

Complementary Control over Habits and Behavioral Vigor by Phasic Activity in the Dorsolateral Striatum

Adam C. G. Crego, Fabián Štoček, Alec G. Marchuk,  James E. Carmichael,  Matthijs A. A. van der Meer, and Kyle S. Smith

Department of Psychological and Brain Sciences, Dartmouth College, Hanover, New Hampshire 03755

Despite clear evidence linking the basal ganglia to the control of outcome insensitivity (i.e., habit) and behavioral vigor (i.e., its behavioral speed/fluidity), it remains unclear whether or how these functions relate to one another. Here, using male Long–Evans rats in response-based and cue-based maze-running tasks, we demonstrate that phasic dorsolateral striatum (DLS) activity occurring at the onset of a learned behavior regulates how vigorous and habitual it is. In a response-based task, brief optogenetic excitation at the onset of runs decreased run duration and the occurrence of deliberative behaviors, whereas midrun stimulation carried little effect. Outcome devaluation showed these runs to be habitual. DLS inhibition at run start did not produce robust effects on behavior until after outcome devaluation. At that time, when the DLS was plausibly most critically required for performance (i.e., habitual), inhibition reduced performance vigor measures and caused a dramatic loss of habitual responding (i.e., animals quit the task). In a second cue-based “beacon” task requiring behavior initiation at the start of the run and again in the middle of the run, DLS excitation at both time points could improve the vigor of runs. Postdevaluation testing showed behavior on the beacon task to be habitual as well. This pattern of results suggests that one role for phasic DLS activity at behavior initiation is to promote the execution of the behavior in a vigorous and habitual fashion by a diverse set of measures.

Key words: habit; vigor; basal ganglia; optogenetics; dorsolateral striatum; plus-maze; goal-directed action; outcome insensitivity; deliberation (VTE)

Significance Statement

Our research expands the literature twofold. First, we find that features of a habitual behavior that are typically studied separately (i.e., maze response performance, deliberation movements, running vigor, and outcome insensitivity) are quite closely linked together. Second, efforts have been made to understand “what” the dorsolateral striatum (DLS) does for habitual behavior, and our research provides a key set of results showing “when” it is important (i.e., at behavior initiation). By showing such dramatic control over habits by DLS activity in a phasic time window, plausible real-world applications could involve more informed DLS perturbations to curb intractably problematic habits.

Introduction

Habits are behavioral routines characterized by their automaticity, consistency, and resistance to change. A key node within the habit-promoting network in the brain is the dorsolateral striatum

(DLS, primate putamen homolog; Yin et al., 2004, 2006; Seger and Spiering, 2011; Smith and Graybiel, 2016; Amaya and Smith, 2018). One line of research has shown that the DLS promotes operant habits that are insensitive to changes in the relationship between a learned behavior and its outcome (Dickinson, 1985; Balleine and Dickinson, 1998; Yin and Knowlton, 2006; Balleine et al., 2009). A second, largely separate line of research has shown that the DLS is important for using response-based (egocentric) running strategies on maze tasks (Packard, 2009), which can also be outcome-insensitive in some, but not all, conditions (De Leonibus et al., 2011; Kosaki et al., 2018). A third parallel line of research has implicated the basal ganglia in promoting movement sequence structure and vigor (Aldridge et al., 2004; Haith et al., 2012; Dudman and Krakauer, 2016; da Silva et al., 2018), with the conclusion that the DLS controls rapid and fluid movement plans (Novak et al.,

Received June 6, 2019; revised Dec. 18, 2019; accepted Dec. 19, 2019.

Author contributions: A.C.G.C. and K.S.S. designed research; A.C.C., F.S., A.G.M., and J.E.C. performed research; K.S.S. contributed unpublished reagents/analytic tools; A.C.G.C. and K.S.S. analyzed data; A.C.G.C., F.S., A.G.M., J.E.C., M.A.A.v.d.M., and K.S.S. wrote the paper.

This work was supported by a National Science Foundation research grant (IOS 1557987) and a National Institutes of Health grant (1R01DA044199) to K.S.S. We thank Kenneth Amaya, Jacqueline Perron-Smith, Elizabeth McNally, Alyssa DiLeo, Alex Brown, Elizabeth Smedley, and Dr. Stephen Chang for assistance.

The authors declare no competing financial interests.

Correspondence should be addressed to Kyle Smith at kyle.s.smith@dartmouth.edu or Adam Crego at Adam.C.Crego.GR@dartmouth.edu.

<https://doi.org/10.1523/JNEUROSCI.1313-19.2019>

Copyright © 2020 the authors

2002; Dudman and Krakauer, 2016). In the rodent, measures of vigor can include fast movement speed; low performance duration; and, in maze running, lack of deliberation-like head movements [vicarious trial and error (VTE)] at choice points.

Given the largely nonoverlapping nature of these research lines, it has become unclear how measures of performance vigor, habit as defined by outcome insensitivity, and egocentric navigation relate to one another. One entry point to studying this issue, used here, is a prominent activity signal in the DLS that occurs during the transition of action to habit. Specifically, there is often a burst of DLS activity (200–500 ms) at the initiation of a well learned movement sequence (Jog et al., 1999; Kubota et al., 2009; Jin and Costa, 2010; Thorn et al., 2010; Smith and Graybiel, 2013a). This signaling event has been found in a range of species and DLS neuronal subtypes (Jog et al., 1999; Kubota et al., 2009; Vicente et al., 2016) and can also occur in wider brain networks (Fujii and Graybiel, 2003; Jin and Costa, 2010, 2015; Fujimoto et al., 2011; Smith and Graybiel, 2013a; Desrochers et al., 2015). Notably, increases in the magnitude of movement-start-related activity in the DLS and its substantia nigra pars compacta input are closely related to behavioral changes reflective of increased vigor (Jin and Costa, 2010; Smith and Graybiel, 2013a; da Silva et al., 2018). Within the DLS in particular, the burst of DLS activity at the onset of a behavior is predictive at the single-trial level of reduced performance duration and reduced deliberations at decision points (Smith and Graybiel, 2013a). In other words, the stronger the DLS activity is at behavior onset, the more vigorous a given maze run is. This DLS pattern also correlates, albeit less so at a trial level, with more runs to a devalued outcome (i.e., to outcome insensitivity in behavior; Smith and Graybiel, 2013a).

In short, loss-of-function studies report a necessary role for the DLS in habits, egocentric navigation, and vigor, but what aspects of these are regulated by DLS activity at behavior onset is unknown; causal manipulations have not been done to evaluate how these aspects of behavior relate to one another. To address these questions, we incorporated an optogenetic approach to stimulate or inhibit the DLS at the start of learned egocentric maze-running behaviors and studied effects on both performance vigor and outcome insensitivity. To pinpoint behavior initiation as a key time point for DLS function, two experiments tested the role of phasic DLS activity in a task that allowed animals to select a full run from the start (a response task) and, separately, a task that required behavior selection at the start of the run and again in the middle of the run (a beacon task).

Materials and Methods

Subjects and surgery

Male Long-Evans rats ($n = 42$ for behavioral training, $n = 3$ for electrophysiology) were individually housed in a dedicated animal vivarium under direct care procedures approved by the Dartmouth College Institutional Animal Care and Use Committee. All rats were obtained at 250–400 g and then maintained on an 85% postsurgical weight for the training and testing duration. Rats were housed on a reverse light/dark cycle, with response task experiments conducted in a dark environment and beacon task experiments conducted in a light environment. Surgical procedures were performed using aseptic techniques under isoflurane anesthesia to allow intracranial injection of viral vectors and implantation of fiber optic guides. Rats received bilateral injections (0.3 μ l volume) into DLS using a microinfusion pump and 33-gauge syringe of a single viral construct: the excitatory channelrhodopsin-2 (ChR2) construct [response task, $n = 8$; beacon task, $n = 10$; adeno-associated virus type 5 (AAV5)-human synapsin (hSYN)-ChR2-enhanced yellow fluo-

rescent protein (EYFP)], the inhibitory eNpHR3.0-halorhodopsin (eNpHR3.0) construct (response task, $n = 8$; AAV5-hSYN-eNpHR3.0-EYFP), or a control vector lacking any opsin transgene (response task, $n = 8$; beacon task, $n = 8$; AAV5-hSYN-EYFP). Bilateral DLS injection and fiber implant coordinates were as follows (in mm): anteroposterior +0.5, mediolateral \pm 4.0, dorsoventral (DV) injection -4.3 from the skull, with fiber implants (200 μ m; ThorLabs or in-house) terminating at DV -3.8 mm. ChR2 expresses a gain-of-function opsin (e.g., artificial sodium channel, acting to activate and depolarize the neuron) at the neuronal cell surface (Allen et al., 2015). In contrast, eNpHR3.0 acts as a loss-of-function opsin (e.g., artificial chloride channel, acting to hyperpolarize the neuron) at the neuronal cell surface (Allen et al., 2015). We nonspecifically targeted DLS neurons using the hSyn promoter given that a majority of projection neurons and interneurons are active at behavior onset (Kubota et al., 2009; Smith and Graybiel, 2013a; O'Hare et al., 2016; Vicente et al., 2016) and play a critical role in habits (Lovinger, 2010; Nelson and Killcross, 2013; Corbit et al., 2014; O'Hare et al., 2018). Fiber implants were permanently affixed to the skull using dental cement and skull screws. Rats received fiber implants and viral vector infusions in a single surgery. Separate rats underwent electrode implants to confirm channelrhodopsin efficacy (see below). All rats had at least 1-week post-surgery recovery time before task exposure.

General experimental design

Maze training. The maze used in all experiments was a custom-built Plexiglas plus maze containing four total maze arms that were identical in length (35.56 cm), width (17.78 cm), and height (33.02 cm) and that were separated by a square center region (diameter 25.4 cm; see Figs. 1A, 4A). Following recovery from surgery, animals were given a habituation session in which free food pellets (45 mg dustless precision pellets: protein 21.3%, fat 3.8%, carbohydrate 54.0%; product #F0165, Bio-Serv) were available at the end arms of the maze. Daily training sessions of 20 trials then proceeded. In a session, rats began in one of two starting arm locations (e.g., either north or south arm in a block design: five trials north arm, five trials south arm, etc., counterbalanced) that were blocked off by a removable insert separating the starting arm from the runway. Once the insert was removed (counterbalanced by direction), it signaled to the rat that a trial had begun. The rat then traversed the starting maze arm to the center decision point, having then a choice of turning left or right. However, only one end arm was baited with a food pellet, and this assignment remained consistent per rat (i.e., response task: turn-direction food pellet; beacon task: cue food pellet; see below). After each trial, the rat would be placed back into the starting arm. There was an interval of \sim 1 min between trials, with the rat waiting in the starting arm location. Therefore, a perfect run would be 20 correct right turns from either a north-arm or south-arm starting location. The assigned turn direction yielding a food pellet was pseudorandomly assigned across animals and groups. Run durations per maze segment, overall run durations, and task events (e.g., start time, trial end, etc.) were recorded by EthoVision XT tracking software (Noldus) via an overhead digital camera connected to a computer. Accuracy of automated behavioral measurements was verified through videotaping and hand scoring of a subset of sessions to verify EthoVision data accuracy. Errors due to poor tracking that were found in EthoVision data postprocessing were corrected by hand-scored analyses directly from individual trial videos. On rare occasions a mechanical issue would occur, and such trials were omitted from analysis. Deliberations were counted as either occurring or not occurring (maximum 1/trial) using an x - y frame of reference for the nose point at the center region of the maze. Deliberations were hand scored, and scorers were blind to group. Deliberations included a clear turn of the head down one maze arm and then a return of the head to forward facing or toward the alternate arm. Training progressed over daily sessions until acquisition of performance accuracy was reached (\geq 75% for 3 d).

