Dorsolateral striatum engagement during reversal learning

Hadley C. Bergstrom, Abby G. Lieberman, Carolyn Graybeal, Anna M. Lipkin, and Andrew Holmes

Laboratory of Behavioral and Genomic Neuroscience, National Institute on Alcoholism and Alcohol Abuse (NIAAA), National Institutes of Health (NIH), Bethesda, Maryland 20852, USA

Most experimental preparations demonstrate a role for dorsolateral striatum (DLS) in stimulus-response, but not outcome-based, learning. Here, we assessed DLS involvement in a touchscreen-based reversal task requiring mice to update choice following a change in stimulus-reward contingencies. In vivo single-unit recordings in the DLS showed reversal produced a population-level shift from excited to inhibited neuronal activity prior to choices being made. The larger the shift, the faster mice reversed. Furthermore, optogenetic photosilencing DLS neurons during choice increased early reversal errors. These findings suggest dynamic DLS engagement may facilitate reversal, possibly by signaling a change in contingencies to other striatal and cortical regions.

[Supplemental material is available for this article.]

The dorsolateral striatum (DLS) is key neural locus for stimulus-response learning and habit formation, and DLS dysfunction is implicated in the pathophysiology of addictions (White 1996; Everitt and Robbins 2005; Yin and Knowlton 2006; Graybiel 2008; Balleine and O’Doherty 2010). In rodent instrumental learning paradigms, extensive training and certain schedules of reinforcement (random interval) favor the generation of DLS-dependent habitual behaviors (Adams 1982; Dickinson 1985) and these behaviors are expedited by alcohol and other drugs of abuse (Nelson and Killcross 2006; Nordquist et al. 2007; DePoy et al. 2013; Gremel and Lovinger 2017).

The involvement of DLS later in training is posited to reflect its role in constructing slow-to-acquire, stimulus-driven motor sequences (Balleine et al. 2009; Dezfouli and Balleine 2012). This contribution contrasts with that of the neighboring dorsomedial striatum (DMS), which is involved in cognitive flexibility and other processes that entail the updating of behavior when outcome values change (Packard and Knowlton 2002; Balleine et al. 2007; Ragozzino 2007; Hart et al. 2018). Hence, current models propose that outcome-sensitive learning, mediated by the DMS and prefrontal cortical regions, dominates performance early in training, and is supplanted by DLS-mediated stimulus-bound performance with further training (Dickinson and Balleine 1995).

Supporting this model, response-related neuronal activity in the caudate (homologue of DMS) emerges before activity in the putamen (homologue of DLS) of nonhuman primates learning a visual learning task (Williams and Eskandar 2006) and the development of habit in human subjects coincides with an increase in BOLD signal in the dorsolateral posterior putamen (Tricomi et al. 2009). Moreover, drug-seeking is disrupted by inactivation of the DLS, but not DMS, following prolonged, but not limited, training in rodents (Zapata et al. 2010; Corbit et al. 2012). However, DLS single-unit activity is evident from the beginning of motor, maze-based and instrumental training (Jog et al. 1999; Barnes et al. 2005; DeCoteau et al. 2007; Tang et al. 2007; Kimchi et al. 2009; Vandaële et al. 2019). Furthermore, lesions or optogenetic inactivation of the DLS in rodents can expedite early acquisition of spatial and stimulus discriminations when choice-outcome associations are still being formed (Bradfield and Balleine 2013; Bergstrom et al. 2018) and pharmacological inactivation of the anterior putamen impair visual reversal learning in marmosets (Jackson et al. 2019).

These prior findings suggest the DLS may exert a greater influence over outcome-based performance than currently appreciated. This led us, in the current study, to test for a potential role for the DLS in outcome-based reversal learning in which mice are required to update choice based on a new stimulus-reward contingency. To that end, we performed in vivo recordings and optogenetic manipulations of the DLS in a touchscreen-based reversal paradigm, a task previously shown to recruit dorsal and ventral striatum (Brigman et al. 2013; Bergstrom et al. 2018; Piantadosi et al. 2019; Radke et al. 2019).

