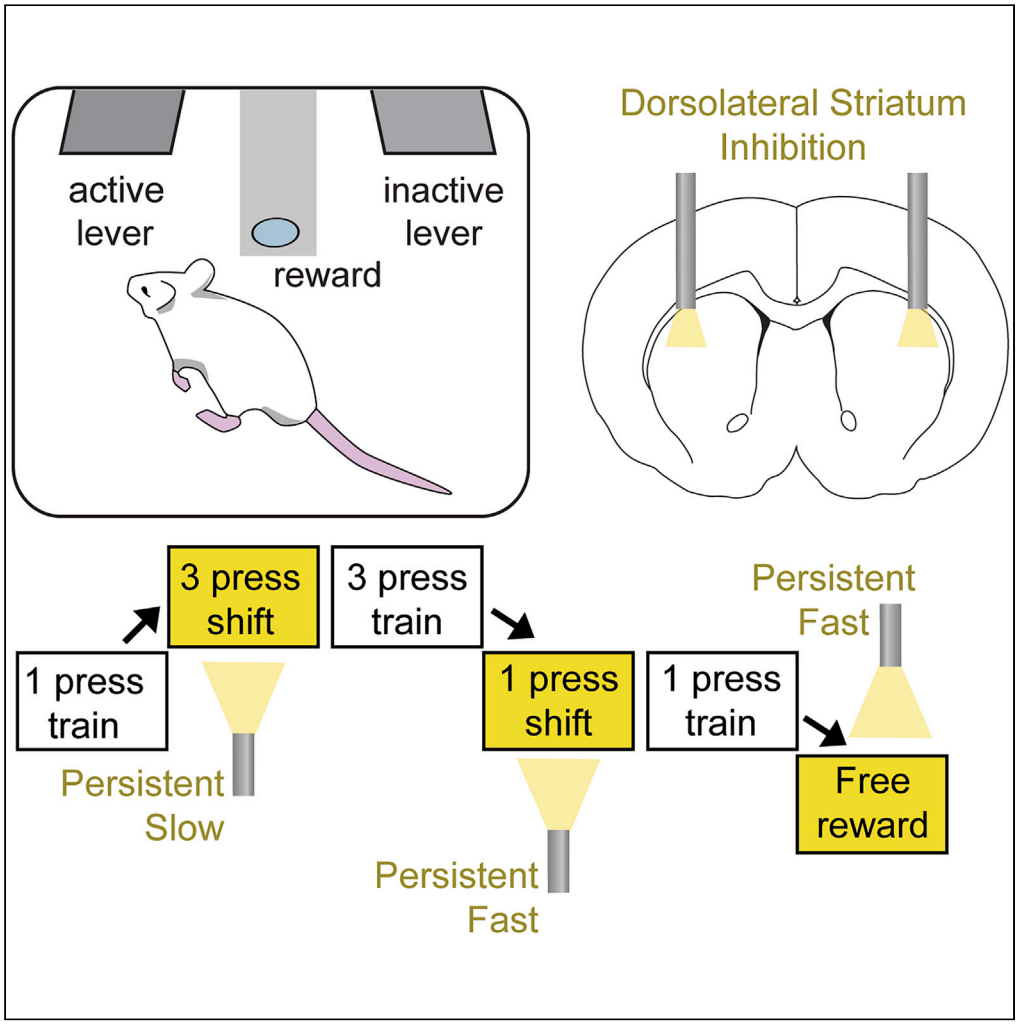
iScience

CellPress
OPEN ACCESS

Article

# A role for the dorsolateral striatum in prospective action control




Adam C.G. Crego, Kenneth A. Amaya, Jensen A. Palmer, Kyle S. Smith

crego.adam@gmail.com (A.C.G.C.)
kyle.s.smith@dartmouth.edu (K.S.S.)

## Highlights

Inhibition of the dorsolateral striatum (DLS) results in perseverative behavior

This perseverative behavior is habit-like and occurs in multiple task changes

DLS inhibition can both decrease and increase speed depending on the task changes

Crego et al., iScience 27, 110044
June 21, 2024 
https://doi.org/10.1016/j.isci.2024.110044

iScience

CellPress
OPEN ACCESS

Article

# A role for the dorsolateral striatum in prospective action control

Adam C.G. Crego,[1,*] Kenneth A. Amaya,[1] Jensen A. Palmer,[1] and Kyle S. Smith[1,2,*]

## SUMMARY

**The dorsolateral striatum (DLS) is important for performing actions persistently, even when it becomes suboptimal, reflecting a function that is reflexive and habitual. However, there are also ways in which persistent behaviors can result from a more prospective, planning mode of behavior. To help tease apart these possibilities for DLS function, we trained animals to perform a lever press for reward and then inhibited the DLS in key test phases: as the task shifted from a 1-press to a 3-press rule (upshift), as the task was maintained, as the task shifted back to the one-press rule (downshift), and when rewards came independent of pressing. During DLS inhibition, animals always favored their initially learned strategy to press just once, particularly so during the free-reward period. DLS inhibition surprisingly changed performance speed bidirectionally depending on the task shifts. DLS inhibition thus encouraged habitual behavior, suggesting it could normally help adapt to changing conditions.**

## INTRODUCTION

Reward-seeking actions can become persistent and difficult to change, often as a result of extensive experience. Behavioral persistence can be observed in several forms, including insensitivity to devaluation of the outcome or insensitivity to changes in action-outcome contingencies.[1] Persistent behavior of this sort is regarded as being driven by stimulus-response or model-free learning, or in other words, by habit. This mode of behavior is retrospective in nature, focusing on continuing what had been done in the past. However, there are also ways in which behaviors can be persistent in a more prospective manner. Such accounts include model-based learning,[2] cognitive fixation,[3] and latent causal inference of action states,[4] in which behaviors can persist essentially because subjects are actively engaged in the behavior and favoring learned experiences over new environmental changes. Although brain areas that support persistence in behavior are nearly always considered to be brain areas for habit, a lack of behavioral change is a negative result and therefore it may be tenuous to ascribe "habit" to all such instances when there are other, more prospective forms of action control could in principal cause persistence too.

One brain area of interest in this regard is the dorsolateral striatum (DLS). The DLS is thought to be a key brain node for habit learning, and is necessary for invigorating actions, organizing them as sequences, and expressing them persistently.[5–15] Accentuation of DLS activity during the initiation of a behavior is linked with how fast, fluid, and persistent that behavior is.[16–22,66,67] The DLS is particularly important for causing extensively trained behaviors to be continued, by habit, when the outcome is delivered noncontingently from the action or when the outcome is devalued.[23,24] DLS activity would thus seem to promote the performance of action routines that have worked previously, with a retrospective focus.

However, the DLS is also implicated in initial action learning and in the adjustment of ongoing behavior.[25–31] Although this might reflect a DLS role in acquiring stimulus-response/model-free learning, it might instead reflect a DLS role in the prospective planning of actions akin to a strategy in which animals behave based on an internal model of task conditions and reward likelihood. Findings in support of the more retrospective habit function of DLS are open to re-interpretation in this domain. Moreover, without details on *how* animals are behaving when they persist in the face of changed action-reward relationships, it can be unclear if they are actually stuck in previously learned routines or if they are actively exploring the new task conditions and learning (see also studies by Dezfouli and Balleine,[2] and Bailey and Mair[25]). Finally, when changes in task conditions or outcome values are used to probe habits, and DLS function, they usually shift from something highly familiar to something new and unfamiliar; in such cases, persistent behaviors despite the task change could be related to a probalistic judgment that prior experiences outweigh the new ones.[3,32] Mechanistic studies of striatal dopamine input and molecular processes have also led to a revisitation of the classical divide of lateral and medial striatum in controlling habit and goal-directed behavior, respectively.[13,33,34]

Here, we sought to take a step in the direction of exploring whether the DLS can contribute to reward-seeking actions in an active, prospective manner. We approached this question by having animals learn a simple free-operant behavior and then challenging them with task rule changes while we inhibited the DLS. Those rules included (1) shifts away from what was initially learned, (2) shifts back toward what was initially learned, and (3) shifts to conditions that require no actions at all for reward. If the DLS supports a prospective strategy of action

[1]Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH 03755, USA
[2]Lead contact
*Correspondence: crego.adam@gmail.com (A.C.G.C.), kyle.s.smith@dartmouth.edu (K.S.S.)
https://doi.org/10.1016/j.isci.2024.110044

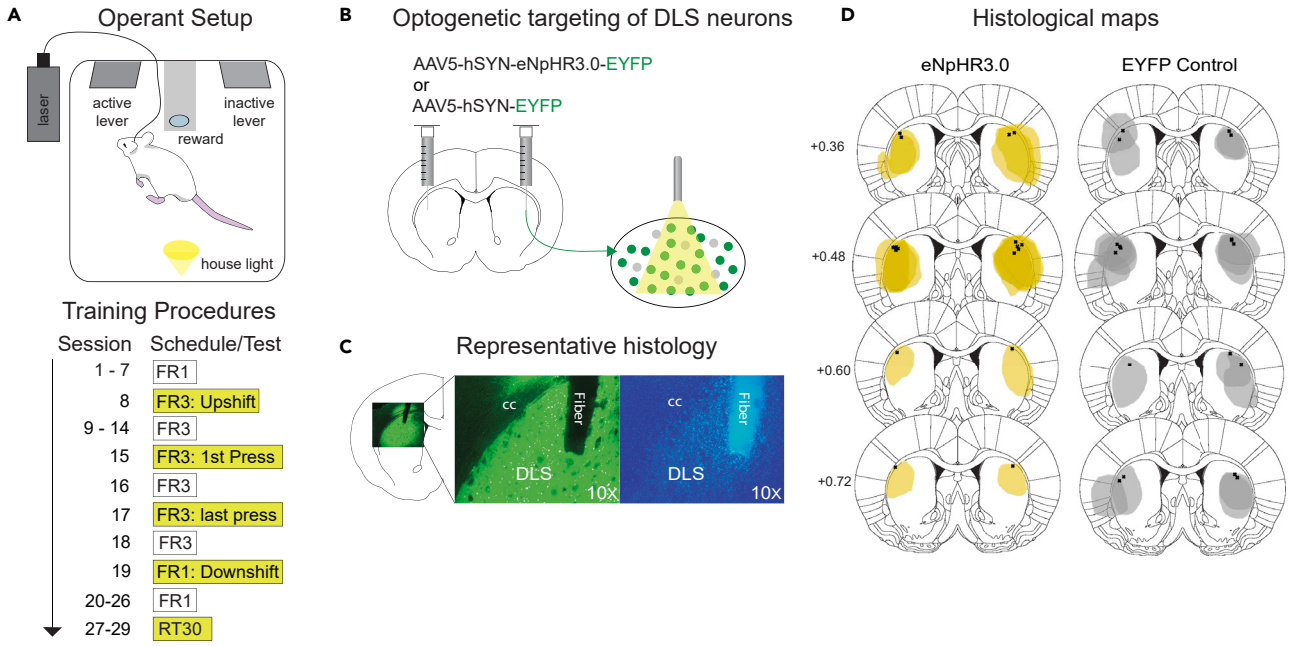



**Figure 1. Operant chamber task**

(A) Top, operant task cartoon illustrating arrangement of the magazine reward port nestled between the active lever (i.e., rewarded) and inactive lever (i.e., unrewarded). Optogenetic setup was external to the chamber, with cables threaded through the top. House light illumination signaled the start of a session. Bottom, training protocol and illumination test days (yellow): contingency tests.

(B) Cartoon representation of bilateral injections of viral constructs into DLS; green dots = EYFP expression, gray dots = no expression.

(C) Example optogenetic expression of DLS section from left to right: EYFP and DAPI fluorescent stain for neuronal health. Fiber placement labeled across sections.

(D) Expression areas (left to right) of eNpHR3.0-EYFP and EYFP control across animals. Expression is denoted per animals as semi-transparent shading overlaid across rats and group. Numbers on the right denote anteroposterior plane in mm relative to the bregma. Outside DLS expression was observed; however, fiber placements denoted as X's suggested illumination targets were exclusive to DLS. FR, fixed-ratio; RT, random-time; EYFP, enhanced yellow fluorescent protein; DAPI, 4′,6′-diamidino-2-phenylindole; eNpHR3.0, halorhodopsin; cc, corpus callosum; DLS, dorsolateral striatum.

selection, several hypotheses are raised. First, inhibition of the DLS should reduce new learning during task shifts. Second, DLS inhibition should not reduce reward-directed actions generally, but rather cause animals to revert to their initially learned strategies rather than develop new strategies. Third, if the shift of task rules is one toward unfamiliarity (i.e., inconsistency with what was learned), the breakdown of prospective planning with DLS inhibition will cause behavior to look maladaptive because it fails to integrate new task rules; conversely, if the task rule shift is toward a rule that is familiar (i.e., consistent with what had been learned), the breakdown of prospective planning will cause behavior to look adaptive because the old learned strategy is now coherent with the task.

Separately, the DLS helps animals perform actions in a rapid and effortful manner. Disruption of the DLS or its dopaminergic input causes slower and more variable performance patterns.[9,17,22,23,25,35–37,66,67] Stimulating the DLS, particularly at the onset of actions, can directly increase performance speed.[17] Whether DLS contributes to performance speed in a manner that serves behavioral exploitation (retrospective) or adaptation (prospective) is unclear. Thus, when DLS activity is engaged, it could be that animals are more rapidly performing what they have previously done or that animals are attuned to the current environment and are rapidly performing actions that best suit it. Thus, we combined measures of action strategy, as aforementioned, with measures of performance speed in order to help tie the two together.