DLS manipulations during maze runs. Three DLS manipulation testing days occurred after criterion accuracy was met. On DLS manipulation days, either 473 or 593.5 nm of light from diode-pumped solid-state lasers (Shanghai Laser & Optics Century Co., Ltd.) was delivered through a fiber patch cord (Thorlabs) connected to an integrated rotary joint beam splitter (Doric Lenses) that allowed for two fiber patch cords to be

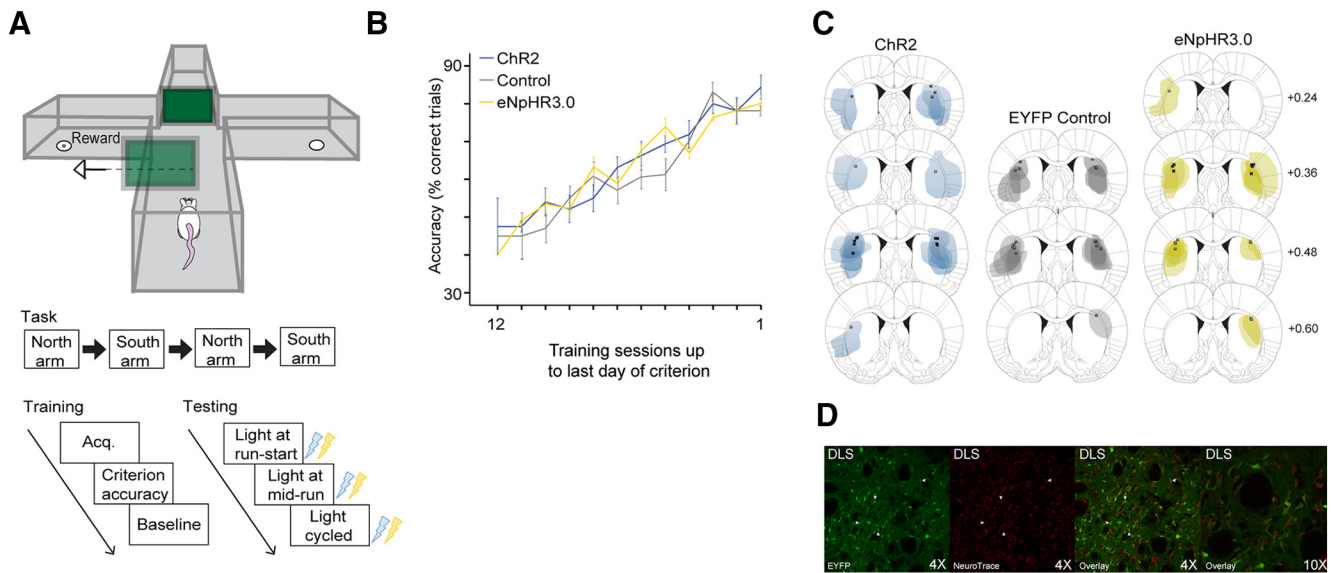


Figure 1. Response task. **A**, Top, Maze task schematic illustrating left-turn example (green, gates). Middle, Example protocol of five-trial blocks with start locations for a session. Bottom left, Session order for task acquisition (through 3 d of $\geq 75\%$ accuracy) and baseline test (fiber cables connected but no light delivery). Bottom right, example DLS manipulation protocol (continuous 0.5 s of light delivery) on test days. **B**, Run accuracy in 12 sessions leading up to the final day of criterion performance, separated by ChR2 (blue), control (gray), and eNpHR3.0 (yellow) groups. Lines and errors show mean \pm SEM. Groups were not different. **C**, Histology maps showing expression areas of ChR2-EYFP, EYFP control, and eNpHR3.0-EYFP across animals. Semitransparent shading denotes expression per animal; shading is overlaid across animals. Numbers denote anteroposterior plane in millimeters relative to bregma. X, fiber implant placements. **D**, Example confocal images of DLS: EYFP, NeuroTrace, low-magnitude overlay, high-magnitude overlay. Arrows show the same neuron examples across labeling.

connected directly to a surgically implanted optic fiber cannula (in-house fiber implants, 200 μm ; Thorlabs) through ceramic sleeves terminating bilaterally into the DLS. Laser delivery was gated either by EthoVision XT software or through a Master-9 pulse generator (A.M.P.I.). Power output was measured to be between 3 and 5 mW at the level of each patch cable ferrule connector per hemisphere before and after each test day. Implants were tested postperfusion to confirm efficacy. On separate manipulation days, laser light was delivered as a 0.5 s pulse gated at the start of a trial, a 0.5 s pulse upon arrival at the middle of the maze, or as a cycled pulse of 0.5 s on/off for the full run duration. A continuous 0.5 s pulse was chosen to capture the DLS temporal activity period seen at maze-running initiation (~ 200 – 500 ms) in the DLS that correlates with maze-running vigor (Smith and Graybiel, 2013a). The cycled 0.5 s on/off parameter was chosen to include illumination at running onset and then more illumination time (i.e., to see whether behavioral changes with illumination at run start were augmented further if more illumination was given during runs) while also minimize heating in the brain that might occur with long-duration light delivery. Days with light delivery at run start or midrun were counterbalanced across groups on both tasks, with cycled light delivery always occurring last.

Outcome devaluation. To operationally define whether maze turn behaviors were driven by habit, we performed an outcome devaluation procedure after the three DLS manipulation days. Rats were given *ad libitum* access to the task food pellets (~ 2.5 g = ~ 50 pellets) for 30 min in their home cage in a separate room from the maze. Afterward, an injection of lithium chloride (0.3 M, 10 ml/kg, i.p.) was given to introduce nausea. Three of these pairings were given at 48 h intervals. Pellet consumption was recorded before and after each procedure. Pellet consumption during task performance was also recorded before and after the devaluation procedure.

Peridevaluation DLS manipulations. Habitual runs were tested operationally using a predevaluation extinction probe day, followed by the outcome devaluation procedure (e.g., three separate pairings of lithium chloride (LiCl) injections following a 20–30 min free intake period in the home cage in a separate context), a postdevaluation extinction (no pellets) probe day, and a postdevaluation (reacquisition) probe day in which pellets could again be earned. DLS manipulations occurred at run start on the predevaluation extinction day, postdevaluation extinction day, and postdevaluation reacquisition day. After the outcome devaluation,

on occasion a rat would refuse to run several trial opportunities in a row (a tendency of the eNpHR3.0 group in the response maze task), in which case the session was concluded. To account for this, data were assessed for trials leading up to the first failed trial opportunity. For analysis of percentage of trials completed within a session before quitting (an effect of outcome devaluation), the first trial a rat failed to complete within 30 s was used (e.g., if it was trial 10 of 20, then the calculation was 50%).

Response task details. One of the two maze tasks used was a response-based maze-running task in which performance is thought to be DLS dependent (Packard and McGaugh, 1996; Chang and Gold, 2004; Palencia and Ragozzino, 2005). For animals undergoing the response task ($n = 24$), training and testing were performed and recorded in a dark environment under a red light to minimize the use of visual cues. Rats running learned to turn in the same direction (i.e., always turn right at a choice point) toward a baited end arm for a single pellet, regardless of maze starting location (i.e., north or south arm). In total, 20 turns (e.g., all right) would yield a food pellet in a daily session. Rats in this response task had the opportunity to correct themselves if they did not enter the correct arm on the first try (i.e., similar to Packard and McGaugh (1996)) and were given a total of 30 s to complete a trial. A circuitous (i.e., multiple arm entries) but ultimately accurate run was recorded as an incorrect trial when calculating overall accuracy due to the lack of a clear response strategy in the run. Trials exceeding 30 s were counted as “misses” and excluded from analysis. In the event of a miss, the session would continue until 20 trials were complete. Daily training sessions were performed until an accuracy criterion was reached ($\geq 75\%$ for 3 d), at which point the DLS manipulations and outcome devaluation proceeded as above.

Beacon task details. The second maze task was a “beacon” task that we derived from the DLS-dependent win-stay and intraenvironment landmark-approach tasks (Packard et al., 1989; McDonald and White, 1993; Sage and Knowlton, 2000; Berke et al., 2009; Kosaki et al., 2015). For this task, one of two removable cues (two vertical XXs or OO) was affixed to one of two maze end arms that could be manually switched. One cue was paired with a food pellet at one end arm, and the other cue was paired with an empty dish in the opposite arm that also stayed consistent across trials. Neither of the cues could be detected until the center region was reached due to the tall opaque maze walls that were covered with a green removable tarp that could be affixed directly to

the maze and easily removed. In this way, rats learned to locate the cue predicting food pellets rather than just guessing and exploring both end arms and not making the association that one of two cues was paired with a food pellet and the other resulted in nothing. In contrast to the response task, in which the run direction could be selected at the beginning of each trial, two major running segments were required for this task: at the start (run to middle) and again at the center (run to pellet-paired cue). All beacon task behavior testing was performed and recorded in a well illuminated environment.

In daily sessions, rats ($n = 18$) began in the same north or south arm locations. Removal of the gate alerted the rat a trial had begun. Sessions consisted of 20 trials per day with a maximum of 20 pellets, with an interval of ~ 1 min between trials. Rats running the beacon task traversed the runway to the center, where the decision (turn left or right) was made based on location of the pellet-paired cue, regardless of the starting maze location (e.g., north or south arm). A block design was also used (five north, five south, five north, five south). Either the XX or OO cue would signify food pellet availability and stay in the same west or east arm for the first 10 trials and then shift to the alternate arm for the remaining 10 trials. All parameters were assigned pseudorandomly across rats. Incorrect turns to the unpaired cue resulted in no food pellet. Unlike the response task, rats that chose the incorrect, unpaired-cue arm were not allowed to correct themselves. For instance, a Plexiglas insert was used to block off the rat from entering the correct arm if it chose the unpaired-cue arm first, resulting in manually placing the rat back in the starting arm location. Training on this beacon task continued until the accuracy criterion, followed by DLS illumination test days, outcome devaluation, and per-devaluation test days as above.

Neuronal spiking data acquisition and processing

Separate male Long–Evans rats ($n = 3$) were used for electrophysiological recording of ChR2-mediated neuronal stimulation, because our stimulation parameters were somewhat different from literature reports on freely behaving animals. These rats were given ChR2-containing viral construct injections into the DLS as above and implanted with an optical fiber (200 μm core, 0.39 NA) surrounded by eight independently moveable tetrodes (VersaDrive-8 Optical, Neuralynx) in either a chronic ($n = 2$; recordings during free behavior) or an acute preparation ($n = 1$; recordings under urethane anesthesia at 1.5 g/kg). Transistor–Transistor Logic (TTL) time stamps from the laser control system and spiking activity (filtered at 600–6000 Hz) were collected using a Digital Lynx SX acquisition system and a pair of HS-18-MM preamplifiers (Neuralynx) with all signals referenced to a skull screw above the cerebellum. During recordings, cycles of blue light deliveries (2–5 mW) were given at 0.5 s pulse duration every minute or as 0.5 s on/off pulses to approximate task conditions. Single sessions included both light delivery protocols when recording stability allowed. Between sessions, tetrodes were lowered and the acquisition of new units was confirmed visually. Units were isolated offline (Offline Sorter, Plexon), plotted (Fig. 1), and analyzed for responsiveness to light delivery (NeuroExplorer, NexTechnologies; SPSS, IBM). A unit was considered responsive if spiking frequency during the light delivery period went beyond a 95% confidence interval of baseline spiking for three or more consecutive 0.02 s time bins. Unit activity was clearly distinguishable from photoelectric artifacts.

Histology

At the termination of each experiment, a lethal dose of anesthesia (sodium pentobarbital) was administered, followed by a transcardial perfusion of 0.9% saline and 4% paraformaldehyde. Brains were put in 20% sucrose solution overnight and frozen at -80°C . Brains were sectioned under a microtome at 30–60 μm , mounted to slides, and coverslipped with a DAPI-containing medium. Fiber placement and zones of EYFP-expressing neurons were assessed using fluorescent microscope analysis (BX53 fluorescent microscope with DP73 camera, Olympus). For estimates of the efficacy of viral infection, brain sections were immunostained to label EYFP (primary and secondary antibodies: rabbit anti-GFP/Alexa Fluor 488 goat anti-rabbit; mouse anti-eYFP/Alexa Fluor 594 goat anti-mouse) and mounted to slides with NeuroTrace (Thermo Fisher Scientific), which is a fluorescent Nissl stain for labeling neurons.

Sections were then analyzed using a confocal microscope (LSM880 with Airyscan laser scanning confocal microscope, Zeiss; Dartmouth College Biology Department Light Microscopy Core Facility). Image acquisition and analysis were quantified using Imaris (Bitplane) and Zen Blue (ZEISS). EYFP, NeuroTrace, and double-labeled neurons were separately counted in pseudorandom sections of DLS using a grid approach from three brains.

Statistical analyses of behavior

Repeated-measures ANOVAs were used to compare differences in dependent behavioral variables [e.g., cumulative maze duration, deliberations (VTEs) at choice point, accuracy, etc.] for within-subject and between-subjects analyses for main effects of the following factors: illumination day (e.g., baseline, run start, midrun, cycled on/off), trial within sessions, virus group (e.g., control vs ChR2 vs eNpHR3.0), and their interactions. These comparisons were conducted separately for four stages based on a priori events of interest: (1) task acquisition; (2) illumination tests after acquisition; (3) across predevaluation and postdevaluation extinction probe days; and (4) during the reacquisition day (the measure of pellets consumed was also compared with an equivalent predevaluation illumination day to confirm devaluation in the task). Due to the a priori interest in a potential trial-level effect within illumination sessions, the degrees of freedom for interaction statistics involving trial in our statistical model were large. If there was a significant main effect for a behavioral variable or interaction between variables for either of the maze tasks (in addition to the postdevaluation reacquisition probe day in both), we used a Tukey-corrected *post hoc* analysis for individual comparisons and individual day comparisons using univariate ANOVAs.