We trained male C57BL/6J mice to learn to accurately discriminate (>85% correct choice) between two visual stimuli presented on a touchscreen sensitive to obtain a food pellet, and then reversed the stimulus-reward contingency and retrained mice to criterion (Fig. 1A,B). In parallel, single-unit DLS activity was recorded via chronically implanted microelectrode arrays (Fitzgerald et al. 2014; Gunduz-Cinar et al. 2019; Halladay et al. 2020) during three test sessions: (1) the final (late) discrimination (4D), (2) the first (early) reversal (4R), and (3) the final (late) reversal (5R) session.

As expected from the results of previous studies in our laboratory, percent correct choice was high at 1D and 4R, but low at 5R, while total errors showed the inverse pattern (Fig. 1C,D). Perseverative errors (an error following an error) were highest at 4R; nonperseverative errors (an error following a correct choice) were also elevated (Fig. 1E,F). Choice and reward-collection and latencies were lowest at 5R (Supplemental Fig. S1). Overall, these
behavioral profiles resemble those previously reported in our laboratory (Izquierdo et al. 2006; Graybeal et al. 2014).

Next, we aligned the in vivo recording data obtained from 203 (n = 64–72 per session) DLS units (average firing rate = 5.95 ± 0.44 Hz) to the time a correct choice or error was made during each recording session (without attempting to segregate neurons into medium spiny neurons versus fast-spiking interneurons). We classified units into those exhibiting phasically increased (“excited”) or decreased (“inhibited”) activity in the 2-sec epoch either prior to or immediately after a choice was made (Fig. 2A). Then, to test whether the recruitment of these units changed in association with performance, we compared the proportion of prechoice and postchoice excited and inhibited activity across the three recording sessions.

Most choice-related neurons were excited during 1D, irrespective of whether their activity was evident pre- or postchoice (Fig. 2B). Strikingly, however, at 4R there was a marked shift from excited to inhibited activity, and specifically so during the prechoice (not postchoice) period. Notably, this preponderance of inhibited activity was largely maintained, though somewhat attenuated, at 1R even though behavioral performance at this stage was comparable to that at 1D, and markedly different from that at 4R. Thus, the shift to prechoice inhibited activity at 4R is not simply due to the preponderance of erroneous responding but, rather, reflects the new task contingencies. In fact, the maintenance of inhibited activity at 4R suggests that DLS units signal a “record” of the change in contingencies throughout reversal even as performance resolves to high-choice accuracy.

The finding that DLS unit activity is sensitive to a change in stimulus-reward contingencies with reversal suggests the DLS may play a dynamic role in modifying behavior in response to the contingency reversal. To explore this possibility in more detail, we asked whether DLS activity at 4R correlated with measures of reversal performance. This revealed a significant inverse relationship between the proportion of DLS units exhibiting prechoice inhibited activity at 4R and the number of sessions to reach reversal criterion (Pearson correlation coefficient r = −0.61, Fisher’s test: P < 0.05); i.e., the more DLS units that were phasically inhibited prior to choice at early reversal, the faster mice eventually relearned the new task contingencies (Fig. 2C). Importantly, further analysis showed this relationship was specific, in that neither prechoice excited activity (r = 0.06, P > 0.05) nor postchoice inhibited (r = 0.30, P > 0.05) or excited (r = −0.41, P > 0.05) activity correlated with sessions to reverse (Supplemental Fig. S2). The degree of inhibited activity at 4R was also unrelated to the number of discrimination sessions to criterion (r = −0.34, P > 0.05), suggesting the correlation with sessions to reverse does not reflect a relationship with learning more generally. In this context, we have previously shown that measures of discrimination and reversal learning load on separate principal components (Izquierdo et al. 2006; Graybeal et al. 2014).

To further investigate the relationship between DLS neuronal activity at 4R and reversal performance, we split mice into slow (n = 6) and fast (n = 5) learners based on a median split (=13 sessions) of the average number of sessions to attain reversal criterion (Fig. 2D). Slow learners made significantly more errors throughout reversal than fast learners, but did not differ in latency to make a choice or collect the reward, consistent with a cognitive but not motivational deficit in this group (Fig. 2E,F). Of note, the two groups were similar in their rates of discrimination learning.