## RESULTS

Rats were trained and tested in task conditions requiring lever pressing behavior to receive reward (Figure 1A). Fixed-ratio 1 (FR1) training occurred first and for 7 days to stamp in learning, in which one press yielded reward. Then, the DLS was inhibited via halorhodopsin as the task was upshifted to FR3 (FR1 → FR3 rule change), which now required 3 presses for reward. Six additional normal FR3 sessions followed. There was then an FR3 test day with DLS inhibition (maintenance day, for comparison), followed by another normal FR3 session. Subsequently, DLS was inhibited as the task was downshifted to FR1 (FR3 → FR1). Then, after 6 normal FR1 days, a random time reinforcement schedule was imposed for 3 days (FR1 → RT30) as the DLS was again inhibited (Figure 1A). A control group received identical treatment but lacked halorhodopsin expression. In all of this, we allowed animals to perform in a self-directed manner in a free operant environment,[1] and we used a battery of behavioral measures to understand what aspects of their behavior and their performance strategies did and did not change during DLS inhibition. One important feature of this design is that animals could continue to use their previously learned behavioral strategy to get

reward, but with the shift in task rules that strategy might no longer be the most optimal for efficient performance. DLS inhibition was always 1-s in duration, and was triggered at the first lever press of a trial. Histology confirmed DLS expression (Figure 1B) and fiber placements (Figure 1C) in both groups (Figure 1D). Previous work[17] showed that roughly 83% of neurons in the DLS express these viral constructs.

### Upshift test day (FR1 → FR3): Animals with DLS inhibition continue prior FR1-like behavior and are slower in performance

For each test day of interest, we analyzed temporal aspects of behavior as well as the structure of actions, which included press amounts as well as action bout types. For action bouts, a 1-press bout was defined as a press of a lever followed by a food magazine entry before another press occurred. A 3-press bout was defined as three sequential presses and then a magazine entry. To normalize for potential press amount differences, we calculated the probabilities of these action bouts occurring in a given session. A ratio of 1-to-3 press bout probabilities was calculated for test days as a key measure of interest in showing to what extent animals' behavioral strategy was FR1-like (i.e., more 1-press bouts) or FR3-like (i.e., more 3-press bouts). This was investigated further by analyzing 1-press and 3-press bout probabilities on the test day as well as on a subsequent no-DLS-inhibition day for comparison. We similarly also analyzed 2-press bouts, 4+-press bouts, and pressing on the inactive lever.

#### *Temporal data: DLS inhibition slows performance*

On the upshift day rats were switched from FR1 to FR3, with DLS inhibition occurring on the first press of the action sequence. The overall time to complete the session was slower for both groups compared to the subsequent non-illumination FR3 day, and within the illumination upshift day the halorhodopsin group was slower than the control group (Figure 2A). This was shown by a significant interaction between day and group (estimate: 0.79, CI: 0.68–0.92, $p = 0.002$), a main effect of day (estimate: 0.63, CI: 0.54–0.73, $p < 0.001$), and no effect of subject group (estimate: 1.27, CI: 0.95–1.70, $p = 0.107$). As contributing factors to overall task completion time, the time spent between lever presses (i.e., inter-press intervals) was greater for halorhodopsin rats on the upshift day (Figure 2B) as shown by a significant interaction between day and group (estimate: 0.89, CI: 0.82–0.96, $p = 0.003$), and a main effect of day (estimate: 0.78, CI: 0.72–0.85, $p < 0.001$) but not group (estimate: 1.14, CI: 0.86–1.53, $p = 0.363$). Similarly, latencies between a magazine entry and the next lever press (Figure 2C) showed a main effect of day (estimate: 0.74, CI: 0.66–0.82, $p < 0.001$) and a significant interaction between day and group (estimate: 0.80, CI: 0.72–0.88, $p < 0.001$), but no effect of group (estimate: 0.89, CI: 0.63–1.27, $p = 0.524$), indicating longer magazine-to-press latencies during DLS inhibition. Lastly, we analyzed the latency between the last FR3 press and the subsequent reward contact within the magazine (Figure 2D), where there was a main effect of day (estimate: 0.90, CI: 0.86–0.94, $p < 0.001$) but not group (estimate: 1.09, CI: 0.89–1.34, $p = 0.406$), and there was not an interaction between day and group (estimate: 0.98, CI: 0.93–1.02, $p = 0.319$). Thus in general, DLS inhibition led to a slowing of behavior chiefly between presses and between a magazine entry and a press, while reward retrieval was less unimpacted. These differences were not due to the groups of animals spending time in different portions of the test chamber, as the majority of time for both groups was spent near the levers and magazine using photobeam analyses (Figure S1).

#### *Action strategy data: DLS inhibition favors 1-press bouts*

We first considered overall press rates on the active lever, which is a typical measure of how habitual responding is when task conditions change. Press rates were lower for DLS inhibition animals. Analysis on the number of lever presses performed per reward (Figure 2E) revealed no main effect of day (estimate: −0.01, CI: −0.05 to −0.02, $p = 0.392$) or day and group interaction (estimate: −0.01, CI: −0.04 to −0.02, $p = 0.458$). However, there was a main between groups effect (estimate: −0.07, CI: −0.13 to −0.01, $p = 0.027$), reflecting fewer presses per reward in the halorhodopsin group.

By this measure of active lever presses, it would appear that rats with DLS inhibition were less habitual in that when the task rules were shifted they pressed fewer times, which was closer to the most task-suitable number of presses to earn rewards in the session (i.e., 90 presses to earn 30-max rewards). The parallel decrease in performance speed, as aforementioned, would similarly support the conclusion that a habitual nature of responding was minimized.

However, the pattern of animals' pressing bouts told a different story. Animals can in principle display a range of action patterns reflective of an underlying strategy or internal model of the task. On the upshift day, as well as on the subsequent day, the majority of actions occurred in 1-press bouts (i.e., press then check the magazine before pressing again). This strategy had been optimal under the learned FR1 conditions. 3-press bouts were fewer in general, though they were the theoretically optimal task strategy in the new FR3 condition. In comparing experimental conditions, animals with DLS inhibition had more 1-press than 3-press bouts as reflected in a roughly doubled 1-to-3 pressing bout ratio relative to controls ($F = 8.74$, $p = 0.012$; Figure 2F). This result shows that DLS inhibition led to poorer adaptation of animals to the new FR3 condition. We next explored these bout types as probabilities for the inhibition day and the subsequent no-inhibition day. For 1-press bouts alone, there was a trend for DLS to elevate them, but there was no day effect (estimate: 0.95, CI: 0.83–1.08, $p = 0.428$), no day and group interaction (estimate: 1.12, CI: 0.92–1.36, $p = 0.246$), and no group effect (estimate: 1.89, CI: 0.89–4.05, $p = 0.099$) (Figure 2G). For 3-press bouts, on the illumination and subsequent days, there was no effect of day (estimate: 0.93, CI: 0.72–1.19, $p = 0.546$) or day and group interaction (estimate: 0.69, CI: 0.38–1.23, $p = 0.203$) (Figure 2H). However, there was a main between groups effect (estimate: 0.19, CI: 0.06–0.61, $p = 0.005$) due to the halorhodopsin group performing significantly fewer 3-press bouts on the illumination day, a difference that carried over to the subsequent day as well.

Animals did also perform 2-press bouts. However, DLS inhibition had no effect on the probability of those (group estimate: 0.97, CI: 0.29–3.23, $p = 0.963$; day estimate: 1.13, CI: 0.86–1.48, $p = 0.380$; interaction estimate: 0.85, CI: 0.57–1.27, $p = 0.431$). 4+ pressing bouts were rare, and they were likewise unaffected by DLS inhibition (group estimate: 212.13, CI: 0.14–312004.01, $p = 0.150$; day estimate: 5.18, CI: 0.60–44.69,

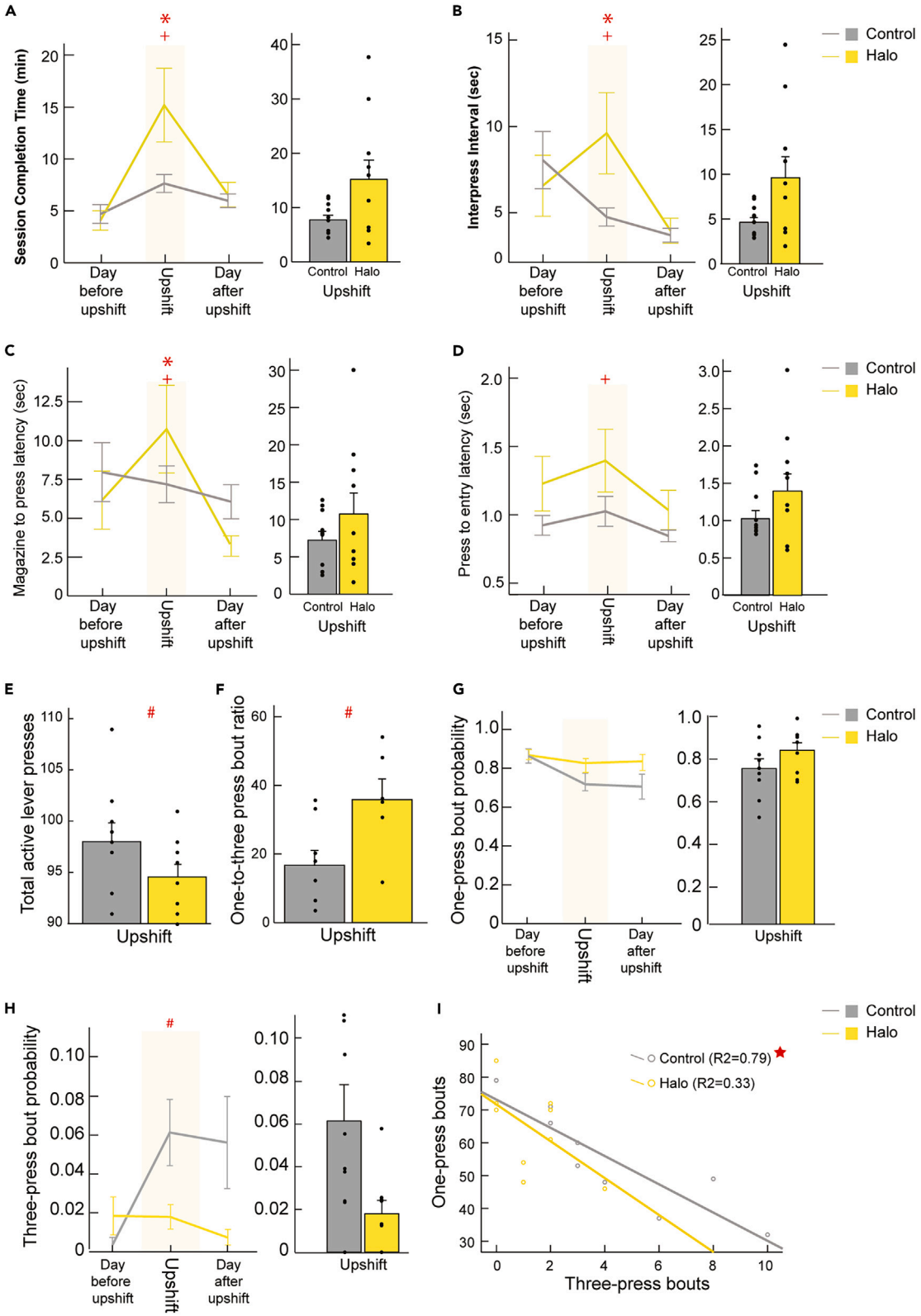

A
Session Completion Time (min)
Day before upshift
Upshift
Day after upshift
Control Halo
Upshift
B
Interpress Interval (sec)
Control
Halo
C
Magazine to press latency (sec)
D
Press to entry latency (sec)
E
Total active lever presses
F
One-to-three press bout ratio
G
One-press bout probability
H
Three-press bout probability
I
One-press bouts
Three-press bouts
Control (R2=0.79)
Halo (R2=0.33)

**Figure 2. Upshift illumination day**

Illumination days denoted by yellow background shading, with days preceding and following the illumination day shown in line plots and illumination day only shown in bar plots with individual data points.

(A) Mean overall time (min) to complete the session.

(B) Mean inter-press intervals (sec).

(C) Mean reward contact (magazine entry) to press latency (sec).

(D) Mean lever press to a magazine entry (sec).

(E) Mean total active lever presses per session on the upshift test day with light delivery.

(F) Ratio of 1-press to 3-press bouts (1-press/3-press per rat) on the upshift test day with light delivery.

(G) Mean 1-press bout probability.

(H) Mean 3-press bout probability.

(I) Scatterplot of 1- and 3-press bout numbers on the upshift day with regression lines and $R^2$ values. Control, gray; eNpHR3.0, yellow. For all figure panels, bars and errors show mean ± SEM; asterisks denote significant comparisons (*$p < 0.05$ day/group interaction; +$p < 0.05$ day effect; #$p < 0.05$ group effect).

$p = 0.135$; interaction estimate: 0.14, CI: 0.01–2.00, $p = 0.149$). Inactive lever presses were very few and increased slightly during the upshift session, but they were unrelated to DLS inhibition (group estimate: 1.55, CI: 0.74–3.23, $p = 0.246$; day estimate: 0.39, CI: 0.22–0.72, $p = 0.002$; interaction estimate: 1.76, CI: 0.83–3.73, $p = 0.138$) (Figure S2).

Overall, pressing behaviors that varied by group tended to be the more objectively task-relevant ones, namely 1- and 3-press bouts. The pressing structure of animals with DLS inhibition showed that they favored doing more of what they had learned to do before the upshift, that being more 1-press bouts relative to 3-press bouts. This 1-press strategy would naturally lead to fewer overall presses in FR3 conditions, given that it involved more regular magazine checks until reward occurred, and thus the bout structure suggested that these animals were more stuck in their prior FR1-like routines if anything, not less, and thus more habitual in that manner of thinking.

