

You might find this additional information useful...

Supplemental material for this article can be found at:

<http://jn.physiology.org/cgi/content/full/00262.2009/DC1>

This article cites 44 articles, 18 of which you can access free at:

<http://jn.physiology.org/cgi/content/full/102/1/475#BIBL>

This article has been cited by 3 other HighWire hosted articles:

Singing-Related Neural Activity Distinguishes Four Classes of Putative Striatal Neurons in the Songbird Basal Ganglia

J. H. Goldberg and M. S. Fee

J Neurophysiol, April 1, 2010; 103 (4): 2002-2014.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)

The Dorsomedial Striatum Reflects Response Bias during Learning

E. Y. Kimchi and M. Laubach

J. Neurosci., November 25, 2009; 29 (47): 14891-14902.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)

Stable Encoding of Task Structure Coexists With Flexible Coding of Task Events in Sensorimotor Striatum

Y. Kubota, J. Liu, D. Hu, W. E. DeCoteau, U. T. Eden, A. C. Smith and A. M. Graybiel

J Neurophysiol, October 1, 2009; 102 (4): 2142-2160.

[\[Abstract\]](#) [\[Full Text\]](#) [\[PDF\]](#)

Updated information and services including high-resolution figures, can be found at:

<http://jn.physiology.org/cgi/content/full/102/1/475>

Additional material and information about *Journal of Neurophysiology* can be found at:

<http://www.the-aps.org/publications/jn>

This information is current as of August 10, 2010 .

Neuronal Correlates of Instrumental Learning in the Dorsal Striatum

Eyal Y. Kimchi,¹ Mary M. Torregrossa,² Jane R. Taylor,^{1,2} and Mark Laubach^{1,3}

¹Interdepartmental Neuroscience Program, Yale University; ²Division of Molecular Psychiatry, Yale University, Connecticut Mental Health Center, Abraham Ribicoff Research Facilities; and ³John B. Pierce Laboratory, New Haven, Connecticut

Submitted 24 March 2009; accepted in final form 11 May 2009

Kimchi EY, Torregrossa MM, Taylor JR, Laubach M. Neuronal correlates of instrumental learning in the dorsal striatum. *J Neurophysiol* 102: 475–489, 2009. First published May 13, 2008; doi:10.1152/jn.00262.2009. We recorded neuronal activity simultaneously in the medial and lateral regions of the dorsal striatum as rats learned an operant task. The task involved making head entries into a response port followed by movements to collect rewards at an adjacent reward port. The availability of sucrose reward was signaled by an acoustic stimulus. During training, animals showed increased rates of responding and came to move rapidly and selectively, following the stimulus, from the response port to the reward port. Behavioral “devaluation” studies, pairing sucrose with lithium chloride, established that entries into the response port were habitual (insensitive to devaluation of sucrose) from early in training and entries into the reward port remained goal-directed (sensitive to devaluation) throughout training. Learning-related changes in behavior were paralleled by changes in neuronal activity in the dorsal striatum, with an increasing number of neurons showing task-related firing over the training period. Throughout training, we observed more task-related neurons in the lateral striatum compared with those in the medial striatum. Many of these neurons fired at higher rates during initiation of movements in the presence of the stimulus, compared with similar movements in the absence of the stimulus. Learning was also accompanied by progressive increases in movement-related potentials and transiently increased theta-band oscillations (5–8 Hz) in simultaneously recorded field potentials. Together, these data suggest that representations of task-relevant stimuli and movements develop in the dorsal striatum during instrumental learning.

INTRODUCTION

The mechanisms through which the striatum is involved in instrumental learning have not been fully established. The lateral striatum has been implicated in stimulus–response habits and the medial striatum has been implicated in goal-directed behavior (Dayan and Balleine 2002). Although several studies have reported alterations in neuronal activity in the lateral striatum during instrumental learning (Barnes et al. 2005; Carelli et al. 1997; Jog et al. 1999; Tang et al. 2007), there have been no direct comparisons of neuronal activity of putative striatal projection neurons across multiple regions of the striatum during instrumental learning. The goal of the present study was to compare activity in the medial and lateral regions of the dorsal striatum as rats learned a simple operant task.

Electrophysiological studies in the nonhuman primate have demonstrated differences between the lateral and medial striatum during associative learning (Brasted and Wise 2004; Pasupathy and Miller 2005; Tremblay et al. 1998; Williams and Eskandar 2006). These studies examined how novel stimu-

lus–reward or stimulus–response mappings altered neuronal firing rates. In some cases, changes in neuronal activity in the medial striatum occurred before changes in the lateral striatum (Williams and Eskandar 2006). However, all of these studies were carried out in animals that had extensive experience with the basic behavioral procedures. Therefore it is not clear whether these studies are relevant for understanding how naïve animals initially learn instrumental tasks.

Lesion studies suggest that the lateral and medial regions of the striatum are differentially involved in instrumental learning (Corbit and Janak 2007; Yin et al. 2004, 2005). Although the results of these behavioral experiments predict different physiological activities in both the lateral and the medial portions of striatum, the predictions have yet to be tested for instrumental learning. To investigate this issue, we modified a paradigm developed by Laubach et al. (2000) to study learning-related changes in the motor cortex. That is, we implanted arrays of electrodes into the medial and lateral regions of the striatum in naïve rats and trained the animals to perform a simple operant task. The task involved animals responding in a “response port” for a sucrose reward that was available after a random interval. Sucrose availability was signaled by an acoustic stimulus and was delivered in an adjacent “reward port.” Rats learned to perform this task within several hundred trials. Devaluation methods (Dickinson et al. 1983) were used to determine whether responding was habitual or goal-directed at two stages of training (after earning either 150 or 700 rewards). By recording neuronal activity (spike trains and field potentials) simultaneously in the medial and lateral regions of the striatum throughout the period of training, we were able to directly observe changes in task-related neuronal activity in the dorsal striatum during instrumental learning.

METHODS

Chronic neuronal ensemble recording in the striatum

Naïve rats ($n = 5$; male Sprague–Dawley rats, ~400–500 g, Charles River Laboratories) were chronically implanted with multi-electrode recording arrays using standard methods as previously described (Kimchi and Laubach 2009). The Institutional Animal Care and Use Committees at the John B. Pierce Laboratory approved all procedures. Arrays were composed of 16 stainless steel wires, arranged in 2×8 configurations with 250- μ m spacing between wires (Neuroline, New York, NY). One array in each rat was placed in the medial striatum (0.2 mm anteroposterior [AP], 2 mm mediolateral [ML], –4.2 mm dorsoventral [DV]) and one in the lateral striatum (0.3 mm AP, 4 mm ML, –4.8 mm DV). The hemispheres for each implant were alternated across animals. Once implanted, multi-electrode arrays were fixed and recorded neuronal activity from the same spatial location for the duration of training.

Neuronal activity was recorded simultaneously from the 32 implanted electrodes using a Plexon Multichannel Acquisition Processor

Address for reprint requests and other correspondence: M. Laubach, John B. Pierce Laboratory, 290 Congress Avenue, New Haven, CT 06519 (E-mail: mlaubach@jbpierce.org).

system. Electrical signals were recorded from all electrodes and processed on-line (using an oscilloscope and audio amplifier) and off-line (using the Plexon Off-line Sorter) to identify the spiking activity of individual striatal neurons. Neuronal signals were amplified $\times 1,000$ – $20,000$. Spike activity was thresholded by voltage and waveforms that crossed the threshold were time-stamped, sampled, and stored at 40 kHz. Unique waveforms were identified on-line and recorded. On-line root-mean-square values, while rats were quietly resting, were typically $15 \mu\text{V}$ (calculated within Plexon On-line Sorter). Waveforms were then processed off-line (Plexon Off-line Sorter) to remove artifacts and sorted into different units using principal component analysis and template-based methods. After processing, units had to meet several criteria to be considered single units: 1) mean peak-to-peak voltage had to be $\geq 100 \mu\text{V}$; 2) signal-to-noise ratio had to be $\geq 3:1$; 3) fewer than 2% of interspike intervals (ISIs) could be $< 2 \text{ ms}$; 4) the mode of the ISI histogram had to be $> 5 \text{ ms}$; and 5) the distribution of maximal waveform points had to be relatively normal (skewness < 0.75). This latter measure ensured that waveforms were isolated from the noise threshold.

Additionally, we evaluated neuronal activity during the drinking period, when animals stood still within the reward port, to ensure that neurons were stationary. This time was defined as 2 s following the onset of delivery of sucrose solution. During drinking, neurons had to fire at least once during 10% of drinking periods and had to have Z-scores < 3.5 on a runs test (Siegel 1956) to be considered stationary. Varying the Z-score criterion from 2 to 4 did not change either the qualitative pattern or the statistical significance of the results. These characteristics were additionally verified by on-line and off-line experimenter assessment. Prior to each behavioral session, wideband signals were recorded for 10 min to assess the quality of the implanted electrodes (sampling at 20 kHz, filtering between 0.5 Hz and 5.9 kHz). During behavior, local field potentials (LFPs; analog filtering only from 0.5 Hz to 5.9 kHz) were recorded from three widely spaced electrodes that lacked clearly resolved units on each array (three medial and three lateral for a total of six per rat). LFPs were amplified $\times 10,000$ and sampled at 1 kHz.

For the results reported in this study, we focused on phasically active neurons from the striatum. We screened for neurons that fired tonically and removed them from the neuronal database. Tonic active neurons fire at $> 10 \text{ Hz}$ and show an irregular distribution of ISIs (Kimura et al. 1990; Yamada et al. 2004). These cells were rarely recorded and comprised $< 3\%$ of all recorded neurons. We excluded 4 cells from lateral striatum (of 262) and 13 from medial striatum (of 333) for the analyses carried out in this study.

We recorded from an average of 0.40 neuron per wire in the lateral striatum and an average of 0.50 neuron per wire in the medial striatum. There was no difference in the mean peak-to-peak amplitude of recorded units in these two striatal regions (lateral: $194 \pm 138 \mu\text{V}$; medial: $206 \pm 137 \mu\text{V}$; mean $\pm$ SD, t -test = 0.28). Therefore our results do not appear to be due to differences in our ability to isolate units in the two portions of the striatum.

Operant task and training procedures

After a 1-wk recovery from electrode implantation, rats were given limited access to food, by allowing them to eat $\leq 20 \text{ g}$ of lab chow in the home cage during a period of 90 min per day. Weights were monitored to ensure that rats maintained about 85% of their initial weights. Behavioral training commenced after 5 days of regulated access to food.

Rats were trained to collect sucrose from a reward port using a series of operant schedules (Fig. 1). Training consisted of 1) one "autoshaping" session ("Auto"), in which sucrose could be collected in the reward port approximately every 60 s, with the time of sucrose delivery chosen from an exponential distribution; 2) one session of response training, using a fixed ratio schedule of reinforcement ("FR1"), in which sucrose was earned 50 times, after each entry into

the response port; 3) one session of response training, using a random-interval schedule of reinforcement ("RI20"), in which sucrose was earned 100 times, after the first entry into the response port at the end of an interval chosen from an exponential distribution with mean of 20 s; and 4) five sessions of response training, using a random-interval schedule of reinforcement ("RI40"), in which sucrose was earned 100 times, after the first entry into the response port at the end of an interval chosen from an exponential distribution with mean of 40 s. In all sessions, rats either had to collect sucrose within 5 s of the activation of the reward port or had to initiate a new sequence of responding in the response port under a renewed schedule to collect the next reward.

Behavioral training took place in a single arena that was specifically constructed for electrical recordings (Med Associates, St. Albans, VT). The floor of the arena was rectangular ($24 \times 30 \text{ cm}$ [depth $\times$ width]). All walls and floor bars were made of acrylic plastic and the long walls sloped diagonally outward. Two "nosepoke" devices, called response ports, flanked a central fluid dispenser, called the reward port. For any given subject, only one response port was active—responses on the inactive response port had no consequence. Entries into the active response port and the reward port were monitored by infrared photobeams (John B. Pierce Laboratory Instruments Shop). Sucrose solution was delivered at a central spout and the pump for the spout was activated 100 ms after entry into the reward port. The pump delivered $60 \mu\text{l}$ of 20% sucrose solution, calibrated by adjusting the duration of pump activation (i.e., $60 \mu\text{l}$, 1.7 s). The pump was silent within the behavioral chamber. Therefore an acoustic noise stimulus (white noise burst, 60 dBa, generated by RP2.1 Processor, TDT Technologies) was presented, using a speaker (ES1, TDT Technologies) located directly above the reward port, to indicate sucrose delivery. The stimulus was presented 100 ms after nosepoke scheduled for reinforcement, i.e., after a nosepoke entry into the response port. The stimulus remained on until either the rat collected a reward or the window to collect a reward had elapsed.

