Neural correlates of stimulus–response and response–outcome associations in dorsolateral versus dorsomedial striatum

Thomas A. Stalnaker[1], Gwendolyn G. Calhoon[2], Masaaki Ogawa[1], Matthew R. Roesch[1] and Geoffrey Schoenbaum[1,3]*

[1] Department of Anatomy and Neurobiology, University of Maryland School of Medicine, Baltimore, MD, USA
[2] Program in Neuroscience, University of Maryland School of Medicine, Baltimore, MD, USA
[3] Department of Psychiatry, University of Maryland School of Medicine, Baltimore, MD, USA

Considerable evidence suggests that there is functional heterogeneity in the control of behavior by the dorsal striatum. Dorsomedial striatum may support goal-directed behavior by representing associations between responses and outcomes (R–O associations). The dorsolateral striatum, in contrast, may support motor habits by encoding associations between stimuli and responses (S–R associations). To test whether neural correlates in striatum in fact conform to this pattern, we recorded single-units in dorsomedial and dorsolateral striatum of rats performing a task in which R–O contingencies were manipulated independently of S–R contingencies. Among response-selective neurons in both regions, activity was significantly modulated by the initial stimulus, providing evidence of S–R encoding. Similarly, response selectivity was significantly modulated by the associated outcome in both regions, providing evidence of R–O encoding. In both regions, this outcome-modulation did not seem to reflect the relative value of the expected outcome, but rather its specific identity. Finally, in both regions we found correlates of the available action–outcome contingencies reflected in the baseline activity of many neurons. These results suggest that differences in information content in these two regions may not determine the differential roles they play in controlling behavior, demonstrated in previous studies.

Keywords: dorsal striatum, habit, goal-directed behavior, electrophysiology, single-unit activity, decision making, learning, rat

INTRODUCTION

The basal ganglia have long been associated with the control of behavior. In particular, the dorsal striatum is thought to support motor habits by encoding associations between sensory stimuli and movements (stimulus–response, or S–R associations) (Squire et al., 1993; Graybiel, 1998; Devan and White, 1999; Jog et al., 1999; Featherstone and McDonald, 2004a,b, 2005; Yin et al., 2004; Atallah et al., 2007; Tang et al., 2007). In theory, S–R associations allow a movement to be triggered directly by a stimulus, without including a representation of the goal or reward that originally reinforced the movement. Such a learning structure is thought to explain the relative imperviousness of habitual responding to changes in goals or reward value (Dickinson, 1985).

More recently, however, neural correlates of reward value have been found in parts of dorsal striatum, and evidence that these value representations modulate actions has also been reported (Kawagoe et al., 1998; Hassani et al., 2001; Lauwereyns et al., 2002; Haruno et al., 2004; Tricomi et al., 2004; Lau and Glimcher, 2007, 2008; Pasquereau et al., 2007; Hori et al., 2009; Ito and Doya, 2009; Kim et al., 2009; Kimchi and Laubach, 2009). These include representations of the value of the action chosen as well as representations of the value of available actions, present whether or not a particular action is chosen. In addition, lesions of the medial part of the dorsal striatum of rats impair actions based on the value of an expected outcome and cause behavior to become more habitual (Yin and Knowlton, 2004; Yin et al., 2005a,b).

Based on this evidence, it has been proposed that representations of reward value are anatomically segregated from strict S–R associations in the dorsal striatum, with the former being confined to the dorsomedial striatum (in rodents; caudate in primates) (Yin and Knowlton, 2006; Balleine et al., 2007; Johnson et al., 2007). Under this idea, dorsomedial striatum stores response–outcome (R–O) associations to drive goal-directed action selection, whereas the dorsolateral striatum (rodents; putamen in primates) stores the traditional sensorimotor (S–R) associations previously ascribed to all of the striatum.

This proposal makes several predictions. First the hypothesis described above predicts that associations between reward value and movements should be represented in dorsomedial but not dorsolateral striatum, whereas associations between cues and responses, independent of reward value, should be represented in neural activity in dorsolateral but not dorsomedial striatum. Additionally, neural activity in dorsomedial striatum at the time an action is executed might be expected to change rapidly to reflect changes in the value of the associated outcome.

To test these predictions, we recorded single-unit activity in the medial and lateral dorsal striatum of rats engaged in a task in which two distinct movements were associated with four outcomes, in a blocked design. Two different manipulations, reward size and reward delay, were used to vary the value of the outcomes. In addition, each trial was cued by one of three odors, the identity of which indicated which of the two movements would lead to

reward: right, left or either (necessitating a choice between the two). Across all blocks, the same three odors always had the same meanings. As a result of this design, S–R associations remained the same across blocks, while the value of each response – and the particular outcome associated with that response – varied from block to block. This allowed us to dissociate neural correlates of the S–R and R–O associations.

Contrary to our expectations, we found that neural activity in the two regions represented S–R and R–O associations to the same extent. These results are inconsistent with the hypothesis that differences in information content in these two regions account for their differential roles in goal-directed and habitual behavior and instead suggest that these roles may be determined by how these areas interact with their downstream targets.

MATERIALS AND METHODS

SUBJECTS

Male Long-Evans rats were obtained at 175–200 g from Charles River Labs, Wilmington, MA. Rats were tested at the University of Maryland School of Medicine in accordance with SOM and NIH guidelines.

SURGICAL PROCEDURES AND HISTOLOGY

Surgical procedures followed guidelines for aseptic technique. Electrodes were manufactured and implanted as in prior recording experiments. Rats had a drivable bundle of 10 25-μm diameter FeNiCr wires (Stablohm 675, California Fine Wire, Grover Beach, CA, USA) chronically implanted in the left hemisphere in the dorsal-most part of the posterior dorsomedial striatum ($n = 5$; 0.4 mm posterior to bregma, 2.6 mm left of midline, and 3.5 mm ventral to the brain surface) or dorsolateral striatum ($n = 4$; 0.7 mm anterior to bregma, 3.6 mm left of midline, and 3.5 mm ventral to the brain surface). Coordinates were identical to those used to make infusions or lesions in studies that have found functional dissociations between these two regions (Yin et al., 2004, 2005b). Prior to implantation, these wires were freshly cut with surgical scissors to extend ~1 mm beyond the cannula and electroplated with platinum (H_2PtCl_6, Aldrich, Milwaukee, WI, USA) to an impedance of ~300 kΩ. Cephalexin (15 mg/kg p.o.) was administered twice daily for 2 weeks post-operatively to prevent infection. At the end of the study, the final electrode position was marked, the rats were euthanized with an overdose of isoflurane and perfused, and the brains were removed from the skulls and processed using standard techniques.

BEHAVIORAL TASK

Recording was conducted in aluminum chambers approximately 18′ on each side with sloping walls narrowing to an area of 12′ × 12′ at the bottom. A central odor port was located above two adjacent fluid wells on a panel in the right wall of each chamber. Two lights were located above the panel. The odor port was connected to an air flow dilution olfactometer to allow the rapid delivery of olfactory cues. Task events were controlled by computer. Port entry and licking was monitored by disruption of photobeams. Odors where chosen from compounds obtained from International Flavors and Fragrances (New York, NY, USA).

The sequence of events in a trial in each block is illustrated in **Figure 1A**. Trials were signaled by illumination of the panel lights inside the box. When these lights were on, nosepoke into the odor port resulted in delivery of the odor cue for 500 ms to a small hemicylinder located behind this opening. One of three different odors was delivered to the port on each trial. At odor offset, the rat had 3 s to make a response at one of the two fluid wells located below the port. One odor indicated that reward would be available at the left well, a second odor indicated that reward would be available at the right well, and a third odor indicated that reward would be available at either well. Odors were presented in a pseudorandom sequence such that the free-choice odor was presented on 7/20 trials and the left/right odors were presented in equal numbers (±1 over 250 trials). In addition, the same odor could be presented on no more than three consecutive trials.

Once the rats were shaped to perform this basic task, we introduced blocks in which we independently manipulated the size of the reward delivered at a given side and the length of the delay preceding reward delivery. Once the rats were able to maintain accurate responding through these manipulations, we began recording sessions. For recording, one well was randomly designated as short (500 ms) and the other long (1 s) at the start of the session (**Figure 1A**: 1st delay block). In the second block of trials these contingencies were switched (**Figure 1A**: 2nd delay block). The length of the delay under long conditions abided the following algorithm. The side designated as long increased by 1 s every time that side was chosen until it became 3 s. If the rat continued to choose that side, the length of the delay increased by 1 s up to a maximum of 7 s. If the rat chose the side designated as long less than 8 out of the last 10 choice trials then the delay was reduced by 1 s to a minimum of 3 s. The reward delay for long forced-choice trials was yoked to the delay in free-choice trials during these blocks. In the third and fourth blocks, we held the delay preceding reward delivery constant (500 ms) while manipulating the size of the expected reward (**Figure 1A**, 1st and 2nd size blocks). The small reward was a 0.05-ml bolus of 10% sucrose solution. For big reward, an additional bolus was delivered after 500 ms. On the third and fourth block, the side with the preferred reward continued to be alternated from block to block. Across the experiment, the number of trials in each block varied non-systematically around 64 trials (SD = 9.7).

SINGLE-UNIT RECORDING

Procedures were the same as described previously (Roesch et al., 2006, 2007). Wires were screened for activity daily; if no activity was detected, the rat was removed, and the electrode assembly was advanced 40 or 80 μm. Otherwise active wires were selected to be recorded, a session was conducted, and the electrode was advanced at the end of the session. Neural activity was recorded using two identical Plexon Multichannel Acquisition Processor systems (Dallas, TX, USA), interfaced with odor discrimination training chambers. Signals from the electrode wires were amplified 20× by an op-amp headstage (Plexon Inc, HST/8o50-G20-GR), located on the electrode array. Immediately outside the training chamber, the signals were passed through a differential pre-amplifier (Plexon Inc, PBX2/16sp-r-G50/16fp-G50), where the single-unit signals were amplified 50× and filtered at 150–9000 Hz. The single-unit signals were then sent to the Multichannel Acquisition Processor box,

where they were further filtered at 250–8000 Hz, digitized at 40 kHz and amplified at 1–32×. Waveforms (>2.5:1 signal-to-noise) were extracted from active channels and recorded to disk by an associated workstation with event timestamps from the behavior computer. Waveforms were not inverted before data analysis.

DATA ANALYSIS AND FIRING RATE EPOCHS
Units were sorted using Offline Sorter software from Plexon Inc (Dallas, TX, USA), using a template matching algorithm. Sorted files were then processed in Neuroexplorer to extract unit times-tamps and relevant event markers. These data were subsequently analyzed in Matlab (Natick, MA, USA). To analyze neural cor-relates of the movement, we examined firing rate from 50 ms after the presentation of the odor to odor port exist, and also from odor port exit to fluid port entry. We performed ANOVAs ($p < 0.05$) on each neuron's firing rate during each of these two epochs, with factors depending on the variable of interest. To match free-choice and forced trials, for each free-choice trial the most recent forced-choice trial in the same direction and block and the next forced-choice trial in the same direction and same block were averaged together. In this way, free- and forced-choice trials were matched for direction, outcome and position in block.

To represent population activity, we first binned the firing rate of each neuron, from the beginning of each trial to the end of each trial. Then we subtracted the baseline firing rate on each trial, defined as that during the 2 s immediately preceding the start of the trial, from all bins (except for the block analysis, for which we did not subtract baseline activity). Note that the rationale for subtracting the baseline firing rate on each trial was that many neurons showed intra-session variability in their baseline activity. As described in the Results, we separately analyzed these variations and found that they appeared to reflect selectivity that developed for a particular block, or, equivalently, for a particular set of action–outcome relationships. Next, we averaged each bin across trials in each condition (each condition means the direction of movement and the identity of the associated outcome – e.g. big left, big right, small left, small right, etc. For outcome and block analysis, this averaging was done separately for the first 10 trials and last 10 trials of each condition; for the S–R analysis, averaging was done across the entire condition). Then we selected the maximum firing rate in any of these bins on forced-choice trials in any condition, and divided all bins in all conditions by that value (i.e. normalized). We performed this normalization in order to collectively analyze neurons with a wide variety of firing rates.

