Behavioral/Cognitive

Cholinergic Interneurons Use Orbitofrontal Input to Track Beliefs about Current State

Thomas A. Stalnaker,[1] Ben Berg,[1] Navkiran Aujla,[1] and Geoffrey Schoenbaum[1,2,3]

[1]National Institute on Drug Abuse Intramural Research Program, Cellular Neurobiology Research Branch, Behavioral Neurophysiology Research Section, Baltimore, Maryland 21224, [2]Department of Anatomy and Neurobiology, University of Maryland, School of Medicine, Baltimore, Maryland 21201, and [3]Solomon H. Snyder Department of Neuroscience, Johns Hopkins University, Baltimore, Maryland 21218

When conditions change, organisms need to learn about the changed conditions without interfering with what they already know. To do so, they can assign the new learning to a new "state" and the old learning to a previous state. This state assignment is fundamental to behavioral flexibility. Cholinergic interneurons (CINs) in the dorsomedial striatum (DMS) are necessary for associative information to be compartmentalized in this way, but the mechanism by which they do so is unknown. Here we addressed this question by recording putative CINs from the DMS in rats performing a task consisting of a series of trial blocks, or states, that required the recall and application of contradictory associative information. We found that individual CINs in the DMS represented the current state throughout each trial. These state correlates were not observed in dorsolateral striatal CINs recorded in the same rats. Notably, DMS CIN ensembles tracked rats' beliefs about the current state such that, when states were miscoded, rats tended to make suboptimal choices reflecting the miscoding. State information held by the DMS CINs also depended completely on the orbitofrontal cortex, an area that has been proposed to signal environmental states. These results suggest that CINs set the stage for recalling associative information relevant to the current environment by maintaining a real-time representation of the current state. Such a role has novel implications for understanding the neural basis of a variety of psychiatric diseases, such as addiction or anxiety disorders, in which patients generalize inappropriately (or fail to generalize) between different environments.

Key words: cholinergic; orbitofrontal; rat; single unit; striatum

---

Significance Statement

Striatal cholinergic interneurons (CINs) are thought to be identical to tonically active neurons. These neurons have long been thought to have an important influence on striatal processing during reward-related learning. Recently, a more specific function for striatal CINs has been suggested, which is that they are necessary for striatal learning to be compartmentalized into different states as the state of the environment changes. Here we report that putative CINs appear to track rats' beliefs about which environmental state is current. We further show that this property of CINs depends on orbitofrontal cortex input and is correlated with choices made by rats. These findings could provide new insight into neuropsychiatric diseases that involve improper generalization between different contexts.

Introduction

Cholinergic interneurons (CINs) in the dorsal striatum, thought to be identical to tonically active neurons (TANs; Inokawa et al.,

2010), comprise a small population (~1% of all striatal neurons; Graveland and DiFiglia, 1985; Oorschot, 2013) that has a wide influence on striatal processing (Kawaguchi et al., 1995). CINs have long been known to respond with a transient pause in activity to stimuli that predict reward or other motivationally relevant outcomes (Kimura et al., 1984; Aosaki et al., 1994; Apicella, 2002; Benhamou et al., 2014). More recent evidence has suggested that outcome-related CIN activity may be modulated by context (Apicella, 2007). For example, CIN activity varies with reward schedule (Shimo and Hikosaka, 2001), action requirements of the current block of trials (Lee et al., 2006), motivational context

(Yamada et al., 2004), or spatial location of expected reward and/or required movement (Shimo and Hikosaka, 2001; Ravel et al., 2006). Based on these findings, it has been hypothesized that CINs (or TANs) signal the "motivational contexts of actions" (Yamada et al., 2004) and may participate in the initiation of actions appropriate for the context (Apicella, 2007).

At the same time, acetylcholine has been hypothesized to modulate plasticity (Calabresi et al., 2000; Wang et al., 2006) and gate corticostriatal inputs (Ding et al., 2010) to reduce interference between old learning and new (Hasselmo and Bower, 1993; Ashby and Crossley, 2011). Consistent with this proposal, interference with cholinergic function in the striatum often causes problems with flexible behavior (Tzavos et al., 2004; Ragozzino et al., 2009; Okada et al., 2014; Aoki et al., 2015). In one such study, Bradfield et al., 2013 used a series of well controlled behavioral tasks to demonstrate that interfering with CIN activity in the posterior dorsomedial striatum (DMS) specifically results in confusion between old learning and new after response-outcome contingencies have been changed. These results suggest that dorsomedial striatal CINs contribute to compartmentalizing associative information related to different states of the environment. However, nothing is known about how CIN firing properties confer this function (e.g., by signaling changes in state, state prediction errors, or by some other mechanism) or how CINs gain access to information about the current state of the environment. Recently, the orbitofrontal cortex (OFC), a prefrontal area, has been proposed to provide a readout of the current location within state space, a function that becomes critical when external stimuli do not define states (Wilson et al., 2014). Thus, one possibility is that state information used by CINs would depend on direct or indirect input from the OFC. Here we addressed this question by recording putative CINs from the DMS in sham and OFC-lesioned rats performing a task consisting of a series of trial blocks, or states, that required the recall and application of contradictory associative information.

Materials and Methods

Subjects. Nine male Long–Evans rats, weighing 175–200 g, were obtained from Charles River Laboratories. Rats were tested at the National Institute on Drug Abuse Intramural Research Program in accordance with School of Medicine and National Institutes of Health guidelines. During testing, rats were given *ad libitum* access to water for 10 min/d after testing. All testing was performed during the light phase.

Behavioral task. Training and recording was conducted in aluminum chambers equipped with a custom-made odor port with two fluid wells beneath it, as described previously (Stalnaker et al., 2010). The fluid wells were connected to fluid delivery lines containing flavored milk (Nesquik brand chocolate or vanilla) diluted 50% with water. Photobeam breaks at the port and wells were monitored and recorded with strobes sent to the recording system.

Rats were trained extensively before electrodes were implanted (see below) and then briefly retrained after implantation to acclimate them to the recording cable. Pre-recording training sessions generally continued until the rat failed to perform a trial for at least 5 min. This resulted in sessions ranging from 150 to 250 trials. This initial shaping phase gradually introduced all elements of the task (described below). Recording was begun when rats could complete five blocks of trials (at least 260 trials) with the recording cable attached.

Each recording session consisted of a series of self-paced trials organized into five blocks. Each trial consisted of the following sequence. The rat initiated a trial by poking into the odor port while the house light was illuminated. Beginning 500 ms after the odor poke, an odor would be delivered for 500 ms. At the end of the odor, the rat could withdraw from the odor port and respond by poking at either the left or right fluid well within 3000 ms. After responding at a fluid well, the rat was required to wait at the well for 500 ms before fluid delivery began. Fluid was delivered in ~0.05 ml boluses, with delivery times separated by 500 ms when multiple boli were delivered (see below). When the rat finished drinking and withdrew from the fluid port, the house light turned off and the trial terminated. If a rat withdrew early at any point in the trial before reward delivery, the house light turned off and the trial terminated.

The identity of the odor delivered on a particular trial specified whether the rat could receive reward at the left well (forced-choice left), the right well (forced-choice right), or either well (free choice). If it responded at the un-rewarded well on forced-choice trials, the house light turned off, the trial terminated, and the subsequent trial delivered the same odor. The identity of the three instructional odors and the response required after each odor remained the same across the entire experiment (training and recording). Odors were presented in a pseudorandom sequence such that the free-choice odor was presented on 35% of trials and the left/right odors were presented in equal numbers on the remaining 65% of trials (±1 over 250 trials). In addition, the same odor could be presented on no more than three consecutive trials.

Rewards consisted of either one or three boli of chocolate or vanilla milk. Response–reward contingencies were consistent within each block of trials, such that the same reward would be delivered for every correct right response, either free- or forced-choice, and a different reward would be delivered for every correct left response, free- or forced-choice. The reward schedule was arranged so that, in each block, reward features available on one side were always paired with the opposite reward features on the other side; thus, when one drop of chocolate milk was available on the left, three drops of vanilla was available on the right, etc., resulting in a total of four different reward combinations. On the first block, which was shorter and used to establish baseline behavior, one of these combinations was chosen randomly. The subsequent four block transitions then followed, in order: (1) a drop-number transition, in which the side with one drop changed to three drops and vice versa, but the side-flavor contingencies remained the same; (2) a flavor transition, in which the side with chocolate changed to vanilla and vice versa, but the side-number contingencies remained the same; (3) another drop-number transition; and (4) another flavor transition. These block transitions were not explicitly signaled; in the last three transitions, they occurred after a minimum of a randomly chosen number of correct trials between 54 and 66, with the additional proviso that rats had to be choosing the larger side at least 6 of the last 10 trials before the switch (rats were choosing the larger side at the end of the designated number of trials anyway on >90% of blocks). The first transition occurred after at least 20 correct trials with the same proviso as above.

Surgical procedures and histology. Surgical procedures followed guidelines for aseptic technique as described previously (Stalnaker et al., 2010). Neurotoxic lesions of left OFC were made using intracerebral infusions of NMDA (12 μg/μl in sterile saline; Sigma) with a glass micropipette and a picospritzer at four sites: 100 nl at +4.0 mm anteroposterior (AP), 3.7 mm mediolateral (ML), and 3.8 mm ventral (V); 100 nl at +4.0 mm AP, 2.2 mm ML, and 3.8 mm V; 100 nl at +3.0 mm AP, 4.2 mm ML, and 5.2 mm V; and 50 nl at +3.0 mm AP, 3.2 mm ML, and 5.2 mm V (coordinates are relative to bregma and the skull surface). In the same surgery, electrodes consisting of drivable bundles of eight 25-μm-diameter NiCr wires (Stablohm 675; California Fine Wire) electroplated with platinum to an impedance of ~300 kΩ, were implanted in the left DMS (−0.4 mm AP, 2.6 mm ML, and 3.5–4.0 mm V to start; coordinates are relative to bregma and the surface of the dura). Electrodes were also implanted in the dorsolateral striatum (DLS) in the same hemisphere of these rats (+0.7 mm AP, 3.6 mm ML, and 3.5 to 4.0 mm V to start). At the end of the study, the final electrode position was marked by passing current through electrodes, the rats were killed with an overdose of isoflurane and perfused with saline, and the brains were removed and processed using standard techniques.

Single-unit recording. Procedures were the same as described previously (Stalnaker et al., 2010). Wires were screened for activity daily; if no activity was detected, the rat was removed and the electrode assembly was advanced 80 μm. Otherwise, a session was conducted, and the electrode was advanced by at least 80 μm at the end of the session. Neural activity was recorded using Plexon Multichannel Acquisition Processor systems,

interfaced with odor discrimination training chambers. Signals from the electrode wires were amplified $20\times$ by an operational amplifier head stage (HST/8o50-G20-GR; Plexon), located on the electrode array. Immediately outside the training chamber, the signals were passed through a differential preamplifier (PBX2/16sp-r-G50/16fp-G50; Plexon), in which the single-unit signals were amplified $50\times$ and filtered at 150–9000 Hz. The single-unit signals were then sent to the multichannel acquisition processor box, in which they were further filtered at 250–8000 Hz, digitized at 40 kHz, and amplified at 1–$32\times$. Waveforms ($>2.5{:}1$ signal-to-noise ratio) were extracted from active channels and recorded with event timestamps sent by the behavioral program.

Data analysis. Units were sorted using Offline Sorter software from Plexon), using a template matching algorithm. Sorted files were then processed in Neuroexplorer to extract unit timestamps and relevant event markers. These data were subsequently analyzed in MATLAB (MathWorks).

For behavioral analyses, we examined choice rates toward the three-drop side on free-choice trials occurring within the last 20 correct trials before block transitions versus those occurring within the first 20 correct trials after block transitions. To make plots of trial-by-trial choice rates (line plots), we first aligned rewarded trials before and after block transitions in all sessions. Then for each trial position relative to the transition (first, second, third, etc., before and after the block transition), we took the proportion of free choices toward the side that delivered the three-drop reward after the transition, ignoring sessions in which a forced-choice trial occurred at that position. The resulting average trial-by-trial choice proportion was smoothed using a three-bin boxcar separately before and after the block switch.