Behavioral devices were interfaced using a digital input–output card (PCI-DIO-96, National Instruments). Behavior was monitored using an infrared camera and videotaped for off-line analysis. An incandescent bulb (4 W at 6 V) was used as the houselight, which was located on the side of the chamber opposite to the response and reward ports. The chamber was placed within a sound-attenuating box (Med Associates) lined with additional sound foam. A fan was located on the inside of the box to provide constant background noise and ventilation. The chamber was placed on a steel plate within a Faraday cage of copper wire for electromagnetic shielding. Protocols were controlled using custom-written software using the Matlab Data Acquisition Toolbox (The MathWorks, Natick, MA) and the freely available Psychophysics Toolbox (Brainard 1997). Chambers were cleaned following each session.

Histology

At the conclusion of the recording sessions, rats were killed with an intraperitoneal injection of pentobarbital ($> 100 \text{ mg/kg}$). Microstimulation lesions were made on some wires to help identify electrode locations. Rats were then perfused with saline followed by 4% formaldehyde. Brains were extracted, stored overnight in 25% sucrose, cut horizontally on a freezing microtome, stained with thionin, dehydrated, mounted, and coverslipped. Electrode tracks were identified using light microscopy and registered to a rat brain atlas (Paxinos and Watson 1998). Three-dimensional models and two-dimensional projections of the rat striatum were constructed using freely available software written for Matlab (available at <http://spikelab.jpierce.org/3DAnatomy>). Electrode placements are illustrated in Fig. 2.

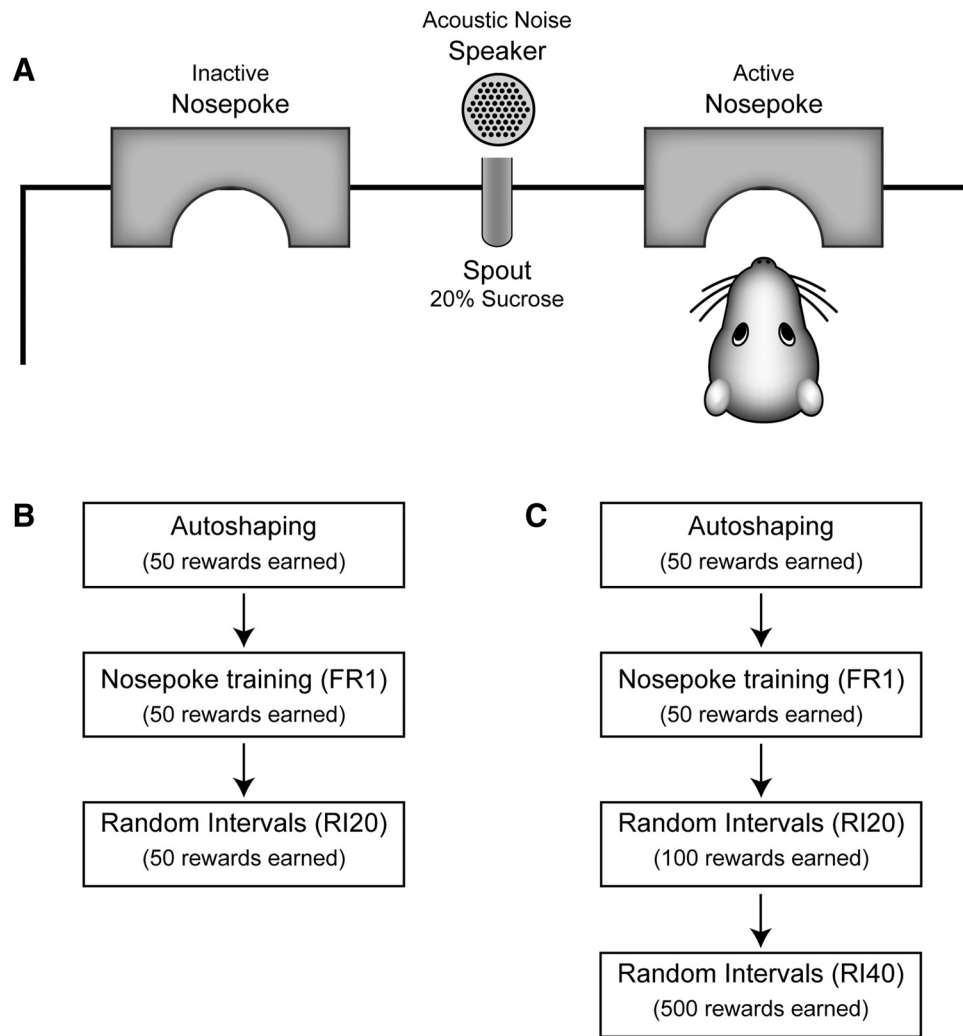


FIG. 1. Task used to study instrumental learning. *A*: layout of the behavioral devices. There were 2 response ports with photobeams to detect entries into the response port. The response ports flanked a central fluid dispenser that delivered 20% sucrose solution into the reward port (here a spout; photobeam not shown). Using operant methods, animals learned to make responses in the active response port to receive sucrose in the reward port. Responses in the inactive port had no consequence. The availability of sucrose was signaled by an acoustic noise stimulus. In physiological experiments, a speaker placed directly above the glass spout delivered an acoustic noise stimulus when sucrose solution was available. In behavioral devaluation experiments, a standard fluid dipper made an acoustic stimulus when activated. *B* and *C*: rats were trained over a series of sessions, using a progression of operant schedules of reinforcement, to make repetitive entries into the active response port. Rats first experienced an autoshaping session, in which sucrose was delivered into the reward port only in the presence of the acoustic stimulus, which occurred approximately every 60 s. There was no requirement for responding in the response port during the autoshaping sessions. Next, rats were required to enter the response port to earn rewards (FR1, "Fixed Ratio of 1"). Finally, rats could earn rewards only when they entered the response port after the end of a pseudorandomly chosen interval (with mean of 20 or 40 s, chosen from an exponential distribution). Sessions were run daily until the rats earned 50 (Autoshaping, FR1) or 100 sucrose rewards (RI sessions). All rats that were implanted with arrays of electrodes ($n = 5$) were trained using the full series of procedures for a total of 700 earned rewards, as shown in *C*. Rats that were trained for devaluation ($n = 32$) were divided into 2 groups, with half of the rats trained until they received 150 rewards (*B*) and the other half trained to 700 rewards (*C*).

Devaluation of sucrose

A separate group of rats ($n = 32$; adult male Sprague–Dawley rats, Charles River Laboratories) was trained on the instrumental task. The Yale University Institutional Animal Care and Use Committee approved all procedures. For these experiments, we used the training procedures described earlier for animals with neuronal recordings and described in Fig. 1. One group of 16 rats was limited to 50 rewards earned on the RI20 training day and received no further training (150 total rewards earned, Fig. 1*B*), after which they were evaluated using devaluation methods. Another group of 16 rats was evaluated using devaluation methods after earning 700 total rewards (Fig. 1*C*).

Behavioral chambers used standard equipment from Med Associates. On the center of one narrow wall, a fluid dispenser (ENV-202M) delivered 60 μ l of 20% sucrose solution. In behavioral chambers, activation of the fluid dipper was accompanied by an

acoustic stimulus intrinsic to the operation of the device and that signaled sucrose availability. Two "nosepoke" devices (ENV-114), called response ports, flanked the central fluid dispenser (ENV-202M), called the reward port. For any given subject, only one response port was active; responses on the inactive response port had no consequence. Chambers were placed within sound-attenuating boxes and a fan was located on the inside of the box to provide constant background noise and ventilation. The house light (ENV-215M) and fan were turned on at the beginning of the behavioral sessions and remained on until the end.

After training, a subset of rats had the sucrose reward devalued by conditioned taste aversion (CTA) training over 6 days ($n = 32$ rats: 16 trained to 150 rewards earned and 16 trained to 700 rewards earned). On days 1 and 3, rats were placed singly in novel plastic cages with free access to the sucrose solution used in instrumental training for 30

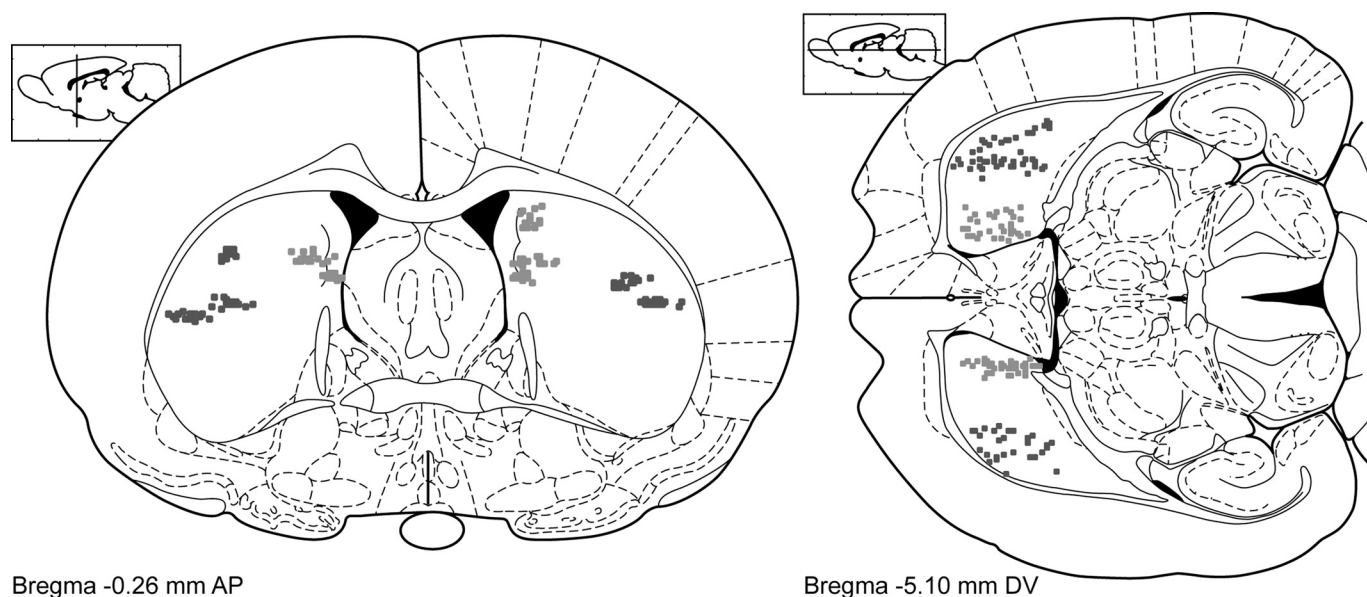


FIG. 2. Localization of recording sites in the dorsal striatum. Electrode tips are depicted as dots. Dark dots represent electrodes in the lateral part of dorsal striatum. Light dots represent electrodes in the medial part of dorsal striatum. Coronal and horizontal sections are shown on the *left* and *right*, respectively, at the level of the striatum that was implanted for this study. Recordings of neuronal activity were made across the anterior to posterior extent of the lateral and medial striatum.

min. The amount of sucrose consumed on each day was measured by weighing the bottles before and after the 30-min consumption period. On day 5, rats were placed in the same operant chamber used in instrumental training to facilitate transfer of the CTA to the instrumental context. In this session, the response ports were removed and 120 presentations of 10 s of sucrose were presented on a fixed-time 20-s (FT20) schedule. The sucrose solution receptacles were weighed before and after these sessions to determine the amount of sucrose consumed.

