in the delay to reinforcement. Behavioural brain research. 2006; 173:217–228. [PubMed: 16884790]

34. Haber SN, Fudge JL, McFarland NR. Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. The Journal of neuroscience: the official journal of the Society for Neuroscience. 2000; 20:2369–2382. [PubMed: 10704511]

35. Corbit LH, Nie H, Janak PH. Habitual alcohol seeking: time course and the contribution of subregions of the dorsal striatum. Biol Psychiatry. 2012; 72:389–395. [PubMed: 22440617]

36. Takahashi Y, Roesch MR, Stalnaker TA, Schoenbaum G. Cocaine exposure shifts the balance of associative encoding from ventral to dorsolateral striatum. Front Integr Neurosci. 2007; 1:11. [PubMed: 18958239]

37. Belin D, Jonkman S, Dickinson A, Robbins TW, Everitt BJ. Parallel and interactive learning processes within the basal ganglia: relevance for the understanding of addiction. Behavioural brain research. 2009; 199:89–102. [PubMed: 18950658]

38. Nishizawa K, Fukabori R, Okada K, Kai N, Uchigashima M, Watanabe M, et al. Striatal indirect pathway contributes to selection accuracy of learned motor actions. The Journal of neuroscience: the official journal of the Society for Neuroscience. 2012; 32:13421–13432. [PubMed: 23015433]

39. Ito R, Robbins TW, Everitt BJ. Differential control over cocaine-seeking behavior by nucleus accumbens core and shell. Nat Neurosci. 2004; 7:389–397. [PubMed: 15034590]

40. Ito R, Dalley JW, Robbins TW, Everitt BJ. Dopamine release in the dorsal striatum during cocaine-seeking behavior under the control of a drug-associated cue. J Neurosci. 2002; 22:6247–6253. [PubMed: 12122083]

41. Ito R, Dalley JW, Howes SR, Robbins TW, Everitt BJ. Dissociation in conditioned dopamine release in the nucleus accumbens core and shell in response to cocaine cues and during cocaine-seeking behavior in rats. J Neurosci. 2000; 20:7489–7495. [PubMed: 11007908]

42. Willuhn I, Burgeno LM, Everitt BJ, Phillips PE. Hierarchical recruitment of phasic dopamine signaling in the striatum during the progression of cocaine use. Proceedings of the National Academy of Sciences of the United States of America. 2012; 109:20703–20708. [PubMed: 23184975]