To categorize putative medium spiny neurons (MSNs), we examined, for each single unit, valley-to-peak width, half-valley width, and average baseline firing rate during the intertrial intervals of all trials. We used the MATLAB *k*-means function with the "city" parameter for "distance" to define three clusters, which separated a cluster with low firing rate and wide waveforms (putative MSNs) and another with high firing rates and narrow waveforms (putative fast-spiking interneurons) and a third in between (undefined). There is no widely accepted method for distinguishing putative CINs from MSNs in rodents. To perform this separation, we examined, for each single unit, mean "CV2," which is a variation on the conventional coefficient of variation (Holt et al., 1996) that has been found to distinguish striatal CINs from other neurons in a juxtacellular recording study in which neurons were immunohistochemically identified (Sharott et al., 2012). The conventional coefficient of variation is defined as the SD divided by the mean of all the interspike intervals across a spike train. The mean CV2 instead calculates the variance between each pair of succeeding interspike intervals and then averages this value across all such pairs across the spike train. Thus slow changes in firing rate across a session have less influence on mean CV2 than on the standard coefficient of variation:

$$CV2 = \text{mean}(2 \times \text{abs}(\text{ISI}_{i+1} - \text{ISI}_i)/(\text{ISI}_{i+1} + \text{ISI}_i)),$$

for $i = 1$ to n, with n indicating the number of interspike intervals across the spike train. We then assigned a unit as a putative CIN if its mean CV2 value was <0.8 and its mean baseline firing rate was <8.0 spikes/s.

For the bin-by-bin block-selectivity indices shown in Figures 3, *a* and *b*, and 9, *a* and *b*: These analyses used all forced-choice trials occurring after the first 20 correct trials (of any kind) in each of the last four blocks of sessions. Forced-choice trials were used because these trials are equally balanced between choices toward each side, and the latter parts of block were used because that was when behavior was stable. Activity on each trial was binned (100 ms) and baseline subtracted (baseline firing rate is the average firing rate across the entire intertrial interval before that trial). The activity of each neuron was peak normalized by dividing all binned firing rates by the firing rate in the 500 ms bin with the highest average firing rate among the following 16 conditions: first 10 or last 10 trials of each of the eight block/direction conditions. For each neuron, we then identified a preferred direction and block based on which condition (4 blocks $\times$ 2 directions) had the highest average firing rate during the choice movement epoch (from odor port exit to fluid port entry) across

the last 10 trials of that condition. We chose this epoch because it was the first time in the trial at which behavior differed between blocks and because we have observed strong movement-related activity in striatal neurons in previous studies using a similar task (Stalnaker et al., 2010). Then we averaged the binned normalized firing rates of the three non-preferred blocks, bin-by-bin, in each direction and subtracted these values from the binned normalized firing rates on the preferred block in each direction. This yielded a bin-by-bin index of how many more neurons fired on the preferred block compared with all other blocks. We tested for significance compared with zero by running a *t* test on this index, using a sliding five-bin average across all neurons in the population and requiring five consecutive sliding five-bin averages to be significant at the designated *p* value (see below). To correct for multiple comparisons, we re-ran the analysis with the condition of all trials shuffled, 1000 times, and then calculated the *p* value at which 5% of the 1000 replicates found at least one instance of significance. That *p* value served as the designated *p* value for that population of neurons ($p = 0.014$ for DMS sham CINs, and $p = 0.021$ for DMS lesion CINs, $p = 0.020$ for DLS sham CINs, and $p = 0.031$ for DLS lesion CINs). In the figures, the bin-by-bin average indices are shown aligned to different events in the trial. The events in the figures are separated by the average time that separated them across the sessions in which those neurons were recorded.

For the block-selectivity indices across particular trial epochs, we calculated the indices as above but across trial epochs instead of 100 ms bins. The trial epochs were as follows: 1500 ms immediately before odor start, 1500 ms immediately after odor start (which encompasses movement), 1500 ms immediately after reward start, and 1500 ms after reward start to 1000 ms after trial end. Indices were calculated for each neuron at each epoch, and then Pearson's correlation coefficients were derived for each pair of epochs across neurons.

For the neural decoding analysis, we used MATLAB code from the Neural Decoding Toolbox (www.readout.info; Meyers, 2013), in some cases modified. For all analyses, we used the zscore_normalize_FP function to preprocess the neural data and the max_correlation_coefficient_CL as the classifier, and we used pseudo-ensembles (meaning that neurons were not recorded in the same sessions). For each resample run, the cross validator selected *Z* neurons from the entire population being tested, with *Z* being the ensemble size. The cross validator then split the data into *n* splits, which means it took *n* trials from each of the conditions to be decoded, for each neuron. It then trained the classifier using $n - 1$ of them and tested the remaining one trial from each condition. Each test trial was classified according to the training condition with which it had the highest correlation across the ensemble of neurons. This process was repeated *n* times such that each trial served once in the test set. Then the ensemble was reselected and the entire process was repeated for *Y* resample runs. *Y* was 50 for the basic block decoder described below and 500 for the suboptimal choice analysis described subsequently. *n* was selected to be as large as possible but still include $>90\%$ of sessions.

For each analysis, we performed a bin-by-bin decoding analysis using 25-neuron ensembles (or the maximum possible number of neurons for analyses in which there were <25 neurons from usable sessions) and an epoch analysis using a range of ensemble sizes from 2 to 25 (or from two to the maximum possible) on three epochs at different points in the trial. To determine significance of the bin-by-bin analysis compared with chance, we repeated the decoding analysis 100 times with shuffled conditions to derive a null distribution of decoding percentages at each bin. Decoding percentage in bins had to meet $p < 0.05$ (one-tailed) and be a part of five consecutive significant bins for the block decoding analysis and 10 consecutive significant bins for the choice-direction decoding analysis. To determine significant differences between decoding in different neural populations in the epoch analysis, we fit a log function (see below) to each observed function of ensemble size, using the MATLAB function *nlinfit*:

$$y = (B_1 \times \ln(x)) + B_2,$$

where *y* is the decoding percentage, *x* is the ensemble size, and B_1 and B_2 were parameters. We then used the MATLAB function *nlparci* to obtain

confidence intervals for the B_1 parameter, which controlled the rise of the function. We found the smallest α level that resulted in non-overlapping confidence intervals between fitted parameters from different populations and then took α^2 as the p value for the comparison between the two. The criterion for significance was set at $p < 0.01$ to correct for the three comparisons (three epochs) for each pair of populations.

The block decoding analysis trained using an epoch at a different point in the trial than the bins it tested. Therefore, this analysis tested for the consistency of the block code across the trial. The testing bins were 500 ms that slid up by 100 ms at each step across the trial. Trial data were first baseline subtracted and aligned at different trial events and were concatenated so that the time between events was equal to the average time between those events in the sessions in which that population was recorded. The test epoch for the first half of the bins occurred at the end of the trial, 1500 ms beginning 500 ms after the beginning of reward delivery. The test epoch for the second half of the bins occurred at the start of the trial, 1500 ms immediately before odor port withdrawal (i.e., choice initiation). For the decoding function of ensemble size, test epochs were the 1000 ms immediately before the odor port withdrawal, the 1000 ms immediately after odor port withdrawal, and 1000 ms immediately after reward delivery.

The block decoding analysis of suboptimal choices followed the same procedures as above, except that only free choices of the small reward occurring after the first 10 correct trials after a number block transition served as the test trials. Thus, only sessions that had at least one such trial in each block could be used. The training set consisted of all correct forced-choice trials excluding those within the first 10 correct trials after number block transition. The control analysis tested only free choices of the large reward occurring after the first 10 correct trials after a number block transition but had the same training set. The rationale for using this training set is that forced-choice trials were equally distributed between choices of the big and small rewards, and hence the set was unbiased with regard to outcome. Thus, we could compare how well this decoder performed on free choices of the small reward with how well it performed with free choices of the large reward, using the same training set.

The block decoding analysis of MSNs yoked to the decoding of CINs followed the same procedures as above, except that, for each ensemble of CINs, the corresponding ensemble of MSNs that was recorded in the same sessions as the CINs was used. The training sets for these yoked MSN ensembles was the same as above (i.e., all correct forced-choice trials excluding those within the first 10 correct trials after a number block transition), but the training epoch was always the same as the testing epoch, which differed from the previous analyses. The rationale for this change was that MSNs in the initial analysis did not appear to show block selectivity that persisted across the trial. Thus, this analysis tests whether they have block selectivity at any one epoch in the trial.

The choice-direction, odor identity, and drop-number decoding analysis trained on the same bins and epochs on which it tested. For the bin-by-bin analysis, we used 200 ms bins advancing by 50 ms across the trial. For the decoding function of ensemble size, test epochs were from 1000 ms before odor delivery to the end of odor delivery, the 1000 ms immediately after odor port withdrawal, and 1000 ms immediately after reward delivery.

Statistics were done using MATLAB, Excel, and Statistica. Planned comparisons were used for testing specific effects of multi-way ANOVAs.

Results

We recorded single-units from the DMS and DLS as rats performed an odor-guided choice task in which blocks of trials defined states that required the recall and application of contradictory associative information. The task, illustrated in Figure 1a, consisted of a series of five trial blocks presented across a single session. Each trial block was made up of choice trials in which a particular milk reward (defined by the number of drops, one or three, and the flavor, chocolate or vanilla) was delivered for a correct left response and a different milk reward for a correct right response. Unsignaled block switches resulted in changes to either the number of drops or the flavor (but never both in a

particular switch) of the rewards available at each well. Block switches were arranged such that each of the last four blocks of sessions had a unique set of response–outcome contingencies, defining a state. Trials were initiated with an odor that signaled whether the trial would be free choice (35% of trials), meaning that either well would yield reward, or forced choice (65% of trials), meaning that only one of the two wells would yield reward. After initial shaping on the task, rats received unilateral sham ($n = 5$) or neurotoxic ($n = 4$; Fig. 1b) lesions of the lateral OFC, and electrode bundles were implanted in the DMS and DLS ipsilateral to the lesions. OFC lesions targeted the ventral and lateral orbital areas and the dorsal and ventral agranular insular regions. This target region includes areas on the dorsal bank of the rhinal sulcus that receive olfactory input from the piriform cortex (Cinelli et al., 1985; Price et al., 1991) and more laterally located insular regions that have direct interactions with the basolateral amygdala (Krettek and Price, 1977; Kita and Kitai, 1990; Shi and Cassell, 1998), while avoiding gustatory regions located in agranular insular cortex posterior to the genu of the corpus callosum (Saper, 1982; Kosar et al., 1986; Krushel and van der Kooy, 1988).

Behavior in the task

During recording, number switches resulted in a rapid and sustained change in choice rate on free-choice trials, which was independent of flavor and lesion status (Fig. 1c,d). Changes in behavior driven by reward number (and the lack thereof for flavor) were similar for the two milk flavors for each individual rat (see scatter plots in Fig. 1c,d). To test this, we performed a mixed ANOVA on the difference in choice rates across block transitions (from the last 20 trials of the previous block to the first 20 of the new block), with transition type (drop number or flavor) as a within-subjects factor and initial flavor (chocolate or vanilla) and group (sham or OFC lesion) as between-subjects factors. This ANOVA showed no effects of group ($F_{(1,194)} = 1.8$; $p = 0.18$) or initial flavor ($F_{(1,194)} = 0.9$; $p = 0.34$). Planned comparisons showed no interaction between group and either the number ($F_{(1,194)} = 2.8$; $p = 0.10$) or flavor transition effect ($F_{(1,194)} = 0.005$; $p = 0.95$).

Neural recordings and cell-type separation

Recordings yielded a total of 1331 well-isolated single units in the DMS, with 538 in shams and 793 with ipsilateral OFC lesions, and 1383 in the DLS, with 624 in sham rats and 759 with ipsilateral OFC lesions (Fig. 1e for recording locations). We categorized putative MSNs using an established cluster analysis that used waveform and firing rate parameters (Thorn and Graybiel, 2014; Fig. 2a, left). There is no widely accepted method for distinguishing putative CINs from MSNs in rodents, and waveform criteria are inadequate for making this distinction (Sharott et al., 2012; Thorn and Graybiel, 2014). Although in theory optogenetic tools could be used to identify CINs, in our hands the Chat–Cre rats that have been produced to date (Witten et al., 2011) do not express Cre in the medial dorsal striatum based on (1) *in situ* hybridization analysis for Cre mRNA and (2) adeno-associated virus-mediated, Cre-dependent GFP transgene expression analysis (data available on request). Therefore, we defined CINs in our dataset based on published juxtacellular recording work in which CINs were histochemically identified (Sharott et al., 2012), as those with an average spike-to-spike CV2 (Holt et al., 1996) of <0.8 and an average baseline firing rate of <8.0 spikes/s. These criteria effectively distinguished CINs ($n = 109$ across groups) from separately defined MSNs, with only 4.6% of the MSN clus-

ter falling below the CIN criteria (Fig. 2a, right). CINs also tended to have wide waveforms even though waveform characteristics were not used to define them (Fig. 2b). The prevalence of putative CINs among all recorded neurons was also not different in shams versus lesions in the DMS (5.2% in shams vs 4.5% in OFC-lesioned rats), but it was slightly lower in the DLS (shams, 2.6%; lesions, 3.8%). In addition, the mean baseline firing rate, mean CV2 value, and mean waveform duration of CINs were unaffected by ipsilateral OFC lesions (t tests comparing these parameters in the DMS: CV2, $t_{(62)} = -0.3$, $p = 0.76$; average baseline firing rate, $t_{(62)} = 0.8$, $p = 0.40$; waveform duration, $t_{(62)} = -1.7$, $p = 0.09$; in the DLS: CV2, $t_{(43)} = -1.3$, $p = 0.20$; average baseline firing rate, $t_{(43)} = -1.2$, $p = 0.24$; waveform duration, $t_{(43)} = -1.1$, $p = 0.27$).