Each group of rats was divided into devalued and control groups. Immediately after each sucrose consumption session, half of each group of rats was injected intraperitoneally with 0.6 M LiCl and the other half with 0.9% NaCl in a volume of 5 ml/kg and are hereafter referred to as “devalued” and “control” rats, respectively. On days 2, 4, and 6, devalued rats received injections of NaCl and control rats received injections of LiCl, such that the injections were not temporally associated with sucrose or daily food. In this manner all animals received an equal number of injections of LiCl, but only the devalued animals formed a CTA for sucrose. There were thus four groups based on training and devaluation status: eight rats trained until 150 rewards and devalued, eight rats trained until 150 rewards and serving as controls, eight rats trained until 700 rewards and devalued, and eight rats trained until 700 rewards and serving as controls.

Habit testing was done following CTA training. Rats were tested for their propensity to enter in the previously active port in an extinction session. These sessions were 5 min in duration and were identical to the random-interval training sessions except no sucrose or acoustic stimulus was presented. The numbers of responses in this test session were normalized for each rat to the number of responses made in the first 5 min of the last random-interval training session. The number of magazine entries for each rat during the habit test was also recorded.

Finally, a sucrose consumption test was carried out. The amount of sucrose consumed in the operant chamber was determined immediately following the habit test to verify the effectiveness of CTA training in the instrumental context. At the end of the habit test, sucrose-filled receptacles were placed in the chamber. Rats were given 120 sucrose presentations on an FT20 schedule as described earlier. The sucrose receptacles were weighed before and after this session to determine the amount of sucrose consumed.

Data from the devaluation experiments were analyzed using a two-way ANOVA and Dunnett’s post hoc test, where appropriate, with SPSS 16 software (SPSS, Chicago, IL).

Analysis of spike activity

Data analysis for physiological experiments was done using Matlab and R (<http://www.R-project.org>). Custom-written software and m files from Plexon were used to analyze neuronal and behavioral data in Matlab. Exploratory and statistical analyses were done using the statistics toolbox for Matlab and a variety of functions in R. Data were exchanged between these programs using the R.matlab library for R. For analyses of neuronal activity, the timestamps from identified single units were aligned to the time of the presentation of the acoustic noise stimulus signaling sucrose availability to create perievent rasters and perievent time histograms. For nonrewarded entries into the response port, neuronal activity was aligned to the time at which the stimulus would have occurred and rewards would have become available (100 ms after entry into the response port). On these “trials,” no stimulus was delivered because the random interval had not yet expired. Response time (RT) was defined as the time from onset of the stimulus until withdrawal from the response port. Movement time (MT) was defined as the time following withdrawal from the response port until entry into the reward port.

To control for the nonspecific effects of movement, we selected a set of nonrewarded entries into the reward port that were matched in duration to the rewarded entries into the reward port. Matching was done recursively within blocks of 50 earned rewards. For each block of data, we carried out the following steps: 1) generated lists of rewarded and nonrewarded movements (withdrawal from response port), 2) measured the duration of the rewarded and nonrewarded movements, 3) calculated the medians for each list, 4) found the closest match to the median rewarded movement in the list of nonrewarded movements (smallest difference in movement duration), 5) stored the times of the best matches for movement duration in new lists, and 6) removed the matching pair of responses from the first list of responses. The process was repeated until all rewarded responses were matched with nonrewarded responses and generated a list of nonrewarded responses with movement durations that were similar to the rewarded responses.

Neuronal activity was analyzed around the times of presentation of the acoustic stimulus and the times of latency-matched head entries in the absence of the stimulus. Based on exploratory analysis of raster plots, it was clear that neurons with response-related firing showed altered firing rates just after stimulus presentation. Therefore we compared neuronal firing rates in two time windows around the stimulus (−30 to +20 and +20 to +70 ms) using signed-rank tests ($P < 0.05$). These windows were chosen to maximize poststimulus/premovement-related processing while keeping the windows of analysis as brief as possible: +20 ms was chosen because this was the earliest time that spiking and LFPs became modulated in the striatum following stimulus onset; +70 ms was chosen because this was the median latency of withdrawal from the response port at the end of training; and −30 ms was chosen to ensure equal time prior to the stimulus. Changes in firing rates around the stimulus were assessed in nonoverlapping blocks of 50 reinforced trials, yielding one block from each of the first 2 days of training, and two blocks from each successive day of training. Extending windows later increased the percentage of neurons modulated overall, whereas shifting them earlier decreased the percentage of neurons modulated overall. Similar analyses were done for neuronal firing rates around the times of withdrawal from the response port, based on the signal from the photobeam inside the response port, and for times of movements that were matched by response duration, as described earlier.

Analysis of modulations in firing rate: empirical fluctuation processes

We also analyzed modulations of neuronal firing rates using structural change tests, based on methods available in the *strucchange* library (Zeileis et al. 2002) for R (<http://www.R-project.org>). The signed-rank test is specific for firing rate changes within a narrow time window, but is less sensitive to longer timescale modulations of firing rate. For this reason we searched for a test that could compare fluctuations in neuronal firing rates to fluctuations expected for random data (i.e., Brownian motion). By using structural change tests, we were able to analyze changes in firing rate without having to assume that firing rates changed during a specific user-defined window around the stimulus. We used several different types of structural change tests, based on the cumulative sums of standardized residuals (Brown et al. 1975), the cumulative sums of ordinary least squares (OLS) residuals (Ploberger and Kramer 1992), and recursive estimates based on the raw data (Ploberger et al. 1989). All three methods gave equivalent results, based on analysis with ANOVA (with $P \gg 0.1$), with respect to the fractions of neurons with task-related modulation in firing rate. As a result, we describe the details on only the simplest of these methods, the recursive estimates test, in the following text.

Following Zeileis et al. (2002), the basic idea of structural change tests is to estimate a simple linear model to predict the spike probability in a given bin i , based on preceding spike probabilities

$$y_i = x_i \beta_i^T + u_i \quad (1)$$

Here, y_i is the predicted spike probability (the dependent variable) for i bins from 1 to n , x_i is the observed spike probability, β_i^T is the set of coefficients (transposed) for preceding bins (to give the linear fit), and u_i is the set of residuals. The null hypothesis H_0 was that there was no structural change, specifically that the residuals from the linear model given in Eq. 1 were unchanged over the series of spike probabilities, i.e., $\beta_i = \beta_0$. An empirical fluctuation process (Ploberger et al. 1989) was measured as

$$y_n(t) = \frac{\sqrt{i}}{\sigma \sqrt{n}} [x^{(i)T} x^{(i)}]^{1/2} [\beta^{(i)} - \beta^{(n)}] \quad (2)$$

where $i = \lfloor k + t(n - k) \rfloor$ with $t \in [0, 1]$. In this study, we used the function *efp* in the *strucchange* library for R to estimate empirical fluctuation processes.

The analysis was run for the peristimulus period (± 0.5 s around stimulus onset, using 1-ms bins) and for the period of sucrose consumption (0–1.0 s after pump activation, using 1-ms bins). The idea for this method was to estimate the spike probability in a given bin based on the history of spiking during previous bins in the peristimulus window. To estimate the significance of a given fluctuation, the maximum value in the empirical fluctuation process [$S_e = \max \|Y_n(t)\|$] was compared with that expected from Brownian motion (or Weiner process), as described in Zeileis (2000). The range of random fluctuation was used to estimate the upper and lower bounds used for significance testing. Because there is an increase in the cumulative level of random fluctuation over time, the upper and lower bounds for random fluctuation increased over the peristimulus period (note the solid gray lines in the *bottom plots* in both panels of Fig. 7). In this study, we used the function *sctest* in the *strucchange* library for R to estimate significance levels for the empirical fluctuation processes.

Assessing learning-related changes in neuronal activity: change-point analysis

We used change-point analysis (Chow 1984) to estimate when there was change in the fractions of neurons that showed significant task-related modulation in firing rate (based on empirical fluctuation processes) or changes in firing rate around the stimulus (based on signed-rank tests of firing rate in narrow windows before and after the stimulus). Methods were used from the *strucchange* library for R (Zeileis et al. 2002). This analysis provides an unbiased method to determine when a time series has undergone a statistically reliable change. A simple linear model, as in Eq. 1, was fit to the data series (fractions of task-modulated neurons over the successive blocks of 50 earned rewards). Then, we calculated the F statistic as follows (Chow 1960)

$$F = \frac{u^T u - e^T e}{e^T e / (n - 2)} \quad (3)$$

where u represents the residuals from fitting a model for every pair of n blocks (local data window) and e represents the residuals from fitting a model over all n blocks (full data). The F statistic follows the χ^2 distribution with 1 degree of freedom and the P values for significance testing in a change-point framework were developed by Hansen (1997). In this study, we carried out change-point analysis using the functions *Fstats* and *breakpoints* in the *strucchange* library for R.

Analysis of local field potentials

LFPs were analyzed by constructing perievent time series for six simultaneously recorded signals (three medial and three lateral per rat). These recordings were taken from electrodes that did not have isolated units. The range of the LFP was defined by the maximum voltage – minimum voltage in the period from 0 to 250 ms after stimulus onset on each trial. Fourier transforms were performed using the Matlab Signal Processing Toolbox *fft* function. Spectral analysis was conducted using bins of 1 ms from the time of the stimulus to +512 ms after the stimulus. Spike-field coherence was computed between perievent spike trains and LFP time series using NeuroSpec 2.0 within Matlab (<http://www.neurospec.org>; Rosenberg et al. 1989). NeuroSpec 2.0 performs multivariate Fourier analysis of time series and point processes. Similar to the LFP analysis, spike field coherence was performed on a single window of 512 ms following the onset of the stimulus, with 1-ms resolution. All spikes during this window from each single unit were used for spike-field coherence analysis. Spikes were treated as point processes within the NeuroSpec code (using type 1 analysis

in the function *sp2a_m1* within NeuroSpec 2.0). Spike-field coherence was performed for all pairs of spiking activity and LFP signals recorded from the same hemisphere and same region of striatum (lateral or medial). Significant spike-field coherence at a given frequency was determined by a coherence value greater than the 95% confidence limit.

RESULTS

Acquisition of the instrumental task

Rats quickly learned the basic behavioral procedure. They showed increased rates of entries into the response port over

the period of training [$F_{(1,29)} = 44.6$, $P \ll 0.001$; Fig. 3A]. They came to respond selectively in the reward port after the stimulus [$F_{(1,62)} = 580.0$, $P \ll 0.001$; Fig. 3B]. That is, rats were increasingly more likely to move from response port to reward port only following presentation of the acoustic stimulus. These changes were accompanied by reductions in response times (RTs) to the stimulus (Fig. 3C), defined as the time taken to withdraw from the response port after stimulus onset, and in movement times (MTs) to the reward port (Fig. 3D), defined as the time following withdrawal from the response port until entry into the reward port. Both of these