Selectivity indices (stimulus index, delay index, size index, and block-selectivity index) were calculated for each neuron by taking the difference between the average normalized firing rates dur-ing an epoch between the conditions of interest. To analyze the evolution of baseline firing rate across sessions and blocks, we averaged normalized firing rate during the epoch from light on to odor port exit in each pair of trials across the block, collapsing across preferred block. We included the last 30 trials of the block previous to the preferred block, the first 50 trials of the preferred block, the last 10 trials of the preferred block, and the first 50 tri-als of the block subsequent to the preferred block. For a control comparison, the same data was calculated for the block with the
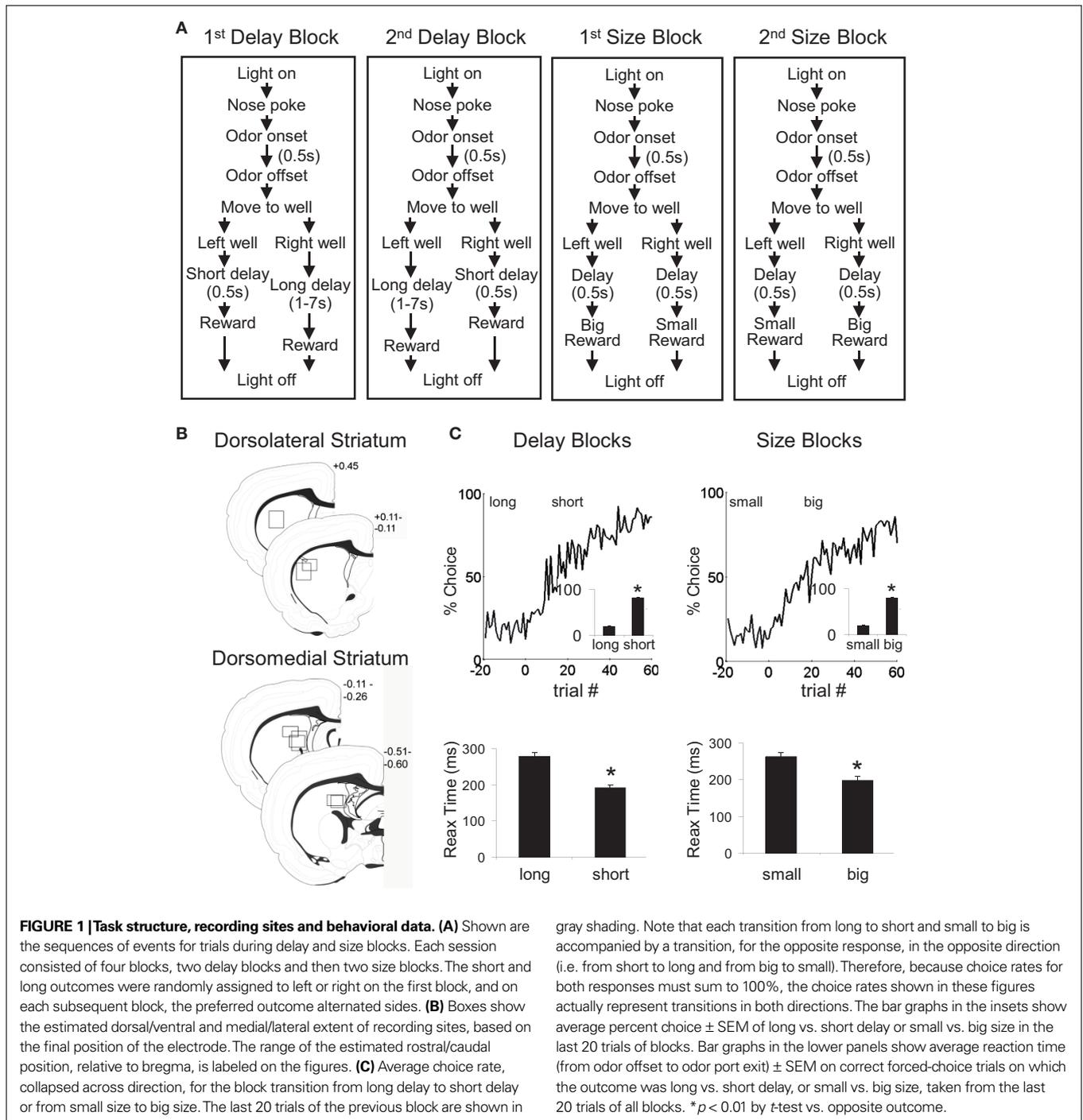
same-value outcomes in the same direction using the other value manipulation. This control block was by definition neither imme-diately before nor immediately after the preferred block. When the preferred block was the first block of the session, it was not included in this analysis. Differences in proportions of neurons in dorsomedial and dorsolateral striatum were tested using Pearson Chi-square tests ($p < 0.05$).

RESULTS
Rats were trained to initiate a trial by nose-poking into a central odor port. After exposure to an odorized air-stream for 0.5 s, they could move down and left to one fluid well or down and right to a second fluid well to receive a reward of sucrose solution. One odor always indicated that reward was available in the left well, while a second odor always indicated that reward was available in the right well ("forced-choice" trials). A third odor indicated that reward was available in either well, necessitating a choice between the two ("free-choice" trials). The size and timing of reward outcomes were manipulated such that, within a block, each movement always led to a particular outcome: either big (two drops), small (one drop), short-delayed (0.5 s), or long-delayed (1–7 s). Thus, across the four blocks in each recording session, each odor-movement (S–R) com-bination was associated with each of the four outcomes. On each side, the outcomes were always presented in an alternating order: high-value, low-value, high-value, low-value (or vice versa). The big reward on one side was paired with the small reward on the other, and the short-delayed reward was paired with the long-delayed reward (see **Figure 1A** for an illustration of the sequence of events of trials in each block).

STIMULUS–RESPONSE AND RESPONSE–OUTCOME CONTINGENCIES MODULATE BEHAVIOR ON FREE- AND FORCED-CHOICE TRIALS
We recorded from dorsolateral striatum in four rats during 74 sessions, and from dorsomedial striatum in five rats in 86 ses-sions. Electrode placements, illustrated in **Figure 1B**, were based on studies that have found a functional dissociation between these two regions (Yin et al., 2004, 2005b). Across all sessions, rats in both groups made the correct response on more than 80% of forced-choice trials, demonstrating that they had accu-rately learned the S–R associations (84% in dorsomedial group, 81% in dorsolateral group, group difference $F_{1,158} = 3.1$, n.s.). Additionally, rats' free- and forced-choice behavior rapidly adapted to the changing R–O contingencies within each block (**Figure 1C**); in the final 20 trials of blocks, rats chose the response associated with short-delayed reward 81% of the time, and that associated with big reward 80% of the time. An ANOVA of these choice rates showed that there was no effect of record-ing group or value manipulation, demonstrating that short and big reward had similar relative values in both groups (recording group $F_{1,158} = 2.0$, n.s.; manipulation $F_{1,158} = 0.4$, n.s.). In addition, by the final 20 forced-choice trials, rats reacted more quickly for the higher value outcome (main effect of value on reaction time, $F_{1,158} = 151$, $p < 0.001$). This effect was present within each value manipulation in each recording group (F's > 13.6, p's < 0.001). Thus, on both free-choice and forced-choice trials, behavior reflected awareness of the S–R and R–O contingencies similarly in the two recording groups.

FIGURE 1 | Task structure, recording sites and behavioral data. (A) Shown are the sequences of events for trials during delay and size blocks. Each session consisted of four blocks, two delay blocks and then two size blocks. The short and long outcomes were randomly assigned to left or right on the first block, and on each subsequent block, the preferred outcome alternated sides. **(B)** Boxes show the estimated dorsal/ventral and medial/lateral extent of recording sites, based on the final position of the electrode. The range of the estimated rostral/caudal position, relative to bregma, is labeled on the figures. **(C)** Average choice rate, collapsed across direction, for the block transition from long delay to short delay or from small size to big size. The last 20 trials of the previous block are shown in gray shading. Note that each transition from long to short and small to big is accompanied by a transition, for the opposite response, in the opposite direction (i.e. from short to long and from big to small). Therefore, because choice rates for both responses must sum to 100%, the choice rates shown in these figures actually represent transitions in both directions. The bar graphs in the insets show average percent choice ± SEM of long vs. short delay or small vs. big size in the last 20 trials of blocks. Bar graphs in the lower panels show average reaction time (from odor offset to odor port exit) ± SEM on correct forced-choice trials on which the outcome was long vs. short delay, or small vs. big size, taken from the last 20 trials of all blocks. *$p < 0.01$ by t-test vs. opposite outcome.

RESPONSE ENCODING

We recorded from a total of 489 neurons in dorsolateral striatum and 587 neurons in dorsomedial striatum. The majority of neurons in both regions had low baseline firing rates (median average baseline firing rate in dorsolateral striatum was 2.5 spikes/s, in dorsomedial striatum, 2.6 spikes/s). We classified neurons as either fast firing or phasically active neurons according to established methods (see Supplementary Material) (Schmitzer-Torbert and Redish, 2008). We did not isolate any neurons with characteristics of tonically active neurons. The pattern of results reported below did not differ between fast firing and phasically active neurons, and so we included all recorded neurons and normalized their firing rates for analysis (see Materials and Methods).

Because we were interested in striatal activity that might influence choices in this task, we defined two epochs that encompassed the time during which the choice must have been made. The first, the odor epoch, began 50 ms after the initial presentation of the odor and ended with the rat's withdrawal from the odor-port. The second, the movement epoch, began with the rat's withdrawal from the odor-port and ended with its entry into one of the two fluid

wells. Note that the odor epoch begins at the earliest point at which an odor-guided choice (left or right) could be made in this task, and the movement epoch begins as that choice begins to manifest itself. Thus neural activity relevant to making or driving the choice behavior must occur during one or both of these two epochs. For each epoch, we defined response-selective neurons as those which showed a significant effect of direction on correct forced-choice trials; that is, such neurons would be selective for either a left or a right movement. By using only forced-choice trials for this analysis, we were able to analyze equal numbers of trials in which each direction was associated with each outcome.

In both dorsolateral and dorsomedial striatum, we found a large proportion of neurons that showed response selectivity for at least one of the four outcomes. Thus these neurons fired significantly more during either the odor period, the response period, or both when the rat moved (or subsequently moved) in one direction versus the other. In dorsolateral striatum, these populations included 147 neurons (30% of all neurons) during the odor epoch and 237 (48% of all neurons) during the movement epoch. Of these, 77 (16% of all neurons) were selective during both epochs. In dorsomedial striatum, these populations included 193 neurons (33% of all neurons) during the odor epoch and 269 (46% of all neurons) during the movement epoch. Of these, 120 (20% of all neurons) were selective during both epochs. The preferred direction of these neurons, defined as the direction in which the highest firing rate occurred, was similar in both areas in both epochs, and there was no strong laterality (dorsolateral: 51% right-preferring during the odor epoch and 64% right-preferring during the movement epoch; dorsomedial: 52% right-preferring during the odor epoch and 56% right-preferring during the movement epoch). In addition, a large proportion of directionally selective neurons showed a significant inhibition of activity prior to and during movement in that neuron's non-preferred direction. An analysis of this activity is shown in Supplementary Material.

Since odor identity was confounded with direction of subsequent movement on forced-choice trials, differential firing in the response-selective populations identified above could have reflected either odor identity or movement direction. The remaining analyses (except for those in the final section of the Results) were carried out on these populations in order to determine which aspect of the response or odor was represented by this activity.

STIMULUS–RESPONSE ENCODING

Within striatum, the dorsolateral region is particularly critical to habitual responding, which is thought to reflect stimulus–response (S–R) associations. If this is due to a special role in encoding S–R associations, then the response-related firing in dorsolateral striatum should be particularly dependent on the stimulus that instructs a particular movement. To test for such encoding, we compared activity of neurons in the previously identified response-selective populations on forced-choice trials with that on matched free-choice trials, which differ in the odor that initiates them. Trials were matched such that they involved the same response for the same outcome and occurred in a similar position within the block. Thus the only obvious factor that differed between them was the identity of the odor cue. Note that this comparison is appropriate for detecting S–R encoding also because free-choice and forced-choice trials

differed in the history of the association between the stimulus and the response. That is, on forced-choice trials, the same odor always signaled the same response across all blocks, whereas on free-choice trials, the odor signaled that a different response should be preferentially made on each block. Thus, activity that allows a mapping of the stimulus to the response based on the learned relationship between the two, as is postulated to occur in S–R encoding, would tend to distinguish these two conditions.

Consistent with the proposal that dorsolateral striatum signals S–R associations, 18 of 147 (12%) response-selective neurons during the odor epoch, and 47 of 237 (20%) response-selective neurons during the movement epoch in dorsolateral striatum exhibited significantly differential firing between these two trial types. Note that these putative S–R encoding neurons were required to show selectivity across all blocks of the session, which means that they signaled a particular S–R conjunction regardless of outcome. Such a pattern is consistent with theoretical accounts of S–R encoding. As illustrated by the examples in **Figures 2A,B** and the population analyses in **Figures 3A and 4A**, these populations included neurons that fired more on forced-choice trials and also neurons that fired more on free-choice trials, suggesting that both kinds of S–R associations were represented (13 of 18 neurons preferred free-choice trials during the odor epoch, and 20 of 47 did so during the response epoch).

However, similar S–R correlates were found in equal or greater numbers in dorsomedial as in dorsolateral striatum. This is evident in the example units in **Figures 2C,D**, and in **Figures 3B and 4B**, which show that response-selective neurons in dorsomedial striatum also exhibited differential firing on response- and outcome-matched forced and free-choice trials. Indeed although the proportion of neurons with differential activity during the movement epoch (54 of 269, or 20%) did not differ from that in dorsolateral striatum (n.s. by Chi-square test), the proportion during the odor epoch (43 of 193, or 22%) was significantly greater than in dorsolateral striatum ($p < 0.05$ by Chi-square test). These neurons were also more likely to prefer the forced-choice trials during the odor epoch (34 of 43 neurons, $p < 0.001$ by Chi-square test), though this was not true during the movement epoch (33 of 54, n.s. by Chi-square test). Overall however the differences between the two regions in S–R encoding were relatively minimal; the mean free-choice/forced-choice selectivity index (see Materials and Methods) of each of these populations (free-choice preferring and forced-choice preferring during each epoch) did not differ significantly between dorsolateral and dorsomedial striatum (see **Table 1**).

