

Dynamic Encoding of Action Selection by the Medial Striatum

Eyal Yaacov Kimchi^{1,2} and Mark Laubach^{1,3}

¹The John B. Pierce Laboratory, New Haven, Connecticut 06519, and ²Interdepartmental Neuroscience Program and ³Department of Neurobiology, Yale University School of Medicine, New Haven, Connecticut 06520

Successful foragers respond flexibly to environmental stimuli. Behavioral flexibility depends on a number of brain areas that send convergent projections to the medial striatum, such as the medial prefrontal cortex, orbital frontal cortex, and amygdala. Here, we tested the hypothesis that neurons in the medial striatum are involved in flexible action selection, by representing changes in stimulus–reward contingencies. Using a novel Go/No-go reaction-time task, we changed the reward value of individual stimuli within single experimental sessions. We simultaneously recorded neuronal activity in the medial and ventral parts of the striatum of rats. The rats modified their actions in the task after the changes in stimulus–reward contingencies. This was preceded by dynamic modulations of spike activity in the medial, but not the ventral, striatum. Our results suggest that the medial striatum biases animals to collect rewards to potentially valuable stimuli and can rapidly influence flexible behavior.

Introduction

An important characteristic of adaptive behavior is the ability to change one's responses to the same stimulus over time. Flexible responding depends on an integration of stimulus, response, and reward information in context-dependent associations. The neuronal basis of this learning involves the frontal cortex, amygdala, hippocampus, and basal ganglia (Wise and Murray, 2000) and in particular the medial striatum of the basal ganglia (Balleine et al., 2007). This part of the striatum is unique in that it receives projections from the medial prefrontal cortex (mPFC) and orbital frontal cortex [OFC; McGeorge and Faull (1989)] and the basal nuclei of the amygdala (McDonald, 1991). These brain areas have previously been shown to be crucial for flexible behavior [mPFC: Ragozzino et al. (1999a,b) and Ferry et al. (2000); OFC: Schoenbaum et al. (2002); amygdala: Stalnaker et al. (2007)].

Flexible association has been studied physiologically in the primate basal ganglia, such as in the caudate nucleus (Apicella et al., 1991; Tremblay et al., 1998; Clarke et al., 2008). The rat medial striatum may be homologous to the primate caudate (Voorn et al., 2004). Lesions or chemical infusions into the medial striatum are known to alter response strategies in mazes (Ragozzino et al., 2002a,b; Palencia and Ragozzino, 2004), discrimination performance (Adams et al., 2001; Featherstone and McDonald, 2005),

and action–outcome learning (Yin et al., 2005). Medial striatal neurons are modulated by stimuli and responses (White and Rebec, 1993; Teagarden and Rebec, 2007). However, there have been no reports of neuronal activity in the rat medial striatum during flexible behavior. Neuronal activity in the primate striatum can change rapidly after changes in stimulus–response pairings (Brasted and Wise, 2004; Pasupathy and Miller, 2005; Williams and Eskandar, 2006). Information relevant to action selection such as choices and rewards can be encoded by the same (Ding and Hikosaka, 2006) or distinct (Lau and Glimcher, 2007) groups of medial striatal neurons.

Previous neurophysiological studies in behaving rats have focused on the ventral striatum, which includes the nucleus accumbens (Peoples et al., 1999; Carelli and Ijames, 2000; Setlow et al., 2003; Janak et al., 2004). Neurons in the ventral striatum are modulated during behavioral sequences that lead to reward collection (Apicella et al., 1991; Carelli and Deadwyler, 1994; Chang et al., 1994; Taha et al., 2007) and are sensitive to the specific outcome of an operant response (Carelli and Deadwyler, 1994; Nicola et al., 2004). Additionally, neurons in the monkey ventral striatum respond to changes in stimulus–reward pairings similarly to neurons in the medial striatum (Tremblay et al., 1998). Surprisingly, however, some behavioral studies have demonstrated that the ventral striatum may be less important than the medial striatum for responses to external stimuli and flexible choice behavior (Amalric and Koob, 1987; Carli et al., 1989; Schoenbaum and Setlow, 2003).

No study has compared medial and ventral striatal neurons during a flexible response task. To study this issue, we developed a novel Go/No-go reaction-time task in which the reward contingency for a given stimulus changed within single experimental sessions. Using simultaneous, multisite recording methods, we found that neurons in the medial, but not the ventral, striatum were modulated immediately after changes in stimulus–reward

Received Oct. 20, 2008; revised Jan. 25, 2009; accepted Jan. 26, 2009.

This work was supported by the National Institutes of Health Medical Scientist Training Program Training Grant 5T32GM07205 (E.Y.K.) and funds from the John B. Pierce Laboratory (M.L.). E.Y.K. and M.L. conceived and designed the experiments, analyzed the data, and wrote this manuscript. E.Y.K. performed the experiments. We thank the Instruments Shop at the John B. Pierce Laboratory for their technical support. We thank Ivan de Araujo, Nicole Horst, Don Katz, Christopher Pittenger, Brian Lau, Daeyeol Lee, Marshall Shuler, Xiao-Jing Wang, and two anonymous reviewers for comments on this manuscript.

The authors declare no competing financial interests.

Correspondence should be addressed to Dr. Mark Laubach, The John B. Pierce Laboratory, 290 Congress Avenue, New Haven, CT 06519. E-mail: mlaubach@jbpierce.org.

DOI:10.1523/JNEUROSCI.5206-08.2009

Copyright © 2009 Society for Neuroscience 0270-6474/09/293148-12\$15.00/0

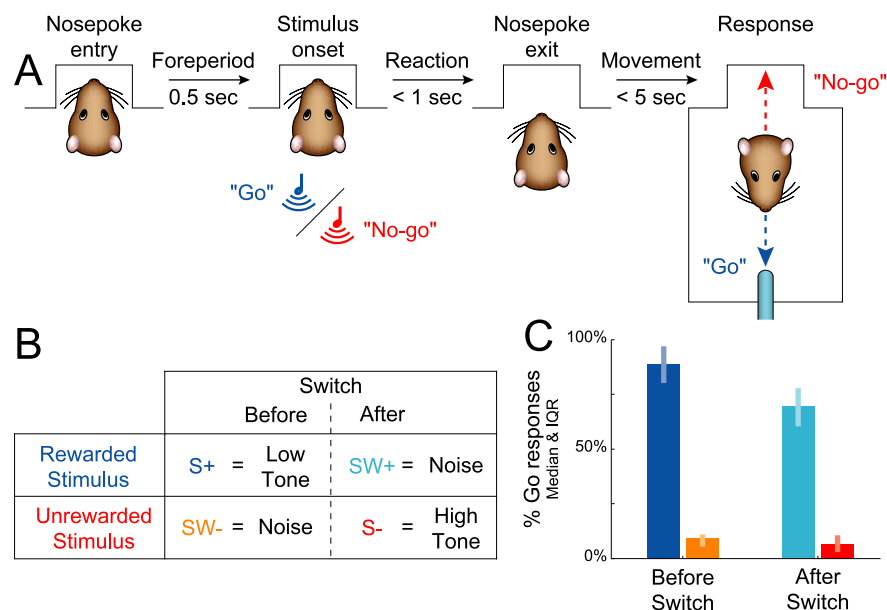


Figure 1. Go/No-go discrimination task. **A**, Rats initiated trials by poking their snouts into a hole (nosepoke entry) on one side of the chamber. They had to maintain this position for 0.5 s until a stimulus was delivered (onset of a tone, noise, or light), withdraw from the hole within 1 s of stimulus onset (nosepoke exit), and then decide whether to cross the chamber to collect a reward. A rapid withdrawal followed by an attempt to collect a reward was termed a "Go response" and resulted in water delivery only on trials with rewarded stimuli (+). If an animal attempted to make a Go response to unrewarded stimuli (−), no water was delivered and the trial repeated (i.e., an unrewarded stimulus was presented on the next trial). **B**, In the first half of a switch session, the rewarded stimulus was a low-frequency (8 kHz) tone that was presented on half of trials (S+). The unrewarded stimulus was called the "switching stimulus" (SW−; noise, shown here, or light). Once rats achieved a criterion level of performance (90% correct over 20 trials), the reward contingencies were changed for the switching stimulus. Responses to it were now rewarded (SW+), and a high-frequency (30 kHz) tone, which was never rewarded, was presented on half of the trials (S−). **C**, Rats were able to adapt their behavior successfully after a switch in reward contingencies. Before the switch, rats made Go responses primarily to S+ (low tone, blue) and minimally to SW− (e.g., noise, orange). Following the switch, rats made Go responses primarily to the now rewarded switching stimulus (SW+, e.g., noise, light blue) and not to the S− (high tone, red). Data are presented from all of the recording sessions (14 rats, 2 recording sessions per rat).

contingencies, and before there were changes in selective behavioral responding in the task.

Materials and Methods

Subjects. Seventeen male Long–Evans rats (Harlan) were individually housed and kept on a 12 h light/dark cycle with lights on at 7:00 A.M. All procedures were approved by the Institutional Animal Care and Use Committee at the John B. Pierce Laboratory. To motivate behavior, rats had restricted access to water for 18 h before behavioral sessions. They were given *ad libitum* access to water at least 1 d every week. Food was always available *ad libitum*. Rats earned about half of their daily water during the behavioral sessions, and the remaining volume of the animal's average water consumption was provided in the home cage, 1–2 h after the behavioral sessions. Rats maintained 90–95% of their free water body weight throughout the training and testing period. Electrodes were implanted in all rats (see below). However, one rat did not survive the surgery and the implanted electrodes in two rats yielded poor recordings. All neuronal analyses are therefore based on the remaining 14 rats.

Operant chamber. Initial training took place in a standard operant chamber (ENV-008, Med Associates). One wall had a central nosepoke hole with an infrared beam to identify nosepoke entry (ENV-114BM, Med Associates). The opposite wall had a metal spout connected to a water pump and a custom-built electronic circuit, which detected contact by measuring resistance between the spout and the floor. Above the spout was a speaker (ES1, TDT Technologies) for presentation of acoustic stimuli (RP2.1 Processor, TDT Technologies). Tones and noise (white noise generated via the TDT DSP System) were presented from the speaker above the waterspout at 65 dB. The speaker and the nosepoke each had a light above them (ENV-221M, Med Associates). A light stimulus consisted of turning on both lights simultaneously. Except for when

lights were used as a brief stimulus, behavior occurred in the dark. Protocols in this chamber were controlled using Med Associates software (MedPC).

