

CALL FOR PAPERS | *Decision Making: Neural Mechanisms*

Hippocampus and subregions of the dorsal striatum respond differently to a behavioral strategy change on a spatial navigation task

Paul S. Regier,^{1*} Seiichiro Amemiya,^{2*} and A. David Redish²

¹Graduate Program in Neuroscience, University of Minnesota, Minneapolis, Minnesota; and ²Department of Neuroscience, University of Minnesota, Minneapolis, Minnesota

Submitted 26 February 2015; accepted in final form 12 June 2015

Regier PS, Amemiya S, Redish AD. Hippocampus and subregions of the dorsal striatum respond differently to a behavioral strategy change on a spatial navigation task. *J Neurophysiol* 114: 1399–1416, 2015. First published June 18, 2015; doi:10.1152/jn.00189.2015.— Goal-directed and habit-based behaviors are driven by multiple but dissociable decision making systems involving several different brain areas, including the hippocampus and dorsal striatum. On repetitive tasks, behavior transitions from goal directed to habit based with experience. Hippocampus has been implicated in initial learning and dorsal striatum in automating behavior, but recent studies suggest that subregions within the dorsal striatum have distinct roles in mediating habit-based and goal-directed behavior. We compared neural activity in the CA1 region of hippocampus with anterior dorsolateral and posterior dorsomedial striatum in rats on a spatial choice task, in which subjects experienced reward delivery changes that forced them to adjust their behavioral strategy. Our results confirm the importance of the hippocampus in evaluating predictive steps during goal-directed behavior, while separate circuits in the basal ganglia integrated relevant information during automation of actions and recognized when new behaviors were needed to continue obtaining rewards.

hippocampus; navigation; neural ensemble; striatum; tetraode

THE PROCESS OF LEARNING and automating actions is driven by multiple but dissociable neural circuits that instantiate different decision making systems (Balleine et al. 2007; Everitt and Robbins 2013; Graybiel 2008; Hikosaka et al. 1995; Jog et al. 1999; Johnson et al. 2007; van der Meer et al. 2012; Miyachi et al. 1997, 2002; O’Keefe and Nadel 1978; Packard and McGaugh 1996; Redish 1999, 2013; Yin and Knowlton 2004). Initial learning and adaptation to changes in the environment are driven by goal-directed systems, in which an organism engages in evaluative and predictive steps, integrating past experience and potential future outcomes (Balleine et al. 2007; Buckner and Carroll 2007; Killcross and Coutureau 2003; van der Meer et al. 2012; Redish 2013). Thus goal-directed behavior is cognitively intensive but flexible, since planning for multiple forthcoming options occurs at or before the time of action selection. Automated behavior is driven by habit-based systems and develops with increasing experience on a task, wherein specific situations trigger specific action chains (Jog et al. 1999; Packard and McGaugh 1996; Smith and Graybiel

2013; Yin and Knowlton 2004, 2006) Because future outcomes are not considered at the time of action selection in the habit system, situations release actions quickly, but once these associations are well established they are difficult to change (e.g., insensitivity to devaluation; Adams 1982; Adams and Dickinson 1981).

Several studies have reported the hippocampus (HC) as a structure important for goal-directed behavior (Johnson and Redish 2007; Maguire and Hassabis 2011; O’Keefe and Nadel 1978; Schacter et al. 2011). HC neurons that are spatially tuned form a cognitive map, allowing for integration of past and potential future experiences in order to plan behavior (O’Keefe and Nadel 1978; McNaughton et al. 2006; Redish 1999; Wikenheiser and Redish 2015a). Neurons in the dorsolateral striatum also respond to spatial cues (Mizumori et al. 2004; Schmitzer-Torbert and Redish 2004, 2008; Yeshenko et al. 2004), but only when spatial cues contain information about obtaining rewards (Berke et al. 2009; Schmitzer-Torbert and Redish 2008). Neural activity in the dorsolateral striatal neurons is related to specific motor movements and actions (Alexander and DeLong 1985a, 1985b; Carelli and West 1991; Cho and West 1997; Jog et al. 1999; Schmitzer-Torbert and Redish 2008), likely ones that have consistently led to reinforcement.

Recent studies have discovered anatomical (Berendse et al. 1992; McGeorge and Faull 1989; Swanson 2000) and functional (Devan et al. 1999; Yin et al. 2004, 2005a; Yin and Knowlton 2004) differences between dorsolateral and dorsomedial striatum. Anterior dorsolateral striatum (aDLS) receives input from motor and sensory areas (Alexander and Crutcher 1990; Berendse et al. 1992; McGeorge and Faull 1989) and regulates motor control and habit-based behaviors (Carelli and West 1991; Cho and West 1997; Hikosaka et al. 1995; Miyachi et al. 1997; Smith and Graybiel 2013). Dorsomedial striatum has been implicated as playing a role in goal-directed behavior (Devan et al. 1999; Gremel and Costa 2013; Yin et al. 2005b; Yin and Knowlton 2004), such as reversal learning (Castañé et al. 2010; Kirkby 1969; Ragozzino 2007; Ragozzino and Choi 2004) and changing strategies (Ragozzino 2007; Ragozzino et al. 2002a, 2002b).

Importantly, recent studies have found anatomical (Berendse et al. 1992; McGeorge and Faull 1989) and functional (Corbit and Janak 2010; Yin et al. 2005b; Yin and Knowlton 2004) differences between anterior (aDMS) and posterior (pDMS)

* P. S. Regier and S. Amemiya contributed equally to this work.

Address for reprint requests and other correspondence: A. D. Redish, Dept. of Neuroscience, 6-145 Jackson Hall, 321 Church St. SE, Univ. of Minnesota, Minneapolis MN 55455 (e-mail: redish@umn.edu).

dorsomedial striatum and, consequently, in their role in behavior. While aDMS receives input from anterior cingulate cortex, dorsal prelimbic area, and some motor/sensory areas, pDMS receives input from the orbitofrontal cortex, ventral prelimbic area, and entorhinal cortex. The aDMS has been postulated to be involved with certain goal-directed behaviors (Clarke et al. 2008; Devan et al. 1999; Ragozzino et al. 2002a, 2002b), but neural correlates in the aDMS of these behaviors have not been found (Kimchi and Laubach 2009; Thorn et al. 2010). The anatomical inputs to pDMS are from structures involved in reversal learning, strategy changing, and action-outcome learning. Interestingly, studies have implicated the pDMS in many of these specific types of learning (Lex and Hauber 2010a, 2010b; Lucantonio et al. 2014; Stalnaker et al. 2012; Yin et al. 2005a, 2005b; Yin and Knowlton 2004).

We therefore hypothesized that just as the aDLS integrates information from sensorimotor areas to translate this information into behavior, the pDMS likely plays more of a role in goal-directed behavior than aDMS, integrating information from goal-oriented cortical areas and translating this information into action. Neural representational comparisons have already been made between aDLS and aDMS (Thorn et al. 2010). However, although the lesion data suggest a stronger role of pDMS than aDMS in these types of goal-oriented learning (Yin et al. 2005b; Yin and Knowlton 2004), to date no one has directly compared neural ensemble recordings from the pDMS. In this article, we report results from simultaneous recordings of aDLS and pDMS.

When rats come to decision points, they sometimes pause, orient toward a goal, and then reorient back and forth. This behavioral phenomenon is termed vicarious trial-and-error (VTE) behavior and has been hypothesized to reflect an underlying search process (Johnson and Redish 2007; van der Meer et al. 2012; Muenzinger 1938; Tolman 1938). VTE primarily occurs during goal-directed behaviors (Gardner et al. 2013; Papale et al. 2012; Schmidt et al. 2013; Tolman 1938). Changes to the reward contingencies within an environment consistently produce an increase in the occurrence of VTE (Blumenthal et al. 2011; Powell and Redish 2014; Schmidt et al. 2013), likely reflecting deliberative behavior as subjects form new or different strategies. Deliberation entails the search and evaluation of potential possibilities (Buckner and Carroll 2007; Daw et al. 2005; Johnson and Redish 2007; Redish 2013). During VTE behaviors, HC neural ensembles sweep forward ahead of the animal toward the potential goals (Gupta et al. 2012; Johnson and Redish 2007). If these sweeps of spatial representation are reflective of the search and evaluation process, then they would not be expected to occur in structures involved in the habit-based components, such as aDLS. Previous experiments have found that aDLS representations do not show forward sweeps (van der Meer et al. 2010), but it remains unknown whether pDMS representations do.

In contrast, in stable environments, as actions become more automated (e.g., habit based), control shifts to sensorimotor circuits capable of encoding action chains (Dezfouli et al. 2014; Graybiel 2008; Yin and Knowlton 2006), which include the aDLS (Alexander and Crutcher 1990; Berendse et al. 1992; Carrelli and West 1991; McGeorge and Faull 1989). Graybiel and colleagues have reported that the development of these action chains on a cued T-maze aligns with the development of preferential firing in dorsolateral striatum at the beginning and

end of their T-maze (task bracketing; Barnes et al. 2005; Jog et al. 1999; Smith and Graybiel 2013; Thorn et al. 2010). Task bracketing is thought to underlie behavioral “chunking,” or the bracketing of a sequence of action chains on a task (Graybiel 1998; Jog et al. 1999; Miller 1956). Smith and Graybiel (2013) recently found that task bracketing was anticorrelated with VTE behaviors. If task bracketing is a consequence of the development of action chains within a habit-based (automated) behavioral decision system, then one would hypothesize that structures involved in the goal-directed (deliberative) system, such as HC and pDMS, should not show task bracketing.

Multiple systems interact to produce appropriate behavioral outputs, each of these systems forming neural circuits that run both in parallel and in conjunction with one another. To better understand how these decision making systems interact, we recorded neural ensembles from three different structures, simultaneously from pDMS and aDLS in six rats and from CA1 in another six rats, on a spatial navigation task that required rats to make decisions based on guidance from internal cues. In the analyzed probe trials, we introduced a change in the reward contingency without any physical change in the environment, forcing a change in behavior. On this task, rats begin each day showing goal-directed behaviors but develop an automated stereotypy of their path through the day (Schmitzer-Torbert and Redish 2004). On encountering this unsignaled change in reward contingency, rats typically return to goal-directed behaviors and reautomate their path under the new contingency (Blumenthal et al. 2011; Gupta et al. 2012; Powell and Redish 2014; Steiner and Redish 2012). This automation/reversal/reautomation allowed us to observe both goal-directed and habit-based behavior in a single session and to measure the neural correlates of these different types of behavior.

METHODS

Subjects

Eleven Fischer Brown Norway rats and one Brown Norway rat were trained to perform a modified version of a Hebb-Williams maze (HWM; Hebb and Williams 1946), similar to the multiple-T left, right, alternate (LRA) task (Blumenthal et al. 2011; Powell and Redish 2014; Steiner and Redish 2012). The maze was a wooden rectangle box with carpeted floor and LEGO brick walls that could be altered to change the internal maze portion (Fig. 1). The internal maze forms a series of low-cost choice points, which we refer to as the navigation sequence. At the end of the navigation sequence, rats came to a high-cost choice point and had to make a left or right turn. If a rat made the correct choice at the choice point, it would receive a food reward (2 unflavored food pellets, 45 mg each; Research Diets, New Brunswick, NJ) at a side feeder location and at a center feeder location (end zone). The pellets were delivered with automatic pellet dispensers (Med Associates, St. Albans, VT). If a rat made an incorrect choice at the choice point, it did not receive any food rewards and had to continue down the return arm to the end zone but would not receive food there. Although returning to the end zone started a new lap whether rewarded or not, rats ran the task continuously for 30 min. Lap identification was used for analysis only—no explicit event signaled the beginning or end of the lap.

Three different reward contingencies were used [left (L), right (R), or alternating (A)]. During training sessions, the reward contingency was held constant through an entire session but changed randomly from session to session. During the experimental phase, each session began with one reward contingency but the reward contingency changed at approximately the halfway point of the session (the reward

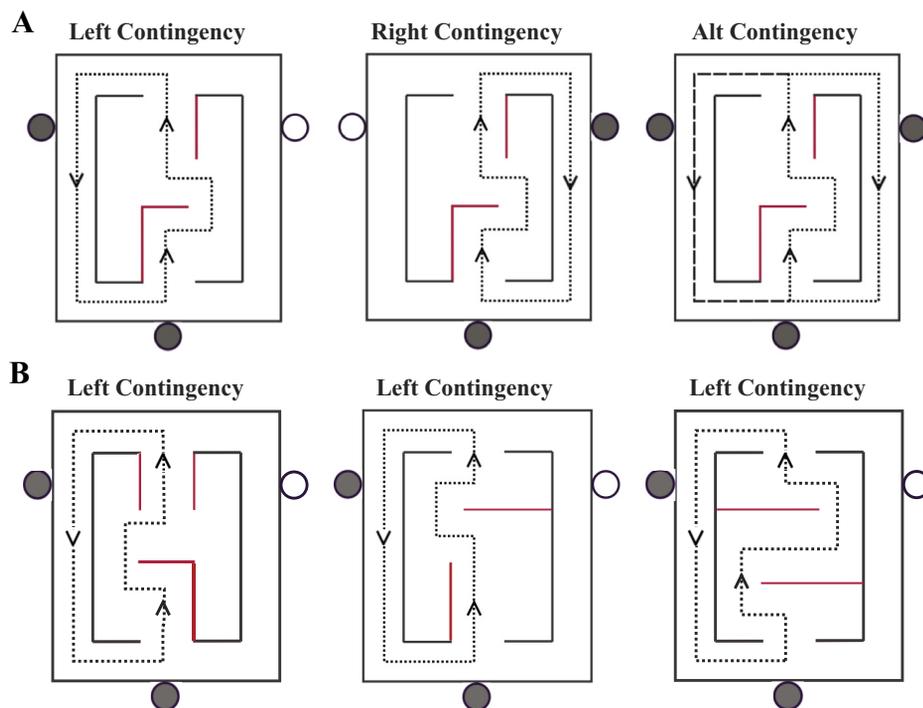


Fig. 1. *A*: 3 contingencies for reward were presented on the Hebb-Williams maze: left, right, and alternation. During training sessions, the contingency remained fixed for the entire session. During test sessions, the contingency was switched at approximately the halfway point of the sessions. Interior walls (labeled in red) changed daily. *B*: examples of different inner maze configurations for a left reward contingency.

contingency switch). Rats ran a subset of the six possible combinations (LR, LA, RL, RA, AL, AR) pseudorandomly. Every session lasted 30 min; rats earned their daily food intake on the task (~12 g/day).