43. Takahashi Y, Roesch MR, Stalnaker TA, Schoenbaum G. Cocaine exposure shifts the balance of associative encoding from ventral to dorsolateral striatum. Front Integr Neurosci. 2007; 1
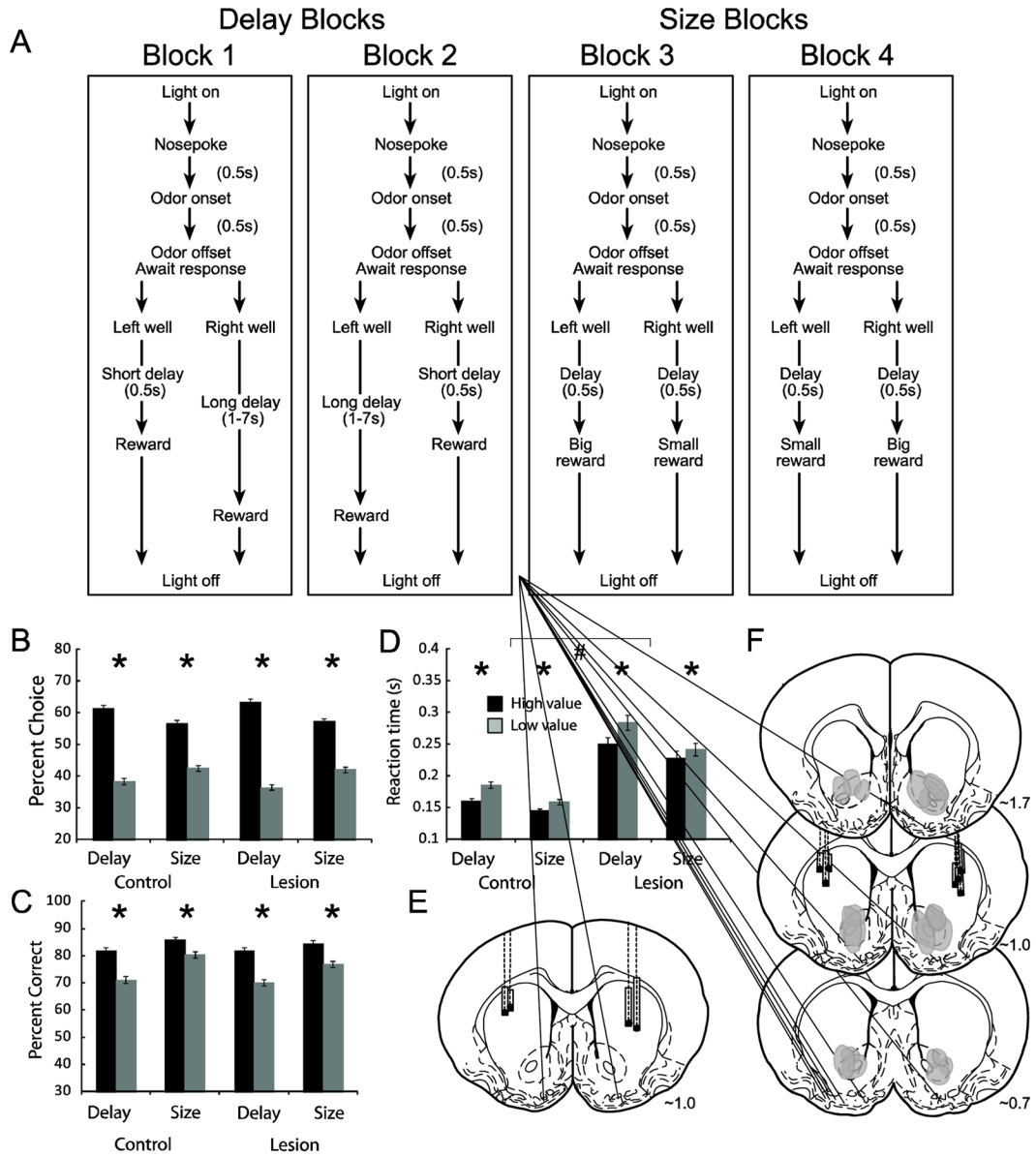
Figure 1.
Task, behavior and recording/lesion locations. **A.** An example of the sequence of events in each trial block. For each recording session, one fluid well was arbitrarily designated as short (500ms delay before reward) and the other designated as long (1–7s delay before reward) (Block 1). After the first block of trials (~60 trials), contingencies unexpectedly reversed (Block 2). With the transition to block 3, the delays to reward were held constant across wells (500ms), but the size of the reward was manipulated. The well designated as 'long' during the previous block now offered 2–3 fluid boli whereas the opposite well offered one bolus. The reward stipulations again reversed in block 4. Free-choice odors signal that either well could be selected for reward, whereas forced-choice odors signaled that reward would only be delivered in the well that the rat was instructed to go to. **B**. The impact of delay length and reward size manipulations on choice behavior during free-choice trials. Percent choice is calculated by taking the number of choices made and divided by the total number of well entries on free-choice trials, multiplied by 100. **C.** Impact of value on