In order to demonstrate that this putative S–R encoding did not represent simple odor encoding, we also calculated a directional selectivity index using free-choice trials, during which the initiating odor is the same but the direction of the response differs, for each putative S–R neuron identified above. The sign of this index was based on the corresponding directional selectivity on forced-choice trials, meaning that a positive free-choice index indicated the same direction of selectivity on free-choice as on forced-choice trials. In dorsolateral striatum, the putative S–R population (both free-choice preferring and forced-choice preferring) identified above during the odor epoch had a mean free-choice directional selectivity index of 0.034 ± 0.038, which is not
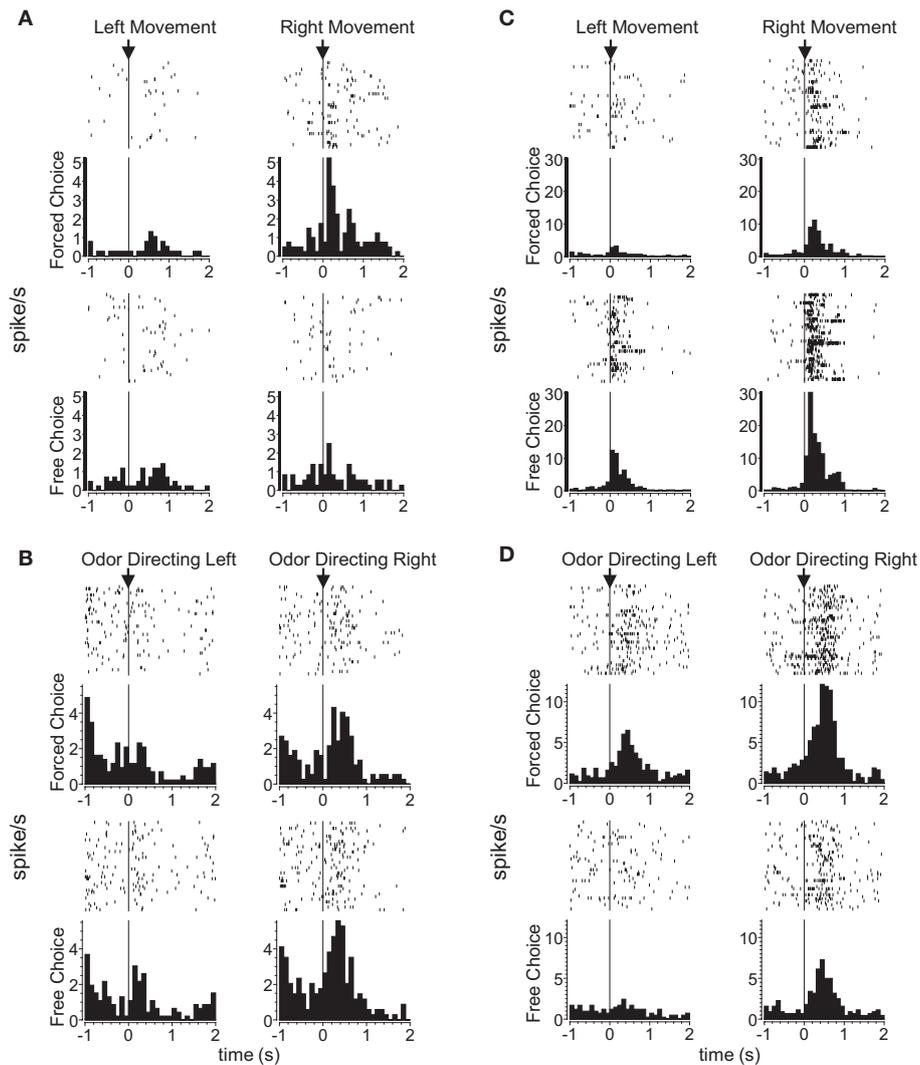
FIGURE 2 | Examples of free-choice/forced-choice-selective single-units recorded from dorsolateral and dorsomedial striatum. Shown are raster plots and time histograms displaying firing rate during correct forced-choice vs. free-choice trials, aligned on the beginning of movement for **(A)** and **(C)**, and at the presentation of the odor for **(B)** and **(D)**. **(A)** and **(B)** are units from dorsolateral striatum; **(C)** and **(D)** are units from dorsomedial striatum. All neurons sh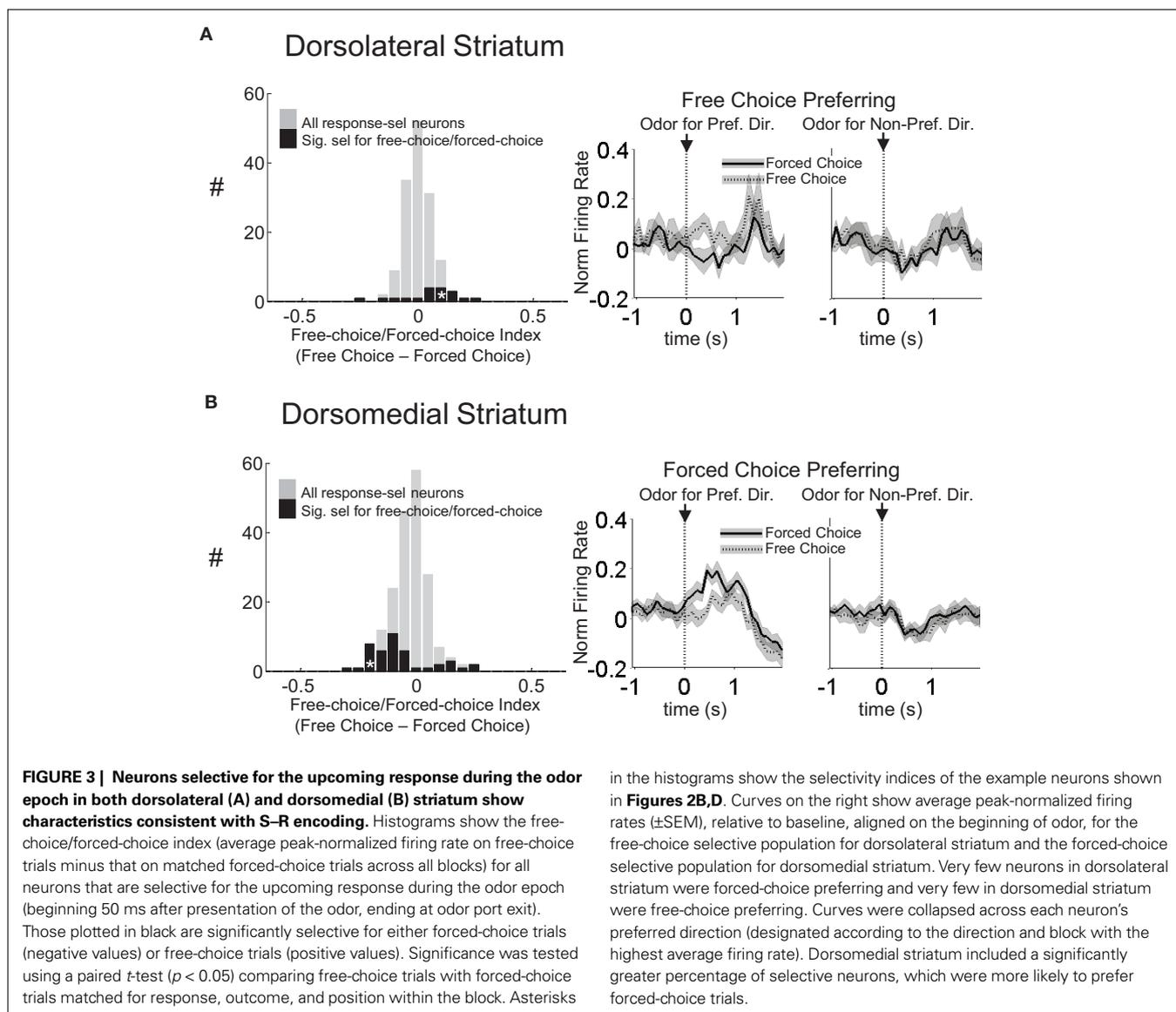own here were significantly selective for the direction of movement that was executed on that trial, in these cases for the rightward movement. Other neurons (the minority) were selective for the leftward movement. Neurons shown in **(A)** and **(D)** were significantly selective for forced-choice trials, while those shown in **(B)** and **(C)** were significantly selective for free-choice trials. For forced-choice trials, only the last ~80 trials are shown so that the raster plots are comparable in size to the free-choice raster plots.

significantly greater than zero ($t_{17} = 0.90$, n.s.). However, the mean free-choice directional selectivity index increased to 0.13 ± 0.023 during the response epoch, which was significantly greater than zero ($t_{46} = 5.7$; $p < 0.001$). The corresponding populations in dorsomedial striatum had mean free-choice directional selectivity indices of 0.096 ± 0.031 during the odor period and 0.19 ± 0.027 during the movement epoch, both of which were significantly greater than zero ($t_{42} = 3.1$; $p < 0.01$; $t_{53} = 7.2$; $p < 0.001$). Thus activity in the putative S–R population was both odor-selective and response-selective, meaning that it responded to a particular S–R conjunction. This was true in both dorsomedial and dorsolateral striatum during the response epoch, but was only true in dorsomedial striatum during the critical odor-sampling period.

Finally, we tested whether putative S–R encoding might be a consequence of differences in reaction time between free- and forced-choice trials. In both areas during both epochs, both free-choice-preferring and forced-choice-preferring neurons were just as likely to be recorded during sessions in which rats responded more quickly on free-choice trials as when they responded more slowly on these trials (n.s. by Chi-square test).

RESPONSE–OUTCOME ENCODING

Within striatum, the dorsomedial region is particularly critical to responding guided by outcome value, which is thought to reflect response–outcome (R–O) associations. If this is due to a special role in encoding R–O associations, then the

FIGURE 3 | Neurons selective for the upcoming response during the odor epoch in both dorsolateral (A) and dorsomedial (B) striatum show characteristics consistent with S–R encoding. Histograms show the free-choice/forced-choice index (average peak-normalized firing rate on free-choice trials minus that on matched forced-choice trials across all blocks) for all neurons that are selective for the upcoming response during the odor epoch (beginning 50 ms after presentation of the odor, ending at odor port exit). Those plotted in black are significantly selective for either forced-choice trials (negative values) or free-choice trials (positive values). Significance was tested using a paired *t*-test ($p < 0.05$) comparing free-choice trials with forced-choice trials matched for response, outcome, and position within the block. Asterisks in the histograms show the selectivity indices of the example neurons shown in **Figures 2B,D**. Curves on the right show average peak-normalized firing rates (±SEM), relative to baseline, aligned on the beginning of odor, for the free-choice selective population for dorsolateral striatum and the forced-choice selective population for dorsomedial striatum. Very few neurons in dorsolateral striatum were forced-choice preferring and very few in dorsomedial striatum were free-choice preferring. Curves were collapsed across each neuron's preferred direction (designated according to the direction and block with the highest average firing rate). Dorsomedial striatum included a significantly greater percentage of selective neurons, which were more likely to prefer forced-choice trials.

response-related firing in dorsomedial striatum should be particularly dependent on the features – particularly the value – of the predicted outcome.

To test this, we compared activity in each epoch on trials involving the four outcome types delivered in our task (i.e. big, small, short-delayed, or long-delayed). Consistent with the proposal that dorsomedial striatum signals R–O associations, many response-selective neurons showed an enhanced response when a particular outcome could be expected to occur in that neuron's preferred direction. An example is shown in **Figure 5A**; this dorsomedial neuron shows a consistent preference for the rightward response with the greatest firing rate on the block in which the short-delayed reward is associated with that response. However, similar R–O correlates were also present in dorsolateral striatum. This is illustrated by the example in **Figure 5B**; this dorsolateral neuron shows a consistent preference for the rightward response with the greatest firing rate when the long-delayed reward is associated with that response. Such correlates were present across the entire population

of response-selective neurons in both dorsomedial and dorsolateral striatum during both the odor epoch and the movement epoch, as shown in the populations responses in **Figures 6 and 7**. Notably, response-selective neurons identified in both epochs tended to fire the most when a particular outcome could be expected to occur in their preferred direction without distinguishing between the other three possible outcomes. This was true in both regions.