All procedures were conducted in accordance with National Institutes of Health guidelines for animal care and approved by the Institutional Animal Care and Use Committee at the University of Minnesota. Care was taken to minimize the number of animals used in these experiments and to minimize suffering.

Surgery

After pretraining on the HWM, rats were chronically implanted with multitetrode hyperdrives [6 rats were implanted with 14-tetrode hyperdrives (made in house, 12 electrodes for recording, 2 for references) targeting the right dorsal HC, 3 rats were implanted with 14-tetrode hyperdrives targeting aDLS and pDMS unilaterally, and 3 rats were implanted with 28-tetrode hyperdrives (made in house, 24

electrodes for recording, 4 for references) targeting the aDLS and pDMS bilaterally]. See Table 1.

Nine rats were initially anesthetized with Nembutal (pentobarbital sodium, 40–50 mg/kg; Abbott Laboratories, North Chicago, IL), and three rats were anesthetized with isoflurane. All rats were maintained on isoflurane (0.5–2% isoflurane vaporized in medical-grade O₂) during the implantation. All rats were situated on a stereotaxic apparatus (Kopf) and received Dual-Cillin (Phoenix Pharmaceutical, St. Joseph, MO) intramuscularly in each hindlimb. The dorsal parts of the rats' heads were shaved and disinfected with alcohol (70% isopropyl) and Betadine (Purdue Rederick, Norwalk, CT), and the skin overlying the scalp was removed. Several jewelers' screws were used to anchor the hyperdrive to the skull, and one of the screws was used as a recording ground. In six rats one craniotomy was opened (HC, targeting CA1); in three rats two craniotomies were opened (unilateral implantation of aDLS and pDMS), and in three rats four craniotomies were opened (bilateral implantation of aDLS and

Table 1. Neuronal activity recordings

Rat ID No.	No. of Tetrodes per Drive (per rat)	Target	No. of Cells and Type	No. of Sessions	No. of Trials per Session
R249	12	CA1	CA1: 280 cells	3	118, 115, 135
R252			265 Pyramidal	3	35, 99, 93
R264			15 Interneurons	4	115, 108, 101, 85
R272				4	62, 58, 110, 82
R282				5	133, 146, 138, 126, 139
R284	12	Right-side aDLS + pDMS	aDLS: 258 cells	4	100, 106, 125, 57
R237			pDMS: 155 cells	6	76, 63, 58, 60, 87, 66
R239				6	88, 83, 74, 82, 82, 82
R247				6	76, 93, 79, 81, 91, 91
	24	Bilateral aDLS + pDMS	aDLS: 411 cells	6	86, 79, 85, 60, 89, 95
R259			pDMS: 218 cells	5	68, 104, 96, 97, 92
R269				6	77, 87, 75, 92, 82, 75
			96 HFN 22 TFN 16TFN		

aDLS, anterior dorsolateral striatum; pDMS, posterior dorsomedial striatum; PFN, phasic-firing neurons, HFN, high-firing neurons, TFN, tonic-firing neurons.

pDMS). Craniotomies were opened with a surgical trephine. The bundles for aDLS were centered at 0.7 mm anterior of bregma and 3.5 mm lateral of midline, and bundles for pDMS were centered at 0.4 mm posterior of bregma and 2.5 mm lateral of midline, in accordance with the study by Yin and Knowlton (2004). The bundles for HC were centered at 3.8 mm posterior of bregma and 3.0 mm lateral of midline.

The craniotomies around the hyperdrive were protected with Silastic (Dow Corning, Midland, MI). Dental acrylic (Perm Reline and Repair Resin, The Hygenic Corporation, Akron, OH) secured the hyperdrive to the skull. Immediately after surgery, all tetrodes were turned down 640 μm . After tetrodes were turned down, rats were given subcutaneous injections (5–10 ml) of sterile saline and oral administration of Tylenol (1 ml). To prevent infections, rats received subcutaneous injections of Baytril (enrofloxacin, 1.1 mg/kg) on the day of surgery and for 7 days after surgery.

Data Collection

After surgery, tetrodes were advanced 40–640 μm per day until reaching the striatum or HC. Initial entry into the HC or striatum was differentiated by observation of the corpus callosum, an area that is electrophysiologically quiet compared with the cortex, HC, and striatum. The HC pyramidal layer was identified by the size and reversal point of sharp-wave ripples, as well as by burst firing by cells in synchrony with the ripple portions of the sharp-wave ripple complexes. The striatum was further identified by the observation of medium spiny neurons, which have long interspike intervals and short bursts of firing.

In nine rats (3 striatum, 6 HC) recording neural activity while running a task was made possible by a 64-channel analog Cheetah system (Neuralynx, Bozeman, MT), and for the other three rats (striatum) a 96-channel digital Cheetah system was used. Spike trains were identified and recorded online with built-in filters, and then clustering of spike trains occurred offline. Neurons were separated into putative cells on the basis of specific waveform properties with KlustaKwik (K. D. Harris) and MClust 3.5 or MClust 4.0 (A. D. Redish). Only clusters with $L_{\text{ratio}} < 0.20$ and isolation distance > 20 were used for analysis.

The position of the rat was monitored with LEDs on the head stage during experimental recording sessions, captured by an overhead camera. Position of the rat was recorded at 60 Hz with a video input to the Cheetah recording system, time-stamping the sampled position of the LEDs. Control of the experiment was performed with in-house code written in MATLAB. Events (e.g., feeder click and food delivery) were recorded and time-stamped by the Cheetah recording system and by MATLAB.

Histology

After the experiment was completed, tetrode locations were marked with small lesions by passing a small amount of anodal current (5 μA for 10 s) through each tetrode. After at least 2 days had passed, rats were anesthetized and perfused transcardially with saline followed by 10% formalin. Brains were stored in formalin followed by 30% sucrose formalin until slicing. Coronal slices were made through the area of the implantation and stained with cresyl violet to visualize tetrode tracks. Locations from the three regions were confirmed histologically (Fig. 2).

Data Analysis

Cell-type classification. Striatal cells were classified into phasic-firing neurons, high-firing neurons, and tonic-firing neurons on the basis of the proportion of time spent in long (> 2 s) interspike intervals (Schmitzer-Torbert and Redish 2004, 2008). HC cells were classified into putative pyramidal neurons and putative interneurons with a threshold of 6 Hz over the recording session. HC cells firing at an

average of < 6 Hz were classified as putative pyramidal neurons, while HC cells firing at an average of > 6 Hz were classified as putative interneurons.

Choice point. The choice point was defined as the top zone of the maze (see Figs. 4 and 6), based on behavioral navigation of the subjects. We defined the choice point as the point where overall path of the animal diverged as rats turned either left or right. As shown in Fig. 4, this point occurred near the end of the navigation sequence.

Firing rate. Firing rate at specific points on the maze over laps was obtained similarly to Thorn et al. (2010). Eight events on the maze were identified [start of the navigation sequence, middle of the navigation sequence, choice point, feeder click, side feeder (enter and exit), return arm, center feeder click, and end zone (start/end of each lap)]. Firing rate was measured over a 2-s time window (± 1 s around each event). Firing rates were z -scored by taking the mean firing rate of each bin and then subtracting the mean firing rate for the rest of the maze divided by the standard deviation of firing for the rest of the maze. z -Scored firing rate was then divided into 500-ms time bins (4 bins for each event) within session over five-lap bins for all structures before and after the contingency switch. A measure of overall firing rate across laps was obtained by taking the mean z -scored firing rate for each five-lap bin, and a measure of overall firing rate across the maze was obtained by taking the mean z -scored firing rate for each 500-ms time bin.

Task-bracketing index. A measure of task bracket-like effects was calculated based on an analysis by Smith and Graybiel (2013), who took the mean firing rate at the start and end of the maze minus the mean firing rate at the auditory cue at the choice point, a measure they called the task-bracket index. Similarly, in the present study, a normalized task-bracketing index was calculated by taking the mean firing rate of the last two bins of the end zone epoch (which marked the end and beginning of each lap on our task) and then subtracting the mean firing rate from the rest of the maze and dividing by the standard deviation of the mean firing rate from the rest of the maze (z -scored the same as presented above). This was done for early laps (1–15) and late laps (16–30) before and after the switch. A two-way ANOVA was performed to test for main effects and interactions, and post hoc tests were performed when there was a significant main effect or interaction and corrected with Bonferroni-Holm techniques.

Tuning curves. A measure of tuning curves was obtained by taking neuronal firing at each time point and position of the rat for each neuron. Tuning curve information was normalized by occupancy to adjust for the amount of time the rat was at each position/time point. The maze was linearized with 1,000 points by creating an ideal path around the maze and then associating tuning curves with those points around the maze. This was done on a lap-by-lap basis.

Correlation of tuning curves. After tuning curve information was obtained for each neuron in each structure (pDMS, aDLS, and CA1) on individual laps, four distributions of correlation coefficients were obtained for each region. The four different distributions (for each condition) were obtained by taking an average of left laps before the switch vs. average of left laps after the switch, an average of right laps before the switch vs. average of right laps after the switch, an average of left vs. average of right laps before the switch, and an average of left vs. average of right laps after the switch. Averages were calculated across laps, and maze locations were preserved.

To determine whether there was a significant difference between regions, we conducted a bootstrap analysis to determine whether the mean correlations were significantly far apart from each other in Euclidean distance. The analysis was applied to each of the four conditions separately. The bootstrap was performed by resampling the total distribution of correlations (across the 2 regions being compared) and redividing the total distribution into sample sizes matching the real sample sizes of the two distributions. (Thus if there were 800 cells providing 800 correlations divided between 500 cells in aDLS and 300 in CA1, we would take the 800 correlations, redivide them into groups of 500 and 300, and then

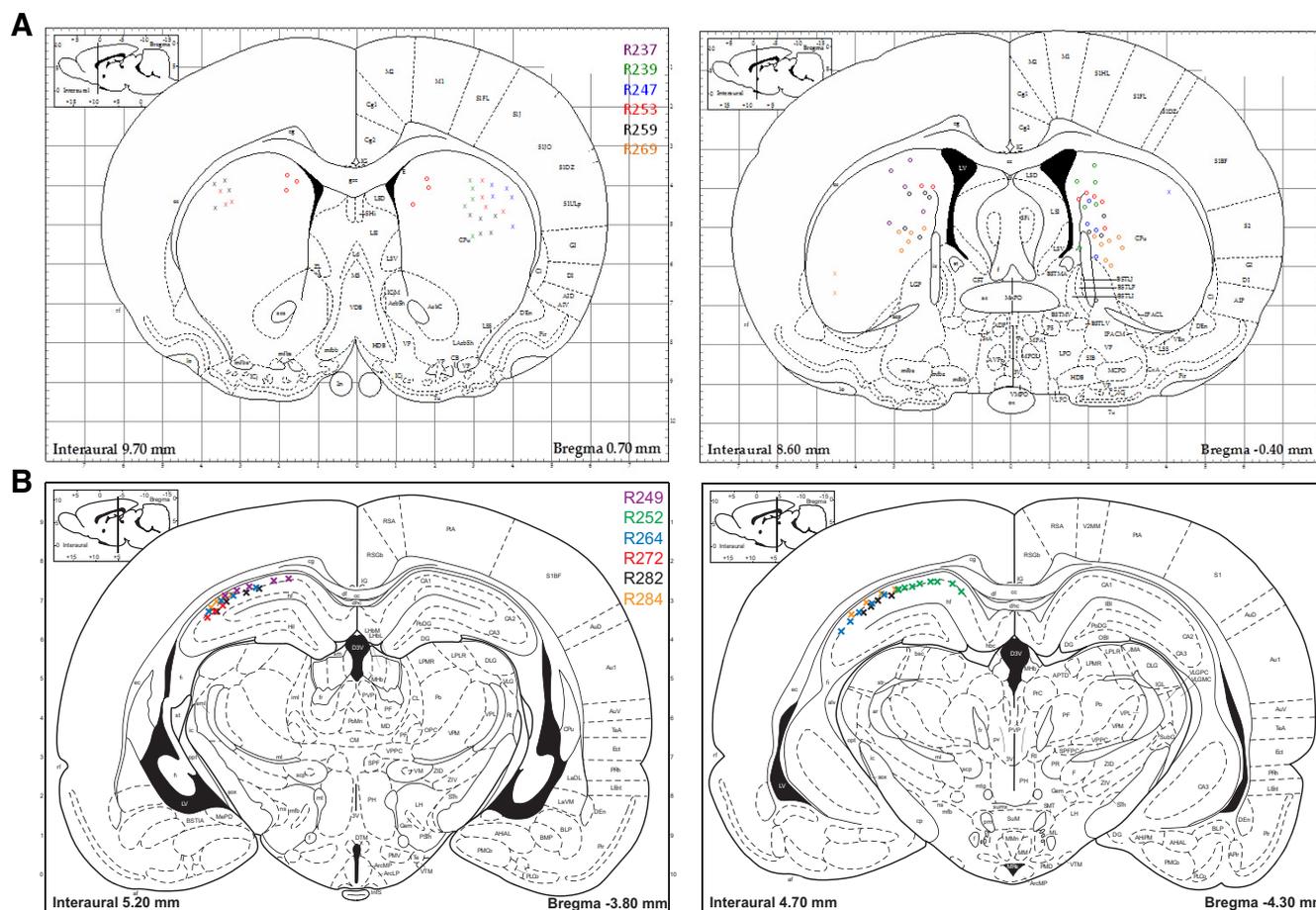


Fig. 2. Histology. *A*: final recording locations in anterior dorsal striatum (*left*) and posterior dorsal striatum (*right*). Dorsomedial striatal locations are marked with circles and dorsolateral striatal locations are marked with \times . The vast majority of dorsomedial striatal locations were confirmed to be in the posterior portion of the dorsomedial striatum. All but 1 dorsolateral striatal location was confirmed to be in the anterior portion of the dorsolateral striatum. *B*: final recording locations in the hippocampus. All hippocampus locations were confirmed to be in the layer of CA1. Colors indicate subject numbers.

find the Euclidean distance between the mean correlations.) One thousand bootstraps were done for each comparison. Finally, we compared each actual distance for each condition to the randomly sampled distribution created for each condition, and a *P* value was obtained by calculating the probability of random samples falling outside of actual mean distance.

