Encoding of Action History in the Rat Ventral Striatum

Yun Bok Kim,1 Namjung Huh,1 Hyunjung Lee,1 Eun Ha Baeg,1 Daeyeol Lee,2 and Min Whan Jung1
1Neuroscience Laboratory, Institute for Medical Sciences, Ajou University School of Medicine, Suwon, Korea; and 2Department of Neurobiology, Yale University School of Medicine, New Haven, Connecticut

Submitted 20 March 2007; accepted in final form 14 October 2007

Kim YB, Huh N, Lee H, Baeg EH, Lee D, Jung MW. Encoding of action history in the rat ventral striatum. J Neurophysiol 98: 3548–3556, 2007. First published October 17, 2007; doi:10.1152/jn.00310.2007. In a dynamic environment, animals need to update information about the rewards expected from their alternative actions continually to make optimal choices for its survival. Because the reward resulting from a given action can be substantially delayed, the process of linking a reward to its causative action would be facilitated by memory signals related to the animal’s previous actions. Although the ventral striatum has been proposed to play a key role in updating the information about the rewards expected from specific actions, it is not known whether the signals related to previous actions exist in the ventral striatum. In the present study, we recorded neuronal ensemble activity in the rat ventral striatum during a visual discrimination task and investigated whether neuronal activity in the ventral striatum encoded signals related to animal’s previous actions. The results show that many neurons modulated their activity according to the animal’s goal choice in the previous trial, indicating that memory signals for previous actions are available in the ventral striatum. In contrast, few neurons conveyed signals on impending goal choice of the animal, suggesting the absence of decision signals in the ventral striatum. Memory signals for previous actions might contribute to the process of updating the estimates of rewards expected from alternative actions in the ventral striatum.

INTRODUCTION

When the animal’s environment is simple and fixed, behaviors critical for survival can be hard-wired and genetically programmed. In contrast, when the environment is complex and dynamic, animals must discover, by trial and error, an optimal behavioral strategy that maximizes positive outcomes, such as food and the opportunity to reproduce. In reinforcement learning (Sutton and Barto 1998), subjective estimates for such as food and the opportunity to reproduce. In reinforcement learning (Sutton and Barto 1998), subjective estimates for rewards expected in the future are referred to as value functions. In a dynamic environment, animals need to update information about the rewards expected from their alternative actions continually to make optimal choices for its survival. Because the reward resulting from a given action can be substantially delayed, the process of linking a reward to its causative action would be facilitated by memory signals related to the animal’s previous actions. Although the ventral striatum has been proposed to play a key role in updating the information about the rewards expected from specific actions, it is not known whether the signals related to previous actions exist in the ventral striatum. In the present study, we recorded neuronal ensemble activity in the rat ventral striatum during a visual discrimination task and investigated whether neuronal activity in the ventral striatum encoded signals related to animal’s previous actions. The results show that many neurons modulated their activity according to the animal’s goal choice in the previous trial, indicating that memory signals for previous actions are available in the ventral striatum. In contrast, few neurons conveyed signals on impending goal choice of the animal, suggesting the absence of decision signals in the ventral striatum. Memory signals for previous actions might contribute to the process of updating the estimates of rewards expected from alternative actions in the ventral striatum.

In contrast, the mechanisms responsible for updating the value functions based on the reward prediction errors are not well understood. Clearly, this updating mechanism has to integrate multiple types of signals, such as value functions and reward prediction errors. In addition, the process of reinforcement learning would be greatly facilitated if memory signals related to the animal’s recent actions are also available in the same anatomical structure involved in updating the value functions because a reward or penalty resulting from a particular action is often revealed after a substantial temporal delay. This problem, referred to as temporal credit assignment, is not trivial for several reasons. First, information about the consequences of a particular action is necessarily delayed in the brain, compared with when the corresponding motor command is issued. Second, a substantial delay can also result from the complex physical properties of the animal’s environment. The problem of temporal credit assignment is particularly challenging when the long-term consequence of an action needs to be assessed. Therefore to update the value function of a chosen action correctly after a temporal delay, memory for previously executed actions must be available in the brain structures that are involved in updating value functions.