To gain further insight into the underlying strategy of behavior, we correlated 1-press and 3-press bouts. Although there is a zero-sum relationship between these bout types across the fixed set of session trials, there is room for more or less correlation to occur. In principle, a strong (negative) correlation would indicate an internal model of "understanding" that the 1-press and 3-press options are linked in some manner for the task. Across control animals, there was indeed a strong and significant negative correlation ($R^2 = 0.79$; $p = 0.001$), meaning the more they performed 3-press bouts the less they performed 1-press bouts (Figure 2I). In contrast, across DLS inhibition animals, there was a weaker correlation ($R^2 = 0.33$) that was not significant ($p = 0.10$). We interpret this difference to indicate that 1- and 3-press bout options were related within the control animals' internal model of the task. The lower correlation in animals with DLS inhibition would indicate that the internal task representation of these animals was more focused on 1-press bouts, which could well have impeded learning of the new FR3 rules. In principle, one might expect a prospective model of the task to incorporate task-relevant 1- and 3-press options as related, while a retrospective habit-like model may contain mostly just the prior options that were used. With DLS inhibition, the latter was seen, suggesting that a prospective model might normally engage the DLS.

To summarize, DLS inhibition slowed performance, biased animals toward a less task-optimal FR1-like pattern of lever pressing that resulted in fewer overall presses, and seemed to prevent animals from forming a relational understanding of 1-press and 3-press bout options in their task model.

## FR3 maintenance day: In the absence of a task change, DLS inhibition is less effective

After the upshift illumination day, all rats were trained under the new FR3 requirement for the next seven days. To test whether DLS inhibition would have a similar effect on ongoing task performance (i.e., when there is no change in the action requirement for reward), a maintenance FR3 day was conducted during which DLS was inhibited at the onset of pressing bouts.

### *Temporal data: Little effect of DLS inhibition*

For task completion time, there was a main effect of day (estimate: 0.79, CI: 0.68–0.92, $p = 0.003$) and a day/group interaction (estimate: 0.73, CI: 0.63–0.86, $p < 0.001$), but not a group effect (estimate: 1.11, CI: 0.81–1.52, $p = 0.534$) (Figure 3A). Halorhodopsin animals were thus slightly slower on the illumination day, while if anything slightly faster on the subsequent day. Inter-press intervals (Figure 3B) followed the same trend, with main effects of day (estimate: 0.90, CI: 0.83–0.96, $p = 0.003$) and a day/group interaction (estimate: 0.92, CI: 0.85–0.98, $p = 0.016$), but no group effect (estimate: 0.93, CI: 0.68–1.27, $p = 0.639$). For the latency between from a magazine entry to a press, (Figure 3C), there was a day/a group interaction (estimate: 0.93, CI: 0.88–0.98, $p = 0.012$); however, neither day (estimate: 0.96, CI: 0.91–1.02, $p = 0.157$) or group (estimate: 0.73, CI: 0.49–1.08, $p = 0.117$) showed an effect. Likewise, the latency between a press and magazine entry (Figure 3D) had a day/group interaction (estimate: 0.91, CI: 0.85–0.98, $p = 0.018$) but not an effect of group (estimate: 1.01, CI: 0.85–1.21, $p = 0.877$) or day (estimate: 0.92, CI: 0.85–0.99, $p = 0.031$). In short, there were some rather weak and inconsistent trends for performance durations with DLS inhibition on this maintenance day. In photobeam analyses, the animals again spent most of their time near the levers and magazine (Figure S1).

### *Action strategy data: Little effect of DLS inhibition*

By receiving the previous 7 FR3 training days, animals had moved further in the direction of behaving in closer alignment with the FR3 rule. 1-press bouts were still dominant but tended to be less numerous than the first day of FR3 experience (upshift day), while 3-press bouts showed the reverse trend and grew higher. In comparing conditions, total lever presses were more than the minimally required 3 per reward

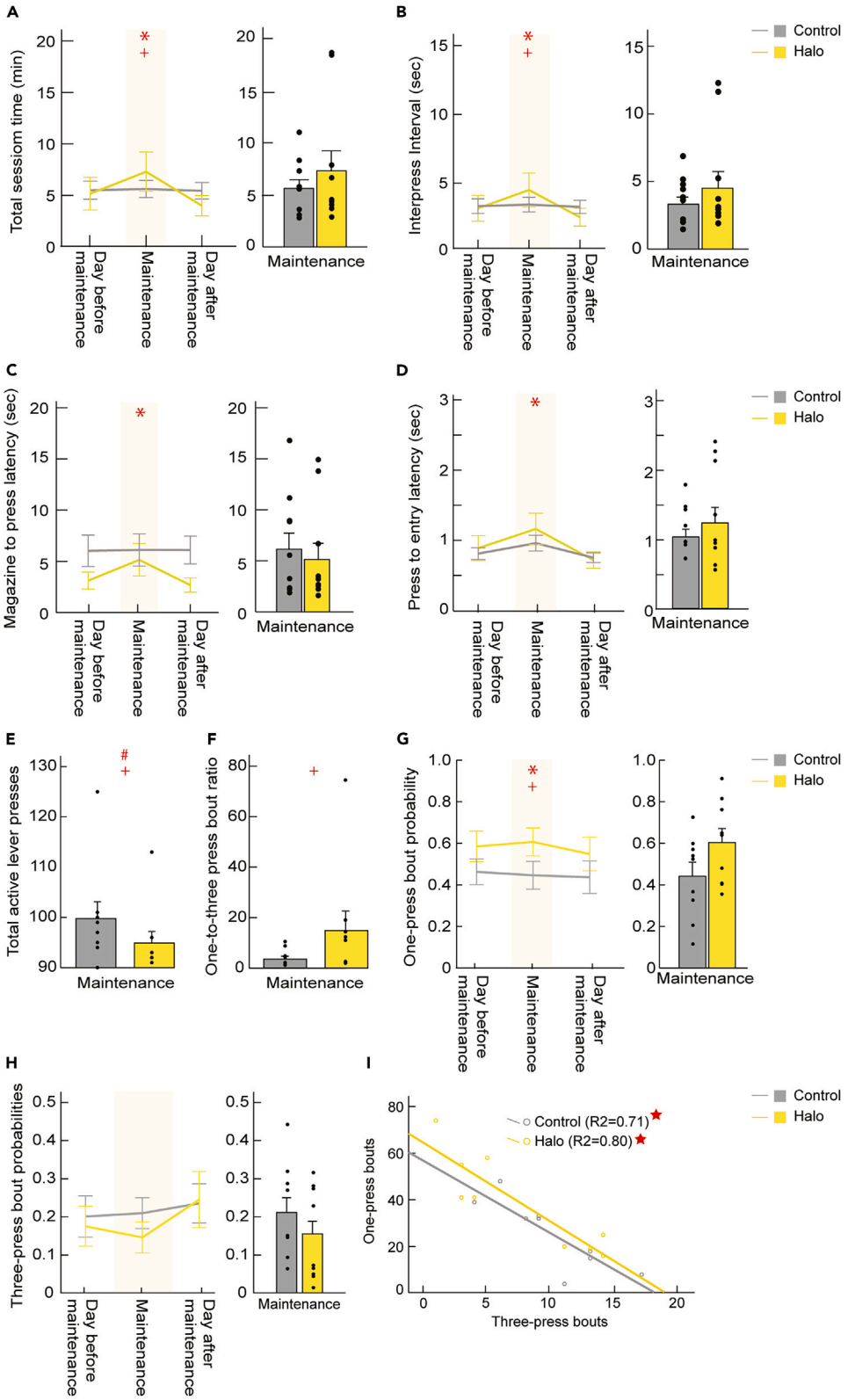

A
Total sessiom time (min)
Day before maintenance
Maintenance
Day after maintenance
Maintenance
B
Interpress Interval (sec)
Control
Halo
C
Magazine to press latency (sec)
D
Press to entry latency (sec)
E
Total active lever presses
F
One-to-three press bout ratio
G
One-press bout probability
H
Three-press bout probabilities
I
One-press bouts
Three-press bouts
Control (R2=0.71)
Halo (R2=0.80)

**Figure 3. FR3 maintenance illumination**

Illumination days denoted by yellow background shading, with days preceding and following the illumination day shown in line plots and illumination day only shown with individual data points.

(A) Mean overall time (min) to complete the session.

(B) Mean inter-press intervals (sec).

(C) Mean reward contact (magazine) to press latency (sec).

(D) Mean lever press to a magazine entry (sec).

(E) Mean total active lever presses per session on the upshift test day with light delivery.

(F) Ratio of 1-press to 3-press bouts (1-press/3-press per rat) on the upshift test day with light delivery.

(G) Mean 1-press bout probability.

(H) Mean 3-press bout probability.

(I) Scatterplot of 1- and 3-press bout numbers on the upshift day with regression lines and $R^2$ values. Control, gray; eNpHR3.0, yellow. For all figure panels, bars and errors show mean ± SEM; asterisks denote significant comparisons (*$p < 0.05$ day/group interaction; +$p < 0.05$ day effect; #$p < 0.05$ group effect).

for both groups, and trended lower for the halorhodopsin group (Figure 3E). There was a main effect of day (estimate: −0.09, CI: −0.15 to −0.02, $p = 0.015$) and group (estimate: −0.35, CI: −0.64 to −0.05, $p = 0.021$), but not a day and group interaction (estimate: 0.01, CI: −0.09 to 0.11, $p = 0.822$). Animals in the halorhodopsin group had a trend toward performing more 1-press bouts preceding, during, and following the day of DLS inhibition. The ratio of 1-to-3 press bouts on this maintenance day was a bit higher in halorhodopsin animals (Wilcoxon test due to non-normal data; W = 29, $p = 0.034$) (Figure 3F). This ratio was reflected in 1-press bout probabilities (Figure 3G), which had a day and group interaction (estimate: 1.51, CI: 1.11–2.07, $p = 0.009$) and group effect (estimate: 2.76, CI: 1.16–6.56, $p = 0.022$), but not day effect (estimate: 1.04, CI: 0.86–1.26, $p = 0.690$). 3-press bouts showed little variation during these testing days (Figure 3H): there was no main effect of group (estimate: 0.37, CI: 0.12–1.13, $p = 0.082$), day (estimate: 1.00, CI: 0.72–1.39, $p = 1.000$), or a day and group interaction (estimate: 0.56, CI: 0.29–1.09, $p = 0.091$). The probability of pressing in 2-press bouts was unaffected (group estimate: 1.14, CI: 0.11–11.39, $p = 0.911$; day estimate: 0.93, CI: 0.70–1.24, $p = 0.632$; interaction estimate: 0.93, CI: 0.61–1.42, $p = 0.744$). Pressing in 4+-press bouts was also unaffected (group estimate: 22.56, CI: 0.05–9999.67, $p = 0.316$; day estimate: 1.04, CI: 0.62–1.77, $p = 0.870$; interaction estimate: 0.46, CI: 0.15–1.40, $p = 0.171$). Although pressing on the inactive lever was very low in frequency, animals with DLS inhibition did it marginally more (group estimate: 7.61, CI: 2.26–25.67, $p = 0.001$; day estimate: 3.67, CI: 1.36–9.91, $p = 0.010$; interaction estimate: 0.19, CI: 0.06–0.58, $p = 0.004$) (Figure S2). Still, they averaged only 2 inactive presses, so we hesitate to draw conclusions.

In assessing the correlation between 1-press and 3-press bouts (Figure 3I), controls maintained a strong and significant negative relationship between these bout types ($R^2 = 0.71$; $p = 0.005$). In stark contrast to the upshift day, animals with DLS inhibition also had a strong negative correlation between the bout types on this maintenance test day ($R^2 = 0.82$; $p = 0.001$). To the extent that this correlation reflects an internal understanding that 1-and 3-bouts are related options, DLS inhibition on this maintenance day did not affect it robustly.

In summary, DLS inhibition created modest and somewhat inconstant changes in performance rate, and did not seem to impact how animals related together the 1-press and 3-press bout options in their internal model of the task. Limited effects of DLS inhibition on reward-seeking behavior in the absence of any task change are consistent with prior studies, and it could relate to other brain areas such as the dorsomedial striatum (DMS) playing a role in performance as well.[17,23,24] Alternatively, it could indicate a relatively insignificant role for the DLS in performing acquired actions when nothing about them needs adjusting.

We additionally looked at the effect of DLS inhibition given at the time of the third press lasting through reward procurement during the FR3 maintenance phase of the study. The logic was that, in particularly speedy animals, some illumination epochs could infiltrate the reward procurement period when illumination onset occurred at the first press. However, this DLS inhibition at the third press produced no changes that bore any similarity with our other test days (Table S1), suggesting that the effects of DLS inhibition occurring at action onset were unrelated to light delivery occurring between the termination of the action sequence and subsequent magazine entry. There were a few data trends across the test day and day after, but nothing of relevance to DLS roles in the task (Table S1).

## Downshift day (FR3 → FR1): DLS inhibition again results in animals favoring initial FR-1 like behaviors, but this time speeds up performance

In the next stage of the experiment, the task rules were again suddenly changed from FR3 back to FR1 and the DLS was inhibited at action onset. This shift to the initially learned FR1 rule creates a scenario in which the new task rules are a change from what was just experienced, but they are also familiar. Recall from the aforementioned text that DLS inhibition during the upshift led to animals favoring their learned FR1-like strategy, as though they were stuck in that routine despite the change in task rules to FR3. If this effect occurred because the animals were poorer at adapting to a change in rules, as though unable to update their model of the task, then DLS inhibition during the downshift here should cause them to be similarly less able to change what they were just doing and have a poorer adaptation to FR1. If, on the other hand, the upshift results occurred because animals reverted to what they initially learned to do, then DLS inhibition during the downshift here should cause them to again do more FR1-like behaviors. The latter effect was seen.