RESULTS

A Change in Behavioral Strategy

We observed the behavior of 12 rats on a complex navigation task that required subjects to recognize a reward contingency and then to recognize and switch to a different reward contingency approximately midway through the session. Behavioral performance results from all 12 rats indicated that only a brief exploratory period was necessary in order to learn the starting contingency (Fig. 3).

When the novel contingency switch event was introduced performance dropped to the expected rate, given that they perseverated in the original contingency but then increased to an accuracy similar to that observed prior to the switch for the rest of the session (Fig. 3). VTE behavior appeared along with this sharp decrease in performance over the first 10 laps after the switch (Fig. 3). The occurrence of VTE behavior after the switch indicated that rats returned to flexible, deliberative

behaviors when they were forced to reevaluate their current strategy.

Neurophysiological Data Sets

We recorded neuronal activity from subregions of the dorsal striatum (simultaneous recording from aDLS and pDMS) and HC (CA1). In the dorsal striatum, we recorded a total of 1,027 neurons. The majority of neurons were putative medium spiny neurons [MSP, 831 (81%) of 1,027 total neurons], 158 (15%) were putative high-firing interneurons (HFN), and 38 (4%) were putative tonic-firing neurons (TFN). By region, we recorded 541 MSP, 96 HFN, and 22 TFN from aDLS and 290 MSP, 62 HFN, and 16 TFN from pDMS (Table 1). In CA1, the majority of neurons were putative pyramidal cells [265 (95%) of 280 total neurons; Table 1]. Recording locations from all three data sets were confirmed histologically (Fig. 2).

Spatial Decoding of Forward Location at Choice Point

During VTE events at the choice point, HC neural activity tends to represent paths ahead of the rat (Gupta et al. 2010; Johnson and Redish 2007) while dorsolateral striatum activity does not (van der Meer et al. 2010). Johnson and Redish (2007) did not find any reliable relationship between the chosen side and the neural activity during these VTE events, but other

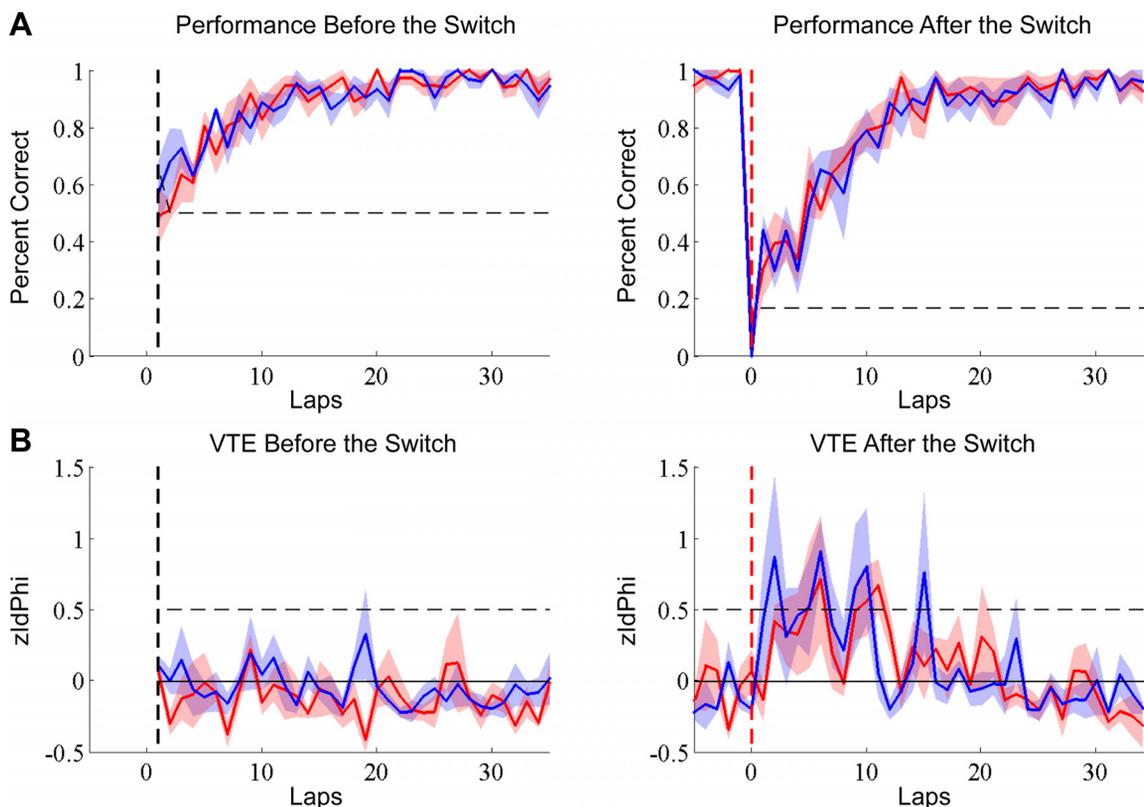


Fig. 3. Behavioral results on the Hebb-Williams maze. *A*: performance before the switch (*left*) was at chance initially but increased across sessions. Performance decreased sharply after the switch was introduced (*right*) but then increased again across laps. *B*: vicarious-trial-and-error (VTE) behavior was not significantly elevated before the switch (*left*), but after the switch was introduced VTE behavior significantly increased for approximately the first 10 laps (*right*). Striatal rats are shown in red and hippocampal rats in blue. VTE is measured as a z-scored integrated angular velocity (zIdPhi). Larger zIdPhi indicates more VTE. See Papale et al. (2012).

studies have found representations of paths to chosen goals during more automated behaviors (Gupta et al. 2010; Wikenheiser and Redish 2015b). To examine the forwardness of spatial representation to compare representations in aDLS, pDMS, and CA1, we used a Bayesian spatial decoding algorithm, which estimates the rat's location from ensemble spiking activity (Zhang et al. 1998).

We decoded position as animals proceeded through a VTE event and found that the HC representation swept ahead of the animal to the two options. Figure 4*B* shows average decoding across a single non-VTE and a single VTE lap. Decoding on both laps shows mostly local decoding with forward components, consistent with previous experiments (Gupta et al. 2012; Johnson and Redish 2007; Wikenheiser and Redish 2015b). Also consistent with those previous experiments, decoding on the non-VTE lap proceeded primarily toward the chosen side, representing the current goal of the animal, while decoding on the VTE lap included decoding to both sides, consistent with previous experiments (Johnson and Redish 2007). This decoding analysis was performed as an average over the entire pass because we were interested specifically in how much decoding went ahead of the animal to the unchosen side. Examination of single theta cycles found that the representation of each side occurred serially, consistent with previous experiments (Johnson and Redish 2007; data not shown).

To examine the forwardness of representation at the choice point in each structure, we calculated the sum of the decoding probability in forward paths from the choice point (Fig. 4). A

one-way ANOVA found a significant effect of regions [$F(3) = 13.525$, $P < 0.0001$]. Multiple comparisons with Bonferroni-corrected paired *t*-tests found higher forwardness in CA1 than in aDLS ($P < 0.0001$) and pDMS ($P < 0.0001$).

To address whether the forward representation reflected the succeeding choice, we calculated the difference of decoding probability between the chosen side and the unchosen side (chosen – unchosen). Because VTE has been found to anticorrelate with automation (Smith and Graybiel 2013), we separately examined the difference of the forward representation in VTE and non-VTE laps. One-sample *t*-tests revealed that CA1 represented the chosen side more than the unchosen side on non-VTE laps ($P < 0.05$), but this difference was not observed on VTE laps. Both striatal sets (aDLS and pDMS) showed higher forwardness on the chosen side more than on the unchosen side regardless of whether the lap included VTE or not ($P < 0.005$). The lack of representation of the chosen side in our HC recordings specifically during VTE laps is consistent with previous experiments (Johnson and Redish 2007).

Development of Ensemble Firing in Dorsolateral Striatum That Tracks Behavioral Performance

To examine whether neuronal ensembles from the different structures dynamically changed along with behavioral performance, we measured the average firing rate of each structure over laps. This measure of general neuronal activity revealed differences between the aDLS and the rest of the structures.

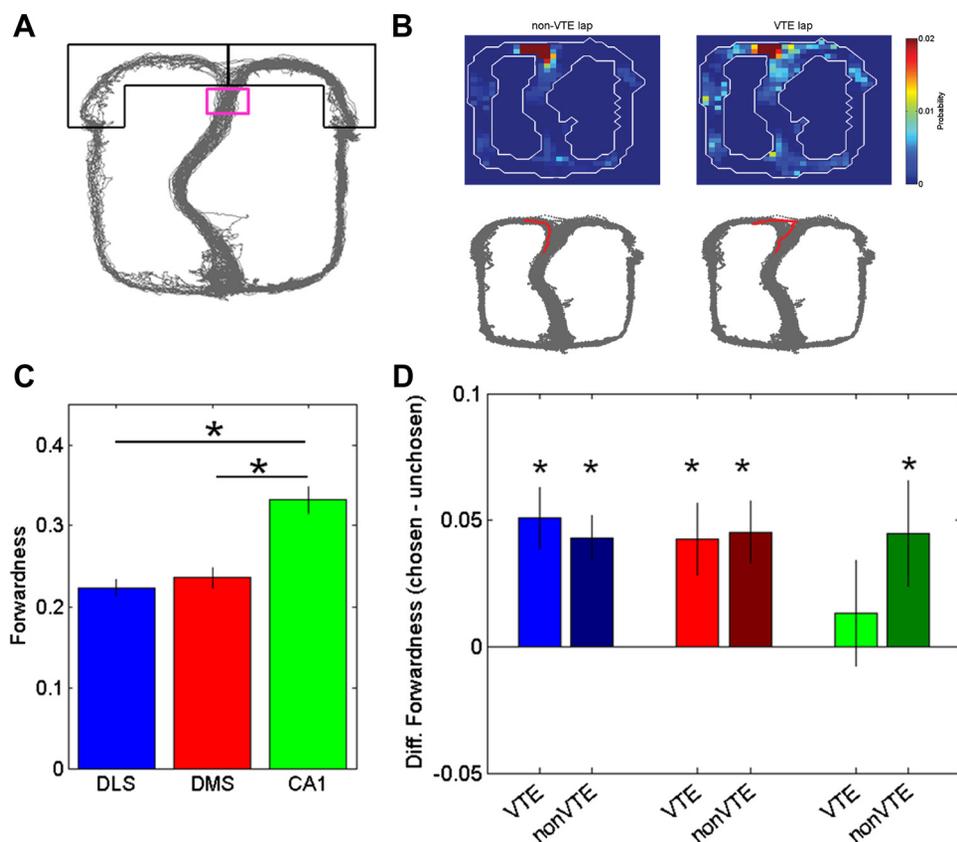


Fig. 4. Forward spatial decoding at choice point. *A*: illustration of choice point and decoded position. Gray lines represent the physical location of the rat. Pink rectangle shows choice point. Black rectangles show decoded area when rats were at choice point. *B*: examples of spatial decoding from hippocampal representations on a non-VTE and a VTE lap. Decoding was done at 125-ms time windows and then averaged over the entire choice point pass. On non-VTE laps decoding included forward representations to the chosen side only, while on VTE laps decoding included forward representations to both sides. *C*: amount of spatial representation forwardness in anterior dorsolateral striatum (aDLS), posterior dorsomedial striatum (pDMS), and CA1. $*P < 0.01$. *D*: difference of forwardness between chosen side and unchosen side on VTE and non-VTE laps. $*P < 0.05$.

Only in aDLS was there an increase in average firing rate before and after the switch (Fig. 5).

To obtain a more accurate measure of where this change was occurring, we adopted a method used by Graybiel and colleagues (Barnes et al. 2005; Jog et al. 1999; Smith and Graybiel 2013; Thorn et al. 2010), measuring firing rate at several maze locations (start zone, navigation sequence, choice point, feeder cue, feeders, return arm, and end zone) and obtained the firing rate ± 1 s around entry into each maze location. We then plotted firing rate around each maze location across laps before and after the switch (for example, see Fig. 6). In the aDLS, there was an increase in firing rate primarily at the end zone, which marked the end and beginning of each lap (Fig. 6). We did not observe noticeable changes at any of the maze locations in any of the other structures.

Development of Task Bracketing Within Session in aDLS but Not pDMS or HC

Development of aDLS firing rate at the beginning and end of action sequences is similar to the effect seen by Graybiel and colleagues (Barnes et al. 2005; Jog et al. 1999; Smith and Graybiel 2013; Thorn et al. 2010), called task bracketing. Graybiel and colleagues observed that firing rate in the aDLS increasingly “bracketed” action chains along with increased experience and better performance across several sessions on a cued T-maze, recently reported to underlie habit-based behavior (Smith and Graybiel 2013). Our task allowed for the development of automated behavior within a single session, behavior that was disrupted by the contingency switch until subjects readjusted to the new contin-

gency and then, again, automated their behavior. Thus we applied their task-bracketing index (see METHODS; compare Smith and Graybiel 2013) to examine whether a development of task bracketing would occur within session, before and after the contingency switch.

We found that a development of task bracketing was evident only in aDLS (Fig. 6). A two-way ANOVA [region (aDLS vs pDMS) \times laps (early vs late)] revealed a significant main effect of region before [$F(1) = 6$, $P = 0.0144$] and after [$F(1) = 5.57$, $P = 0.0184$] the switch and a significant interaction of region \times laps [$F(1) = 4.43$, $P = 0.0354$] before the switch. No significant task bracketing was found in the HC recordings.