Results

Response task

Training and testing

Rats were first trained on a DLS-dependent plus-maze response task (Chang and Gold, 2004; Palencia and Ragozzino, 2005), in which a specific turn (e.g., left turn) at a choice point was always paired with an appetitive food pellet regardless of starting location (Fig. 1A). Training proceeded until three consecutive days of $\geq 75\%$ accuracy were achieved. Run accuracy in sessions leading up to the final day of criterion performance showed no main effect of group ($F_{(2,243)} = 1.558, p = 0.213$) or day/group interaction ($F_{(22,243)} = 1.349, p = 0.144$), but a main effect of day ($F_{(11,243)} = 44.861, p < 0.001$; Fig. 1B). This indicated that rat groups learned the task similarly before optogenetic interventions. A series of test days (Fig. 1A) followed, during which the DLS was optogenetically stimulated (ChR2) or inhibited (eNpHR3.0) for a continuous 0.5 s at the onset of maze runs, for a continuous 0.5 s during the middle (maze center) of runs, or on a continuous 0.5 s on/0.5 s off cycle for the duration of runs. These manipulations were compared with a control condition in which animals were treated identically but lacked DLS opsin expression. Histology confirmed accurate placement of viral constructs and fibers in the DLS (Fig. 1C). Confocal imaging (GFP immunostain/NeuroTrace) revealed that an estimated 83% of DLS neurons expressed the viral molecules (Fig. 1D). Separate electrophysiological recordings confirmed that continuous 0.5 s ChR2-mediated stimulation increased spiking activity in the DLS (Fig. 2); similar eNpHR3.0-mediated inhibition has been reported (Gradinaru et al., 2009; Smith et al., 2012; Chang et al., 2017).

The averaged cumulative duration of runs (total time to complete a trial) was markedly affected by DLS manipulations (Fig. 3A). This was shown by a within-subject effect of day ($F_{(2,9,1014)} = 4.828, p = 0.003$) and a day/group interaction ($F_{(5,8,1014)} = 3.897, p < 0.001$), but no interactions of day/trial ($F_{(55,1014)} = 0.934, p = 0.613$) or day/group/trial ($F_{(110,1014)} = 1.081, p = 0.276$). Similarly, there was a main between-groups effect ($F_{(2,338)} = 17.992, p < 0.001$) on all three illumination days: run start ($F_{(2,451)} =$

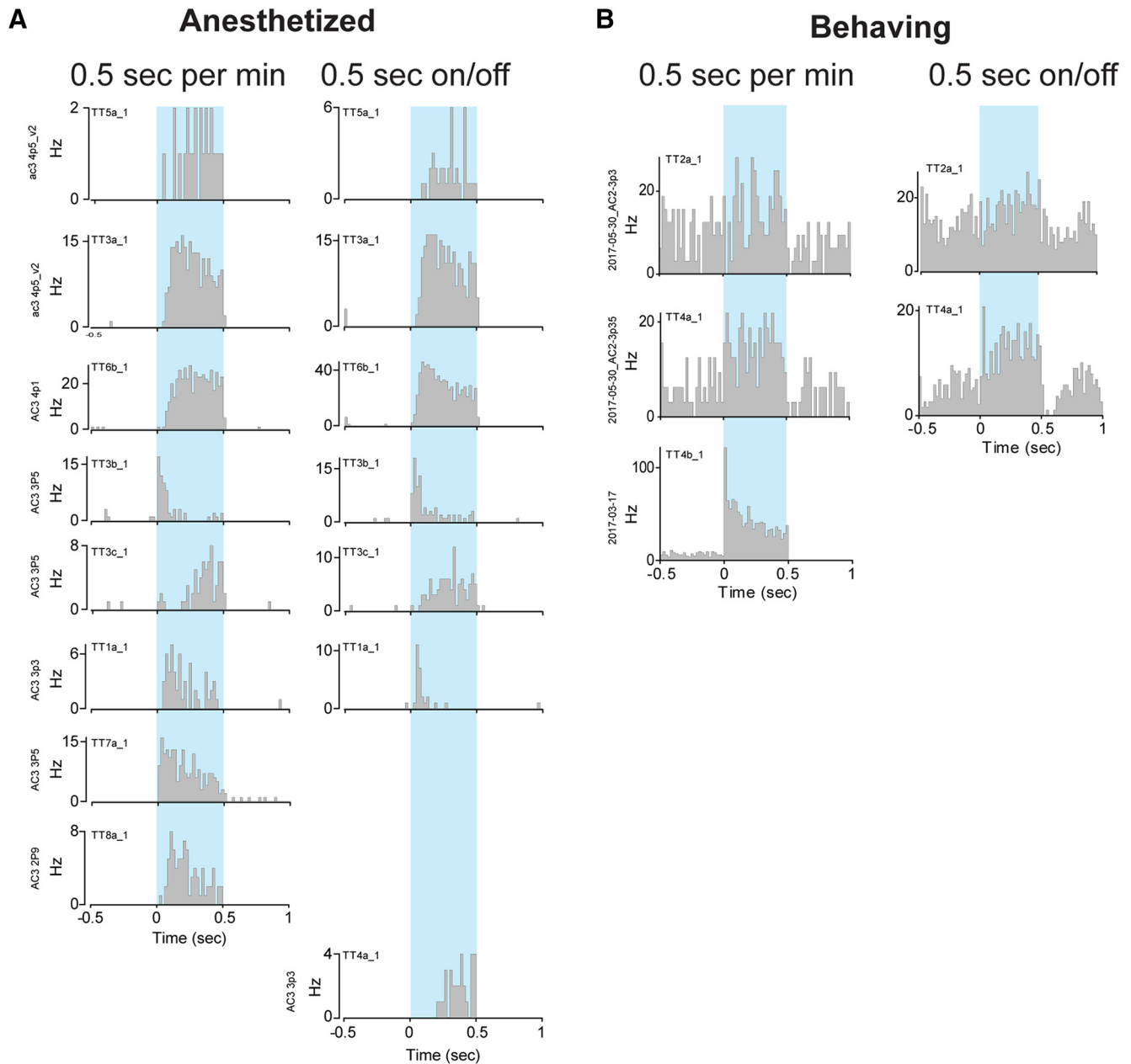


Figure 2. ChR2-evoked stimulation of DLS neurons with continuous 0.5 s blue light. Histograms (0.02 s bins) of all single recorded units that were responsive to blue light. Firing rate (Hz) is shown 0.5 s before, during, and after light delivery. **A**, Units recorded under anesthesia ($n = 9$). **B**, Units recorded during free behavior ($n = 3$). **A, B**, Columns show the same unit recorded when light was delivered for 0.5 s every 1 min (left columns) and when light was delivered in a 0.5 s on/off cycle (right columns). Gaps denote units that were not stable or were not detected during that light delivery protocol. Although not all units responded to light immediately, they nonetheless exhibited response excitation as confirmation of ChR2-mediated stimulation of the DLS. Units with activity inhibited by light were not observed.

20.733, $p < 0.001$), midrun ($F_{(2,447)} = 7.663$, $p < 0.001$), and cycled ($F_{(2,449)} = 34.748$, $p < 0.001$). *Post hoc* comparisons for run-start illumination showed that the ChR2 group was faster than controls ($p < 0.001$) and eNpHR3.0 rats ($p < 0.001$) but the eNpHR3.0 group was not significantly faster than controls ($p = 0.751$). Midrun stimulation was no different from controls ($p = 0.326$), but inhibition resulted in slower trials against controls ($p = 0.033$) and ChR2 rats ($p < 0.001$). Cycled illumination *post hoc* comparisons showed bidirectional effects of stimulation and inhibition, with stimulation speeding up performance and inhibition slowing it: ChR2/control ($p < 0.001$), eNpHR3.0/control ($p < 0.001$), and ChR2/eNpHR3.0 ($p < 0.001$).

The number of deliberative head movements (VTEs) at the choice point in the center of the maze followed a similar pattern (Fig. 3B) to cumulative duration. We found a within-subject effect of day ($F_{(3,1014)} = 3.423$, $p = 0.017$) and day/group interaction ($F_{(6,1014)} = 5.858$, $p < 0.001$), but no interaction of day/trial ($F_{(57,1014)} = 0.963$, $p = 0.555$) or day/group/trial ($F_{(114,1014)} = 1.030$, $p = 0.402$). Similar to run duration changes, there was a main between-groups effect ($F_{(2,338)} = 11.023$, $p < 0.001$). Days showing significant effects were run-start illumination ($F_{(2,451)} = 22.148$, $p < 0.001$) and cycled illumination ($F_{(2,449)} = 25.975$, $p < 0.001$), with *post hoc* comparisons revealing the ChR2 group deliberated less than the control ($p < 0.001$) and eNpHR3.0 ($p <$

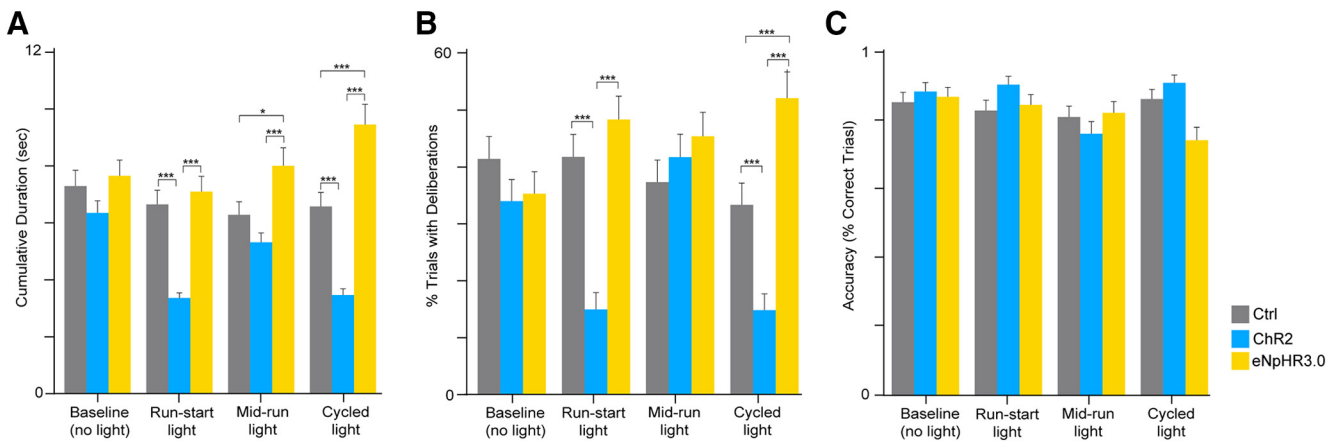


Figure 3. Changes in response maze behavior with DLS manipulations. **A**, Duration of maze runs (seconds). DLS stimulation reduced run duration if given at run start or cycled. DLS inhibition could increase run duration marginally when given midrun and when cycled. **B**, Percentage of trials with deliberations. Increases and decreases in deliberation generally followed increases and decreases in run duration. **C**, Percentage of correct running across the baseline and illumination test days. DLS manipulations did not affect run accuracy. Bars and errors show mean \pm SEM; asterisks denote significant *post hoc* comparisons (* $p < 0.05$, *** $p < 0.001$); lack of asterisks denotes lack of significance.

0.001) groups but showing no difference between eNpHR3.0 and controls ($p = 0.422$). Deliberation behavior on cycled illumination days was similar but with cycled inhibition carrying a significant deliberation-increasing effect (ChR2/control: $p < 0.001$; eNpHR3.0/control: $p < 0.001$; ChR2/eNpHR3.0: $p < 0.001$). Thus, deliberations were reduced with run-start or cycled DLS stimulation, as was run duration. Midrun stimulation and inhibition did not affect deliberations. Deliberations were increased only with cycled DLS inhibition.

The accuracy of runs was not affected by DLS run-start stimulation (Fig. 3C), suggesting that the use of task rules to inform behavior was unaffected. There were no main within-subject effects of day ($F_{(3,1017)} = 2.087$, $p = 0.100$), nor were there interactions of day/group ($F_{(6,1017)} = 1.641$, $p = 0.133$), day/trial ($F_{(57,1017)} = 0.881$, $p = 0.722$), day/group/trial ($F_{(114,1017)} = 1.084$, $p = 0.268$), or a main effect between groups ($F_{(2,339)} = 1.073$, $p = 0.343$). This result shows that the vigor of runs on a DLS-dependent response task is powerfully modulated by DLS activity, especially at the initial onset of behavior. We were surprised that accuracy was unaffected, given reports described above that response-based maze navigation requires the DLS.