### *Temporal data: Animals with DLS inhibition are fast*

For either of the aforementioned possible results, a decrease in performance speed was predicted to result from DLS inhibition. Strikingly, however, rats with DLS inhibition were exceptionally fast in their completion of the task on this downshift day (Figure 4A). Looking across the

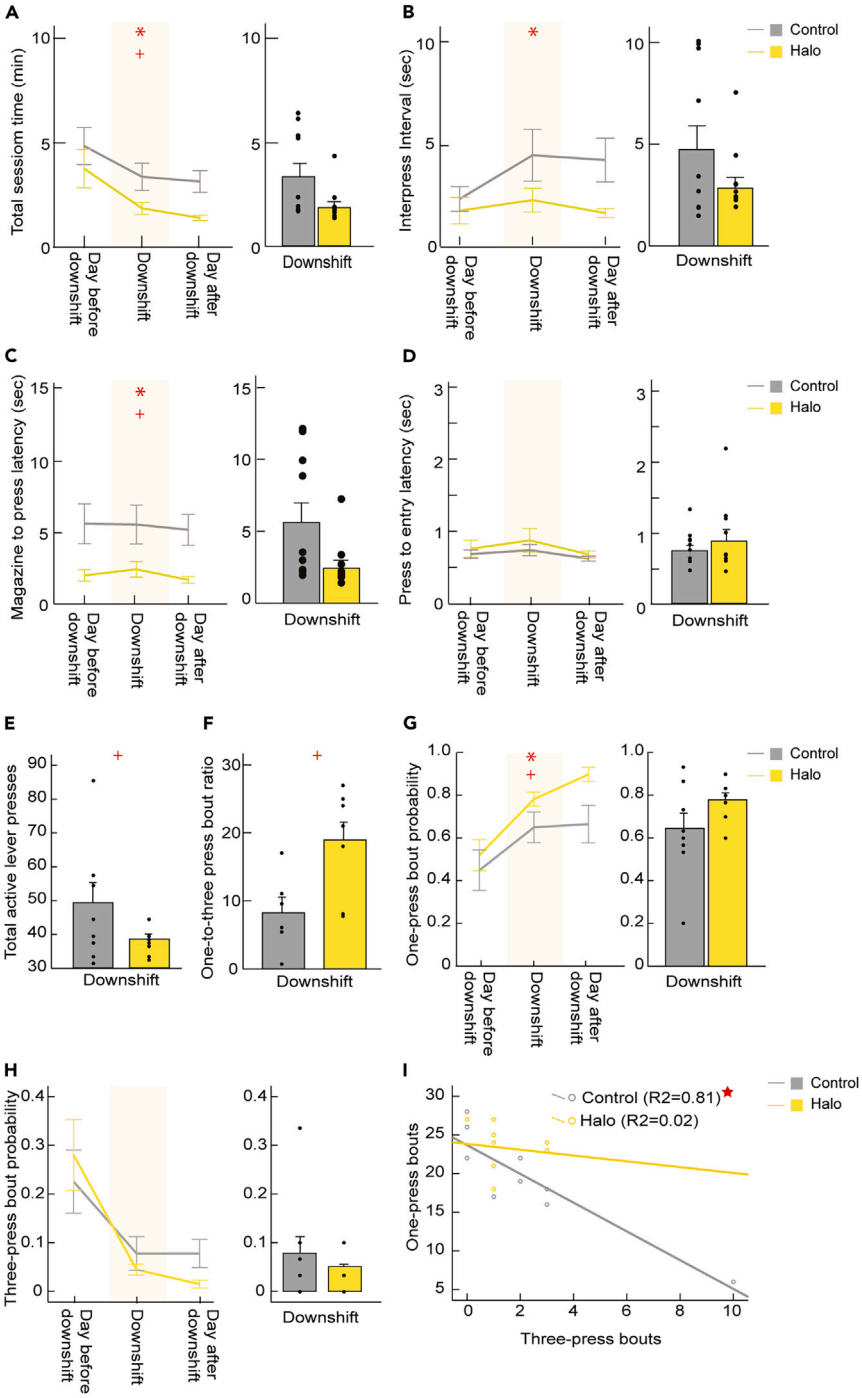



**Figure 4. Downshift illumination day**

Illumination days denoted by yellow background shading, with days preceding and following the illumination day shown in line plots and illumination day only shown in bar plots with individual data points.

(A) Mean overall time (min) to complete the session.

(B) Mean inter-press intervals (sec).

(C) Mean reward contact (magazine) to press latency (sec).

(D) Mean lever press to a magazine entry (sec).

(E) Mean total active lever presses per session on the upshift test day with light delivery.

(F) Ratio of 1-press to 3-press bouts (1-press/3-press per rat) on the upshift test day with light delivery.

**Figure 4. *Continued***
(G) Mean 1-press bout probability.
(H) Mean 3-press bout probability.
(I) Scatterplot of 1- and 3-press bout numbers on the upshift day with regression lines and $R^2$ values. Control, gray; eNpHR3.0, yellow. For all figure panels, bars and errors show mean ± SEM; asterisks denote significant comparisons (*$p < 0.05$ day/group interaction; +$p < 0.05$ day effect; #$p < 0.05$ group effect).

last FR3 day and the downshift test day, there were significant effects of day (estimate: 0.88, CI: 0.83–0.92, $p < 0.001$) and a day/group interaction (estimate: 0.90, CI: 0.85–0.95, $p < 0.001$). There was no group effect (estimate: 0.87, CI: 0.66–1.16, $p = 0.349$). Mean inter-press intervals followed a similar trend (Figure 4B). While there were no effects of day (estimate: 1.02, CI: 0.96–1.08, $p = 0.467$) or group (estimate: 0.68, CI: 0.34–1.37, $p = 0.281$), there was a significant day/group interaction (estimate: 0.91, CI: 0.83–0.98, $p = 0.018$). Latency between magazine entries and the next press (Figure 4C) revealed no main effect of day (estimate: 1.00, CI: 0.95–1.06, $p = 0.882$), which related to some pre-existing group differences; however, we did find significant effects for group (estimate: 0.44, CI: 0.22–0.89, $p = 0.023$) and a day/group interaction (estimate: 0.86, CI: 0.80–0.93, $p < 0.001$). Lastly, mean time between a press and subsequent reward contact (Figure 4D) did not show any effects of day (estimate: 0.93, CI: 0.84–1.02, $p = 0.126$), group (estimate: 1.11, CI: 0.81–1.51, $p = 0.525$), or a day/group interaction (estimate: 0.99, CI: 0.86–1.13, $p = 0.865$). Thus, the robustly sped-up task completion time in halorhodopsin rats reflected shorter time between presses and shorter time to transition from magazine entry back to pressing. These were the same task moments that were most affected (and oppositely affected) by DLS inhibition during the earlier FR1→FR3 upshift test. As before, the majority of time for both groups was spent near the levers and magazine judging by photobeam breaks (Figure S1), so physical positioning seemed not to clearly underlie group differences.

### *Action strategy data: DLS inhibition again favors 1-press bouts*

Animals with DLS inhibition on this downshift day pressed significantly fewer times, reflective of greater performance efficiency given that only one press was needed per reward (Figure 4E). There was a main effect of day (estimate: −0.09, CI: −0.15 to −0.02, $p = 0.015$) and group (estimate: −0.35, CI: −0.64 to −0.05, $p = 0.021$), but no day and group interaction (estimate: 0.01, CI: −0.09 to 0.11, $p = 0.822$). Underlying this effect was a significantly elevated FR1-like strategy of behavior: animals with DLS inhibition had an over doubled 1-to-3 press bout ratio compared to controls ($F = 6.98$, $p = 0.022$; Figure 4F). In examining each bout types separately, the DLS inhibition group performed more 1-press bouts on the illumination day (Figure 4G). This was reflected by a main effect of group (estimate: 2.76, CI: 1.16–6.56, $p = 0.022$) and a day/group interaction (estimate: 1.51, CI: 1.11–2.07, $p = 0.009$), but no day effect (estimate: 1.04, CI: 0.86–1.26, $p = 0.690$). There was a trend toward fewer 3-press bouts in rats with DLS inhibition (Figure 4H), but there was not a main effect of day (estimate: 1.00, CI: 0.72–1.39, $p = 1.000$), a day/group interaction (estimate: 0.56, CI: 0.29–1.09, $p = 0.091$), or a group effect (estimate: 0.37, CI: 0.12–1.13, $p = 0.082$). The probability of 2-press bouts did not diverge between groups on the downshift inhibition day, but there was a difference on the day after with the DLS inhibition group doing a bit fewer (group estimate: 4699.64, CI: 15.02–1470782.16, $p = 0.004$; day estimate: 1.35, CI: 0.89–2.06, $p = 0.160$; interaction estimate: 0.34, CI: 0.17–0.67, $p = 0.002$). 4+-press bouts were few, declined during this period, and were unaffected by DLS inhibition (group estimate: 0.00, CI: 0.00–8403377.86, $p = 0.600$; day estimate: 0.32, CI: 0.12–0.84, $p = 0.021$; interaction estimate: 1.55, CI: 0.11–20.97, $p = 0.741$). Inactive lever presses were rare as usual and did not change (group estimate: 5.90, CI: 0.32–109.07, $p = 0.233$; day estimate: 0.00, CI: 0.00–inf, $p = 0.979$; interaction estimate: 281554219.88, CI: 0.00–inf, $p = 0.979$) (Figure S2).

As with the upshift day, DLS inhibition also appeared to occlude animals' understanding of 1-press and 3-press bouts being related, as this correlation was low ($R^2 = 0.02$; $p = 0.75$) compared to control animals in which it was significantly higher ($R^2 = 0.81$; $p < 0.001$) (Figure 2I). Thus, for both task shift days, animals favored 1-press bouts and lost an internal relatedness of 1- and 3-press bouts. However, while DLS inhibition slowed performance on the upshift day, it sped up performance on the downshift day.

## RT30 days: DLS inhibition leads to persistent and rapid performance

Changes in the timing and frequency of lever-press actions during upshift and downshift days could possibly be related to greater/lesser probability of reward occurring given the action. This is difficult to discern in FR schedules, as pressing directly produces reward delivery. Thus, as a final experimental phase we exposed animals to a shift from FR1 to a random-time 30 (RT30) schedule. Under RT30, rewards were delivered independently of pressing every ~30 s. This schedule affords an opportunity to evaluate how changes in lever press behavior might relate to how animals experience the relationships between the action and reward,[38–40] which we tested for three sequential RT30 days with DLS illumination. Concerning hypotheses, the prior tests indicated that DLS inhibition increased the use of the initially acquired 1-press-bout strategy, suggesting a DLS role in prospectively adapting behavior to task changes that occurred beyond that initial FR1 stage. Here, during RT, if the DLS can carry such a prospective function, DLS inhibition ought to cause animals to exhibit poorer adaptation—i.e., favoring 1-press bouts again rather than more optimally reducing pressing altogether.

### *Temporal data: Animals with DLS inhibition are fast*

Animals with DLS inhibition exhibited a remarkably faster performance rate compared to controls. Inter-press-intervals expectedly rose in control animals over RT30 days, while these intervals in animals with DLS inhibition remained low (i.e., faster). This was seen in a day by group interaction (estimate: 0.70, CI: 0.52–0.0.94, $p = 0.019$) (Figure 5A). Animals with DLS inhibition also exhibited a trend toward a shorter time between a magazine entry and continuation of pressing behavior, but the interaction of day and group was not significant (estimate: 0.98, CI: 0.82–1.17, $p = 0.804$) (Figure 5B). Similarly, they showed lower press-to-magazine entry times; day by group interaction; estimate: 0.83,

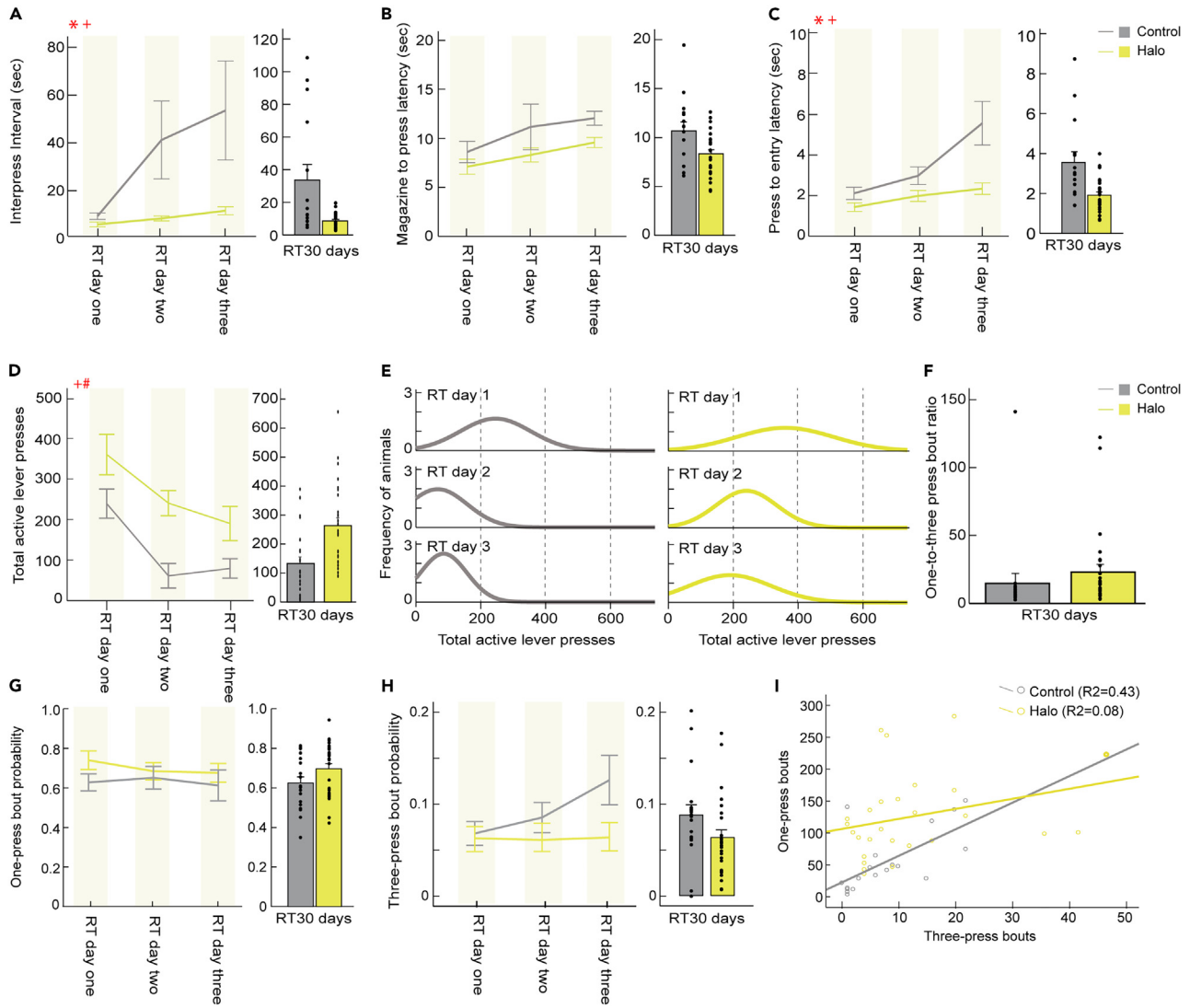



**Figure 5. Random time-30 days**

Light delivery on each day in line plots. Line plots show each illumination day; bars show averaged data over the 3 illumination days with individual data points.