Development of task bracketing in the aDLS tracked behavioral performance, such that as a subject’s performance increased within session so did task bracketing in the aDLS (a Bonferroni-Holm-corrected paired t -test showed that aDLS late laps $>$ aDLS early laps before the switch, $P = 0.013$). When performance decreased after the change of reward contingency was introduced, so too did task bracketing in the aDLS (a Bonferroni-Holm-corrected paired t -test showed that aDLS late laps before the switch $>$ aDLS early laps after the switch, $P = 0.001$). These effects were not observed in any of the other structures. On late laps, aDLS task bracketing was higher than pDMS task bracketing before ($P < 0.0001$) and after ($P = 0.009$) the switch of reward contingency.

To investigate whether this was a general effect, we applied the task-bracketing index to the other maze locations (Fig. 7). In the HC, there were no significant increases or decreases at any of the additional locations. In the striatum, there were no additional significant interactions, such that the rate of firing did not develop or decline over laps in either the aDLS or the

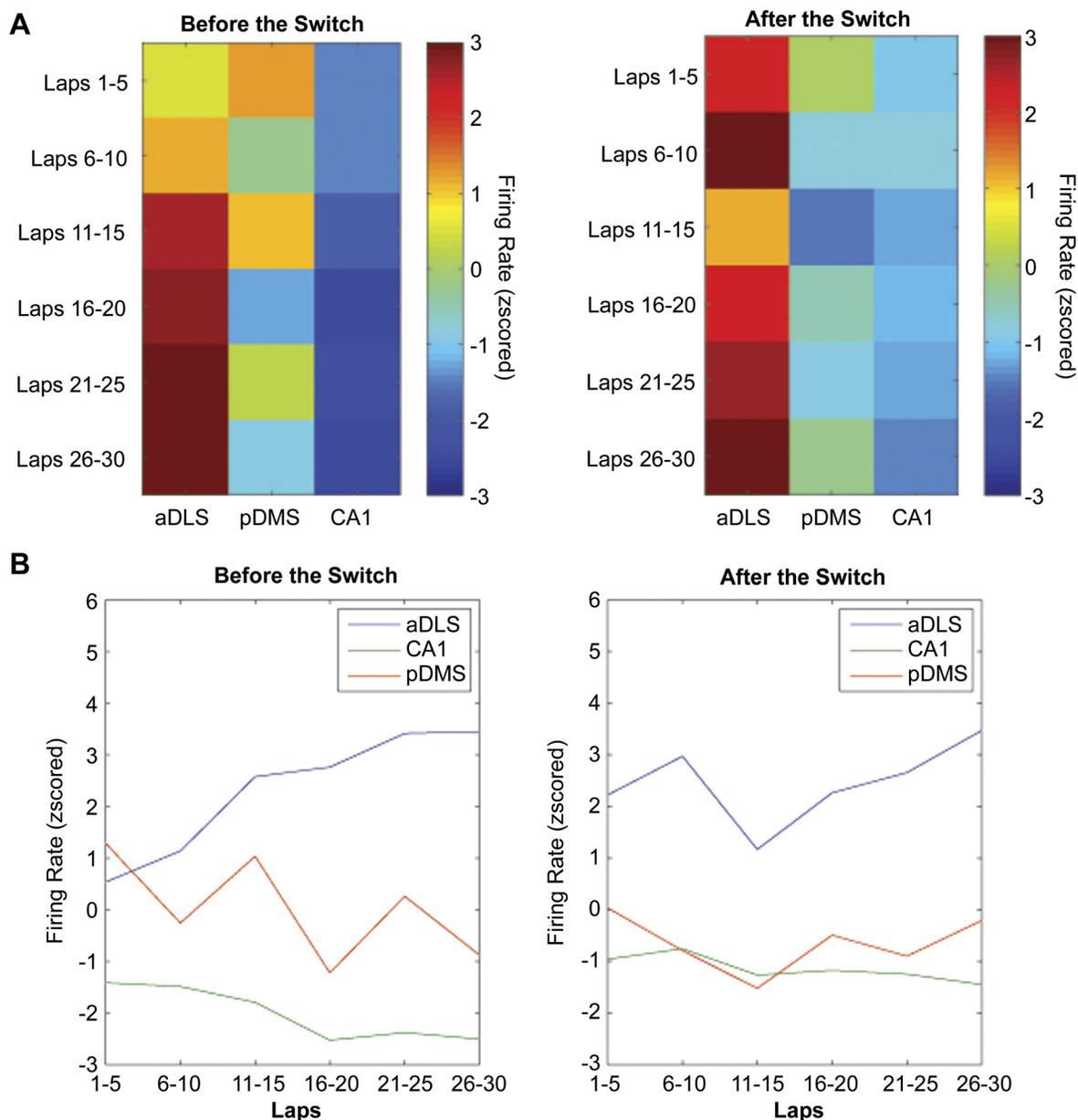


Fig. 5. Overall firing rate across laps. *A*: color plots of z-scored firing rate when averaging all maze locations together and plotting across 5-lap bins before (*left*) and after (*right*) the switch for aDLS, pDMS, and CA1. Only in the aDLS was there a change in average firing rate across laps, with general activity increasing across laps before the switch. This firing rate of aDLS stayed high for first laps after the switch but then decreased and increased again to levels prior to the switch. *B*: line plots of z-scored firing rate across laps for each region show the same effect as in *A*.

pDMS; however, there were several instances where one region or another showed overall higher task bracketing at a given location. Specifically, a main effect of region was found at the navigation sequence, where pDMS firing rate was greater than aDLS, both before [$F(1) = 9.71$, $P = 0.0019$] and after [$F(1) = 4.52$, $P = 0.0336$] the switch, at the choice point both before [$F(1) = 10.32$, $P = 0.0013$] and after [$F(1) = 30.74$, $P < 0.0001$] the switch, and at the feeder entry both before [$F(1) = 48.05$, $P < 0.0001$] and after [$F(1) = 77.35$, $P < 0.0001$] the switch. In addition to the overall elevated firing of aDLS compared with pDMS at the end zone (see above), a main effect of region was found at the feeder exit both before [$F(1) = 77.49$, $P < 0.0001$] and after [$F(1) = 96.41$, $P < 0.0001$] the switch. These additional findings indicate that, while pDMS showed elevated firing at events in the middle of the maze and aDLS

showed elevated firing at the feeder exit and end zone, only at the end zone (the maze location that marked the beginning and end of each lap) was there a development of aDLS task bracketing.

Individual Neurons from Different Regions Displayed Distinct Patterns of Firing Rate Response

To understand how the different structures were responding to the reward contingency change in the task, we examined how the firing pattern of individual neurons changed with the contingency change (Figs. 8–10). To do this, we created tuning curves from individual neurons by linearizing the maze and measuring neuronal response on a lap-by-lap basis. In aDLS, cell responses tended to be biased to one side of the maze or the

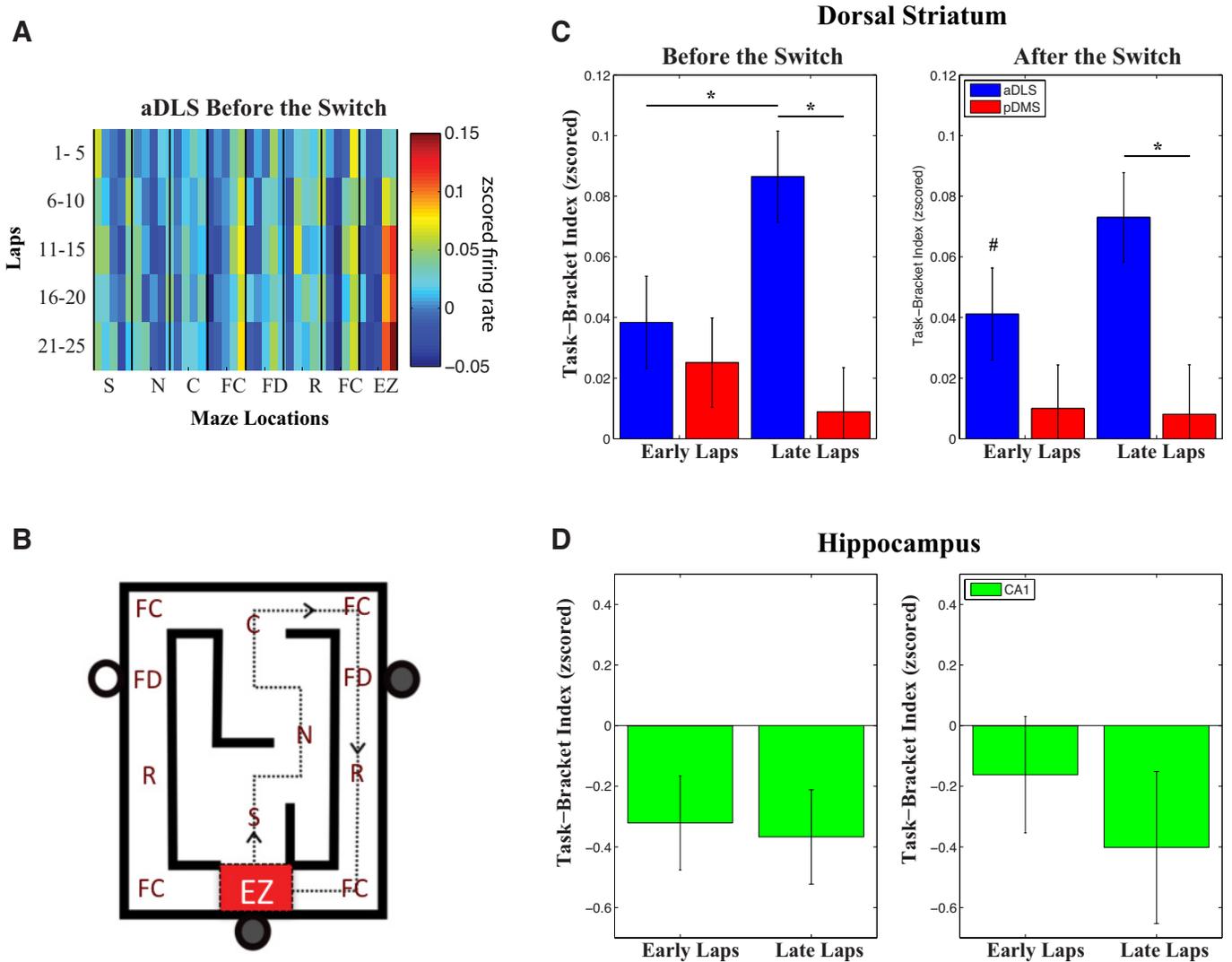


Fig. 6. A development of task bracketing was only observed in the aDLS. *A*: example color plot in aDLS before the switch in reward contingency for firing rate at maze locations. S, start zone; N, navigation sequence; C, choice point; FC, feeder cue; FD, feeder; R, return rail; EZ, end zone. *B*: template of maze locations showing where firing rate was obtained. Red box at end zone indicates where general increase in firing rate was occurring in aDLS neurons. *C* and *D*: task-bracketing index for striatal (*C*) and hippocampal (*D*) neurons before (*left*) and after (*right*) the switch. Only in aDLS did task-bracketing index increase on late laps before the switch, decrease on early laps after the switch, then increase again on late laps after the switch. *Significant difference. #Significant difference comparing aDLS early laps after the switch to aDLS late laps before the switch.

other (Fig. 8). In pDMS, firing rate appeared to be altered by the change of contingency, introduced approximately midway through the session, regardless of maze side (Fig. 9). In CA1, the location of the animal on the maze appeared to govern the neuronal response (Fig. 10).

To investigate whether there were differences of firing patterns of individual neurons between structures, we calculated two correlations of firing patterns across changes between goal (left- or right-side responding) and changes between reward contingency (pre- and postswitch) for each neuron. Specifically, a correlation coefficient was obtained for before vs. after contingency switch laps and for left vs. right laps. To control for lap side and contingency switch, we analyzed correlation coefficients in four different conditions (left laps before the switch vs. left laps after the switch, right laps before the switch vs. right laps after the switch, left vs. right laps before the switch, and left vs. right laps after the switch). We plotted these on a two-dimensional plane in order to observe

interactions of specific side of the maze and laps that came before and after the change of contingency. aDLS neurons displayed different patterns of firing for left and right laps more than pDMS and HC.

Correlation coefficients of aDLS cells were consistently lower when correlating firing rates on left vs. right laps, with a broad range of correlation coefficients for before- vs. after-switch firing rates in all four conditions (Fig. 11). Compared with pDMS and CA1, correlation coefficients for left vs. right firing rates were lowest in aDLS ($P < 0.0001$, compare Figs. 12 and 13, see Fig. 14). The HC had several place cells with low correlations between left and right laps; however, these cases all occurred when the place field was on a return path. Whenever a place cell was located on the navigation sequence, it appeared that firing rate was consistent between left and right laps as well as before and after the switch (Fig. 10).