Figure 1. DLS neuronal recordings in a pairwise discrimination and reversal touchscreen task. (A) Mice had microarray electrode arrays unilaterally implanted in the DLS (dots depict estimated location of arrays) and were trained to discriminate between two visual stimuli (“fan” and “marble”) presented on a touch-sensitive screen to obtain a food pellet. The stimulus-reward contingency was then reversed, and mice retrained to criterion. DLS neuronal activity was recorded on three sessions: late discrimination (LD), early reversal (ER), and late reversal (LR). (C) Percent correct choice was high at 1D and 4R, and low at 4R (repeated ANOVA stage effect: F(2,24) = 129.56, P < 0.01, followed by Newman-Keuls post hoc tests). Errors were low at 1D and 4R, and high 4R, whether measured as total errors (F(2,24) = 26.41, P < 0.01) (D), perseverative errors (an error following an error) (F(2,24) = 26.33, P < 0.01) (E) or nonperseverative errors (an error following a correct) (F(2,24) = 15.43, P < 0.01) (F). For corresponding choice and reward-collection latencies, see Supplemental Figure S1. A = 13 mice. Data are means ± SEM. (*) P < 0.05.
(overall = 12.0 ± 2.1 SEM, slow = 10.0 ± 1.3, fast = 13.5 ± 6.6 sessions to criterion, t-test: P < 0.05), again indicating the differences in reversal were specific to that problem. We then compared DLS unit activity at $R^*$ between the two groups. There was a significantly smaller proportion of prechoice inhibited DLS units in the slow, relative to the fast, learner group ($t_{(9)} = 3.49$, $P < 0.01$) (Fig. 2G). These data are again consistent with a positive relationship between inhibited activity in DLS neurons at early reversal and the efficiency of subsequent reversal learning.

The results of our recording data led us to perform a causal test for the contribution of the DLS to reversal performance. To do so, we used in vivo optogenetics to silence DLS neurons as mice performed the task (Fig. 3A). Stereotaxic surgery was used to bilaterally transfet DLS neurons with adeno-associated virus containing the inhibitory opsin, archaerhodopsin (ArchT, rAAV8/CAG-ArchT-GFP) or a control construct (rAAV8/CAG-GFP), and chronically implant optic-fibers directed at the DLS (Bergstrom et al. 2018; Sengupta and Holmes 2019). Green light (561 nm, 7 mW) was shone on the DLS during each trial throughout reversal, beginning at the initiation of the trial through to reward collection (correct trials) or 3-sec postchoice (error trials) (light on duration averaged 7–10 sec). In accordance with our previous study of DLS-photosilencing effects on discrimination (Bergstrom et al. 2018), performance was analyzed by binning sessions into mutually exclusive early, mid and late phases. This was done by equally subdividing the number of sessions to criterion for each mouse into the first, second, and third groups of sessions (e.g., nine sessions to criterion = three early, three mid, three late sessions). For session totals indivisible by three, the additional session(s) accrued to the late, then mid phase. The cumulative values in each phase...
were calculated and either expressed as a total value (errors) or average by phase (%correct and latencies).

We found that DLS-silencing did not affect the number of sessions to attain reversal criterion or percent correct responding during any stage of reversal (Fig. 3B,C) (choice and reward-collection latencies were also unaffected, Supplemental Fig. S3A, B). Photosilencing did, however, significantly increase the total number of errors ($P<0.05$), specifically (at least statistically) the number of perseverative errors ($P<0.05$), at $1^\text{st}$R but not the later reversal stages (Fig. 3D–F). Given the specificity of the effect to the early testing stage, this silencing-induced increase in errors cannot be due to a nonspecific increase in touchscreen responding. Instead, these data are consistent with an impairment in early reversal performance as a result of DLS silencing. This deficit likely failed to manifest as a decrease in overall percent correct responding because this measure was already near floor levels in GFP controls. As such, these data are consistent with a recent study demonstrating that pharmacological inhibition of the anterior putamen causes visual reversal impairments in nonhuman primates (Jackson et al. 2019), and contrast with the ability of DLS photosilencing to facilitate early discrimination (Bergstrom et al. 2018).