Thus, neurons in both dorsomedial and dorsolateral striatum appear to encode the association between a response and a particular outcome. To quantify this, we analyzed the difference in firing rate for the preferred response when it was associated with the high value (big or short-delayed) vs. the low-value outcome (small or long-delayed), for each of the two manipulations. We performed this analysis separately for the odor epoch and the movement epoch and found similar results. During the odor period, encoding of the upcoming response was modulated by the value for at least one of the value manipulations (size or delay) in 56 of 147 response-selective neurons (38%) in dorsolateral striatum
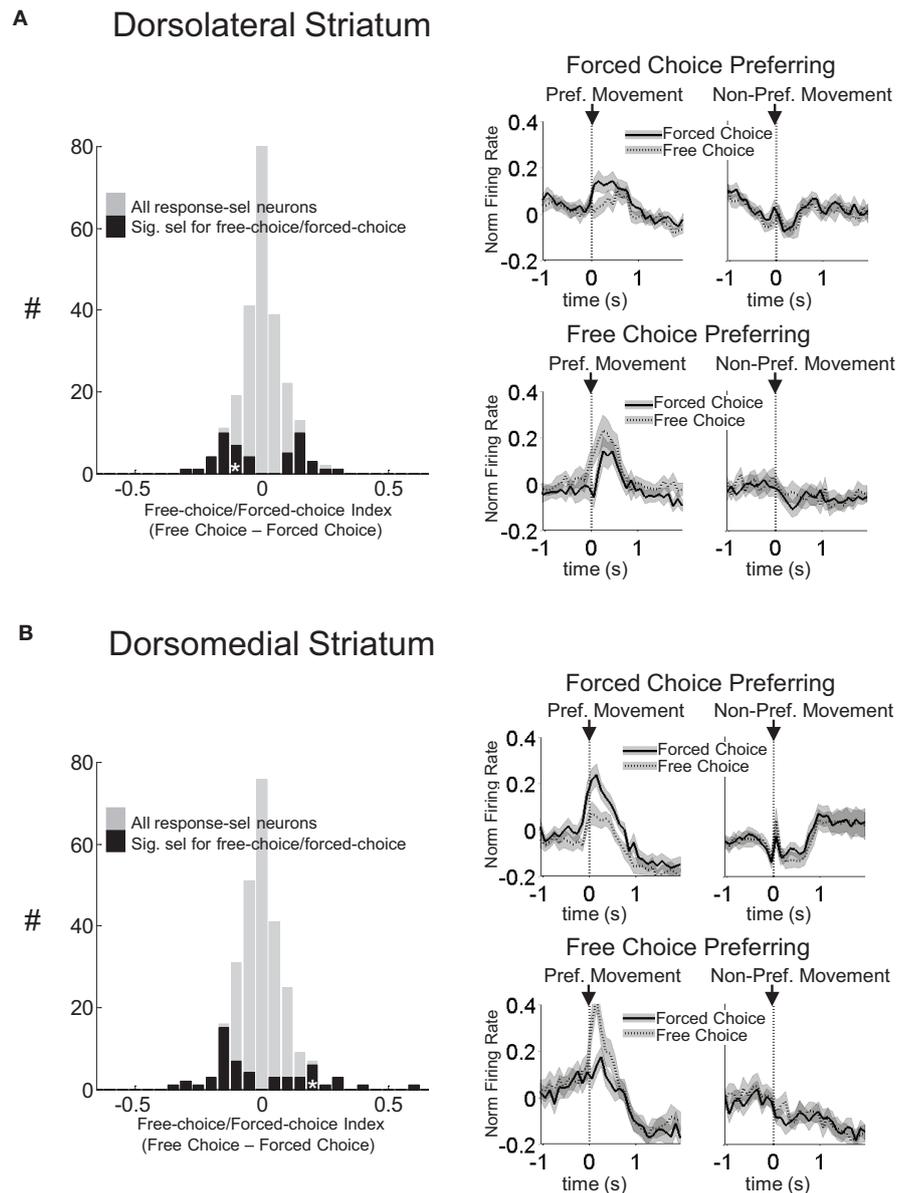
FIGURE 4 | Neurons selective for the response during the movement epoch in both dorsolateral (A) and dorsomedial (B) striatum show characteristics consistent with S–R encoding. Histograms on the left show the free-choice/forced-choice index (average peak-normalized firing rate on free-choice trials minus that on matched forced-choice trials across all blocks) for all response-selective neurons. Those plotted in black are significantly selective for either forced-choice trials (negative values) or free-choice trials (positive values). Significance was tested using a paired *t*-test (*p* < 0.05) comparing free-choice trials with forced-choice trials matched for response, outcome, and position within the block. Asterisks in the histograms show the selectivity indices of the example neurons shown in **Figures 2A,C**. Curves on the right show average peak-normalized firing rates (±SEM), relative to baseline, aligned on the beginning of movement towards reward, for each significantly selective population. Curves were collapsed across each neuron's preferred direction (designated according to the direction and block with the highest average firing rate).

and 68 of 193 response-selective neurons in dorsomedial striatum (35%). These proportions did not differ from each other by Chi-square test. During the movement period, encoding of the upcoming response was modulated by value in 74 of 237 response-selective neurons (31%) in dorsolateral striatum, and 87 of 269 response-selective neurons (32%) in dorsomedial striatum. These proportions were also not significantly different from each other (Chi-square test, n s.).

A comparison of delay- and size-encoding, presented in **Figures 6C,D and 7C,D**, illustrates that the neural populations representing delay and size were largely non-overlapping. In other words, neurons that were selective for reward delay in a particular direction were not similarly selective for reward size in that direction, and vice versa. This is evident in the bimodal distributions of neurons with significant outcome modulation of the firing in their preferred direction, represented by the colored points in **Figures 6 and 7**. Neurons

Table 1 | Shown are the mean selectivity indices (±SEM) for each of the analyses reported in the paper for dorsomedial and dorsolateral striatum, and the results of the *t*-test comparing these means. Only for the block-selectivity index was there a significant difference between the two areas.

| Epoch	Area	Stimulus index (S–R Analysis)		Outcome index (R–O Analysis)				Block-selectivity index (collapsed across all block preferences)
		Forced-choice preferring	Free-choice preferring	Big preferring	Small preferring	Short preferring	Long preferring	
Odor	Dorsomedial striatum	−0.13 ± 0.01	0.16 ± 0.02	0.31 ± 0.03	0.31 ± 0.04	0.35 ± 0.03	0.27 ± 0.03	0.15 ± 0.01
	Dorsolateral striatum	−0.12 ± 0.04	0.12 ± 0.02	0.34 ± 0.03	0.29 ± 0.03	0.32 ± 0.02	0.32 ± 0.04	0.11 ± 0.009
	Comparison (by *t*-test)	$t_{37} = -0.4$ $p = 0.71$	$t_{20} = 1.6$ $p = 0.12$	$t_{30} = 0.7$ $p = 0.48$	$t_{28} = 0.2$ $p = 0.82$	$t_{44} = 0.9$ $p = 0.40$	$t_{23} = 1.3$ $p = 0.22$	$t_{209} = 3.1$ $p = 0.002$
Movement	Dorsomedial striatum	−0.15 ± 0.01	0.20 ± 0.03	0.39 ± 0.03	0.34 ± 0.02	0.33 ± 0.02	0.36 ± 0.03	
	Dorsolateral striatum	−0.14 ± 0.01	0.16 ± 0.01	0.32 ± 0.04	0.37 ± 0.03	0.33 ± 0.04	0.35 ± 0.03	
	Comparison (by *t*-test)	$t_{58} = -0.8$ $p = 0.44$	$t_{39} = 1.5$ $p = 0.14$	$t_{37} = 1.4$ $p = 0.17$	$t_{43} = 0.8$ $p = 0.41$	$t_{42} = 0.1$ $p = 0.92$	$t_{45} = 0.8$ $p = 0.71$	
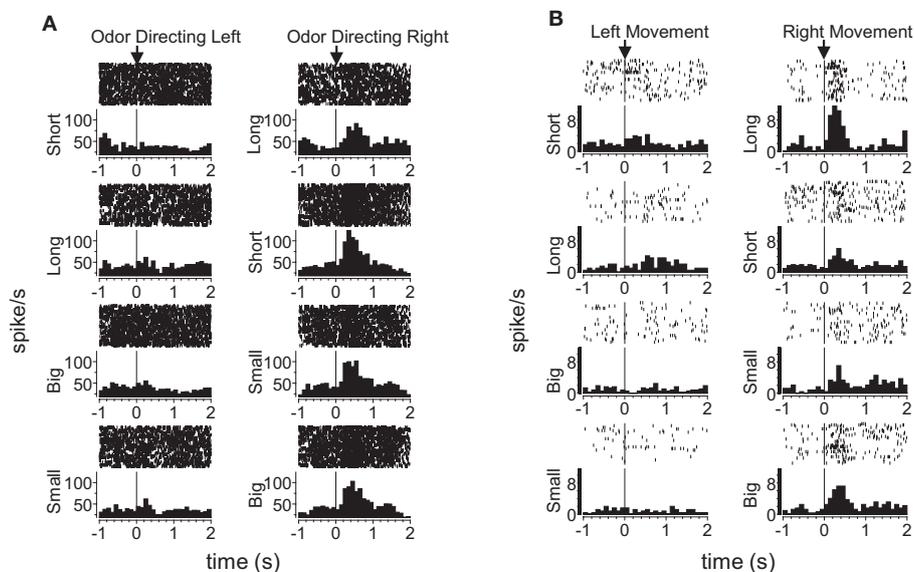
FIGURE 5 | Examples of outcome-selective single-units recorded from dorsomedial and dorsolateral striatum. Shown are raster plots and time histograms displaying firing rate during correct forced-choice trials, aligned on the beginning of odor presentation in **(A)** and on the beginning of movement towards reward in **(B)**. These two units, from dorsomedial **(A)** and dorsolateral **(B)** striatum, are both response-selective and show the greatest activity when a particular outcome is expected in their preferred direction. The unit in **(A)** responded most when the short-delayed outcome could be expected to result from the rightward movement, and the unit in **(B)** responded most when the long-delayed outcome could be expected to result from the rightward movement.

were no more likely to be significantly selective for both manipulations, as represented by the yellow points in **Figures 6 and 7**, than would be expected by chance alone. In fact, during the odor epoch, only 2 of 56 outcome-modulated neurons in dorsolateral striatum (versus 3 expected by chance, n.s. by Chi-square test) and 1 of 68 neurons in dorsomedial striatum (versus 3.6 expected by chance, n.s. by Chi-square test) preferred both high-value outcomes or both low-value outcomes.. The corresponding numbers during the movement epoch were 4 of 74 outcome-modulated neurons in dorsolateral striatum (versus 3.3 expected by chance, n.s. by Chi-square test) and 3 of 87 outcome-modulated neurons in dorsomedial striatum (versus 4.2 expected by chance, n.s. by Chi-square test).
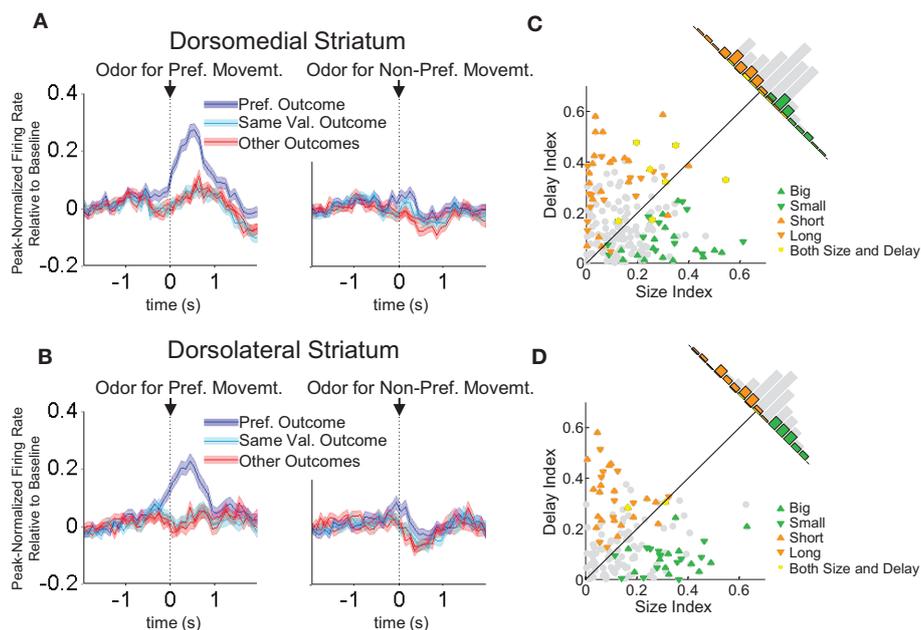
FIGURE 6 | Neurons selective for the upcoming response during the odor epoch are modulated by outcome in both dorsomedial and dorsolateral striatum. Curves in **(A)** and **(B)** show average peak-normalized firing rates (±SEM), relative to baseline, during the last 20 forced-choice trials of each block, aligned on the beginning of odor presentation. Populations included all neurons selective for the upcoming response during the odor epoch (193 out of 587 recorded in dorsomedial; 147 out of 489 recorded in dorsolateral). Curves were collapsed across each neuron's preferred direction and preferred outcome (designated according to the direction and block with the highest average firing rate). Preferred outcomes were equally distributed across the four outcomes. Scatter plots in **(C)** and **(D)** show the delay modulation index vs. the size modulation index for each response-selective neuron. Colored points indicate neurons that were significantly selective for the size modulation, the delay modulation, or both. Bar graphs show the difference between the two indices for each neuron. To the extent that outcome-modulated responses reflect the value of the response, colored points should congregate around the diagonal, the colored bars should peak in the center, and the number of neurons significantly modulated by both manipulations should exceed chance. In fact, however, in both regions colored points are significantly removed from the diagonal and neurons modulated by both manipulations are no more frequent than chance. Thus separate populations of neurons encode each response–outcome conjunction. Delay modulation index = absolute value of the difference between normalized firing rates during preferred directional response on delay block 1 and delay block 2. Size modulation index is the corresponding difference for size blocks.