Previous experiments have reported that HC place cells sometimes change their place fields (remap) or the firing rate

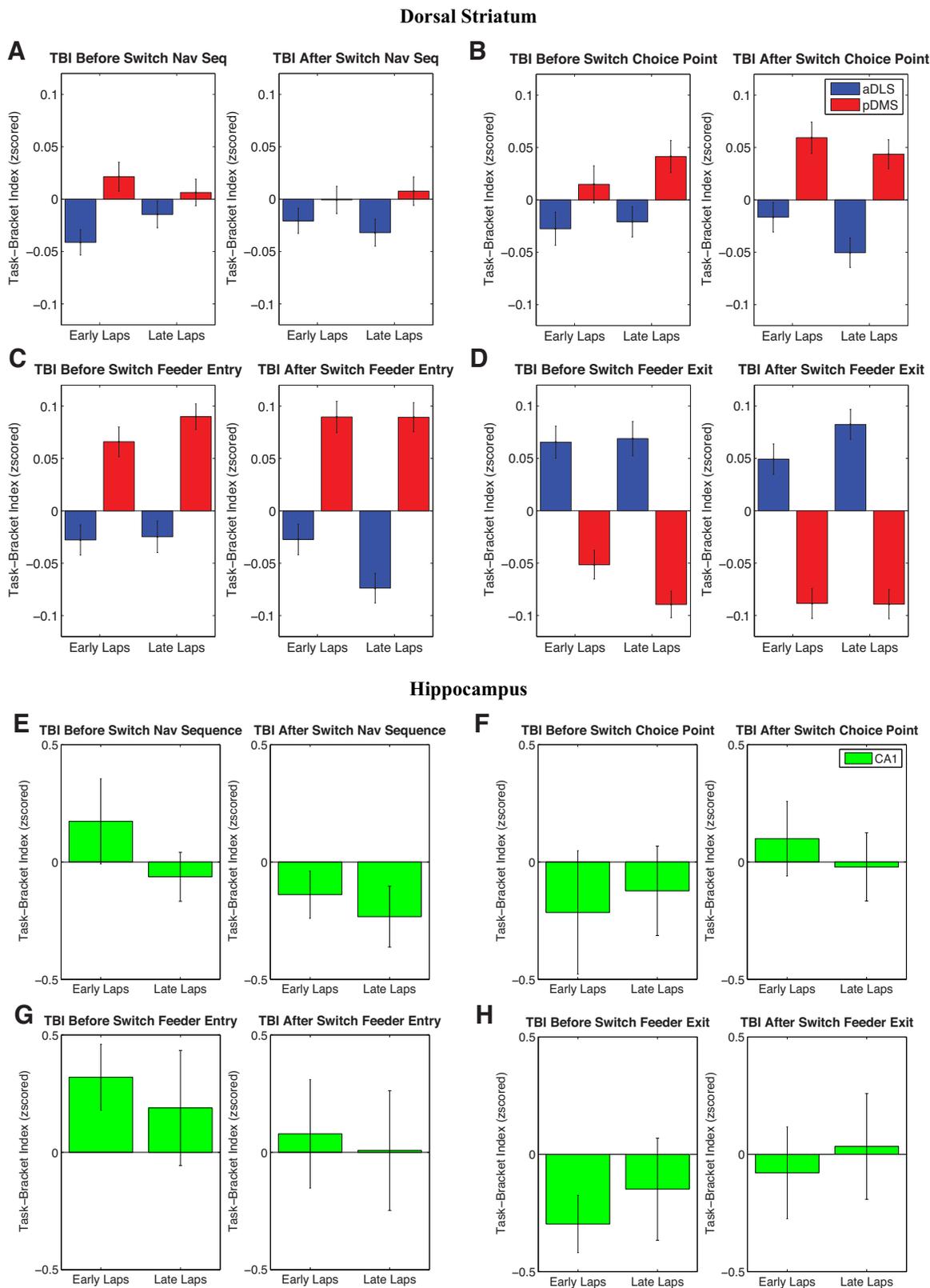


Fig. 7. Task-bracketing index (TBI) at additional maze locations. *A–D*: task bracketing in striatum at the navigation sequence (*A*), choice point (*B*), feeder entry (*C*), and feeder exit (*D*) before and after the contingency switch. Task bracketing was greater overall in pDMS at the navigation sequence, choice point, and feeder entry. Task bracketing measures were greater overall in aDLS at the feeder exit. There were no significant increases or decreases in any of these other maze locations. *E–H*: task bracketing in hippocampus at the navigation sequence (*E*), choice point (*F*), feeder entry (*G*), and feeder exit (*H*). No significant differences from zero were found in CA1.

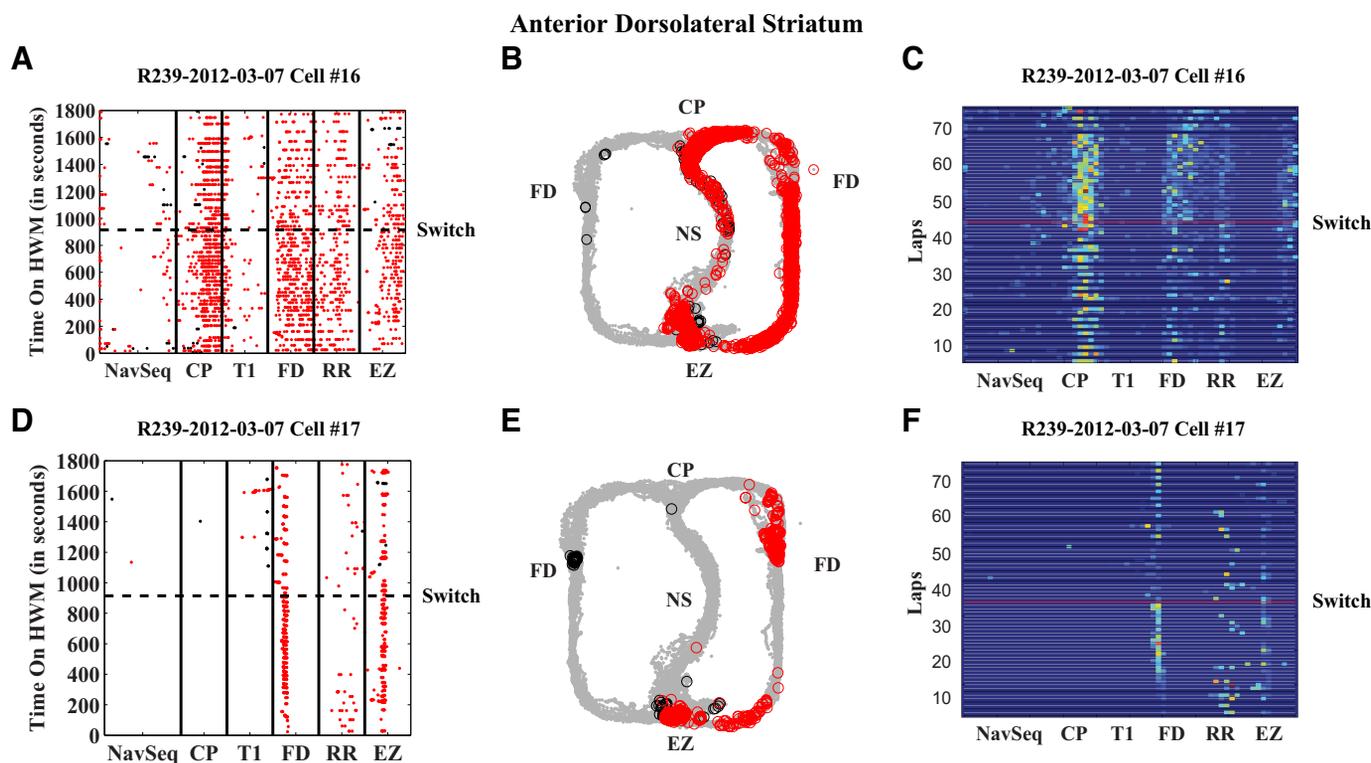


Fig. 8. Example cells from aDLS. First example cell raster (A) shows this cell firing primarily on right laps (red dots) even at the navigation sequence and choice point, clearly shown when plotted on the rat's tracking data (B). Tuning curves (C) used for the correlation analysis show a right-side bias for this cell. Warmer colors indicate higher firing at that location, and cooler colors indicate lower firing. Lighter horizontal lines indicate right laps, while darker horizontal lines indicate left laps. Red horizontal line is the switch of contingency lap. Second example cell raster (D), scatterplot on tracking data (E), and tuning curve color plot (F) show similar phenomena. HWM, Hebb-Williams maze.

within a given field (rate modulation) on shared paths. These changes are called “splitter cells” because they split the cell's spatial response by a nonspatial trajectory component (Catanese et al. 2014; Wood et al. 2000). In CA1, we observed only a few examples of place field remapping on the navigation sequence (2/17 CA1 cells with place fields on the navigation sequence), consistent with our observations of tuning curve plots. The other cells whose place fields were at the same location showed minimal rate modulation either between right and left turns [ratio of firing rate, 1.77 ± 1.32 SD, not significant (NS)] or between before and after the switch of reward contingency (1.92 ± 0.73 SD, NS). For the splitter cell and rate modulation analyses, HC and striatal (see below) results were analyzed by comparing the ratio of firing rate to 1, since an absence of change in firing rate would equal a 1-to-1 ratio. *t*-Tests were corrected with Bonferroni statistics [$0.05/n$ (regions)].

To compare striatal responses to task and behavioral changes, we applied these same splitter-cell analyses to the striatal data. In both aDLS and pDMS, almost half of the cells on the navigation sequence showed some trajectory-related modulation. In aDLS 15 of 37 cells changed their maze-related response on the navigation sequence, and in pDMS 7 of 16 cells changed their response. A χ^2 -test found a trend of a main effect in the number of splitter cells in aDLS, pDMS, and HC but did not reach significance [$\chi^2(2) = 5.106$, $P = 0.08$]. Of the half that did not change their maze responses, aDLS showed more rate modulation than either HC or pDMS between left and right turns [ratio of firing rate, 2.63 ± 1.85 SD, $t(21) = 4.14$, $P < 0.008$] while pDMS showed less (ratio of firing rate,

1.25 ± 0.25 SD, NS). A one-way ANOVA showed a main effect comparing ratio of firing rate in all regions between left and right laps [$F(2) = 3.17$, $P = 0.05$]. On the other hand, pDMS appeared to show more rate modulation between before and after the switch of reward contingency. Surprisingly, higher variation in pDMS resulted in nonsignificant rate modulation (2.37 ± 2.06 SD pDMS, NS), while rate modulation in aDLS was significant [1.97 ± 1.11 SD aDLS, $t(15) = 3.47$, $P < 0.008$]; however, a one-way ANOVA did not show a main effect when comparing ratio of firing rate between all regions before and after the switch of reward contingency [$F(2) = 0.35$, $P = 0.71$].

During probe sessions, we introduced a change to the task requirements for reinforcement that brought about a change in behavioral strategy (Fig. 3). We would expect neuronal response to reflect this strategy shift if an area was partially mediating this behavioral change. Observation of individual neurons found that pDMS neurons showed different firing rates before and after the switch (Fig. 12). Results from the correlation analysis showed that firing rate of before- vs. after-switch laps was altered more in pDMS than in aDLS and CA1 (lower correlations, Fig. 14, for each comparison $P < 0.0001$). Although the firing rates of many pDMS neurons were equally correlated on left vs. right laps and before- vs. after-switch laps, a number of pDMS neurons were less correlated on before- vs. after-switch laps than left vs. right laps (Fig. 12) and significantly more so than any other structure (Fig. 14). This was the case, even when subjects were performing the same actions during the same sequences, as evidenced by example cells showing differential firing patterns before and

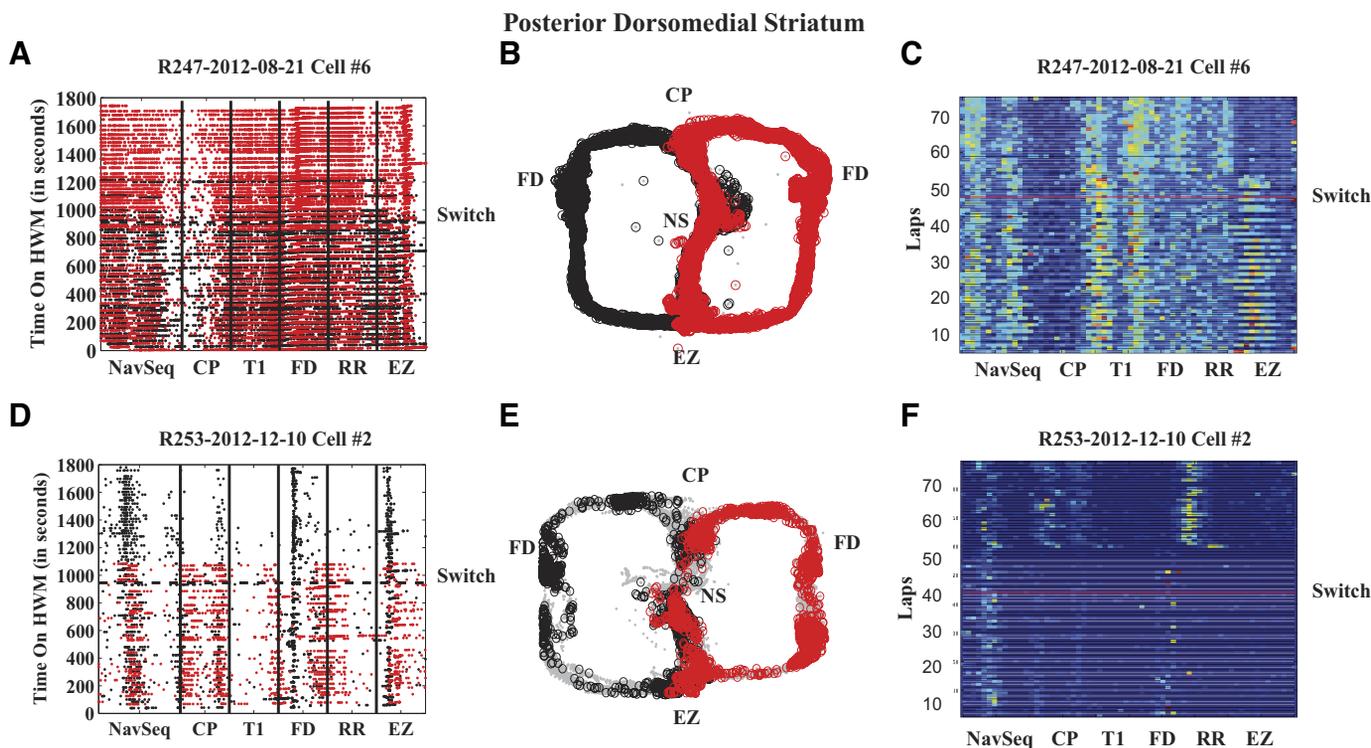


Fig. 9. Example cells in the pDMS. First example cell raster (*A*) shows this cell firing on both right (red dots) and left (black dots) laps at all maze locations, which is reflected by the spikes plotted on the tracking data (*B*). Tuning curves (*C*) used for the correlation analysis show a right-side bias for this cell at the end zone that decreases dramatically after the contingency switch (to a rightward reward contingency). At other maze locations, there are other noticeable changes in tuning curve information. *D* and *E*: second example cell raster (*D*) and neuronal spikes plotted on the tracking data (*E*). The tuning curve color plot (*F*) shows that tuning curve information increases at several maze locations, staying consistent at the navigation sequence. NS, navigation sequence; CP, choice point; T1, turn after choice point; FD, side feeder; RR, return rail; EZ, end zone. Tuning curve plot: warmer colors indicate higher firing at that location, and cooler colors indicate lower firing. Lighter horizontal lines indicate right laps, and darker horizontal lines indicate left laps. Red horizontal line is the switch of reward contingency lap.

after the switch on a location of the maze where the rat would perform similar movements (Fig. 12).

