



Published in final edited form as:

Cell Rep. 2018 May 22; 23(8): 2264–2272. doi:10.1016/j.celrep.2018.04.081.

## Dorsolateral Striatum Engagement Interferes with Early Discrimination Learning

Hadley C. Bergstrom<sup>1,5,\*</sup>, Anna M. Lipkin<sup>1</sup>, Abby G. Lieberman<sup>1</sup>, Courtney R. Pinard<sup>1</sup>, Ozge Gunduz-Cinar<sup>1</sup>, Emma T. Brockway<sup>1</sup>, William W. Taylor<sup>1</sup>, Mio Nonaka<sup>1</sup>, Olena Bukalo<sup>1</sup>, Tiffany A. Wills<sup>2</sup>, F. Javier Rubio<sup>3</sup>, Xuan Li<sup>3</sup>, Charles L. Pickens<sup>1</sup>, Danny G. Winder<sup>2,4</sup>, and Andrew Holmes<sup>1</sup>

<sup>1</sup>Laboratory of Behavioral and Genomic Neuroscience, National Institute on Alcohol Abuse and Alcoholism, NIH, Bethesda, MD, USA

<sup>2</sup>Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN, USA

<sup>3</sup>Behavioral Neuroscience Research Branch, National Institute on Drug Abuse, NIH, Baltimore, MD, USA

<sup>4</sup>Vanderbilt Center for Addiction Research, Nashville, TN, USA

### SUMMARY

In current models, learning the relationship between environmental stimuli and the outcomes of actions involves both stimulus-driven and goal-directed systems, mediated in part by the DLS and DMS, respectively. However, though these models emphasize the importance of the DLS in governing actions after extensive experience has accumulated, there is growing evidence of DLS engagement from the onset of training. Here, we used *in vivo* photosilencing to reveal that DLS recruitment interferes with early touchscreen discrimination learning. We also show that the direct output pathway of the DLS is preferentially recruited and causally involved in early learning and find that silencing the normal contribution of the DLS produces plasticity-related alterations in a PL-DMS circuit. These data provide further evidence suggesting that the DLS is recruited in the construction of stimulus-elicited actions that ultimately automate behavior and liberate cognitive resources for other demands, but with a cost to performance at the outset of learning.

### In Brief

This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

\*Correspondence: habergstrom@vassar.edu.

<sup>5</sup>Lead Contact

#### SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures and six figures and can be found with this article online at <https://doi.org/10.1016/j.celrep.2018.04.081>.

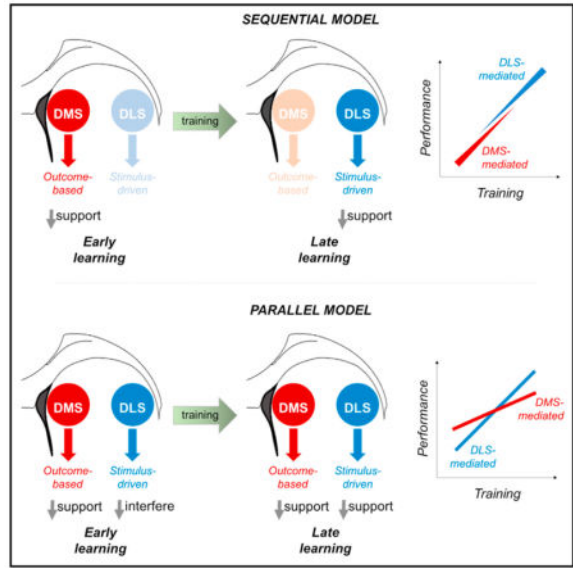
#### AUTHOR CONTRIBUTIONS

H.C.B., A.M.L., A.G.L., C.R.P., O.B., O.G.-C., E.T.B., W.W.T., M.N., T.A.W., and C.L.P. collected and analyzed data. A.H. and D.G.W. designed experiments. F.J.R. and X.L. provided training. All authors contributed to the writing of the manuscript.

#### DECLARATION OF INTERESTS

The authors declare no competing interests.

What is the contribution of the DLS in early discrimination learning? Bergstrom et al. show using in vivo optogenetics, fluorescence in situ hybridization, and brain-wide activity mapping that silencing the DLS facilitates early discrimination learning, drives activity in a parallel PL-DMS circuit, and preferentially recruits the DLS “direct” output pathway.



## INTRODUCTION

The dorsolateral striatum (DLS) and dorsomedial striatum (DMS) support the ability to learn to detect, discriminate, and engage with environmental stimuli that lead to rewards and avoid those that do not. Current theories emphasize DMS support of goal-directed learning that is rapidly acquired and responsive to changes in choice outcomes, gradually supplanted by DLS-mediated, stimulus-driven, habitual performance as experience accrues and choice outcome becomes predictable (Balleine et al., 2009; Corbit et al., 2012; Friend and Kravitz, 2014; Graybiel, 2008; Surmeier, 2013).

In now classic studies, DLS lesions or inactivation prevented the development of outcome-impervious habit behavior in a rat lever-press operant task, whereas conversely, DMS inactivation or NMDA receptor blockade rendered rats insensitive to changes in outcome value in a two-choice, two-outcome instrumental lever-press task (Yin et al., 2005a, 2005b, 2006). The implication of these studies, that the DMS and DLS generate parallel forms of learning in tandem, is elegantly demonstrated by work showing that learning context can gate the ability of mice to toggle between goal-directed and habitual performance (Gremel and Costa, 2013). Moreover, there are data from a wide range of tasks showing that the DLS exhibits robust task-related activity from early training (Brigman et al., 2013; Kim et al., 2013; Kimchi et al., 2009; Stalnaker et al., 2010; Thorn et al., 2010).

Indeed, it has long been posited that stimulus-driven and goal-directed systems develop in tandem and can varyingly compete or cooperate for control over actions at certain points in learning (Balleine and Ostlund, 2007; Balleine et al., 2007; Bradfield and Balleine, 2013;

Corbit and Janak, 2007; Daw et al., 2005; Dickinson and Balleine, 1995; Smith and Graybiel, 2013; Thorndike, 1911; Vicente et al., 2016; Yin et al., 2006). Some of the clearest evidence supporting this model stems from the observation that DLS lesioning can in some instances expedite learning (Bradfield and Balleine, 2013), possibly by shifting control to goal-directed learning systems (for related discussion, see Daw et al., 2005; Yin et al., 2004).