The recording chamber was similar to the standard operant box but modified for electrical recordings (Med Associates). All walls and floor bars were made of acrylic plastic; the long walls sloped diagonally outward as they went up. A copper shielded electrostatic speaker (ES1, TDT Technologies) was used to present the tones and noise. Light stimuli were made from LEDs covered with sand-blasted plastic discs that were 1 inch in diameter (Pierce Laboratory Instruments Shop). The waterspout was made of glass and licking was registered by break of an infrared beam in front of the spout (Pierce Laboratory Instruments Shop). A 32-channel commutator (Plexon) was mounted in the center of the ceiling of the chamber. This device connected cables from the implanted probes to the recording system. The chamber was placed on a steel plate within a sound attenuation chamber lined with sound foam, itself surrounded by a Faraday cage. Behavioral protocols in the recording chamber were controlled using a digital input/output card (PCI-DIO-96, National Instruments) run by the Matlab Data Acquisition Toolbox (Mathworks) and the freely available Psychophysics Toolbox (Brainard, 1997).

Initial training. After handling and fluid restriction, rats were placed in the operant chamber with the nosepoke hole closed. They underwent 1 week of simple tone conditioning. Sessions lasted for 1 h each day. On the first day, an 8 kHz tone (6 s, 65 dB) was presented approximately every 30 s, and 120 μ l of water was delivered at the spout (S+). Starting on the second day, water delivery occurred only if the rats licked the spout during the S+. On one-third of trials, no tone was presented and we measured the probability that rats would lick at the metal spout in the absence of the conditioned stimulus. Once rats licked consistently more on tone trials than silent probe trials over at least two sessions, they were advanced to nosepoke training.

Nosepoke training. As shown in Figure 1, rats initiated presentations of stimuli by inserting their snout into the nosepoke hole, breaking the infrared beam inside it. Rats had to maintain this nosepoke position for a set time (foreperiod) to receive a stimulus. The foreperiod was increased from 25 to 500 ms within sessions across the first 3 d of nosepoke training. If rats withdrew from the nosepoke early, no stimuli were presented and responses were not rewarded. No time-out or other form of negative reinforcement was used to reduce the occurrence of premature responses and these trials were not included in behavioral or physiological analyses. Once a stimulus was presented, rats had to withdraw from the nosepoke within a set time [reaction time (RT); limit shortened to 1 s] and cross to the other side of the chamber (30 cm) to make contact at the spout within a set time (movement time; limit shortened to 5 s) to collect water (80 μ l). We use the term "Go response" to describe nosepoke withdrawals within 1 s of stimulus onset that were followed by contact with the waterspout within 5 s. Stimuli remained on until a Go response at the waterspout, a new nosepoke, or the time window for a response had elapsed. Trials were self-paced, initiated by the rats' nosepokes, and sessions lasted ~90 min.

Discrimination training. After initial nosepoke training, rats were advanced to a simple Go/No-go discrimination task. On any trial, during the nosepoke, the rat could receive one of two types of stimuli. Only one stimulus was presented on each trial. If the rat received an S+ (8 kHz tone), it could make a Go response to collect a water reward. If the animal

made a No-go response to S+, there was no punishment or indication that this was incorrect (although there is an implicit missed opportunity to collect a reward). If a rat received an S− (30 kHz tone), Go responses were incorrect and did not lead to rewards. Instead, a Go response forced a repetition of S− on the next trial. Once rats were trained, they typically reinitiated a new trial with a nosepoke after receiving S−. New stimuli were pseudorandomly selected with the constraint that no stimulus type could be selected more than three times in a row.

Stimulus blocks. Rats were trained on blocks of familiar discrimination pairs for ~4 months (~75 sessions per rat). In any pair, one stimulus was rewarded and one stimulus was unrewarded. Throughout training, the reward status of two stimuli never changed (S+: 8 kHz tone was always rewarded, S−: 30 kHz tone was never rewarded). The reward statuses of the other stimuli (SW) were switched within sessions (SW+ when rewarded, SW− when unrewarded). In the sessions presented in this study, a noise burst was primarily used as the flexible, switching stimulus. Over training, rats also received practice with another switching stimulus (lights) as well as sessions that did not involve switching (i.e., stimuli–reward contingencies did not change within the session). Data with lights as the SW stimulus was qualitatively similar to that obtained with noise as the SW stimulus.

In each “switch session,” rats made discriminations between S+ (always rewarded 8 kHz tone) and SW− (unrewarded switching noise) for the first half of the session. Then, the stimulus–reward contingencies were changed. The switching stimulus was now SW+ (same noise, now rewarded) and S− was used as the unrewarded stimulus (never rewarded 30 kHz tone). Switches in reward contingencies were not signaled and occurred after 50 correct responses after achieving criterion on the initial discrimination. The criterion for accurate discrimination was 18 correct trials within a window of 20 stimuli (90% accuracy).

Electrode implantation. Before surgery, a rat was initially anesthetized with 4% vaporized halothane. The rat was then injected intraperitoneally with ketamine (100 mg/kg) and xylazine (10 mg/kg). Smaller supplementary doses of ketamine were given as needed to suppress responses to toe pinches. The scalp was shaved and the rat was placed in a stereotaxic apparatus using blunt 45° ear bars to prevent eardrum rupture. Eyes were covered with ophthalmic antibiotic ointment to prevent desiccation. The scalp was disinfected with iodine, incised, and retracted to expose the skull surface. Lambda and bregma were leveled and bilateral craniotomies were made to implant multielectrode arrays.

Multielectrode arrays were composed of 50 μ m stainless steel wires, coated with Teflon, and spaced in 2×8 , 4×4 , or $3 \times 3 \times 2$ configurations with 250 μ m spacing between wires (Neuroline). The wires were connected to Omnetics connectors (0.025 or 0.050). All rats but one were implanted with 16 wire arrays in each hemisphere; the remaining rat received a 16-wire array in one hemisphere and an 8-wire array in the other hemisphere ($3 \times 3 \times 2$ configuration). Neuronal activity was monitored as electrodes were lowered into the medial striatum. Arrays were placed dorsally (centered at 0 mm AP, 2.2 mm ML, −4 mm DV) or ventrally (centered at 0.4 mm AP, 1.8 mm ML, −6 mm DV): four rats received bilateral medial implants, two rats received bilateral ventral implants, and eight rats received one medial and one ventral implant in opposite hemispheres.

After electrode implantation, the craniotomy was covered with cyanoacrylate (Slo-Zap) and cyanoacrylate accelerator (Zip-Kicker). The scalp

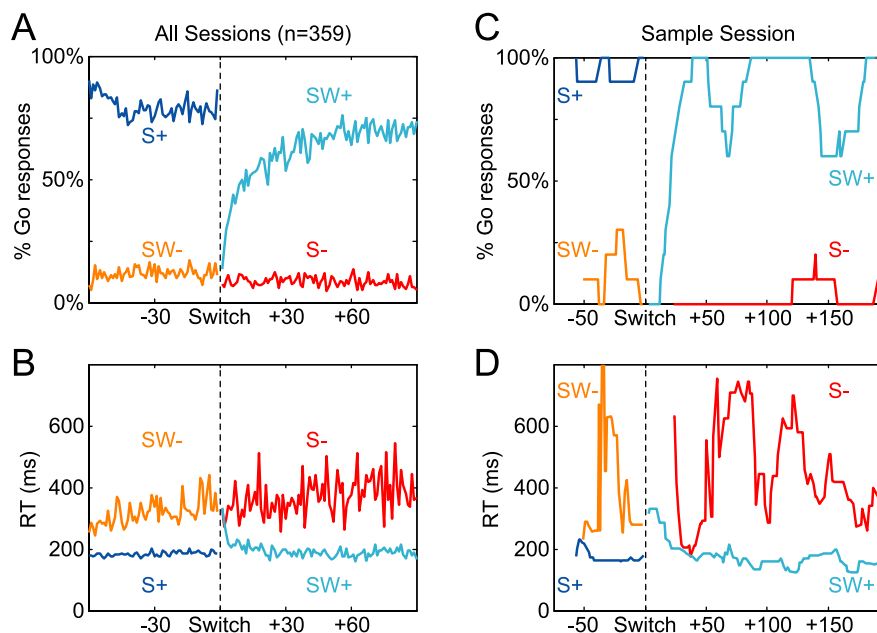


Figure 2. Effects of stimulus value on behavioral performance. **A**, The proportions of trials with Go responses are shown for all behavioral sessions (359 training and recording sessions from 17 rats). Rats were initially more likely to make Go responses to the rewarded stimulus (S+, the low-frequency tone, blue) than the switching stimulus (SW−, orange). After the change stimulus–reward contingencies (vertical dashed line), animals made Go responses to the switching stimulus (SW+, light blue) and not to the unrewarded stimulus (S−, high-frequency tone, red). **B**, Group data for the animals’ reaction times are shown. Specifically, the median RT for each trial relative to the switch in action selection is plotted for all behavioral sessions. Rats initiated responses more quickly to rewarded (S+, SW+) than unrewarded stimuli (SW−, S−). **C**, The proportion of Go responses is shown for an example session (same colors as in **A**). The proportion of Go responses is plotted as a 10-trial running mean where a Go response is scored as 1, and a No-go response is scored as 0. After the switch in reward contingencies, the rat began to respond consistently to SW+. **D**, Corresponding reaction time data for the same example session (as a 10-trial running median). The rat initiated responses more quickly to rewarded than unrewarded stimuli.

was covered with methyl methacrylate (AM Systems). Wound margins were daubed with antibiotic ointment (Vetropolylin). Rats received one subcutaneous injection of buprenorphine (0.03 mg/kg) and oral enrofloxacin (34 mg/L) in their drinking water for 1 week. After recovery from surgery (one week), animals were fluid restricted to reinitiate behavioral sessions with neurophysiological recordings.

Recordings. Rats were briefly anesthetized with 2% halothane to connect and tape headstage cables to the implanted connectors. Rats recovered spontaneous motor activity within a few minutes and were then placed in a separate acrylic chamber for at least half an hour before behavioral recording. During this time, spike sorting was performed and spontaneous recordings of neuronal activity were made, including recordings of wideband signals from each electrode (sampling at 20 kHz; analog filtering from 0.5 Hz to 5.9 kHz).

Signals from the implanted electrodes were fed into the recording system (Plexon Multichannel Acquisition Processor) and amplified 1000–20,000 times. Spike activity was isolated using a voltage threshold. Waveforms that crossed the threshold were sampled, timestamped, and stored at 40 kHz. Unique waveforms were identified online. Online root-mean-square values, while rats were resting, were typically 20 μ V (calculated within the Plexon OnLine Sorter). Waveforms were then processed off-line (Plexon OffLine 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 at least 100 μ V. (2) Signal-to-noise ratio had to be at least 3:1. (3) Fewer than 2% of interspike intervals could be <2 ms. (4) The mode of the interspike interval histogram had to be >5 ms. (5) Baseline firing rates had to be <30 Hz. (6) The distribution of maximal waveform points had to be relatively normal (skewness <0.75). This latter measure ensured that waveforms from putative single-units were effectively isolated from the noise threshold.