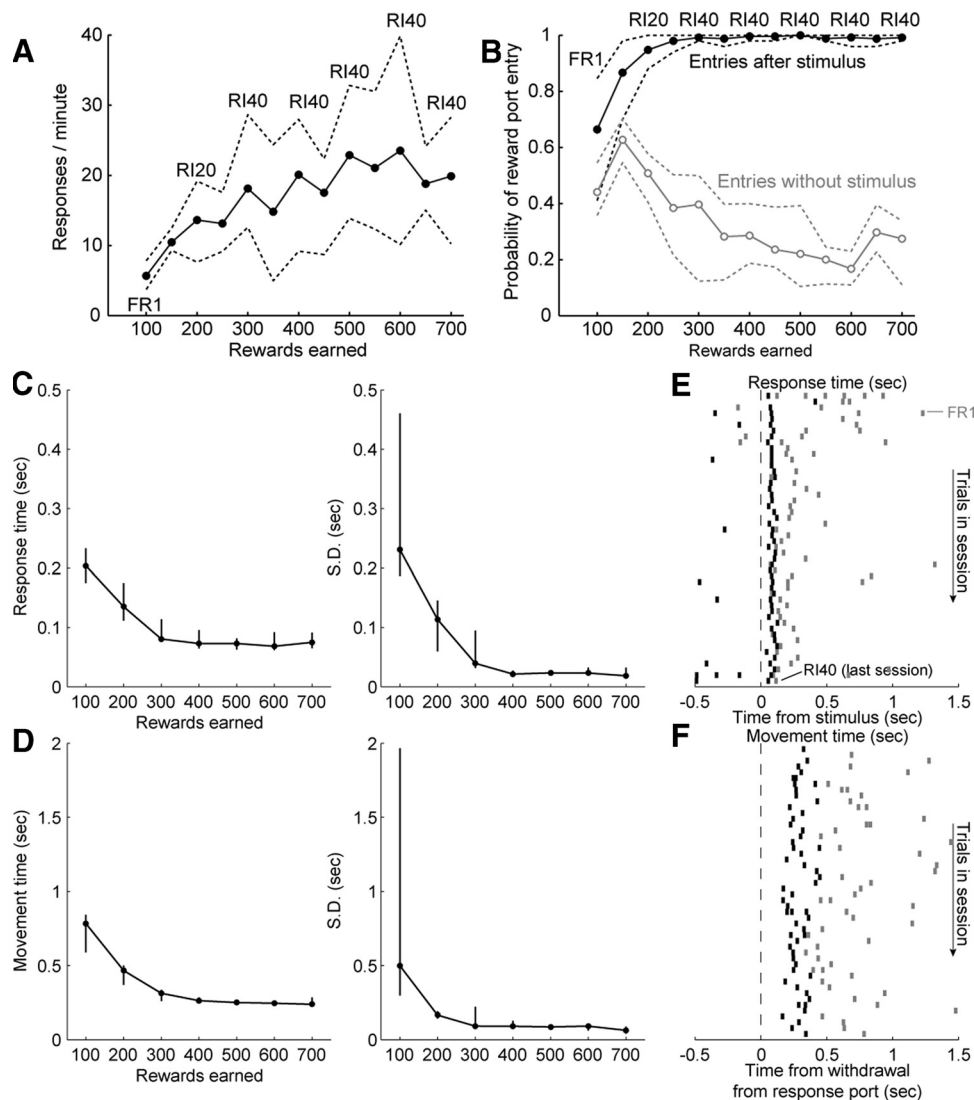


FIG. 3. Acquisition of the instrumental task. A: the mean rate of responding per minute in the response port is plotted as a function of the number of sucrose rewards earned. Bands (dashed lines) represent the 95% confidence interval around the means (estimated using bootstrap methods). The types of operant schedule used for each block of rewards are indicated within the plot. The rate of responding in the response port increased with training, reaching an asymptotic level after about 500 rewards were earned (3rd session with RI40 schedule of reinforcement). B: the probabilities of entry into the reward port in the presence and absence of the acoustic noise stimulus are plotted as a function of the number of rewards earned. As in A, the bands indicate the 95% confidence intervals. Rats were increasingly more likely to enter the reward port following the noise stimulus (black) compared with control responses in the absence of the stimulus (gray). C: response times (RTs) to the noise stimulus were measured as the time between the onset of the stimulus and withdrawal from the response port. The mean (left) and SD (right) of the RTs decreased over training, especially during the initial sessions (FR1, RI20, RI40; lines depict medians and interquartile ranges across subjects). D: movement time (MT) to the spout, from response port withdrawal to reward port entry, also decreased during training. There was a very large decrease in the variability of these movements between the 1st (FR1) and 2nd (RI20) sessions. E and F: response and MTs are shown for a single rat. Raster plots show the timing of response time from onset of the stimulus to withdrawal from the response port (top panel) and MT from withdrawal from the response port to entry into the reward port (i.e., contact with the spout, bottom panel). Data from the first 50 rewarded trials from the first session (FR1) are shown as gray tick marks. Data from the last 50 rewarded trials from the last session (RI40) are shown as black tick marks. Rewarded trials are ordered from early to late in each session, with the early part of each session shown at the top of each raster.

measures were reduced in a progressive manner over the training sessions [RT: $F_{(1,29)} = 33.0$, $P < 0.001$; MT: $F_{(1,29)} = 45.2$, $P < 0.001$]. RTs were reduced from 0.20 ± 0.05 s (median $\pm$ interquartile range) in the FR1 session to 0.13 ± 0.06 s in the RI20 session. Average response times to the stimulus remained <0.1 s over subsequent sessions with the RI40 schedule. The time taken to collect sucrose was also reduced from 0.78 ± 0.25 s in the FR1 session to 0.47 ± 0.13 s in the RI20 session. Rapid movements to the reward port (movement times <0.5 s) were observed throughout subsequent sessions with the RI40 schedule. These effects are shown for a single subject in Fig. 3, *E* and *F*. Last, nosepoke responses in the inactive port were rare, with rates of $<1/\text{min}$ throughout training (data not shown).

Devaluation experiments

To determine whether our task led to the formation of habitual responding, two further groups of 16 rats were trained until they earned either 150 or 700 rewards (Fig. 1). After training, each group of rats was divided into devalued and control subgroups. The animals in the devalued groups experienced a conditioned taste aversion (CTA) for sucrose. CTA training produced a significant decrease in sucrose consumption only in devalued animals [$F_{(1,27)} = 6.75$, $P = 0.015$; data prior to habit testing not shown], with no effect of training duration.

Following devaluation, rats were tested for their propensity to enter the response port during an extinction session. There was no significant effect of devaluation [$F_{(1,28)} = 0.059$, $P > 0.05$] or the duration of training [$F_{(1,28)} = 0.076$, $P > 0.05$] on the relative proportion of entries into the response port versus the last day of training (Fig. 4*A*). This result is consistent with entries into the response port being habitual. By contrast, the absolute number of entries into the reward port during the extinction sessions was significantly reduced in the devalued animals, with a significant effect of devaluation [$F_{(1,28)} = 22.31$, $P < 0.001$], no effect of training [$F_{(1,28)} = 0.687$, $P > 0.05$], and no interaction between devaluation and training [$F_{(1,28)} = 0.076$, $P > 0.05$]. The effect of devaluation was similar for both proportions and absolute number of entries. These results indicate that entries into the reward port remained goal-directed throughout training (Fig. 4*B*). Sucrose consumption was also measured in the instrumental context immediately following the extinction session (Fig. 4*C*). Devaluation reduced free consumption of sucrose only in the devalued animals. Two-way ANOVA revealed a significant effect of devaluation [$F_{(1,28)} = 4.43$, $P < 0.05$] and no effect of training [$F_{(1,28)} = 0.135$, $P > 0.05$]. Together, these results provide evidence for the rapid formation of habitual responding in the response port and for the persistence of goal-directed responding in the reward port.

Learning-related changes in neuronal activity

To examine neuronal correlates of instrumental learning, we implanted five rats with arrays of microwire electrodes in the lateral and medial regions of the striatum in opposite hemispheres (Fig. 2). The animals were then trained using the full behavioral procedures described in Fig. 1*C*, for a total of 700 earned rewards. Neuronal correlates of learning were assessed

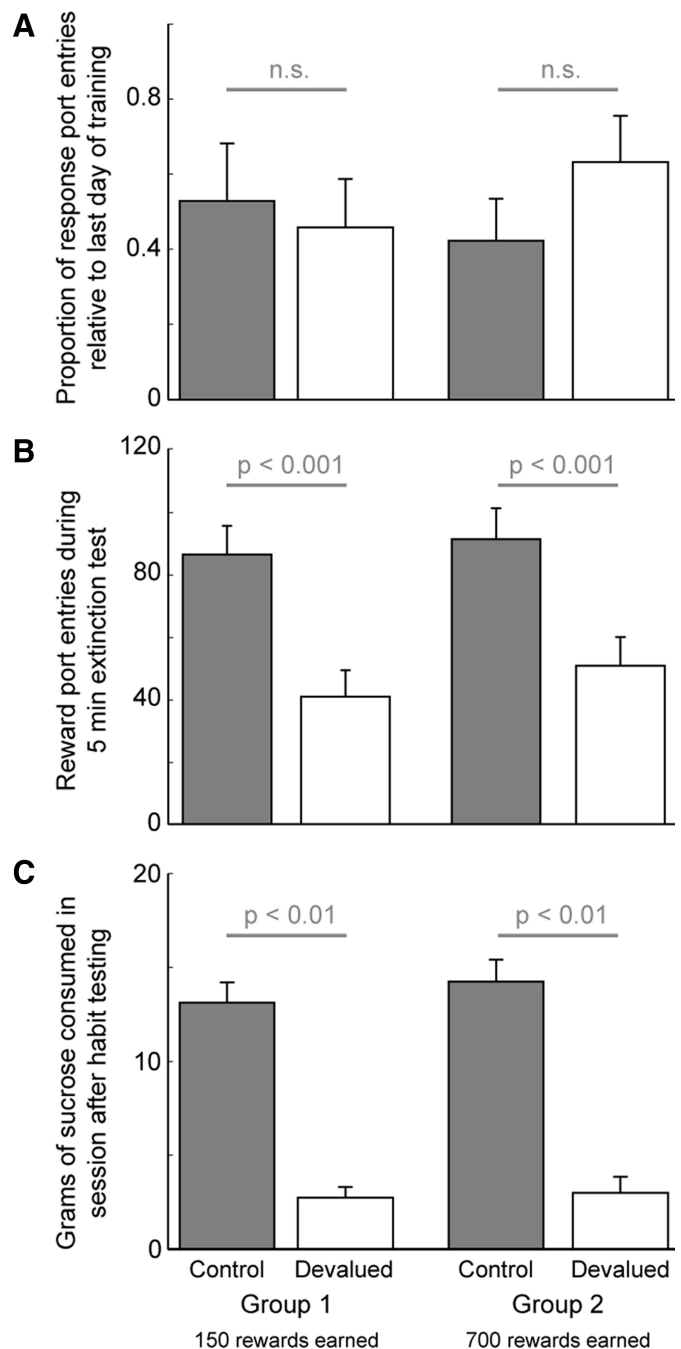


FIG. 4. Effects of devaluation on instrumental behavior. *A*: devalued (white) and control (gray) animals made an equal number of head entries in the response port during extinction testing. Data are expressed as the proportion of entries into the response port relative to the last day of training (means $\pm$ SE). *B*: entries into the reward port during extinction testing were significantly reduced in the devalued animals (white) compared with control animals (gray). Data are depicted as the raw number of entries into reward ports during the 5-min extinction session (means $\pm$ SE). The effect of devaluation was similar when the data were expressed as a proportion of the last training day. *C*: sucrose consumption after the extinction session decreased in devalued rats (white) compared with controls (gray).

for 258 neurons from the lateral striatum and 320 neurons from the medial striatum.

On some electrodes, we were able to record neuronal activity across the series of training sessions. In such cases, we observed increasing modulations of firing rates over the period

of training (Fig. 5A). Within ensembles of simultaneously recorded neurons, we observed prominent modulations of neuronal firing rates at the end of training compared with the initial training sessions (Fig. 5B). The timing of task-related modulations in firing rates was assessed using population averages (Fig. 6). As recently reported (Berke 2008), we found that population averages from the striatum were weakly modulated (Fig. 6A). However, by squaring the average response of each neuron prior to averaging, we were able to detect the main epoch of modulated firing rate in the task, i.e., the time when rats withdrew from the response port and moved to the reward port (Fig. 6B). Importantly, the timing of peaks in the population averages closely matched the latencies of the animals' movements between the ports (see boxplots in each panel in Fig. 6). There was no consistent pattern of modulation in either region of striatum during the subsequent reward consumption (not shown).