CIN activity reflects the current block, or state, in the task

To test the degree to which CINs were responsive to state, we first calculated a bin-by-bin block-selectivity index across the trial, separating the preferred direction of each neuron from the anti-preferred direction [see Materials and Methods; DMS CINs were evenly split between those preferring the ipsilateral ($n = 14$) and contralateral ($n = 14$) sides]. This index was based on the difference between normalized activity on the preferred block of each neuron and that on the nonpreferred blocks in a given direction. The preferred block and direction were defined by which block and direction had maximum activity during the choice-movement epoch (from odor port exit to fluid port entry). This index used forced-choice trials occurring after the first 20 correct trials in each block. Averaged across the entire population of CINs, the index in the preferred direction was significantly greater than zero from near the beginning of the trial to the end (Fig. 3a). To test whether the current block was represented at the level of individual single units, we tested the correlations between block-selectivity indices (in the preferred direction) calculated for four individual epochs across the trial for each neuron. Single-unit indices were significantly correlated between all pairs of epochs (Fig. 3c, right; average $R^2 = 0.51$, all p values <0.0001; Table 1), including between the very first pre-odor epoch and the very last post-reward epoch (Fig. 3c, left). Sustained block selectivity across individual single units could not have resulted from particular units re-
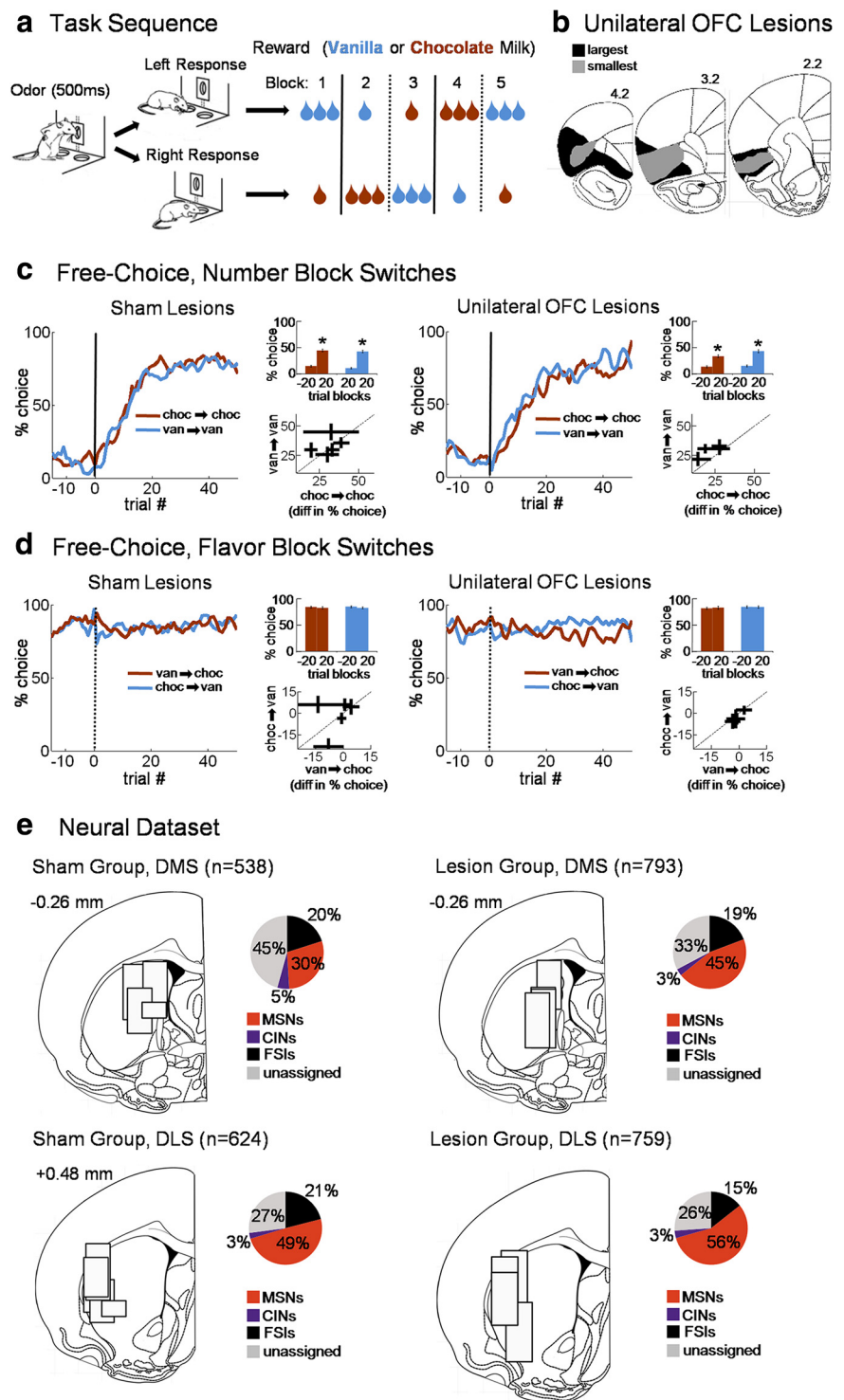

Figure 1. The task, lesions, behavior, and recording locations. The task (**a**) had four different blocks, or *states*, defined by the set of available response–reward contingencies. Trials began with an instructional odor, indicating a free-choice or forced-choice, after which rats responded at one of the two fluid wells for 1 or 3 drops of chocolate or vanilla milk. Reward contingencies were stable across blocks of ~60 trials, but at unsignaled transitions the number of drops or flavor changed on both sides (only 1 of the 4 possible block sequences is shown). Unilateral neurotoxic lesions of orbitofrontal cortex (OFC) (**b**) were made in one group of rats (numbers are millimeters from bregma). Groups with sham lesions or unilateral lesions of OFC were similarly sensitive to changes in the number of drops (**c**), and similarly insensitive to changes in the flavor (**d**). Line figures show average trial-by-trial choice rates across transitions; bar graphs summarize these data by showing average choice rates in the last 20 trials of the previous block versus the first 20 of the new block; scatter plots show rat-by-rat difference scores (choice rate after minus choice rate before) with length of lines showing SEs. **e**, The approximate locations of recordings and proportions of putative cell types in each of the four groups. The width of each box represents 1 mm. FSI, Fast-spiking interneuron. *$p < 0.001$.

a Cell Type Separation
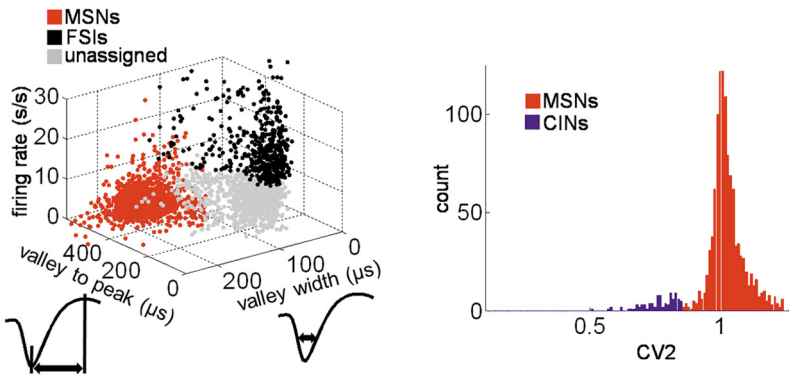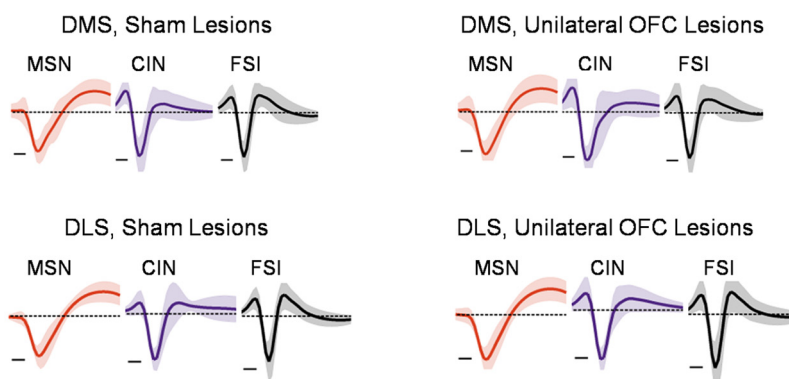

b Average Waveforms

Figure 2. Cell-type separation. ***a***, MSNs were first separated from other neurons [including fast-spiking interneurons (FSIs)] using a three-dimensional cluster analysis on all recorded units (left). Subsequently, putative CINs were selected by taking the average spike-to-spike CV2 of every unit's entire spike-train. The criteria for CINs was CV2 <0.85 and average firing rate <8 spikes/s, which were based on measurements from histochemically confirmed CINs reported in published juxtacellular recordings. These criteria effectively discriminated CINs from separately defined MSNs (right; CV2 distribution, low-firing rate MSNs versus putative CINs). The upper limit of the CIN criterion is outside of the 1-a confidence interval for low-firing rate MSNs with a = 0.028. ***b***, Resulting average waveforms for each cell type in each recording group (the dotted line represents the 0 voltage level. Scale bar, 100 ms. Shading represents SE). Note that CINs were not selected based on waveform, whereas MSNs and FSIs were. FSIs were not analyzed for this report.

block to predict from which block a different set of unknown trials had come, based on which training set was maximally correlated with each test trial. This training-testing procedure was repeated across all splits of the data. This analysis used all correct trials, both free and forced choice in either direction, excluding the first 20 correct trials in each block to avoid trials during the switch in behavior. To test for the consistency of the block code across the trial, the decoder was trained using a different part of the trial than that used for testing, with a training epoch late in the trial used to test sliding epochs early in the trial, and a training epoch early in the trial used to test sliding epochs late in the trial. Using this approach, we found that block decoding in CINs became significant soon after lights on and continued to be significant through the end of the reward period (Fig. 4). CIN decoding performance was significantly better in the early, pre-choice period of the trial than the 25-neuron MSN ensembles, suggesting that CINs were the leading source of state information in the striatum (for all decoding statistics, see Table 3).

Miscoding of the current state is associated with suboptimal choices

Goal-directed choices are in theory based on knowledge of the current set of available response–outcome contingencies, which are associated with the current state. Thus, signaling the current state might be useful in recalling current contingencies and driving the appropriate choice. If CIN state encoding were part of this process, then miscoding of the current state by CINs should predict inappropriate or suboptimal choices. We tested for this by constructing a block decoder that tested only free choices of the small reward, i.e., suboptimal choices, excluding those in the first 10 trials after number switches. As the training set for this decoder, we used all correct forced-choice trials (also excluding the first 10 after number switches) because these trials were evenly distributed between receipt of the big and small rewards available in each block and hence would be unbiased with regard to outcome.