(A) Mean inter-press intervals time (sec).

(B) Mean reward contact (magazine) to press latency (sec).

(C) Mean press-to-magazine entry (sec).

(D) Mean total active lever presses.

(E) Smoothed histogram denoting frequency of animals on each RT day by their total pressing levels, by group.

(F) Ratio of 1-press to 3-press bouts (1-press/3-press per rat) on the upshift test day with light delivery.

(G) Mean 1-press bout probability.

(H) Mean 3-press bout probability.

(I) Scatterplot of 1- and 3-press bout numbers on the upshift day with regression lines and $R^2$ values. Control, gray; eNpHR3.0, yellow. For all figure panels, bars and errors show mean ± SEM; asterisks denote significant comparisons (*$p < 0.05$ day/group interaction; +$p < 0.05$ day effect; #$p < 0.05$ group effect).

CI: 0.70–0.97, $p = 0.021$) (Figure 5C). In total, the RT30 sessions were characterized in control animals by a slowing of actions over the days, while DLS inhibition caused animals to maintain a relatively speedier performance. In photobeam analyses, both groups spent their time near the levers and magazine, which was particularly true of the halorhodopsin group (Figure S1).

### *Action strategy data: DLS inhibition results in more 1-press bouts*

Animals with DLS inhibition appeared quite persistent in their lever pressing behavior over the RT days. Both groups had very high levels of pressing on the first RT30 day, due likely to the longer session time available to them. While both groups declined pressing over the two

subsequent RT30 days, animals with DLS inhibition pressed considerably more than controls throughout this experimental phase (Figure 5D). Analysis of this overall pressing revealed both a group (estimate: 4.27, CI: 0.32–8.22, $p = 0.034$) and session effect (estimate: −2.74, CI: −3.42 to −2.07, $p < 0.001$), but not their interaction (estimate: −0.17, CI: −1.53 to 1.19, $p = 0.806$). In looking at the distribution of pressing across animals, controls clustered into lower levels of pressing as the RT30 days went on, while halorhodopsin animals maintained higher pressing rates and their pressing was more distributed (i.e., more variable) (Figure 5E).

As the animals with DLS inhibition performed more lever presses overall, their tendency was to perform 1-press bouts. Concerning probabilities, which normalizes across different overall press rates, both groups preferred 1-press bouts when they engaged with the lever (day by group interaction only; estimate: 1.22, CI: 1.03–1.46, $p = 0.024$) and the 1/3 bout ratio was no different between them ($F = 0.34$, $p = 0.571$) (Figures 5F and 5G). The occurrence of 3-press bouts was infrequent, and their probability was not different between experimental groups (estimate: 0.92, CI: 0.67–1.25, $p = 0.577$) (Figure 5H). This meant that DLS inhibition led to more overall presses, but when that was normalized, the groups were similar in favoring 1-press bouts. During the RT30 days, 2-press bouts were not significantly different in probability between groups (group estimate: 0.03, CI: 0.00–4.53, $p = 0.165$; day estimate: 0.69, CI: 0.45–1.04, $p = 0.078$; interaction estimate: 1.38, CI: 0.88–2.19, $p = 0.164$). The probability of 4+ press bouts was less for the DLS inhibition group, but the occurrences of these bouts was still very infrequent (group estimate: 527.27, CI: 523.17–531.41, $p < 0.001$; day estimate: 2.70, CI: 2.68–2.73, $p < 0.001$; interaction estimate: 0.54, CI: 0.53–0.54, $p < 0.001$). Inactive lever pressing also increased similarly in both groups during this period, though they were fractionally less than $^1/_4$ of total presses (group estimate: 0.98, CI: 0.44–2.16, $p = 0.958$; day estimate: 0.99, CI: 0.71–1.36, $p = 0.928$; interaction estimate: 0.98, CI: 0.68–1.41, $p = 0.906$) (Figure S2). This uptick in inactive lever presses could reflect task exploration or confusion during the RT phase of the study, but DLS inhibition did not affect it.

Similar to the effects on prior test days, however, was the lack of relatedness between 1-press and 3-press bouts per animal; these bouts were essentially uncorrelated for DLS inhibition animals ($R^2 = 0.081$) compared to controls ($R^2 = 0.33$) across the RT30 days, although neither correlation reached significance (Figure 5I). As the number of actions was much less constrained during RT30, such that doing more of one and not necessarily doing the other, the correlation between bout types flipped to a positive one.

We next looked at the probability of a lever press action just prior to, and just after, reward delivery. Probabilities were individually set per animal as a rounded mean inter-press interval over the session. In both groups, actions were more likely to occur prior to reward delivery than after reward delivery, and pressing in both time windows declined over the course of the sessions, as befitting the RT30 schedule. However, animals with DLS inhibition performed significantly more interactions both before and after reward delivery. In analyses, pre-reward action probability (Figure 6A) did not reveal a group effect (estimate: 0.73, CI: 0.28–1.91, $p = 0.521$) but there were significant effects for both session (estimate: 0.52, CI: 0.39–0.68, $p < 0.001$) and the group/session interaction (estimate: 2.01, CI: 1.15–3.52, $p = 0.015$). The probability of an action just after reward (Figure 6B) followed an identical pattern with no group effect (estimate: 0.63, CI: 0.17–2.40, $p = 0.503$) but significant effects for both session (estimate: 0.50, CI: 0.37–0.68, $p < 0.001$) and the group/session interaction (estimate: 2.17, CI: 1.19–3.95, $p = 0.012$). The trends were also somewhat reflected within the peri-reward time window (+/− 1 s) (Figure 6D), in which presses per minute was significant for session (estimate: −4.28, CI: −5.94 to −2.62, $p < 0.001$), but was not significant for group (estimate: 6.63, CI: −3.60 to 16.86, $p = 0.204$) or the group/session interaction (estimate: 0.33, CI: −2.99 to 3.64, $p = 0.846$). These data indicated no major differences between groups in *when* actions would occur, but when they did occur. Overall press rate between groups were unlikely to be related to differences in extra positive reinforcement gained in DLS inhibition animals by virtue of their greater overall action frequency, as these animals distributed their increased pressing before reward as well as after reward (rather than just before reward). Results are similarly unlikely to be explained by DLS inhibition improving the timing of actions to reflect a ∼30 s reward delivery schedule; in such a case we would have expected more of a piling up of actions just preceding reward delivery in the DLS inhibition group more so than in controls.

Further analysis was undertaken for both action probability and lever press rate with regard to DLS illumination timing (Figure 6E). For this, lever press rates outside the laser delivery time window were considered distinct from those occurring within the light administration time period, and a press that triggered the illumination itself was considered to occur outside of the illumination period. Analyses showed a significant effect of pressing rates between within-illumination time vs. outside-of-illumination time (estimate: −15.47, CI: −20.87 to −10.08, $p < 0.001$), but no effects were found for group (estimate: 1.28, CI: −4.90 to 7.47, $p = 0.684$), session (estimate: −0.98, CI: −3.07 to 1.11, $p = 0.357$), or any interactions between group and session (estimate: −0.07, CI: −2.67 to 2.54, $p = 0.960$), group and illumination time (estimate: 4.46, CI: −2.27 to 11.19, $p = 0.194$), session and illumination time (estimate: −3.19, CI: −7.37 to 0.99, $p = 0.134$), or a group/session/illumination time interaction (estimate: −0.90, CI: −6.12 to 4.31, $p = 0.734$). Concerning action probability, this was calculated as a lever press occurring during the illumination period (binary: yes/no) as each lever press resulted in the triggering of a laser pulse during the RT session. We did not find any significant effects for group (estimate: 0.78, CI: 0.32–1.88, $p = 0.582$), session (estimate: 1.04, CI: 0.85–1.27, $p = 0.690$), or a group/session interaction (estimate: 1.03, CI: 0.80–1.33, $p = 0.806$) (Figure 6F). Thus, actions tended to be performed more frequently within the illumination period regardless of experimental group.

Collectively, animals with DLS inhibition during the RT sessions were likely to perform more lever press actions (particularly 1-press bouts), perform these behavioral routines more rapidly, and to disassociate 1-press and 3-press options in their presumed internal model of the task conditions. Persistence of action during RT and related conditions is a classic demonstration of habitual control over behavior; here, it occurred with DLS inhibition. Although some studies have found more rapid adjustments to action-outcome contingency changes in instrumental tasks during DLS inhibition, such as to a new schedule of omission in which actions are punished by reward loss, these studies typically involve a history of task rule changes leading to an interval schedule ahead of this contingency degradation test day (e.g., a study by Yin et al.[24]). Differences in task history here versus in other work are likely relevant to how DLS inhibition affects behavior when task rules are altered.

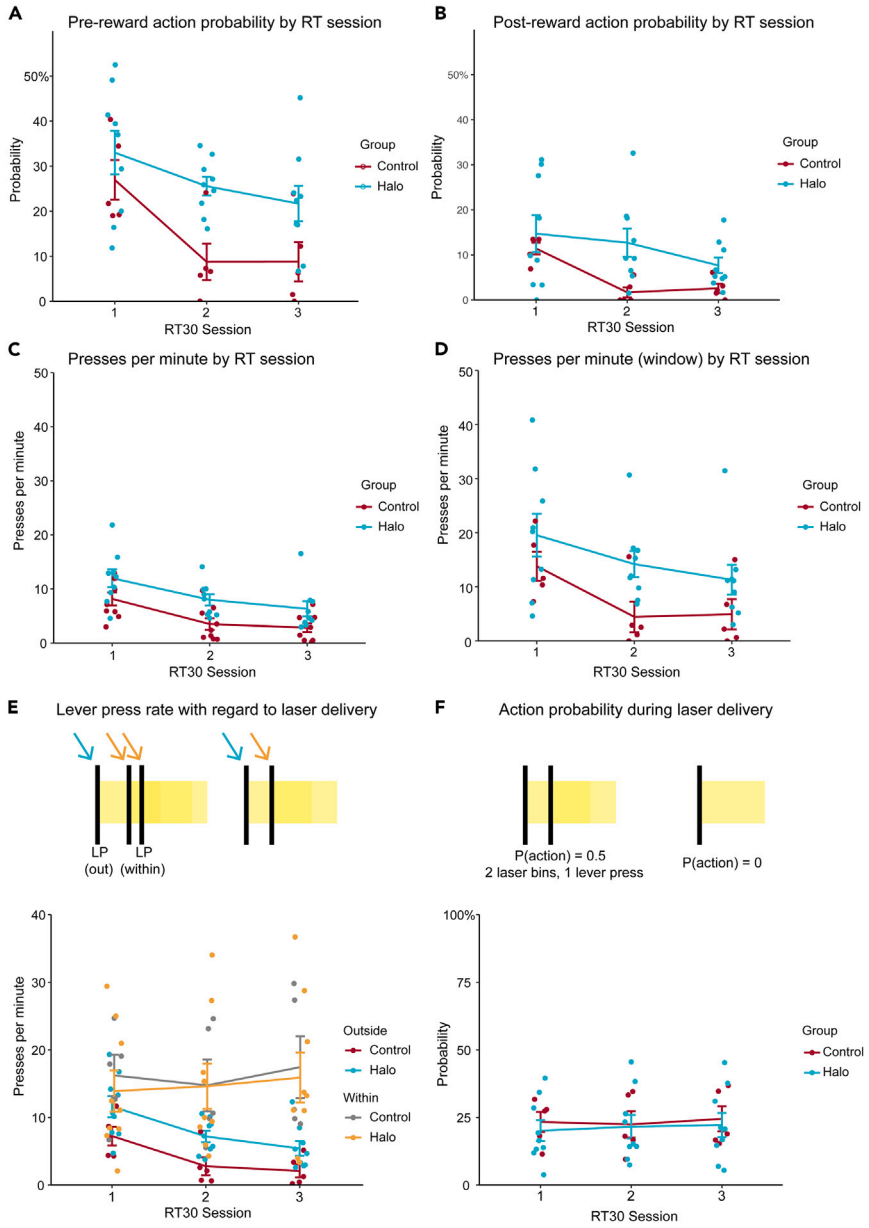



**Figure 6. RT30 illumination sessions**

1-s continuous DLS illumination on every press (reset by magazine entry).

(A) Pre-reward action probability by RT session.

(B) Post-reward action probability by RT session.

(C) Presses per minute by RT session.

(D) Presses per minute (illumination/laser 1-s time window) by RT session.

(E) Lever press rate comparison without versus within laser delivery time window.

(F) Action probability during laser delivery; action probability (set per animal as its rounded mean inter-press interval per session).

For all figure panels, bars and errors show mean +/- SEM.