To assess the significance of these changes in neuronal activity, we tested for significant modulations in each neuron's task-related activity by using a structural change test (Zeileis et al. 2002), using a criterion of $P < 0.05$ (Fig. 7). We used an

empirical fluctuation process to determine whether a neuron had a statistically significant modulation in its average firing rate during the peristimulus epoch (± 0.5 s, 1-ms bins). This method is illustrated for two neurons that were significantly modulated (Fig. 7, A and B). The fractions of neurons across the entire population that showed modulations in firing rate were summarized over blocks of 50 earned rewards (Fig. 8, A and B). As shown in Fig. 8A, the fractions of modulated neurons in both regions of the striatum increased with training [$F_{(1,13)} = 7.6$, $P < 0.001$]. More neurons were modulated in the lateral striatum compared with the medial striatum throughout the training period [$F_{(1,1)} = 10.05$, $P < 0.01$]. Similar effects were found for neuronal activity synchronized to the time when fluid was delivered from the spout within the reward port, with significantly more neurons modulated during drinking (0–1.0 s after onset of fluid) in the lateral striatum compared with the medial striatum [Fig. 8B; effect of training: $F_{(1,13)} = 3.54$, $P < 0.02$; effect of area: $F_{(1,1)} = 59.62$, $P \ll 0.001$]. In both areas and both task epochs, the major change in task-related modulations occurred at the earliest stage of training, during the session using the FR1 schedule of reinforcement.

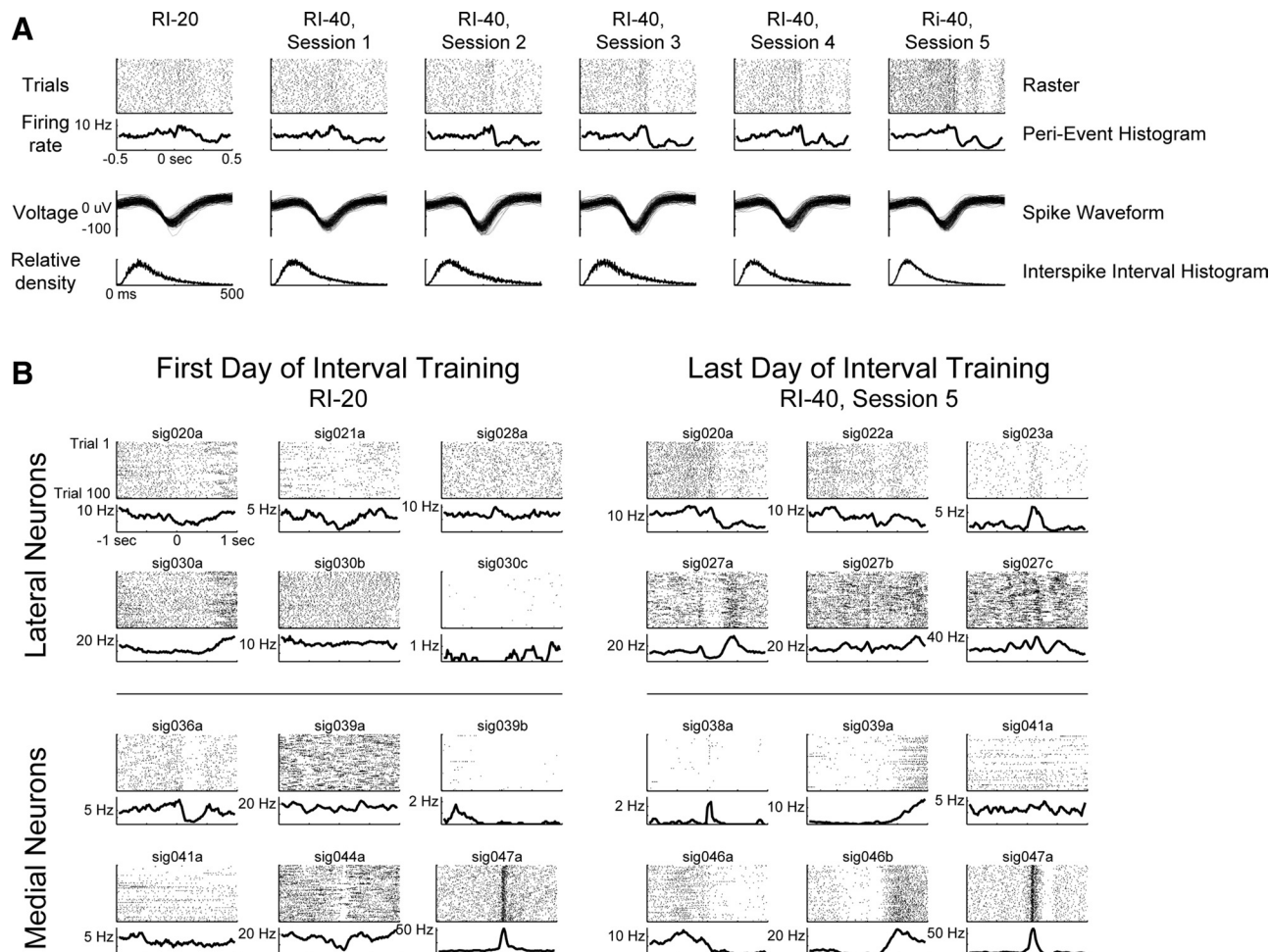


FIG. 5. Examples of learning-related changes in striatal activity. A: spike activity, recorded from a single electrode, is shown for the series of training sessions (columns) with random-interval procedures. Firing rates changed progressively, becoming increasingly modulated in each subsequent training session. The 1st row in each column contains raster plots. The 2nd row shows the average perievent histograms. Perievent histograms were generated for display purposes by convolving the spike trains (1-ms bins) with Gaussian windows (20 ms SD). The 3rd row shows the spike waveforms. The 4th row shows the interspike intervals (1-ms bins). B: changes in striatal firing rates were also apparent over ensembles of simultaneously recorded neurons in the medial and lateral striatum. Here, we show activity from all neurons recorded in one rat in the first and last sessions of random interval training. Neuronal activity is more visibly modulated at the end of training, particularly sharply around the stimulus in the lateral striatal neurons.

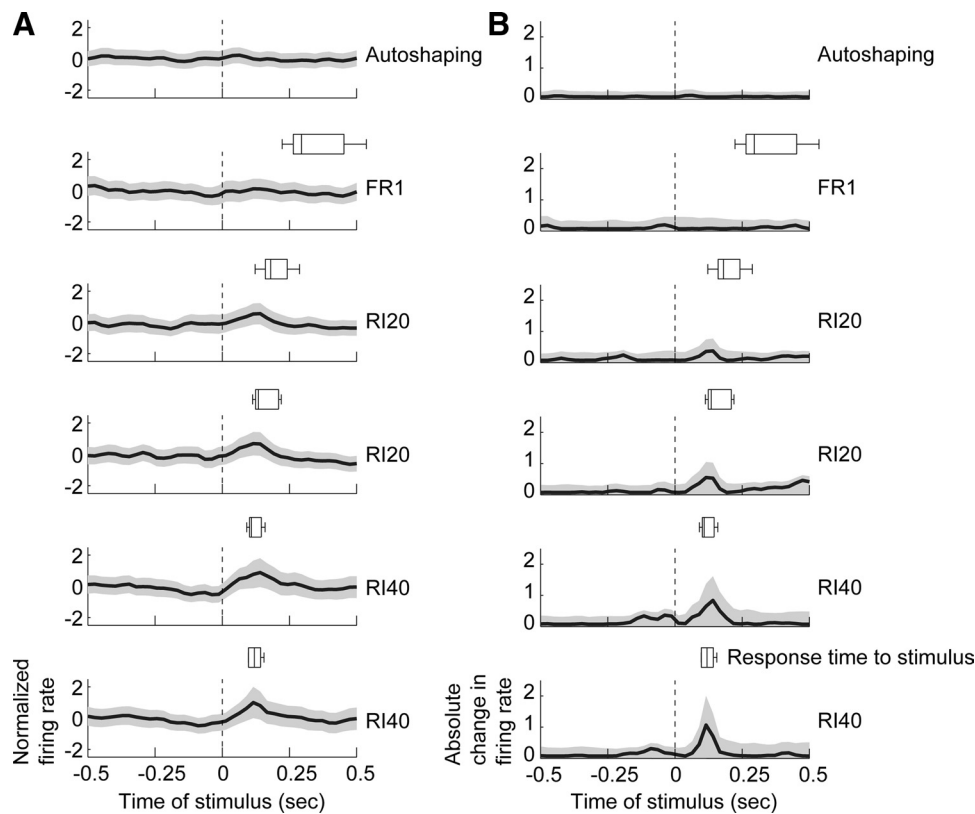


FIG. 6. Learning-related changes in population averages from the striatum. *A*: population averages of normalized firing rates are shown for all striatal neurons (lateral and medial) during the initial training sessions (VI60, FR1, RI20, RI40). The timing of withdrawals from the response port are shown above each population average as boxplots, except for the autoshaping session. Perievent histograms were generated by convolving the spike trains (± 0.5 s, 1-ms bins) with Gaussian windows (25 ms SD) and decimating by a factor of 25. *B*: absolute changes in population activity are shown for the neurons in *A*. Here, firing rates were normalized to have a mean of zero and were squared prior to averaging. This analysis was done to account for changes in firing rates independent of whether the change was an increase or decrease in firing rate. Together, these plots show that consistent modulation of population activity occurred exclusively during the withdrawal from the response port.

Many neurons fired just after the onset of the stimulus, when the rats moved from the response port to the reward port (Fig. 8, *C* and *D*). By plotting the activity of these neurons synchronized to the time when rats entered the reward port, it was clear that the neurons were modulated during this reward-collection behavior and not during consumption of sucrose (Fig. 9). Subsequent analysis of data from single neurons using a narrow window around stimulus onset established that neurons in the lateral, but not the medial, striatum fired in response to the stimulus [$F_{(1,1)} = 3.1$, $P < 0.03$] and that this difference in firing rate was sensitive to the extent of training [$F_{(1,13)} = 21.62$, $P < 0.001$; Fig. 8*C*]. Change-point analysis (Chow 1984) found that the proportion of neurons modulated around the stimulus in the lateral striatum was significantly elevated after 150 rewards were earned and that neurons in the medial striatum became significantly modulated around the stimulus only after >500 rewards were earned. An equivalent analysis was carried out for activity aligned to withdrawal from the response port and this showed that more neurons in the lateral striatum fired during withdrawal from the response port [$F_{(1,1)} = 30.6$, $P < 0.001$]; however, response-related firing was not altered over the period of training [$F_{(1,13)} = 1.51$, $P > 0.2$; Fig. 8*D*].

Response-related activity was influenced by the stimulus

To dissociate the changes in firing rates from the animals' movements between the response and reward ports, we

examined whether firing rates were sensitive to the presence of the acoustic stimulus. We selected a set of head entries, made in the absence of the stimulus, that were matched in duration to head entries in the presence of the stimulus. This analysis revealed that many neurons fired differently at the time of the stimulus compared with the latency-matched data (Fig. 9*A*). Likewise, the neurons fired differently during rewarded and nonrewarded entries into the reward port, despite the animals making similar movements in the presence and absence of the stimulus (Fig. 9*B*). These results suggest that neuronal activity in lateral striatum is not simply a correlate of movement. Rather, response-related firing is influenced by the stimulus that predicts reward availability.

Animals made very few slow responses (fewer than three responses per session with movement durations greater than >1 s). Therefore it was not possible for us to assess whether neuronal activity was modulated by the stimulus in the absence of movement. However, we did carry out recordings in all animals under anesthesia prior to perfusion. We presented a range of acoustic stimuli (noise bursts, tones, frequency sweeps) using standard methods and measured neuronal responses to the stimuli using peristimulus histograms. We did not find any cells in the lateral or medial regions of the striatum that were responsive to the acoustic stimuli (data not shown).