Several lines of anatomical and physiological evidence suggest that the striatum might play a key role in the process of updating value functions. First, the convergence of cortical inputs and dopaminergic projections to the striatum provides the anatomical substrate necessary for integrating reward prediction errors and value functions. Inputs from cortical areas related to planning and execution of motor responses, such as the prefrontal cortex, may provide the signals related to the actions chosen by the animal and their value functions (Baeg et al. 2003; Barraclough et al. 2004; Lee et al. 2007; Leon and Shadlen 1999; Watanabe 1996), whereas dopamine neurons provide reward prediction errors (Schultz 1998). Second, striatal neurons modulate their activity according to a variety of factors related to the value functions and actions chosen by the animal, suggesting that the striatum plays an important role in action selection (Cromwell and Schultz 2003; Kawagoe et al. 1998; Nicola et al. 2004; Samejima et al. 2005). However, it is not known whether neurons in the striatum encode signals related to the animal’s previous actions.

In the present study, we investigated the time course of signals related to the rat’s goal selection behavior in the ventral striatum (VS) during a visual discrimination task. Our task did not allow the animal to make its choices freely, and hence it could not be determined whether striatal signals related to the selected goals reflect the animal’s choice or simply the move-
ment direction. Nevertheless, in this paper, we refer to the direction of animal’s movement toward a goal as “choice” for the sake of brevity. Rats were rewarded with a constant amount of water for visiting the lit side of a figure-8-shaped maze. This made it possible to distinguish choice-related signals from the signals related to motivational significance of the animal’s behavior because correct actions were always rewarded with the same reward. The results show that many VS neurons modulate their activity according to the animal’s goal choice in the previous trial, suggesting that neural signals related to previous actions necessary for updating value functions exist in the VS.

METHODS
Animal preparations
The experimental protocol was approved by the Ethics Review Committee for Animal Experimentation of the Ajou University School of Medicine. Experiments were performed with young male Sprague-Dawley rats (~9–11 wk old, 250–330 g, n = 3). Animals were individually housed in the colony room and initially allowed free access to food and water. Once behavioral training began, animals were restricted to 30 min access to water after finishing one behavioral session per day. Experiments were performed in the dark phase of 12-h light/dark cycle.

Behavioral task
The animals were trained to perform a visual discrimination task. This was an imperative (forced-choice) task in which the animals were rewarded with water (0.05 ml) only for visiting the side of a figure-8-shaped maze indicated by the visual cue. The overall dimension of the maze was 90 × 50 cm, and the width of the track was 9–13 cm. It was elevated 40 cm from the floor with 5 cm high walls along the entire track. The visual cue was delivered by one of the two green light-emitting diodes (diameter: 5 mm) that were located above the upper left and upper right arms of the maze (8 cm above the maze floor and 3 cm lateral to the midline; Fig. 1A). The sequence of light signals (left vs. right) was chosen randomly across trials. After visiting one of the goal locations, the animal was required to return to the starting location at the center of the maze by completing the remaining track (Fig. 1A). When the animal entered the central section of the maze, one of the visual cues was turned on and the next trial began immediately. The visual cue was extinguished 2 s after its onset, or when the animal exited the central section of the maze, whichever came first. The animals performed the task for ~30–40 trials per day. Presentation of a visual stimulus and the delivery of water were triggered by infrared light beam sensors along the maze.