## DISCUSSION

Behavioral persistence—the insensitivity of behavior to environmental change—is usually thought to reflect habits that result from stimulus-response or model-free learning. Brain areas like the DLS that are important for behavioral persistence are thus identified as key brain nodes for habits. However, there have long been arguments that behavioral persistence could also arise in principle from a prospective, planning type of process.[2–4,41,42] Moreover, in studies linking the DLS with behavioral persistence, it is often unclear to what extent such persistence is

reflexive and "mindless" versus a reflection of an active underlying process based in judgements and inference about action-outcome probabilities. Leaning heavily into behavioral persistence as a marker of habits has led to a problem for the field in interpreting the function of an area like the DLS that seem to promote it; it is intuitive to call the DLS a stimulus-response learning machine, but it may not be. By using task rule shifts and action strategy analyses, our findings come down in favor of the DLS being a brain area that can (at least sometimes) help animals to prospectively plan their behavior in changing environmental conditions, a role inconsistent with the DLS being strictly a site for stimulus-response or model-free learning (i.e., habit). In the context of prior work, the DLS can certainly encourage behavioral persistence when one looks at persistence in a certain manner, such as continued action engagement when actions are no longer optimal; but when we look here at action details in a way that can tease apart prospectively oriented vs. retrospectively oriented strategies, there is the potential for the DLS to control behavior in a way that is not via habit and, therefore, a potential way for the DLS to contribute to behaviors that seem persistent via a surprisingly forward-looking mode of function. This idea is an implication rather than a proof of the present results, and much additional work will be needed to flesh it out. However, it indicates that habits can be facilitated by DLS inhibition in some manner.

One intriguing way to reconcile our results with those of prior studies on the DLS is that the DLS participates in creating action learning clusters or states that can be segmented apart, and which can be selected by the DLS as best befits the environment. This notion builds from a recent effort[4] using latent cause inference modeling frameworks[43] to understand habits as resulting from animals creating a latent cluster of initial learning, then creating a different cluster for new learning when circumstances change such as from reward devaluation as a second cluster. Animals then choose to behave based on which cluster befits the testing conditions. In that case, when animals encounter reward devaluation in different non-task contexts, they choose to continue to behave as though reward were not devalued during a subsequent test in the task context (i.e., by habit) because the test context was most similar to the prior learning context when reward had been valued. The results of our study support the possibility of latent learning clusters being relevant to understanding how instrumental behaviors are learned and selected, and in understanding the functional properties of habit-related brain areas like the DLS.

In our case, the task changes occurred in the same environment that animals acquired FR1 learning in. As a result, the latent inference framework would require animals to segment the initial FR1 learning apart from the new learning (e.g., FR3), and to select behaviors that become adapted to the new learning conditions (e.g., FR3 task rules) because that is the actively experienced task circumstance. This is analogous in Pavlovian conditioning to how animals learn extinction rules when they are imposed on them, but do not forget the initial learned behavior. Here, specifically, the behavior of control animals suggests that they learn an FR1-like strategy at first, but they then adjust this strategy when task circumstances change to FR3, back to FR1, and then to RT30. By observing that DLS inhibition caused animals to revert to their initially learned FR1-like behavior suggests it was never forgotten, but rather remained latent as a behavioral option. Moreover, as animals with DLS inhibition were less able to set aside that FR1 learning when the task shifts occurred, it suggests that the DLS normally helps animals set aside prior learning to prospectively plan and engage with new conditions as they arise. Speculatively, because DLS inhibition at the FR3→FR1 downshift caused animals to favor FR1-like behavior rather than their most recent FR3-like behavior, it might suggest that an effect of DLS inhibition is to collapse latent states into the initially learned or most strongly learned one (FR1) rather than the most recent one. Another habit promoting area, the infralimbic cortex in rat, seems to instead carry the function of promoting the most recently acquired strategy of behavior, such that inhibition of it uncovers the most recent strategy.[44]

The notion that the DLS helps cluster out learning contexts could be relevant to understanding many studies showing the consequences of DLS inactivation. For example, without a DLS, animals become poor at integrating new task information to change behavior by exhibiting impairments in action extinction learning,[27] in action reversal learning,[45] and in tracking changing cue locations to find goals.[46] In these cases, DLS inhibition makes animals appear stuck in their behavior, or inflexible. In other cases, animals appear less stuck and more flexible, such as when DLS inhibition helps animals integrate new changes in outcome value or some action-outcome relationship changes.[23,24,47] Just as our results here show that animals with DLS inhibition can look more or less stuck in behavior depending on what we are asking them to do as experimenters, so too one might view the broader literature on DLS function. However, the literature findings might be more cohesively understandable if animals lacking DLS function were struggling to contextualize old information from new information to select their behavior. Sometimes this might result in a failure of new information to impact learned behavior as though it were ignored (e.g., for extinction or reversal learning impairment), while other times it might result in all information being collapsed together (e.g., for increased outcome devaluation sensitivity).

The fact that the FR1 task was nominally easier for animals to perform is of note as FR1-like behavior was always favored by animals during DLS inhibition. The DLS has long been recognized as an area important for actions to be expressed quickly and repeatedly, or in other terms, effortfully and vigorously. The parallel changes in performance duration and performance strategy that we observed with DLS inhibition shows that these factors interact with one another. However, there is reason to suggest that effort alone is not a parsimonious explanation of the data, more specifically that DLS inhibition is not simply causing animals to do what is easiest for them in 1-press bouts. First, the 1-press behavioral pattern of animals during FR3 and RT30 conditions was objectively more effortful than alternatives. Specifically, the three-press bouts during FR3, and a reduction of actions in general during RT30, require less effort than a "press-magazine-press-magazine-press-magazine" routine of animals with DLS inhibition on these days. Moreover, the greater reliance on 1-press bouts occurred mainly during task shifts rather than during the maintenance day when conditions were stable despite equivalent DLS inhibition. Also, DLS inhibition had markedly different effects on measures of performance durations at different task phases. Animals were predictably slower during the FR1-FR3 upshift with an inhibited DLS. Such changes in performance duration can be evident as animals simply perform a learned behavior, but we have found that speed changes resulting from DLS inhibition are far more pronounced when there is a change in task requirements.[17,48] Similarly robust changes in task performance time were not observed during the FR3 maintenance day, in which there was no task shift. Additionally, animals

with DLS inhibition were *faster* during the FR3-FR1 downshift and also during the RT30 shift, which is difficult to explain if DLS inhibition were reducing effort. Also, animals with DLS inhibition pressed many more times than controls during the RT30 days, thus exerting more effort overall. Finally, other types of action bouts (i.e., 2-press bouts and 4+-press bouts) were affected very little by DLS inhibition, whereas the task-relevant bouts of 1- and 3-presses were. If DLS inhibition were simply reducing effort, one might expect to see a reduction in 2-press and 4+-press bouts along with reduced 3-press bouts. Still, one must conclude that effort interacts with action strategy as concerns DLS activity, and studies using task changes with equivalent and reversed effort components (e.g., from FR3 to FR1 to FR3) will be needed to understand if the DLS promotes prospective action planning in a manner that relates to how difficult that action plan is.

It is also worth considering that the 1-press strategy that was favored by animals with DLS inhibition involved an increase in reward magazine checking behaviors between each press on FR3 and RT30 days. Magazine checking could plausibly reflect an active strategy of exploring task conditions (e.g., a study by Bryant et al.[49]). However, if the increased levels of magazine checking during DLS inhibition reflected task and reward exploration, we would have expected it to be related to improved learning (i.e., better adaptation to the task rules) on that day and/or the following day. Instead, the checking behavior seemed to be a component of a rather inflexible FR1-like pattern of behavior in the animals. Compulsive checking behavior is an aspect of obsessive compulsive disorder (OCD), which may have some relevance here as well.

Regarding the curious increase in performance speed during downshift and RT30 days with DLS inhibition, one explanation may lie in the fact that the task demands were increased during the upshift (i.e., more actions required for reward), while they were in principle decreased during the downshift and RT30 days (i.e., fewer actions required for reward). In this way, normally the DLS promotes increased action rates, particularly when action demands increase (upshift), but that it can also encourage decreased action rates when task demands decrease (downshift, RT30). A similar possibility relates to the comparability of animals' behavioral strategy to efficient goal procurement. When goal procurement was less compatible with those (upshift), performance time increased. When goal procurement was more compatible with those (downshift, RT), performance time decreased. These possibilities assume that FR1-like behavior being performed during RT30 was a type of behavior that, while not accurate according to the task rules we set up, could still be considered useful for animals as a means to the reward.

With a few exceptions, changes in performance duration tended to occur when animals were moving to a lever press (i.e., the onset of an action), whether the prior event was another press or a magazine entry. By contrast, less consistent changes in performance durations were seen at the plausible time of action termination (e.g., from a lever press to a magazine entry). This result is consistent with findings showing that DLS activity, and the activity of its dopaminergic input, at the onset of an action sequence is positively correlated with how fast that action sequence is.[16–19,21,36,50] What is inconsistent with the literature is our finding that, during the downshift and RT30 days, DLS inhibition sped up performance. To our knowledge, in prior work, including our own,[17] DLS inhibition always reduces speed. As aforementioned, we speculate the difference could be related to prior work having task conditions remain stable or shift to something new (e.g., a new contingency or a newly devalued reward), much like our upshift day here, while the downshift day and plausibly RT30 days were viewed as a return to the familiar in the animals' repertoire.

Regarding the timing of DLS inhibition, it occurred here as the initial lever press action was taken, not at the moment the decision to press was made (which we found impossible to identify). This timing raises the possibility that animals are pressing more or less to get or to avoid DLS inhibition, as though it carried value itself. However, self-stimulation/avoidance seems unlikely. DLS inhibition lacked strong effects during the FR3 maintenance test days, which would not be the case if the inhibition carried some value. Also, the total number of presses during DLS inhibition was less during the upshift day, while it was more during RT days, which would not be expected if animals were attempting to self-stimulate on both occasions. Further regarding inhibition timing, it could in short trials have continued into the point of reward consumption. A subset of DLS neurons are known to be active during reward delivery,[51,52] which could have been affected. Similarly, there is evidence in slice preparations of rebound excitation at the offset of halorhodopsin-mediated neuronal inhibition. We have not observed rebound excitation in awake, behaving animals using similar inhibition protocols as here,[44,53] but if it were occurring, it would be happening around reward consumption for many of the FR trials. However, there was a lack of any clear effect of DLS inhibition when it was explicitly given after the action and during reward consumption, a similar lack of robust effect when it was given during maintenance sessions, and a lack of any consistent reward-related light delivery during RT30 sessions; these observations suggest that reward-related signaling was not responsible for the full set of DLS inhibition effects we report.

Regarding the use of a within-subject testing design, we do not believe that effects observed on later test days (e.g., downshift) were due to prior test experiences (e.g., upshift). For example, one concern would be if animals in the Halo group would have favored FR1-like responding during the downshift day even if no light were delivered, because they had the earlier upshift experience. However, if so, a favoring of the FR1-like strategy should also have been seen at other times, such as the day preceding the downshift day, which was not the case (Figure 4G), nor was there a similar spike in FR1-like responding on the maintenance day with light delivery. Finally, the upshift and downshift days had opposite effects of DLS inhibition on performance speed, despite having equivalent effects on the use of an FR1 strategy, suggesting that the downshift day was not simply a reproduction of the earlier upshift day.

These results here diverge from what it is seen with inhibition of a small population of DLS cholinergic interneurons, during which animals are better at adapting their behavior to changes in task rules.[38] With neuron-wide DLS inhibition here, animals did seemingly the opposite by favoring the initially learned FR1 style of behavior and appearing as more inflexible. Differences in DLS manipulation methods between these studies aside, it could be that cholinergic neurons and non-cholinergic neurons within the DLS carry opponent functions. These results also may diverge from those of a recent study finding that activity in DLS neurons reflected previously experienced maze running experience better than forthcoming running goals,[54] suggesting they represent retrospectively oriented (i.e., previously learned) information. Here, it would appear conversely that prior experience (i.e., FR1-driven behavior) was dominant when the DLS was actually offline. Reconciling these results will require a better understanding of cell type-specific processes occurring in the DLS.

Within the broader striatum, there is considerable interest in how medial and lateral divisions toggle modes of behavior that are goal-directed (thought to be a medial function) or habitual (thought to be a lateral function). These could be parallel functions, dictated by the ebb and flow of activity in medial vs. lateral domains. Alternatively, medial and lateral striatum could interface with one another, such as through midbrain-striatal spiraling connections that march from medial to lateral or within the striatum through collaterals and interneurons, which would allow for one striatal domain to directly affect activity in the other.[13,55–57] Typically medial and lateral striatum are thought to oppose one another in control over behavior through one of these routes, but there is also recent evidence that the medial striatum can promote habitual behaviors in a manner that had historically been assigned to the lateral striatum (e.g., studies by Malvaez and Wassum,[13] and Ambrosi and Lerner[55]). We add to these new observations our current findings in which the DLS can promote goal-directed behaviors, which has historically been the domain of the medial striatum. Thus, it is plausible that a mechanism by which DLS inhibition increased habits in our data is the DMS habit-promoting function.

This work points to a larger challenge in how to study habits. Typically, habits are inferred from a negative result (i.e., lack of sensitivity to reward devaluation or action-outcome contingency change). Recently, reward devaluation studies have uncovered problems associated with using this method to determine a behavior as a habit, as the context of the devaluation procedure itself carries great weight over observed results.[58,59] Although measures of vigor and skill have proven to be a helpful positive identifier of habits, one can also find non-habitual behaviors that look just as skillful as habits.[1,60,61] Further, while manipulations to reduce DLS function can render animals more flexible in some situations, it can also render them more inflexible in other situations as noted previously and observed in the present study. We are inclined to suggest that measures of how flexible vs. inflexible animals are in response to environmental change may have limited utility in understating DLS roles in habits, at least if viewed as a binary condition. There are simply too many reasons why behavioral inflexibility might arise. There is a clear need for positive behavioral identifiers of habits, which may be aided by exploring new task designs and by looking in fine detail at how animals behave in them.[5,9,11,13,34,62,63] A focus on action strategies may be of benefit as well, as we attempt here; animals that persist in a behavior by habit ought to persist in such a way that is indistinguishable from what they learned to do, beyond just whether they persist or not. In broader terms, here, the DLS was a top-down target for understanding notions of behavioral persistence. One can also take a bottom-up approach to discover how exactly the DLS controls behavioral details and apply that knowledge upwards to understand how those behavioral details can explain the role of the DLS in measures of behavioral persistence. This includes studies showing DLS control over speed and effort domains of behavior with explanatory power for understanding how these precise behavioral functions might lead to habit-like performance.[21,22,66,67] Linking these top-down and bottom-up approaches could be of considerable scientific benefit to resolving the psychological function of areas like the DLS.