forced-choice trials for short vs. long delay and big vs. small reward. **D.** Reaction times (odor offset to nose unpoke from odor port) on forced-choice trials comparing short vs. long delay trials and big vs. small reward trials. High value = short and large. Low value = long and small. **E–F.** Location of recording sites and unilateral lesions based on histology for sham (E) and lesioned rats (F). Recordings and lesions were performed in the same hemisphere (3 lefts; 4 rights). Filled gray boxes mark the locations of electrodes based on histology and initial recording site. Black dot marks the bottom of the recording tract. Transparent gray areas mark lesions for each animal. Shown are representative slices at 1.7, 1.0 and 0.7 anterior to bregma taken from Paxinos and Watson (1997). Asterisks indicate planned comparisons revealing statistically significant differences (t test, *p*<0.05). # indicates a main effect of lesion in the ANOVA (p < 0.05). Error bars indicate standard error of the mean (SEM).
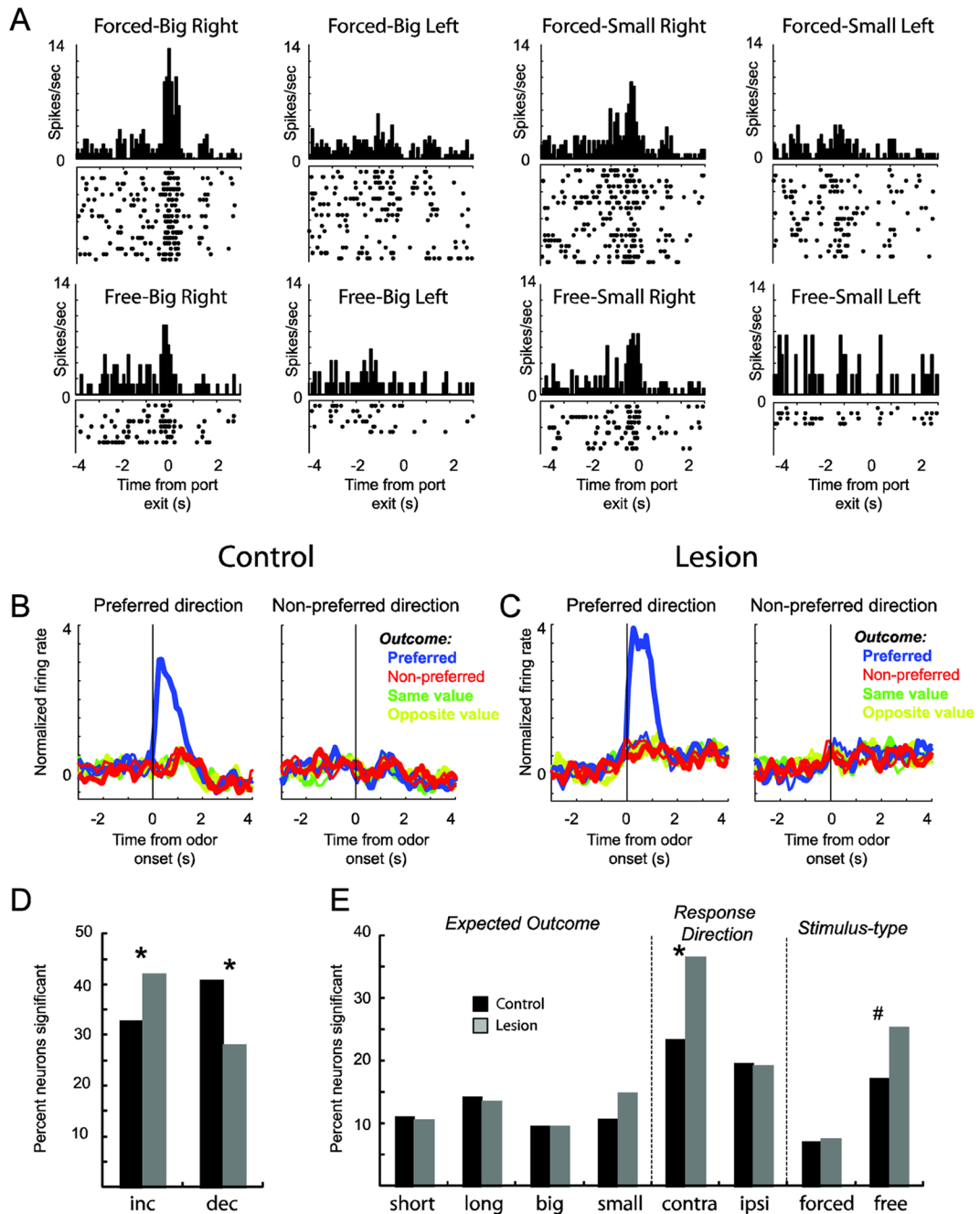
Figure 2. Stimulus and response encoding in DS was enhanced after VS lesions
A. Single cell example showing activity during size blocks during forced (top row) and free (bottom row) choice performance. **B–C**. Average firing rate over all 352 and 457 neurons for controls (b) and lesions (c) for free- or forced-choices trials, depending on which elicited the strongest response. All trials are referenced to the trial-type that elicited the maximal firing during the time starting 100 ms after odor onset to entry into the fluid well. See text for more detail. Data were normalized by subtracting the mean and dividing by the standard deviation (Z score). Blue = preferred outcome; red = non-preferred outcome; green = same value and yellow = opposite value; thin = first 5 trials; thick = last 5 trials. **D**. Percent of neurons that showed significant increases or decreases, averaged across trial-type, during the