Electrophysiological recording
A microdrive array (Neuro-hyperdrive, Kopf Instruments, Tujunga, CA) loaded with 12 tetrodes was implanted in the left or right VS (1.9 mm A, 1.0 mm L, 6.5–8.0 mm V from bregma) under deep anesthesia with sodium pentobarbital (50 mg/kg body wt). Tetrodes were fabricated by twisting four strands of polyimide-insulated nichrome wires (H. P. Reid, Palm Coast, FL) together and gently heated to fuse the insulation without short-circuiting the wires (final overall diameter: ~40 μm). The electrode tips were cut and gold-plated to reduce impedance to 0.3–0.6 MΩ measured at 1 kHz. After ~7–10 days of recovery from surgery, tetrodes were gradually advanced toward the VS (maximum 320 μm/day). Once the tetrodes entered the intended recording region, they were advanced only 20–40 μm/day. When new unit signals that were different from those recorded in the previous day were obtained, recordings were made without advancing the tetrode. The identity of unit signals was determined based on the clustering pattern of spike waveform parameters (Fig. 2), averaged spike waveforms, baseline discharge frequencies, autocorrelograms, and interspike interval histograms (Baeg et al. 2007). Unit signals were collected via a headstage of complementary metal oxide semiconductor (CMOS) operational amplifier (Neuralynx, Bozeman, MT), amplified with the gain of 5,000–10,000, band-pass-filtered between 0.6 and 6 kHz, digitized at 32 kHz, and stored on a SUN4u workstation using Cheetah data acquisition system (Neuralynx). Single units were isolated by examining various two-dimensional projections of the relative amplitude data in all channels of each tetrode (Fig. 2A), and manually applying boundaries to each subjectively identified unit cluster using custom software (Xclust, M. Wilson). Spike width was also used as an additional feature of spike waveforms for unit isolation. Only those clusters that were clearly separable from each other and from background noise throughout the recording session were included in the analysis. The head position of the animals was also recorded at 60 Hz by tracking an array of light-emitting diodes mounted on the headstage. Unit signals were recorded with the animals placed on a pedestal (resting period) for ~10 min before and after experimental sessions to examine the stability of recorded unit signals. Unstable units were excluded from the analysis. When recordings were completed, small marker lesions were made by passing an electrolytic current (50 μA, 30 s, cathodal) through one channel of each tetrode and recording locations were verified histologically as previously described (Baeg et al. 2001).

Data analysis
UNIT CLASSIFICATION. We separated recorded units, using a nonhierarchical k-means clustering algorithm (SPSS 10.0), into two groups.
based on average firing rate, spike duration, and the ratio between the peak and valley amplitudes of a filtered spike waveform (Fig. 2B). Of 523 units that were subject to analysis (≥0.1 Hz), the type 1 neurons \((n = 483)\) had low firing rates \((2.5 ± 0.1 \text{ and } 2.6 ± 0.1 \text{ Hz during resting and running periods, respectively})\), long spike durations \((257.6 ± 4.5 \text{ µs})\), and high peak-to-valley ratios \((1.57 ± 0.01)\). In contrast, the type 2 neurons \((n = 40)\) had relatively high firing rates \((26.3 ± 2.1 \text{ and } 27.9 ± 1.9 \text{ Hz during resting and running periods, respectively})\), short spike durations \((201.6 ± 13.7 \text{ µs})\), and low peak-to-valley ratios \((1.39 ± 0.07)\). The type 1 and 2 neurons correspond to putative medium spiny neurons (MSNs) and local interneurons, respectively (Wilson 2004). Although both types of neurons were included in the analyses, essentially the same results were obtained when putative local neurons were excluded from the analyses (data not shown).

BEHAVIORAL STAGES. To determine the time course of neural signals related to different variables manipulated in this study, we divided each trial into four stages, and the mean spike rate in each stage was analyzed separately. These four stages correspond to response selection, approach to reward, reward consumption, and return to the center of the maze (Fig. 1A). The response selection stage corresponds to the central section of the maze. The reward approach stage spans the period between the end of the response selection stage and the time of the animal’s arrival at one of the reward sites. The reward consumption stage was the time period in which the animal stayed in the reward site. The return stage started as the animal departed from a reward site and ended when the animal entered the central section of the maze. Average durations of the response selection, reward approach, reward consumption, and return stages were \(0.81 ± 0.12, 1.26 ± 0.04, 4.76 ± 0.37, \) and \(4.78 ± 0.59 \text{ s, respectively}.\) Each trial consisted of all four behavioral stages beginning with the response selection stage.

ANALYSIS OF CHOICE-RELATED ACTIVITY. To test how the neural signals related to the animal’s choice changed across different stages, we applied a multiple linear regression analysis in which the mean firing rate \(S_t\) of a neuron during a particular behavioral stage of a trial \(t\) was given by a linear function of the animal’s behavioral choice in the same trial \(C_t\) and the three previous trials \(C_{t-1}, C_{t-2},\) and \(C_{t-3}\) as the following

\[
S_t = a_1 + a_2C_t + a_3C_{t-1} + a_4C_{t-2} + a_5C_{t-3} + e_t
\]

where \(a_j\)’s are regression coefficients and \(e_t\) represents the error term.