We further addressed the question of whether delay-selective neurons were selective for the reward size manipulation, and vice versa, by calculating the mean delay-selective index for each size-selective population, and the mean size-selective index for each delay-selective population. As shown in **Table S1** in Supplementary Material, once we corrected for multiple comparisons, none of these means differed from zero. Thus none of the outcome-selective populations from the two epochs and brain areas were selective for outcome based on value. Consistent with this finding, the outcome-modulated populations did not seem to signal value even within the particular manipulation that drove the differential activity; equal numbers of neurons fired to the high and the low value within each value manipulation (see **Table 2** for the complete numerical breakdown). Further, a comparison of the magnitude of selectivity indices for each of the four outcome-selective populations (collapsed across direction) between dorsolateral and dorsomedial striatum revealed no significant differences (see **Table 1**). Thus in both dorsolateral striatum and dorsomedial striatum during both the odor epoch and the movement epoch, outcome-selective populations were divided evenly between those selective for each of the four outcomes – high and low-value outcomes for each value manipulation. Because these neurons were also response-selective, this pattern suggests that

these neurons signaled a particular response when it predicted a particular idiosyncratic outcome rather than signaling the relative value of a particular response.

STIMULUS–RESPONSE–OUTCOME ENCODING

Given the predominance of S–R and R–O encoding in both dorsomedial and dorsolateral striatum, one might expect that at least some neurons in these areas would encode S–R–O conjunctions. We looked for such neurons in two ways. First, we examined the overlap between putative S–R populations and R–O populations. Close to 40% of all S–R neurons identified above turned out to be either size- or delay-selective. In dorsomedial striatum, 17 of 43 (40%) S–R neurons identified during the odor epoch were outcome-selective; 20 of 54 (37%) identified during the movement epoch were outcome-selective. In dorsolateral striatum, 6 of 18 (33%) identified during the odor epoch were outcome-selective; 18 of 47 (38%) identified during the movement epoch were outcome-selective. These percentages were not different than would be expected by chance given the proportions of S–R and R–O in the response-selective population (n.s. by Chi-square test). Note that the presence of outcome encoding in S–R populations does not mean that S–R encoding *depended* on the outcome that was
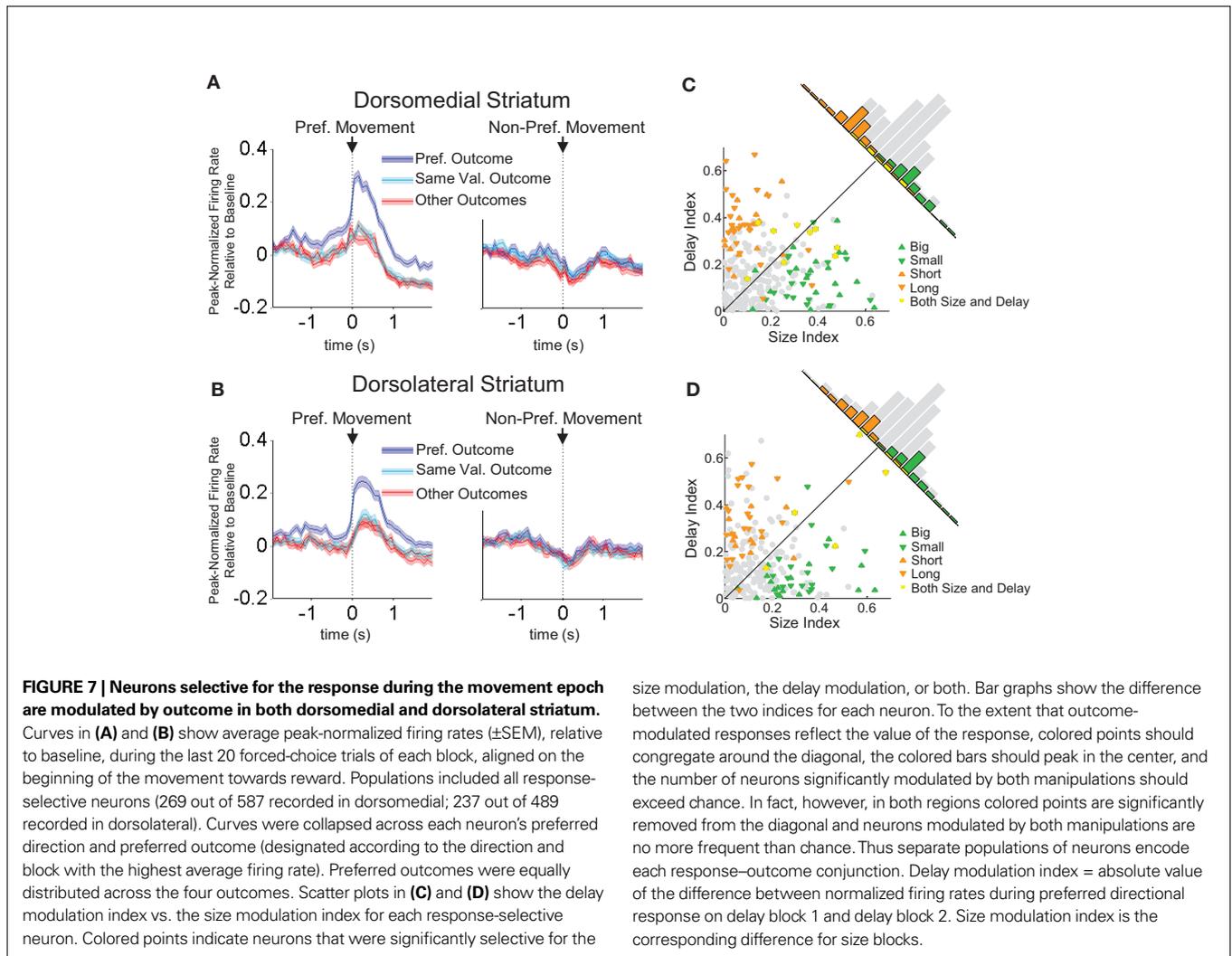
FIGURE 7 | Neurons selective for the response during the movement epoch are modulated by outcome in both dorsomedial and dorsolateral striatum. Curves in **(A)** and **(B)** show average peak-normalized firing rates (±SEM), relative to baseline, during the last 20 forced-choice trials of each block, aligned on the beginning of the movement towards reward. Populations included all response-selective neurons (269 out of 587 recorded in dorsomedial; 237 out of 489 recorded in dorsolateral). Curves were collapsed across each neuron's preferred direction and preferred outcome (designated according to the direction and block with the highest average firing rate). Preferred outcomes were equally distributed across the four outcomes. Scatter plots in **(C)** and **(D)** show the delay modulation index vs. the size modulation index for each response-selective neuron. Colored points indicate neurons that were significantly selective for the size modulation, the delay modulation, or both. Bar graphs show the difference between the two indices for each neuron. To the extent that outcome-modulated responses reflect the value of the response, colored points should congregate around the diagonal, the colored bars should peak in the center, and the number of neurons significantly modulated by both manipulations should exceed chance. In fact, however, in both regions colored points are significantly removed from the diagonal and neurons modulated by both manipulations are no more frequent than chance. Thus separate populations of neurons encode each response–outcome conjunction. Delay modulation index = absolute value of the difference between normalized firing rates during preferred directional response on delay block 1 and delay block 2. Size modulation index is the corresponding difference for size blocks.

Table 2 | Shown are the numbers of response-selective neurons that were significantly selective for each of the four outcomes, or for two of the outcomes. "Same-value-preferring" refers to neurons that preferred both high-value outcomes (big and short) or both low-value outcomes (small and long). "Opposite-value-preferring" refers to neurons that preferred a high-value outcome (big or short) in one manipulation and a low-value outcome (small or long) in the other. Neurons preferring two same-value outcomes were no more frequent than predicted by chance (by Chi-square test, $p < 0.01$).

Epoch	Area	Size-modulated		Delay-modulated		Both size- and delay-modulated		Total outcome-modulated (% of total response-selective neurons)
		Big-pref.	Small-pref.	Short-pref.	Long-pref.	Same-value-pref.	Opp-value-pref.	
Odor	Dorsomedial striatum	18	16	26	15	1	6	68 (35%)
	Dorsolateral striatum	14	14	20	10	2	0	56 (38%)
Movement	Dorsomedial striatum	20	25	26	25	3	6	87 (32%)
	Dorsolateral striatum	19	29	18	22	4	1	74 (31%)

available. In fact, S–R encoding was consistent across all blocks. Rather, it suggests that outcome-selectivity rode on top of S–R selectivity in many neurons.

Secondly, we compared activity of the response-selective populations during forced-choice trials with that during matched free-choice trials, collapsing across preferred outcome just as in

Figures 6 and 7. In this way, we sought to determine whether the outcome-selectivity that was present during forced-choice trials also depended on the preceding stimulus. Indeed, as shown in **Figure 8**, outcome-selectivity that was apparent in the population-averaged activity on forced-choice trials largely disappeared in matched free-choice trials. Note that because rats made very few choices
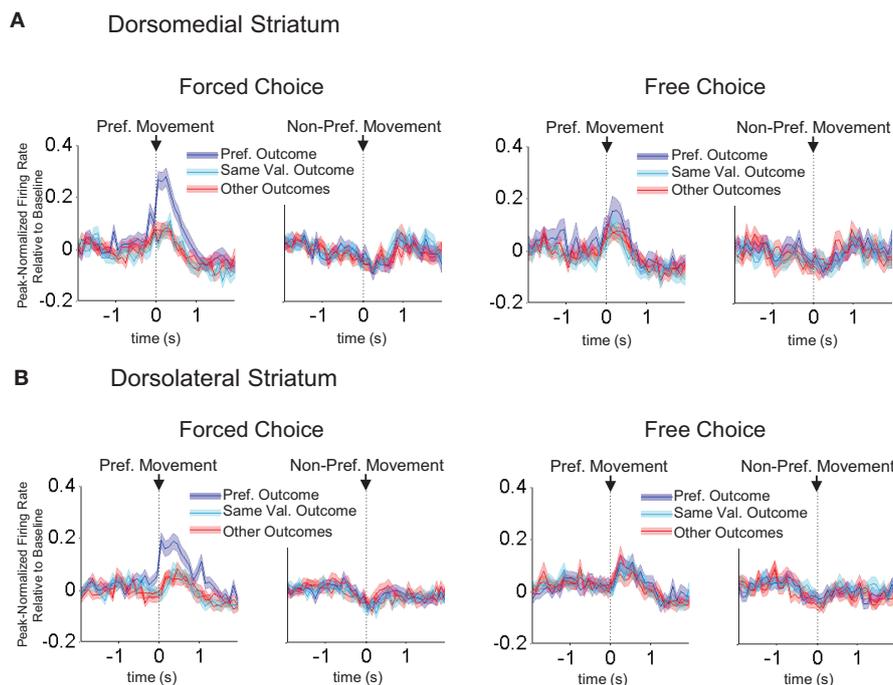
FIGURE 8 | Outcome-selectivity in both dorsomedial and dorsolateral striatum depends on the preceding stimulus. Curves in **(A)** and **(B)** show average peak-normalized firing rates (±SEM), relative to baseline, on matched forced-choice and free-choice trials, aligned on the beginning of the movement towards reward. Curves were collapsed across each neuron's preferred direction and preferred outcome (designated according to the direction and block with the highest average firing rate on forced-choice trials). Only the latter half of free-choice trials in each block, along with forced-choice trials matched for direction, outcome and position within the block, are included. Only response-selective neurons from sessions in which all conditions had at least two free-choice trials are included, resulting in 87 neurons in dorsomedial striatum and 122 neurons in dorsolateral striatum.

of the low-value outcome after they had learned the R–O contingencies, we had to exclude many sessions in which there were insufficient free-choice trials. For this reason we were also unable to perform a neuron-by-neuron analysis of outcome-selectivity on free-choice trials. However, in the sessions with enough free-choice trials, the average outcome selectivity index (collapsed across outcome) on free-choice trials was 0.069 ± 0.031 in dorsomedial striatum, which was significantly less than the 0.21 ± 0.017 during forced-choice trials in these same neurons ($p < 0.001$ by t-test, $t_{172} = 4.0$). Similarly, the outcome selectivity on free-choice trials in dorsolateral striatum was 0.026 ± 0.024, significantly less than the 0.21 ± 0.017 during forced-choice trials in those same neurons ($p < 0.001$ by t-test, $t_{242} = 6.4$). Thus, in both regions, activity to a great extent reflected S–R–O associations.