analysis epoch compared to baseline (t test; $p < 0.05$). **E.** Height of each bar indicates the percent of neurons that showed a main effect or interaction effect of outcome (short, long, big and small), direction (contra and ipsilateral) and odor-type (free and forced choice odors). * = $p < 0.05$; # = 0.11; chi-square.
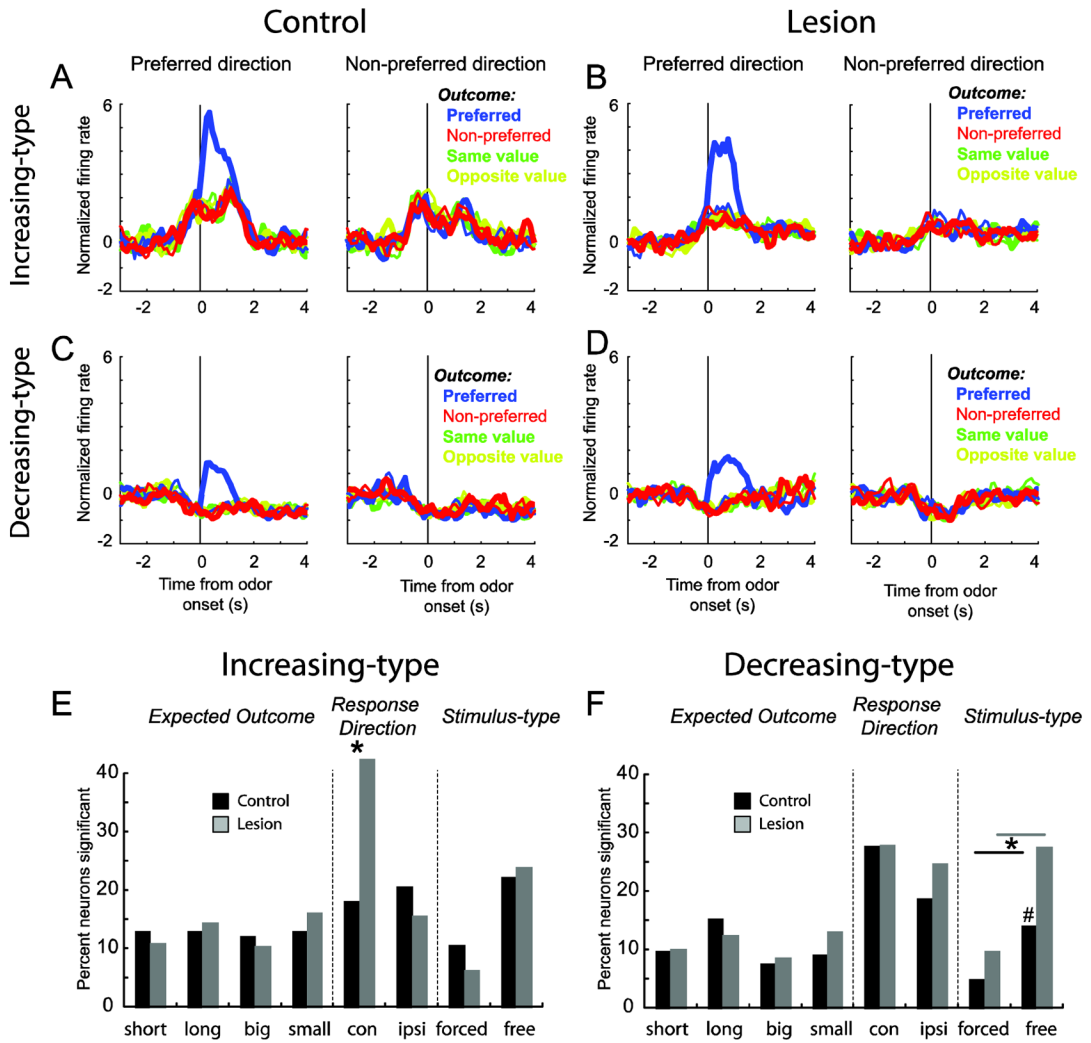
Figure 3. Stimulus and response encoding in decreasing- and increasing-type DS neurons was enhanced after VS lesions