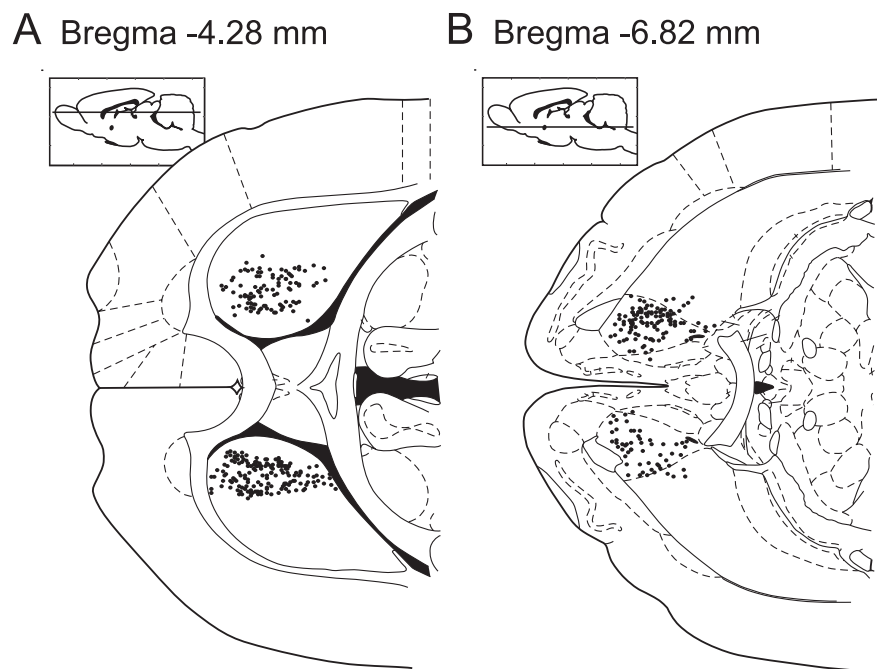


Figure 3. Locations of recordings. The locations of the recording sites are depicted using horizontal sections. **A**, 341 neurons were localized to the medial striatum (black dots). **B**, 211 neurons were localized to the ventral striatum, including the core of the nucleus accumbens. There were no significant differences in waveform sizes or overall firing rates between the medial and ventral neurons (rank-sum test, $p > 0.05$). Atlas figures are adapted (Paxinos and Watson, 1998).

We limited our analysis to 30 correct trials before and after the change in stimulus–reward contingencies. Inspection of behavioral records indicated that the animals performed the task in a consistent manner during this period. This also allowed us to compare activity around changes in stimulus–reward contingencies with activity in other, equal durations of the session. Furthermore, we required that a neuron’s spike waveforms and interspike interval statistics were consistent throughout the behavioral session. We also evaluated neuronal activity during the drinking period, a time epoch that was unrelated to our analysis of action selection, to ensure that firing rates were generally stationary over trials. During the drinking period, neurons had to fire at least once on 10% of trials and have a z-score of <4 on a runs test (Siegel, 1956).

Histology. At the conclusion of the recording sessions, rats were killed with an intraperitoneal dose of pentobarbital (>100 mg/kg). Rats were then perfused with chilled saline followed by 4% formaldehyde. Brains were extracted and placed into solutions of 25% sucrose for cryoprotection. Brains were cut on a frozen sliding microtome, stained with thionin, dehydrated in an increasing concentration of ethanol followed by xylenes, and mounted and coverslipped on gelatin subbed slides. Electrode holes were identified using light microscopy and plotted onto a rat brain atlas (Paxinos and Watson, 1998). Three-dimensional models of the rat striatum were constructed using freely available software written for Matlab by E. Y. Kimchi: <http://spikelab.jbpierce.org/3DAnatomy>.

Identifying neuronal modulations. Data analysis was done using Matlab (Mathworks) and R (<http://www.R-project.org>). Timestamps from identified single units were aligned to the time of the stimulus to create perievent rasters and perievent histograms (bin size: 1 ms, time-window: -1 to $+1$ s around the stimulus). We initially evaluated neuronal modulations at the time of the trigger stimulus by comparing firing rates in two windows surrounding the stimulus (100 ms before and after stimulus onset). Activity in these two windows was compared using the sign-rank test ($p < 0.05$). For initial analyses, we only analyzed data from correctly performed trials with reaction times >100 ms.

In addition, we evaluated neuronal modulations over the entire pre-stimulus epoch using a structural change test (Chow, 1960). The test was done as follows: (1) a cumulative sum histogram was calculated for data series (the average normalized firing rate), (2) a linear model was fit to full data window, (3) a series of linear models were fit to smaller data

windows, 0.5 s in duration, that were moved over the data series in 1 ms increments, (4) the F statistic (Chow, 1960) was calculated as the sum of squared residuals between the coefficients for a given data window and the coefficients for the full data window (see supplemental materials for details, available at www.jneurosci.org), and (5) the F statistic was evaluated using standard values for the F distribution (Hansen, 1997) with a criterion of $p < 0.05$. This analysis was done using the strucchange library for R (Zeileis et al., 2002).

Decoding neuronal activity related to components of action selection. Decoding methods were used to compare neuronal activity with various components of action selection. Specifically, we assessed whether neuronal activity contained information about changes in stimulus–reward contingencies, reward collection behavior, or variability in reaction times. Decoding methods were well suited for this study, compared with other methods such as ANOVA or ROC analysis, because decoding methods provide measures of the confidence of classification (i.e., posterior probabilities) on a trial-by-trial basis.

Sets of trials surrounding the switch in stimulus–reward contingencies were selected for decoding analysis. Neuronal activity on the last 30 SW– trials was compared with that on the first 30 SW+ trials after presentation of S– (unrewarded tone stimulus). The goal was to quantify when striatal neurons reflected this change in stimulus–reward contingencies. Such sensitivity determines what the appropriate behavioral response should be, regardless of what it actually will be. The stimulus–reward contingency for each trial was classified using firing rate, measured during a 0.6 s epoch starting at onset of the stimulus. Firing rates for two types of trials (e.g., SW– vs SW+) were compared using a probabilistic classifier (which is known as naive Bayes or the MAP classifier, where MAP stands for maximum a posteriori) (for review, see John and Langley, 1995; Domingos and Pazzani, 2004). We used the implementation of naive Bayes in the *e1071* library for R. Additional analysis was done using alternative measures of neuronal activity, e.g., temporal changes in firing rate measured using wavelet methods (Laubach, 2004), and using other types of classifiers, including an unsupervised method for cluster analysis. The results reported here were consistent across methods, and so, for clarity, we report the simplest methods in the main body of the paper.

The naive Bayes classifier was based on nonparametric kernel density estimation. Training data were used to estimate the relative density of firing rates for each type of trial (e.g., firing rates on trials with SW– and SW+). Standard methods (the density function from the stats library for R) and Gaussian smoothing kernels were used for kernel density estimation. Testing data were evaluated by measuring the most likely trial type based on the neuron’s firing rate and the density estimates for firing rate obtained from the training data. For each trial, we estimated the posterior probability that the trial occurred before or after the switch in stimulus–reward contingencies. Leave-one-out cross-validation was used. Based on simulations of random data, the minimum number of trials needed to obtain reliable estimates of successful classification (see the next paragraph) depended on having at least 30 trials of each type.

To control for nonspecific effects over the experimental sessions (Prokopenko et al., 2004), we compared neuronal activity from the first 30 SW+ trials to activity from the next 30 SW– trials. This analysis was done for activity on trials with the same stimulus and reward contingencies. Since stimulus–reward contingencies were unchanged, any decoding would be due to nonspecific changes in neuronal firing, such as changes in motivation.