Despite these prior findings, the contribution of the DLS to learning remains to be described in certain choice paradigms, using techniques that allow control of neuronal activity specifically when choices are made (as opposed to pre-test lesions and inactivations). In the present study, we sought to extend the literature by examining the contribution of the DLS to performance on a visual discrimination task using a touchscreen-based platform that has close analogs in human cognitive test batteries and therefore lends itself to translational studies of cognitive impairment in psychiatric illness (Horner et al., 2013). We hypothesized that attenuating the activity of the DLS, using optogenetics, around the time mice were making responses in a discrimination learning task would alter performance at specific, most likely early, stages of learning.

Our findings show that photosilencing DLS neurons as mice acquired discriminated choices expedites early learning and improves the use of information about prior choice outcomes. We also show that the direct output pathway of the DLS is preferentially recruited and causally involved in choice learning, while the indirect output pathway exerts a significant, though less prominent, influence on performance. Furthermore, through analysis of systems-wide patterns of *Arc* expression at different stages of learning, we reveal how DLS photosilencing produces adaptations in a prelimbic cortex-DMS circuit known to act as a substrate for goal-directed learning.

Taken together, these data suggest the DLS may begin forming stimulus-driven associations from the earliest stages of learning of a discriminated choice, but with the cost of temporarily retarding performance, in line with current theories (Bradfield and Balleine, 2013; Daw et al., 2005). More generally, our findings add to a growing body of evidence that the DLS and DMS are engaged in a parallel, rather than sequential, manner as learning develops, with potential implications for understanding how dorsal striatal dysfunction contributes to neuropsychiatric conditions such as addiction.

## RESULTS AND DISCUSSION

### DLS Photosilencing Expedites Early Learning

To examine the role of DLS in choice learning, we used a task wherein mice learned, in a stage-wise manner, to respond to one of two visual stimuli on a touch-sensitive monitor (DePoy et al., 2013; Graybeal et al., 2011; Izquierdo et al., 2006). A biased design was adopted, in which mice were rewarded for choosing the stimulus (“fan”) that was visually non-preferred (30%–40% unconditioned choice versus 60%–70% unconditioned choice for the unrewarded “marble” stimulus) at the start of training (Figure 1A). Mice learned to reliably (>85% correct in two consecutive sessions) obtain food reward within  $\sim 11 \pm 1$  daily sessions ( $583 \pm 37$  trials) (Figures 1B and 1C). To silence DLS neuronal activity in this task,

we virally expressed the light-gated opsin arch-aerhodopsin (ArchT; rAAV8/CAG-ArchT-GFP) and confirmed the efficacy of the virus using *ex vivo* slice electrophysiology (Figures 1D and 1E). Behaviorally, we first showed that shining green light in the DLS did not alter performance in an open field or in a real-time place-preference assay, measures of motor and reward-related behavior, respectively (Figures 1F, 1G, and S1).

In the learning task, we found that photosilencing the DLS in each trial, as mice made a choice on the touchscreen (light on from stimuli presentation through reward collection < 10 s on average), significantly improved correct responding and reduced error rates during early learning (and quickened late-stage response latency), without affecting reward retrieval latencies or overall response number at any stage (Figures 1H–1J and S1). In fact, the effects of silencing were already evident at the very start of testing, with a significantly lower error rate in the photosilenced group in both the first (GFP =  $72.4 \pm 4.6$  errors, ArchT =  $55.4 \pm 3.8$  errors;  $t[14] = 2.9$ ,  $p < 0.05$ ) and second (GFP =  $70.3 \pm 5.7$  errors, ArchT =  $45.4 \pm 5.7$  errors;  $t[14] = 3.0$ ,  $p < 0.01$ ) sessions, as well as higher correct responding in the second session (GFP =  $10.4 \pm 1.7$  errors, ArchT =  $16.1 \pm 1.1$  errors;  $t[14] = 2.6$ ,  $p < 0.05$ ). Furthermore, by analyzing the trial-by-trial sequence of choices, we found that DLS silencing increased “win-stay” and “lose-shift” behavior during early learning (Figures 1K and 1L). These data show that silencing the DLS significantly facilitated early choice learning, associated with improved use of prior trial outcomes to guide choice.

In contrast to these effects of silencing at choice, early learning was unaffected by DLS silencing at reward collection (Figure S1). There remained the possibility that rather than improving early learning, DLS silencing altered performance by changing the perception of the two stimuli. For example, because as noted above, mice show an unlearned perceptual preference for the non-rewarded “marble” stimulus at the start of training (meaning that percentage correct choice begins below chance), any effect that caused a reduction in this bias would increase relative choice of the rewarded “fan” stimulus, leading to an apparent rapid improvement in learning (because percentage correct choice would be at chance, not below). To discount this possibility, we trained a cohort with the stimulus-reward contingencies switched (“marble” rewarded; initial performance now well above chance): if DLS silencing interfered with perception, then percentage correct choice would have decreased to chance levels (“impaired” early learning), which was not the case (Figure S2). Finally, to exclude the possibility that the absence of effects of DLS silencing at later learning was due to loss of the efficacy of inhibition after repeated photosilencing, we showed that silencing only once learning criterion was fully attained also failed to affect performance (Figure S2).

Next, we further examined the behavioral effects of DLS silencing on early learning by video-tracking patterns of movement in the operant chamber as a function of separate zones we defined in the vicinity of the touchscreen or reward collection magazine or the transition area between these two zones (Figure 2). This revealed that DLS silencing slowed and decreased movement in the transition zone located during early learning and also reduced the number of visits (and slightly but significantly reduced the speed of movement) in the touchscreen choice zone (Figure 2).

One interpretation of these changes is that facilitation of early learning resulting from DLS silencing was reflected in choices that were executed in a more motorically economical and deliberative manner (Redish, 2016). But why would silencing this region produce such an effect? With accrued experience, the DLS is posited to be a substrate for the concatenation of action sequence into a streamlined stimulus-driven response (Dezfouli et al., 2014; Graybiel, 1998; Jin and Costa, 2010). However, early in training, the engagement of the DLS may lead to the generation of multiple, individual actions that are not as yet organized into an effective “meta-action.” Thus, DLS silencing may have reduced redundant action generation to decrease overall movement as choices were executed effectively. In this context, one interesting avenue for future work could involve limiting DLS silencing to specific behavioral actions during the period of choice deliberation in order to begin dissecting the precise role of the DLS in mediating early discrimination performance.