DISCRIMINANT ANALYSIS. We used a linear discriminant analysis with a leave-one-out cross-validation procedure to examine further how reliably the information about the animal’s behavioral choice in a given trial can be inferred from the neuronal ensemble activity during the response selection stage of each trial. In this analysis, a single trial was removed, and a linear discriminant function was generated based on the neuronal ensemble activity in the remaining trials separated according to the animal’s choice in the current trial (trial lag = 0) or in each of the three previous trials (trial lag = 1–3). The removed trial was then classified based on this discriminant function. This procedure was repeated for all trials and the percentage of correct classification was calculated for each session. We then used a \(t\)-test to determine whether this is significantly different from the
chance level (50% correct classification) across the entire sessions. This analysis was applied separately to the neuronal ensemble activity during the first and last 500 ms of the response selection stage as well as the activity during the entire response selection stage.

We also used a discriminant analysis to test further whether the activity during the response selection stage was related more closely to the visual cue or the animal’s choice in the previous trial. In this analysis, the discriminant function was determined based on the activity in all correct trials, and the percentage of trials in which the error trials were classified correctly according to the animal’s previous choice was computed. This analysis was performed on the activity of individual neurons as well as ensemble activity. The statistical significance was evaluated at the level of 0.05, unless noted otherwise, and all the data are expressed as means ± SE.

**RESULTS**

**Behavioral and neuronal database**

Rats were trained in a visual discrimination task (Fig. 1A) until they performed correctly in >70% of the trials for three consecutive days. Recordings began once the animals reached this criterion. On average, the rats performed 35.4 ± 0.3 trial/session, and the average rate of correct trials was 90.3 ± 0.8%. While the animals were performing the task, a total of 572 well-isolated and stable neurons were recorded in the VS (Fig. 2A). To improve the reliability of the analyses further, neurons with the overall activity lower than 0.1 spikes/s were excluded from the analysis (n = 49). Thus a total of 523 units were included in the analyses described in the following text.

**Signals related to animal’s choice**

Consistent with the findings from previous studies in rats (e.g., Chang et al. 2002; Daw 2003; Lavoie and Mizumori 1994; Mulder et al. 2004; Shibata et al. 2001; Woodward et al. 1999), neurons in the VS displayed diverse patterns of activity during the task performance. Elevated neuronal activities were observed across all behavioral stages and over the entire maze, as illustrated by the firing rate maps in Fig. 3, which were constructed as described previously (Song et al. 2005). Furthermore, many of these neurons displayed modulations in their activity according to the animal’s choice in the current or previous trials. For example, the neuron illustrated in Fig. 3A modulated its activity reliably according to the animal’s choice in the previous trial during the response selection stage (Fig. 3A, previous trial), whereas its activity did not show differential activity related to the animal’s upcoming choice in the same trial (Fig. 3A, current trial). The remaining example neurons shown in Fig. 3 modulated their activity according to the animal’s choice in the current trial during the reward approach (Fig. 3B), reward consumption (Fig. 3C), or return (Fig. 3D) stages. Therefore the activity of these neurons re-

FIG. 3. Examples of choice-specific neurons. A–D: neurons selectively responding to animal’s choice made in the previous and/or current trials during various behavioral stages are shown. Color maps represent the spatial profile of the firing rate (number of spikes/occupancy time) for each VS unit. Red indicates maximum firing rate (from top to bottom: 38.0, 64.3, 16.6, and 15.0 Hz) and dark blue indicates no firing. In the spike raster plots, each tick mark indicates an incidence of unit discharge and each row represents 1 trial. The raster plots are aligned at the beginning of a particular behavioral stage that is also indicated by the arrows in the corresponding firing rate map. Trials were divided into 2 groups depending on the animal’s choice (left: black, right: orange) in the previous (left) or current (right) trial, and spike density functions are shown in corresponding colors below each raster plot. Spike density functions were generated by applying a Gaussian kernel (σ = 50 ms) to the corresponding spike trains.
reflected the animal’s choice after its chosen action has been executed.