Due to issues of statistical power, we did not attempt to analyze differ-

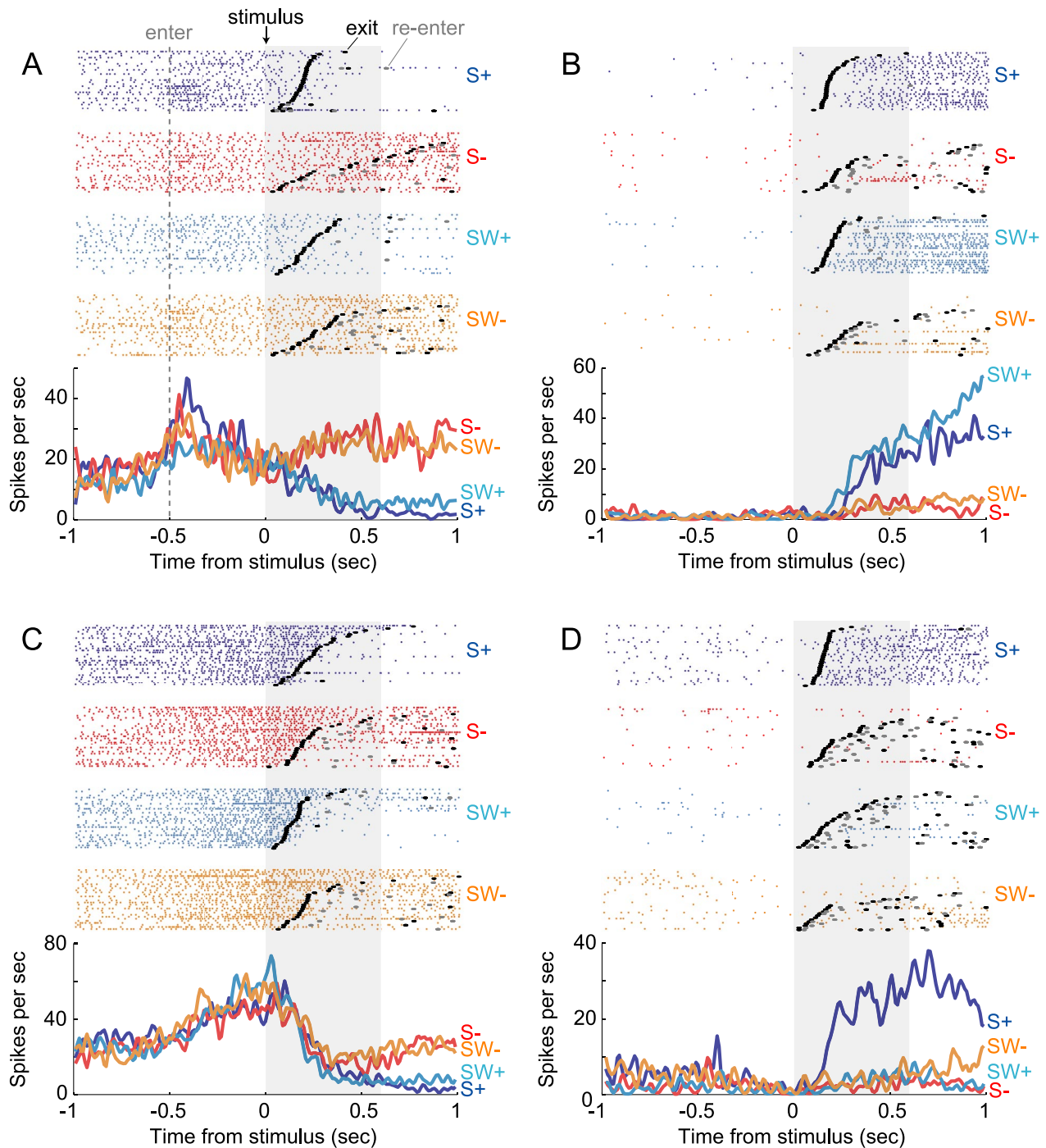


Figure 4. Neuronal activity related to components of action selection. Four examples of neurons that varied with various components of the task are shown. Activity is aligned to the onset of the stimuli. Rasters are shown in the upper portion of each panel, are separated by stimulus–reward pairing, and are sorted by reaction time (shortest RTs shown at the bottom). The bottom panels show peri-event histograms for each stimulus–reward condition and depict spike-density functions measured using a Gaussian kernel (width of 10 ms). **A**, This neuron was a leading decoder of a change in stimulus–reward contingency, with accuracy >80%. The neuron was modulated during the delay period (−0.5–0 s) but showed a reduced firing rate after rewarded stimuli. In contrast, it fired persistently throughout the reaction time epoch after unrewarded stimuli. Thus, the neuron fired in a very different manner to the same stimulus depending on the value (SW+ vs SW−). This activity represents what type of response would be rewarded, regardless of what response is actually made. **B**, A neuron that fired during Go responses to S+ and SW+ and decoded the change in the stimulus–reward contingency, with accuracy >80%. **C**, A neuron that strongly varied with the reaction time, but not with the stimulus–reward contingency (accuracy <60%). **D**, A neuron that fired only during locomotor behavior triggered by the low frequency tone. This neuron did not vary with the change in stimulus–reward contingencies (accuracy <60%) and was recorded during a session when the rat did not rapidly switch its behavior. Such an activity pattern was relatively rare.

ences between correct and error responses. The number of error trials is relatively low and to address this we would need to increase the number of trials that were analyzed before and after the switches in stimulus–reward contingencies. This would complicate the results by introducing longer-term variability into the analyses, which we were already explicitly measuring. As such, we have deferred error analysis for this data set.

Quantification of decoding results. All decoding results were summarized using receiver operating characteristic (ROC) analysis (using the verification library for R) and information theory (Krippendorff, 1986). ROC analysis is based on plotting, for a given type of trial, the fraction of true positives (e.g., fraction of actual SW+ trials correctly predicted to be SW+) versus the fraction of false positives (e.g., fraction of actual SW−

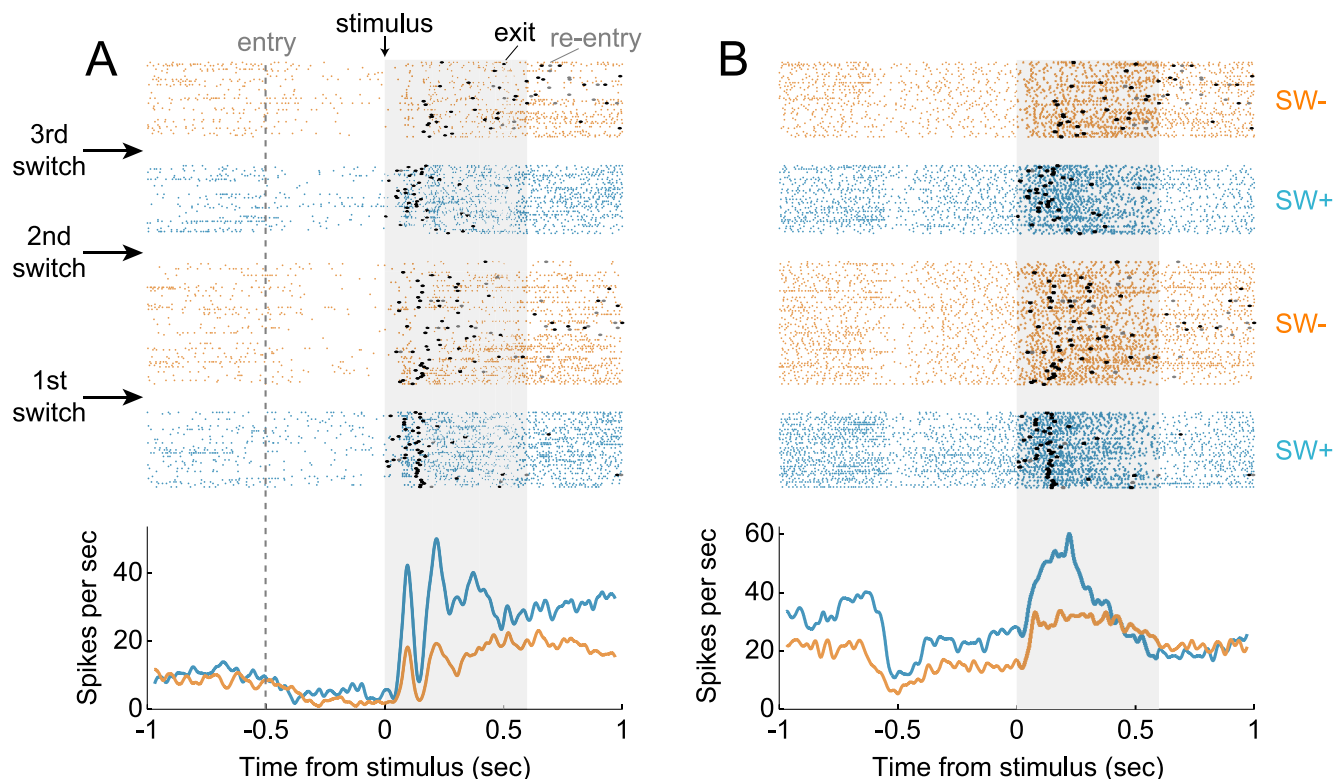


Figure 5. Reproducibility of neuronal changes to repeated switches in action selection. Two medial striatal neurons are shown that were recorded simultaneously in an animal that went through three switches in stimulus–reward associations. The neurons provide evidence that the rat striatum is reproducibly altered after multiple changes in stimulus–reward contingency. This session used a light stimulus as the switching stimulus. The beginning of the session is at the bottom of the rasters, with later trials in the session above. Rasters in light blue depict portions of the session when the switching stimulus was rewarded (SW+). Rasters in orange depict portions of the session when the switching stimulus was unrewarded (SW−). **A**, This neuron fired more after the stimulus and again during response initiation to SW+ compared with SW−. **B**, This neuron fired at an overall higher rate in blocks when the switching stimulus was rewarded (SW+) compared with blocks when the switching stimulus was not rewarded (SW−). This activity can represent whether a Go response to a switching stimulus would be rewarded, regardless of what the stimulus or response will be on a particular trial.

trials incorrectly predicted to be SW+). Significant levels of discrimination were then estimated by calculating the area under the ROC curve and the significance of the area under the curve was measured using a Wilcoxon rank-sum test. Essentially, this procedure measured the probability that the area under the curve was significantly greater than that expected for random data (Mason and Graham, 2002).

Information theory was used to calculate the mutual information between predictions of trial type (e.g., predicted SW− or SW+) made by the naive Bayes decoder and the actual trial type (e.g., actually SW− or SW+). Mutual information was assessed for the confusion matrix, a matrix with two columns for the actual class types and two rows for the predicted class types. The significance of the level of mutual information was assessed using a χ^2 test (Krippendorff, 1986). Detailed summaries of our use of these analyses are available (Laubach et al., 2000; Narayanan et al., 2005) and further details are provided in the supplemental material, available at www.jneurosci.org. For both the ROC and information theory measures, significant decoding was assessed using a criterion of $p < 0.05$.

Dynamics of changes in neuronal activity across trials. Significant differences in Go responding and neuronal predictions of a rewarded Go response (switching stimulus now SW+) were estimated using the methods described above for measuring significant changes in firing rate (i.e., *scstest*) and for measuring the timing of such changes (i.e., *breakpoints*). Data were compared for blocks of 30 trials before and after a switch in stimulus–reward contingencies. Behavioral responding was measured for each trial as a 0 (No-go response) or 1 (Go response). Neuronal predictions of responding were measured either using the posterior probabilities from the naive Bayes decoder or using a thresholded version of the posterior probability, with 0 for predictions of No-go responding and 1 for predictions of Go responding. We obtained equivalent results for the estimation of the timing of neuronal changes using the raw pos-

terior probabilities and the thresholded probabilities. As above, we obtained equivalent results using classic *F*-statistic (Chow, 1960), Bayesian (Barry and Hartigan, 1993), and other (Zeileis et al., 2003) methods for structural change analysis.

To ensure that the results were specific to changes in stimulus context, we also only analyzed data from 10 of that 14 rats that made Go responses to the switching stimulus after the first presentation of the unrewarded tone stimulus (S−) and that had neurons that fired during the period of Go responding. Two of the 14 rats switched spontaneously and were excluded, as their behavior was not comparable to the rest of animals. These animals may have discovered the switch in stimulus–reward contingencies by randomly checking for rewards at the time that the stimulus block changed and so may not have experienced a change in stimulus context. Two other animals did not have neurons that were sensitive to action selection, and behavioral data from these animals was excluded from further analysis.

Time course of changes in neuronal activity within a trial. A “moving window” analysis was used to quantify the time course of changes in neuronal activity after a change in stimulus–reward contingencies. This analysis explored when, within a trial, neuronal sensitivity developed. Firing rates were measured using a 0.6 s time-window. The window was stepped in 0.05 s increments over the period from 0.6 s before to 0.4 s after the onset of the switching stimulus (SW−/SW+ noise). The first window in the series represented the firing rate from 0.1 s before entry into the nosepoke until the time of stimulus onset (0.6 s later, referenced as 0 s). The final window represented the firing rate from 0.4 to 1.0 s after stimulus onset, when rats moved to collect rewards or initiated new trials.