Neurons in pDMS were less likely to change their neuronal response based on the lap side, compared with aDLS neurons (Fig. 14). Unlike aDLS, which has been shown to respond to specific actions, such as taking a left turn or arriving at the left feeder, results from the present study indicate that pDMS neurons may encode specific strategies, reflecting the current action-reward contingency. Although some reports have suggested that HC neuronal activity changes in response to a behavioral strategy shift, we did not observe this—our results indicated that neurons in the CA1 did not remap to new locations after the contingency change (Fig. 13). HC cells did show some rate modulation (Fyhn et al. 2007) on the navigation sequence across the contingency changes, but the modulation levels were not significant. Thus, unlike pDMS neurons, the changes in CA1 tuning curves were likely a consequence in the change in behavior (being more likely to go right vs. left or vice versa) and did not reflect the behavioral strategy change. This is likely due to the way that the animals were trained to expect multiple reward contingencies within a single environment (Fuhs 2006).

DISCUSSION

Current theories of decision making suggest that there are multiple decision making processes, including goal-oriented (action-outcome, deliberative) and habit-based (chunked action

chains, procedural) that accomplish tasks by different information processing algorithms instantiated through different interacting anatomical structures. This hypothesis implies that different structures should provide different representations and those different representations should reflect information processing differently. Current theories have suggested that HC plays an important role in the goal-oriented system while aDLS plays an important role in the habit-based system. Although lesion data have suggested a role for pDMS in the goal-oriented system, neural ensembles therein have not been explored.

We found marked differences in aDLS, pDMS, and HC neuronal responses to a reward contingency change on a spatial navigation task on which rats automated their behavior, reverted to a goal-oriented decision process, and then reautomated their behavior. Neuronal firing in pDMS reflected changes in reward contingency, more so than either aDLS or HC (CA1) neurons. In contrast, aDLS developed firing at the beginning and end of laps that tracked behavioral performance on the task (task bracketing). CA1 tuning curves did not appear to change with changes in behavior. Instead, CA1 neurons displayed typical place cell activity at different locations on the maze that remained stable within sessions, consistent with a cognitive map that had already been formed for this well-learned task.

Looking at how information changed through the decision making process, neuronal ensembles recorded from CA1

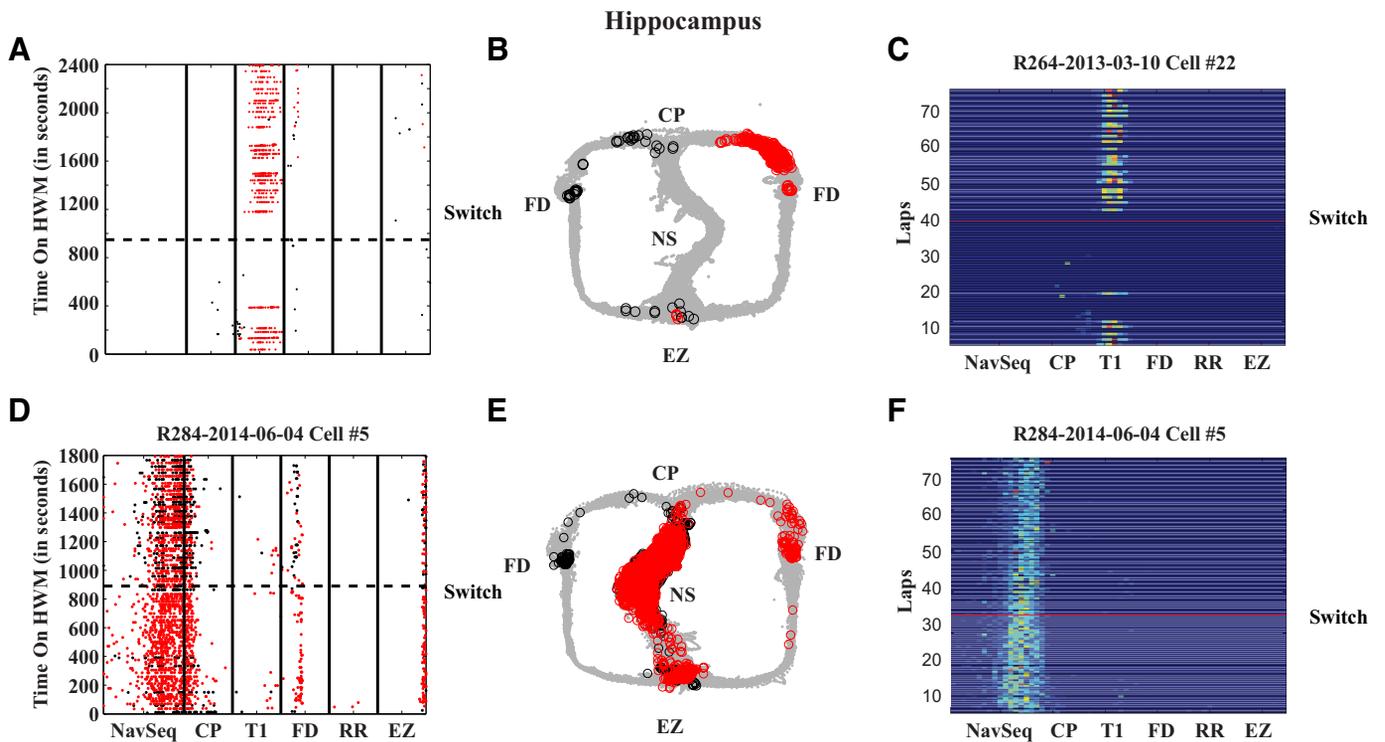


Fig. 10. Example cells in the CA1. First example cell raster (A) shows this cell firing primarily on right laps (red dots), reflected by the neuronal spikes on the rat's tracking data (B) and the tuning curve plot (C). Second example cell raster (D), and neuronal spikes on the tracking data (E), and tuning curve color plot (F) show a cell that fires primarily on the navigation sequence for left and right laps. Tuning curve plot: warmer colors indicate higher firing at that location, and cooler colors indicate lower firing. Lighter horizontal lines indicate right laps, and darker horizontal lines indicate left laps. Red horizontal line is the switch of reward contingency lap.

showed more forward representation at the choice point compared with aDLS and pDMS. Importantly, CA1 represented sides equally on VTE laps but not non-VTE laps, while aDLS and pDMS represented the chosen side more than the unchosen

side on all laps, suggesting that CA1 reflected the searching process itself, while the striatal representations reflected the selected action. On non-VTE laps, we expect that the rat was already aware of its target destination, and thus CA1 ensembles

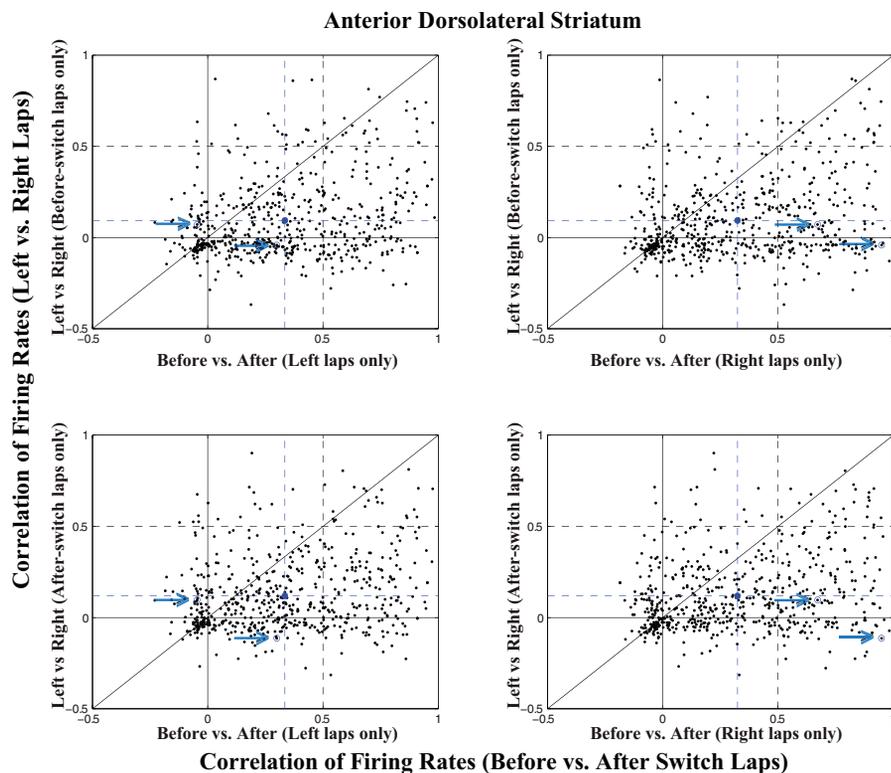
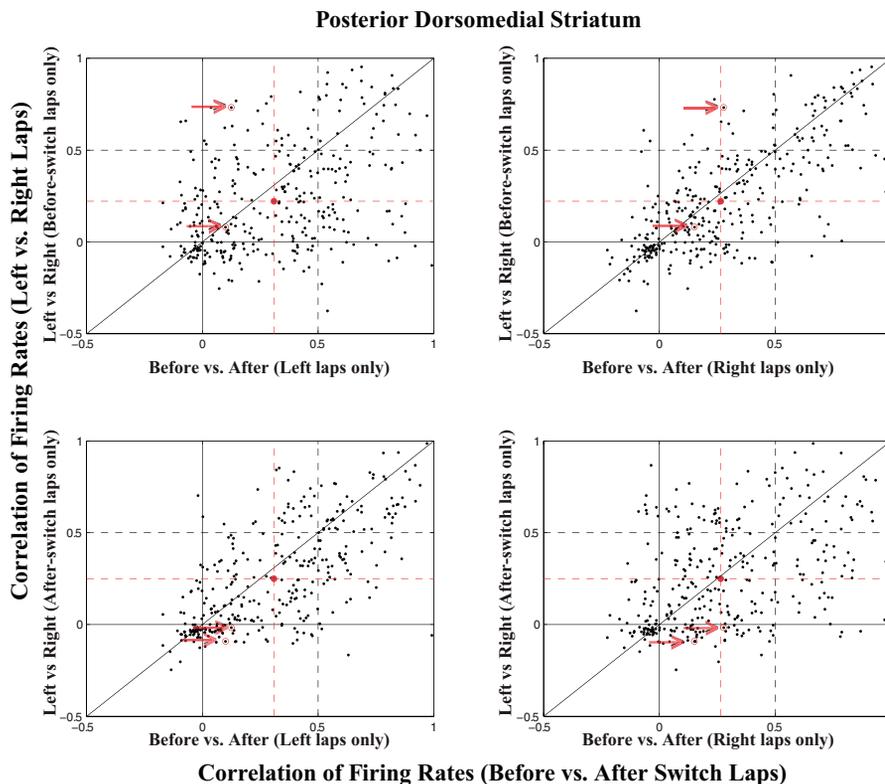


Fig. 11. Correlations of firing rates for aDLS were obtained by taking an average firing rate from the constructed tuning curves (e.g., Fig. 8, C and F). This was done for left laps before the switch (BS) vs. left laps after the switch (AS), right laps BS vs. right laps AS, left vs. right laps BS, and left vs. right laps AS. Correlations were then plotted for left laps BS vs. left laps AS \times left vs. right laps BS (top left), right laps BS vs. right laps AS \times left vs. right laps BS (top right), left laps BS vs. left laps AS \times left vs. right laps AS (bottom left), and right laps BS vs. right laps AS \times left vs. right laps AS (bottom right). In all plots, the means of the correlations of firing rates are plotted in filled-in blue circles, the mean for before- vs. after-switch laps is plotted as a vertical blue dashed line, and the mean for left vs. right laps is plotted as a horizontal blue dashed line. Example cells from Fig. 8 are indicated with blue circles and a blue arrow. Correlations of firing rates in the aDLS were significantly less on left vs. right laps compared with pDMS and CA1 (Fig. 14).