Analysis of run durations in the major maze segments (latency to enter maze center from run initiation, maze center, end arm; Table 1) indicated that changes in vigor were not related strictly to changes in the initiation of the run when DLS perturbation occurred, nor were they only related to changes in run duration in the middle segment (when deliberations would occur, as below). Thus, full run durations were reduced by DLS stimulation applied at run start (and by cycled stimulation, which included a run-start time point), but not by midrun stimulation. DLS inhibition at run start did nothing, but did increase run duration when applied midrun and particularly when cycled.

Devaluation

We next examined how the change in run vigor produced by run-start manipulations related to habit by formal definition through an outcome devaluation procedure. This procedure tests for the way changes in the value of the expected outcome and received outcome were used in behavior (Fig. 4). A more (or less) habitual run would exhibit independence from (or sensitivity to) changes in the outcome value (i.e., outcome insensitivity).

We first established a baseline by giving animals a probe test under extinction conditions (no pellets given) with DLS light delivery at run start, followed by a normal retraining reminder session without any light delivery (Fig. 4A). Then, free intake of the food pellets in home cages was paired with nauseogenic LiCl until an aversion to the pellets developed. Each rat received three pairings, and all rats rejected all pellets by the last pairing in their home cages without any group differences (Fig. 4B). Rats were then returned to the task and exposed to a postdevaluation extinction probe test followed by a postdevaluation reacquisition (i.e., pellets returned) test, both with DLS light being delivered at run start.

All animals ran to the devalued goal location in the postdevaluation extinction probe test at a level of accuracy that was comparable with the predevaluation extinction test (Fig. 4C). There were no significant differences in run accuracy between extinction sessions for day ($F_{(1,123)} = 3.187$, $p = 0.077$), group ($F_{(2,123)} = 0.819$, $p = 0.443$), or day/group interaction ($F_{(2,123)} = 0.562$, $p = 0.572$). There was also no difference in accuracy (when rats did run) during the postdevaluation reacquisition day ($F_{(2,75)} = 1.309$, $p = 0.276$). Although animals were accurate when they did run, we found a more sensitive measure of outcome sensitivity to be task “quitting,” in which animals neglected to even start a run (Fig. 4D). Statistically, between predevaluation and postdevaluation extinction days (i.e., extinction sessions occurring before and after outcome devaluation), there was a main within-subject effect of day ($F_{(1,24)} = 64.444$, $p < 0.001$), a day/group interaction ($F_{(2,24)} = 35.008$, $p < 0.001$), and a main between-groups effect ($F_{(2,24)} = 35.008$, $p < 0.001$). *Post hoc* comparisons showed this to be due to eNpHR3.0 rats failing to initiate runs more often (ChR2/eNpHR3.0, $p < 0.001$; control/eNpHR3.0, $p < 0.001$; ChR2/control, $p = 0.735$). The same trend, although nonsignificant, occurred during reacquisition ($F_{(1,9)} = 0.117$, $p = 0.077$). We interpret this as evidence that the initiation of the run was the time at which animals did or did not process the change in expected outcome value. In this case, control and ChR2 animals were habitual, running on nearly all opportunities to do so in the postdevaluation extinction session. During reacquisition, both groups similarly reduced their choice to run by $\sim 50\%$, confirming that they could integrate the devalued outcome only after it was experienced on the task. In stark contrast, rats with run-start

Table 1. Statistics for response task maze segments

Maze segment	Main comparisons	Baseline	Illumination at run start	Illumination mid run	Cycled illumination
Latency to enter maze center	Within-subject effects: main effect day, $F_{(2,9,1011)} = 2.316, p = 0.077$; interactions: day/group, $F_{(5,7,1011)} = 1.678, p = 0.127$; day/trial, $F_{(54,6,1011)} = 1.019, p = 0.439$, or day/group/trial, $F_{(109,1011)} = 1.137, p = 0.171$; between-subjects effects: main effect group $F_{(2,337)} = 53.425, p < 0.001$	Control versus Chr2 (SE = 0.334, $p = 0.622$), control versus eNpHR3.0 (SE = 0.335, $p < 0.001$), and Chr2 versus eNpHR3.0 (SE = 0.336, $p < 0.001$)	Control versus Chr2 (SE = 0.265, $p = 0.289$), control versus eNpHR3.0 (SE = 0.265, $p < 0.001$), and Chr2 versus eNpHR3.0 (SE = 0.269, $p < 0.001$)	Control versus Chr2 (SE = 0.311, $p = 0.992$), control versus eNpHR3.0 (SE = 0.316, $p < 0.001$), and Chr2 versus eNpHR3.0 (SE = 0.320, $p < 0.001$)	Control versus Chr2 (SE = 0.262, $p = 0.087$), control versus eNpHR3.0 (SE = 0.267, $p < 0.001$), and Chr2 versus eNpHR3.0 (SE = 0.266, $p < 0.001$)
Maze center	Within-subject effects: main effect day, $F_{(2,7,1011)} = 4.996, p = 0.003$; interactions: day/group $F_{(5,4,1011)} = 0.748, p = 0.597$; day/trial, $F_{(51,1011)} = 0.862, p = 0.743$; or day/group/trial, $F_{(102,1011)} = 1.011, p = 0.456$; between-subjects effects: Main effect group, $F_{(2,337)} = 14.269, p < 0.001$	Control versus Chr2 (SE = 0.460, $p = 0.002$), control versus eNpHR3.0 (SE = 0.462, $p = 0.002$), and Chr2 versus eNpHR3.0 (SE = 0.463, $p = 0.998$)	Control versus Chr2 (SE = 0.375, $p < 0.001$), control versus eNpHR3.0 (SE = 0.373, $p = 0.017$), and Chr2 versus eNpHR3.0 (SE = 0.380, $p = 0.074$)	Control versus Chr2 (SE = 0.347, $p < 0.001$), control versus eNpHR3.0 (SE = 0.353, $p = 0.016$), and Chr2 versus eNpHR3.0 (SE = 0.357, $p = 0.748$)	Control versus Chr2 (SE = 0.408, $p = 0.002$), control versus eNpHR3.0 (SE = 0.41601, $p = 0.701$), and Chr2 versus eNpHR3.0 (SE = 0.415, $p < 0.001$)
End arm	Within-subject effects: main effect day, $F_{(3,996)} = 4.411, p = 0.004$; interactions: day/group, $F_{(6,966)} = 3.619, p < 0.001$; day/trial, $F_{(57,966)} = 1.069, p = 0.341$; or day/group/trial, $F_{(114,966)} = 1.149, p = 0.147$; between-subjects effects: main effect group, $F_{(2,332)} = 2.402, p = 0.092$	Control versus Chr2 (SE = 0.394, $p = 0.667$), control versus eNpHR3.0 (SE = 0.397, $p = 0.756$), and Chr2 versus eNpHR3.0 (SE = 0.398, $p = 0.264$)	Control versus Chr2 (SE = 0.327, $p = 0.006$), control versus eNpHR3.0 (SE = 0.326, $p = 0.926$), and Chr2 versus eNpHR3.0 (SE = 0.332, $p = 0.020$)	Control versus Chr2 (SE = 0.383, $p = 0.691$), control versus eNpHR3.0 (SE = 0.392, $p = 0.730$), and Chr2 versus eNpHR3.0 (SE = 0.397, $p = 0.999$)	Control versus Chr2 (SE = 0.406, $p = 0.018$), control versus eNpHR3.0 (SE = 0.413, $p = 0.055$), and Chr2 versus eNpHR3.0 (SE = 0.412, $p < 0.001$)

Repeated-measures ANOVA was conducted for each maze segment by day, with trial and group (i.e., virus) as factors. Tukey *post hoc* corrections were performed for main effects in subsequent columns (i.e., baseline day, etc.). Significant comparisons appear in bold.

DLS inhibition chose not to run on the vast majority of trials in the postdevaluation extinction test, showing that they were goal directed in behavior. It is highly unlikely that the behavioral changes caused by eNpHR3.0 during the postdevaluation testing days were due to motor impairment, because identical manipulations (in the same rats and same test conditions) did not produce this same effect before devaluation.

On each test day (predevaluation and postdevaluation probe; postdevaluation reacquisition), DLS stimulation at run start again produced a pattern of vigor improvement as measured first by reduced run duration (Fig. 4E). Unlike during predevaluation sessions, DLS inhibition at run start produced a sizeable increase in run duration (Fig. 4E). The eNpHR3.0-mediated increase in duration occurred before outcome devaluation during the extinction session as well as in postdevaluation sessions. For run duration, there was a main within-subject effect of day ($F_{(1,82)} = 25.139, p < 0.001$) and day/group interaction ($F_{(2,82)} = 7.471, p < 0.001$), but no interaction of day/trial ($F_{(19,82)} = 0.826, p = 0.671$) nor day/group/trial ($F_{(22,82)} = 1.001, p = 0.473$). There were main between-groups effects ($F_{(2,82)} = 22.703, p < 0.001$) on all three illumination days: predevaluation extinction ($F_{(2,202)}$

$= 41.428, p < 0.001$; eNpHR3.0/control, $p < 0.001$; eNpHR3.0/Chr2, $p < 0.001$; control/Chr2, $p = 0.237$), postdevaluation extinction ($F_{(2,82)} = 18.935, p < 0.001$; eNpHR3.0/control, $p < 0.001$; eNpHR3.0/Chr2, $p < 0.001$; control/Chr2, $p = 0.010$), and postdevaluation reacquisition ($F_{(2,48)} = 14.366, p < 0.001$; eNpHR3.0/control, $p < 0.001$; eNpHR3.0/Chr2, $p < 0.001$; control/Chr2, $p = 0.122$). Figure 4F illustrates how different animals given DLS inhibition or stimulation at run start compared with controls in their run duration (percentage change), showing run-start manipulations during normal testing, as above, and during the devaluation testing phases.

Deliberative head movements were affected similarly to run duration. For deliberation head movements (Fig. 4G), there were no main within-subject effects across predevaluation extinction versus postdevaluation extinction days ($F_{(1,83)} = 3.084, p = 0.083$), nor were there interactions: day/group ($F_{(2,83)} = 0.592, p = 0.555$), day/trial ($F_{(19,83)} = 0.541, p = 0.935$), and day/group/trial ($F_{(21,83)} = 0.676, p = 0.845$). Still, there was a main between-groups effect ($F_{(2,83)} = 15.612, p < 0.001$) on all three illumination days: predevaluation extinction ($F_{(2,203)} = 32.478, p < 0.001$; control/Chr2, $p = 0.050$, control/eNpHR3.0, $p <$