ENCODING OF THE AVAILABLE RESPONSE–OUTCOME ASSOCIATION

The response–outcome correlates described above occur as the response is being made and depend on the direction of that response. A different kind of outcome encoding, which has been called "action-value" encoding, has also been reported to occur in primate striatal neurons, in which an available outcome (or its value) is encoded regardless of whether the associated response is chosen (Kawagoe et al., 1998; Lauwereyns et al., 2002; Samejima et al., 2005; Lau and Glimcher, 2007, 2008; Ito and Doya, 2009; Kim et al., 2009). This neural correlate can occur before the response is taken, and therefore it could be used to drive response selection. In the present task,

available action-values remain constant during each block but vary between blocks. Therefore, in order to detect action-value correlates, we employed a two-way ANOVA with block and direction as factors. We analyzed firing within a pre-response epoch, which extended from the beginning of the trial to the beginning of the response. We looked for neurons whose firing rate showed a significant effect of block, but did not depend on the direction of the response that was made on that trial. We found 99 of 489 neurons in dorsolateral striatum (20%) and 112 of 587 neurons in dorsomedial striatum (19%) met these criteria. Note that because the only factors that changed systematically between blocks were the action-values and, relatedly, the action–outcome contingencies, these block-selective neurons were by definition responsive either to action-values or to action–outcomes. Example neurons of this type, shown in **Figures 9A,B**, and the population responses, shown in **Figures 10A,B**, illustrate that these neurons tended to show an elevated baseline firing rate in one particular block, rather than a phasic response during the trial. Thus, the baseline firing rate in these neurons was higher when particular response–outcome combinations were available in a block, irrespective of which response was actually chosen on a particular trial. Furthermore, like the R–O correlates described above, this shift seemed to be driven by the identity of available outcomes rather than their general value. Thus, in the population responses in **Figure 10**, the block in which the same-valued outcomes were available in the same directions as in the preferred block did not show an elevated baseline firing rate.
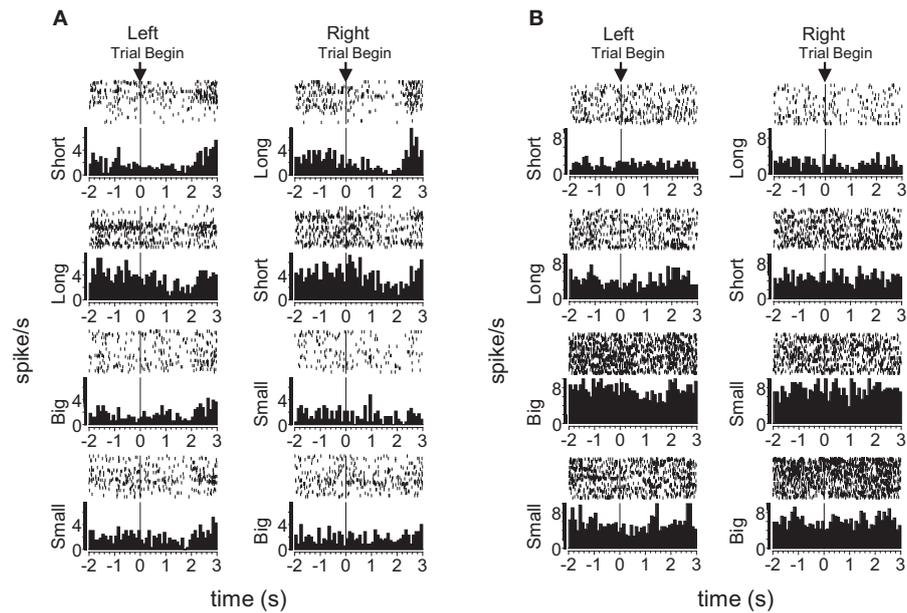
FIGURE 9 | Examples of block-selective single-units recorded from dorsolateral and dorsomedial striatum. Shown are raster plots and time histograms displaying firing rate during correct forced-choice trials, aligned on the beginning of the trial. Each row includes the trials from one block of the session. The unit shown in **(A)**, from dorsolateral striatum, shifted its baseline firing rate in the block with long-delayed outcomes on the left and short-delayed outcomes on the right. The unit shown in **(B)**, from dorsomedial striatum, shifted its baseline firing rate in the block with big outcomes on the left and small outcomes on the right. Blocks are shown in the temporal order in which they occurred.

As shown in **Table 3**, elevated activity was distributed evenly among the four kinds of blocks in both dorsomedial and dorsolateral striatal populations, and the percentage of all neurons that showed these correlates did not differ between dorsolateral and dorsomedial striatum (Chi-square test, ns). However, when we calculated a block-selectivity index (see Materials and Methods) for each neuron, we found that the mean index was slightly larger in dorsomedial than in dorsolateral striatum (see **Table 1**).

To further test whether this shift in baseline firing rate actually reflected the outcomes that were available during that block, we calculated how the shift developed across the preferred block. As shown in **Figure 11**, the shift in baseline firing developed in both regions as the rat learned the new response–outcome contingencies within a block and returned to its original level during the following block. In the comparison block, in which the value of reward available in the preferred well was similar, the baseline firing rate did not change significantly across the block. Thus like the R–O correlates described earlier, the shift in baseline firing identified here reflected not the value of the outcome but its specific idiosyncratic characteristics. Importantly this comparison also suggests that the baseline shift was not simply a recording artifact, because it began and ended systematically at the beginning of particular blocks. Also supporting this conclusion is the observation that preferred blocks occurred as often in the middle two blocks as in the first or last blocks, as would be expected from recording artifacts that appeared at the beginning or end of the session (in dorsolateral striatum, 43 of 99 neurons preferred one of the two middle blocks; in dorsomedial 43 of 112 neurons did so).

DISCUSSION

Clinical, behavioral and neurophysiological evidence has long pointed to an important role of dorsal striatum in motor control (Denny-Brown and Yanagisawa, 1976; Flowers, 1976; Knowlton et al., 1996; Graybiel, 1998; Jog et al., 1999; Packard and Knowlton, 2002; Barnes et al., 2005). Consistent with this idea, we found here that more than half of all neurons in dorsal striatum were selective for the movement that was performed on a given trial, either during or before the movement itself. These neurons typically showed a phasic increase in firing during or before one of the two trained movements and a slight inhibition during the movement in the opposite direction. In many neurons, this inhibition was statistically significant both immediately before and during the movement, and therefore could reflect a functionally important inhibition of the competing learned response (see Supplementary Material).

The specific function of dorsal striatum in motor control is often thought to involve automatic, habitual or stimulus-driven behavior (Packard et al., 1989; Packard and McGaugh, 1996; Featherstone and McDonald, 2004a,b, 2005; Tang et al., 2007; Balleine et al., 2009). In this conception, the dorsal striatum promotes the acquisition (Carelli et al., 1997; Nakamura and Hikosaka, 2006) and/or stores (Atallah et al., 2007) S–R associations, which allow a sensory stimulus to trigger a movement or series of movements whenever it is encountered. Consistent with this idea, we found evidence of S–R encoding in nearly 20% of response-selective neurons across the dorsal striatum. In these neurons, movement selectivity depended on the identity of the stimulus that instructed that movement and the history of the association between that stimulus and the movement. Thus a neuron that fired for a response on

forced-choice trials, which were cued by one odor, fired significantly less or significantly more when the rat made the same response on free-choice trials, which were cued by a different odor. This selectivity was maintained across different blocks, during which different outcomes were presented for each response, and is therefore consistent with outcome-independent S–R representations. Stimulus-dependent encoding such as this has not been found in other interconnected brain regions, such as orbitofrontal cortex and ventral striatum (Feierstein et al., 2006; Roesch et al., 2006, 2009). Insofar as it has been tested in these regions, response-selective

FIGURE 10 | Baseline firing rates reflect available outcomes in both dorsolateral (A) and dorsomedial (B) striatum. Curves show average peak-normalized firing rates (±SEM) during the last 20 forced-choice trials of each block, aligned on the beginning of the trial, collapsed across each neuron's preferred block (designated according to the block with the highest average firing rate). Neurons in these populations showed an elevated baseline firing rate during their preferred block. Populations included all neurons with a significant effect of block, but no effect of direction. (99 out of 489 recorded in dorsolateral; 112 out of 587 recorded in dorsomedial).

encoding seems to be identical regardless of the stimulus that instructs the response. Thus the current result would be consistent with proposals that the dorsal striatum plays a specialized role in encoding S–R associations.

Of course alternative interpretations of the meaning of this putative S–R encoding are possible. For example, activity that distinguishes free- and forced-choice trials could reflect the differential use of general decision-making processes during the two kinds of trials. Although such an interpretation is impossible to rule out in the context of the current experiment, we would argue that interpreting the two kinds of trials in terms of the differential relationships between stimuli and responses is more straightforward and parsimonious.

While in theory the S–R encoding that underlies habits should not include representations of expected outcomes, many studies have found that striatal encoding of movements is strongly modulated by expected outcomes (Hollerman et al., 1998; Hassani et al., 2001; Haruno et al., 2004; Tricomi et al., 2004; Delgado, 2007; Pasquereau et al., 2007; Lau and Glimcher, 2008; Tanaka et al., 2008; Hori et al., 2009). These have included recordings made in both the caudate nucleus and the putamen in non-human primates, in which various aspects of the expected outcome modulate encoding before and during movements made to obtain those outcomes (Hassani et al., 2001; Pasquereau et al., 2007; Lau and Glimcher, 2008; Hori et al., 2009). This outcome-modulation has been interpreted as encoding the value of the action taken and possibly mediating goal-directed behavior, either by allowing the evaluation of actions or by modulating the performance of actions. Consistent with these reports, we found that expected outcomes modulated activity in over 30% of response-selective neurons across dorsal striatum, and the population activity showed a strong outcome-dependency. Furthermore, we found that this outcome encoding was in large part inseparable from the stimulus encoding identified earlier. In many neurons, activity depended on stimuli, responses and outcomes, such that they encoded the S–R–O conjunction.

In contrast to previous findings in dorsal striatum and medial prefrontal cortex, outcome-dependency did not reflect generic *value* (Luk and Wallis, 2009). Rather, movement encoding seemed to incorporate a representation of the specific idiosyncratic features of the outcome that could be expected to result from that movement – that is, these neurons represented the R–O contingencies present in a particular block of trials. In previous studies, value has typically been manipulated within a single dimension – either reward size or reward probability – and thus it may have been impossible to distinguish encoding of value from that of outcome identity *per se*. Our task, in contrast, used two qualitatively different value manipulations, which may have allowed the emergence of outcome-related as opposed to purely
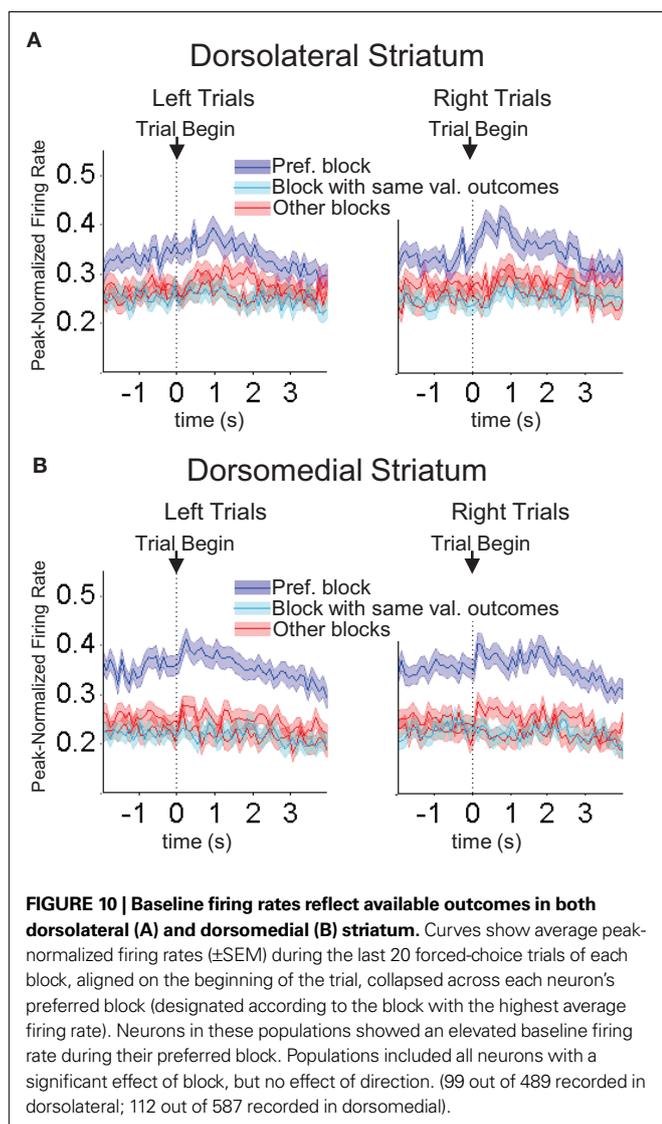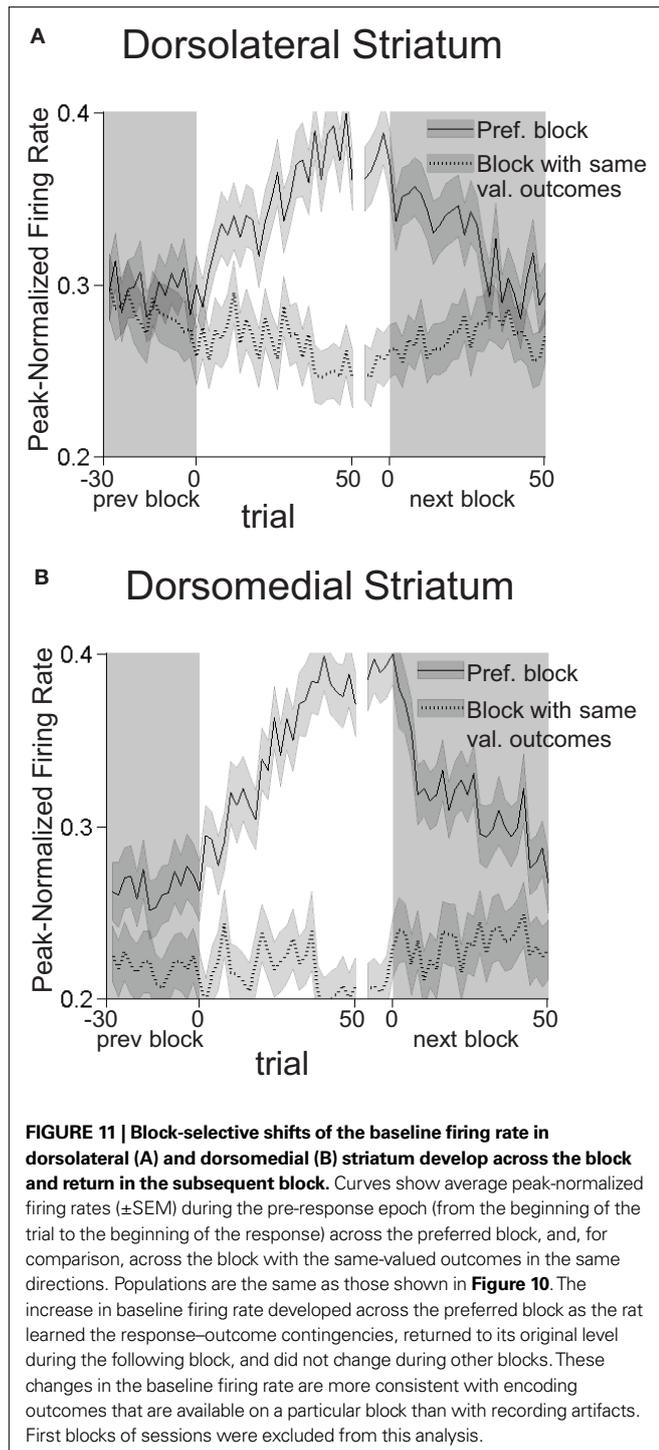
Table 3 | Shown are the number of block-selective neurons that preferred each of the four blocks. In parentheses is the percentage of all block-selective neurons in that region.