A–D. Average firing rate over all increasing (A–B) and decreasing (C–D) type neurons for controls (A and C) and lesions (B and D) for free- or forced-choice trials, depending on which elicited the strongest response. All trials are referenced to the trial-type that elicited the maximal firing during the time starting 100 ms after odor onset to entry into the fluid well. See text for more detail. Data were normalized by subtracting the mean and dividing by the standard deviation (Z score). Blue = preferred outcome; red = non-preferred outcome; green = same value and yellow = opposite value; thin = first 5 trials; thick = last 5 trials. **E–F**. Height of each bar indicates the percent of neurons that showed a main or interaction effect of outcome (short, long, big and small), direction (contra and ipsilateral) and odor-type (free and forced choice odors). * p < 0.05; # = 0.057; chi-square.
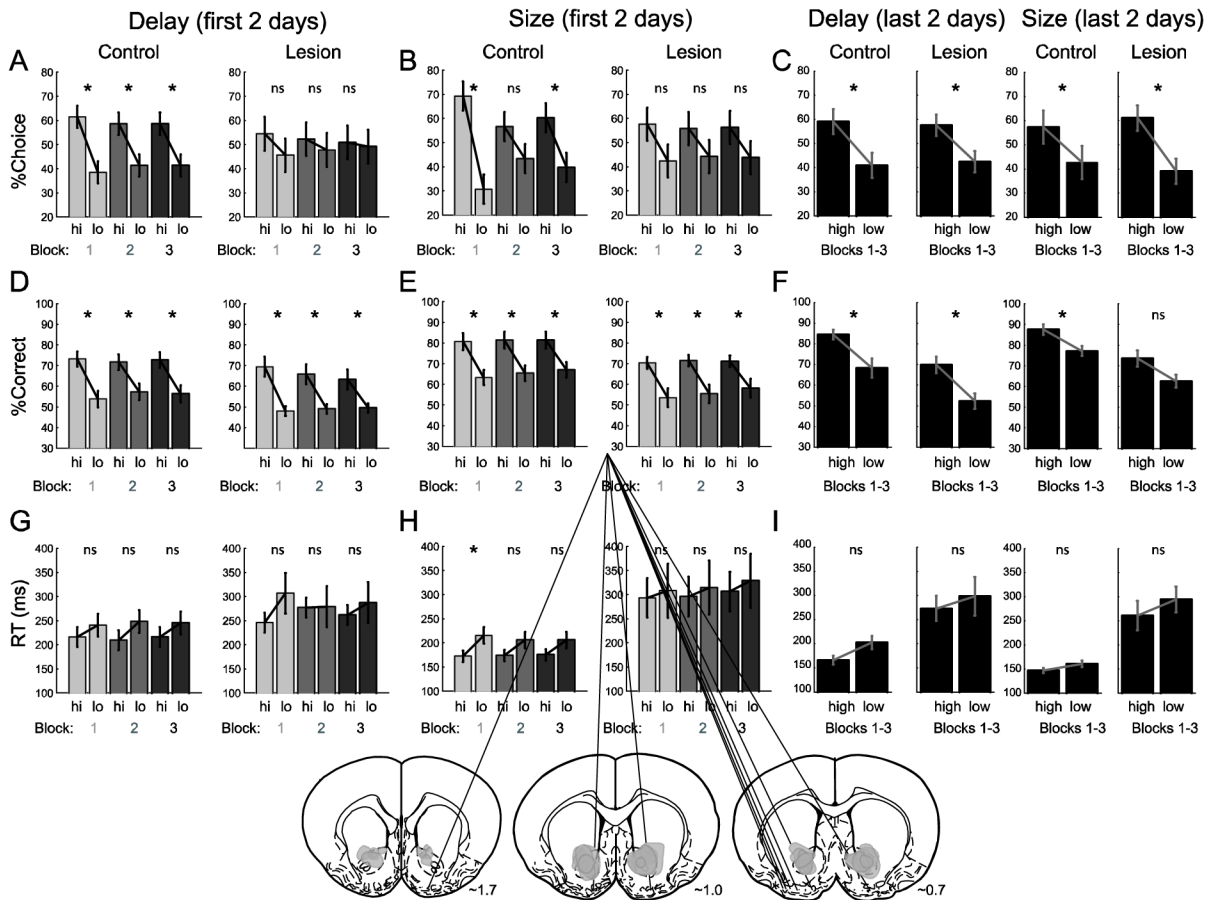
Figure 4. VS lesions caused temporary impairments of reward-guided decision-making
Percent choice on free-choice trials (top row), percent correct scores on forced-choice trials (second row) and reaction time on forced-choice trials (bottom row; port exit minus odor offset) for controls (n = 5) and lesions (n = 5) during the first and last 2 days of testing. Each day rats performed 3 trial blocks of either size or delay across 4 days for each manipulation. During the first 2 days of testing scores were broken down by the three blocks to demonstrate that VS lesions impaired all three trial blocks. Gray areas mark lesions for each animal. Shown are representative slices at 1.7, 1.0 and 0.7 anterior to bregma taken from Paxinos and Watson (1997). High (hi) = short and large. Low (lo) = long and small. Asterisks indicate planned comparisons revealing statistically significant differences (t test, *p*<0.05). Error bars indicate SEM.