The results of this analysis, shown in Figure 5a, demonstrate that block decoding by CINs on suboptimal choices was not only inaccurate, it was actually significantly below chance, even early in the trial before the choice was made. This below-chance encoding was a function of ensemble size: the more neurons that were included in the ensemble, the further below chance decoding became. As shown in Figure 5b, the below-chance decoding on suboptimal trials occurred because CINs tended to miscode the block as if they had the opposite arrangement of drop numbers but the same flavors (e.g., if the correct block delivered one drop vanilla left/three drops chocolate right, the opposite number block would be three drops vanilla left/one drop chocolate
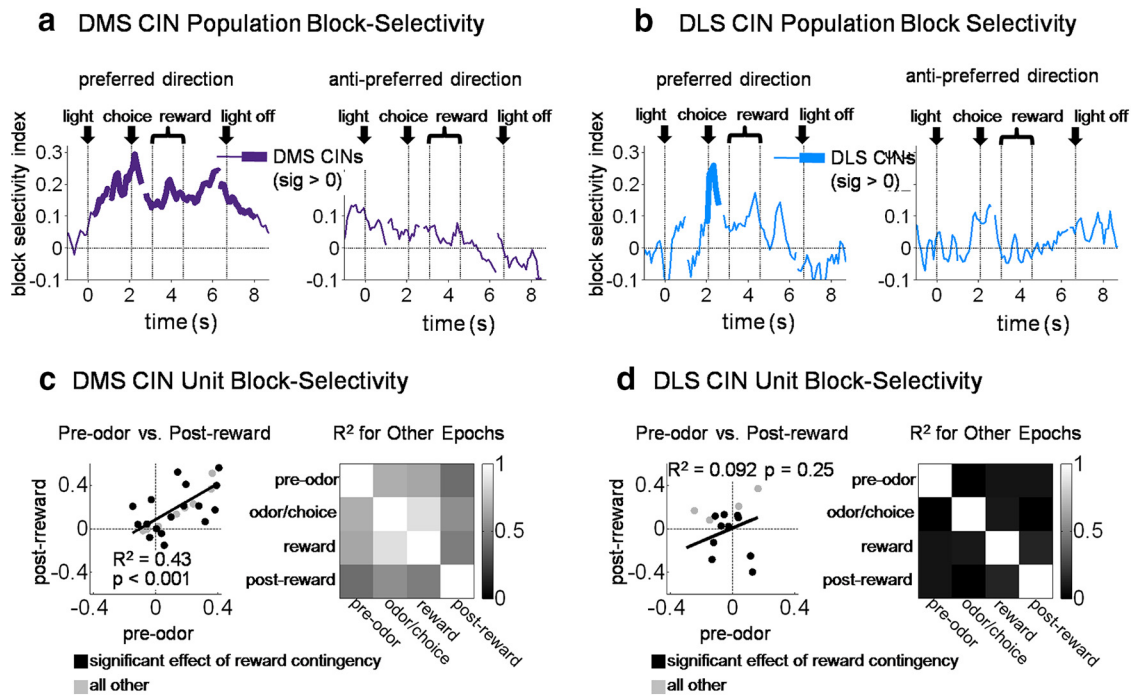
sponding to particular trial events, because no individual event defined a block. Thus, block selectivity encoded by individual CINs represented information that was synthesized across trials, i.e., information best described as the current state of the environment.

Block selectivity appeared to be specific to CINs recorded in the DMS. As shown in Figure 3b, in the 16 putative CINs recorded in the DLS, we only observed significant block-selective activity during the movement epoch, which is the epoch used to define the preferred block and direction. Both before and after this movement epoch, block selectivity was not significantly different from zero in the DLS CIN population. Furthermore, as shown in Figure 3d, block-selectivity indices of DLS CINs taken individually were weakly and mostly insignificantly correlated between epochs occurring across the trial (average $R^2 = 0.11$; only one of the six pairs, reward vs post-reward, had a $p < 0.05$; Table 2).

To test whether DMS CIN block selectivity actually reflected information about the current block and to examine state information held by CINs more closely, we constructed a decoding algorithm for current block and tested 25-neuron pseudo-ensembles randomly selected from all CINs (see Materials and Methods). The decoder used sets of training trials from each

a DMS CIN Population Block-Selectivity

b DLS CIN Population Block Selectivity

c DMS CIN Unit Block-Selectivity

d DLS CIN Unit Block-Selectivity

Figure 3. DMS CINs track the current state (i.e., block) across trials. Plots in *a* and *b* show a bin-by-bin average index of block-selectivity for the entire DMS (*n* = 26) and DLS (*n* = 16) CIN populations, respectively, averaged by each neuron's preferred direction (defined as direction with highest average firing rate during the choice movement; left) or anti-preferred direction (right). Only correct forced-choice trials after the first 20 correct trials in blocks were included in these averages. Bins were aligned to multiple trial events separated by the average time between them. Block-selectivity scores of individual DMS CINs were strongly correlated between different epochs across the trial (*c*), but this was not true of DLS CINs (*d*). *c*, *d*, Left, Scores for pre-odor epoch versus post-reward epoch; other pairs of epochs are summarized on the right. All R^2 values for DMS CINs were significant at $p < 0.001$, average $R^2 = 0.61$. For DLS CINs, none of thx R^2 values were significant at $p < 0.05$ (*p* values > 0.13); average $R^2 = 0.07$.

Table 1. Correlations between block-selectivity indices in pairs of epochs in CINs recorded in DMS sham rats (related to Fig. 3*c*)

Epochs	Before odor	Odor/choice	Reward
Odor/choice	$R^2 = 0.45$ $p < 0.0001$		
Reward	$R^2 = 0.49$ $p < 0.0001$	$R^2 = 0.62$ $p < 0.0001$	
After reward	$R^2 = 0.47$ $p < 0.0001$	$R^2 = 0.46$ $p < 0.0001$	$R^2 = 0.54$ $p < 0.0001$

Table 2. Correlations between block-selectivity indices in pairs of epochs in CINs recorded in DLS in sham rats (related to Fig. 3*d*)

Epochs	Before odor	Odor/choice	Reward
Odor/choice	$R^2 = 0.01$ $p = 0.73$		
Reward	$R^2 = 0.00$ $p = 0.81$	$R^2 = 0.08$ $p = 0.29$	
After reward	$R^2 = 0.19$ $p = 0.092$	$R^2 = 0.00$ $p = 1.00$	$R^2 = 0.35$ $p < 0.05$

Figure 4. DMS CIN pseudo-ensembles decode the current state (i.e., block) across the trials. The top plot shows block decoding accuracy of a pseudo-ensemble of 25 DMS CINs versus MSNs across the trial. All rewarded trials after the first 20 correct trials in blocks were used in the decoder, regardless of direction. Bottom plots show decoding accuracy as a function of ensemble size for three 1 s epochs across the trial. **$p < 0.01$.

right). Accordingly, an analysis of reaction times on suboptimal choice trials suggested that these choices reflected a mistaken, perhaps optimistic, prediction that the large reward would be delivered (Fig. 5*c*). Notably, this miscoding was specific to CINs (MSNs coded and miscoded the block at chance levels across the trial; Fig. 5*a,b*), and it was specific to suboptimal choices (free choices of the big outcome showed no such miscoding in either CINs or MSNs; Fig. 6). The strong association between coding of state by CINs and the rats' behavior, particularly on suboptimal choices, suggests that CINs were not signaling external cues or
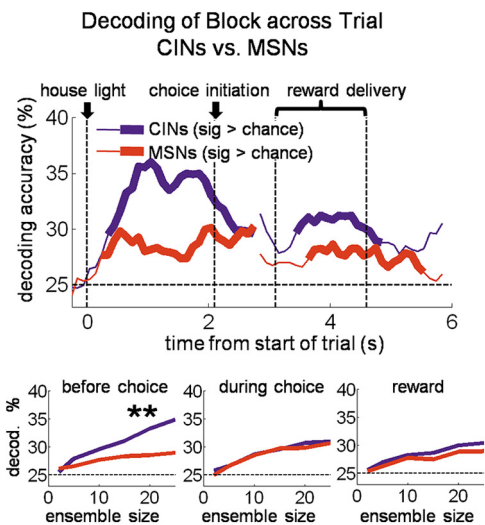
contexts but instead reflected the rats' actual real-time beliefs about the current task state.

Of course, for CIN state coding to influence choices, it must do so by influencing MSNs (which provide the sole output of the striatum). Although the MSNs did not independently represent state like the CINs, we hypothesized that CIN activity might still serve to bias or gate activity in particular groups of MSNs during

Table 3. Parameter estimates for decoding accuracy functions

Condition	CINs or MSNs	Before choice	During choice	Reward
Block decoding accuracy all trials, shams (Fig. 4)	CINs	3.0 ± 1.1	2.4 ± 1.0	2.1 ± 0.4
	MSNs	**1.4 ± 0.7**	1.8 ± 1.0	1.4 ± 0.5
Accuracy on suboptimal choices, shams (Fig. 5a)	CINs	−10.6 ± 4.1	−7.3 ± 2.7	−5.4 ± 3.9
	MSNs	**1.6 ± 1.2**	**2.3 ± 0.5**	**1.9 ± 2.8**
Miscoding on suboptimal choices, shams (Fig. 5b)	CINs	11.2 ± 8.2	17.3 ± 10.0	9.5 ± 4.8
	MSNs	**−0.2 ± 3.1**	**2.5 ± 5.3**	**−1.3 ± 2.9**
Accuracy on optimal choices, shams (Fig. 6a)	CINs	**3.2 ± 3.3**	**1.5 ± 3.3**	**0.8 ± 2.7**
	MSNs	1.0 ± 5.1	2.2 ± 1.8	2.2 ± 1.8
Miscoding on optimal choices, shams (Fig. 6b)	CINs	**−3.0 ± 1.7**	**−1.9 ± 0.8**	**−1.7 ± 2.3**
	MSNs	−1.2 ± 3.2	−2.1 ± 2.0	−1.8 ± 2.9
% Yoked, suboptimal trials (Fig. 7a)	MSNs	−0.2 ± 1.9	24.7 ± 6.9	−3.3 ± 1.9
Accuracy of yoked MSNs, suboptimal trials (Fig. 7b)	MSNs	3.1 ± 1.5	−11.9 ± 0.8	−5.9 ± 1.6
% Yoked, optimal trials (Fig. 7c)	MSNs	4.0 ± 2.8	2.9 ± 2.2	2.1 ± 2.4
Accuracy of yoked MSNs, optimal trials (Fig. 7d)	MSNs	6.3 ± 1.8	7.5 ± 2.3	8.6 ± 1.6
Choice-direction decoding accuracy, shams (Fig. 8)	CINs	11.8 ± 0.57	16.0 ± 5.2	13.0 ± 2.0
	MSNs	**5.8 ± 1.7**	11.6 ± 3.3	**7.6 ± 2.5**
Block decoding accuracy all trials, OFC lesions (Fig. 10)	CINs	0.3 ± 0.4	0.4 ± 2.0	0.4 ± 0.4
Accuracy on suboptimal choices, OFC lesions (Fig. 12a)	CINs	−1.2 ± 3.0	0.14 ± 1.1	2.5 ± 2.7
Miscoding on suboptimal choices, OFC lesions (Fig. 12b)	CINs	−4.8 ± 2.0	0.1 ± 2.8	−1.0 ± 3.5
Choice-direction decoding accuracy, OFC lesions (Fig. 12d)	CINs	**7.6 ± 2.8**	14.6 ± 2.4	**5.7 ± 2.8**
	Sham or lesion	Before odor	During odor	After odor
Odor identity decoding accuracy (Fig. 11a)	Sham	0.3 ± 1.3	3.6 ± 1.1	11.7 ± 2.1
	OFC lesion	**4.7 ± 2.2**	**10.5 ± 2.8**	11.1 ± 2.2
	Sham or lesion	Before reward	During reward	After reward
Reward amount decoding accuracy (Fig. 11b)	Sham	3.5 ± 2.6	3.3 ± 0.9	9.1 ± 4.1
	OFC lesion	6.5 ± 1.8	**13.5 ± 2.6**	13.0 ± 2.0

Decoding accuracy as a function of ensemble size curves were fit to a log function with two parameters (see Materials and Methods). Shown are the estimates (± 95% confidence intervals) for the parameter, B_1, which controls the rise of the function and thus provides a measure of how much information about the decoded variable is on average contained in that population of neurons. Estimates and confidence intervals were estimated using the MATLAB function *nlparci*. Significance was determined by finding the smallest α level that resulted in non-overlapping confidence intervals between fitted parameters from the two different populations, with $p =$ (minimum non-overlapping α)^2 and significance criterion at $p < 0.01$. Significance is indicated by the bold font for the following comparisons: (1) for MSNs on all trials and on suboptimal choices versus CINs in the corresponding condition; (2) for MSNs and CINs on optimal choices versus the corresponding population on suboptimal choices; and (3) for OFC lesion conditions versus corresponding sham conditions.

particular parts of the trial. For example, we have previously found in a task similar to that used here that dorsal striatal MSNs are selective for particular response–outcome associations, i.e., the rules that govern appropriate responding in particular blocks. This selectivity is strongest in the period leading up to and during the choice (Stalnaker et al., 2010). It is possible that CIN state coding might serve to select the set of MSNs encoding the associations or rules appropriate for the current block, allowing execution of the appropriate choice. Such control over MSN activity would be particularly apparent on suboptimal trials, because, in these trials, the state represented by the CINs diverges from the experimenter-defined state of the task. To test this idea, we constructed DMS MSN ensembles restricted to neurons that were recorded in the same sessions as CINs (2.6 ± 2.2 MSNs per CIN) and then tested how block decoding by these ensembles related to CIN block decoding. For the MSN decoder, we trained and tested at the same trial epoch, as opposed to using a training epoch in a different part of the trial from the test epoch, as before. This change allowed us to ask whether MSNs signaled the state appropriately within individual trial epochs rather than across the trial, as CINs appeared to do. The results of this analysis, illustrated in Figure 7, show that, on suboptimal trials, MSN state decoding was tightly coupled to CIN state decoding. Interestingly, this coupling was present specifically surrounding the time of the choice, precisely when the information about the rules signaled by the MSNs in our previous study would be helpful to guide behavior. On optimal choice trials, MSNs had a much smaller tendency to be yoked to the CIN-encoded state in the period right before the choice. These results suggest that CINs have a strong influence on

the activity of MSNs around the same time that MSNs best represent the associative rules in the different blocks. This effect was primarily evident when the rat made a suboptimal choice, as if that information were driving these choices. When choices were consistent with the current state (i.e., optimal choices), the linkage of CINs and MSNs was weaker, perhaps reflecting the parallel involvement of other neural systems in such default behavior.