	Big left/ small right	Small left/ big right	Short left/ long right	Small left/ big right	Total	Percent of all neurons
Dorsomedial striatum	36 (32%)	30 (27%)	18 (16%)	28 (25%)	112	19%
Dorsolateral striatum	22 (22%)	24 (24%)	32 (32%)	21 (21%)	99	20%

FIGURE 11 | Block-selective shifts of the baseline firing rate in dorsolateral (A) and dorsomedial (B) striatum develop across the block and return in the subsequent block. Curves show average peak-normalized firing rates (±SEM) during the pre-response epoch (from the beginning of the trial to the beginning of the response) across the preferred block, and, for comparison, across the block with the same-valued outcomes in the same directions. Populations are the same as those shown in **Figure 10**. The increase in baseline firing rate developed across the preferred block as the rat learned the response–outcome contingencies, returned to its original level during the following block, and did not change during other blocks. These changes in the baseline firing rate are more consistent with encoding outcomes that are available on a particular block than with recording artifacts. First blocks of sessions were excluded from this analysis.

encoding, in which R–O contingencies seemed to be signaled regardless of the movement that was actually chosen. Like previous reports of action-value encoding in primates and rats (Lauwereyns et al., 2002; Samejima et al., 2005; Lau and Glimcher, 2007, 2008; Ito and Doya, 2009; Kim et al., 2009; Kimchi and Laubach, 2009), the activity we observed occurred before the action was chosen. However it did not appear in general as a phasic increase, but rather as an upward shift in baseline firing rate that developed in particular trial blocks and diminished in the subsequent block. As was the case for phasic changes in firing described earlier, the encoding of available outcomes did not appear to reflect the *value* of the available actions. Instead, it appeared to represent the idiosyncratic outcomes associated with the two specific responses in a block. Although we observed this kind of activity in both dorsomedial and dorsolateral striatum, one of the few significant differences between the two was the stronger selectivity found in dorsomedial compared to dorsolateral striatum. Because such activity is postulated to provide a basis for making choices, this could reflect the greater involvement of dorsomedial striatum in supporting goal-directed choices.

The interpretation that dorsal striatal representations of values (in previous primate studies) or R–O contingencies (in the present study) might underlie goal-directed behavior rests on the assumption that animals were in fact engaging in goal-directed behavior during these recordings. However, because such studies, including ours, have not typically obtained direct evidence that animals are using knowledge of expected outcomes to drive or modulate their behavior, animals could in theory be using habitual, stimulus-driven behavior, even during rapid switches in choice used here or elsewhere (but see Kimchi et al., 2009). Under this interpretation, apparent representations of the value of chosen actions in striatum – R–O correlates – could instead represent reward-induced modulation of the strength of (or effects of arousal on) S–R encoding.

Several pieces of evidence argue against this interpretation in the present study. First, the use of multiple outcomes that are frequently switched would tend to maintain a reliance on goal-directed behavior as opposed to habitual behavior (Holland, 2004), which develops preferentially after overtraining with invariant contingencies (Dickinson, 1985). Second, rats showed significant changes in choice behavior and reaction time very quickly after outcomes were switched at block transitions (within 10–20 trials), whereas habitual S–R encoding would be expected to develop more slowly, by trial and error. Finally and most importantly, nearly half of outcome-selective neurons fired more when responses were associated with one of the *less* valuable outcomes. This result contradicts the idea that such firing reflects outcome-induced modulation of the strength of S–R encoding or the effects of arousal on response encoding, since in these explanations one would expect a greater neuronal response for the response associated with the more valuable outcome. Thus it seems likely that outcome-dependent encoding in dorsal striatum reflects a true representation of the expected outcome.

It is important to note that the co-existance of S–R and R–O information in dorsomedial and dorsolateral striatum does not necessarily contradict recent behavioral accounts dissociating the functions of these two sub-regions. Indeed, ample evidence

value-related encoding. Notably in this same task, the ventral striatum shows evidence of value encoding across manipulations, suggesting that dorsal striatum may be somewhat unique in representing outcome features independently of the value of that outcome (Roesch et al., 2009).

In addition to outcome modulation of the activity encoding the chosen movement, we also observed evidence of a different kind of outcome encoding, similar to what has been called "action-value"

exists to support the idea that dorsomedial and dorsolateral striatum play different roles in instrumental learning and decision-making. As noted earlier, lesion and pharmacological manipulations of dorsomedial striatum in rats have been found to selectively impair goal-directed behavior while leaving habitual behavior intact or enhanced (Yin and Knowlton, 2004; Yin et al., 2005a,b), whereas similar manipulations of dorsolateral striatum have impaired habitual behavior and revealed more goal-directed behavior (Yin et al., 2004, 2006; Balleine et al., 2009). Additionally, evidence from both rodents and primates have suggested a temporal dissociation between these two regions: early procedural learning, which would tend to remain more goal-directed, seems to depend more on dorsomedial striatum or caudate, while performance of well-established procedural learning, which would tend to be more habitual, may depend more on dorsolateral striatum or putamen (Miyachi et al., 1997, 2002; Yin et al., 2009). Indeed, recent work comparing neuronal activity in dorsomedial and dorsolateral striatum during performance of a habitual instrumental behavior has shown significantly greater plasticity and movement related firing in dorsolateral regions (Kimchi et al., 2009). By contrast, in a related study, activity in dorsomedial striatum was particularly sensitive to overt changes in the likelihood of reward (Kimchi and Laubach, 2009). Other recent studies have suggested that the two regions may cooperate to control some aspects of behavior, with the dorsolateral striatum supporting stimulus-based action selection and the dorsomedial striatum supporting evaluation of actions based on their relationship to outcomes (Corbit and Janak, 2007; Shiflett et al., 2010). Still other studies report important differences in oscillatory rhythms or in vulnerability to chronic stress across the medio-lateral extent of striatum in rodents (Berke et al., 2004; Dias-Ferreira et al., 2009).

The simplest interpretation of these studies is that sub-regions within dorsal striatum encode different kinds of information, with dorsolateral regions signaling S–R associations and dorsomedial regions signaling R–O associations. However, our findings provide evidence against this simple interpretation. Instead we found surprisingly similar kinds of encoding in both regions of dorsal striatum, both in encoding of S–R associations and in encoding of R–O associations. This is particularly notable in light of the task we used, which included elements of well-established, over-trained learning (S–R associations on forced-choice trials) as well as elements of new R–O learning in each block. In fact, to the extent that encoding was different across the two regions, putative S–R encoding was in evidence earlier in the trial – during the odor period – in dorsomedial striatum, which would be the opposite result to that predicted by the behavioral evidence.

It is of course possible that the lack of differential S–R or R–O encoding in dorsomedial vs. dorsolateral striatum in the current experiment is the result of the particular behavioral paradigm that we used. This paradigm differs in a number of ways from those used in the experiments that have differentiated the function of these two sub-regions of dorsal striatum. For example, in previous experiments, instrumental behaviors such as lever-pressing or chain-pulling were used instead of nose-poking, and the stimuli cuing instrumental behaviors were not explicitly presented as they were in the present experiment. Additionally, in the

current experiment we did not explicitly test whether particular behavioral responses were supported by habitual vs. outcome-guided bases. However, the hypothesis that dorsomedial and dorsolateral striatum support different associative structures to guide instrumental behavior is a broad hypothesis, rather than one tied to the specific instrumental paradigms used. Furthermore, data suggest that these two regions of dorsal striatum maintain different associative structures even when those structures are not actively driving behavior. For example, lesions or pharmacological inactivations of dorsomedial striatum cause behavior that is normally devaluation-sensitive (i.e. outcome-guided) to become devaluation-insensitive (i.e. habitual). This seems to suggest that the associative structures underlying habitual behavior, presumably encoded by remaining parts of striatum, are maintained even under conditions in which those behaviors do not normally support behavior. Under this account, one would expect to find differential encoding of habitual vs. outcome-related associations in dorsomedial vs. dorsolateral striatum.

There are a number of other potential interpretations that could account for the lack of differential encoding that we found in dorsomedial versus dorsolateral striatum. One particularly intriguing possibility is that dorsomedial striatum might support what have been called "model-based" methods of driving behavior, while dorsolateral striatum might support "model-free" reinforcement learning (Daw et al., 2005). This idea could account for the results of the previous lesion and inactivation studies that have found dissociations in the devaluation sensitivity of instrumental behaviors supported by the two sub-regions of dorsal striatum. At the same time, however, model-free and model-based accounts would involve similar associative representations, which could account for the similar encoding that we found in the two regions. For example, both model-free and model-based methods might involve S–R associations, but might arrive at them through different computational processes.

A second possibility might be that differences in connectivity and output patterns, rather than information content, might determine the roles of these two striatal sub-regions. This could involve gross differences in the anatomical projections of these two regions, as suggested by the notion of parallel loops involving different part of the basal ganglia (Alexander et al., 1986, 1990; Groenewegen et al., 1990). Indeed, even if projection patterns from medial and lateral dorsal striatum are partially overlapping (Hedreen and DeLong, 1991; Joel and Weiner, 1994; Haber et al., 2000), more subtle differences in output could explain differential functionality. For example, projections of different neural populations to the same downstream areas could allow information in one sub-region to directly oppose the same information signaled by the other sub-region. Notably, this simple model could be easily implemented in the neural circuitry within striatum or in downstream areas and would explain the behavioral results described above. Assuming that encoding happens more rapidly in dorsomedial striatum or anterior caudate in primates, as suggested by some data (Miyachi et al., 1997, 2002; Pasupathy and Miller, 2005; Williams and Eskandar, 2006; Kimchi and Laubach, 2009; Yin et al., 2009), initial behavior would be based on the value of the outcome associated with the responses (i.e. goal-directed). Later, as information represented

in dorsolateral striatum becomes stronger, behavior could come under the control of associations between antecedent cues and the response (i.e. habitual).