How do the results of these optogenetic manipulations align with our recording data? One possibility is that the shift to inhibited activity in DLS units prior to early reversal choices reflects the suppression of the old contingency to enable formation of the new. If this were the case, however, DLS silencing would be predicted to result in a facilitation of reversal performance (e.g., by removing competition between the old and new memories, or liberating other structures, such as the DMS, to support reversal learning unhindered), rather than the deficit observed. An alternative interpretation is that the change in DLS unit activity signals that the contingencies have changed, and this signal is utilized to update behavior to reflect the new contingency. This would explain the positive relationship between size of the inhibited DLS population at early reversal and the subsequent rate of reversal, as well as the impairment in early reversal caused by DLS silencing. However, given the effect of silencing was limited to the early reversal stage, this putative contribution of the DLS is not necessary for the contingencies to be reversed in full. Instead, it may serve to catalyze the ability of other regions to fulfill this function, such as the DMS and PFC. In turn, this predicts that augmenting DLS activity, e.g., via optogenetic photostimulation, could expedite early reversal learning.

In sum, the current data provide preliminary support for a contribution of the DLS to reversal learning in mice, such that a dynamic increase in the prechoice inhibition of DLS units during a change in stimulus-reward contingencies may facilitate the ability to modify behavior accordingly. These findings add to a growing body of evidence supporting a contribution of the DLS to forms of learning that utilize outcome information to guide behavior (Dezfouli et al. 2014; Bergstrom et al. 2018; Vandaële et al. 2019).

Acknowledgments

Research supported by the NIAAA Intramural Research Program.

References

Adams CD. 1982. Variations in the sensitivity of instrumental responding to reinforcer devaluation. J Exp Psychol 34: 77–98. doi:10.1080/14640748208400878

Balleine BW, O’Doherty JP. 2010. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. Neuropsychopharmacology 35: 48–69. doi:10.1038/npp.2009.131

Balleine BW, Delgado MR, Hikosaka O. 2007. The role of the dorsal striatum in reward and decision-making. J Neurosci 27: 8161–8165. doi:10.1523/JNEUROSCI.1554-07.2007

Balleine BW, Liljeholm M, Ostlund SB. 2009. The integrative function of the basal ganglia in instrumental conditioning. Behav Brain Res 199: 43–52. doi:10.1016/j.bbr.2008.10.034

Barnes TD, Kubota Y, Hu D, Jin DZ, Graybiel AM. 2005. Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. Nature 437: 1158–1161. doi:10.1038/nature04053