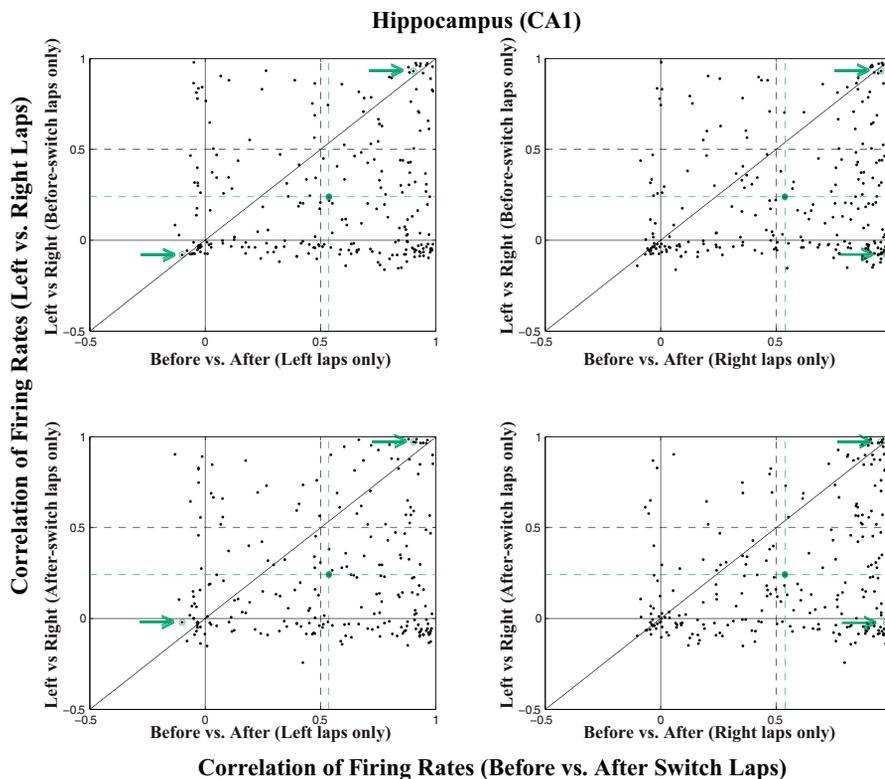
Fig. 12. Correlations of firing rates for pDMS were obtained by taking an average firing rate from the constructed tuning curves (Fig. 9, C and F). This was done for left laps before the switch (BS) vs. left laps after the switch (AS), right laps BS vs. right laps AS, left vs. right laps BS, and left vs. right laps AS. Correlations were then plotted for left laps BS vs. left laps AS \times left vs. right laps BS (top left), right laps BS vs. right laps AS \times left vs. right laps BS (top right), left laps BS vs. left laps AS \times left vs. right laps AS (bottom left), and right laps BS vs. right laps AS \times left vs. right laps AS (bottom right). In all plots, the means of the correlations of firing rates are plotted in filled-in red circles, the mean for before- vs. after-switch laps is plotted as a vertical red dashed line, and the mean for left vs. right laps is plotted as a horizontal red dashed line. Example cells from Fig. 9 are indicated with red circles and a red arrow. Correlations of firing rates in the pDMS were significantly less on before- vs. after-switch laps compared with aDLS and CA1 (Fig. 14).



reflected the path to the current goal of the rat (Gupta et al. 2012; Wikenheiser and Redish 2015a, 2015b) However, on VTE laps, we expect that the rat was deliberating over multiple possibilities and the CA1 ensembles reflected the search process examining both goals. This distinction is consistent with previous observations in HC neural ensembles (Johnson and

Redish 2007). Consistent with previous experiments, aDLS ensembles did not show strong forward representations (van der Meer et al. 2010), and what little forward information they did represent reflected the chosen option. Interestingly, although pDMS has been identified as playing an important role in goal-oriented (flexible, deliberative) decision processes (Lex

Fig. 13. Correlations of firing rates for CA1 were obtained by taking an average firing rate from the constructed tuning curves (Fig. 10, C and F). This was done for left laps before the switch (BS) vs. left laps after the switch (AS), right laps BS vs. right laps AS, left vs. right laps BS, and left vs. right laps AS. Correlations were then plotted for left laps BS vs. left laps AS \times left vs. right laps BS (top left), right laps BS vs. right laps AS \times left vs. right laps BS (top right), left laps BS vs. left laps AS \times left vs. right laps AS (bottom left), and right laps BS vs. right laps AS \times left vs. right laps AS (bottom right). In all plots, the means of the correlations of firing rates are plotted in filled-in green circles, the mean for before- vs. after-switch laps is plotted as a vertical green dashed line, and the mean for left vs. right laps is plotted as a horizontal green dashed line. Example cells from Fig. 10 are indicated with green circles and a green arrow. Correlations of firing rates in the CA1 were significantly more on before- vs. after-switch laps compared with aDLS and pDMS and more on left vs. right laps compared with aDLS (Fig. 14).



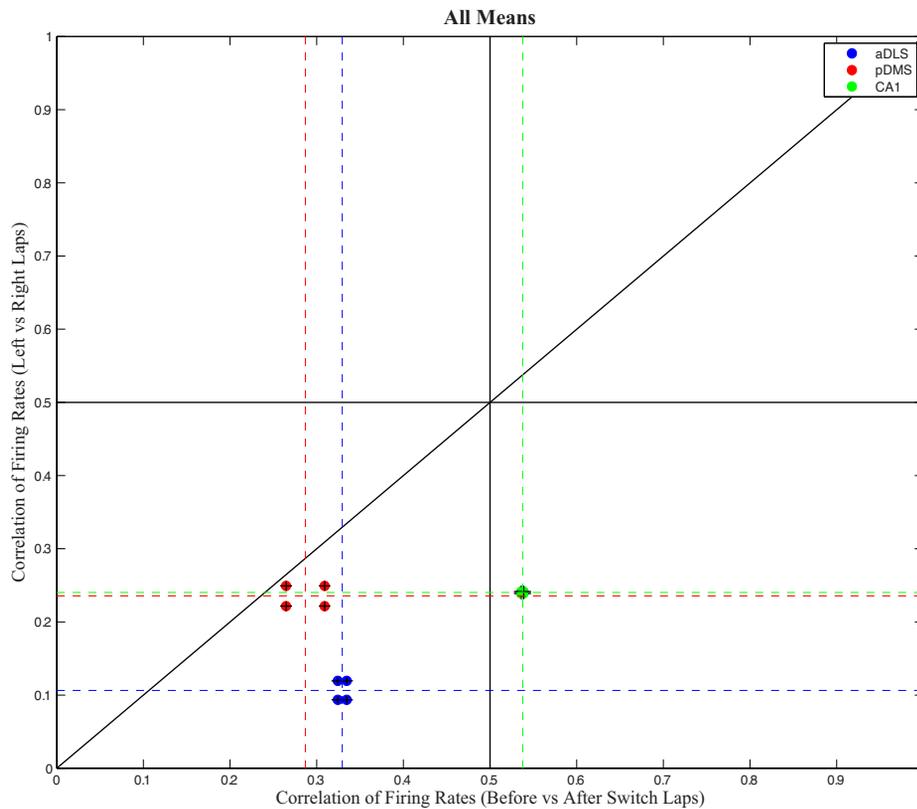


Fig. 14. All means from aDLS, pDMS, and CA1 correlations of firing rate (Figs. 11–13). Means of aDLS correlations were significantly lower on left vs. right laps compared with pDMS and CA1 (for each comparison, $P < 0.0001$). Means of pDMS correlations were significantly lower on before- vs. after-switch laps compared with aDLS and CA1 (for each comparison, $P < 0.0001$).

and Hauber 2010a, 2010b; Yin et al. 2005a, 2005b; Yin and Knowlton 2004), the pDMS ensembles appeared more like aDLS ensembles than HC ensembles, with limited forward representations and a preference for the chosen side, even on VTE laps.

Recognizing a change in the environment and adjusting behavior appropriately is essential for survival. Different cortical substrates are involved in different environmental/reward-related behavioral changes. For example, adjusting behavioral strategies, reversal learning, and contingency degradation are all mediated by different cortical areas (Corbit et al. 2002; Lex and Hauber 2010a, 2010b; Lucantonio et al. 2014; Ragozzino 2007; Schoenbaum et al. 2002). Common among the evaluation of each change to the environment is the necessity of flexibly associating the outcome with the preceding action. In the HWM-LRA task, this entails the recognition that an action no longer leads to reward and the subsequent adjustment of behavior. Several recent studies have implicated the pDMS and its related input structures mediating these sorts of reward-related behavioral shifts (Corbit et al. 2002; Izquierdo et al. 2004; Killcross and Coutureau 2003; Lex and Hauber 2010a, 2010b; Shiflett et al. 2010; Yin et al. 2005a, 2005b), making pDMS/orbitofrontal and pDMS/prelimbic circuits important for altering behavior when an unexpected variation occurs in the environment. Recent studies indicate that cortical areas may evaluate state changes (for reviews, see Lucantonio et al. 2014; Ragozzino 2007; Torregrossa et al. 2008) and pDMS may integrate information from cortical areas into appropriate actions (Kimchi and Laubach 2009; Stalnaker et al. 2012). Results from the present study support this idea, with pDMS neuronal patterns reflecting different behavioral strategies more than aDLS or HC neurons.

In stable reward delivery contingencies, as the animal realizes that the same actions consistently lead to desired outcomes, goal-directed behavior typically transitions to more automated behaviors. Goal-directed behavior is cognitively intensive, since planning for future outcomes is a computationally expensive operation that must occur before action selection. Automating behavior is a way to optimize benefit from an environment. Thus situations (e.g., stimuli) associated with actions that have consistently led to reinforcement eventually come to release appropriate action chains (Adams 1982; Daw et al. 2006; Dezfouli et al. 2014; van der Meer et al. 2012). These stimulus-action associations are cached and controlled by sensorimotor circuits in the basal ganglia, such as the aDLS (Everitt and Robbins 2013; Graybiel 1998, 2008; Hikosaka et al. 2002; Miyachi et al. 2002; Packard and McGaugh 1996; Yin and Knowlton 2006).

Previous studies have reported the reorganization of neuronal activity in the aDLS with increased experience on a task (Barnes et al. 2005; Jog et al. 1999; van der Meer et al. 2010; Thorn et al. 2010). Graybiel and colleagues have reported increases in neuronal activity across several training states at the beginning and end of action sequences (task bracketing) on a cued T-maze, recently reported to underlie habit-based behavior (Smith and Graybiel 2013). We found that task bracketing in the aDLS developed along with increased performance within the task but was disrupted when a change of reward contingency was introduced, correlating with both decreased performance and increased revaluation.

Studies in primates suggest anatomical and functional differences between the rostral/caudal regions of the caudate and putamen (Miyachi et al. 1997, 2002), similar to the dorsolateral vs. dorsomedial striatal differences in rodents (Devan et al.

1999; Yin and Knowlton 2004, 2006), with recent rodent studies suggesting more of a role of posterior dorsomedial (pDMS) than anterior dorsomedial (aDMS) in flexible learning (Yin et al. 2005b; Yin and Knowlton 2004). In primates, a recent study reported that the head of the caudate is important for flexible learning and the tail is important for more stable information processing (Kim et al. 2014); our results suggest more stable processing in the rodent aDLS and more flexible processing in the rodent pDMS.

Previous work has reported that both HC (O'Keefe and Nadel 1978; Redish 1999) and aDLS (Mizumori et al. 2004; Schmitzer-Torbert 2004; Yeshenko et al. 2004) neurons were spatially tuned on a task in which spatial cues provided information about how to obtain rewards; however, on a spatial task in which spatial cues did not provide information about how to obtain rewards, only HC neurons were spatially tuned (Berke et al. 2009; Wikenheiser and Redish 2011) while aDLS neurons were not (Berke et al. 2009; Schmitzer-Torbert and Redish 2008). We found that HC and aDLS neurons responded in a similar fashion to spatial context in the present study, such that many neurons in both structures were spatially tuned to one side of the maze or the other. Neurons in the aDLS had more of a tendency to respond differently to left and right laps compared with neurons in any of the other structures, even on the navigation sequence. In contrast, HC place fields only differentiated left and right laps when the fields were on the return arms. This is inconsistent with previous studies that have found rate modulation and splitter cells (Ferbinteanu and Shapiro 2003; Frank et al. 2000; Wood et al. 2000) and may be due to differences of training and proficiency on the tasks, since the differentiation of activity depending on context correlates to task performance (Ferbinteanu and Shapiro 2003). Further differentiation of aDLS and HC was evident in the task-bracketing index measure, where only aDLS (and not HC) neurons developed task bracketing. Interestingly, pDMS did not develop task bracketing either, suggesting that among these three structures aDLS plays a unique role in the action chunking of the habit-based (procedural) decision system.

Our results suggest that separate circuits in the basal ganglia integrate relevant cortical information during automation of actions and the recognition of when new behaviors are needed to continue obtaining rewards. Subregions of the dorsal striatum, such as the aDLS and pDMS, integrated different information, with aDLS neurons developing bracketing patterns of firing along with behavioral performance and pDMS correlating with changes in the reward delivery contingency. HC neurons played a different role entirely, with an already available cognitive map (on this well-learned maze), on which search processes could play out. Interestingly, these search processes were not seen in either aDLS or pDMS.

ACKNOWLEDGMENTS

Present address of P. S. Regier: Center for Studies of Addiction, Department of Psychiatry, University of Pennsylvania, Philadelphia, PA.

GRANTS

Funding for this work was provided by National Institutes of Health (NIH) Grants MH-080318 and DA-030672 (A. D. Redish), a training fellowship on NIH T32 Grant DA-007234 (P. S. Regier), and Japan Society for the Promotion of Science (JSPS) KAKENHI-11J06508 (S. Amemiya).

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

Author contributions: P.S.R., S.A., and A.D.R. conception and design of research; P.S.R. and S.A. performed experiments; P.S.R., S.A., and A.D.R. analyzed data; P.S.R., S.A., and A.D.R. interpreted results of experiments; P.S.R., S.A., and A.D.R. prepared figures; P.S.R., S.A., and A.D.R. drafted manuscript; P.S.R., S.A., and A.D.R. edited and revised manuscript; P.S.R., S.A., and A.D.R. approved final version of manuscript.