Taken together, this initial set of experiments demonstrated behavioral changes resulting from DLS silencing that are not only consistent with the active involvement of the DLS from the earliest stages of learning but suggest that this engagement normally interacts with other systems, with the effect of impeding the initial acquisition of rewarded choice. The early recruitment of the DLS in this task fits well with prior studies that have concluded that the DLS is capable of supporting learning in a wide variety of assays (Gremel and Lovinger, 2016; Gremel and Costa, 2013; Kim et al., 2013; Kimchi et al., 2009; Stalnaker et al., 2010; Thorn et al., 2010; Yin et al., 2004, 2005a, 2005b, 2006b, 2009). It does remain possible, however, that rather than reflecting an effect of removing the involvement of the DLS in early discrimination learning, the early improvement in performance stems from the attenuation of a competing stimulus-reward relationship formed prior to the discrimination, during pre-training. In principal, these possibilities could be parsed by silencing DLS during discrimination in mice that had received no prior experience in the touchscreen apparatus; though it is unclear if the discrimination could even be learned under such conditions.

### Specific Striatal Output Pathways Contribute to Learning

Our next step was to ask whether this contribution of the DLS to learning was mediated at the level of specific (direct versus indirect) output pathways that are segregated by anatomical connectivity (somatomotor versus prefrontal) and molecular phenotype (dopamine receptor D1; *Drd1* versus *Drd2* or adenosine A2a receptor; *Adora2a*) (Gerfen and Surmeier, 2011). To this end, we used fluorescence *in situ* hybridization to quantify the relative number of DLS *Drd1*- or *Drd2*-expressing cells that were activated at different stages of learning. We used the expression levels of *Arc* mRNA (*Arc*), as a readout of learning-induced synaptic plasticity, rather than neuronal activation more generally (as would indicated by with other IEG markers such as c-Fos (Shepherd and Bear, 2011) (Figures 3A–3D). To test the relative expression levels of *Arc* in *Drd1* or *Drd2* DLS neurons following activity, we first quantified *Arc* expression following stimulation with a strong chemical (pentylentetrazole) induction stimulus and confirmed no differences between cell types (Figure S3).

We found that although the overall number of activated (*Arc*-positive) DLS cells was unaltered with training, the activated proportion that were *Arc/Drd1* coexpressing was high

during early learning and then significantly decreased from early to late learning, whereas the proportion that were *Arc/Drd2* coex-pressing did not change across stages (Figures 3E and 3F). These data suggest the direct output pathway of the DLS is preferentially recruited at early learning and that, as learning progresses, activity becomes more evenly distributed across the output pathways, though there remained a significant *Drd1* bias even at late learning (Figure 3G). A comparable shift from the direct to indirect pathway has previously been found to parallel the transition from outcome-based to stimulus-driven behavior in other behavioral settings (Furlong et al., 2015; Shan et al., 2014), suggesting that direct-pathway predominance early in training may be a phenomenon that is common across tasks, reflecting the strong action-drive (“go signal”) that is attributed to this pathway (Cui et al., 2014; Friend and Kravitz, 2014).

If early direct-pathway recruitment is important to DLS mediation of learning in the choice task, then disrupting it should affect performance. To test for this, we selectively silenced the DLS direct pathway by expressing rAAV5/EF1a-DIO-eArch3.0-eYFP in *Drd1*-Cre transgenic mice and shining green light at the time of choice. We found that silencing the pathway significantly improved correct responding, reduced error rates, and improved both win-stay and lose-shift behavior during early learning, as well as reducing response latencies at all stages (Figures 3H–3L and S4). These data show that restricting silencing to the direct pathway by and large mimicked the learning-facilitating effects of pathway-nonspecific silencing and fit with our observation that the direct pathway is preferentially recruited in the DLS during early learning.

Although providing support for the functional importance of the direct pathway, these data do not speak to the possible contribution of the indirect pathway to performance in the choice task. This is a particularly pertinent issue in view of growing evidence that indirect-pathway striatal neurons do not simply act as a “stop signal,” as traditional models proposed, but exert a more nuanced influence on action learning (Cui et al., 2014; Friend and Kravitz, 2014; Tecuapetla et al., 2016). We therefore selectively silenced the indirect pathway by shining green light at choice on DLS cells expressing rAAV5/EF1a-DIO-eArch3.0-eYFP in *Adora2a*-cre BAC transgenic mice (Figure 3M).

This manipulation decreased early-stage error rates but, unlike direct-pathway silencing, did not alter early correct responding and slowed, rather than quickened, mid-stage response latencies (Figures 3N–3Q and S4). Overall, the weaker effects of indirect-pathway silencing on early learning align with the *Arc* mRNA expression results that suggested a preponderant direct-pathway role at this stage. Moreover, the pattern of effects that were produced by silencing the indirect pathway mirrored those of direct-pathway silencing for some measures (reduced error rates) but differed (no change in percentage correct choice, win-shift, lose-stay) or were opposite for others (slower mid-stage choice latencies). These data add to growing evidence for a more complex (both cooperative and divergent) relationship between the pathways in guiding complex actions than traditionally thought (Cui et al., 2014; Surmeier, 2013; Tecuapetla et al., 2016).



## Silencing the DLS Leads to Systems-Level Reorganization of Arc Expression

Our finding that DLS silencing promoted early choice learning, together with the fact that performance was established to a high level without the normal contribution of the DLS, implies that the recruitment of other learning systems are capable of supporting the behavior. To provide some initial insight into the identity of alternative systems that are recruited when mice were forced to learn this choice task with the DLS silenced, we trained mice with concomitant DLS silencing and then immunohistochemically stained neurons in multiple cortical, striatal, and amygdaloid regions for Arc (Figure 4A). As a first step in analyzing these Arc data, we looked for general patterns (i.e., irrespective of silencing condition or learning stage) of coordinated activity across regions by performing cross-correlations of Arc counts. We saw a very strong correlation between the number of Arc-stained neurons in the DMS and the prelimbic subregion of medial prefrontal cortex (PL) (Fisher transformation,  $r = 0.74$ ,  $df[27]$ ,  $p < 0.05$  after nine-comparison Bonferroni correction) (Figures 4B and 4C). No significant correlations were evident for any other brain regions, including for the DMS and other regions, such as the somatomotor cortex (SM) (Figures 4D, 4E, and S5). These data strongly suggest close functional integration between the PL and DMS in choice learning, as has been previously documented for various forms of outcome-based learning (Balleine et al., 2009).