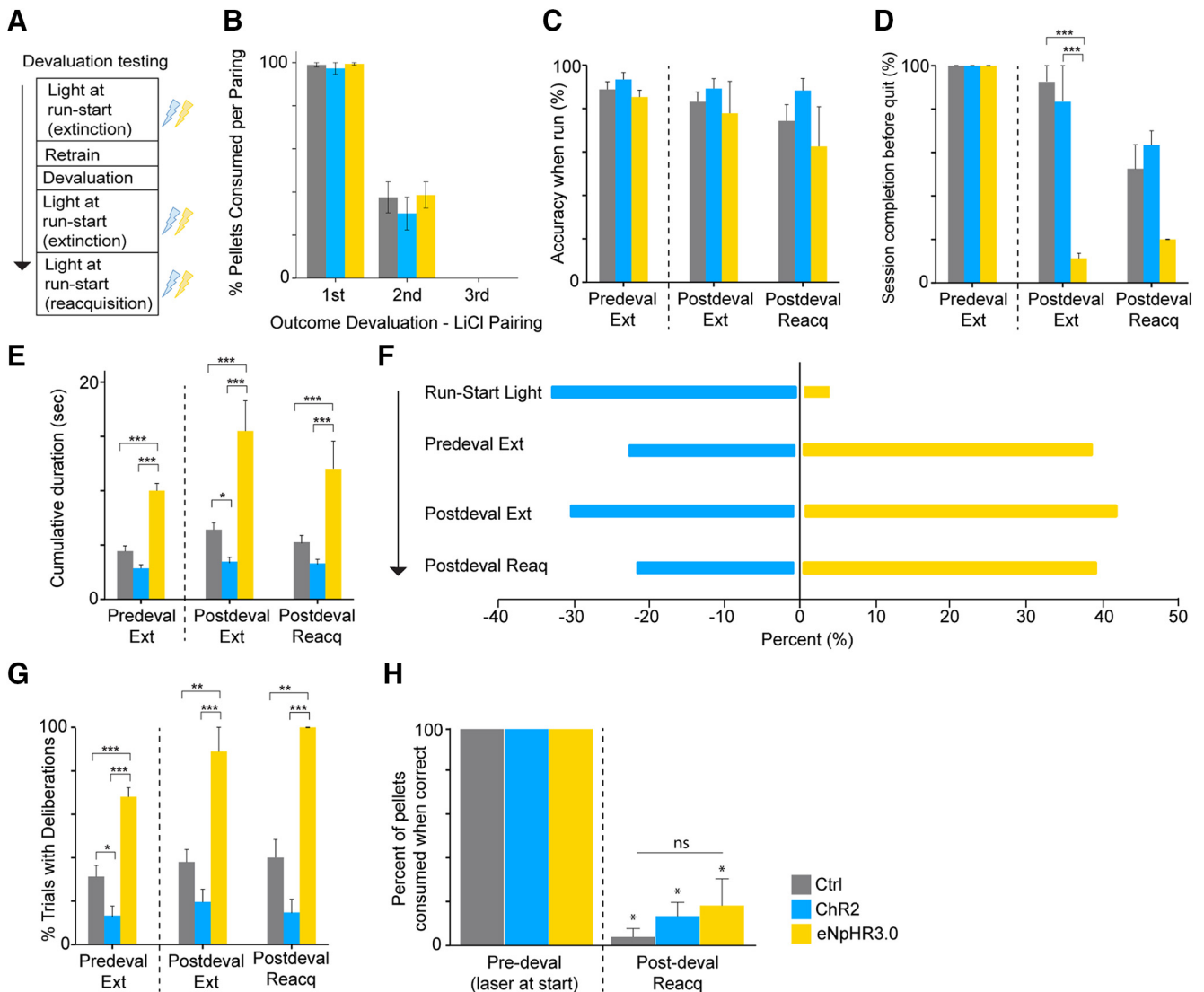


Figure 4. Response maze behavior related to outcome devaluation. **A**, Devaluation testing timeline. Light was given at run start in a predevaluation extinction session (followed by a no-light retraining session) and postdevaluation extinction and reacquisition (pellets delivered) sessions. Gray, controls; blue, Chr2; yellow, eNpHR3.0. **B**, Pellet consumption during each of the three pellet/LiCl pairings to cause devaluation. All animals rejected pellets by the third pairing. **C**, Accuracy of maze running in trials in which runs occurred. No group differences. **D**, Percentage of trials in each session completed before quitting the task. Animals ran all trials before outcome devaluation. After devaluation, animals ran by habit in Chr2 and control groups, but animals with DLS inhibition ran very few trials, showing devaluation sensitivity. **E**, Average cumulative duration of trial runs (seconds). DLS made animals faster on each day, whereas DLS inhibition made animals slower. **F**, Percentage change in cumulative duration relative to controls. Blue, Chr2 group versus control; yellow, eNpHR3.0 versus control. **G**, Percentage of trials with deliberation. These again mirrored changes in run duration. **H**, Percentage of pellets consumed in task (before task quitting), which was reduced across groups and trials from predevaluation to postdevaluation with illumination at run start. Groups did not differ in reduced pellet consumption after devaluation. Bars and errors show mean \pm SEM; asterisks denote significant *post hoc* comparisons ($*p < 0.05$, $**p < 0.01$, $***p < 0.001$); lack of asterisks denotes lack of significance. ns, Not significantly different.

0.001; Chr2/eNpHR3.0, $p < 0.001$), postdevaluation extinction: ($F_{(2,83)} = 10.765$, $p < 0.001$; control/eNpHR3.0, $p = 0.006$; Chr2/eNpHR3.0, $p < 0.001$), and postdevaluation reacquisition ($F_{(2,50)} = 11.928$, $p < 0.001$; control/eNpHR3.0, $p = 0.003$; Chr2/eNpHR3.0, $p < 0.001$).

When animals did run the task during reacquisition after outcome devaluation, they rejected nearly all of the pellets, as seen in a comparison of this day with the normal predevaluation run-start illumination day (Fig. 4H). Within-subject analysis showed an effect of day ($F_{(1,356)} = 2776.301$, $p < 0.001$) and interactions [day/trial ($F_{(14,356)} = 7.981$, $p < 0.001$), day/group ($F_{(2,356)} = 7.981$, $p < 0.001$), day/group/trial ($F_{(25,356)} = 4.434$, $p < 0.001$)] but no main between-subjects effect of group on the predevaluation day ($F_{(2,262)} = 0$) or reacquisition ($F_{(2,94)} = 2.177$, $p =$

0.119). Groups did not differ in their pellet rejection, confirming that the devaluation memory generalized equivalently to the task and indicating that the tendency of rats with DLS inhibition to quit the task was not related to them having a stronger pellet aversion. Again, changes in run duration were not limited to behavior at the initial maze segment when illumination was given (Table 2).

In short, DLS stimulation at run start increased run vigor similarly before and after pellet devaluation. At the same time, the frequency of neglecting to run the task, run accuracy when they did run, and level of pellet rejection in the Chr2 group were all similar to controls after devaluation. There may have been a ceiling for increasing how habitual (outcome-insensitive) the runs were, although there was still “room” for vigor improvement of

Table 2. Statistics for response task with devaluation

Maze segment	Main comparisons	Predevaluation extinction	Postdevaluation extinction	Predevaluation reacquisition
Latency to enter maze center	Within-subject effects: main effect day , $F_{(1,82)} = 4.438, p < 0.038$; interactions: day/group, $F_{(2,82)} = 1.056, p = 0.353$; day/trial, $F_{(19,82)} = 0.302, p = 0.998$; or day/group/trial, $F_{(22,82)} = 0.554, p = 0.941$; between-subjects effects: main effect group $F_{(2,82)} = 4.446, p = 0.015$	Control versus ChR2 (SE = 0.436, $p = 0.286$), control versus eNpHR3.0 (SE = 0.367, $p = 0.004$), and ChR2 versus eNpHR3.0 (SE = 0.402, $p < 0.001$)	Control versus ChR2 (SE = 0.401, $p = 0.196$), control versus eNpHR3.0 (SE = 0.750, $p = 0.109$), and ChR2 versus eNpHR3.0 (SE = 0.773, $p = 0.014$)	Between-subjects effects: main effect group, $F_{(2,49)} = 1.167, p = 0.320$; main effect trial, $F_{(13,52)} = 1.847, p = 0.060$; group/trial interaction, $F_{(21,52)} = 0.850, p = 0.649$
Maze center	Within-subject effects: main effect day $F_{(1,82)} = 35.097, p < 0.001$; interactions: day/group , $F_{(2,82)} = 11.352, p < 0.001$; day/trial, $F_{(19,82)} = 0.764, p = 0.741$; or day/group/trial, $F_{(22,82)} = 0.873, p = 0.628$; between-subjects effects: main effect group , $F_{(2,82)} = 25.837, p < 0.001$	Control versus ChR2 (SE = 0.453, $p = 0.793$), control versus eNpHR3.0 (SE = 0.381, $p < 0.001$), and ChR2 versus eNpHR3.0 (SE = 0.418, $p < 0.001$)	Control versus ChR2 (SE = 0.359, $p = 0.129$), control versus eNpHR3.0 (SE = 0.670, $p < 0.001$), and ChR2 versus eNpHR3.0 (SE = 0.690, $p < 0.001$)	Between-subjects effects: main effect group , $F_{(2,49)} = 37.673, p < 0.001$; main effect trial, $F_{(12,49)} = 3.794, p < 0.001$; group/trial interaction , $F_{(14,49)} = 4.483, p < 0.001$ Control versus ChR2 (SE = 0.238, $p = 0.488$), control versus eNpHR3.0 (SE = 0.391, $p < 0.001$), and ChR2 versus eNpHR3.0 (SE = 0.391, $p < 0.001$)
End arm	Within-subject effects: main effect day, $F_{(1,83)} = 3.799, p = 0.055$; interactions: day/group, $F_{(2,83)} = 1.226, p = 0.299$; day/trial, $F_{(19,83)} = 0.518, p = 0.947$; or day/group/trial, $F_{(21,83)} = 0.934, p = 0.551$; between-subjects effects: main effect group , $F_{(2,83)} = 9.130, p < 0.001$	Control versus ChR2 (SE = 0.700, $p = 0.630$), control versus eNpHR3.0 (SE = 0.590, $p < 0.001$), and ChR2 versus eNpHR3.0 (SE = 0.646, $p < 0.001$)	Control versus ChR2 (SE = 0.646, $p = 0.068$), control versus eNpHR3.0 (SE = 1.208, $p = 0.044$), and ChR2 versus eNpHR3.0 (SE = 1.244, $p = 0.002$)	Between-subjects effects: main effect group , $F_{(2,47)} = 7.389, p = 0.002$; main effect trial, $F_{(12,47)} = 0.654, p = 0.785$; group/trial interaction, $F_{(14,47)} = 1.054, p = 0.421$ Control versus ChR2 (SE = 0.594, $p = 0.154$), control versus eNpHR3.0 (SE = 0.964, $p < 0.001$), and ChR2 versus eNpHR3.0 (SE = 0.959, $p = 0.002$)

Repeated-measures ANOVA was conducted for each maze segment by day, with trial and group (i.e., virus) as factors in comparing predevaluation versus postdevaluation extinction days. Tukey *post hoc* corrections were made for main effects in subsequent columns (i.e., predevaluation and postdevaluation extinction). Predevaluation reacquisition day was analyzed separately in the last column. Significant comparisons are in bold.

those runs [i.e., the lowest possible cumulative duration in any rat for (1) postdevaluation extinction (ChR2, 2.3 s; control, 4.2 s) and (2) reacquisition (ChR2, 2.3 s; control, 4.0 s)]. In contrast, DLS inhibition at run start produced a profound increase in outcome sensitivity and reduction of vigor when animals did run. The reduced vigor occurred mainly when pellets were omitted (extinction probe) or devalued, suggesting a large difference in DLS inhibition effects on behavior when outcome values are stable versus reduced. Plausibly, continued running during the sessions, when they are reduced, is mostly DLS dependent, which is why inhibition produced such robust effects. Specifically, on each test day, including the predevaluation probe, the runs of eNpHR3.0 animals with run-start inhibition (when they did run) were far slower, and deliberations occurred on nearly every trial. Moreover, these eNpHR3.0 animals mainly quit the task after outcome devaluation when DLS illumination was given at run start, showing a marked reduction in how habitual their runs were, and this reduction was in close alignment with the reduction in performance vigor.

Beacon task

Training and testing

The results from the response maze experiment raised the question of whether the improvement in running vigor that was observed with run-start DLS stimulation occurred because that was the point in time at which a full run routine was selected and set in motion. To explore this possibility, we used a separate cohort of rats and the beacon plus-maze task described in Materials and Methods. In this task, rats learned to navigate toward a visual cue for a food pellet (Fig. 5A). The logic was to split the maze run into two segments, one which required

animals to traverse the start arm of the maze to the center to locate the visual cue, and a second which required them to traverse the arm containing that cue to receive food pellets. We focused on comparing ChR2-mediated DLS stimulation with controls. The same training and optogenetic manipulation protocol used for the above response task was used for the beacon task (Fig. 5A). Histology confirmed accurate viral vector and fiber placements in the DLS (Fig. 5B). Training on this task led to run accuracy improving over sessions leading up to the final day of criterion performance, and groups did not differ in this initial task learning (Fig. 5C). There was no main effect of group ($F_{(1,258)} = 1.018, p = 0.314$) or day/group interaction ($F_{(28,158)} = 0.914, p = 0.574$), but there was a main effect of day ($F_{(21,253)} = 5.514, p < 0.001$).

Increased running vigor was replicated in this beacon task condition with DLS stimulation at run start. Remarkably, in contrast to the response task, midrun DLS stimulation was suddenly also effective at improving overall running vigor on this task. Enhancing DLS activity at this point in the run, which was when the cue was located and approached, also increased vigor at levels comparable with the increase caused by run-start stimulation. Cycled stimulation was essentially identical as well.