Firing rates in the series of time-windows were then evaluated using decoding methods. Naive Bayes classifiers were trained and tested with data from each time-window. Identical methods to those described

above were used. Two measures of the success of classification (area under the ROC curve, mutual information measured from the confusion matrix) were noted for each time-window. The average values of mutual information were then plotted for the series of time-windows. Plots were made for all neurons in the medial and ventral striatum, and for all neurons that significantly decoded the change in stimulus value during the blocks of ± 30 trials around the switch in stimulus–reward contingencies. Changes in the data series were analyzed using change-point analysis, as described above. Two points in the data series were noted: the first time-window that had a significant level for the F (or Chow) statistic and the time-window that had the maximum level for the F statistic (the change-point).

Results

Go/No-go discrimination task

To determine if the medial striatum is sensitive to the reward value of stimuli, neuronal activity was recorded as rats ($n = 14$) performed a Go/No-go discrimination task (Fig. 1A). The sessions started with an 8 kHz tone serving as the rewarded stimulus (S+), and a noise burst or light serving as the unrewarded stimulus (SW−) (Fig. 1B). Initially, rats made Go responses predominantly to the S+ stimulus and minimally to the SW− stimulus (Fig. 1C). After receiving an unrewarded stimulus, rats tended to perform a new trial immediately. After rats demonstrated 90% accuracy on this discrimination (see Materials and Methods), the reward value of the previously unrewarded stimulus was switched, now serving as a rewarded stimulus (SW+) (Fig. 1B). A 30 kHz tone served as the unrewarded stimulus in this part of the session (S−). Rats adjusted their behavior overall to respond selectively to SW+ (rewarded switched stimulus, Fig. 1C).

During recordings, rats performed the task with a median accuracy of 83% (IQR 81–86%, 28 sessions, 2 sessions per rat, one using a noise burst and one using lights as the switching stimulus). The stimuli and switches were familiar to the animals, which successfully adjusted their behavior (90% accuracy) after a change in action selection with a median of 10.5 errors [interquartile range (IQR) 6–22.5 errors] within 50.5 trials (IQR 27.5–84). Trial by trial response and RT data are shown in Figure 2 for all sessions and for one example session. Distributions of RT and time until response are shown in supplemental Figure S1 (available at www.jneurosci.org as supplemental material) for the noise sessions analyzed below.

Spike activity in the medial and ventral striatum

Spike activity was recorded from 552 single neurons during 28 sessions. The neurons had a median firing rate of 4.97 Hz (IQR 1.65–10.10 Hz) during nontrial portions of the sessions (grooming or chamber exploration). Few neurons (3%, 19/552) could be classified as tonically active (Kimura et al., 1990). Recording sites were localized to the medial (341 neurons) or ventral (211 neurons) portions of the striatum (Fig. 3). There were no significant differences in waveform sizes or overall firing rates between the medial and ventral neurons (rank-sum test, $p > 0.05$).

Modulations of spike activity around stimulus onset

Many neurons (73%, 403/552) were modulated during the pre-stimulus epoch (± 1 s of onset), with modulation assessed using a structural change test (Zeileis et al., 2002). Neuronal response properties were heterogeneous, with some neurons showing modulations in firing rate at the presentation of the stimulus and others being modulated later in the reaction time epoch (Fig. 4). ~20% of these neurons (83/403) were modulated during a narrower time window around the stimulus, measured by comparing firing rates in 100 ms epochs before and after the stimulus

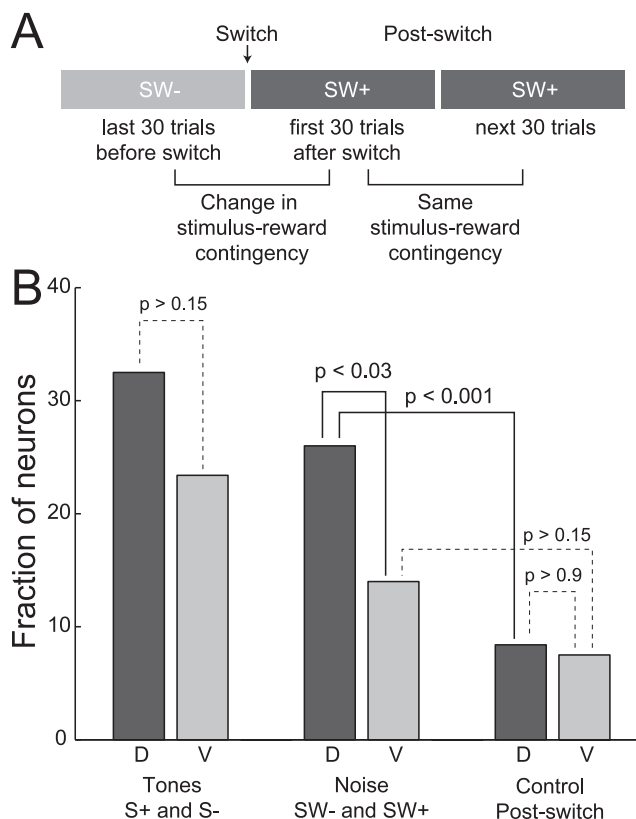


Figure 6. Stimulus–reward contingencies can be decoded with medial striatal neuronal activity. **A**, The organization of trials for the decoding analysis is shown. To assess neuronal sensitivity to changes in stimulus–reward contingency, firing rates on the last 30 trials before the switch were compared with the first 30 trials after the switch (SW− vs SW+). The first trial after the switch was defined as the first presentation of SW+ after an S−. To control for nonspecific effects over the experimental sessions, we compared neuronal activity from the first 30 SW+ trials to activity from the next 30 SW+ trials. **B**, The fractions of neurons that significantly discriminated between trial types are shown by area. The middle columns demonstrate that significantly more medial neurons were sensitive to a switch in the stimulus–reward contingency than ventral neurons (40/154 medial, dark gray bar, vs 15/107 ventral, light gray bar, Proportions test: $\chi^2 = 4.73$, $p < 0.03$). For the medial, but not ventral, neurons this was significantly greater than changes observed in the control period (40 vs 13 medial neurons; $\chi^2: 15.41$, $df = 1$, $p < 10^{-3}$; 15 vs 8 ventral neurons; $\chi^2: 1.75$, $p > 0.15$). Ventral neurons could, however, significantly decode stimulus–reward contingencies for the tone stimuli, which had fixed reward values throughout training. Approximately equal proportions of neurons in the medial and ventral striatum decoded these stimuli (ventral: S+ vs S−: 25 of 107, >23.4%, compared with SW− vs SW+ as above; $\chi^2: 3.87$, $df = 1$, $p < 0.05$). These results, together, suggest that changes in action selection to a stimulus with flexible reward value are represented by neurons in the medial, but not the ventral, striatum.

(signed rank test, $p < 0.05$). For most neurons modulated within this window around the stimulus, activity increased in the 100 ms after the presentation of the stimulus (69%, 57/83). Average population activity by area and trial type is displayed in supplemental Figure S2, available at www.jneurosci.org as supplemental material. Further analysis is only for sessions using a noise stimulus as the switching stimulus SW ($n = 261$ neurons: 154 medial, 107 ventral).

Decoding changes in components of action selection

Approximately 21% of striatal neurons (55 of 261) were sensitive to the change in the stimulus–reward contingency. These results were based on decoding analyses. Neurons were considered to decode the change in stimulus–reward contingencies during the peri-switch block if a Wilcoxon rank-sum test on the area under the ROC curve and a χ^2 test on the decoding results confusion

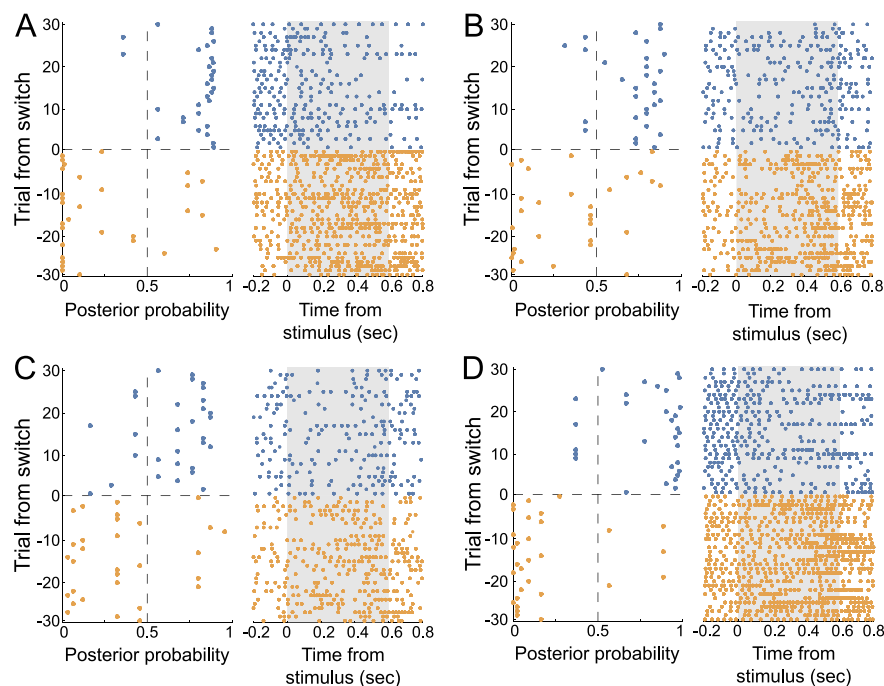


Figure 7. Rapid decoding of action selection after a change in stimulus–reward contingency. Neurons that were leading decoders of stimulus–reward contingency in one experimental session are shown. Posterior probabilities from the decoding (SW– vs SW+) are shown on the left for trials with unrewarded switching stimuli (SW–, orange) and with rewarded switching stimuli (SW+, light blue). The probabilities of Go responding changed from low to high after the first presentation of the unrewarded tone (S– presented at trial 0, compare trials –30 to –1 with trials 1 to 30). Neuronal firing patterns associated with each posterior probability are shown on the right. The time epoch that was analyzed with the decoding method is highlighted by the gray boxes in the raster plot (stimulus onset at 0 to 600 ms after).

matrix were significant at $p < 0.05$. Examples of task-related neuronal activity, recorded in the medial striatum, are shown in Figure 4. Rasters were sorted by the reaction time, to reveal neurons that varied with the speed of responding in the task. (Larger versions of the raster plots and plots of the neurons in A and B around both the stimulus and the subsequent response are shown in figures in the supplemental materials, available at www.jneurosci.org.)