We used a linear regression analysis to quantify the effects of the animal’s choices in four successive trials on neural activity (see METHODS). The results based on all trials are shown in Fig. 4A (64 sessions, 523 units). The same analysis was also repeated using only the correct trials (Fig. 4B), but this was done only for the sessions with ≥20 correct trials (39 sessions, 307 units). The animal’s choice in the current trial significantly affected activity in more than a quarter of neurons in the reward approach (25.4% for all and 21.2% for correct trials), reward consumption (31.6 and 31.9%, for all and correct trials, respectively), and return (27.2 and 23.8%, for all and correct trials, respectively) stages. In contrast, in the response selection stage, <3% of neurons were significantly modulated by animal’s choice in the current trial (i.e., impending behavioral choice), which was below the level expected by chance (alpha = 0.05, horizontal lines in Fig. 4). On the other hand, the animal’s choice in the immediately preceding trial (Ct-1) modulated the activity of many neurons (17.0 and 16.9% for all and correct trials, respectively) in the response selection stage. The proportion of such neurons then decreased gradually during the subsequent reward approach (11.7 and 10.1%, for all and correct trials, respectively), reward consumption (7.5% for both), and return (5.4 and 8.8%, for all and correct trials, respectively) stages. Nevertheless, the number of neurons that significantly modulated their activity according to the animal’s choice in the previous trial during the reward approach and reward consumption stages was significantly higher than the level expected by chance (binomial test, P < 0.05). This was also true for the return stage when only correct trials were included in the analysis. Regarding animal’s choice in two or three trials ago (Ct-2 and Ct-3), it modulated only small numbers of neurons in all behavioral stages. None of these numbers were significantly different from the level expected by chance (binomial test, P > 0.05).

During the reward approach stage, the activity of some neurons was significantly influenced by the animal’s previous choice as well as the animal’s choice in the current trial (e.g., Fig. 3B). Among those neurons showing the significant effect of the animal’s current choice on their activity during the reward approach stage (n = 133), 38 neurons (28.6%) also modulated their activity according to the animal’s choice in the previous trial, when all trials were included in the analysis. This was significantly higher than expected when the effects of the animal’s choices in two successive trials were assumed to be independent (χ² test, P < 0.001). In the analysis that included only the correct trials, the number of such neurons was 18 (of 65 neurons, 27.7%), which was also significantly higher than expected by chance (χ² test, P < 0.001). Thus if a neuron significantly modulated its activity according to the animal’s choice in a given trial, this neuron was also likely to modulate its activity according to the animal’s choice in the previous trial during the reward approach stage. This provides additional evidence that the effect of the animal’s behavior in the previous trial on neural activity during the reward approach stage was not due to chance.

We also tested whether the activity seemingly related to the animal’s previous choice might have simply resulted from the variability in the animal’s trajectory during the response selection stage that was systematically related to the animal’s previous choice. This analysis was performed using the coefficients for the first principal component of the animal’s trajectory (see METHODS) that accounted for 61.0% of the total variance of the lateral head position during the response selection stage. We found that the previous choice significantly influenced neuronal activity during the response selection stage more frequently (n = 73 of 523, 14.0%) than the component coefficient (n = 31, 6.0%). Moreover, the coefficients for the previous choice (α1, Eq. 2) estimated with and without the first principal component (P1) were highly correlated (r = 0.953). Therefore the effect of previous choice on neural activity in the response selection stage cannot be entirely explained by the variability in the animal’s trajectory. Because the position of the rat was determined by tracking an array of diodes mounted on rat’s head, it is possible that part of animal’s body was outside the central section of the maze (i.e., return stage) although the animal position was determined to be in the response selection stage. To test whether this factor influenced the results, we performed the same analysis using only the data from the second half of the response selection period (i.e., the last 6 of 12 spatial bins) where rat’s body is expected to be fully contained in the central section. This analysis yielded essentially the same result (data not shown).