First, the DLS stimulations reduced run duration (Fig. 6A). There was no main within-subject effect of day ($F_{(1,7,954)} = 1.190, p = 0.299$), nor were there interactions of day/trial ($F_{(31,7,954)} = 1.102, p = 0.323$) or day/group/trial ($F_{(31,7,954)} = 0.565, p = 0.974$), but there was a day/group interaction ($F_{(1,7,954)} = 7.727, p < 0.001$). Between-groups comparison showed no difference on the baseline day that lacked illumination ($F_{(1,320)} = 0.739, p = 0.390$) but significant differences on run-start ($F_{(1,320)} = 13.729,$

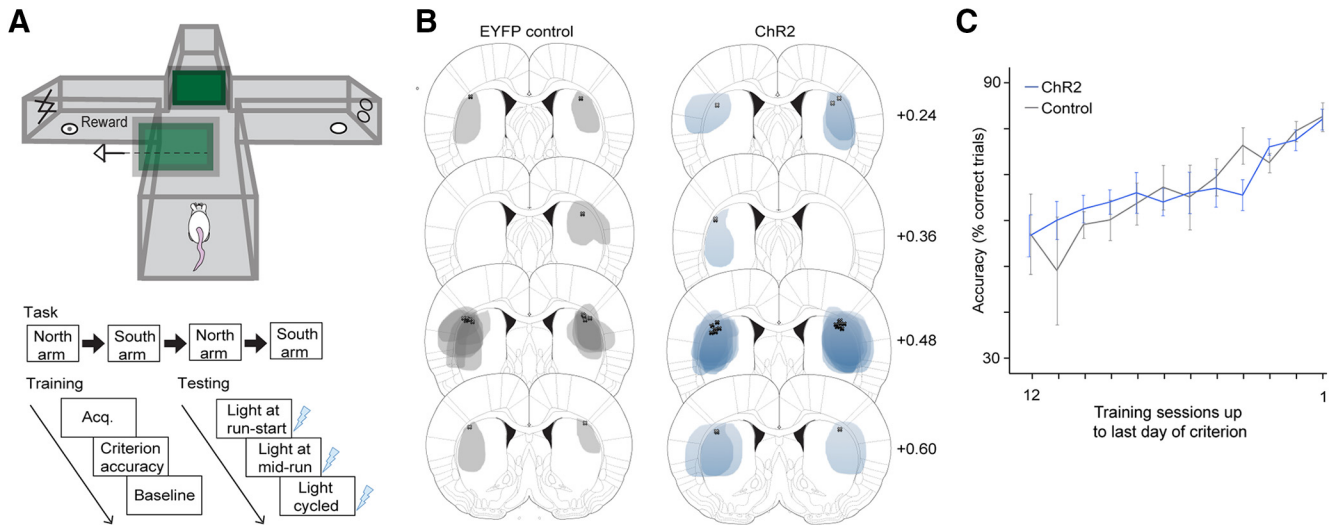


Figure 5. Beacon task. **A**, Maze setup and schematics for trial blocks, training, and DLS manipulation testing as in Figure 1A. **B**, Histological maps showing control and ChR2 expression, as in Figure 1C. **C**, Run accuracy in 12 sessions leading up to the final day of criterion performance, separated by ChR2 (blue) and control (gray) groups (no group differences). Lines and errors show mean \pm SEM.

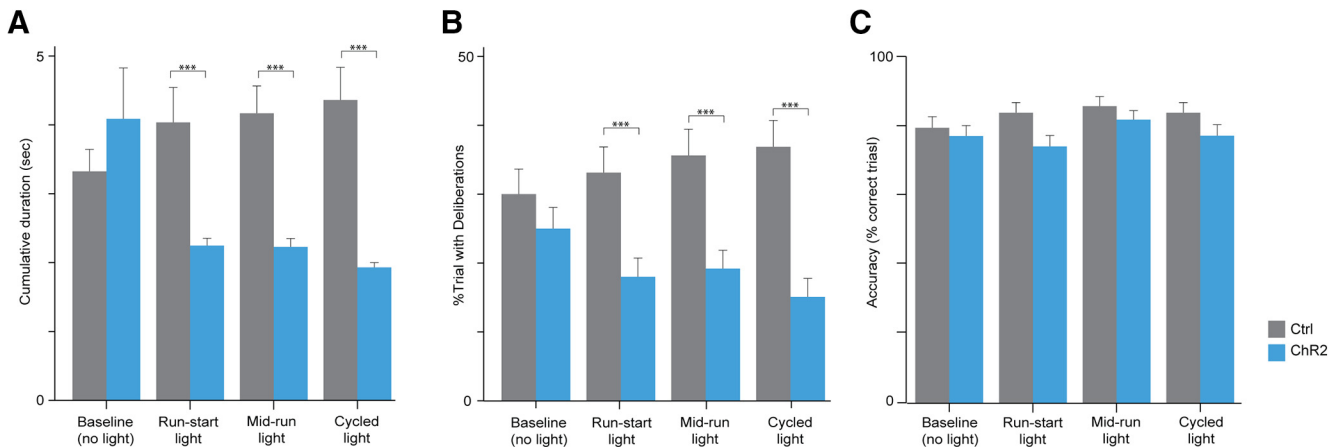


Figure 6. Changes in beacon maze-running behavior with DLS manipulation. **A**, Average cumulative duration of maze runs (seconds). DLS stimulation reduced run duration whether given at run start, midrun, or cycled. Blue, ChR2; gray, controls. **B**, Percentage of trials with deliberation, which DLS stimulation also reduced. **C**, Percentage of trials run correctly, which did not differ between groups. Asterisks denote significant *post hoc* comparisons (** $p < 0.001$).

$p < 0.001$), midrun ($F_{(1,319)} = 30.453, p < 0.001$), and cycled illumination days ($F_{(1,319)} = 23.895, p < 0.001$).

Second, DLS stimulations also reduced deliberations at the center choice point of the maze (Fig. 6B). There was no main within-subject effect of day ($F_{(3,954)} = 0.229, p = 0.876$) nor interactions of day/trial ($F_{(57,954)} = 1.282, p = 0.082$) or day/group/trial ($F_{(57,954)} = 0.823, p = 0.822$), but there was a day/group interaction ($F_{(3,954)} = 2.622, p = 0.050$). Similar to run duration, there was no between-groups difference on the baseline day ($F_{(1,320)} = 1.111, p = 0.293$), but there were significant group differences during illumination at run start ($F_{(1,320)} = 11.381, p < 0.001$), midrun ($F_{(1,319)} = 18.376, p < 0.001$), and cycled ($F_{(1,319)} = 17.367, p < 0.001$) days. Thus, each type of DLS stimulation reduced deliberations.

As with the response task, no DLS manipulation affected running accuracy on this beacon task (Fig. 6C). There was no main within-subject effect for day ($F_{(3,954)} = 1.531, p = 0.205$), nor interactions: day/group ($F_{(3,954)} = 0.703, p = 0.550$), day/trial ($F_{(57,954)} = 0.791, p = 0.868$), day/group/trial ($F_{(57,954)} = 1.107, p = 0.277$). Thus, DLS stimulation at run start and midrun could

decrease run duration, duration variation across trials, and deliberations at the choice point but did not affect the animals' use of task rules to dictate running choices.

In contrast to the response task, we observed that the effect of enhanced vigor that occurred across the illumination sessions did relate to running changes in specific parts of the task, specifically the latency of animals to enter the maze center from the start of the trial. For this measure, there was a main interaction effect of day/group ($F_{(2,4,954)} = 5.117, p = 0.004$) and a between-subjects effect of group ($F_{(1,318)} = 21.728, p < 0.001$; Table 3).

Devaluation

Our outcome devaluation protocol then followed (Fig. 7A), in which LiCl was paired with the food pellets to establish an aversion (Fig. 7B). Following the outcome devaluation, all animals ran this task habitually by running just as accurately as before and electing to run nearly all trials (Fig. 7C,D). Analyses included comparison of extinction days (predevaluation and postdevaluation) and a comparison of groups within the reacquisition day.

Table 3. Statistics for beacon task maze segments

Maze segment	Main comparisons	Baseline	Illumination at run start	Illumination midrun	Cycled illumination
Latency to enter maze center	Within-subject effects: main effect day, $F_{(2,4,954)} = 0.246, p = 0.819$; interactions: day/group, $F_{(2,4,954)} = 5.117, p = 0.004$; day/trial, $F_{(45,2,954)} = 1.207, p = 0.169$; or day/group/trial, $F_{(45,2,954)} = 0.859, p = 0.733$; between-subjects effects: main effect group, $F_{(1,318)} = 21.728, p < 0.001$	Between-subjects effects: main effect group, $F_{(1,320)} = 0.550, p = 0.459$; main effect trial, $F_{(19,320)} = 1.329, p = 0.162$; group/trial interaction, $F_{(19,320)} = 0.415, p = 0.987$	Between-subjects effects: main effect group, $F_{(1,320)} = 14.668, p < 0.001$; main effect trial, $F_{(19,320)} = 0.737, p = 0.779$; group/trial interaction, $F_{(19,320)} = 0.683, p = 0.836$	Between-subjects effects: main effect group, $F_{(1,319)} = 24.632, p < 0.001$; main effect trial, $F_{(19,319)} = 0.620, p = 0.891$; group/trial interaction, $F_{(19,319)} = 0.319, p = 0.997$	Between-subjects effects: main effect group, $F_{(1,319)} = 22.864, p < 0.001$; main effect trial, $F_{(19,319)} = 0.524, p = 0.951$; group/trial interaction, $F_{(19,319)} = 0.498, p = 0.963$
Maze center	Within-subject effects: main effect day, $F_{(1,4,954)} = 2.631, p = 0.090$; interactions: day/group, $F_{(1,4,954)} = 3.610, p = 0.042$; day/trial, $F_{(27,5,954)} = 0.668, p = 0.901$; or day/group/trial, $F_{(27,5,954)} = 0.899, p = 0.615$; between-subjects effects: main effect group, $F_{(2,337)} = 14.269, p < 0.001$	Between-subjects effects: main effect group, $F_{(1,320)} = 2.974, p = 0.086$; main effect trial, $F_{(19,320)} = 0.620, p = 0.891$; group/trial interaction, $F_{(19,320)} = 0.780, p = 0.731$	Between-subjects effects: main effect group, $F_{(1,320)} = 1.314, p = 0.253$; main effect trial, $F_{(19,320)} = 1.117, p = 0.332$; group/trial interaction, $F_{(19,320)} = 0.739, p = 0.778$	Between-subjects effects: main effect group, $F_{(1,319)} = 3.719, p = 0.055$; main effect trial, $F_{(19,319)} = 1.119, p = 0.330$; group/trial interaction, $F_{(19,319)} = 0.999, p = 0.462$	Between-subjects effects: main effect group, $F_{(1,320)} = 1.744, p = 0.188$; main effect trial, $F_{(19,320)} = 0.809, p = 0.696$; group/trial interaction, $F_{(19,320)} = 1.025, p = 0.431$
End arm	Within-subjects effects: main effect day, $F_{(1,1,954)} = 1.194, p = 0.279$; interactions: day/group, $F_{(1,1,954)} = 1.970, p = 0.160$; day/trial, $F_{(20,5,954)} = 0.705, p = 0.843$; or day/group/trial, $F_{(20,5,954)} = 0.737, p = 0.790$; between-subjects effects: main effect group, $F_{(1,318)} = 0.326, p = 0.569$	Between-subjects effects: main effect group, $F_{(1,320)} = 1.474, p = 0.226$; main effect trial, $F_{(19,320)} = 0.768, p = 0.745$; group/trial interaction, $F_{(19,320)} = 0.765, p = 0.749$	Between-subjects effects: main effect group, $F_{(1,320)} = 3.815, p = 0.052$; main effect trial, $F_{(19,320)} = 1.163, p = 0.288$; group/trial interaction, $F_{(19,320)} = 0.477, p = 0.970$	Between-subjects effects: main effect group, $F_{(1,319)} = 6.223, p = 0.013$; main effect trial, $F_{(19,319)} = 1.242, p = 0.221$; group/trial interaction, $F_{(19,319)} = 0.855, p = 0.640$	Between-subjects effects: main effect group, $F_{(1,319)} = 0.249, p = 0.618$; main effect trial, $F_{(19,319)} = 1.232, p = 0.229$; group/trial interaction, $F_{(19,319)} = 0.394, p = 0.990$

Repeated-measures ANOVA conducted for each maze segment by day, with trial and group (i.e., virus) as factors. Tukey *post hoc* corrections were made for main effects in subsequent columns (i.e., baseline day, etc.). Significant comparisons are in bold.

For performance accuracy, between predevaluation and postdevaluation extinction days, there was no main within-subject effect of day ($F_{(1,214)} = 2.135, p = 0.145$), day/group interaction ($F_{(1,214)} = 0.169, p = 0.682$), or between-groups effect ($F_{(1,214)} = 0.464, p = 0.497$). On the reacquisition day, there was also no effect of group ($F_{(1,270)} = 0.035, p = 0.852$). Intriguingly, rats also continued to run accurately in the reacquisition test that followed. For run quitting, between predevaluation and postdevaluation extinction days, there was no main within-subject effect of day ($F_{(1,30)} = 0.108, p = 0.746$), day/group interaction ($F_{(1,30)} = 0.034, p = 0.854$), or between-groups effect ($F_{(1,30)} = 0.460, p = 0.504$); or between-groups effect for reacquisition ($F_{(1,16)} = 0.450, p = 0.513$). This level of persistence in running could reflect an incentive attraction to the maze cue itself rather than the paired outcome (Morrison et al., 2015; Smedley and Smith, 2018).