Next, we examined Arc expression to gain insight into which brain regions might be undergoing compensatory task-related neuronal recruitment and plasticity as a consequence of having to learn the task with the DLS silenced (Steward et al., 1998). Results indicated that when the DLS was silenced throughout training, the number of Arc-positive neurons in the PL and DMS was markedly reduced at the late learning stage, compared with non-silenced controls (Figures 4F and 4G). There was no effect of silencing on these regions at early learning, nor was there an effect on any of the other regions examined, including the SM, insular cortex, nucleus accumbens, and amygdala (Figures 4H–4K and S5). These data offer evidence of a highly specific alteration in the PL-DMS circuit as a result of DLS-silencing.

It is tempting to speculate that this alteration represents the augmented recruitment of a circuit comprising these regions to resolve choice learning in the absence of a normally functioning DLS. Although it may appear paradoxical that increased PL-DMS engagement would manifest as a smaller, rather than larger, neuronal population at late learning, prior studies in other regions have reported decreases in the size of a task-related neuronal ensemble coincident with improvements in learning (Komiyama et al., 2010). Transcriptional and post-translational regulation of Arc has been associated with both long-term potentiation (LTP) and long-term depression (LTD) processes (Bramham et al., 2010). Therefore, another interpretation is that the reduction in Arc-expressing neurons in the PL/DMS reflects compensatory changes in synaptic plasticity within this pathway, rather than recruitment per se. However, formally testing these interpretations would require more direct measurements of plasticity, for example by measuring synaptic currents and LTP/LTD in PL-DMS neurons in mice that had learned the discrimination task with the DLS silenced throughout (Yin et al., 2009). Further studies would also be needed to fully explore potential functional changes in other regions, stemming from DLS silencing (e.g., downstream

pallidal and nigral output areas (Lee et al., 2016), that could contribute to the learning effects of silencing the DLS.

The suggestion that DLS silencing results in the compensatory engagement of a PL-DMS circuit, acting as an alternative system to subserve choice learning, assumes that the DMS and its inputs from PL functionally support this task, which has not been demonstrated. Therefore, to establish a role for the DMS in this task, we expressed rAAV8/CAG-ArchT-GFP to silence the region during learning, using the same procedure that we used for DLS silencing (Figure 4L). We found that DMS silencing at choice produced quite widespread deficits across learning: reductions in mid- and late-stage correct responding and win-stay behavior and accompanying increases in error rates (Figures 4M–4P and S6). Decrements in performance were not evident at the early stage, but it should be borne in mind that performance at this stage in the control group is already essentially at floor at ~30% correct responding, which reflects the aforementioned unlearned perceptual bias for the unrewarded stimulus in this stimulus pairing. As such, it is unclear if DMS silencing failed to affect early learning or the task is simply unable to detect such an effect. The question of whether specific silencing of the direct and indirect pathways of the DMS would affect learning also remains one that should be addressed in future work.

The main conclusion from this set of experiments is that with the DLS silenced, choice learning led to a high correlated pattern of Arc activity, specifically involving two brain regions, the PL and DMS, which are known key mediators of goal-directed types of performance. Furthermore, with the normal operation of the DMS compromised through photosilencing, performance became error prone and less responsive to prior outcomes, a profile that is characteristic of a DLS-like stimulus-driven strategy. Another important observation is that although DMS silencing disrupted performance, excellent (criterion) levels of choice proficiency were still attained with training. Along similar lines, though initial learning was significantly affected by, in this case, expedited DLS silencing, mice went on to display accurate choice behavior that was no different from that seen in controls.

The broader implication is that there appears to be an essential redundancy between the DMS and DLS, and the goal-directed and stimulus-driven learning strategies they respectively support in this visual discrimination paradigm. That is, although both contribute to discrimination learning in critical ways, disruption of one system does not necessarily preclude criterion performance in the task. This is entirely in keeping with current models (Bradfield and Balleine, 2013; Daw et al., 2005) and prior work that has drawn similar conclusions using a diverse range of functional measures and manipulations from instrumental, maze-based, and skill-learning paradigms (see Introduction and citations therein).

Redundant learning systems would serve to safeguard behaviors that are crucial to fitness (e.g., exploiting resources that reliably produce food, while minimizing time- and resource-costly exploratory foraging). Indeed, it would be advantageous if the two parallel systems could functionally substitute for each other earlier in learning, before performance had reached final criterion levels. To test for this, we trained a cohort of mice to high but still sub-criterion levels of performance (~70% correct) and then photosilenced the DMS. Results



showed that silencing failed to disrupt behavior in these partially trained mice (Figure S6), demonstrating choice learning was protected against a loss of DMS function, presumably because, by this point in training, learning had been instantiated by alternate systems, including the DLS.

### **Concluding Remarks**

Long-standing theories propose the parallel recruitment of stimulus-driven and goal-directed learning systems, mediated in part by the DLS and DMS, respectively (Balleine and Ostlund, 2007; Bradfield and Balleine, 2013; Dickinson and Balleine, 1995; Thorndike, 1911). Current models also emphasize the importance of the DLS in governing actions after extensive experience has accumulated about the relationship between stimuli and rewards or their absence. However, our findings provide evidence that, at least in the context of choice learning, the DLS begins to exert an influence over behavior from the outset of learning. This influence likely reflects the nascent construction of stimulus-elicited meta-actions that can ultimately automate behavior and liberate cognitive resources for other demands (Dezfouli et al., 2014), but with a cost to performance. Our study offers a new appreciation for how parallel striatal systems can interact at various points across discrimination learning.

## **EXPERIMENTAL PROCEDURES**

See Supplemental Experimental Procedures for detailed procedures.

### **Subjects**

Unless stated otherwise, subjects were adult male C57BL/6J mice.

### **Behavioral Task**

All discrimination learning was conducted using the Bussey-Saksida Touch Screen System (Horner et al., 2013).

### **Viral Infusion and Optical Fiber Implantation**

For pathway-non-specific DLS photosilencing, mice were bilaterally infused with rAAV8/CAG-ArchT-GFP or rAAV8/CAG-GFP in the DLS. A ceramic ferrule assembly was implanted above the viral injection site.

### **Effects of DLS Silencing on Ex Vivo Neuronal Activity**

Brain slices containing the dorsal striatum were prepared for recording from C57BL/6J mice after bilateral infusion of rAAV8/CAG-ArchT-GFP into the DLS.

### **Effects of DLS Silencing in the Open Field and Real-Time Place Preference Tests**

Mice were bilaterally infected with rAAV8/CAG-ArchT-GFP and tested in the open-field and real-time place preference (RTPP) tests.