To test whether the activity of VS neurons related to the animal’s previous choice might contribute to the process of updating action value functions, we applied a regression model that includes the interaction between the animal’s choice and reward in the previous trial (see METHODS, Eq. 3) to neuronal activity during the response selection stage. The numbers of neurons that significantly modulated their activity according to the animal’s choice in the previous trial, its reward and their interaction were 69 (of 523 neurons, 13.2%), 35 (6.7%), and 38 (7.3%), respectively. Hence the percentage of neurons carrying the signals related to the animal’s previous choice per se was higher than that of neurons that showed the interaction between the animal’s choice and reward in the previous trial.

These results indicate that once the animal made a behavioral choice (i.e., making a left or right turn from the central section), the choice signal was reflected in neural activity in the reward approach stage, and it persisted through the reward consumption and return stages. When the animal returned to the central section (the response selection stage) and was ready...
to begin a new trial, the choice signal was somewhat reduced but still persisted. Thus neural activity in the response selection stage carried signals related to the animal’s behavioral choice in the previous trial \((C_{t-1})\). Once the animal made its choice in the new trial, signals related to the animal’s new choice modulated the neural activity in the VS strongly during the reward approach stage, but at the same time, the previous choice signal still modulated neural activity, albeit to a lesser degree. The signals related to the animal’s previous choice \((C_{t-1})\) were further diminished as the animal proceeded to the reward consumption and return stages. By the time animal proceeded further to the response selection stage in the next trial \((C_{t+1})\), there was no longer any signal related to the signals related to the animal’s previous choice \((C_{t-1})\).

**Choice signal in different phases of the response selection stage**

The lack of neural signals related to impending behavioral choice is somewhat at odds with the results from previous studies suggesting the role of the VS in action selection (Nicola 2007; Nicola et al. 2004; Pennartz et al. 1994; Setlow et al. 2003; Taha and Fields 2006). Thus we further tested the possibility that signals related to the animal’s upcoming action are conveyed largely during a particular phase of the response selection stage. In particular, we tested whether such signals are confined to a late phase of the response selection stage, immediately before the reward approach stage began. Therefore we applied the same multiple regression analysis described in the preceding text separately for the activity during the first and last 500 ms of the response selection stage. As shown in Fig. 5A, this analysis yielded essentially the same results as the analysis based on the activity during the entire response selection stage. Few neurons carried signals related to the impending behavioral choice \((C_t)\), whereas significant proportions of neurons (first 500 ms: 16.8%, \(P < 0.001\); last 500 ms: 16.1%, \(P < 0.001\), binomial test) modulated their activities according to the behavioral choice in the previous trial \((C_{t-1})\). Choices made two or three trials ago \((C_{t-2} \text{ or } C_{t-3})\) did not significantly influence neuronal activity during these time periods.

Whether the activity of VS neurons carried signals related to the animal’s choice in the current trial was further investigated by applying a linear discriminant analysis to neuronal ensemble activity during the response selection stage (8.86 &pm; 4.65 unit/session, 64 sessions). This analysis revealed that during the entire response selection stage as well as its first or the last 500-ms interval, neuronal ensemble activity did not encode any information about the animal’s upcoming choice \((n = 64\) sessions, \(t\)-test, \(P > 0.05\); Fig. 5B). In contrast, significantly more trials were correctly classified according to the animal’s choice in the previous trial on average, 67.2 \pm 2.1, 67.1 \pm 2.0, and 66.5 \pm 2.4% of the trials were correctly classified for the entire response selection stage, the first 500 ms of the response selection stage, and the last 500 ms interval, respectively (Fig. 5B), and all of these values were significantly higher than the chance level \((t\)-test, \(P < 0.001\)). Choices made two trials before \((C_{t-2})\) were correctly discriminated by ensemble activity slightly more frequently than 50% (53.5 \pm 1.2, 52.9 \pm 1.5, and 52.9 \pm 1.4%, respectively; Fig. 5B). Nevertheless, these values were significantly higher than the chance level for the entire response selection stage and for its last 500 ms interval \((t\)-test, \(P = 0.002\) and \(P < 0.020\), respectively). To test whether the effect of the animal’s choice made two trials before \((C_{t-2})\) might be mediated by any serial correlation in the animal’s choices in successive trials, this analysis was repeated after excluding the sessions in which the animal’s successive choices were significantly correlated \((n = 9\) sessions). The level of correct discrimination was still significantly higher than the chance level (entire response selection period, 52.5 \pm 1.2%, \(t\)-test, \(P = 0.039\), suggesting that the effect of the animal’s choice made two trials before did not result from the correlation in the animal’s successive choices. Choices made three trials before \((C_{t-3})\) could not be discriminated reliably by ensemble activity in any of these intervals. In summary, these results confirm that VS neuronal activity during the response selection stage conveys information about the previous behavioral choices but not about the animal’s upcoming behavioral choice.