However, there were still improvement effects on the vigor of runs generated by DLS stimulation in our devaluation procedure (Fig. 7E,F). We analyzed run duration for our predevaluation and postdevaluation extinction probe days and discovered no main within-subject effects for day ($F_{(1,214)} =$

$0.514, p = 0.474$) or interactions of day/trial ($F_{(19,214)} = 1.571, p = 0.065$) and day/group/trial ($F_{(19,214)} = 0.990, p = 0.475$). There was a day/group interaction ($F_{(1,214)} = 5.644, p = 0.018$), and there was a main between-groups effect for the predevaluation extinction ($F_{(1,297)} = 15.331, p < 0.001$) and postdevaluation extinction ($F_{(1,215)} = 7.671, p = 0.006$) days. In the predevaluation and postdevaluation extinction probe tests, we found run-start stimulation of the DLS reduced run duration compared with controls (Fig. 7E,F). In addition, we analyzed the single postdevaluation reacquisition day alone and found a similar between-groups difference ($F_{(1,270)} = 4.455, p = 0.036$).

An identical analysis was performed for deliberations yielding a similar trend (Fig. 7G), in which there was no main within-subject effect of day ($F_{(1,214)} = 0.314, p = 0.576$) or interactions of day/group ($F_{(1,214)} = 0.183, p = 0.669$), day/trial ($F_{(19,214)} = 0.672, p = 0.844$), or day/group/trial ($F_{(19,214)} = 0.931, p = 0.545$). However, there was a strong main between-groups effect for the predevaluation extinction ($F_{(1,297)} = 11.142, p < 0.001$) and postdevaluation extinction ($F_{(1,215)} = 16.901, p < 0.001$)

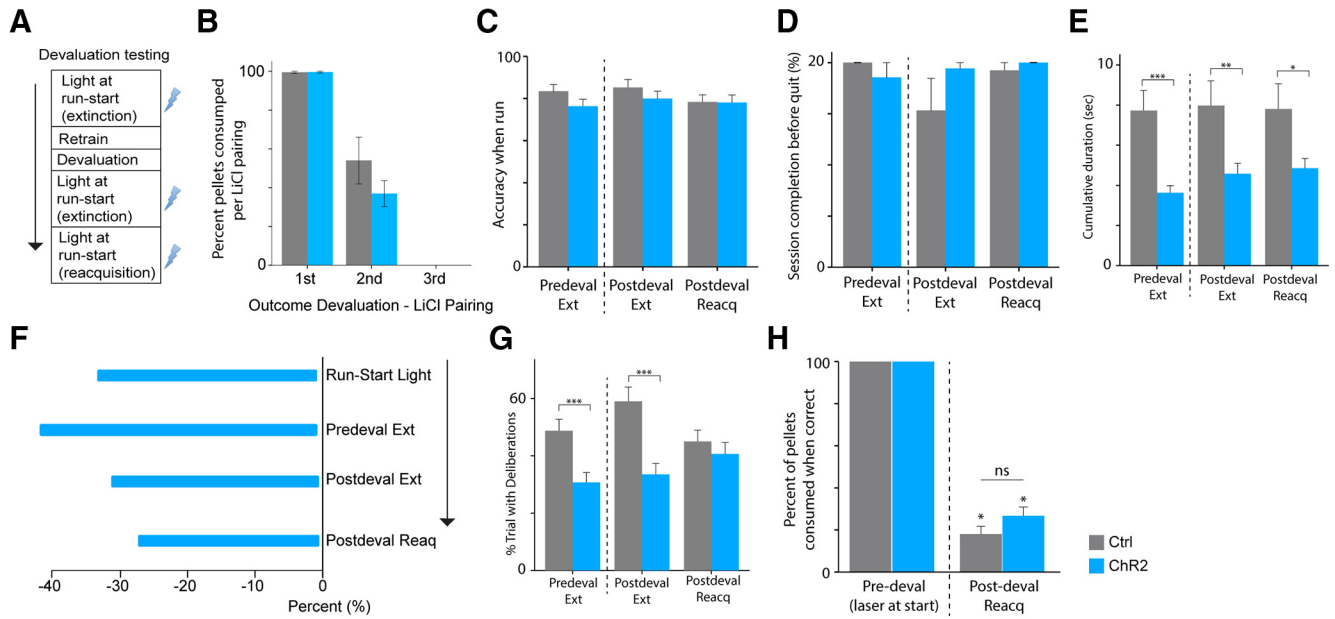


Figure 7. Beacon task related to outcome devaluation. **A**, Devaluation testing timeline, as in Figure 4. **B**, Pellet consumption during each pellet/LiCl pairing day to devalue rat groups. Consumption fell to zero in both groups. **C**, Performance accuracy when rats ran, which was high and stable, suggesting habitual performance by both groups. Blue, Chr2; gray, controls. **D**, Percentage of trials completed before quitting. Both groups ran nearly all trial opportunities. **E**, Average cumulative duration of trial runs (seconds). DLS stimulation reduced it on each day. **F**, Percentage change in cumulative duration in animals with Chr2-mediated stimulation relative to controls (blue). **G**, Percentage of trials with deliberation, which DLS stimulation reduced. **H**, Percentage of pellets consumed on task (before quitting a run), which was reduced equivalently across groups from predevaluation to postdevaluation. Asterisks denote significant *post hoc* comparisons (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). ns, Not significantly different.

Table 4. Statistics for beacon task with devaluation

Maze segment	Main comparisons	Predevaluation extinction	Postdevaluation extinction	Predevaluation reacquisition
Latency to enter maze center	Within-subject effects: main effect day, $F_{(1,214)} = 3.466, p = 0.064$; interactions: day/group, $F_{(1,214)} = 9.102, p = 0.003$; day/trial, $F_{(19,214)} = 1.315, p = 0.176$; or day/group/trial, $F_{(19,214)} = 1.239, p = 0.228$; between-subjects effects: main effect group, $F_{(1,214)} = 27.996, p < 0.001$	Between-subjects effects: main effect group, $F_{(1,297)} = 24.951, p < 0.001$; main effect trial, $F_{(19,297)} = 0.393, p = 0.990$; group/trial interaction, $F_{(19,297)} = 0.296, p = 0.999$	Between-subjects effects: main effect group, $F_{(1,215)} = 12.873, p < 0.001$; main effect trial, $F_{(19,215)} = 0.601, p = 0.903$; group/trial interaction, $F_{(19,215)} = 0.714, p = 0.802$	Between-subjects effects: main effect group, $F_{(1,270)} = 11.417, p < 0.001$; main effect trial, $F_{(19,270)} = 0.448, p = 0.979$; group/trial interaction, $F_{(19,270)} = 0.299, p = 0.998$
Maze center	Within-subject effects: main effect day, $F_{(1,214)} = 8.022, p = 0.005$; interactions: day/group, $F_{(1,214)} = 0.010, p = 0.921$; day/trial, $F_{(19,214)} = 1.088, p = 0.365$; or day/group/trial, $F_{(19,214)} = 0.686, p = 0.831$; between-subjects effects: main effect group, $F_{(1,214)} = 0.072, p = 0.789$	Between-subjects effects: main effect group, $F_{(1,297)} = 0.156, p = 0.693$; main effect trial, $F_{(19,297)} = 0.761, p = 0.752$; group/trial interaction, $F_{(19,297)} = 0.632, p = 0.881$	Between-subjects effects: main effect group, $F_{(1,215)} = 0.053, p = 0.819$; main effect trial, $F_{(19,215)} = 1.165, p = 0.290$; group/trial interaction, $F_{(19,215)} = 0.557, p = 0.933$	Between-subjects effects: main effect group, $F_{(1,270)} = 0.392, p = 0.532$; main effect trial, $F_{(19,270)} = 0.969, p = 0.498$; group/trial interaction, $F_{(19,270)} = 0.744, p = 0.772$
End arm	Within-subjects effects: main effect day, $F_{(1,214)} = 0.789, p = 0.375$; interactions: day/group, $F_{(1,214)} = 0.174, p = 0.677$; day/trial, $F_{(19,214)} = 1.440, p = 0.110$; or day/group/trial, $F_{(19,214)} = 1.578, p = 0.064$; between-subjects effects: main effect group, $F_{(1,214)} = 0.936, p = 0.334$	Between-subjects effects: main effect group, $F_{(1,297)} = 0.288, p = 0.592$; main effect trial, $F_{(19,297)} = 0.911, p = 0.569$; group/trial interaction, $F_{(19,297)} = 1.005, p = 0.454$	Between-subjects effects: main effect group, $F_{(1,215)} = 0.716, p = 0.398$; main effect trial, $F_{(19,215)} = 1.011, p = 0.450$; group/trial interaction, $F_{(19,215)} = 0.740, p = 0.775$	Between-subjects effects: main effect group, $F_{(1,269)} = 0.515, p = 0.473$; main effect trial, $F_{(19,269)} = 0.928, p = 0.549$; group/trial interaction, $F_{(19,269)} = 0.629, p = 0.883$

Repeated-measures ANOVA was conducted for each maze segment by day, with trial and group (i.e., virus) as factors in comparing predevaluation versus postdevaluation extinction days. Tukey *post hoc* corrections were made for main effects in subsequent columns (i.e., predevaluation and postdevaluation extinction). Predevaluation reacquisition day was analyzed separately in the last column. Significant comparisons are in bold.

illumination days. There was no between-groups effect during postdevaluation reacquisition ($F_{(1,270)} = 0.785, p = 0.376$). These behavior changes were not limited to particular maze-running segments (Table 4).

The consumption/rejection of the devalued pellets themselves was unaffected by DLS stimulations when comparing consumption during postdevaluation reacquisition with the normal prior predevaluation test day (Fig. 7H). There were effects of day

($F_{(1,627)} = 1726.655, p < 0.001$) and interactions of day/trial ($F_{(19,627)} = 2.875, p < 0.001$) and day/group/trial ($F_{(19,627)} = 1.829, p = 0.017$); however, there was no day/group interaction ($F_{(1,356)} = 2.667, p = 0.103$) and no main effect of group on the predevaluation day ($F_{(1,337)} = 0$) or reacquisition day ($F_{(1,290)} = 1.944, p = 0.164$).

Discussion

Whether a learned behavior occurs as a habit or as a purposeful goal-directed action is thought to be dictated by the level of participation among competing corticobasal ganglia neural systems that function to promote one or the other strategy (Killcross and Coutureau, 2003; Yin and Knowlton, 2006; Balleine et al., 2009; Packard, 2009). The habit system, involving the DLS, is linked in one line of research to promoting behaviors that are carried out reflexively and independently from the value of their outcome, and in another line of research to promoting behaviors that are vigorous and fluid. Our results help show that there is a clear link between the vigor of a behavior and how outcome sensitive it is, and that phasic DLS activity controls both in a powerful manner (Graybiel, 2008; Smith and Graybiel, 2013b, 2014). We specifically show that phasic DLS signaling at the time when behaviors are selected or initiated promotes behavioral vigor (i.e., lower performance duration time and reduced deliberative head movements). This is true of response-based maze-running behavior in which the response can be fully selected from the outset, as well as a cue-based maze-running task in which a response must be selected at two points during the maze run. We also show that the same phasic signaling is critical for controlling how habitual (i.e., outcome insensitive) behaviors are in these tasks. These results show that previously known phasic DLS activity during the start of a behavior, seen in many species and tasks (Smith and Graybiel, 2016), can directly control how vigorous and habitual those behaviors turn out to be. The fact that habits can be readily controlled by the brain during such a short, precise time window could be used to an advantage in designing intervention strategies for humans with otherwise treatment-resistant compulsive behaviors. Moreover, by showing how typically different measures for habits—maze-running vigor and outcome insensitivity—relate to one another in the brain, these results could help inform future scientific efforts to study habits at a more holistic level than is typical. Few studies have compared outcome relatedness of behavior with maze-running measures, but the results here support a general view that maze runs [whether purely egocentric (response task) or cue-driven (beacon task)] can be quite habitual in the devaluation-insensitivity sense (Sage and Knowlton, 2000; De Leonibus et al., 2011; Smith and Graybiel, 2016; Kosaki et al., 2018).

Concerning forms of behavioral vigor, there are dissociations between response vigor, which is related to movement onset latency, and performance vigor, which is related instead to ongoing movements (Wang et al., 2013). Performance vigor has been linked to DLS function (Novak et al., 2002; Dudman and Krakauer, 2016). Consistent with this, our findings show that run duration and deliberations were affected by run-start manipulations of the DLS in the response task, and those effects were not tied specifically to run onset latency, suggesting a key role in dictating performance vigor rather than response vigor. However, our beacon maze results did link run onset latency to some DLS stimulation effects on overall running time, suggesting some potential role for the DLS in response vigor as well. It will be worthwhile to conduct a comparative study on the mesolimbic system, which has also been thought to control effortful and vig-

orous behavior (Niv et al., 2007; Floresco, 2015; Bailey et al., 2016; Salamone et al., 2016; Berke, 2018).