The neuron in Figure 4A was a leading decoder of a change in stimulus–reward contingency, predicting trials before and after the switch with accuracy $>80\%$. The neuron fired at an increased rate, starting when the animal entered the nosepoke aperture. On trials with rewarded stimuli (S+ and SW+), the neuron exhibited a reduced firing rate, starting at stimulus onset. On trials with unrewarded stimuli (S– and SW–), the neuron maintained its firing rate from the delay period. This information predicts whether a Go response would be rewarded, and not necessarily what the response will actually be. Approximately 21% of striatal neurons (55 of 261) contained information about the change in stimulus–reward contingencies. The neuron in Figure 4B varied with Go and No-go responding, becoming active some time after the rat started to withdraw from the nosepoke aperture (locomotion is approximately from ~ 250 ms to ~ 1000 ms after the onset of the stimulus). The neuron fired during reward collection behavior to both S+ and SW+. This neuron significantly decoded the change in stimulus–reward contingency, with accuracy $>80\%$. The neuron in Figure 4C showed strong changes in activity at the time of withdrawal from the nosepoke (following the pattern in the reaction times over trials), but did not decode the change in stimulus–reward contingency (accuracy $<60\%$ correct). The neuron in Figure 4D fired only during locomotor be-

havior triggered by the low frequency tone. This neuron did not vary with the change in stimulus–reward contingencies (accuracy $<60\%$ correct) and, crucially, was recorded during a session when the rat did not rapidly switch its behavior.

Reproducibility of neuronal changes

In one subject, neuronal activity was recorded during a session in which there were multiple switches in action selection in the same experimental session. Two switch sensitive neurons from this animal are shown in Figure 5. The neuron in A had greater activity immediately after stimulus and during the reward collection response when the switching stimulus was rewarded (SW+). The neuron in B had greater activity throughout the trial during blocks when responses to the switching stimulus were rewarded (SW+). These neurons provide evidence that rat striatal neurons reproducibly change their activity after changes in stimulus–reward contingencies.

Comparison of neurons in medial and ventral striatum

Significantly more neurons in the medial striatum (40 of 154 neurons, 26.0%) decoded changes in the stimulus–reward contingencies compared with neurons in the ventral striatum (15 of 107 neurons, 14.0%; Proportions test: $\chi^2 = 4.73$, $p < 0.03$) (Fig. 6). Fewer neurons in both areas discriminated between trials from the first and second blocks of 30 trials after the switch in stimulus–reward contingencies (13 of 154 neurons, 8.4%, in medial striatum; 8 of 107 neurons, 7.5%, in ventral striatum). As stimulus–reward contingencies were fixed during the postswitch block (all trials were potentially rewarded), this period of the task served as a control for spontaneous fluctuations in neuronal firing and for effects of motivational changes over the course of the session. More neurons varied around the switch in the medial (40 vs 13; $\chi^2: 15.41$, $df = 1$, $p < 10^{-4}$), but not the ventral (15 vs 8; $\chi^2: 1.75$, $p > 0.1$), striatum. These results suggest that neurons in the medial, but not the ventral, striatum were sensitive to changes in the value of the flexible stimulus.

To assess how neurons fired in response to stimuli with fixed stimulus–reward associations (i.e., the tones), we used the same analysis methods as above to decode Go responses made to the S+ and S– stimuli during the peri-switch block (± 30 trials around the change in the value of SW). This analysis revealed that more neurons in both parts of the striatum fired differentially during Go and No-go responses made to the tones than during similar responses to SW (Fig. 6B). Significantly more neurons in the ventral striatum were differentially activated by the tones (25 of 107, $>23.4\%$) compared with SW (13 of 107, 8.5%, as above). A proportions test showed that these numbers of neurons were significantly different ($\chi^2: 3.87$, $df = 1$, $p < 0.05$). Therefore, while changes in action selection to the switching stimulus resulted in altered neuronal activity only in the medial striatum, neurons throughout the medial striatum fired differentially during Go and No-go responses made to the tone stimuli.

Dynamics of the representation of stimulus–reward contingencies across trials

Neuronal activity changed rapidly after a change in stimulus–reward contingencies. Plots of posterior probabilities from the decoding analysis are shown in Figure 7 for four simultaneously recorded neurons. These neurons rarely predicted Go responses before the switch and significantly predicted Go responses on most trials after the switch. There was an abrupt change in the posterior probabilities after the first presentation of S− (trial 0). Raster plots are shown immediately to the right of the posterior probabilities, allowing for a comparison of the raw firing patterns associated with predictions of Go and No-go responses. For these neurons, spike activity during the reaction time epoch occurred at a different rate to the unrewarded (SW−) and rewarded (SW+) stimulus.

To compare behavioral and neuronal measures at the group level, we plotted the fraction of rats that made Go responses and the fractions of neurons that predicted that a Go response would be rewarded for trials around the change in stimulus–reward contingency (SW− to SW+) (Fig. 8A,C). These plots clearly showed that neurons stably predicted that Go responses would be rewarded before the animals consistently made such Go responses. The change-point for the neurons' decoding of the stimulus–reward contingency was on the first trial with the noise stimulus after the presentation of the unrewarded tone (Chow's F statistic: 1104.6, df: 1, p : $<10^{-15}$) (Fig. 8C). In contrast, the change-point for the behavioral data was between the fourth and fifth trial after the first presentation of the unrewarded tone. Reaction times changed faster than Go responding (change-point at trial 2 with SW+; Chow's F statistic: 66.76, df: 1, p : $<10^{-10}$) (Fig. 8B), but still occurred some time after the first presentation of the unrewarded tone. These results suggest that the switch in stimulus–reward contingencies led to changes in neuronal activity in the striatum that preceded changes in behavior.

Change-point analysis revealed a similar time course of altered neuronal activity in the medial and ventral after the switch in stimulus–reward contingencies. That is, the 40 medial and 15 ventral neurons that significantly decoded the switch in contingencies changed activity on the first trial into the postswitch block (Fig. 8D). This result, together with our data on the fractions of neurons that changed around the switch (Fig. 6), suggests a difference in the quantity, but not the quality, of neurons that dynamically encode action selection in the medial and ventral striatum.

Time course of changes in neuronal activity relative to the stimulus within trials

A “moving window” analysis (described in Materials and Methods) (Fig. 9A) showed that the decoding of changes in stimulus–

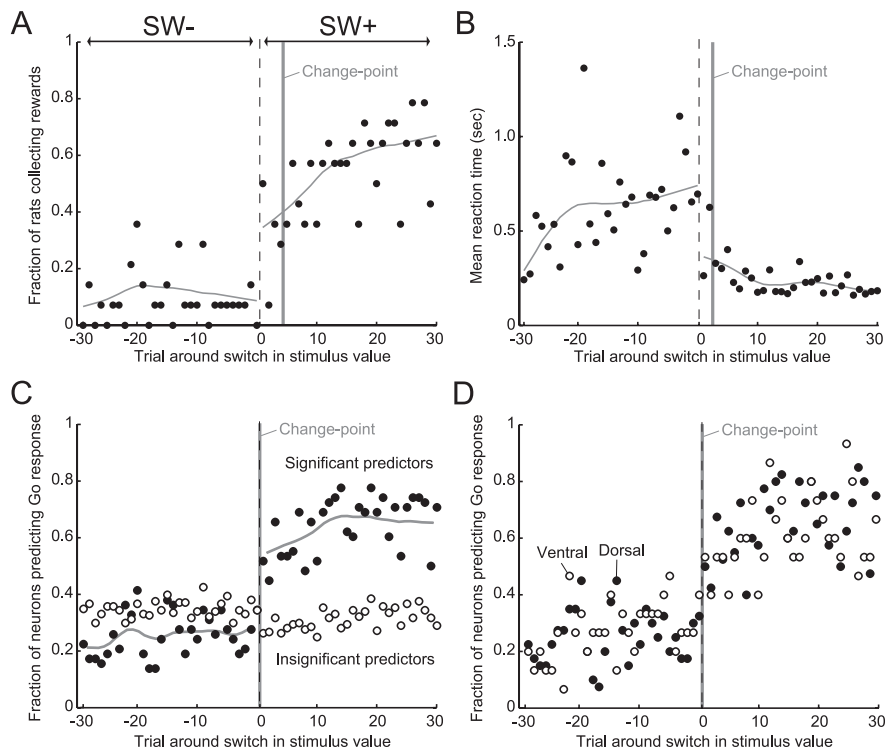


Figure 8. Dynamics of neuronal and behavioral measures of action selection after changes in stimulus–reward contingency. **A**, The plot shows the proportion of rats that collected rewards on trials with switching stimuli (SW) surrounding a switch in reward contingencies (SW− to SW+). The left part of the plot shows trials when the switching stimulus was unrewarded (SW−). The right part of the plot shows trials when the switching stimulus was rewarded (SW+). The behavioral data here is from 10 rats that were used for decoding analysis (rats that changed based on stimulus context and that were accurate after the switch in stimulus–reward contingency). The change point for all panels was calculated using a structural change test as described in Materials and Methods. **B**, Mean reaction times on trials before and after the switch in reward contingencies. **C**, **D**, The fractions of neurons that successfully predicted a rewarded stimulus on each trial surrounding a switch in reward contingencies. Before the switch, ~20% of neurons predicted a rewarded stimulus (incorrect predictions). After the switch, ~70% of neurons predicted a rewarded stimulus (correct predictions). **C**, Predictions of reward stimuli based on firing rates of neurons in the striatum (both medial and ventral) are shown for neurons that were significant predictors (black) and those that were not (white). Change-point analysis identified the first trial after the switch as containing significantly different information for significant predictors. There was no systematic increase in information in nonpredictive neurons at the time of the switch. **D**, Results are depicted for significant predictors of the change in stimulus–reward contingency from medial (black) and ventral (white) striatum. The results suggest that there are neurons in both parts of the striatum with similar dynamic across trials. However, the proportion of such neurons is significantly greater in the medial striatum (see Fig. 6).

reward contingencies was due to differences in firing rates in the medial, but not the ventral, striatum that started around the time of the stimulus (denoted by the left vertical dashed line in Fig. 9B, labeled “First significant deviation in F statistic”) and reached a maximum level at ~0.2 s after stimulus onset (denoted by the right vertical dashed line in Fig. 9B, labeled “Change-point”). This was during the period of action selection, before when rats either turned and walked across the operant chamber to collect rewards or initiated new trials. Differences in activity did not occur before the stimulus, and are therefore not likely to have been due to activity related to response initiation. At these early points in the trials, significantly more neurons in the medial striatum varied with the switch in stimulus–reward contingencies (Fig. 9C).