Obviously, additional work is necessary to test these speculative explanations; however our results highlight the need to combine single-unit recording with behavioral work, even when behavioral results seem crystal clear, in order to fully understand how information processing in different parts of a circuit generates behavioral effects. The critical functions of these two regions could not have been inferred from our single-unit data, but neither is the behavioral data sufficient to fully understand how those critical functions arise.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/neuroscience/integrativeneuroscience/paper/10.3389/fnint.2010.00012/

REFERENCES

Alexander, G. E., Crutcher, M. D., and DeLong, M. R. (1990). Basal ganglia-thalamocortical circuits: parallel substrates for motor, oculomotor, "prefrontal" and "limbic" functions. *Prog. Brain Res.* 85, 119–146.

Alexander, G. E., DeLong, M. R., and Strick, P. L. (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.* 9, 357–381.

Atallah, H. E., Lopez-Paniagua, D., Rudy, J. W., and O'Reilly, R. C. (2007). Separate neural substrates for skill learning and performance in the ventral and dorsal striatum. *Nat. Neurosci.* 10, 126–131.

Balleine, B. W., Delgado, M. R., and Hikosaka, O. (2007). The role of the dorsal striatum in reward and decision-making. *J. Neurosci.* 27, 8161–8165.

Balleine, B. W., Liljeholm, M., and Ostlund, S. B. (2009). The integrative function of the basal ganglia in instrumental conditioning. *Behav. Brain Res.* 199, 43–52.

Barnes, T. D., Kubota, Y., Hu, D., Jin, D. Z., and Graybiel, A. M. (2005). Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. *Nature* 437, 1158–1161.

Berke, J. D., Okatan, M., Skurski, J., and Eichenbaum, H. B. (2004). Oscillatory entrainment of striatal neurons in freely moving rats. *Neuron* 43, 883–896.

Carelli, R. M., Wolske, M., and West, M. O. (1997). Loss of lever press-related firing of rat striatal forelimb neurons after repeated sessions in a lever pressing task. *J. Neurosci.* 17, 1804–1814.

Corbit, L. H., and Janak, P. H. (2007). Inactivation of the lateral but not medial dorsal striatum eliminates the excitatory impact of Pavlovian stimuli on instrumental responding. *J. Neurosci.* 27, 13977–13981.

Daw, N. D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 8, 1704–1711.

Delgado, M. R. (2007). Reward-related responses in the human striatum. *Ann. N. Y. Acad. Sci.* 1104, 70–88.

Denny-Brown, D., and Yanagisawa, N. (1976). The role of the basal ganglia in the initiation of movement. *Res. Publ. Assoc. Res. Nerv. Ment. Dis.* 55, 115–149.

Devan, B. D., and White, N. M. (1999). Parallel information processing in the dorsal striatum: relation to hippocampal function. *J. Neurosci.* 19, 2789–2798.

Dias-Ferreira, E., Sousa, J. C., Melo, I., Morgado, P., Mesquita, A. R., Cerqueira, J. J., Costa, R. M., and Sousa, N. (2009). Chronic stress causes frontostriatal reorganization and affects decision-making. *Science* 325, 621–625.

Dickinson, A. (1985). Actions and habits: the development of behavioral autonomy. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* 308, 67–78.

Featherstone, R. E., and McDonald, R. J. (2004a). Dorsal striatum and stimulus-response learning: lesions of the dorsolateral, but not dorsomedial, striatum impair acquisition of a simple discrimination task. *Behav. Brain Res.* 150, 15–23.

Featherstone, R. E., and McDonald, R. J. (2004b). Dorsal striatum and stimulus-response learning: lesions of the dorsolateral, but not dorsomedial, striatum impair acquisition of a stimulus-response-based instrumental discrimination task, while sparing conditioned place preference learning. *Neuroscience* 124, 23–31.

Featherstone, R. E., and McDonald, R. J. (2005). Lesions of the dorsolateral striatum impair the acquisition of a simplified stimulus-response dependent conditional discrimination task. *Neuroscience* 136, 387–395.

Feierstein, C. E., Quirk, M. C., Uchida, N., Sosulski, D. L., and Mainen, Z. F. (2006). Representation of spatial goals in rat orbitofrontal cortex. *Neuron* 51, 495–507.

Flowers, K. A. (1976). Visual "closed-loop" and "open-loop" characteristics of voluntary movement in patients with Parkinsonism and intention tremor. *Brain* 99, 269–310.

Graybiel, A. M. (1998). The basal ganglia and chunking of action repertoires. *Neurobiol. Learn. Mem.* 70, 119–136.

Groenewegen, H. J., Berendse, H. W., Wolters, J. G., and Lohman, A. H. (1990). The anatomical relationship of the prefrontal cortex with the striatopallidal system, the thalamus and the amygdala: evidence for a parallel organization. *Prog. Brain Res.* 85, 95–116; discussion 116–118.

Haber, S. N., Fudge, J. L., and McFarland, N. R. (2000). Striatonigrostriatal pathways in primates form an ascending spiral from the shell to the dorsolateral striatum. *J. Neurosci.* 20, 2369–2382.

Haruno, M., Kuroda, T., Doya, K., Toyama, K., Kimura, M., Samejima, K., Imamizu, H., and Kawato, M. (2004). A neural correlate of reward-based behavioral learning in caudate nucleus: a functional magnetic resonance imaging study of a stochastic decision task. *J. Neurosci.* 24, 1660–1665.

Hassani, O. K., Cromwell, H. C., and Schultz, W. (2001). Influence of expectation of different rewards on behavior-related neuronal activity in the striatum. *J. Neurophysiol.* 85, 2477–2489.

Hedreen, J. C., and DeLong, M. R. (1991). Organization of striatopallidal, striatonigral, and nigrostriatal projections in the macaque. *J. Comp. Neurol.* 304, 569–595.

Holland, P. C. (2004). Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *J. Exp. Psychol. Anim. Behav. Process* 30, 104–117.

Hollerman, J. R., Tremblay, L., and Schultz, W. (1998). Influence of reward expectation on behavior-related neuronal activity in primate striatum. *J. Neurophysiol.* 80, 947–963.

Hori, Y., Minamimoto, T., and Kimura, M. (2009). Neuronal encoding of reward value and direction of actions in the primate putamen. *J. Neurophysiol.* 102, 3530–3543.

Ito, M., and Doya, K. (2009). Validation of decision-making models and analysis of decision variables in the rat basal ganglia. *J. Neurosci.* 29, 9861–9874.

Joel, D., and Weiner, I. (1994). The organization of the basal ganglia-thalamocortical circuits: open interconnected rather than closed segregated. *Neuroscience* 63, 363–379.

Jog, M. S., Kubota, Y., Connolly, C. I., Hillegaart, V., and Graybiel, A. M. (1999). Building neural representations of habits. *Science* 286, 1745–1749.

Johnson, A., van der Meer, M. A., and Redish, A. D. (2007). Integrating hippocampus and striatum in decision-making. *Curr. Opin. Neurobiol.* 17, 692–697.

Kawagoe, R., Takikawa, Y., and Hikosaka, O. (1998). Expectation of reward modulates cognitive signals in the basal ganglia. *Nat. Neurosci.* 1, 411–416.

Kim, H., Sul, J. H., Huh, N., Lee, D., and Jung, M. W. (2009). Role of striatum in updating values of chosen actions. *J. Neurosci.* 29, 14701–14712.

Kimchi, E. Y., and Laubach, M. (2009). Dynamic encoding of action selection by the medial striatum. *J. Neurosci.* 29, 3148–3159.

Kimchi, E. Y., Torregrossa, M. M., Taylor, J. R., and Laubach, M. (2009). Neuronal correlates of instrumental learning in the dorsal striatum. *J. Neurophysiol.* 102, 475–489.

Knowlton, B. J., Mangels, J. A., and Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science* 273, 1399–1402.

Lau, B., and Glimcher, P. W. (2007). Action and outcome encoding in the primate caudate nucleus. *J. Neurosci.* 27, 14502–14514.

Lau, B., and Glimcher, P. W. (2008). Value representations in the primate striatum during matching behavior. *Neuron* 58, 451–463.

Lauwereyns, J., Watanabe, K., Coe, B., and Hikosaka, O. (2002). A neural correlate of response bias in monkey caudate nucleus. *Nature* 418, 413–417.

Luk, C. H., and Wallis, J. D. (2009). Dynamic encoding of responses and outcomes by neurons in medial prefrontal cortex. *J. Neurosci.* 29, 7526–7539.

Miyachi, S., Hikosaka, O., and Lu, X. (2002). Differential activation of monkey striatal neurons in the early and late stages of procedural learning. *Exp. Brain Res.* 146, 122–126.

Miyachi, S., Hikosaka, O., Miyashita, K., Karadi, Z., and Rand, M. K. (1997). Differential roles of monkey striatum in learning of sequential hand movement. *Exp. Brain Res.* 115, 1–5.

Nakamura, K., and Hikosaka, O. (2006). Facilitation of saccadic eye movements

by postsaccadic electrical stimulation in the primate caudate. *J. Neurosci.* 26, 12885–12895.

Packard, M. G., Hirsh, R., and White, N. M. (1989). Differential effects of fornix and caudate nucleus lesions on two radial maze tasks: evidence for multiple memory systems. *J. Neurosci.* 9, 1465–1472.

Packard, M. G., and Knowlton, B. J. (2002). Learning and memory functions of the basal ganglia. *Annu. Rev. Neurosci.* 25, 563–593.

Packard, M. G., and McGaugh, J. L. (1996). Inactivation of hippocampus or caudate nucleus with lidocaine differentially affects expression of place and response learning. *Neurobiol. Learn. Mem.* 65, 65–72.

Pasquereau, B., Nadjar, A., Arkadir, D., Bezard, E., Goillandeau, M., Bioulac, B., Gross, C. E., and Boraud, T. (2007). Shaping of motor responses by incentive values through the basal ganglia. *J. Neurosci.* 27, 1176–1183.

Pasupathy, A., and Miller, E. K. (2005). Different time courses of learning-related activity in the prefrontal cortex and striatum. *Nature* 433, 873–876.

Roesch, M. R., Calu, D. J., and Schoenbaum, G. (2007). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat. Neurosci.* 10, 1615–1624.

Roesch, M. R., Singh, T., Brown, P. L., Mullins, S. E., and Schoenbaum, G. (2009). Ventral striatal neurons encode the value of the chosen action in rats deciding between differently delayed or sized rewards. *J. Neurosci.* 29, 13365–13376.

Roesch, M. R., Taylor, A. R., and Schoenbaum, G. (2006). Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. *Neuron* 51, 509–520.

Samejima, K., Ueda, Y., Doya, K., and Kimura, M. (2005). Representation of action-specific reward values in the striatum. *Science* 310, 1337–1340.

Schmitzer-Torbert, N. C., and Redish, A. D. (2008). Task-dependent encoding of space and events by striatal neurons is dependent on neural subtype. *Neuroscience* 153, 349–360.

Shiflett, M. W., Brown, R. A., and Balleine, B. W. (2010). Acquisition and performance of goal-directed instrumental actions depends on ERK signaling in distinct regions of dorsal striatum in rats. *J. Neurosci.* 30, 2951–2959.

Squire, L. R., Knowlton, B., and Musen, G. (1993). The structure and organization of memory. *Annu. Rev. Psychol.* 44, 453–495.

Tanaka, S. C., Balleine, B. W., and O'Doherty, J. P. (2008). Calculating consequences: brain systems that encode the causal effects of actions. *J. Neurosci.* 28, 6750–6755.

Tang, C., Pawlak, A. P., Prokopenko, V., and West, M. O. (2007). Changes in activity of the striatum during formation of a motor habit. *Eur. J. Neurosci.* 25, 1212–1227.

Tricomi, E. M., Delgado, M. R., and Fiez, J. A. (2004). Modulation of caudate activity by action contingency. *Neuron* 41, 281–292.

Williams, Z. M., and Eskandar, E. N. (2006). Selective enhancement of associative learning by microstimulation of the anterior caudate. *Nat. Neurosci.* 9, 562–568.

Yin, H. H., and Knowlton, B. J. (2004). Contributions of striatal subregions to place and response learning. *Learn. Mem.* 11, 459–463.

Yin, H. H., and Knowlton, B. J. (2006). The role of the basal ganglia in habit formation. *Nat. Rev. Neurosci.* 7, 464–476.

Yin, H. H., Knowlton, B. J., and Balleine, B. W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *Eur. J. Neurosci.* 19, 181–189.

Yin, H. H., Knowlton, B. J., and Balleine, B. W. (2005a). Blockade of NMDA receptors in the dorsomedial striatum prevents action-outcome learning in instrumental conditioning. *Eur. J. Neurosci.* 22, 505–512.

Yin, H. H., Ostlund, S. B, Knowlton, B. J., and Balleine, B. W. (2005b). The role of the dorsomedial striatum in instrumental conditioning. *Eur. J. Neurosci.* 22, 513–523.

Yin, H. H., Knowlton, B. J., and Balleine, B. W. (2006). Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. *Behav. Brain Res.* 166, 189–196.

Yin, H. H., Mulcare, S. P., Hilario, M. R., Clouse, E., Holloway, T., Davis, M. I., Hansson, A. C., Lovinger, D. M., and Costa, R. M. (2009). Dynamic reorganization of striatal circuits during the acquisition and consolidation of a skill. *Nat. Neurosci.* 12, 333–341.