DMS CIN activity predicts the upcoming choice

We tested the idea suggested by the above analysis that CIN state coding (and miscoding) might influence choices by constructing a decoder that predicted the direction of the current choice on all free-choice trials across sessions. Consistent with the idea, DMS CIN ensembles were highly accurate in predicting the upcoming choice well before choice movement began and earlier than MSNs (Fig. 8). Note that this decoder included trials immediately after number-block switches, when rats were still choosing the side with the small reward most of the time. Thus, the high level of accuracy of CIN ensembles (~95% immediately before choice execution) could not be simply a "bias" signal based on which side had the large reward in any given block (across all free-choice trials, rats only chose the side with the large reward 77% of the time). Information about the upcoming choice encoded by CINs could not be based on any stimulus in the environment and instead could only reflect an internal belief about what the state was. This is additional evidence that CINs were involved in the process by which current beliefs about task-state were translated into choices.
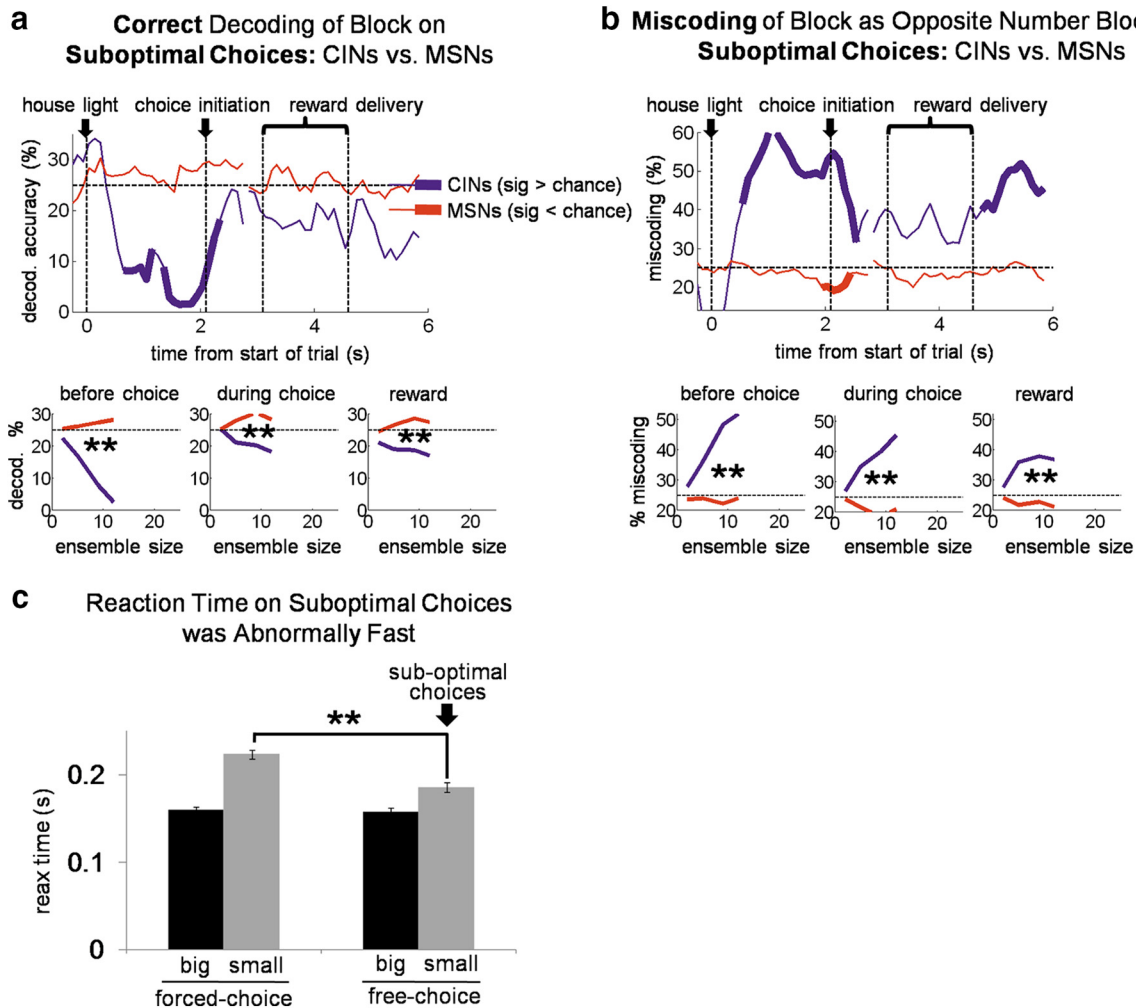
Figure 5. On suboptimal (i.e., inappropriate) choices, DMS CINs miscoded the block as the one that would have been appropriate for that choice. Top plots show percentage accurate block decoding (*a*) and percentage miscoding as the opposite drop-number block (*b*), by a pseudo-ensemble of 12 CINs or MSNs on suboptimal choices. Twelve CINs were recorded in sessions with enough such free-choices. Bottom plots show accurate decoding and miscoding as a function of ensemble size for three 1 s epochs across the trial. *c*, Bars show average (± SE) reaction time by trial type. Reaction times were faster for big reward than small reward, showing that faster reaction times occurred when more valuable rewards were predicted. However, free-choices of the small reward, suboptimal choices, were still significantly faster than forced-choices of the small reward, suggesting an inaccurate (and perhaps optimistic) prediction as to the number of drops to be received on those trials. **$p < 0.01$.

DMS CIN state and choice information depends on the OFC

However, where does the state information encoded by CINs come from? We tested whether OFC was critical to internally cued state representations in DMS CINs by examining CINs in rats with ipsilateral lesions of the OFC. Our analysis found that indeed OFC lesions resulted in significantly weaker block selectivity across the trial in the DMS CINs, without having the same effect on DLS CINs. As illustrated in Figure 9a, block-selectivity indices in the DMS in OFC-lesioned rats were not significantly different from zero during early and late trial periods when DMS CINs in shams maintained significant selectivity [DMS CINs in lesioned rats were evenly split between those preferring the contralateral side ($n = 18$) and those preferring the ipsilateral side ($n = 18$)]. Accordingly, selectivity scores of individual DMS CINs in lesioned rats were mostly uncorrelated between pairs of epochs across the trial period (Fig. 9b; average $R^2 = 0.09$, only two of six R^2 had $p < 0.05$; Table 4). DLS CINs in lesioned rats did not show this pattern; if anything, they were more often correlated across pairs of epochs than their counterparts in sham rats (Fig. 9d; average $R^2 = 0.20$, four of six R^2 had $p < 0.05$; Table 5). The block decoding analysis also showed the dependence of DMS CIN

block information on the OFC. OFC lesions resulted in block decoding in DMS CINs that was at chance (and significantly less than that in shams) both before and after the choice period (Fig. 10). Thus, information about the state was essentially eliminated from the DMS CINs by ipsilateral OFC lesions. Notably, this occurred although CINs in lesioned rats continued to respond to many external events in each trial. Indeed, as shown in Figure 11, DMS CINs in lesioned rats were actually significantly better than CINs in sham rats at decoding the odor identity and the number of reward drops during presentation of those stimuli compared to DMS CINs in sham rats. This again suggests that information about current state encoded by CINs in sham rats did not emerge from responsivity to a variety of separate trial events; on the contrary, to the extent that CINs carried information about individual trial events, they would be impaired at tracking current state (and vice versa). Indeed, the loss of OFC-derived state information may have left the CINs free to respond to more generally stable, state-spanning trial events.

Finally, OFC lesions also decoupled CIN activity from choice behavior. As shown in Figure 12, *a* and *b*, decoding of block in DMS CINs did not show the strong miscoding on
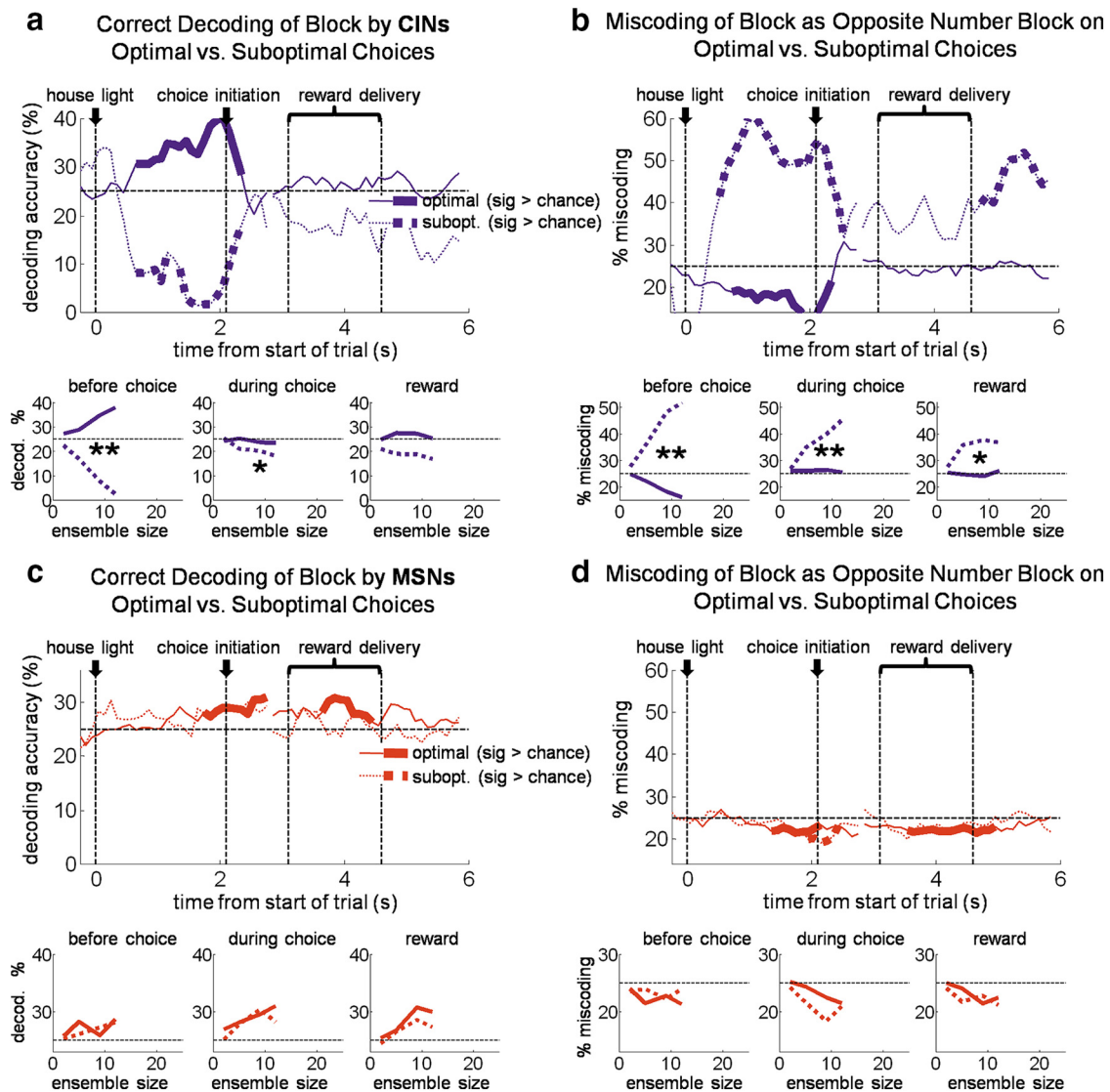
Figure 6. Decoder performance on free-choices of the big reward (i.e., optimal choices) compared with that on free-choices of the small reward (i.e., suboptimal choices). *a*, *c*, Top plots show block decoding accuracy of a pseudo-ensemble of 12 DMS CINs (*a*) or 12 DMS MSNs (*c*), during sliding 500 ms epochs across the trial. The same procedures were used as for the analysis shown in Figure 5, except that only free-choices of the big reward were tested. For comparison, data from small free-choices (suboptimal choices), shown in Figure 5, is also plotted here in dotted lines. Bottom plots show decoding accuracy as a function of ensemble size for the same three epochs shown in Figure 4. *b*, *d*, Top plots show percentage of test trials in which the block decoder from the corresponding plot (in *a* and *c*, respectively) misclassified the block as the one with the opposite drop-number rewards but the same flavor. Bottom plots show the misclassification percentage as a function of ensemble size. *$p < 0.05$ **$p < 0.01$.

suboptimal choices; instead, decoding was at chance, although reaction times on suboptimal choices were relatively fast, just as they were in the sham group (Fig. 12c). Likewise, as shown in Figure 12d, choice-direction decoding in CINs was also significantly impaired by ipsilateral OFC lesions, both before and after the choice period. The reason for this decoupling may be that the intact OFC in the contralateral hemisphere could drive both reaction times and choice behavior. This implies that, when state information was needed to determine what the choice would be (or had been), OFC lesions degraded the representation of this information in the ipsilateral CINs. This is again consistent with the idea that the OFC is needed to maintain a consistent representation of state and choice, especially during time periods when this information must be maintained internally. Retrospective choice information (i.e., information about which choice had led to a particular outcome) might be especially important to promote

learning of associations between the choice and its consequences (i.e., credit assignment). Thus, the loss of OFC-derived state information could have effects not only on the acute coding of state by DMS CINs but also on striatum-mediated associative learning that depends on state and choice information.