### Limitations of the study

The current study requires extension in several domains, which limit the extent to which conclusions can be drawn about how substantial the role of the DLS is in prospectively guiding flexible behavior. Of the key factors, the first is to establish how generalizable the findings are to other conditions; work must be done to show similar effects of DLS inhibition on tasks that shift from hard to easy, or that involve task shifts of equal effort. Generalization must also be assessed for the female sex, and for other types of rewards and actions. Second, given the ever-expanding appreciation of cell-type-specific functions in the striatum, DLS manipulations that target particular cells (rather than pan-neuronal here) will be invaluable. Third, while the optogenetic approach we used limited manipulations to task performance time, it will be important to different performance moments of task initiation vs. execution vs. completion vs. reward consumption in a thorough manner. Finally, as discussed, there is highly likely to be interaction, whether direct or indirect, with DMS-related circuitry; given some evidence that the DMS can encourage habits, it is a speculative but compelling notion that DLS inhibition influence or unmask the DMS which in part carries out the behaviors.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Surgical methods
  - Apparatus
  - Optogenetic experimental methods
  - Experimental and behavioral design
  - Histological methods
- QUANTIFICATION AND STATISTICAL ANALYSIS

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2024.110044.

## ACKNOWLEDGMENTS

We thank Dr. Neil Winterbauer, Dr. Stephen Chang, Dr. Alyssa DiLeo, Alex Brown, and Matthew Betz for assistance. This work was supported by an NSF research grant to K.S.S. (IOS 1557987), an NIH grant to K.S.S. (R01DA04419), and an NIH grant to K.A.A. (F99NS115270).

## AUTHOR CONTRIBUTIONS

K.S.S. and A.C.G.C.: designed research; A.C.G.C.: performed research; K.S.S., A.C.G.C., J.A.P., and K.A.A.: analyzed data and wrote the paper.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 14, 2023
Revised: March 20, 2024
Accepted: May 17, 2024
Published: May 21, 2024

## REFERENCES

1. Balleine, B.W., Liljeholm, M., and Ostlund, S.B. (2009). The integrative function of the basal ganglia in instrumental conditioning. Behav. Brain Res. *199*, 43–52.
2. Dezfouli, A., and Balleine, B.W. (2012). Habits, action sequences and reinforcement learning. Eur. J. Neurosci. *35*, 1036–1051.
3. Tolman, E.C. (1932). Purposive Behavior in Animals and Men (Century/Random House UK).
4. Garrett, N., Allan, S., and Daw, N.D. (2023). Model based control can give rise to devaluation insensitive choice. Add. Neurosci. *6*, 100070.
5. Amaya, K.A., and Smith, K.S. (2018). Neurobiology of habit formation. Curr. Opin. Behav. Sci. *20*, 145–152.
6. Barker, J.M., Taylor, J.R., Barker, J.M., and Taylor, J.R. (2014). Habitual alcohol seeking: modeling the transition from casual drinking to addiction. Neurosci. Biobehav. Rev. *47*, 281–294.
7. Corbit, L.H., and Janak, P.H. (2016). Habitual alcohol seeking: neural bases and possible relations to alcohol use disorders. Alcohol Clin. Exp. Res. *40*, 1380–1389.
8. Gasbarri, A., Pompili, A., Packard, M.G., and Tomaz, C. (2014). Habit learning and memory in mammals: Behavioral and neural characteristics. Neurobiol. Learn. Mem. *114*, 198–208.
9. Graybiel, A.M. (2008). Habits, rituals, and the evaluative brain. Annu. Rev. Neurosci. *31*, 359–387.
10. Knowlton, B.J., and Patterson, T.K. (2018). Habit formation and the striatum. Curr. Top. Behav. Neurosci. *37*, 275–295.
11. Lerner, T.N. (2020). Interfacing behavioral and neural circuit models for habit formation. J. Neurosci. Res. *98*, 1031–1045.
12. Lovinger, D.M., and Gremel, C.M. (2021). A circuit-based information approach to substance abuse research. Trends Neurosci. *44*, 122–135.
13. Malvaez, M., and Wassum, K.M. (2018). Regulation of habit formation in the dorsal striatum. Curr. Opin. Behav. Sci. *20*, 67–74.
14. Seger, C.A. (2018). Corticostriatal foundations of habits. Curr. Opin. Behav. Sci. *20*, 153–160.
15. Yin, H.H., and Knowlton, B.J. (2004). Contributions of striatal subregions to place and response learning. Learn. Mem. *11*, 459–463.
16. Barnes, T.D., Kubota, Y., Hu, D., Jin, D.Z., and Graybiel, A.M. (2005). Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. Nature *437*, 1158–1161.
17. Crego, A.C.G., Štoček, F., Marchuk, A.G., Carmichael, J.E., van der Meer, M.A.A., and Smith, K.S. (2020). Complementary control over habits and behavioral vigor by phasic activity in the dorsolateral striatum. J. Neurosci. *40*, 2139–2153.
18. Jin, X., and Costa, R.M. (2010). Start/stop signals emerge in nigrostriatal circuits during sequence learning. Nature *466*, 457–462.
19. Jog, M.S., Kubota, Y., Connolly, C.I., Hillegaart, V., and Graybiel, A.M. (1999). Building neural representations of habits. Science *286*, 1745–1749.
20. Regier, P.S., Amemiya, S., and Redish, A.D. (2015). Hippocampus and subregions of the dorsal striatum respond differently to a behavioral strategy change on a spatial navigation task. J. Neurophysiol. *114*, 1399–1416.
21. Smith, K.S., and Graybiel, A.M. (2013). A Dual Operator View of Habitual Behavior Reflecting Cortical and Striatal Dynamics. Neuron *79*, 608.
22. Yttri, E.A., and Dudman, J.T. (2016). Opponent and bidirectional control of movement velocity in the basal ganglia. Nature *533*, 402–406.
23. Yin, H.H., Knowlton, B.J., and Balleine, B.W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. Eur. J. Neurosci. *19*, 181–189.
24. Yin, H.H., Knowlton, B.J., and Balleine, B.W. (2006). Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. Behav. Brain Res. *166*, 189–196.
25. Bailey, K.R., and Mair, R.G. (2006). The role of striatum in initiation and execution of learned action sequences in rats. J. Neurosci. *26*, 1016–1025.
26. Bergstrom, H.C., Lipkin, A.M., Lieberman, A.G., Pinard, C.R., Gunduz-Cinar, O., Brockway, E.T., Taylor, W.W., Nonaka, M., Bukalo, O., Wills, T.A., et al. (2018). Dorsolateral striatum engagement interferes with early discrimination learning. Cell Rep. *23*, 2264–2272.
27. Goodman, J., Ressler, R.L., and Packard, M.G. (2017). Enhancing and impairing extinction of habit memory through modulation of NMDA receptors in the dorsolateral striatum. Neuroscience *352*, 216–225.
28. Packard, M.G., and McGaugh, J.L. (1992). Double dissociation of fornix and caudate nucleus lesions on acquisition of two water maze tasks: Further evidence for multiple memory systems. Behav. Neurosci. *106*, 439–446.
29. Turner, K.M., Svegborn, A., Langguth, M., McKenzie, C., and Robbins, T.W. (2022). Opposing roles of the dorsolateral and dorsomedial striatum in the acquisition of skilled action sequencing in rats. J. Neurosci. *42*, 2039–2051.
30. Yin, H.H., Mulcare, S.P., Hilário, M.R.F., Clouse, E., Holloway, T., Davis, M.I., Hansson, A.C., Lovinger, D.M., Costa, R.M., and Costa, R.M. (2009). Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. Nat. Neurosci. *12*, 333–341.
31. Yu, C., Gupta, J., Chen, J.F., and Yin, H.H. (2009). Genetic deletion of A2A adenosine receptors in the striatum selectively impairs habit formation. J. Neurosci. *29*, 15100–15103.
32. Dickinson, A., Balleine, B., Watt, A., Gonzalez, F., and Boakes, R.A. (1995). Motivational control after extended instrumental training. Anim. Learn. Behav. *23*, 197–206.
33. Seiler, J.L., Cosme, C.V., Sherathiya, V.N., Schaid, M.D., Bianco, J.M., Bridgemohan,

A.S., and Lerner, T.N. (2022). Dopamine signaling in the dorsomedial striatum promotes compulsive behavior. Curr. Biol. *32*, 1175–1188.e5.
34. van Elzelingen, W., Warnaar, P., Matos, J., Bastet, W., Jonkman, R., Smulders, D., Goedhoop, J., Denys, D., Arbab, T., Willuhn, I., and Willuhn, I. (2022). Striatal dopamine signals are region specific and temporally stable across action-sequence habit formation. Curr. Biol. *32*, 1163–1174.e6.
35. Cromwell, H.C., and Berridge, K.C. (1996). Implementation of action sequences by a neostriatal site: a lesion mapping study of grooming syntax. J. Neurosci. *16*, 3444–3458.
36. da Silva, J.A., Tecuapetla, F., Paixão, V., and Costa, R.M. (2018). Dopamine neuron activity before action initiation gates and invigorates future movements. Nature *554*, 244–248.
37. Panigrahi, B., Martin, K.A., Li, Y., Graves, A.R., Vollmer, A., Olson, L., Mensh, B.D., Karpova, A.Y., and Dudman, J. (2015). Dopamine Is Required for the Neural Representation and Control of Movement Vigor. Cell *162*, 1418–1430.
38. Amaya, K.A., and Smith, K.S. (2021). Spatially restricted inhibition of cholinergic interneurons in the dorsolateral striatum encourages behavioral exploration. Eur. J. Neurosci. *53*, 2567–2579.
39. Balleine, B., and Killcross, S. (1994). Effects of ibotenic acid lesions of the nucleus accumbens on instrumental action. Behav. Brain Res. *65*, 181–193.
40. Balleine, B.W., Killcross, A.S., and Dickinson, A. (2003). The effect of lesions of the basolateral amygdala on instrumental conditioning. J. Neurosci. *23*, 666–675.
41. Amaya, K.A., and Smith, K.S. (2024). Alternative approaches to understanding habit learning in the dorsolateral striatum. In Habits - Their Definition, Neurobiology and Role in Addiction, Y. Vandaele, ed. (Springer Nature Press).
42. Berridge, K.C. (2021). Comment on Vandaele and Ahmed: Rethinking habits in addiction. Neuropsychopharmacology *46*, 687–688.
43. Gershman, S.J., Norman, K.A., and Niv, Y. (2015). Discovering latent causes in reinforcement learning. Curr. Opin. Behav. Sci. *5*, 43–50.
44. Smith, K.S., Virkud, A., Deisseroth, K., and Graybiel, A.M. (2012). Reversible online control of habitual behavior by optogenetic perturbation of medial prefrontal cortex. Proc. Natl. Acad. Sci. USA *109*, 18932–18937.
45. Jackson, S.A.W., Horst, N.K., Axelsson, S.F.A., Horiguchi, N., Cockcroft, G.J., Robbins, T.W., and Roberts, A.C. (2019). Selective role of the putamen in serial reversal learning in the marmoset. Cerebr. Cortex *29*, 447–460.
46. Kosaki, Y., Poulter, S.L., Austen, J.M., and McGregor, A. (2015). Dorsolateral striatal lesions impair navigation based on landmark-goal vectors but facilitate spatial learning based on a "cognitive map. Learn. Mem. *22*, 179–191.
47. Jonkman, S., Pelloux, Y., and Everitt, B.J. (2012). Differential roles of the dorsolateral and midlateral striatum in punished cocaine seeking. J. Neurosci. *32*, 4645–4650.
48. Crego, A.C.G., Chang, S.E., Butler, W.E., and Smith, K.S. (2016). Optogenetic research in behavioral neuroscience: insights into the brain basis of reward learning and goal-directed behavior. In Optogenetics: From Neuronal Function to Mapping & Disease Biology (Cambridge University Press).
49. Bryant, K.G., Singh, B., and Barker, J.M. (2022). Reinforcement history dependent effects of low dose ethanol on reward motivation in male and female mice. Front. Behav. Neurosci. *16*, 875890.
50. Thorn, C.A., Atallah, H., Howe, M., and Graybiel, A.M. (2010). Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. Neuron *66*, 781–795.
51. Schmitzer-Torbert, N., and Redish, A.D. (2004). Neuronal activity in the rodent dorsal striatum in sequential navigation: separation of spatial and reward responses on the multiple T task. J. Neurophysiol. *91*, 2259–2272.
52. Smith, K.S., and Graybiel, A.M. (2016). Habit formation coincides with shifts in reinforcement representations in the sensorimotor striatum. J. Neurophysiol. *115*, 1487–1498.
53. Chang, S.E., Smedley, E.B., Stansfield, K.J., Stott, J.J., and Smith, K.S. (2017). Optogenetic inhibition of ventral pallidum neurons impairs context-driven salt-seeking. J. Neurosci. *37*, 5670–5680.
54. Cunningham, P.J., Regier, P.S., and Redish, A.D. (2021). Dorsolateral striatal task initiation bursts represent past experiences more than future action plans. J. Neurosci. *41*, 8051–8064.
55. Ambrosi, P., and Lerner, T.N. (2022). Striatonigrostriatal circuit architecture for disinhibition of dopamine signaling. Cell Rep. *40*, 111228.
56. Haber, S.N., Fudge, J.L., and McFarland, N.R. (2000). Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. J. Neurosci. *20*, 2369–2382.
57. Ikeda, H., Saigusa, T., Kamei, J., Koshikawa, N., and Cools, A.R. (2013). Spiraling dopaminergic circuitry from the ventral striatum to dorsal striatum is an effective feed-forward loop. Neuroscience *241*, 126–134.
58. Amaya, K.A., Stott, J.J., and Smith, K.S. (2020). Sign-tracking behavior is sensitive to outcome devaluation in a devaluation context-dependent manner: implications for analyzing habitual behavior. Learn. Mem. *27*, 136–149.
59. Bouton, M.E., Allan, S.M., Tavakkoli, A., Steinfeld, M.R., and Thrailkill, E.A. (2021). Effect of context on the instrumental reinforcer devaluation effect produced by taste-aversion learning. J. Exp. Psychol. Anim. Learn. Cogn. *47*, 476–489.
60. Vandaele, Y., and Janak, P.H. (2023). Lack of action monitoring as a prerequisite for habitual and chunked behavior: Behavioral and neural correlates. iScience *26*, 105818.
61. Garr, E., and Delamater, A.R. (2019). Exploring the relationship between actions, habits, and automaticity in an action sequence task. Learn. Mem. *26*, 128–132.
62. Schreiner, D.C., Renteria, R., and Gremel, C.M. (2020). Fractionating the all-or-nothing definition of goal-directed and habitual decision-making. J. Neurosci. Res. *98*, 998–1006.
63. Watson, P., O'Callaghan, C., Perkes, I., Bradfield, L., and Turner, K. (2022). Making habits measurable beyond what they are not: a focus on associative dual-process models. Neurosci. Biobehav. Rev. *142*, 104869.
64. Kuznetsova, A., Brockhoff, P.B., and Christensen, R.H.B. (2017). lmerTest package: tests in linear mixed effects models. J. Stat. Software *82*, 1–26.
65. Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. J. Stat. Software *67*, 1–48.
66. Rueda-Orozco, P.E., and Robbe, D. (2015). The striatum multiplexes contextual and kinematic information to constrain motor habits execution. Nat. Neurosci. *18*, 453–460.
67. Jurado-Parras, M.T., Safaie, M., Sarno, S., Louis, J., Karoutchi, C., Berret, B., and Robbe, D. (2020). The dorsal striatum energizes motor routines. Curr. Biol. *30*, 4362–4372.