REFERENCES

- Adams CD.** Variations in the sensitivity of instrumental responding to reinforcer devaluation. *Q J Exp Psychol B* 34: 77–98, 1982.
- Adams CD, Dickinson A.** Instrumental responding following reinforcer devaluation. *Q J Exp Psychol B* 33: 109–121, 1981.
- Alexander GE, Crutcher MD.** Functional architecture of basal ganglia circuits: neural substrates of parallel processing. *Trends Neurosci* 13: 266–271, 1990.
- Alexander GE, DeLong MR.** Microstimulation of the primate neostriatum. I. Physiological properties of striatal microexcitable zones. *J Neurophysiol* 53: 1401–1416, 1985a.
- Alexander GE, DeLong MR.** Microstimulation of the primate neostriatum. II. Somatotopic organization of striatal microexcitable zones and their relation to neuronal response properties. *J Neurophysiol* 53: 1417–1430, 1985b.
- Balleine BW, Delgado MR, Hikosaka O.** The role of the dorsal striatum in reward and decision-making. *J Neurosci* 27: 8161–8165, 2007.
- Barnes TD, Kubota Y, Hu D, Jin DZ, Graybiel AM.** Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature* 437: 1158–1161, 2005.
- Berendse HW, Graaf YG, Groenewegen HJ.** Topographical organization and relationship with ventral striatal compartments of prefrontal corticostriatal projections in the rat. *J Comp Neurol* 316: 314–347, 1992.
- Berke JD, Breck JT, Eichenbaum H.** Striatal versus hippocampal representations during win-stay maze performance. *J Neurophysiol* 101: 1575–1587, 2009.
- Blumenthal A, Steiner A, Seeland K, Redish AD.** Effects of pharmacological manipulations of NMDA-receptors on deliberation in the Multiple-T task. *Neurobiol Learn Mem* 95: 376–384, 2011.
- Buckner RL, Carroll DC.** Self-projection and the brain. *Trends Cogn Sci* 11: 49–57, 2007.
- Carelli RM, West MO.** Representation of the body by single neurons in the dorsolateral striatum of the awake, unrestrained rat. *J Comp Neurol* 309: 231–249, 1991.
- Castañe A, Theobald DE, Robbins TW.** Selective lesions of the dorsomedial striatum impair serial spatial reversal learning in rats. *Behav Brain Res* 210: 74–83, 2010.
- Catanese J, Viggiano A, Cerasti E, Zugaro MB, Wiener SI.** Retrospectively and prospectively modulated hippocampal place responses are differentially distributed along a common path in a continuous T-maze. *J Neurosci* 34: 13163–13169, 2014.
- Cho J, West MO.** Distributions of single neurons related to body parts in the lateral striatum of the rat. *Brain Res* 756: 241–246, 1997.
- Clarke HF, Robbins TW, Roberts AC.** Lesions of the medial striatum in monkeys produce perseverative impairments during reversal learning similar to those produced by lesions of the orbitofrontal cortex. *J Neurosci* 28: 10972–10982, 2008.
- Corbit LH, Janak PH.** Posterior dorsomedial striatum is critical for both selective instrumental and Pavlovian reward learning. *Eur J Neurosci* 31: 1312–1321, 2010.
- Corbit LH, Ostlund SB, Balleine BW.** Sensitivity to instrumental contingency degradation is mediated by the entorhinal cortex and its efferents via the dorsal hippocampus. *J Neurosci* 22: 10976–10984, 2002.
- Daw ND, Niv Y, Dayan P.** Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci* 8: 1704–1711, 2005.
- Daw ND, Niv Y, Dayan P.** Actions, policies, values and the basal ganglia. In: *Recent Breakthroughs in Basal Ganglia Research*, edited by Bezdard E. New York: Nova Science, 2006, p. 91–106.
- Devan BD, McDonald RJ, White NM.** Effects of medial and lateral caudate-putamen lesions on place- and cue-guided behaviors in the water maze: relation to thigmotaxis. *Behav Brain Res* 100: 5–14, 1999.

- Dezfouli A, Lingawi NW, Balleine BW.** Habits as action sequences: hierarchical action control and changes in outcome value. *Philos Trans R Soc Lond B Biol Sci* 369: 20130482, 2014.
- Everitt BJ, Robbins TW.** From the ventral to the dorsal striatum: devolving views of their roles in drug addiction. *Neurosci Biobehav Rev* 37: 1946–1954, 2013.
- Ferbinteanu J, Shapiro ML.** Prospective and retrospective memory coding in the hippocampus. *Neuron* 40: 1227–1239, 2003.
- Frank LM, Brown EN, Wilson M.** Trajectory encoding in the hippocampus and entorhinal cortex. *Neuron* 27: 169–178, 2000.
- Fuhs MC.** *Space and Context in the Rodent Hippocampal Region* (PhD dissertation). Pittsburgh, PA: Carnegie Mellon Univ., 2006.
- Fyhn M, Hafting T, Treves A, Moser MB, Moser EI.** Hippocampal remapping and grid realignment in entorhinal cortex. *Nature* 446: 190–194, 2007.
- Gardner RS, Uttaro MR, Fleming SE, Suarez DF, Ascoli GA, Dumas TC.** A secondary working memory challenge preserves primary place strategies despite overtraining. *Learn Mem* 20: 648–656, 2013.
- Graybiel AM.** The basal ganglia and chunking of action repertoires. *Neurobiol Learn Mem* 70: 119–136, 1998.
- Graybiel AM.** Habits, rituals, and the evaluative brain. *Annu Rev Neurosci* 31: 359–387, 2008.
- Gremel CM, Costa RM.** Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nat Commun* 4: 2264, 2013.
- Gupta AS, van der Meer MA, Touretzky DS, Redish AD.** Hippocampal replay is not a simple function of experience. *Neuron* 65: 695–705, 2010.
- Gupta AS, van der Meer MA, Touretzky DS, Redish AD.** Segmentation of spatial experience by hippocampal θ sequences. *Nat Neurosci* 15: 1032–1039, 2012.
- Hebb D, Williams K.** A method of rating animal intelligence. *J Gen Psychol* 34: 59–65, 1946.
- Hikosaka O, Nakamura K, Sakai K, Nakahara H.** Central mechanisms of motor skill learning. *Curr Opin Neurobiol* 12: 217–222, 2002.
- Hikosaka O, Rand MK, Miyachi S, Miyashita K.** Learning of sequential movements in the monkey: process of learning and retention of memory. *J Neurophysiol* 74: 1652–1661, 1995.
- Izquierdo A, Suda RK, Murray EA.** Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. *J Neurosci* 24: 7540–7548, 2004.
- Jog MS, Kubota Y, Connolly CI, Hillegaart V, Graybiel AM.** Building neural representations of habits. *Science* 286: 1745–1749, 1999.
- Johnson A, van der Meer MA, Redish AD.** Integrating hippocampus and striatum in decision-making. *Curr Opin Neurobiol* 17: 692–697, 2007.
- Johnson A, Redish AD.** Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *J Neurosci* 27: 12176–12189, 2007.
- Killcross S, Coutureau E.** Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb Cortex* 13: 400–408, 2003.
- Kim HF, Ghazizadeh A, Hikosaka O.** Separate groups of dopamine neurons innervate caudate head and tail encoding flexible and stable value memories. *Front Neuroanat* 8: 120, 2014.
- Kimchi EY, Laubach M.** The dorsomedial striatum reflects response bias during learning. *J Neurosci* 29: 14891–14902, 2009.
- Kirkby R.** Caudate nucleus lesions and perseverative behavior. *Physiol Behav* 4: 451–454, 1969.
- Lex B, Hauber W.** Disconnection of the entorhinal cortex and dorsomedial striatum impairs the sensitivity to instrumental contingency degradation. *Neuropsychopharmacology* 35: 1788–1796, 2010a.
- Lex B, Hauber W.** The role of dopamine in the prefrontal cortex and the dorsomedial striatum in instrumental conditioning. *Cereb Cortex* 20: 873–883, 2010b.
- Lucantonio F, Caprioli D, Schoenbaum G.** Transition from “model-based” to “model-free” behavioral control in addiction: involvement of the orbitofrontal cortex and dorsolateral striatum. *Neuropharmacology* 76B: 407–415, 2014.
- Maguire EA, Hassabis D.** Role of the hippocampus in imagination and future thinking. *Proc Natl Acad Sci USA* 108: E39, 2011.
- McGeorge AJ, Faul RL.** The organization of the projection from the cerebral cortex to the striatum in the rat. *Neuroscience* 29: 503–537, 1989.
- McNaughton BL, Battaglia FP, Jensen O, Moser EI, Moser MB.** Path integration and the neural basis of the “cognitive map.” *Nat Rev Neurosci* 7: 663–678, 2006.
- van der Meer MA, Johnson A, Schmitzer-Torbert NC, Redish AD.** Triple dissociation of information processing in dorsal striatum, ventral striatum, and hippocampus on a learned spatial decision task. *Neuron* 67: 25–32, 2010.
- van der Meer M, Kurth-Nelson Z, Redish AD.** Information processing in decision-making systems. *Neuroscientist* 18: 342–359, 2012.
- Miller GA.** The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychol Rev* 63: 81–97, 1956.
- Miyachi S, Hikosaka O, Lu X.** Differential activation of monkey striatal neurons in the early and late stages of procedural learning. *Exp Brain Res* 146: 122–126, 2002.
- Miyachi S, Hikosaka O, Miyashita K, Kárádi Z, Rand MK.** Differential roles of monkey striatum in learning of sequential hand movement. *Exp Brain Res* 115: 1–5, 1997.
- Mizumori SJ, Yeshenko O, Gill KM, Davis DM.** Parallel processing across neural systems: implications for a multiple memory system hypothesis. *Neurobiol Learn Mem* 82: 278–298, 2004.
- Muenzinger KF.** Vicarious trial and error at a point of choice. I. A general survey of its relation to learning efficiency. *J Genet Psychol* 53: 75–86, 1938.
- O’Keefe J, Nadel L.** *The Hippocampus as a Cognitive Map*. Oxford, UK: Clarendon, 1978.
- Packard MG, McGaugh JL.** Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. *Neurobiol Learn Mem* 65: 65–72, 1996.
- Papale AE, Stott JJ, Powell NJ, Regier PS, Redish AD.** Interactions between deliberation and delay-discounting in rats. *Cogn Affect Behav Neurosci* 12: 513–526, 2012.
- Powell NJ, Redish AD.** Complex neural codes in rat prefrontal cortex are stable across days on a spatial decision task. *Front Behav Neurosci* 8: 120, 2014.
- Ragozzino ME.** The contribution of the medial prefrontal cortex, orbitofrontal cortex, and dorsomedial striatum to behavioral flexibility. *Ann NY Acad Sci* 1121: 355–375, 2007.
- Ragozzino ME, Choi D.** Dynamic changes in acetylcholine output in the medial striatum during place reversal learning. *Learn Mem* 11: 70–77, 2004.
- Ragozzino ME, Jih J, Tzavos A.** Involvement of the dorsomedial striatum in behavioral flexibility: role of muscarinic cholinergic receptors. *Brain Res* 953: 205–214, 2002a.
- Ragozzino ME, Ragozzino KE, Mizumori SJ, Kesner RP.** Role of the dorsomedial striatum in behavioral flexibility for response and visual cue discrimination learning. *Behav Neurosci* 116: 105–115, 2002b.
- Redish AD.** *Beyond the Cognitive Map: From Place Cells to Episodic Memory*. Cambridge, MA: MIT Press, 1999.
- Redish AD.** *The Mind Within the Brain: How We Make Decisions and How Those Decisions Go Wrong*. Oxford, UK: Oxford Univ. Press, 2013.
- Schacter DL, Guerin SA, St Jacques PL.** Memory distortion: an adaptive perspective. *Trends Cogn Sci* 15: 467–474, 2011.
- Schmidt B, Papale A, Redish AD, Markus EJ.** Conflict between place and response navigation strategies: effects on vicarious trial and error (VTE) behaviors. *Learn Mem* 20: 130–138, 2013.
- Schmitzer-Torbert N, Redish AD.** 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, 2004.
- Schmitzer-Torbert NC, Redish AD.** Task-dependent encoding of space and events by striatal neurons is dependent on neural subtype. *Neuroscience* 153: 349–360, 2008.
- Schoenbaum G, Nugent SL, Saddoris MP, Setlow B.** Orbitofrontal lesions in rats impair reversal but not acquisition of go, no-go odor discriminations. *Neuroreport* 13: 885–890, 2002.
- Shiflett MW, Brown RA, Balleine BW.** Acquisition and performance of goal-directed instrumental actions depends on ERK signaling in distinct regions of dorsal striatum in rats. *J Neurosci* 30: 2951–2959, 2010.
- Smith KS, Graybiel AM.** A dual operator view of habitual behavior reflecting cortical and striatal dynamics. *Neuron* 79: 361–374, 2013.
- Stalnaker TA, Calhoun GG, Ogawa M, Roesch MR, Schoenbaum G.** Reward prediction error signaling in posterior dorsomedial striatum is action specific. *J Neurosci* 32: 10296–10305, 2012.
- Steiner AP, Redish AD.** The road not taken: neural correlates of decision making in orbitofrontal cortex. *Front Neurosci* 6: 131, 2012.
- Swanson LW.** Cerebral hemisphere regulation of motivated behavior. *Brain Res* 886: 113–164, 2000.
- Thorn CA, Atallah H, Howe M, Graybiel AM.** Differential dynamics of activity changes in dorsolateral and dorsomedial striatal loops during learning. *Neuron* 66: 781–795, 2010.

- Tolman EC.** The determiners of behavior at a choice point. *Psychol Rev* 45: 1–41, 1938.
- Torregrossa MM, Quinn JJ, Taylor JR.** Impulsivity, compulsivity, and habit: the role of orbitofrontal cortex revisited. *Biol Psychiatry* 63: 253–255, 2008.
- Wikenheiser AM, Redish AD.** Changes in reward contingency modulate the trial-to-trial variability of hippocampal place cells. *J Neurophysiol* 106: 589–598, 2011.
- Wikenheiser AM, Redish AD.** Decoding the cognitive map: ensemble hippocampal sequences and decision making. *Curr Opin Neurobiol* 32: 8–15, 2015a.
- Wikenheiser AM, Redish AD.** Hippocampal theta sequences reflect current goals. *Nat Neurosci* 18: 289–294, 2015b.
- Wood ER, Dudchenko PA, Robitsek RJ, Eichenbaum H.** Hippocampal neurons encode information about different types of memory episodes occurring in the same location. *Neuron* 27: 623–633, 2000.
- Yeshenko O, Guazzelli A, Mizumori SJ.** Context-dependent reorganization of spatial and movement representations by simultaneously recorded hippocampal and striatal neurons during performance of allocentric and egocentric tasks. *Behav Neurosci* 118: 751–769, 2004.
- Yin HH, Knowlton BJ.** Contributions of striatal subregions to place and response learning. *Learn Mem* 11: 459–463, 2004.
- Yin HH, Knowlton BJ.** The role of the basal ganglia in habit formation. *Nat Rev Neurosci* 7: 464–476, 2006.
- Yin HH, Knowlton BJ, Balleine BW.** Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur J Neurosci* 19: 181–189, 2004.
- Yin HH, Knowlton BJ, Balleine BW.** Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning: striatum and action. *Eur J Neurosci* 22: 505–512, 2005a.
- Yin HH, Ostlund SB, Knowlton BJ, Balleine BW.** The role of the dorsomedial striatum in instrumental conditioning: striatum and instrumental conditioning. *Eur J Neurosci* 22: 513–523, 2005b.
- Zhang K, Ginzburg I, McNaughton BL, Sejnowski TJ.** Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells. *J Neurophysiol* 79: 1017–1044, 1998.