Discussion

These results have several major implications. First, they show that DMS CINs have a special role in keeping track of the current state. Individual DMS CINs represented state from the start to the end of the trials within each block. The finding that CINs code state is consistent with evidence that interference with cholinergic function in the striatum often causes problems with flexible behavior (Tzavos et al., 2004; Ragozzino et al., 2009; Okada et al., 2014; Aoki et al., 2015) and with very specific evidence that this reflects a role for DMS CINs in integrating new learning with old
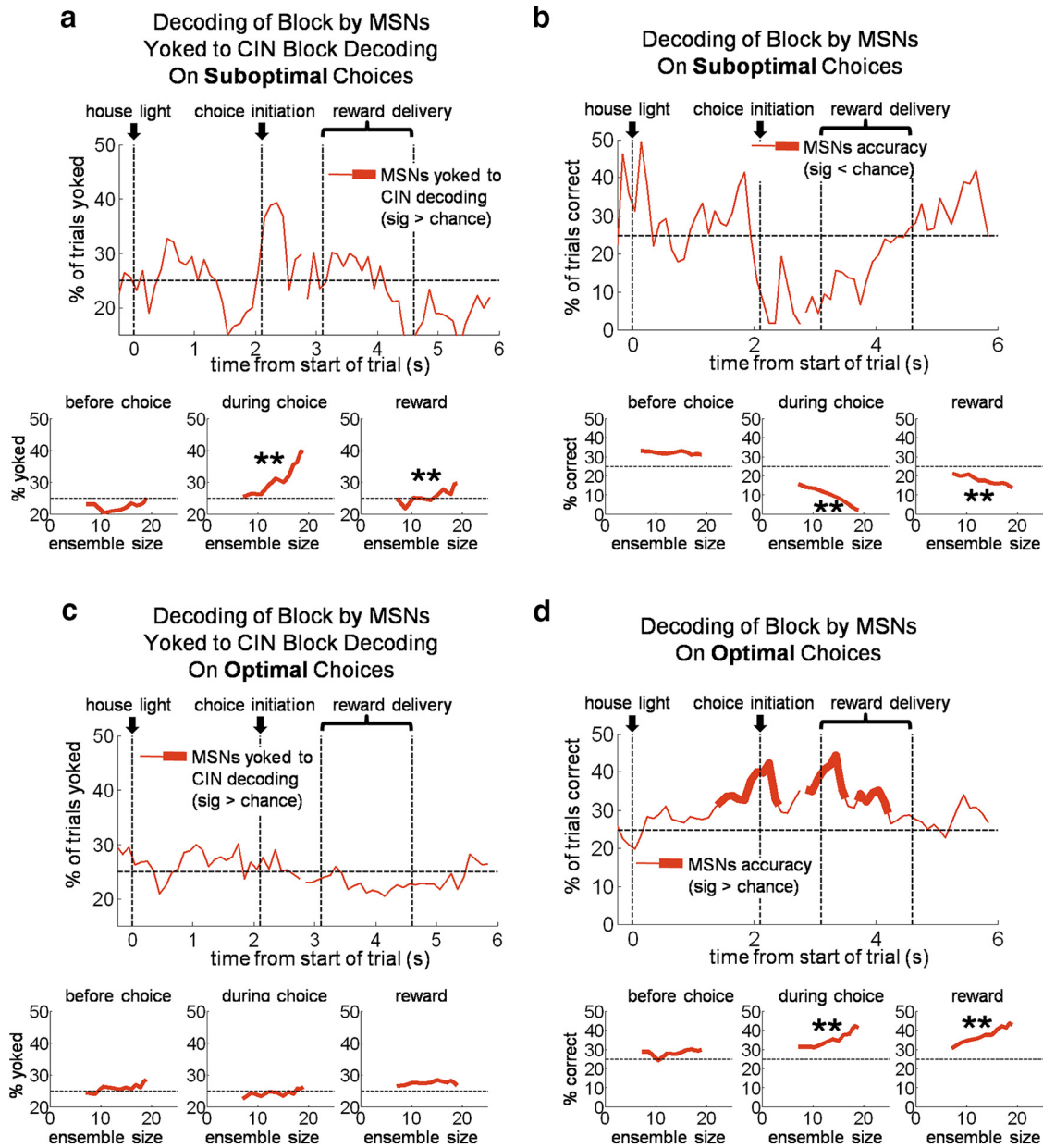
Figure 7. MSNs block decoding is "yoked" to CIN decoding during execution of suboptimal choices. *a–d*, Top plots show block decoding accuracy of pseudo-ensembles of MSNs recorded in the same sessions as CINs. *a*, *c*, MSN decoding yoked to CIN decoding on suboptimal choices, meaning the percentage of trials on which MSN ensembles identified the same block as the CIN ensembles. *b*, *d*, Block decoding accuracy of the same MSN ensembles as *a* and *c* (i.e., not yoked to CINs). Bottom plots show decoding accuracy as a function of ensemble size for three 500 ms epochs across the trial. (starting 1000 ms before choice initiation, 100 ms before choice initiation, and at first reward delivery, respectively). **$p < 0.01$ compared with 0.

Table 4. Correlations between block-selectivity indices in pairs of epochs in DMS CINs recorded in OFC-lesioned rats (related to Fig. 9c)

Epochs	Before odor	Odor/choice	Reward
Odor/choice	$R^2 = 0.14$ $p = 0.027$		
Reward	$R^2 = 0.27$ $p = 0.0011$	$R^2 = 0.10$ $p = 0.057$	
After reward	$R^2 = 0.00$ $p = 0.71$	$R^2 = 0.01$ $p = 0.55$	$R^2 = 0.00$ $p = 0.82$

Table 5. Correlations between block-selectivity indices in pairs of epochs in DLS CINs recorded in OFC-lesioned rats (related to Fig. 9d)

Epochs	Before odor	Odor/choice	Reward
Odor/choice	$R^2 = 0.03$ $p = 0.37$		
Reward	$R^2 = 0.19$ $p < 0.05$	$R^2 = 0.21$ $p < 0.05$	
After reward	$R^2 = 0.03$ $p = 0.36$	$R^2 = 0.50$ $p < 0.0001$	$R^2 = 0.26$ $p < 0.01$

without interference (Bradfield et al., 2013). Bradfield et al. showed that rats in which CIN function was compromised learned an initial set of associations normally but then failed to appropriately separate new learning for those materials from the initial learning. Instead, the rats seemed to average the two learning episodes together, as if unable to separate the different learning states or contexts. Our data provide a clear explanation for this result—that CINs signal the subject's internal

Decoding of **Choice Direction** on Free-choice Trials
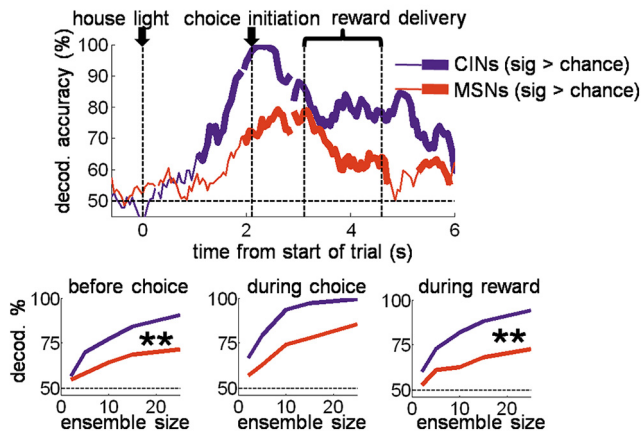

Figure 8. DMS CINs also coded the direction of free-choice before and after it occurred better than MSNs. Top plot shows choice-direction decoding accuracy of a pseudo-ensemble of 25 CINs or MSNs. Only free-choice trials across entire sessions were used; therefore significant decoding of direction before the choice could only reflect an internal intention or state. Bottom plots show decoding accuracy as a function of ensemble size for three 1 s epochs across the trial. ** $p < 0.001$.

belief regarding current state in real time. The loss of such a signal would result in new learning being confused or combined with previous learning. Importantly, the mechanism demonstrated here differs from much more restricted alternative proposals in which CINs signal state changes or so-called "state errors." Our data indicate a much more pervasive role for CINs in representing state in real time. This would put CINs in a position to modulate behavior directly, although appropriate state recall, and also indirectly through state rep-

resentation and even state creation during learning episodes. In other words, when learning occurs, the ultimate effect of that learning on behavior may depend as much on the state represented by CINs as on the efficacy of the learning.
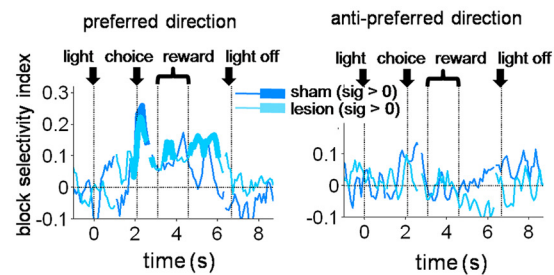
CIN activity was also uniquely related to the direction of choice, especially at times when this information could only be internally derived rather than stimulus driven. The combination of integrative state information and trial-specific choice information suggests an involvement in translating beliefs about the current state into appropriate action selection. This gives credence and specificity to suggestions that CINs (or TANs) provide motivational contexts for actions (Yamada et al., 2004) and that they may be involved in initiation (or reporting) of appropriate actions (Blazquez et al., 2002; Lee et al., 2006; Apicella, 2007; Yarom and Cohen, 2011). The mechanism through which DMS CINs use state information to "tag" or otherwise influence synaptic events and potentially influence choice behavior remains an open question (Ashby and Crossley, 2011). Recent work has pointed to cholinergic gating of plasticity and activity at corticostriatal synapses (Wang et al., 2006; Ding et al., 2010) and to the close interplay and dependence between cholinergic and dopaminergic activity in the striatum (Morris et al., 2004; Wang et al., 2006; Goldberg and Reynolds, 2011). One interesting idea is that dopaminergic and cholinergic activity could provide related but complementary pieces of information that modulate learning: whereas dopamine provides a reward error signal, acetylcholine might keep track of current state (Goldberg and Reynolds, 2011; Bradfield et al., 2013).