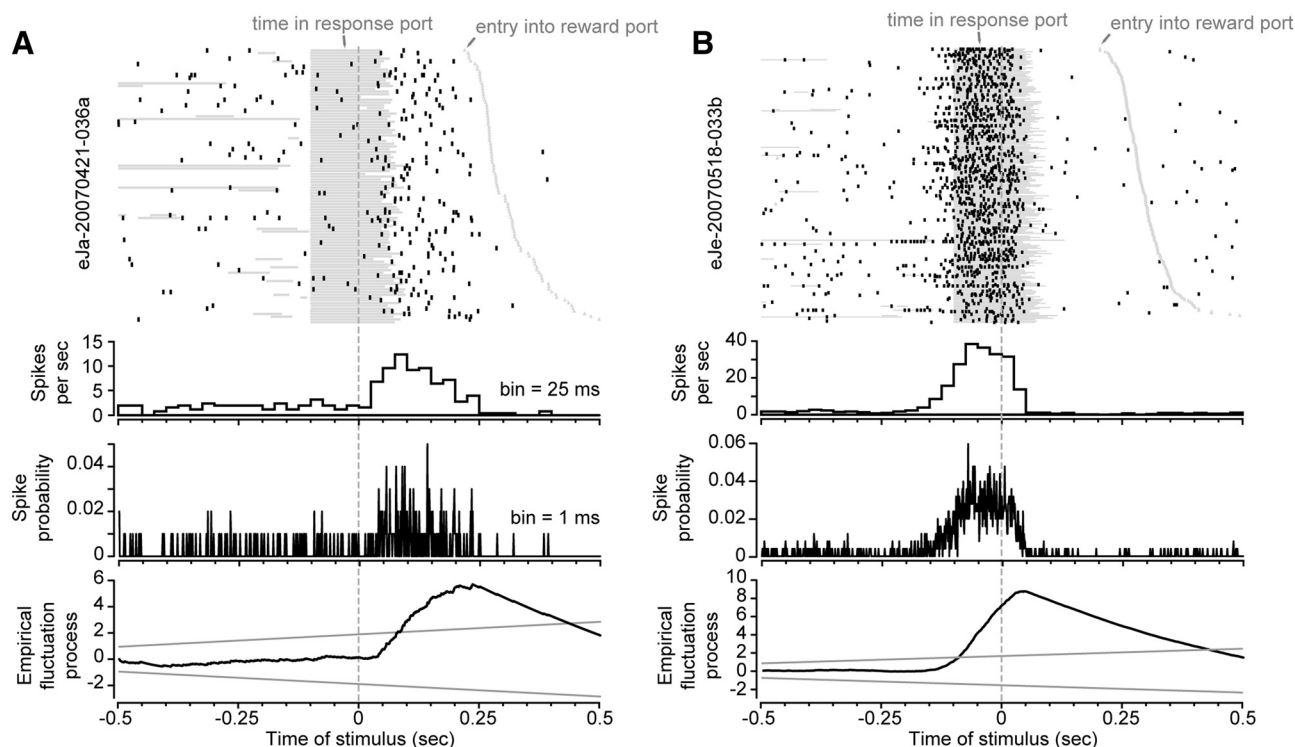


FIG. 7. Task-related modulations of firing rate. *A*: neuronal activity is shown for a neuron with significant modulation of firing rate around the reward-predictive stimulus and increased firing rate during movement from the response port to the reward port. *B*: neuronal activity is shown for another neuron with significant modulation around the reward-predictive stimulus. This neuron fired during entries into the response port and did not fire during movement from the response port to the reward port. In both panels, the *top plots* are raster plots and show the timing of spikes around the onset of the noise stimulus. The periods when the rats were in the response port are depicted as horizontal gray bars (some of which are from previous responses and occurred before the time of the stimulus). The timings of subsequent entries into the reward port are depicted as gray triangles (starting ~ 0.25 s after the stimulus). The 2nd and 3rd plots from the *top* show the average firing rate (measured in 25-ms bins) and the spike probability (measured in 1-ms bins), respectively. The *bottom plots* show the results of the empirical fluctuation analysis, with the statistic $y_n(t)$ (see Eq. 2) plotted on the ordinate axis and time around the stimulus plotted on the abscissa. The black lines depict deviations from the expected value of the firing rate and the gray lines depict the boundaries of the fluctuations that would be expected based on a random diffusion process (Brownian motion).

Changes in LFPs and spike-field coherence

LFPs recorded simultaneously with spike data were visibly modulated following the stimulus (Fig. 10). In both medial and lateral striatum, the stimulus evoked a complex LFP waveform characterized by an initial series of positive fluctuations (20–100 ms) followed by a wider negative fluctuation (100–200 ms) (Fig. 10, *A* and *B*). There was a sharpening of the initial positivity and deepening of the negativity over the course of training. The mean LFP range (maximum voltage–minimum voltage from 0 to 250 ms after the stimulus) increased with training [$F_{(1,412)} = 316.9$, $P \ll 0.001$, Fig. 10C], but did not differ by area [$F_{(1,412)} = 1.5$, $P = 0.21$]. There was no consistent change in LFP range for movements unaccompanied by the stimulus and that thus did not lead to a reward [$F_{(1,382)} = 0.7$, $P = 0.41$].

Fourier transforms of the LFPs revealed power in three main bands (Fig. 10D): a low-frequency band (< 5 Hz), a theta-frequency band (5–8 Hz), and a gamma-frequency band (30–45 Hz). We did not observe significant peaks in the power spectra in higher frequencies. Power in the theta frequency was higher in the lateral than that in the medial striatum [$F_{(1,412)} = 26.4$, $P \ll 0.001$] and varied significantly with training [$F_{(1,412)} = 30.0$, $P \ll 0.001$, Fig. 10E], increasing in the FR1 session early in training. Power in the gamma frequency was higher in the medial than that in the lateral striatum [$F_{(1,412)} =$

18.7, $P < 0.001$], but did not change significantly with training [$F_{(1,412)} = 0.0$, $P = 0.96$, data not shown].

Analysis of coherence between the spike and LFP activity showed that neurons fired spikes in phase with LFP oscillations that occurred at low frequencies (< 5 Hz). The proportion of significant spike-field pairs was greater in the lateral than that in the medial striatum (740/1,761 = 42% lateral vs. 697/2,430 = 29% medial; proportions test, $P \ll 0.001$). Over training, the proportion of significant spike-field pairs was greater in the lateral than that in the medial striatum [$F_{(1,132)} = 28.6$, $P \ll 0.001$] and increased with training [$F_{(1,132)} = 35.5$, $P \ll 0.001$, Fig. 10F]. Crucially, there was no consistent training-related change in spike-field coherence during exits from the response port in the absence of the stimulus [$F_{(1,122)} = 3.4$, $P = 0.07$].

Summary

In summary, we trained rats to perform a simple operant task. We found that four measures of instrumental behavior were altered during task acquisition: 1) the rates of responding (Fig. 3A), 2) the selectivity of movements to the spout (Fig. 3B), 3) the latency of withdrawal from the response port in response to the stimulus (Fig. 3C), and 4) the speed of movement from response port to reward port in response to the stimulus (Fig. 3D). Behavioral studies, done using identical

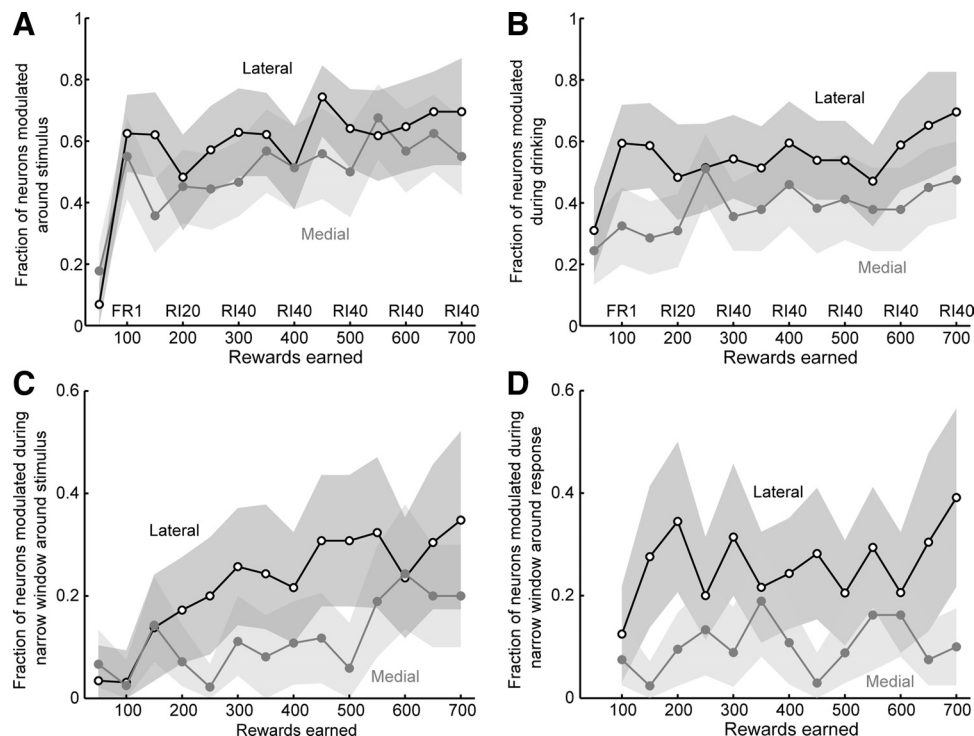


FIG. 8. Neuronal correlates of task acquisition. *A*: the fractions of neurons with task-related modulations of firing rate are plotted as a function of the number of rewards earned. Neurons in both the medial (gray) and lateral (black) striatum showed increases in task-related modulations over the period of learning and more neurons in the lateral striatum were modulated during learning. *B*: the fractions of neurons with modulations of firing rate during sucrose consumption are plotted as a function of the number of rewards earned. Neurons in both the medial (gray) and lateral (black) striatum showed learning-related increases in neuronal modulation and, as above, more neurons in the lateral striatum were modulated during learning. *C*: the fractions of neurons with modulations of firing rate during a narrow time window around the onset of the acoustic noise stimulus (-30 to $+20$ ms compared with $+20$ to $+70$ ms) are plotted as a function of the number of rewards earned. Significantly more lateral neurons fired in response to the acoustic noise stimulus compared with medial neurons and the fractions of stimulus-related activations in the lateral striatum increased with training. *D*: the fractions of neurons with modulations of firing rate during the withdrawal from the response port (± 50 ms) are plotted as a function of the number of rewards earned. Significantly more lateral neurons fired during this behavior compared with medial neurons and the fractions of response-related neurons did not increase over the period of training.

schedules of reinforcement, established that entry into the response port was habitual from early in training and that entry into the reward port remained goal-directed throughout training (Fig. 4). These changes were accompanied by a significant increase in the fraction of striatal neurons, especially in the lateral region, that fired after the stimulus during movements directed at the reward port (Fig. 8). Simultaneously recorded LFPs showed pronounced learning-related changes in the size of modulations following the stimuli and there were increased theta oscillations during learning (Fig. 10). Spiking activity was increasingly correlated with the LFPs at very low frequencies (<5 Hz). As before, these changes were greater in the lateral striatum than those in the medial striatum.

DISCUSSION

We demonstrate for the first time that during the initial learning of an instrumental task, neuronal activity changes progressively in the dorsal striatum, especially in the lateral region. Task-related firing occurred selectively, during reward-collection behavior, as animals moved from a response port to a reward port. This movement-related activity was modulated by the acoustic stimulus that signaled presentation of the sucrose reward. These changes in spike activity were accompanied by increasing modulations of striatal local field potentials (LFPs) during movement to the reward port and by increasing coherence between spikes and LFPs at low LFP

frequencies (<5 Hz). These results suggest that acquisition of the task resulted in a large-scale reorganization of neuronal processing in the dorsal striatum.

Implications for the role of the dorsal striatum in instrumental learning

The changes in striatal activity that we observed might reflect 1) the formation of stimulus–response (Daw et al. 2005) or stimulus–reward associations (Corbit and Janak 2007), 2) sensorimotor learning as reflected by the large reduction in movement latencies between the response and reward ports (Cohen and Nicolelis 2004; Costa et al. 2004; Laubach et al. 2000; Tang et al. 2007; Yin et al. 2009), or 3) changes in energy expenditures and the development of efficient movements during the task (Desmurget and Turner 2008). Further experiments are needed to examine these possibilities. An important new result from our study is the finding that movement latencies were dramatically reduced during learning, which suggests that instrumental learning may occur at the same time as sensorimotor learning, and might even involve some of the same neural structures (e.g., lateral striatum).