## STAR★METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---|---|---|
| Bacterial and virus strains | | |
| AAV5-hSYN-eNpHR3.0-EYFP | UNC Gene Therapy Center Vector Core | Halorhodopsin |
| AAV5-hSYN-EYFP | UNC Gene Therapy Center Vector Core | Control |

### RESOURCE AVAILABILITY

#### Lead contact

Further information and requests should be directed to the lead contact: Dr. Kyle S. Smith (kyle.s.smith@dartmouth.edu).

#### Materials availability

This study did not generate new unique reagents.

#### Data and code availability

- Data reported in this paper may be shared by the lead contact upon request.
- This paper does not report original code.
- Any additional information required to analyze the data reported in this paper is available from the lead contact upon request.

### EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

All procedures were approved by the Dartmouth College Institutional Animal Care and Use Committee. Male Long Evans rats ($n$ = 18) were housed individually in a dedicated animal vivarium, with daily health checks from the veterinary staff. All rats used were at 250–400 g and were maintained on an 85% post-surgical weight for training and testing. Rats were housed on a reverse light-dark cycle. All male Long Evans rats used in this study were purchased from Charles River Laboratories.

### METHOD DETAILS

#### Surgical methods

Surgeries were performed using aseptic techniques under isoflurane anesthesia: intracranial injection of viral vectors, followed by intracranial implantation of fiber optic guides. Rats received bilateral injections (0.3 μL) into DLS of a single viral construct: eNpHR3.0/halorhodopsin ($n$ = 9, AAV5-hSYN-eNpHR3.0-EYFP) or control ($n$ = 9, AAV5-hSYN-EYFP) using a 33-gauge syringe and microinfusion pump. Bilateral DLS coordinates for viral injections were, in mm: AP +0.5, ML −/+ 4.0, DV injection −4.3 from skull, with fiber implants (200 μm, ThorLabs or in-house) terminating at DV -3.8 mm. Fiber implants were permanently affixed to the skull using dental cement and skull screws. Rats were given at least 2 weeks from surgery to the start of experimentation, which began with 7 initial training days. Thus, at least 3 weeks elapsed between surgery and the first DLS inhibition day.

#### Apparatus

Instrumental conditioning was carried out in eight identical modular operant chambers (Med Associates) in a dedicated experimental room. All chambers were housed in their own sound attenuated box and closed during all training sessions. Chambers were equipped with two identical levers that were retractable. Both levers were inserted upon the start of a session and remained inserted until the completion of the session. For each animal, one lever was active (paired with reward) and the other inactive. In between the levers was a magazine food cup receptacle, where sucrose reward pellets would drop upon the specified Med PC macro script program. At the top back region of the chambers, a house light would illuminate to indicate a session had begun and remain on for the duration of the experiment, but the remainder of the testing room was in the dark. A photobeam within the magazine acted as a reset when quantifying specific measures of lever pressing (e.g., inter-trial-intervals, lever press bouts, etc.). Three other photobeams were spaced across the chamber floor. All sessions began with house light illumination and the insertion of two levers.

#### Optogenetic experimental methods

Optical patch cords with a 0.2 μm fiber core and 0.39 NA were used for light administration (Thorlabs, Newton, NJ; Doric Lenses, Quebec City, Canada). On days where DLS was manipulated, a 593.5 nm light from DPSS lasers (Shanghai Laser & Optics Century Co., Ltd.) was delivered through these fiber patch cords that were connected to a rotary-joint beam splitter (Thorlabs, Newton, NJ; Doric Lenses, Quebec City,

Canada). This allowed for two fiber patch cables to be connected directly to the surgically implanted optic fiber cannula (in-house fiber implants, 200 μm, ThorLabs) by ceramic sleeves. Laser illumination delivery was gated and triggered by TTL pulses through designed Med PC programs, to individual operant chamber boxes. Power output was measured as 3–5 mW directly from each patch cable ferrule connector prior to and after laser test days. Fiber implants were also tested post-perfusion to confirm their patency.

Rats were habituated to the optogenetic cables prior to any experimental days. On separate manipulation days, laser light was delivered into the DLS as a continuous 1-s pulse triggered by the inflection of a pre-defined paired reward lever press. The 1-s illumination duration was chosen to allow animals to behave as they would during that 1-s period, and 1-s captured well a typical duration between lever deflection and magazine entry. Any unpaired lever pressing did not result in laser illumination delivery. Operant behavior and task events (e.g., lever presses, time, etc.) were recorded by Med PC software via a designated computer. Accuracy of automated behavioral measurements was verified through videotaping and hand scoring of a subset of sessions.

## Experimental and behavioral design

Rats were first trained on a fixed-ratio 1 (FR1) schedule of reward, where a predefined active lever (i.e., consistent throughout the duration of the experiment) was paired with reward delivery. An inactive lever was paired with nothing. Lever assignment was counterbalanced across groups. FR1 training sessions were administered for 7 sequential days, wherein rats learned a contingency that a single lever press resulted in the delivery of a single reward pellet within the magazine port. This learned behavior was then challenged by a change in the action requirement for the reward, which was a surprise upshift in requirement from one to three presses for reward (FR1 → FR3; "upshift day"). This sudden shift now required rats to press the previously assigned paired-rewarded lever three times. The DLS was optogenetically inhibited (eNpHR3.0) on only the first press of the active lever during this upshift day.

A series of 7 FR3 training days followed without further DLS inhibition to evaluate how animals adjusted to the new task rules. Then, the DLS was inhibited during an FR3 "maintenance day". This day was conducted identically to the upshift day, and it was followed by a single non-illumination FR3 day for comparison. Next, illumination was given on the last press leading to reward procurement during FR3, followed by a non-illumination FR3 day.

Following this, the task was then suddenly returned back to the originally learned FR1 schedule (FR3 → FR1; "downshift day"). DLS illumination was time-locked to the first press after a magazine entry (i.e., subsequent pressing would not yield illumination until a magazine entry occurred again). This downshift day allowed us to understand the effects of DLS inhibition when FR1-based performance was the learned rule to adapt against (upshift day) versus when it was a new rule to adapt toward (downshift day).

A series of 7 FR1 training days followed without DLS inhibition to again establish performance on that FR1 schedule. Then, and finally, rats were subjected to a surprise introduction of a random-time 30 (RT30) schedule in which reward was delivered entirely independently from any pressing. This shift to RT30 was designed to entirely remove the action requirement for reward, allowing us to examine how DLS inhibition might affect the frequency, distribution, and pattern of learned actions when they were no longer necessary. This procedure is similar to one we recently published to probe behavioral exploration vs. behavioral fixity in understanding DLS function[38] (see also: Balleine and Killcross[39]; Balleine et al.[40]). Illumination was applied during this RT30 period, and it was paired with every lever press (i.e., previously paired-rewarded lever) regardless of magazine entries.

All training and experimental sessions would end after 60 min or once 30 rewards were earned, with the exception of an RT30 (random time) day always ending in 30 min. We opted for this within-subject design for several reasons. We found it desirable to have multiple test days per animal to aid comparison of trends across test days. Additionally, in the spirit of the Reduce/Reuse/Replace NIH guidelines, we found this design to be a desirable solution for having a robust enough dataset without requiring a very large number of subjects to run each test day independently. It was also important to run cohorts of animals such that each cohort spans the various experimental conditions; we were able to do this having the two Control/Halo conditions using the within-subject approach.

We deliberately designed the study so that the initial 1-press strategy could still be used with success during FR3 or RT30; it just wasn't optimal anymore. This creates a favorable situation because we sought to test the hypothesis that the DLS helps select and regulate self-generated action strategies. Strategies are thus relatively self-chosen beyond the constraint of needing to press the lever, rather than being imposed directly by experimenters. It can be instead common to incorporate a new action during task shifts. However, doing this would bias animals toward faster adoption of the optimal strategy because it would be the only workable one, thus pushing animals toward prospective action planning rather quickly as well as introducing learning variables related to reward omissions when old behaviors are repeated. In other words, we created a condition in which animals would change their strategy not because they were losing rewards but rather because they were attending to the rules and adapting.

## Histological methods

At the termination of each cohort, lethal doses of anesthesia (sodium pentobarbital) were administered. This was followed by a transcardial perfusion of 0.9% saline and 4% paraformaldehyde. Brains were fixed further in 20% sucrose solution overnight and frozen at −80°C. Brains were sectioned at 30–60 μm using a microtome, and later mounted to slides and cover slipped with a DAPI-containing medium. Fiber placement and EYFP expressing neuron spread were analyzed using a fluorescent microscope (Olympus BX53 fluorescent microscope with DP73 camera).

## QUANTIFICATION AND STATISTICAL ANALYSIS

Lever deflections and pellet magazine entries were recorded through Med PC. Operant behavioral measures were individually extracted from Med PC files using a custom Python script. Statistical testing was conducted using R (R Core Team, 2016). All graphs were made using R (R; "ggplot2") or SPSS, and stylized in Adobe Illustrator.

Individual generalized linear mixed models (R; "lme4") were created for each dependent behavioral variable (Gamma distribution for temporal measures: overall time (min), inter-press intervals (sec), reward contact to lever press (sec), lever press to mag entry (sec); Binomial distribution for action measures averaged: 1-press bouts, and 3-press bouts) for within- and between-subject analyses for main effects day (illumination day and subsequent non-laser day), virus group (control vs. eNpHR3.0), and the interaction between these variables. By comparing across 2 test days (light delivery day and the following day) and 2 groups (halorhodopsin and control), an interaction indicates a change in group effects as a function of test day. These comparisons were conducted separately based on *a priori* events of interest for DLS function on separate illumination days (e.g., upshift, maintenance, downshift, RT30). Reported statistics include parameter estimates, 95% confidence intervals, and *p*-values (R; "lmerTest",[64]). Separately, a linear mixed model was created to analyze the effects of day, group, and their interaction on presses made per reward. Linear mixed models were fit by maximum likelihood and t-tests used Satterthwaite approximations of degrees of freedom (R; "lmerMod"). Linear models were analyzed with package lme4 from CRAN.[65] R code was: $Y \sim X1 * X2 + (1 | Rat)$. In cases where two models didn't fail to converge (6A and 6B), it was: $Y \sim X1 * X2 + (1 + X2 | Rat)$. Reported statistics include parameter estimates (β values), 95% confidence intervals, and *p*-values. Also, per-animal performance ratios were calculated of 1-press bouts to 3-press bouts and analyzed for group effects on the illumination days by one-way ANOVA or, when data were non-normally distributed, Wilcoxon rank sign test. Finally, correlation analyses were performed on 1-press vs. 3-press bouts to gain a sense of how much they related to one another in animals' performance patterns.

For behavior during the RT30 phase of the experiment, measures of action probability just prior to and just following reward delivery were also analyzed. Time windows for these analyses were individually set as the mean inter-press interval, rounded up. For example, if an animal averaged 2.3s between presses, the pre- and post-reward delivery windows were set as 3s before and 3s after reward delivery. Models of action probability were generalized linear mixed models with fixed effects of group, session, and their interaction and random intercepts for individual start points on session 1. Reported statistics include odds ratios, 95% confidence intervals, and *p*-values. Lever press rates (as presses per minute) as predicted by group assignment and session were assessed both overall and within a peri-reward time window (+/− 1 s). Linear mixed models with fixed effects of group, session, and their interaction and random intercepts for individual rats were used here to predict respective dependent variables, either overall presses per minute or presses per minute within the specified time window. Lever press rates and action probabilities were then assessed with respect to laser delivery, with lever press rates outside of laser delivery being considered distinct from lever press rates during laser delivery. Action probability was calculated as the probability of a lever press occurring during laser administration as a simple binary (yes/no) as each lever press triggered laser delivery during this phase of the experiment. A linear mixed model with fixed effects of laser bin, group, session, and interaction terms with random intercepts for individual rats was used to predict presses per minute made. Importantly, Session was re-centered such that the final RT-30 session is the comparison point for the main effects of group and laser bin. To model action probability during laser delivery, a generalized linear mixed model with fixed effects of group, session, and their interaction with random intercepts for individual rat starting points was used to predict probability of action. Reported statistics include parameter estimates (or odds ratios for GLMMs), 95% confidence intervals, and *p*-values. Significance symbols in graphs are *($p < 0.05$ Day/Group interaction); +($p < 0.05$ Day effect); #($p < 0.05$ Group effect).