Another important question is how CIN state encoding influences activity in MSNs, the output neurons of the striatum. From our analysis, it appears that state information is not as concentrated in MSNs as it is in CINs (except perhaps during choice
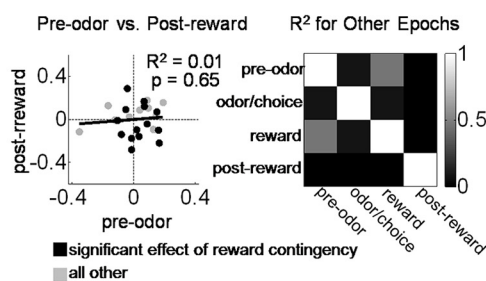

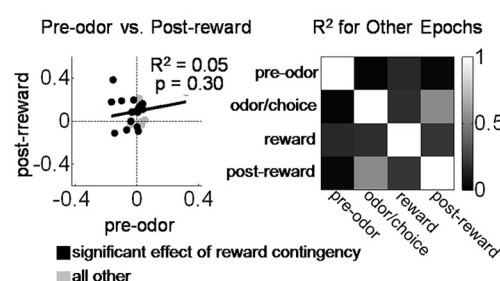**Figure 9.** OFC lesions eliminate state encoding in DMS CINs. As in Figure 3, plots in *a* and *b* show a bin-by-bin average index of block-selectivity for the entire DMS (*a*) and DLS (*b*) CIN population in OFC-lesioned rats compared with those in sham rats. Block-selectivity scores of individual CINs in *c* and *d* show that OFC lesions eliminated the consistency of block-selectivity across the trial in DMS CINs (*c*) without affecting that in DLS CINs (*d*). Left, Scores in a pre-odor epoch versus post-reward epoch; other pairs of epochs are summarized on the right.
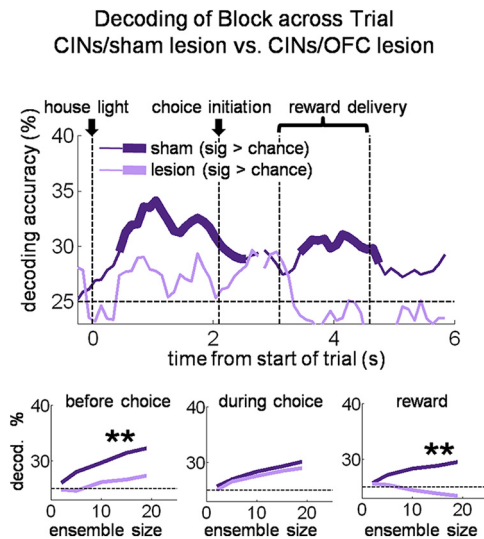
Decoding of Block across Trial
CINs/sham lesion vs. CINs/OFC lesion

Figure 10. OFC lesions eliminate state encoding in DMS CINs. As in Figure 4, top plot shows block decoding accuracy of a pseudo-ensemble of 19 DMS CINs in OFC-lesioned rats (only 19 neurons were recorded in blocks with a sufficient number of trials) versus those in sham rats, also with a pseudo-ensemble of 19 neurons for comparison. Bottom plots show decoding accuracy as a function of ensemble size for three 1 s epochs across the trial. **$p < 0.01$.

opaque to us in this experiment. However, our finding that MSNs recorded simultaneously with the CINs seem to fire in a way that reflects the CIN-encoded state, particularly around the choice point, is consistent with such an influence. We can only speculate on why this occurs so prominently on suboptimal choice trials. One novel possibility is that this reflects the unique importance of DMS CINs in the flexible control of behavior, which may be particularly strongly driving behavior on these trials; their role (and thus coding) may be much weaker on optimal choice trials when presumably other systems, including those driving habitual behavior, may operate in parallel to support behavior.

In this regard, it is also worth commenting on the specificity of our result. We observed state correlates in DMS CINs but not in DLS CINs recorded in the same rats. This specificity is consistent with previous behavioral results, in which the role of the CINs in integrating new learning with old was specific to the DMS (Bradfield et al., 2013). However before jumping to the conclusion that the segregation of information into states is a unique function of DMS CINs, it is worth pointing out that, in both the earlier behavioral work and in our recording task, the relevant information that must be segregated (the rules that differ across blocks or training episodes) reflected changing response–outcome associations. As has now been well established, the DMS is particularly important for behavior that reflects such rules, whereas the DLS is not (Yin et al., 2004, 2005, 2006; Valentin et al., 2007; Tanaka et al., 2008; Tricomi et al., 2009). Instead, the DLS seems to be more important for behavior that reflects stimulus–response associations. The stimulus–response associations did not change in our task. Indeed, from the perspective of stimulus–response rules, our experiment consisted of but a single state or trial block. Thus, our failure to see block correlates in the DLS CINs in our task may simply reflect the fact that our blocks did not differ in the type of information that DLS CINs use to identify different states. This would be notable, because we have failed to see marked differ-
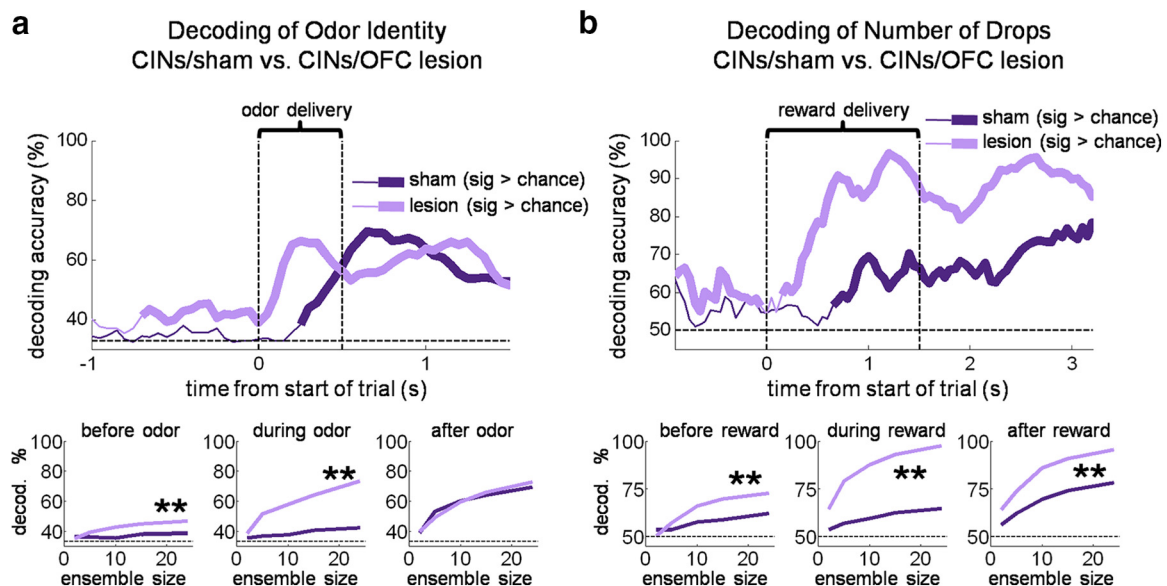
execution). Furthermore, it appears that state information held by MSNs is not as closely related to choice as it is in CINs. The relative restriction of state representations to the CINs suggests that the DMS does not broadcast state information to downstream areas. Instead, its role may be mostly local, perhaps influencing the appropriate recall of specific task variables (cues, responses, anticipated outcomes) or associative information (left→big/chocolate, right→small, vanilla) in specific subpopulations of MSNs with which they interact. To really test this idea, one would have to determine which set of MSNs receive input from a particular CIN or set of CINs, information that is

Figure 11. DMS CINs recorded in rats with OFC lesions decode trial stimuli better than those in sham rats. Top plots show odor identity (**a**) and reward number (**b**) decoding accuracy of a pseudo-ensemble of 24 DMS CINs in shams versus OFC lesions, during sliding 200 ms epochs across the trial. All correct trials were included and therefore for odor identity, three odors had to be decoded (the forced-choice left, forced-choice right, and free-choice odors). Note that because of correction trials and the pseudorandom sequence of trials, some limited information about the upcoming trial-type could be derived before odor delivery began. Bottom plots show decoding accuracy as a function of ensemble size for three epochs each: for odor decoding, 1 s ending with the beginning of odor delivery, the 500 ms of odor delivery, and the 1 s immediately following odor delivery; for reward number decoding, 1 s ending with the beginning of reward delivery, the period of reward delivery and consumption (1500 ms beginning with delivery of the first drop of reward), and 1 s immediately after the end of the reward delivery epoch. **$p < 0.01$.
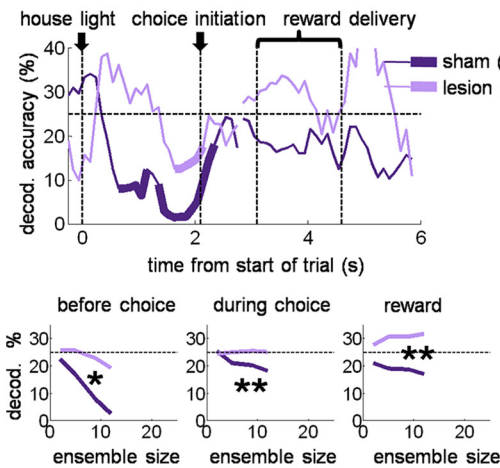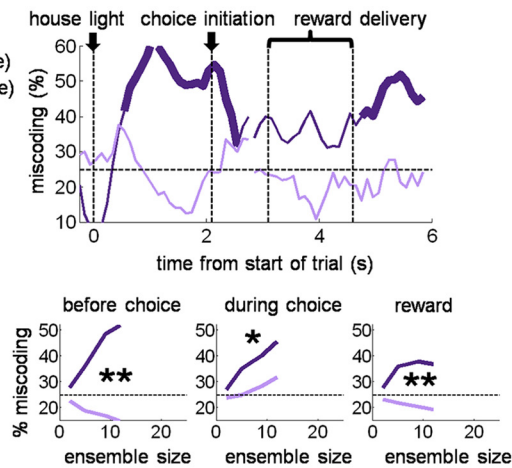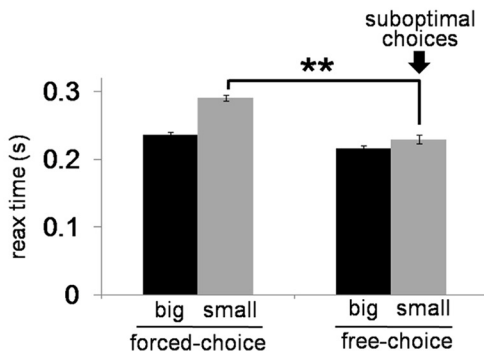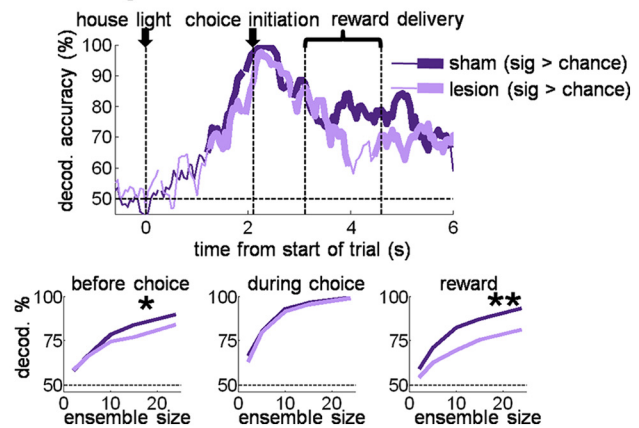
Figure 12. Ipsilateral OFC lesions eliminate the relationship between free-choices of the small reward and miscoding of the block in DMS CINs. Top plots show percentage accurate block decoding (**a**) and percentage miscoding as the opposite drop-number block (**b**), by a pseudo-ensemble of 12 CINs in sham rats or ipsilateral OFC-lesioned rats on suboptimal choices. Bottom plots show accurate decoding and miscoding as a function of ensemble size for three 1 s epochs across the trial. **c**, Bars show average (±SE) reaction time by trial type in lesioned rats. The pattern of reaction times was the same as in sham rats (Fig. 5); that is, they were faster for big reward than small reward, but free-choices of the small reward, suboptimal choices, were similarly fast to those when a big reward was expected, suggesting an inaccurate prediction as to the number of drops to be received on those trials. **d**, Top plot shows choice-direction decoding accuracy of a pseudo-ensemble of 24 CINs in sham rats or ipsilateral OFC-lesioned rats. Only free-choice trials were used; therefore significant decoding of direction before the choice could only reflect an internal intention or state. Bottom plots show decoding accuracy as a function of ensemble size for three 1 s epochs across the trial. **p < 0.01.

ences in the rules represented at the level of individual MSNs in the DMS versus the DLS (Stalnaker et al., 2010).

Second, these results join other findings (Bradfield et al., 2015) that support the recent proposal that the critical contribution of the OFC to processing in other areas is to signal state information (Wilson et al., 2014). Consistent with this idea, CIN state information in the DMS was almost completely dependent on the ipsilateral OFC, although other aspects of CIN response properties were unaffected or even enhanced. Dependence on the OFC is in accord with long-standing evidence that associative information signaled by the OFC is strongly context dependent (Thorpe et al., 1983; Schoenbaum et al., 1999; Wallis et al., 2001; Simmons and Richmond, 2008; Kennerley et al., 2011; Young and Shapiro, 2011) and with more recent reports that OFC neurons may signal context directly (Farovik et al., 2015). Importantly, the states in this task were not perceptually distinct from each other, i.e., the task state was not explicitly identified by an external cue or even the sequential order of the blocks, which changed each day. Instead, to derive the current state, rats had to

remember the properties of the reward that had been received on the right and those received on the left on recent trials and integrate these pieces of information when re-engaging with the task at the start of each trial. OFC has been repeatedly implicated in the ability to integrate information from different sources and use it to make inferences (McDannald et al., 2011; Gallagher et al., 1999; Jones et al., 2012). Here this ability would be critical for disambiguating states that are perceptually similar (Wilson et al., 2014). Notably, this was true for both block information and internal information reflecting the response selected, before and after the choice itself. Thus, the OFC influenced the ability to maintain information relevant to both learning (i.e., state) and ongoing behavior (i.e., choice) in DMS CINs. Although the precise pathways through which the OFC may influence CIN activity is uncertain, the interaction of the OFC with DMS CINs could be one mechanism through which the OFC influences behavioral flexibility.