Bergstrom HC, Lipkin AM, Lieberman AG, Pinard CR, Gunduz-Cinar O, Brockway ET, Taylor WW, Nonaka M, Bukalo O, Wills TA et al. 2018.
Dorsolateral striatum engagement interferes with early discrimination learning. Cell Rep 23: 2264–2272. doi:10.1016/j.celrep.2018.04.081
Bradfield LA, Balleine BW. 2013. Hierarchical and binary associations compete for behavioral control during instrumental biconditional discrimination. J Exp Psychol Anim Behav Process 39: 2–13. doi:10.1037/a0030941
Brigman JL, Daut RA, Wright T, Gunduz-Cinar O, Graybeal C, Davis MI, Jiang Z, Saksida LM, Jinde S, Pease M, et al. 2013. GluN2B in corticostriatal circuits governs choice learning and choice shifting. Nat Neurosci 16: 1101–1110. doi:10.1038/nn.3457
Corbit LH, Nie H, Janak PH. 2012. Habitual alcohol seeking: time course and the contribution of subregions of the dorsal striatum. Biol Psychiatry 72: 389–395. doi:10.1016/j.biopsych.2012.02.024
DeCoteau WE, Thorn C, Gibson DJ, Courtencanche R, Mitra P, Kubota Y, Graybiel AM. 2007. Oscillations of local field potentials in the rat dorsal striatum during spontaneous and instructed behaviors. J Neurophysiol 97: 3803–3805. doi:10.1152/jn.00108-2007
DePoy L, Daut R, Brigman JL, MacPherson K, Crowley N, Gunduz-Cinar O, Dickson A, Balleine BW. 2014. Habits as action sequences: Dorsal striatum erodes with extended training. Proc Natl Acad Sci U S A 110: 14783–14788. doi:10.1073/pnas.1308198110
Dezfouli A, Lingawi NW, Balleine BW. 2014. Habits, action sequences and reinforcement learning. Eur J Neurosci 38: 1036–1051. doi:10.1111/ejns.2012.00850.x
Dezfouli A, Balleine BW. 2012. Habits, action sequences and reinforcement learning. Eur J Neurosci 38: 1036–1051. doi:10.1111/ejns.2012.00850.x
Dickinson A, Balleine BW. 1995. Overtraining and the motivational control of instrumental action. Anim Learn Behav 22: 197–206. doi:10.3758/BF03199935
Everitt BJ, Robbins TW. 2005. Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. Nat Neurosci 8: 1481–1489. doi:10.1038/nn1579
Fitzgerald PJ, Whittle N, Flynn SM, Graybeal C, Pinard CL, Cinar R, Saksida LM, Kunos G, et al. 2013. Chronic alcohol produces neureadaptations to prime dorsal striatal learning. Proc Natl Acad Sci U S A 110: 14783–14788. doi:10.1073/pnas.1308198110
Graybiel AM. 2007. Oscillations of local potentials in the rat dorsal striatum. Science.286.5445.1745–1749. doi:10.1126/science.286.5445.1749
Kimchi EV, Torregrossa MM, Taylor JR, Laubach M. 2009. Neuronal correlates of instrumental learning in the dorsal striatum. J Neurophysiol 102: 475–489. doi:10.1152/jn.00262.2009
Nelson A, Killcross S. 2006. Amphetamine exposure enhances habit formation. J Neurosci 26: 3805–3812. doi:10.1523/JNEUROSCI.4305-05.2006
Norquist RE, Voon P, de Moolij-van Malsen JG, Joosten RN, Pennartz CM, Vanderschuren LJ. 2007. Augmented reinforcer value and accelerated habit formation after repeated amphetamine treatment. Eur Neuropsychopharmacol 17: 532–540. doi:10.1016/j.euro.2006.12.005
Packard MG, Knowlton BJ. 2002. Learning and memory functions of the basal ganglia. Annu Rev Neurosci 25: 563–593. doi:10.1146/annurev.neuro.25.112701.142937
Packard PT, Lieberman AG, Pickens CL, Bergstrom HC, Holmes A. 2019. A novel multichoice touchscreen paradigm for assessing cognitive flexibility in mice. Learn Mem 26: 24–30. doi:10.1177/1077709818792288
Radke AK, Kocharian A, Covey DP, Lovinger DM, Chee JR, Mateo Y, Holmes A. 2019. Contributions of nucleus accumbens dopamine to cognitive flexibility. Eur J Neurosci 50: 2023–2035. doi:10.1111/ejn.14152
Ragozzino ME. 2007. The contribution of the medial prefrontal cortex, orbitofrontal cortex, and dorsomedial striatum to behavioral flexibility. Ann N Y Acad Sci 1121: 355–375. doi:10.1196/annals.1401.013
Sengupta A, Holmes A. 2019. A discrete dorsal raphe to basal amygdala 5-HT circuit calibrates aversive memory. Neurosci 103: 489–505 e487. doi:10.1016/j.neuro.2019.05.029
Tang C, Pawlik AP, Prokopenko V, West MO. 2007. Changes in activity of the striatum during formation of a motor habit. Eur J Neurosci 25: 1212–1227. doi:10.1111/j.1460-9586.2007.05353.x
Tricomi E, Balleine BW, O’Doherty JP. 2009. A specific role for posterior dorsolateral striatum in human habit learning. Eur J Neurosci 29: 2225–2232. doi:10.1111/j.1460-9586.2009.06796.x
Vandaele Y, Mahajan NR, Ottenheimer DJ, Richard JM, Mysore SP, Janak PH. 2019. Distinct recruitment of dorsomedial and dorsolateral striatum erodes with extended training. Elife 8: e49536. doi:10.7554/eLife.49536
White NM. 1996. Addictive drugs as reinforcers: multiple partial actions on cognitive flexibility. Eur J Neurosci 8: 81–94. doi:10.1111/j.1460-9586.1996.tb03586.x
Williams ZM, Eskandar EN. 2006. Selective enhancement of associative learning by microstimulation of the anterior caudate. Nat Neurosci 9: 1121–1127. doi:10.1038/nn1579
Woolf CJ, Mackay CE, Salter MW. 2012. Scientific opportunity: the end of neuroplasticity? Nat Rev Neurosci 13: 348–353. doi:10.1038/nrn3139
Zapata A, Minney VL, Shippenberg TS. 2010. Shift from goal-directed to habitual cocaine seeking after prolonged experience in rats. J Neurosci 30: 15457–15463. doi:10.1523/JNEUROSCI.4072-10.2010
Received April 3, 2020; accepted in revised form June 19, 2020.