### Effects of DLS Silencing at Choice on Learning

Mice were pre-trained on the touchscreen and then bilaterally infected with either rAAV8/CAG-ArchT-GFP or rAAV8/CAG-GFP to silence the DLS. For silencing, light was delivered at trial initiation through choice to reward collection (correct trial) and at trial initiation through choice to 3 s post-choice (incorrect trial).

### Effects of DLS Silencing at Reward

The procedures were identical to those described above, except that light was delivered at reward collection.

### Effects of DLS Silencing at Choice on Early Learning, with the Alternate Stimulus Rewarded

The procedures were identical to those described above, except that the responses at the “marble” stimulus produced a food reward, whereas responses at the “fan” stimulus (i.e., “errors”) produced no food reward. Mice were tested during early learning.

### Effects of DLS Silencing at Choice on Learning Criterion

The procedures were identical to those described above for silencing at choice, except that silencing was restricted to the three sessions after discrimination training without silencing, to attain at least 85% correct performance on two consecutive test days (the late learning stage) (Figures 1B and 1C).

### Effects of DLS Silencing on Spatiotemporal Measures of Learning

The procedures were identical to those described above for silencing at choice. Behavior was recorded during early, mid, and late stages using an overhead infrared camera and videos analyzed offline (Figure 2).

### Assessment of Learning-Related DLS Output-Pathway Recruitment

Mice were given two discrimination sessions (early group), and another group was trained to criterion (late group), then prepared for fluorescence *in situ* hybridization.

### Comparison of Induced Arc Expression in the DLS Direct and Indirect Pathways

Hemizygous *B6.Cg-Tg(Drd1a-tdTomato)6Calak/J* × hemizygous *Tg(Drd2-EGFP)S118Gsat* (GENSAT) dual-reporter mice (C57BL/6J background) were injected with pentylentetrazole and prepared for immunohistochemistry.

### Effects of DLS Direct Pathway Silencing on Learning

Male *Drd1*-Cre-positive and Cre-negative mice were used for testing (Gerfen et al., 2013). Following pre-training, the DLS was bilaterally infected with rAAV5/EF1a-DIO-eArch3.0-eYFP, and optical fibers were implanted. The procedures were identical to those described above for silencing at choice.

### Effects of DLS Indirect-Pathway Silencing on Learning

Male *Adora2a*-Cre-positive and Cre-negative mice were used for testing. Following pre-training, the DLS was bilaterally infected with rAAV5/EF1a-DIO-eArch3.0-eYFP and optical fibers implanted. The procedures were identical to those described above for silencing at choice.

### Assessment of System-Level Adaptations due to DLS Silencing

The procedures were identical to those described above for silencing at choice. One group was given two discrimination sessions (early group), and another group was trained to criterion (late group) and prepared for Arc immunohistochemistry.

### Effects of DMS Silencing at Choice on Learning

Mice were pre-trained on the touchscreen and then bilaterally infected with either rAAV8/CAG-ArchT-GFP or rAAV8/CAG-GFP and had optical fibers implanted into the DMS. Green light was shone during choice.

### Effects of DMS Silencing at Choice on Mid-stage Learning

The procedures were identical to those described above for silencing at choice, except that silencing was restricted to the mid-learning stage.

### Statistical Methods

Independent t tests were used for all statistical comparisons. In the case of multiple comparisons, either the false discovery rate (Benjamini and Hochberg, 1995) or Bonferroni correction was used to control type I error.

### Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

### Acknowledgments

This research was supported by the National Institute on Alcohol Abuse and Alcoholism (NIAAA) Intramural Research Program, the Henry Jackson Foundation for the Advancement of Military Medicine (CNRM-70-2578), and the Department of Defense in the Center for Neuroscience and Regenerative Medicine. D.G.W is supported by NIH awards AA019455 and DA042475. T.A.W. is supported by K99/R00 AA022651. The authors thank Drs. Yavin Shaham and Bruce Hope (NIDA) for valuable technical advice on the methods for fluorescence in situ hybridization.

### References

- Balleine BW, Ostlund SB. Still at the choice-point: action selection and initiation in instrumental conditioning. *Ann N Y Acad Sci.* 2007; 1104:147–171. [PubMed: 17360797]
- Balleine BW, Delgado MR, Hikosaka O. The role of the dorsal striatum in reward and decision-making. *J Neurosci.* 2007; 27:8161–8165. [PubMed: 17670959]
- Balleine BW, Liljeholm M, Ostlund SB. The integrative function of the basal ganglia in instrumental conditioning. *Behav Brain Res.* 2009; 199:43–52. [PubMed: 19027797]
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B.* 1995; 57:289–300.