Understanding how neural systems keep track of state could also be important for treating psychiatric diseases, such as addic-

tion or anxiety disorders, in which patients generalize (or fail to generalize) inappropriately between different contexts. For example, in posttraumatic stress disorder, treatment strategies often address the original learning when perhaps what has gone awry is the ability to keep separate information learned in different states when moving between similar environments, such as from Baghdad to Baltimore. Distinguishing the wartime and civilian contexts, which share many features, may often be a matter of maintaining the appropriate internal state, particularly when it is necessary to bridge gaps between external markers that clearly distinguish the two. A pathological weakening of this mechanism would likely not prevent separation of these states under most circumstances, but it might occasionally result in momentary, transient recall of the inappropriate state. This would result in expression of behaviors and emotional responses from one context in the other. Viewed from this perspective, understanding how these internal states are recognized and represented in the brain is a critical question facing the field, because it may be key for both temporarily preventing symptomatic events and permanently altering the underlying learned responses in a host of neuropsychiatric diseases. Our study suggests that CINs, with input from the OFC, are critical to this process.

References

Aoki S, Liu AW, Zucca A, Zucca S, Wickens JR (2015) Role of striatal cholinergic interneurons in set-shifting in the rat. J Neurosci 35:9424–9431. CrossRef Medline

Aosaki T, Tsubokawa H, Ishida A, Watanabe K, Graybiel AM, Kimura M (1994) Responses of tonically active neurons in the primate's striatum undergo systematic changes during behavioral sensorimotor conditioning. J Neurosci 14:3969–3984. Medline

Apicella P (2002) Tonically active neurons in the primate striatum and their role in the processing of information about motivationally relevant events. Eur J Neurosci 16:2017–2026. CrossRef Medline

Apicella P (2007) Leading tonically active neurons of the striatum from reward detection to context recognition. Trends Neurosci 30:299–306. CrossRef Medline

Ashby FG, Crossley MJ (2011) A computational model of how cholinergic interneurons protect striatal-dependent learning. J Cogn Neurosci 23:1549–1566. CrossRef Medline

Benhamou L, Kehat O, Cohen D (2014) Firing pattern characteristics of tonically active neurons in rat striatum: context dependent or species divergent? J Neurosci 34:2299–2304. CrossRef Medline

Blazquez PM, Fujii N, Kojima J, Graybiel AM (2002) A network representation of response probability in the striatum. Neuron 33:973–982. CrossRef Medline

Bradfield LA, Bertran-Gonzalez J, Chieng B, Balleine BW (2013) The thalamostriatal pathway and cholinergic control of goal-directed action: interlacing new with existing learning in the striatum. Neuron 79:153–166. CrossRef Medline

Bradfield LA, Dezfouli A, van Holstein M, Chieng B, Balleine BW (2015) Medial orbitofrontal cortex mediates outcome retrieval in partially observable task situations. Neuron 88:1268–1280. CrossRef Medline

Calabresi P, Centonze D, Gubellini P, Pisani A, Bernardi G (2000) Acetylcholine-mediated modulation of striatal function. Trends Neurosci 23:120–126. CrossRef Medline

Cinelli AR, Ferreyra-Moyano H, Barragan E (1987) Reciprocal functional connections of the olfactory bulbs and other olfactory related areas with the prefrontal cortex. Brain Res Bull 19:651–661. CrossRef Medline

Ding JB, Guzman JN, Peterson JD, Goldberg JA, Surmeier DJ (2010) Thalamic gating of corticostriatal signaling by cholinergic interneurons. Neuron 67:294–307. CrossRef Medline

Farovik A, Place RJ, McKenzie S, Porter B, Munro CE, Eichenbaum H (2015) Orbitofrontal cortex encodes memories within value-based schemas and represents contexts that guide memory retrieval. J Neurosci 35:8333–8344. CrossRef Medline

Gallagher M, McMahan RW, Schoenbaum G (1999) Orbitofrontal cortex and representation of incentive value in associative learning. J Neurosci 19:6610–6614. Medline

Goldberg JA, Reynolds JN (2011) Spontaneous firing and evoked pauses in the tonically active cholinergic interneurons of the striatum. Neuroscience 198:27–43. CrossRef Medline

Graveland GA, DiFiglia M (1985) The frequency and distribution of medium-sized neurons with indented nuclei in the primate and rodent neostriatum. Brain Res 327:307–311. CrossRef Medline

Hasselmo ME, Bower JM (1993) Acetylcholine and memory. Trends Neurosci 16:218–222. CrossRef Medline

Holt GR, Softky WR, Koch C, Douglas RJ (1996) Comparison of discharge variability in vitro and in vivo in cat visual cortex neurons. J Neurophysiol 75:1806–1814. Medline

Inokawa H, Yamada H, Matsumoto N, Muranishi M, Kimura M (2010) Juxtacellular labeling of tonically active neurons and phasically active neurons in the rat striatum. Neuroscience 168:395–404. CrossRef Medline

Jones JL, Esber GR, McDannald MA, Gruber AJ, Hernandez A, Mirenzi A, Schoenbaum G (2012) Orbitofrontal cortex supports behavior and learning using inferred but not cached values. Science 338:953–956. CrossRef Medline

Kawaguchi Y, Wilson CJ, Augood SJ, Emson PC (1995) Striatal interneurones: chemical, physiological and morphological characterization. Trends Neurosci 18:527–535. CrossRef Medline

Kennerley SW, Behrens TE, Wallis JD (2011) Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. Nat Neurosci 14:1581–1589. CrossRef Medline

Kimura M, Rajkowski J, Evarts E (1984) Tonically discharging putamen neurons exhibit set-dependent responses. Proc Natl Acad Sci U S A 81:4998–5001. CrossRef Medline

Kita H, Kitai ST (1990) Amygdaloid projections to the frontal cortex and the striatum in the rat. J Comp Neurol 298:40–49. CrossRef Medline

Kosar E, Grill HJ, Norgren R (1986) Gustatory cortex in the rat. I. Physiological properties and cytoarchitecture. Brain Res 379:329–341. CrossRef Medline

Krettek JE, Price JL (1977) Projections from the amygdaloid complex to the cerebral cortex and thalamus in the rat and cat. J Comp Neurol 172:687–722. CrossRef Medline

Krushel LA, van der Kooy D (1988) Visceral cortex: integration of the mucosal senses with limbic information in the rat agranular insular cortex. J Comp Neurol 270:39–54, 62–63. CrossRef Medline

Lee IH, Seitz AR, Assad JA (2006) Activity of tonically active neurons in the monkey putamen during initiation and withholding of movement. J Neurophysiol 95:2391–2403. Medline

McDannald MA, Lucantonio F, Burke KA, Niv Y, Schoenbaum G (2011) Ventral striatum and orbitofrontal cortex are both required for model-based, but not model-free, reinforcement learning. J Neurosci 31:2700–2705. CrossRef Medline

Meyers EM (2013) The neural decoding toolbox. Front Neuroinform 7:8. CrossRef Medline

Morris G, Arkadir D, Nevet A, Vaadia E, Bergman H (2004) Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. Neuron 43:133–143. CrossRef Medline

Okada K, Nishizawa K, Fukabori R, Kai N, Shiota A, Ueda M, Tsutsui Y, Sakata S, Matsushita N, Kobayashi K (2014) Enhanced flexibility of place discrimination learning by targeting striatal cholinergic interneurons. Nat Commun 5:3778. CrossRef Medline

Oorschot DE (2013) The percentage of interneurons in the dorsal striatum of the rat, cat, monkey and human: a critique of the evidence. Basal Ganglia 3:19–24. CrossRef

Price JL, Carmichael ST, Carnes KM, Clugnet M-C, Kuroda M, Ray JP (1991) Olfactory input to the prefrontal cortex. In: Olfaction: a model system for computational neuroscience (Davis J, Eichenbaum H, eds), pp 101–120. Cambridge, MA: Massachusetts Institute of Technology.

Ragozzino ME, Mohler EG, Prior M, Palencia CA, Rozman S (2009) Acetylcholine activity in selective striatal regions supports behavioral flexibility. Neurobiol Learn Mem 91:13–22. CrossRef Medline

Ravel S, Sardo P, Legallet E, Apicella P (2006) Influence of spatial information on responses of tonically active neurons in the monkey striatum. J Neurophysiol 95:2975–2986. CrossRef Medline

Saper CB (1982) Convergence of autonomic and limbic connections in the insular cortex of the rat. J Comp Neurol 210:163–173. CrossRef Medline

Schoenbaum G, Chiba AA, Gallagher M (1999) Neural encoding in orbito-

frontal cortex and basolateral amygdala during olfactory discrimination learning. J Neurosci 19:1876–1884. Medline

Sharott A, Doig NM, Mallet N, Magill PJ (2012) Relationships between the firing of identified striatal interneurons and spontaneous and driven cortical activities *in vivo*. J Neurosci 32:13221–13236. CrossRef Medline

Shi CJ, Cassell MD (1998) Cortical, thalamic, and amygdaloid connections of the anterior and posterior insular cortices. J Comp Neurol 399: 440–468. CrossRef Medline

Shimo Y, Hikosaka O (2001) Role of tonically active neurons in primate caudate in reward-oriented saccadic eye movement. J Neurosci 21: 7804–7814. Medline

Simmons JM, Richmond BJ (2008) Dynamic changes in representations of preceding and upcoming reward in monkey orbitofrontal cortex. Cereb Cortex 18:93–103. CrossRef Medline

Stalnaker TA, Calhoon GG, Ogawa M, Roesch MR, Schoenbaum G (2010) Neural correlates of stimulus-response and response-outcome associations in dorsolateral versus dorsomedial striatum. Front Integr Neurosci 4:12. CrossRef Medline

Tanaka SC, Balleine BW, O'Doherty JP (2008) Calculating consequences: brain systems that encode the causal effects of actions. J Neurosci 28: 6750–6755. CrossRef Medline

Thorn CA, Graybiel AM (2014) Differential entrainment and learning-related dynamics of spike and local field potential activity in the sensorimotor and associative striatum. J Neurosci 34:2845–2859. CrossRef Medline

Thorpe SJ, Rolls ET, Maddison S (1983) The orbitofrontal cortex: neuronal activity in the behaving monkey. Exp Brain Res 49:93–115. Medline

Tricomi E, Balleine BW, O'Doherty JP (2009) A specific role for posterior dorsolateral striatum in human habit learning. Eur J Neurosci 29:2225–2232. CrossRef Medline

Tzavos A, Jih J, Ragozzino ME (2004) Differential effects of M(1) muscarinic receptor blockade and nicotinic receptor blockade in the dorsomedial striatum on response reversal learning. Behav Brain Res 154:245–253. CrossRef Medline

Valentin VV, Dickinson A, O'Doherty JP (2007) Determining the neural substrates of goal-directed learning in the human brain. J Neurosci 27: 4019–4026. CrossRef Medline

Wallis JD, Anderson KC, Miller EK (2001) Single neurons in prefrontal cortex encode abstract rules. Nature 411:953–956. CrossRef Medline

Wang Z, Kai L, Day M, Ronesi J, Yin HH, Ding J, Tkatch T, Lovinger DM, Surmeier DJ (2006) Dopaminergic control of corticostriatal long-term synaptic depression in medium spiny neurons is mediated by cholinergic interneurons. Neuron 50:443–452. CrossRef Medline

Wilson RC, Takahashi YK, Schoenbaum G, Niv Y (2014) Orbitofrontal cortex as a cognitive map of task space. Neuron 81:267–279. CrossRef Medline

Witten IB, Steinberg EE, Lee SY, Davidson TJ, Zalocusky KA, Brodsky M, Yizhar O, Cho SL, Gong S, Ramakrishnan C, Stuber GD, Tye KM, Janak PH, Deisseroth K (2011) Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated reinforcement. Neuron 72:721–733. CrossRef Medline

Yamada H, Matsumoto N, Kimura M (2004) Tonically active neurons in the primate caudate nucleus and putamen differentially encode instructed motivational outcomes of action. J Neurosci 24:3500–3510. CrossRef Medline

Yarom O, Cohen D (2011) Putative cholinergic interneurons in the ventral and dorsal regions of the striatum have distinct roles in a two choice alternative association task. Front Syst Neurosci 5:36. CrossRef Medline

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. CrossRef Medline

Yin HH, Knowlton BJ, Balleine BW (2005) Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. Eur J Neurosci 22:505–512. CrossRef Medline

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. CrossRef Medline

Young JJ, Shapiro ML (2011) Dynamic coding of goal-directed paths by orbital prefrontal cortex. J Neurosci 31:5989–6000. CrossRef Medline