Our data reveal two time courses of changes in neural activity in the dorsal striatum. The first is an abrupt increase in gross neural modulation around the FR1 session, most evident on longer timescales within a trial. The second time course of change is progressive over training: transitions in neural activ-

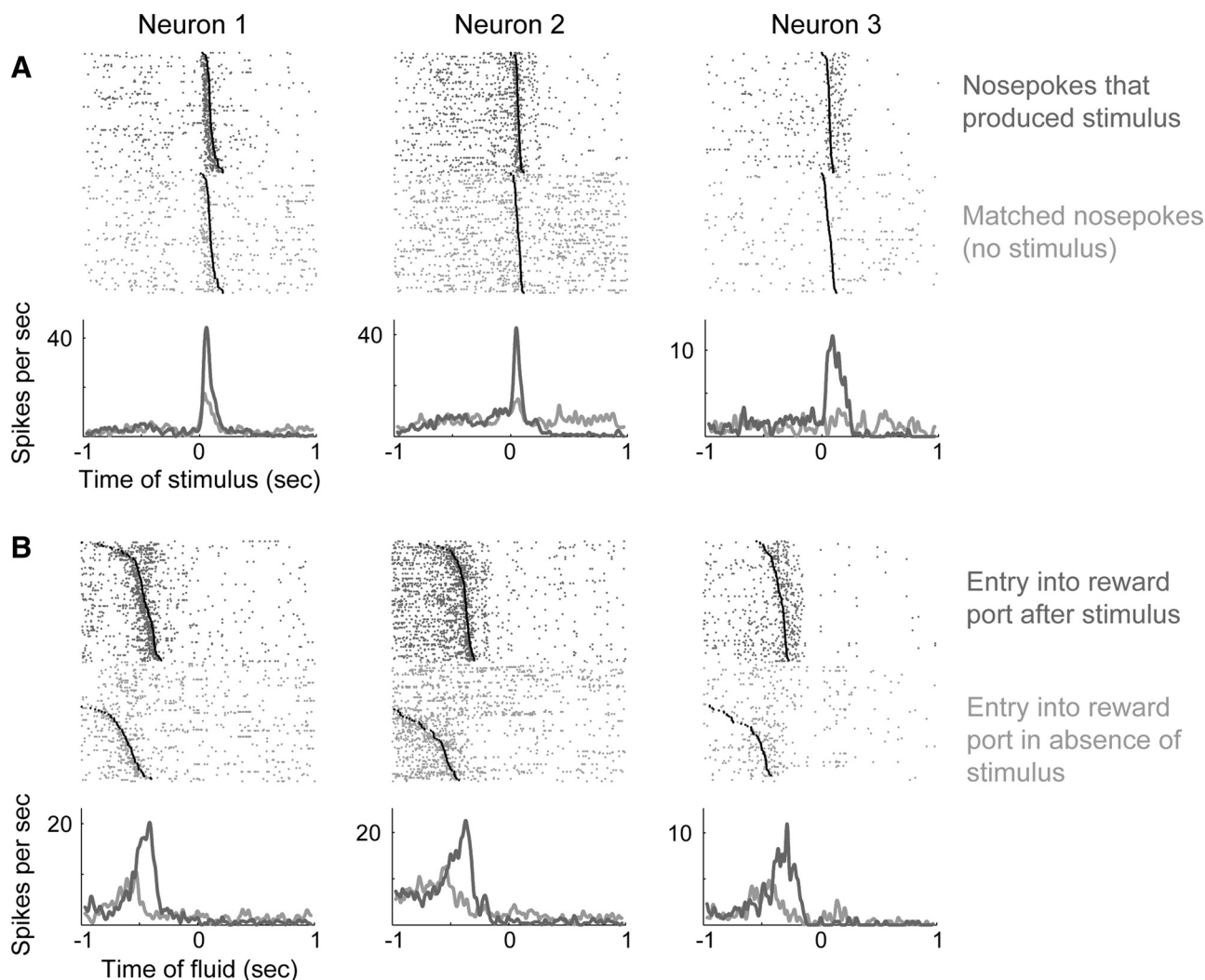


FIG. 9. Stimulus-dependent modulations of response-related activity in the dorsal striatum. *A*: lateral striatal neurons fired differently during entries into the response port that produced or did not produce the acoustic noise stimulus (rank-sum test, $P < 0.05$). Each column is a different neuron recorded in a different session. Neuronal activity is aligned to the stimulus (dark gray) or to the time when the stimulus would have occurred (absence of stimulus, light gray). Withdrawals from the response port are indicated in the raster plots by black circles. Perievent histograms were generated for display purposes by convolving the spike trains (1-ms bins) with Gaussian windows (20 ms SD). *B*: neuronal activity is plotted around the time of entry into the reward port for the same cells shown in *A*. The cells fired immediately after the stimulus and during movement to the reward port and were not further modulated during the acquisition or consumption of sucrose.

ity, both spikes and LFP, became more closely aligned with the stimulus and response initiation. These changes may be related both to changes in action sequences and to discriminative processing, as revealed behaviorally by matching changes in both movement speed and response selectivity. Thus neural activity in the dorsal striatum may be characterized by rapid behavioral engagement with poorly temporally organized neural activity followed by progressive selective refinement and temporal coordination.

Our results support a classic study by Carelli and West (1997), who found changes in lateral striatal neurons during the initial acquisition of a tone-controlled lever-press task, and more recent work by Tang et al. (2007), who showed learning-related changes in the lateral striatum during the acquisition of a conditioned head-movement task. Our results also support recent work on procedural learning in the T maze that highlighted changes in neuronal firing rates and theta-band oscillations in the striatum during initial task acquisition (DeCoteau

et al. 2007; Jog et al. 1999). However, our results do not support predictions that can be made from behavioral studies done by Yin et al. (2004, 2005). These studies predict that there should be pronounced changes in neuronal activity in the medial, but not the lateral, striatum early in learning.

Complicating this conclusion, we surprisingly did not find a training stage when nosepoke responses were sensitive to devaluation. In a strict formulation, instrumental learning theory stipulates that responses are learned through a transition from goal-directed to habit-driven behavior. We may have been unable to capture such a transition due to the precise way in which we tested devaluation sensitivity. That is, goal-directed behavior may have occurred even earlier in training or animals may have been making too few nosepoke responses early on to measure reliable decreases (Fig. 3*A*). In either case, modulations of neural activity were never significantly greater in the medial than those in the lateral striatum. However, if instrumental conditioning does not require a phase of goal-

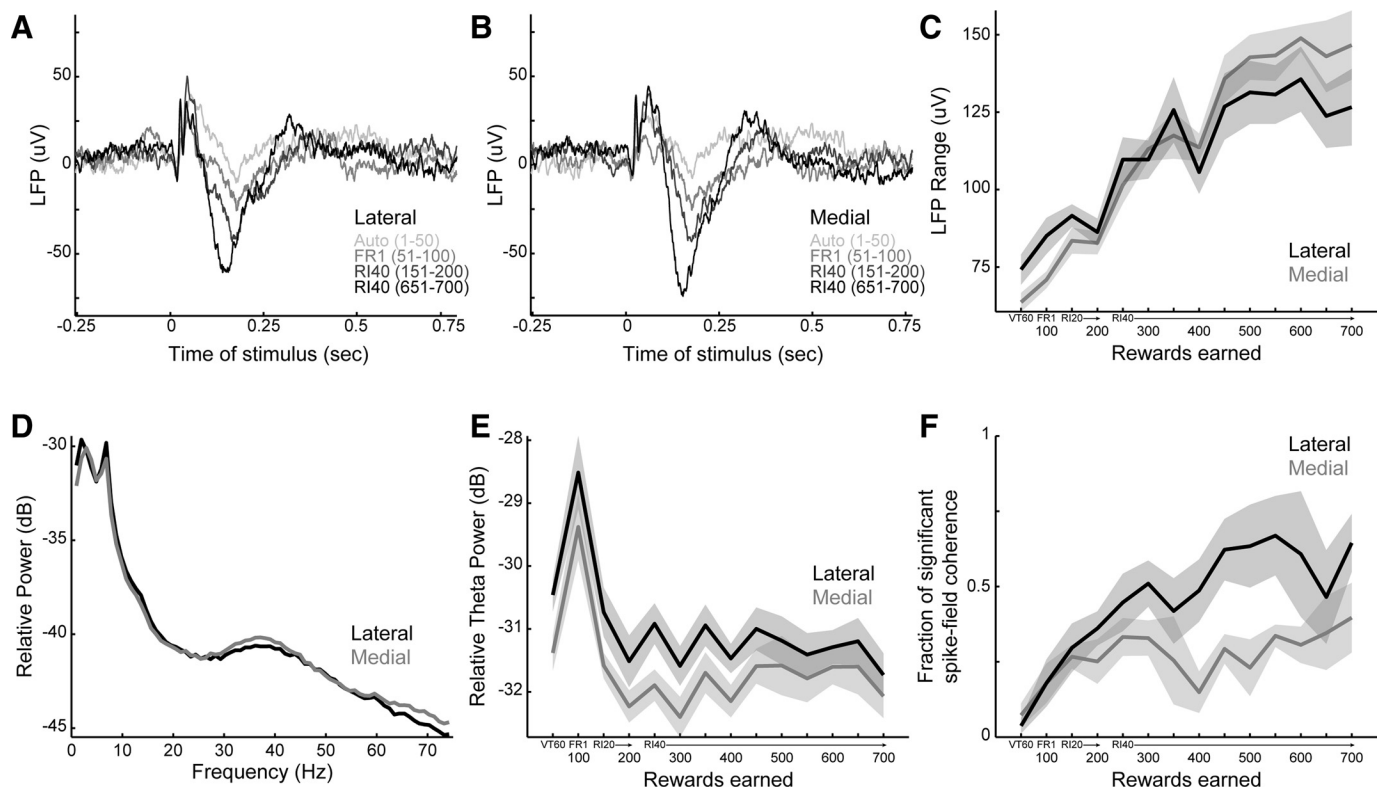


FIG. 10. Learning-related changes in local field potentials (LFPs). *A*: mean LFPs recorded from lateral electrodes at the time of the stimulus. Four blocks of 50 trials are depicted (autoshaping: lightest color, FR1, first block of RI40 training; last block of RI40 training: darkest color). Numbers in parentheses indicate the rewards accumulated in the block. The recorded potentials are characterized by an initial series of positive fluctuations followed by a wider negative fluctuation. Shaded areas depict the SE. *B*: mean local field potentials recorded from medial striatum at the time of the stimulus. Conventions are as in *A*. *C*: the range of LFP fluctuations (maximum–minimum voltage, 0–250 ms after the stimulus) increased with training, but did not differ by area. *D*: power spectra of the LFP signals following the stimulus (0–512 ms) are depicted across all blocks of response training. Power was noted in 3 bands: a low-frequency band (<5 Hz), a theta-frequency band (5–8 Hz), and a gamma-frequency band (30–45 Hz). *E*: power in the theta frequency was higher in medial striatum and varied significantly with training. *F*: significant spike/LFP coherence was found only at low frequencies (<5 Hz; data not shown). The proportion of significant spike/LFP pairs was greater in the lateral than that in medial striatum and increased with training.

directed behavior and our protocol did not sufficiently involve one, it is possible that the medial striatum would otherwise have been more engaged, consistent with specific behavioral predictions (Yin et al. 2004, 2005). We believe that further comparative studies within the striatum similar to those in this study and Yin et al. (2009) may help shed light on when medial striatum is engaged in early learning. Recent work by this group (Yin et al. 2009), using a rotorod task (as in Costa et al. 2004), reported that the medial striatum is the primary site of changes in neuronal activity early in sensorimotor learning, in a task in which it would be somewhat more difficult to devalue the outcome.

An important difference between our study and the behavioral work by Yin et al. (2004, 2005), however, is the time point when assessments are made about the roles of the lateral and medial regions of the striatum in learning. Yin's studies were based on extinction tests, done after the completion of training and after the use of devaluation methods. By contrast, our study involved recording neuronal activity during the actual acquisition of the instrumental task. Our approach is thus similar to that of Carelli and West (1997), Laubach et al. (2000), and Tang et al. (2007). By recording during learning, we were able to directly assess learning-related neuronal activity in the medial and lateral regions of the striatum. The fact that we did not find early changes in the medial striatum, as predicted in the work of Yin and colleagues, may be due to

differences between the tasks that were used (e.g., our task involved repetitive head entries into a response port) or to the possibility that responding in extinction after devaluation depends on different regions of the striatum (and frontal cortex) than those areas that are involved in actually learning the task. Although we made training procedures for physiological and behavioral experiments as identical as possible, there were slight differences between them. However, none of these differences was of major significance for comparisons between the experiments.