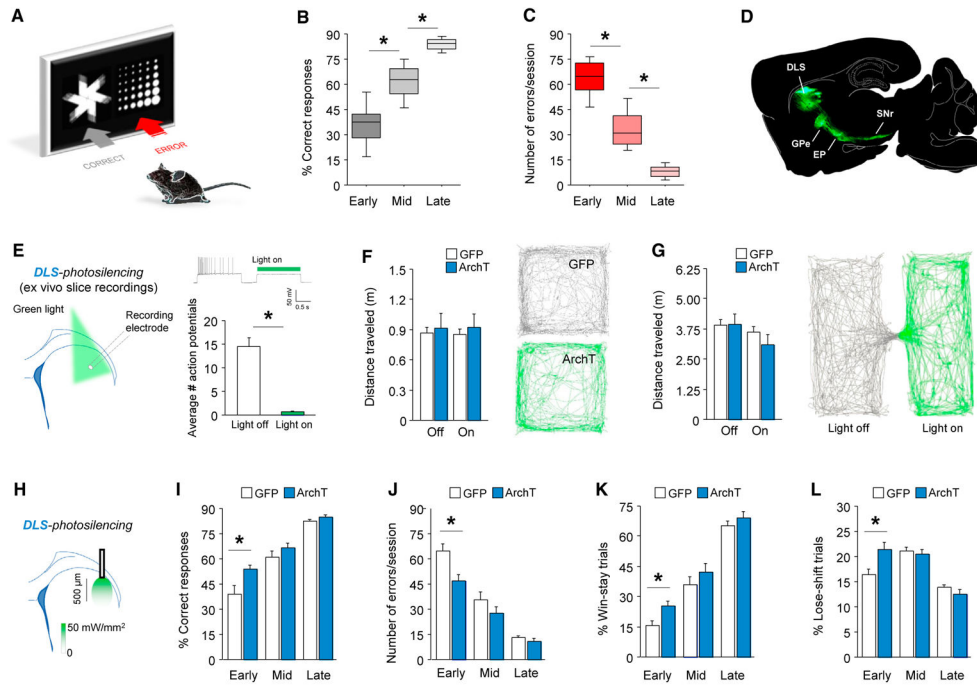
- Bradfield LA, Balleine BW. Hierarchical and binary associations compete for behavioral control during instrumental biconditional discrimination. *J Exp Psychol Anim Behav Process.* 2013; 39:2–13. [PubMed: 23316974]
- Bramham CR, Alme MN, Bittins M, Kuipers SD, Nair RR, Pai B, Panja D, Schubert M, Soule J, Tiron A, Wibrand K. The Arc of synaptic memory. *Exp Brain Res.* 2010; 200:125–140. [PubMed: 19690847]
- Brigman JL, Daut RA, Wright T, Gunduz-Cinar O, Graybeal C, Davis MI, Jiang Z, Saksida LM, Jinde S, Pease M, et al. GluN2B in corticostriatal circuits governs choice learning and choice shifting. *Nat Neurosci.* 2013; 16:1101–1110. [PubMed: 23831965]
- Corbit LH, Janak PH. Inactivation of the lateral but not medial dorsal striatum eliminates the excitatory impact of Pavlovian stimuli on instrumental responding. *J Neurosci.* 2007; 27:13977–13981. [PubMed: 18094235]
- 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]
- Cui G, Jun SB, Jin X, Luo G, Pham MD, Lovinger DM, Vogel SS, Costa RM. Deep brain optical measurements of cell type-specific neural activity in behaving mice. *Nat Protoc.* 2014; 9:1213–1228. [PubMed: 24784819]
- Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci.* 2005; 8:1704–1711. [PubMed: 16286932]
- DePoy L, Daut R, Brigman JL, MacPherson K, Crowley N, Gunduz-Cinar O, Pickens CL, Cinar R, Saksida LM, Kunos G, et al. Chronic alcohol produces neuroadaptations to prime dorsal striatal learning. *Proc Natl Acad Sci U S A.* 2013; 110:14783–14788. [PubMed: 23959891]
- Dezfouli A, Lingawi NW, Balleine BW. Habits as action sequences: hierarchical action control and changes in outcome value. *Philos Trans R Soc Lond B Biol Sci.* 2014; 369:20130482. <https://doi.org/10.1098/rstb.2013.0482>. [PubMed: 25267824]
- Dickinson A, Balleine B. Motivational control of instrumental action. *Curr Dir Psychol Sci.* 1995; 4:162–167.
- Friend DM, Kravitz AV. Working together: basal ganglia pathways in action selection. *Trends Neurosci.* 2014; 37:301–303. [PubMed: 24816402]
- Furlong TM, Supit AS, Corbit LH, Killcross S, Balleine BW. Pulling habits out of rats: adenosine 2A receptor antagonism in dorsomedial striatum rescues meth-amphetamine-induced deficits in goal-directed action. *Addict Biol.* 2015; 22:172–183. [PubMed: 26515740]
- Gerfen CR, Surmeier DJ. Modulation of striatal projection systems by dopamine. *Annu Rev Neurosci.* 2011; 34:441–466. [PubMed: 21469956]
- Gerfen CR, Paletzki R, Heintz N. GENSAT BAC Cre-recombinase driver lines to study the functional organization of cerebral cortical and basal ganglia circuits. *Neuron.* 2013; 80:1368–1383. [PubMed: 24360541]
- Graybeal C, Feyder M, Schulman E, Saksida LM, Bussey TJ, Brigman JL, Holmes A. Paradoxical reversal learning enhancement by stress or prefrontal cortical damage: rescue with BDNF. *Nat Neurosci.* 2011; 14:1507–1509. [PubMed: 22057192]
- Graybiel AM. The basal ganglia and chunking of action repertoires. *Neurobiol Learn Mem.* 1998; 70:119–136. [PubMed: 9753592]
- Graybiel AM. Habits, rituals, and the evaluative brain. *Annu Rev Neurosci.* 2008; 31:359–387. [PubMed: 18558860]
- Gremel CM, Costa RM. Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nat Commun.* 2013; 4:2264. [PubMed: 23921250]
- Gremel CM, Lovinger DM. Associative and sensorimotor cortico-basal ganglia circuit roles in effects of abused drugs. *Genes Brain Behav.* 2016; 16:71–85. [PubMed: 27457495]
- Horner AE, Heath CJ, Hvoslef-Eide M, Kent BA, Kim CH, Nilsson SR, Alsjö J, Oomen CA, Holmes A, Saksida LM, Bussey TJ. The touchscreen operant platform for testing learning and memory in rats and mice. *Nat Protoc.* 2013; 8:1961–1984. [PubMed: 24051959]
- Izquierdo A, Wiedholz LM, Millstein RA, Yang RJ, Bussey TJ, Saksida LM, Holmes A. Genetic and dopaminergic modulation of reversal learning in a touchscreen-based operant procedure for mice. *Behav Brain Res.* 2006; 171:181–188. [PubMed: 16713639]