Regardless, the changes in running vigor that resulted from DLS perturbations showed an intriguing correspondence to how outcome insensitive (i.e., habitual) those runs were. Results from both the response task and the beacon task showed that, to our surprise, animals were normally quite habitual despite not having received a large amount of training. Thus, further DLS stimulation hit a “ceiling” for increasing habits above normal. Still, DLS stimulation could produce improvements in running vigor in both tasks above and beyond the outcome independence of the behavior, highlighting the potential importance for looking at both measures when studying the brain basis of habits.

Dampening of DLS activity in the response task at the same run-start time point did not robustly reduce running vigor until food pellets were removed during the predevaluation extinction probe test or after outcome devaluation (although cycled DLS inhibition was more effective during reinforced sessions before devaluation). This result suggests that running vigor was rather impervious to DLS disruption at the onset of behavior but was highly sensitive to being reduced when there was a “reason” to reduce behavior based on the reduction of outcome values. In other words, the task periods in which continued running behavior was likely most DLS controlled (i.e., habitual despite lack of pellets available or devalued pellets) were those in which DLS inhibition produced the clearest effect on behavior. This might suggest that parallel circuits for goal-directed behavior could compensate for DLS inhibition at those times (i.e., when the food pellets were still valuable and present) but could not when the pellets were devalued or absent and habits were required to maintain performance. This set of DLS inhibition results is clearly reminiscent of the habit loss (i.e., outcome sensitivity gain) shown in prior DLS loss-of-function studies (Yin et al., 2004, 2006) and potentially pinpoints those results to compromised DLS signaling specifically at the beginning of behavior.

We also note that run accuracy was unaffected by any DLS perturbation. This lack of change in accuracy is curious given that versions of both of these maze tasks have been shown to be DLS dependent (Packard et al., 1989; McDonald and White, 1993; Sage and Knowlton, 2000; Chang and Gold, 2004; Palencia and Ragozzino, 2005; Berke et al., 2009; Kosaki et al., 2015). It is possible that either this phasic DLS signaling does not relate to the use of task rules to perform accurately or other neural systems were able to compensate for DLS activity perturbations to accomplish accurate runs.

Despite the reasonably close correspondence to DLS activity at behavior initiation, levels of running vigor, and levels of outcome insensitivity (i.e., habit), we raise one possibility that they may not always be aligned and are thus important to compare going forward. A prior study (Smith and Graybiel, 2013a) found that maze-running vigor increases during habit formation, just as outcome insensitivity develops and DLS activity becomes aligned to the beginning of the maze run. However, the magnitude of DLS activity could be compared at the single-run level with different aspects of the running behavior. It was found that the closest correlation of DLS activity was with running vigor, whereas measures of outcome insensitivity (i.e., whether animals approached or avoided a devalued goal) were not closely related to trial-level DLS activity. Although we did not vary DLS manipulations to give them on some trials but not others to truly compare vigor and habit measures, doing so could potentially reveal whether vigor/outcome insensitivity are always related (always co-occurring, reflecting a single underlying habit process) or can

be dissociated (not always co-occurring, reflecting dissociable processes). Similarly, given that these tasks left little room for increases in outcome insensitivity, as control animals ran habitually to devalued food pellets, further work will be needed to determine whether goal-directed actions can be made into habits by phasic DLS activity, and whether that corresponds to a joint increase in run vigor and outcome insensitivity.

Finally, given the nonspecific targeting of DLS neurons, it will be important to next resolve how different striatal cell types and basal ganglia pathways serve habits and vigor similarly or differently. For example, although studies have found that both direct-pathway and indirect-pathway neurons, and fast-firing interneurons, of the striatum are engaged at the beginning of a behavioral sequence (Kubota et al., 2009; Smith and Graybiel, 2013a; O'Hare et al., 2016; Vicente et al., 2016) and are involved in habits (Lovinger, 2010; Nelson and Killcross, 2013; Corbit et al., 2014; O'Hare et al., 2018), these studies and others also document dissociations among striatal cell populations in how they represent ongoing behavior, which raises the need to study striatal cells individually as well in this experimental context.

References

- Aldridge JW, Berridge KC, Rosen AR (2004) Basal ganglia neural mechanisms of natural movement sequences. *Can J Physiol Pharmacol* 82:732–739.
- Allan BD, Singer AC, Boyden ES (2015) Principles of designing interpretable optogenetic behavior experiments. *Learn Mem* 22:232–238.
- Amaya KA, Smith KS (2018) Neurobiology of habit formation. *Curr Opin Behav Sci* 20:145–152.
- Bailey MR, Simpson EH, Balsam PD (2016) Neural substrates underlying effort, time, and risk-based decision making in motivated behavior. *Neurobiol Learn Mem* 133:233–256.
- Balleine BW, Dickinson A (1998) Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* 37:407–419.
- Balleine BW, Liljeholm M, Ostlund SB (2009) The integrative function of the basal ganglia in instrumental conditioning. *Behav Brain Res* 199:43–52.
- Berke JD (2018) What does dopamine mean? *Nat Neurosci* 21:787–793.
- Berke JD, Breck JT, Eichenbaum H (2009) Striatal versus hippocampal representations during win-stay maze performance. *Journal of Neurophysiology* 101:1575–1587.
- Chang Q, Gold PE (2004) Inactivation of dorsolateral striatum impairs acquisition of response learning in cue-deficient, but not cue-available, conditions. *Behav Neurosci* 118:383–388.
- Chang SE, Smedley EB, Stansfield KJ, Stott JJ, Smith KS (2017) Optogenetic inhibition of ventral pallidum neurons impairs context-driven salt seeking. *J Neurosci* 37:5670–5680.
- Corbit LH, Nie H, Janak PH (2014) Habitual responding for alcohol depends upon both AMPA and D2 receptor signaling in the dorsolateral striatum. *Front Behav Neurosci* 8:301.
- da Silva JA, Tecuapetla F, Paixão V, Costa RM (2018) Dopamine neuron activity before action initiation gates and invigorates future movements. *Nature* 554:244–248.
- De Leonibus E, Costantini VJ, Massaro A, Mandolesi G, Vanni V, Luvisetto S, Pavone F, Oliverio A, Mele A (2011) Cognitive and neural determinants of response strategy in the dual-solution plus-maze task. *Learn Mem* 18:241–244.
- Desrochers TM, Amemori K, Graybiel AM (2015) Habit learning by naive macaques is marked by response sharpening of striatal neurons representing the cost and outcome of acquired action sequences. *Neuron* 87:853–868.
- Dickinson A (1985) Actions and habits: the development of behavioural autonomy. *Philos Trans R Soc Lond B Biol Sci* 308:67–78.
- Dudman JT, Krakauer JW (2016) The basal ganglia: from motor commands to the control of vigor. *Curr Opin Neurobiol* 37:158–166.
- Floresco SB (2015) The nucleus accumbens: an interface between cognition, emotion, and action. *Annu Rev Psychol* 66:25–52.
- Fujii N, Graybiel AM (2003) Representation of action sequence boundaries by macaque prefrontal cortical neurons. *Science* 301:1246–1249.
- Fujimoto H, Hasegawa T, Watanabe D (2011) Neural coding of syntactic structure in learned vocalizations in the songbird. *J Neurosci* 31:10023–10033.
- Gradinaru V, Mogri M, Thompson KR, Henderson JM, Deisseroth K (2009) Optical deconstruction of parkinsonian neural circuitry. *Science* 324:354–359.
- Graybiel AM (2008) Habits, rituals, and the evaluative brain. *Annu Rev Neurosci* 31:359–387.
- Haith AM, Reppert TR, Shadmehr R (2012) Evidence for hyperbolic temporal discounting of reward in control of movements. *J Neurosci* 32:11727–11736.
- Jin X, Costa RM (2010) Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature* 466:457–462.
- Jin X, Costa RM (2015) Shaping action sequences in basal ganglia circuits. *Curr Opin Neurobiol* 33:188–196.
- Jog MS, Kubota Y, Connolly CI, Hillegaart V, Graybiel AM (1999) Building neural representations of habits. *Science* 286:1745–1749.
- Killcross S, Coutureau E (2003) Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb Cortex* 13:400–408.
- Kosaki Y, Poulter SL, Austen JM, 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.
- Kosaki Y, Pearce JM, McGregor A (2018) The response strategy and the place strategy in a plus-maze have different sensitivities to devaluation of expected outcome. *Hippocampus* 28:484–496.
- Kubota Y, Liu J, Hu D, DeCoteau WE, Eden UT, Smith AC, Graybiel AM (2009) Stable encoding of task structure coexists with flexible coding of task events in sensorimotor striatum. *J Neurophysiol* 102:2142–2160.
- Lovinger DM (2010) Neurotransmitter roles in synaptic modulation, plasticity and learning in the dorsal striatum. *Neuropharmacology* 58:951–961.
- McDonald RJ, White NM (1993) A triple dissociation of memory systems: hippocampus, amygdala, and dorsal striatum. *Behav Neurosci* 107:3–22.
- Morrison SE, Bamkole MA, Nicola SM (2015) Sign tracking, but not goal tracking, is resistant to outcome devaluation. *Front Neurosci* 9:468.
- Nelson AJ, Killcross S (2013) Accelerated habit formation following amphetamine exposure is reversed by D1, but enhanced by D2, receptor antagonists. *Front Neurosci* 7:76.
- Niv Y, Daw ND, Joel D, Dayan P (2007) Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology (Berl)* 191:507–520.
- Novak KE, Miller LE, Houk JC (2002) The use of overlapping submovements in the control of rapid hand movements. *Exp Brain Res* 144:351–364.
- O'Hare JK, Ade KK, Sukharnikova T, Van Hooser SD, Palmeri ML, Yin HH, Calakos N (2016) Pathway-specific striatal substrates for habitual behavior. *Neuron* 89:472–479.
- O'Hare J, Calakos N, Yin HH (2018) Recent insights into corticostriatal circuit mechanisms underlying habits: invited review for current opinions in behavioral sciences. *Curr Opin Behav Sci* 20:40–46.
- Packard MG (2009) Exhumed from thought: basal ganglia and response learning in the plus-maze. *Behav Brain Res* 199:24–31.
- Packard MG, McGaugh JL (1996) Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. *Neurobiol Learn Mem* 65:65–72.
- Packard MG, Hirsh R, White NM (1989) Differential effects of fornix and caudate nucleus lesions on two radial maze tasks: evidence for multiple memory systems. *J Neurosci* 9:1465–1472.
- Palencia CA, Ragozzino ME (2005) The contribution of NMDA receptors in the dorsolateral striatum to egocentric response learning. *Behav Neurosci* 119:953–960.
- Sage JR, Knowlton BJ (2000) Effects of US devaluation on win-stay and win-shift radial maze performance in rats. *Behav Neurosci* 114:295–306.
- Salamone JD, Pardo M, Yohn SE, López-Cruz L, SanMiguel N, Correa M (2016) Mesolimbic dopamine and the regulation of motivated behavior. *Curr Top Behav Neurosci* 27:231–257.
- Seger CA, Spiering BJ (2011) A critical review of habit learning and the basal ganglia. *Front Syst Neurosci* 5:66.
- Smedley EB, Smith KS (2018) Evidence for a shared representation of sequential cues that engage sign-tracking. *Behav Processes* 157:489–494.

- Smith KS, Graybiel AM (2013a) A dual operator view of habitual behavior reflecting cortical and striatal dynamics. *Neuron* 79:361–374.
- Smith KS, Graybiel AM (2013b) Using optogenetics to study habits. *Brain Res* 1511:102–114.
- Smith KS, Graybiel AM (2014) Investigating habits: strategies, technologies and models. *Front Behav Neurosci* 8:39.
- Smith KS, Graybiel AM (2016) Habit formation. *Dialogues Clin Neurosci* 18:33–43.
- Smith KS, Virkud A, Deisseroth K, Graybiel AM (2012) Reversible online control of habitual behavior by optogenetic perturbation of medial prefrontal cortex. *Proc Natl Acad Sci U S A* 109:18932–18937.
- Thorn CA, Atallah H, Howe M, Graybiel AM (2010) Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. *Neuron* 66:781–795.
- Vicente AM, Galvão-Ferreira P, Tecuapetla F, Costa RM (2016) Direct and indirect dorsolateral striatum pathways reinforce different action strategies. *Curr Biol* 26:R267–R269.
- Wang AY, Miura K, Uchida N (2013) The dorsomedial striatum encodes net expected return, critical for energizing performance vigor. *Nat Neurosci* 16:639–647.
- Yin HH, Knowlton BJ (2006) The role of the basal ganglia in habit formation. *Nat Rev Neurosci* 7:464–476.
- Yin HH, Knowlton BJ, Balleine BW (2004) Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur J Neurosci* 19:181–189.
- Yin HH, Knowlton BJ, Balleine BW (2006) Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. *Behav Brain Res* 166:189–196.