Discussion

We trained rats to perform a discrimination task that requires flexible responding to familiar stimuli. Rats quickly adjusted their behavior when the reward value of a given stimulus switched. This rapid, flexible behavior enabled us to assess the sensitivity of striatal neurons to changes in action selection. We found that a significant frac-

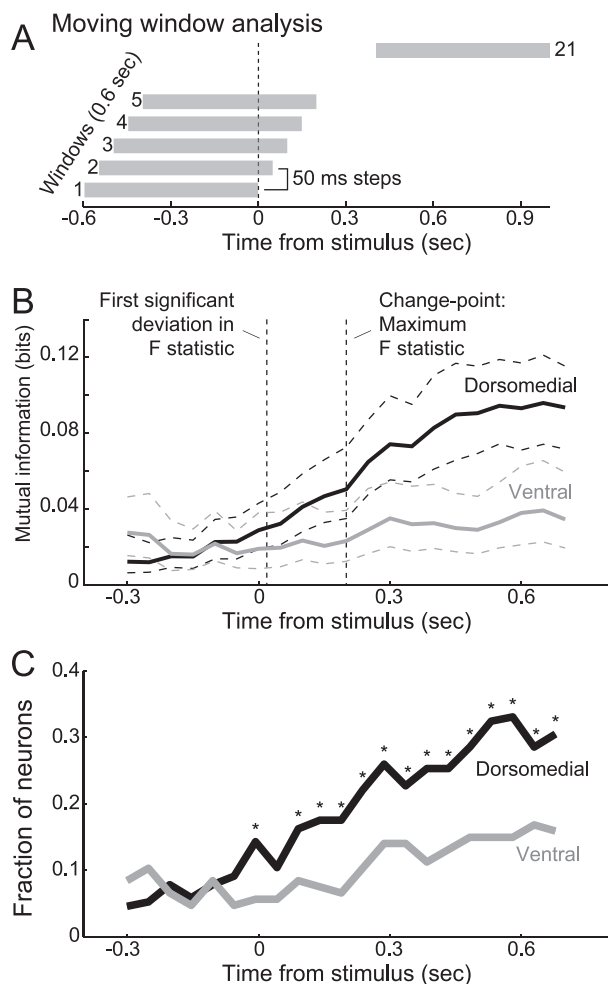


Figure 9. Time course of changes in neuronal activity within a trial. **A**, A “moving window” analysis was used to quantify the time course of changes in neuronal activity after a change in stimulus–reward contingencies. Firing rates were measured using a 0.6 s time-window that was stepped in 0.05 s increments over the period from 0.6 s before to 0.4 s after the onset of the switching stimulus (SW−/SW+ noise). A probabilistic classifier was trained and tested with firing rate from each neuron in each time-window. **B**, Decoding of changes in stimulus–reward contingencies was due to information that began to grow at the time of the stimulus (denoted by the left vertical dashed line in Fig. 9B, labeled “First significant deviation in F statistic”) and reached a maximum level at ~0.2 s after stimulus onset (denoted by the right vertical dashed line in Fig. 9B, labeled “Change-point”). More information was accumulated across the population of medial striatum neurons (black line \pm 95% confidence interval) than ventral striatum neurons (gray line \pm 95% confidence interval). Results are graphed at the center of each window step. **C**, The proportion of neurons that decode stimulus–reward contingencies for flexible stimuli (SW− vs SW+) increases earlier and more robustly in the medial striatum (black line) than ventral striatum (gray line). Asterisks indicate time-points at which the proportions between medial and ventral striatum differed significantly (Proportions test: $\chi^2 p < 0.05$).

tion of neurons in the medial, but not the ventral, striatum rapidly tracked switches in stimulus–reward contingencies, which is crucial for adapting flexible behavior. These changes in neuronal activity occurred before changes in behavioral performance, and suggest a role of the striatum in driving adaptive behavior. Our results are the first demonstration, outside of the hippocampus (Wirth et al., 2003), of modulations of neuronal activity that precede changes in behavioral decision making.

Relationship of the striatum to frontal cortex

Our findings demonstrate that rapid neuronal sensitivity to changes in the behavioral significance of auditory and visual stimuli is not a unique feature of the primate striatum (Wise and

Murray, 2000). Even without contributions from a true granular prefrontal cortex (Preuss, 1995), the rat striatum has the ability to rapidly track action selection related to sensory stimuli. Although the task we used is different in detail from tasks more recently used with monkeys, our study is analogous to the work of Hikosaka and colleagues (Itoh et al., 2003; Ding and Hikosaka, 2006; Nakamura and Hikosaka, 2006), wherein animals repeatedly adapt to familiar changes in stimuli–reward associations. These tasks potentially allow animals to anticipate how they will respond to a particular stimulus, even without knowing what stimulus will be specifically offered on a given trial. Itoh et al. (2003) have interpreted their findings as reflecting the active bias of lateralized eye movements. Other groups have reported similar learning-related changes in task-related activity in the primate striatum (Brasted and Wise, 2004; Pasupathy and Miller, 2005; Williams and Eskandar, 2006). These studies suggest that some striatal neurons are involved in the rapid reconfiguration of motor plans.

Lesions of frontal and striatal areas have been shown to impair stimulus–response learning (Winocur and Eskes, 1998). Evidence for changes in neurons in the frontal cortex during associative operant conditioning has been reported previously (Schoenbaum et al., 1998). Neuroimaging studies show that multiple areas in frontal cortex are activated during the learning of arbitrary stimulus–response associations (Toni and Passingham, 1999). If similar activations occur in rat frontal cortex during our task, then we would expect to observe widespread changes in striatal firing during learning, both in the medial and ventral striatum.

Relationship of the striatum to other subcortical structures

Both the medial and ventral striatum receive dopaminergic input, from the substantia nigra pars compacta and ventral tegmental area respectively. If these inputs were similar, then there would be similar sensitivity to reinforcement learning between the medial and ventral striatum. Additionally, information is postulated to spiral or progress from the ventral striatum to dorsal striatum, via interconnections between the striatum and the dopaminergic inputs (Haber et al., 2000). Other subcortical inputs that project to the striatum, such as the amygdala (Kelley et al., 1982; Muramoto et al., 1993) and thalamic nuclei (Oyoshi et al., 1996; Komura et al., 2001), may contribute additional reward information to both regions. The amygdala in particular projects relatively widely to most of the striatum, although less so to the lateral striatum (McDonald, 1991). This could also explain reward-related similarities in medial and ventral striatal physiology while contrasting them with the lateral striatum.

Role of medial striatum in decision making

In contrast to consistently rewarded stimuli, medial and ventral striatal neurons had different sensitivities to switches in the reward contingencies of flexibly rewarded stimuli. For switching stimuli, learning-related neuronal activity was more prominent and more closely tied to changes in task performance in the medial than ventral striatum. Sensitivity to changes in stimulus–reward contingencies emerged earlier both within and across trials in the medial striatum. As flexible behavior requires online tracking of context and reward history in the selection of appropriate action, changes in value-related information might be provided preferentially to the medial striatum, either by a particular confluence of cortical inputs (McGeorge and Faull, 1989; Laubach, 2005) or by a presumably unique population of subcortical inputs communicating information such as reward prediction errors (Schultz, 1998).

Striatal neurons may bias animals to attempt reward collection after a change in stimulus–reward contingencies

When there was a switch in stimulus–reward contingencies, it typically took our rats several trials before they responded consistently to the newly rewarded stimulus. In contrast, neuronal activity within the medial striatum changed almost immediately. This neuronal sensitivity grew both across and within trials as behavior progressed. It emerged heterogeneously but in increasing proportion across the neuronal population. The difference between earlier neuronal sensitivity and later behavioral expression across trials suggests that a substantial population of striatal neurons may be required to select a particular action (Frank and Claus, 2006). Neurons in the medial striatum may represent changing action selection under the influence of multiple signals, perhaps derived from frontal cortex, and acting over multiple time-scales (Fusi et al., 2007).

Through connections with frontal cortex, the striatum may prepare or bias downstream targets of the basal ganglia to process forthcoming sensory stimuli. Such a role for the basal ganglia has been formalized in several theoretical models (Frank et al., 2001; Lo and Wang, 2006). Poststimulus changes in neuronal activity may serve as instructive signals that evaluate the outcome of the current trial and improve behavioral performance on future trials (Houk and Wise, 1995). This concept is supported by a study in which microstimulation on the current trial helped bias responses for the next trial (Nakamura and Hikosaka, 2006). Alternatively, these signals may reflect the values of the animals' actions (Lau and Glimcher, 2007). In this way, the medial striatum may be part of both a preparative and evaluative system for the control of purposeful behavior, contrasting with a more habit like lateral striatum (Yin and Knowlton, 2006) or a more reinforcement focused ventral striatum (Carelli and Deadwyler, 1994). Our data supports a theory that the medial striatum has a specialized role in the selection of actions during flexible behavior.