- Jin X, Costa RM. Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature*. 2010; 466:457–462. [PubMed: 20651684]
- Kim H, Lee D, Jung MW. Signals for previous goal choice persist in the dorsomedial, but not dorsolateral striatum of rats. *J Neurosci*. 2013; 33:52–63. [PubMed: 23283321]
- Kimchi EY, Torregrossa MM, Taylor JR, Laubach M. Neuronal correlates of instrumental learning in the dorsal striatum. *J Neurophysiol*. 2009; 102:475–489. [PubMed: 19439679]
- Komiyama T, Sato TR, O'Connor DH, Zhang YX, Huber D, Hooks BM, Gabbito M, Svoboda K. Learning-related fine-scale specificity imaged in motor cortex circuits of behaving mice. *Nature*. 2010; 464:1182–1186. [PubMed: 20376005]
- Lee HJ, Weitz AJ, Bernal-Casas D, Duffy BA, Choy M, Kravitz AV, Kreitzer AC, Lee JH. Activation of direct and indirect pathway medium spiny neurons drives distinct brain-wide responses. *Neuron*. 2016; 91:412–424. [PubMed: 27373834]
- Redish AD. Vicarious trial and error. *Nat Rev Neurosci*. 2016; 17:147–159. [PubMed: 26891625]
- Shan Q, Ge M, Christie MJ, Balleine BW. The acquisition of goal-directed actions generates opposing plasticity in direct and indirect pathways in dorsomedial striatum. *J Neurosci*. 2014; 34:9196–9201. [PubMed: 25009253]
- Shepherd JD, Bear MF. New views of Arc, a master regulator of synaptic plasticity. *Nat Neurosci*. 2011; 14:279–284. [PubMed: 21278731]
- Smith KS, Graybiel AM. A dual operator view of habitual behavior reflecting cortical and striatal dynamics. *Neuron*. 2013; 79:361–374. [PubMed: 23810540]
- Stalnaker TA, Calhoun 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.
- Steward O, Wallace CS, Lyford GL, Worley PF. Synaptic activation causes the mRNA for the IEG Arc to localize selectively near activated postsynaptic sites on dendrites. *Neuron*. 1998; 21:741–751. [PubMed: 9808461]
- Surmeier DJ. Neuroscience: To go or not to go. *Nature*. 2013; 494:178–179. [PubMed: 23354047]
- Tecuapetla F, Jin X, Lima SQ, Costa RM. Complementary contributions of striatal projection pathways to action initiation and execution. *Cell*. 2016; 166:703–715. [PubMed: 27453468]
- Thorn CA, Atallah H, Howe M, Graybiel AM. Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. *Neuron*. 2010; 66:781–795. [PubMed: 20547134]
- Thorndike, EL. *Animal Intelligence: Experimental Studies*. Macmillan; 1911.
- Vicente AM, Galvão-Ferreira P, Tecuapetla F, Costa RM. Direct and indirect dorsolateral striatum pathways reinforce different action strategies. *Curr Biol*. 2016; 26:R267–R269. [PubMed: 27046807]
- Yin HH, Knowlton BJ, Balleine BW. Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur J Neurosci*. 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. *Eur J Neurosci*. 2005a; 22:505–512. [PubMed: 16045503]
- Yin HH, Ostlund SB, Knowlton BJ, Balleine BW. The role of the dorsomedial striatum in instrumental conditioning. *Eur J Neurosci*. 2005b; 22:513–523. [PubMed: 16045504]
- Yin HH, Knowlton BJ, Balleine BW. Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. *Behav Brain Res*. 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]

### Highlights

- Photosilencing the DLS improves early discrimination learning
- Selective DLS direct output-pathway photosilencing improves early learning
- DLS photosilencing alters learning-related PL-DMS activation
- Selective photosilencing of the DMS disrupts learning





### Figure 1. DLS Photosilencing Expedites Early Discriminated Choice Learning

(A–C) Touchscreen visual discrimination training (A) increased correct responding ( $F_{2,50} = 185.25$ ,  $p < 0.01$ ,  $n = 26$ ) (B) and decreased error rates ( $F_{2,50} = 161.06$ ,  $p < 0.01$ ) (C).

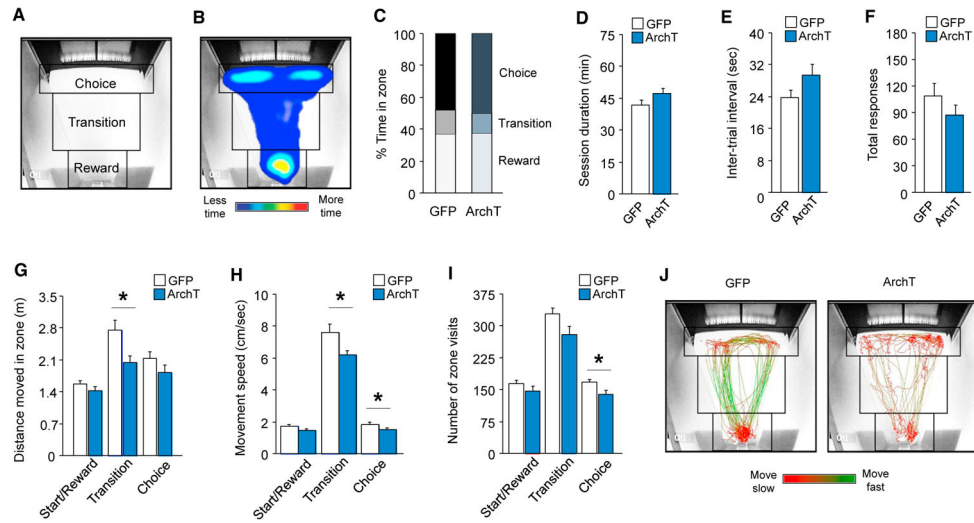
(D) GFP-labeled DLS AAV-expression. EP, endopeduncular nucleus; GPe, external globus pallidus; SNr, substantia nigra reticulata.

(E) Green light reduced DLS neuronal activity *ex vivo* ( $t[80] = 4.68$ ,  $p < 0.01$ ,  $n = 7$ ).

(F and G) DLS photosilencing did not affect open-field locomotion (F) or real-time place preference (G).

(H–L) DLS photosilencing (H) increased early-stage percentage correct responding ( $t[14] = 2.40$ ,  $p = 0.05$ ,  $n = 7-9$ ) (I), decreased errors ( $t[14] = 2.91$ ,  $p < 0.01$ ) (J), and increased percentage win-stay ( $t[14] = 2.92$ ,  $p < 0.05$ ) (K) and percentage lose-shift ( $t[14] = 2.96$ ,  $p < 0.05$ ) (L) behavior.

\* $p < 0.05$ . Data are mean  $\pm$  SEM.



### Figure 2. Effects of DLS Photosilencing Effects on Spatiotemporal Parameters in the Touchscreen Task

(A) Definitions of choice, transition, and reward zones.

(B) Representative heatmap of cumulative time spent in each zone.

(C) Percentage distribution of time in each zone did not differ between GFP and ArchT groups.

(D) Average time to session completion did not differ between GFP and ArchT groups.

(E) Average inter-trial interval did not differ between GFP and ArchT groups.

(F) Total responses made did not differ between GFP and ArchT groups.