Although entries into the response port became habitual early in training (after just 150 earned rewards), reward-collection behavior (movement from the response port to the reward port) was goal-directed throughout the period of training. The effects of devaluation on reward collection behavior are not commonly reported, but similar dissociations have been observed by others (Nelson and Killcross 2006). This portion of behavior is actually more closely related to when we observed changes in neural activity in the lateral striatum. Therefore the changes we observed in the lateral striatum occurred in association with an action that remained goal-directed throughout training. Although not explicitly tested behaviorally by Yin and colleagues, the result is not expected based on the learning of habits and actions and suggests that the lateral striatum may have an important role in both of these forms of learning.

Differences between lateral and medial regions of the striatum

The differences we observed between the lateral and the medial regions of the striatum could arise for several reasons. There are fundamentally different temporal mechanisms of striatal plasticity in each area (Partridge et al. 2000) and there are also different sources of dopaminergic inputs to the lateral and medial regions of the striatum (Gerfen et al. 1987). The amygdala and orbital frontal cortex (OFC) have also been implicated in mediating the conditioned properties of stimuli (Pickens et al. 2003) and similar changes in learning-related neuronal activity have been found in these areas (Tye et al. 2008). However, the amygdala and OFC do not project to the lateral striatum. Instead, these areas innervate the medial and ventral regions of the striatum (Groenewegen and Trimble 2007). Through these connections, the ventral striatum may access a system of striatonigrostriatal pathways that influence dopaminergic transmission in the lateral striatum, as recently described in primates (Haber et al. 2000) and rodents (Ikemoto 2007). It is also possible that reward-related information mediated by the amygdala and OFC is sent to the lateral striatum from the most lateral portion of OFC, which has recently been shown to innervate the region of lateral striatum where we made our recordings (Schilman et al. 2008). Further study is needed to determine whether the learning-related changes in striatal activity we observed originate in the striatum itself or are derived from more broadly organized network activity, such as interactions with the prefrontal cortex (Houk and Wise 1995).

Parallel learning-related changes in spike activity and field potentials

Learning was accompanied by progressive increases in two parallel measures of neuronal activity in the dorsal striatum. Single neurons showed increases in task-modulated firing rates over the period of training. Simultaneously, movement-related potentials and transiently increased theta-band oscillations (5–8 Hz) developed in LFP recordings during the same period that was associated with increased firing by the striatal neurons, i.e., during reward-collection behavior. Analysis of spike-field coherence showed that the task-related spike activity occurred in phase with fluctuations of the LFP. Because these changes in LFP were observed across multiple electrodes, we suggest that learning induced coordinated changes in the activity of large numbers of striatal neurons (Darbin and Wichmann 2008) or coordinated changes in the activity of inputs to the striatum from brain areas such as the amygdala, which has been shown to develop altered patterns of LFP activity during learning (Bauer et al. 2007).

Selective activation of striatal neurons during reward-collection behavior

Our results suggest that representations of task-relevant stimuli and movements develop in the dorsal striatum during instrumental learning. The critical moment in the task was the time when rats initiate movements to the reward port to check for the availability of rewards. At this moment, most task-related neurons became activated and their firing rates depended on the presence of the reward-predictive stimulus. This

result is relevant for the interpretation of a recent study that examined how striatal neurons fired in relation to the current value of stimuli used in a go/no-go reaction time task (Kimchi and Laubach 2009). This study found that the same task epoch, during reward-collection behavior, was associated with major modulation of striatal neurons. The results reported here suggest that striatal activations during reward-collection behavior, which have been observed in many previous studies of the striatum, are not present in task-naïve animals and develop during the very earliest stages of instrumental learning.

ACKNOWLEDGMENTS

We thank J. Quinn for valuable intellectual contributions in the development of this research.

GRANTS

This research was supported by National Institutes of Health Medical Scientist Training Program Training Grant 5T32-GM-07205 to E. Y. Kimchi; DA-011717, RL-AA-017537, and DA-016556 to J. R. Taylor; and DA-022812 to M. M. Torregrossa; funds from the John B. Pierce Laboratory to M. Laubach; and a Tourette's Syndrome Association grant to J. R. Taylor.

REFERENCES

- 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.
- Bauer EP, Paz R, Paré D. Gamma oscillations coordinate amygdalo-rhinal interactions during learning. *J Neurosci* 27: 9369–9379, 2007.
- Berke JD. Uncoordinated firing rate changes of striatal fast-spiking interneurons during behavioral task performance. *J Neurosci* 28: 10075–10080, 2008.
- Brainard DH. The Psychophysics Toolbox. *Spat Vis* 10: 433–436, 1997.
- Brasted PJ, Wise SP. Comparison of learning-related neuronal activity in the dorsal premotor cortex and striatum. *Eur J Neurosci* 19: 721–740, 2004.
- Brown RL, Durbin J, Evans JM. Techniques for testing the constancy of regression relationships over time. *J R Stat Soc B* 37: 149–163, 1975.
- Carelli RM, Wolske M, West MO. Loss of lever press-related firing of rat striatal forelimb neurons after repeated sessions in a lever pressing task. *J Neurosci* 17: 1804–1814, 1997.
- Chow GC. Tests of equality between sets of coefficients in two linear regressions. *Econometrica* 28: 591–605, 1960.
- Chow GC. Random and changing coefficient models. In: *Handbook of Econometrics*, edited by Griliches Z, Intriligator MD. Amsterdam: Elsevier, 1984, p. 1213–1245.
- Cohen D, Nicolelis MA. Reduction of single-neuron firing uncertainty by cortical ensembles during motor skill learning. *J Neurosci* 24: 3574–3582, 2004.
- Corbit LH, Janak PH. Inactivation of the lateral but not medial dorsal striatum eliminates the excitatory impact of Pavlovian stimuli on instrumental responding. *J Neurosci* 27: 13977–13981, 2007.
- Costa RM, Cohen D, Nicolelis MA. Differential corticostriatal plasticity during fast and slow motor skill learning in mice. *Curr Biol* 14: 1124–1134, 2004.
- Darbin O, Wichmann T. Effects of striatal GABA_A-receptor blockade on striatal and cortical activity in monkeys. *J Neurophysiol* 99: 1294–1305, 2008.
- Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci* 8: 1704–1711, 2005.
- Dayan P, Balleine BW. Reward, motivation, and reinforcement learning. *Neuron* 36: 285–298, 2002.
- DeCoteau WE, Thorn C, Gibson DJ, Courtemanche R, Mitra P, Kubota Y, Graybiel AM. Learning-related coordination of striatal and hippocampal theta rhythms during acquisition of a procedural maze task. *Proc Natl Acad Sci USA* 104: 5644–5649, 2007.
- Desmurget M, Turner RS. Testing basal ganglia motor functions through reversible inactivations in the posterior internal globus pallidus. *J Neurophysiol* 99: 1057–1076, 2008.
- Dickinson A, Nicholas D, Adams CD. The effect of the instrumental training contingency on susceptibility to reinforcer devaluation. *Q J Exp Psychol B* 35: 35–51, 1983.

- Gerfen CR, Herkenham M, Thibault J. The neostriatal mosaic: II. Patch- and matrix-directed mesostriatal dopaminergic and non-dopaminergic systems. *J Neurosci* 7: 3915–3934, 1987.
- Groenewegen HJ, Trimble M. The ventral striatum as an interface between the limbic and motor systems. *CNS Spectr* 12: 887–892, 2007.
- Haber SN, Fudge JL, McFarland NR. Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *J Neurosci* 20: 2369–2382, 2000.
- Hansen BE. Approximate asymptotic p values for structural-change tests. *J Bus Econ Stat* 15: 60–67, 1997.
- Houk JC, Wise SP. Distributed modular architectures linking basal ganglia, cerebellum, and cerebral cortex: their role in planning and controlling action. *Cereb Cortex* 5: 95–110, 1995.
- Ikemoto S. Dopamine reward circuitry: two projection systems from the ventral midbrain to the nucleus accumbens-olfactory tubercle complex. *Brain Res Rev* 56: 27–78, 2007.
- Jog MS, Kubota Y, Connolly CI, Hillegaart V, Graybiel AM. Building neural representations of habits. *Science* 286: 1745–1749, 1999.
- Kimchi EY, Laubach M. Dynamic encoding of action selection by the medial striatum. *J Neurosci* 29: 3148–3159, 2009.
- Kimura M, Kato M, Shimazaki H. Physiological properties of projection neurons in the monkey striatum to the globus pallidus. *Exp Brain Res* 82: 672–676, 1990.
- Laubach M, Wessberg J, Nicolelis MA. Cortical ensemble activity increasingly predicts behaviour outcomes during learning of a motor task. *Nature* 405: 567–571, 2000.
- Nelson A, Killcross S. Amphetamine exposure enhances habit formation. *J Neurosci* 26: 3805–3812, 2006.
- Partridge JG, Tang KC, Lovinger DM. Regional and postnatal heterogeneity of activity-dependent long-term changes in synaptic efficacy in the dorsal striatum. *J Neurophysiol* 84: 1422–1429, 2000.
- Pasupathy A, Miller EK. Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature* 433: 873–876, 2005.
- Paxinos G, Watson C. *The Rat Brain in Stereotaxic Coordinates* (4th ed.). San Diego, CA: Academic Press, 1998.
- Pickens CL, Saddoris MP, Setlow B, Gallagher M, Holland PC, Schoenbaum G. Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. *J Neurosci* 23: 11078–11084, 2003.
- Ploberger W, Kramer W. The CUSUM test with OLS residuals. *Econometrica* 60: 271–285, 1992.
- Ploberger W, Kramer W, Kontrus K. A new test for structural stability in the linear regression model. *J Econometr* 40: 307–318, 1989.
- Rosenberg JR, Amjad AM, Breeze P, Brillinger DR, Halliday DM. The Fourier approach to the identification of functional coupling between neuronal spike trains. *Prog Biophys Mol Biol* 53: 1–31, 1989.
- Schilman EA, Uylings HB, Galis-de Graaf Y, Joel D, Groenewegen HJ. The orbital cortex in rats topographically projects to central parts of the caudate-putamen complex. *Neurosci Lett* 432: 40–45, 2008.
- Siegel S. *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill, 1956.
- Tang C, Pawlak AP, Prokopenko V, West MO. Changes in activity of the striatum during formation of a motor habit. *Eur J Neurosci* 25: 1212–1227, 2007.
- Tremblay L, Hollerman JR, Schultz W. Modifications of reward expectation-related neuronal activity during learning in primate striatum. *J Neurophysiol* 80: 964–977, 1998.
- Tye KM, Stuber GD, de Ridder B, Bonci A, Janak PH. Rapid strengthening of thalamo-amygdala synapses mediates cue-reward learning. *Nature* 453: 1253–1257, 2008.
- Williams ZM, Eskandar EN. Selective enhancement of associative learning by microstimulation of the anterior caudate. *Nat Neurosci* 9: 562–568, 2006.
- Yamada H, Matsumoto N, Kimura M. Tonically active neurons in the primate caudate nucleus and putamen differentially encode instructed motivational outcomes of action. *J Neurosci* 24: 3500–3510, 2004.
- 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, Mulcare SP, Hilario MR, Clouse E, Holloway T, Davis MI, Hansson AC, Lovinger DM, Costa RM. Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nat Neurosci* 12: 333–341, 2009.
- Yin HH, Ostlund SB, Knowlton BJ, Balleine BW. The role of the dorso-medial striatum in instrumental conditioning. *Eur J Neurosci* 22: 513–523, 2005.
- Zeileis A. p values and alternative boundaries for CUSUM tests (Working Paper 78). In: *Adaptive Information Systems and Modelling in Economics and Management Science*. Vienna: Vienna Univ. of Economics and Business Administration, 2000.
- Zeileis A, Leisch F, Hornik K, Kleiber C. *strucchange*: an R package for testing for structural change in linear regression models. *J Stat Software* 7: 1–38, 2002.