References

- Adams S, Kesner RP, Ragozzino ME (2001) Role of the medial and lateral caudate-putamen in mediating an auditory conditional response association. *Neurobiol Learn Mem* 76:106–116.
- Amalric M, Koob GF (1987) Depletion of dopamine in the caudate nucleus but not in nucleus accumbens impairs reaction-time performance in rats. *J Neurosci* 7:2129–2134.
- Apicella P, Ljungberg T, Scarnati E, Schultz W (1991) Responses to reward in monkey dorsal and ventral striatum. *Exp Brain Res* 85:491–500.
- Balleine BW, Delgado MR, Hikosaka O (2007) The role of the dorsal striatum in reward and decision-making. *J Neurosci* 27:8161–8165.
- Barry D, Hartigan JA (1993) A Bayesian analysis for change point problems. *J Am Stat Assoc* 88:309–319.
- Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10:433–436.
- Brasted PJ, Wise SP (2004) Comparison of learning-related neuronal activity in the dorsal premotor cortex and striatum. *Eur J Neurosci* 19:721–740.
- Carelli RM, Deadwyler SA (1994) A comparison of nucleus accumbens neuronal firing patterns during cocaine self-administration and water reinforcement in rats. *J Neurosci* 14:7735–7746.
- Carelli RM, James SG (2000) Nucleus accumbens cell firing during maintenance, extinction, and reinstatement of cocaine self-administration behavior in rats. *Brain Res* 866:44–54.
- Carli M, Jones GH, Robbins TW (1989) Effects of unilateral dorsal and ventral striatal dopamine depletion on visual neglect in the rat: a neural and behavioural analysis. *Neuroscience* 29:309–327.
- Chang JY, Sawyer SF, Lee RS, Woodward DJ (1994) Electrophysiological and pharmacological evidence for the role of the nucleus accumbens in cocaine self-administration in freely moving rats. *J Neurosci* 14:1224–1244.
- Chow GC (1960) Tests of equality between sets of coefficients in two linear regressions. *Econometrica* 28:591–605.
- Clarke HF, Robbins TW, Roberts AC (2008) Lesions of the medial striatum in monkeys produce perseverative impairments during reversal learning similar to those produced by lesions of the orbitofrontal cortex. *J Neurosci* 28:10972–10982.
- Ding L, Hikosaka O (2006) Comparison of reward modulation in the frontal eye field and caudate of the macaque. *J Neurosci* 26:6695–6703.
- Domingos P, Pazzani M (2004) On the optimality of the simple Bayesian classifier under zero-one loss. *Mach Learn* 29:103–130.
- Featherstone RE, McDonald RJ (2005) Lesions of the dorsolateral or medial striatum impair performance of a previously acquired simple discrimination task. *Neurobiol Learn Mem* 84:159–167.
- Ferry AT, Lu XC, Price JL (2000) Effects of excitotoxic lesions in the ventral striatopallidal–thalamocortical pathway on odor reversal learning: inability to extinguish an incorrect response. *Exp Brain Res* 131:320–335.
- Frank MJ, Claus ED (2006) Anatomy of a decision: striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal. *Psychol Rev* 113:300–326.
- Frank MJ, Loughry B, O'Reilly RC (2001) Interactions between frontal cortex and basal ganglia in working memory: a computational model. *Cogn Affect Behav Neurosci* 1:137–160.
- Fusi S, Asaad WF, Miller EK, Wang XJ (2007) A neural circuit model of flexible sensorimotor mapping: learning and forgetting on multiple time-scales. *Neuron* 54:319–333.
- Haber SN, Fudge JL, McFarland NR (2000) Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *J Neurosci* 20:2369–2382.
- Hansen BE (1997) Approximate asymptotic p values for structural-change tests. *J Bus Econ Stat* 15:60–67.
- Houk JC, Wise SP (1995) Distributed modular architectures linking basal ganglia, cerebellum, and cerebral cortex: their role in planning and controlling action. *Cereb Cortex* 5:95–110.
- Itoh H, Nakahara H, Hikosaka O, Kawagoe R, Takikawa Y, Aihara K (2003) Correlation of primate caudate neural activity and saccade parameters in reward-oriented behavior. *J Neurophysiol* 89:1774–1783.
- Janak PH, Chen MT, Caulder T (2004) Dynamics of neural coding in the accumbens during extinction and reinstatement of rewarded behavior. *Behav Brain Res* 154:125–135.
- John GH, Langley P (1995) Estimating continuous distributions in Bayesian classifiers. In: *Proceedings of the eleventh conference on uncertainty in artificial intelligence*, pp 338–345. San Mateo, CA: Morgan Kaufman.
- Kelley AE, Domesick VB, Nauta WJ (1982) The amygdalostratial projection in the rat—an anatomical study by anterograde and retrograde tracing methods. *Neuroscience* 7:615–630.
- Kimura M, Kato M, Shimazaki H (1990) Physiological properties of projection neurons in the monkey striatum to the globus pallidus. *Exp Brain Res* 82:672–676.
- Komura Y, Tamura R, Uwano T, Nishijo H, Kaga K, Ono T (2001) Retrospective and prospective coding for predicted reward in the sensory thalamus. *Nature* 412:546–549.
- Krippendorff K (1986) *Information theory: structural models for qualitative data*. Thousand Oaks, CA: Sage.
- Lau B, Glimcher PW (2007) Action and outcome encoding in the primate caudate nucleus. *J Neurosci* 27:14502–14514.
- Laubach M (2004) Wavelet-based processing of neuronal spike trains prior to discriminant analysis. *J Neurosci Methods* 134:159–168.
- Laubach M (2005) Who's on first? What's on second? The time course of learning in corticostriatal systems. *Trends Neurosci* 28:509–511.
- Laubach M, Wessberg J, Nicolelis MA (2000) Cortical ensemble activity increasingly predicts behaviour outcomes during learning of a motor task. *Nature* 405:567–571.
- Lo CC, Wang XJ (2006) Cortico-basal ganglia circuit mechanism for a decision threshold in reaction time tasks. *Nat Neurosci* 9:956–963.
- Mason S, Graham N (2002) Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: statistical significance and interpretation. *Q J Roy Meteor Soc* 128:2145–2166.
- McDonald AJ (1991) Organization of amygdaloid projections to the prefrontal cortex and associated striatum in the rat. *Neuroscience* 44:1–14.
- McGeorge AJ, Faull RL (1989) The organization of the projection from the cerebral cortex to the striatum in the rat. *Neuroscience* 29:503–537.
- Muramoto K, Ono T, Nishijo H, Fukuda M (1993) Rat amygdaloid neuron responses during auditory discrimination. *Neuroscience* 52:621–636.
- Nakamura K, Hikosaka O (2006) Facilitation of saccadic eye movements by

- postsaccadic electrical stimulation in the primate caudate. *J Neurosci* 26:12885–12895.
- Narayanan NS, Kimchi EY, Laubach M (2005) Redundancy and synergy of neuronal ensembles in motor cortex. *J Neurosci* 25:4207–4216.
- Nicola SM, Yun IA, Wakabayashi KT, Fields HL (2004) Cue-evoked firing of nucleus accumbens neurons encodes motivational significance during a discriminative stimulus task. *J Neurophysiol* 91:1840–1865.
- Oyoshi T, Nishijo H, Asakura T, Takamura Y, Ono T (1996) Emotional and behavioral correlates of mediodorsal thalamic neurons during associative learning in rats. *J Neurosci* 16:5812–5829.
- Palencia CA, Ragozzino ME (2004) The influence of NMDA receptors in the medial striatum on response reversal learning. *Neurobiol Learn Mem* 82:81–89.
- Pasupathy A, Miller EK (2005) Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature* 433:873–876.
- Paxinos G, Watson C (1998) The rat brain in stereotaxic coordinates, Ed 4. San Diego: Academic.
- Peoples LL, Uzviak AJ, Gee F, West MO (1999) Tonic firing of rat nucleus accumbens neurons: changes during the first 2 weeks of daily cocaine self-administration sessions. *Brain Res* 822:231–236.
- Preuss TM (1995) Do rats have prefrontal cortex? The Rose-Woolsey-Akert program reconsidered. *J Cog Neurosci* 7:1–24.
- Prokopenko VF, Pawlak AP, West MO (2004) Fluctuations in somatosensory responsiveness and baseline firing rates of neurons in the lateral striatum of freely moving rats: effects of intranigral apomorphine. *Neuroscience* 125:1077–1082.
- Ragozzino ME, Detrick S, Kesner RP (1999a) Involvement of the prelimbic-infralimbic areas of the rodent prefrontal cortex in behavioral flexibility for place and response learning. *J Neurosci* 19:4585–4594.
- Ragozzino ME, Wilcox C, Raso M, Kesner RP (1999b) Involvement of rodent prefrontal cortex subregions in strategy switching. *Behav Neurosci* 113:32–41.
- Ragozzino ME, Jih J, Tzavos A (2002a) Involvement of the medial striatum in behavioral flexibility: role of muscarinic cholinergic receptors. *Brain Res* 953:205–214.
- Ragozzino ME, Ragozzino KE, Mizumori SJ, Kesner RP (2002b) Role of the medial striatum in behavioral flexibility for response and visual cue discrimination learning. *Behav Neurosci* 116:105–115.
- Schoenbaum G, Setlow B (2003) Lesions of nucleus accumbens disrupt learning about aversive outcomes. *J Neurosci* 23:9833–9841.
- Schoenbaum G, Chiba AA, Gallagher M (1998) Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nat Neurosci* 1:155–159.
- Schoenbaum G, Nugent SL, Saddoris MP, Setlow B (2002) Orbitofrontal lesions in rats impair reversal but not acquisition of go, no-go odor discriminations. *Neuroreport* 13:885–890.
- Schultz W (1998) Predictive reward signal of dopamine neurons. *J Neurophysiol* 80:1–27.
- Setlow B, Schoenbaum G, Gallagher M (2003) Neural encoding in ventral striatum during olfactory discrimination learning. *Neuron* 38:625–636.
- Siegel S (1956) Nonparametric statistics for the behavioral sciences. New York: McGraw-Hill.
- Stalnaker TA, Franz TM, Singh T, Schoenbaum G (2007) Basolateral amygdala lesions abolish orbitofrontal-dependent reversal impairments. *Neuron* 54:51–58.
- Taha SA, Nicola SM, Fields HL (2007) Cue-evoked encoding of movement planning and execution in the rat nucleus accumbens. *J Physiol* 584:801–818.
- Teagarden MA, Rebec GV (2007) Subthalamic and striatal neurons concurrently process motor, limbic, and associative information in rats performing an operant task. *J Neurophysiol* 97:2042–2058.
- Toni I, Passingham RE (1999) Prefrontal-basal ganglia pathways are involved in the learning of arbitrary visuomotor associations: a PET study. *Exp Brain Res* 127:19–32.
- Tremblay L, Hollerman JR, Schultz W (1998) Modifications of reward expectation-related neuronal activity during learning in primate striatum. *J Neurophysiol* 80:964–977.
- Voorn P, Vanderschuren LJ, Groenewegen HJ, Robbins TW, Pennartz CM (2004) Putting a spin on the dorsal-ventral divide of the striatum. *Trends Neurosci* 27:468–474.
- White IM, Rebec GV (1993) Responses of rat striatal neurons during performance of a lever-release version of the conditioned avoidance response task. *Brain Res* 616:71–82.
- Williams ZM, Eskandar EN (2006) Selective enhancement of associative learning by microstimulation of the anterior caudate. *Nat Neurosci* 9:562–568.
- Winocur G, Eskes G (1998) Prefrontal cortex and caudate nucleus in conditional associative learning: dissociated effects of selective brain lesions in rats. *Behav Neurosci* 112:89–101.
- Wirth S, Yanike M, Frank LM, Smith AC, Brown EN, Suzuki WA (2003) Single neurons in the monkey hippocampus and learning of new associations. *Science* 300:1578–1581.
- Wise SP, Murray EA (2000) Arbitrary associations between antecedents and actions. *Trends Neurosci* 23:271–276.
- Yin HH, Knowlton BJ (2006) The role of the basal ganglia in habit formation. *Nat Rev Neurosci* 7:464–476.
- Yin HH, Knowlton BJ, Balleine BW (2005) Blockade of NMDA receptors in the medial striatum prevents action-outcome learning in instrumental conditioning. *Eur J Neurosci* 22:505–512.
- Zeileis A, Leisch F, Hornik K, Kleiber C (2002) Strucchange: an R package for testing for structural change in linear regression models. *J Stat Softw* 7:1–38.
- Zeileis A, Kleiber C, Krämer W, Hornik K (2003) Testing and dating of structural changes in practice. *Comput Stat Data Anal* 44:109–123.