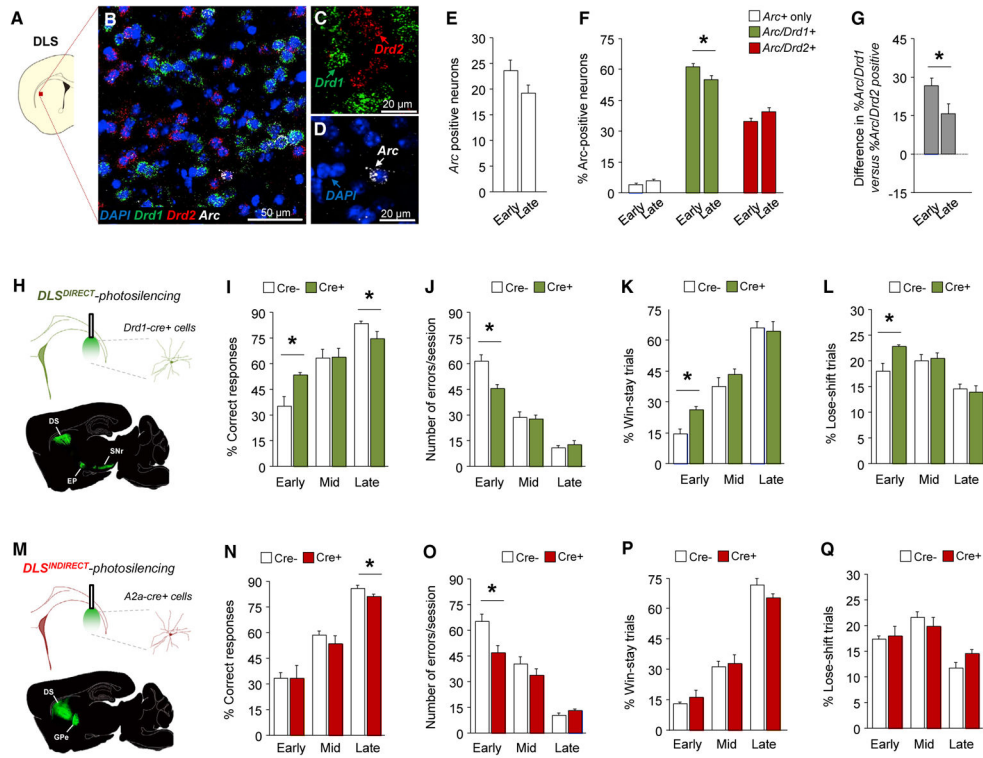
(G) The ArchT group traveled less distance in the transition zone than GFP controls ( $t[14] = 2.81$ ,  $p < 0.05$ ).

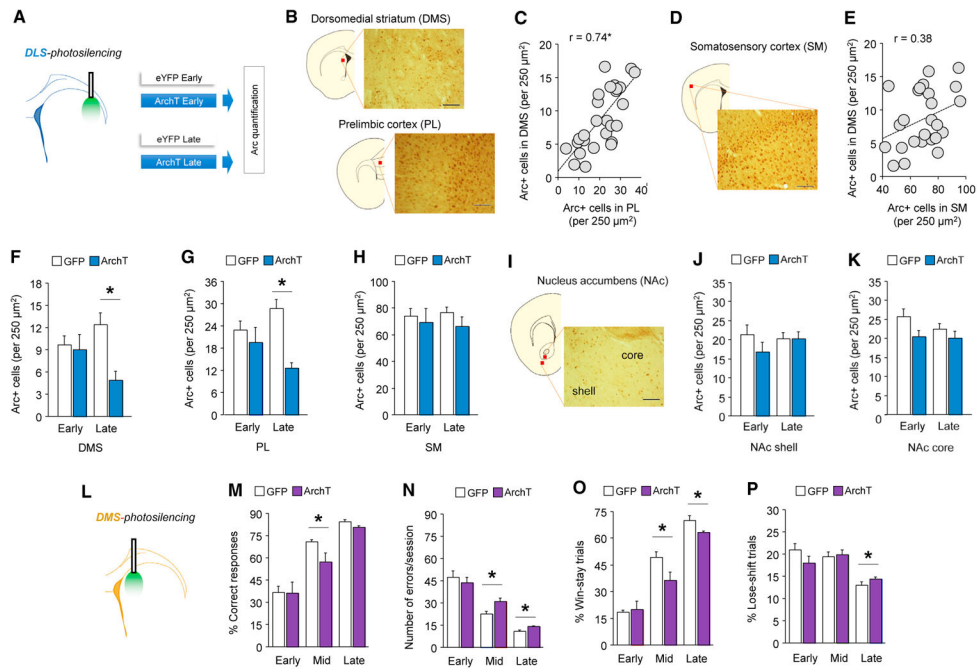
(H) The ArchT group moved more slowly in the transition ( $t[14] = 2.30$ ,  $p < 0.05$ ) and choice ( $t[14] = 2.18$ ,  $p < 0.05$ ) zones than GFP controls.

(I) The ArchT group visited the choice zone less frequently than GFP controls ( $t[14] = 2.38$ ,  $p < 0.05$ ).

(J) Representative movement speed traces in GFP and ArchT groups (maximum green = 7.5 cm/s, maximum red = 1.5 cm/s).

\* $p < 0.05$ . Data are mean  $\pm$  SEM.





#### Figure 4. System-Level Adaptations to DLS Photosilencing

(A) Effects of DLS photosilencing on number of Arc<sup>+</sup> neurons in various brain regions was quantified after early or late learning.

(B–E) Irrespective of DLS photosilencing or learning stage, Arc<sup>+</sup> neuronal number was highly correlated between DMS and PL (B and C), but not DMS and SM

(D and E) (data from  $n = 5–9$  mice,  $n = 6$  sections per mouse,  $n = 650–790$  DMS cells,  $n = 1,600–1,772$  PL cells).

(F–P) DLS photosilencing decreased late-learning Arc expression in DMS (F) ( $t[11] = 3.76$ ,  $p < 0.01$ ,  $n = 59$ ) and PL (G) ( $t[11] = 5.88$ ,  $p < 0.01$ ), not SM or NAc (H–K). DLS-photosilencing decreased late-learning Arc expression in DMS (F,  $t(11) = 3.76$ ,  $p < 0.01$ ,  $n = 5–9$ ) and PL ( $t(11) = 5.88$ ,  $p < 0.01$ ) (G), not SM (H), NAc shell or NAc core (I–K) DMS-photosilencing (L) impaired mid/late-learning on multiple measures (M, % correct mid =  $t(13) = 2.22$ ,  $p < 0.05$ ; N, errors mid =  $t(13) = 2.43$ ,  $p < 0.05$ , late =  $t(13) = 2.56$ ,  $p < 0.05$ ; O, % win-stay trials mid =  $t(13) = 2.30$ ,  $p < 0.05$ , late =  $t(13) = 2.13$ ,  $p = .05$ ;  $n = 7–8$ ) and increased late-stage lose-shift behavior ( $t(14) = 2.58$ ,  $p < 0.05$ ) (P).

\* $p < 0.05$ . Data are mean  $\pm$  SEM. Scale bar = 50  $\mu\text{m}$ .