**Signals related to cue versus choice**

The results described so far suggest that signals related to the animal’s previous action are encoded by a substantial fraction of VS neurons during the response selection stage. However, the animals in our study produced relatively small number of errors, resulting in relatively high correlation between sensory cues and behavioral choices. We therefore tested the possibility that the neural activity seemingly related to the animal’s previous choice actually encodes the information about the impending behavioral choice \((C_{t-1})\). Results are shown separately for the same 3 different intervals used in \(A\). **•**, corresponds to the chance level (50%); **••**, correct classification significantly higher than the chance level. The error bars indicate SE.
neuronal activity can be significantly modulated by either, neither or both of the previous choice and cue, this analysis forced the error trials to be classified according to either the previous choice or the previous cue. When the discriminant analysis was based on activities of single neurons (n = 523), 57.3% of error trials (1,062 of 1,855) were classified according to the previous choice. Similarly, when we used neuronal ensemble activity for the discriminant analysis, 67.1% of all error trials (141 of 210) were classified according to the previous choice. Both of these values were significantly higher than expected by chance (binomial test, P < 0.001). Therefore signals related to the animal’s previous choices are unlikely to result from the influence of the sensory cue in the previous trial.

DISCUSSION

Coding of action history in the VS

The results from the present study show that some neurons in the VS carry memory signals for previous actions. Multiple regression analysis on individual neuronal activity revealed that about one-sixth of VS neurons significantly changed their activities during the response selection stage depending on the animal’s choice in the previous trial. Similar results were obtained with the discriminant analysis applied to the ensemble activity. The effect of the animal’s previous choice on neural activity was found for the initial as well as last 500-ms interval during the response selection stage, suggesting that the memory signal for the previous choice is maintained throughout the entire response selection stage. Although this effect diminished gradually over time during the following behavioral stages, a statistically significant number of neurons showed signals related to the animal’s choice in the previous trial during the subsequent stages (reward approach, reward consumption, and return). Interestingly, the activity of many neurons in the VS encoded the signals related to the animal’s previous choice as well as the signals related to the animal’s choice in the current trial, suggesting that memory signals related to two different actions are maintained in the VS simultaneously.

The optimal behavioral strategy for the animal during the visual discrimination task used in this study was relatively simple and only required the animal to move to the direction indicated by the visual cue. Therefore once the animals were fully trained, there was no further need for learning, and the outcomes from their choices were always fully predictable. The fact that the neurons in the VS still encoded the signals related to the animal’s previous choices therefore suggests that such signals are transmitted to the VS even when they are no longer needed. Because the animal’s environment can always change unpredictably, the presence of signals related to the animal’s previous choices in the VS might be useful when the contingencies between the animal’s actions and reward change. It is also possible that such neural signals might play a role in preventing extinction.

Role of memory signals related to previous choices

The exact role of choice-related memory signals observed in the VS is not clear. Nevertheless, these signals can potentially play an important role in bridging the temporal gap between the time when the animal decides to take a particular action and the time when the outcome of such an action is revealed. Theoretical studies have proposed two different mechanisms to solve this problem of so-called temporal credit assignment (Fig. 1B). Both of these mechanisms have been originally proposed as a possible means to account for the fact that a conditional stimulus can acquire the ability to predict the occurrence of an unconditional stimulus even when there is a temporal gap between them (Sutton and Barto 1990). However, similar mechanisms might be used to bind the signals related to the animal’s action and its consequences. First, the sensory representation of the animal’s action might be augmented by a series of transient signals that are known as complete serial compound stimuli or tapped delay lines (Montague et al. 1996; Pan et al. 2005; Sutton and Barto 1990) (Fig. 1B, TDL) or by signals that gradually decay with a particular time constant (Suri and Schultz 1999). Second, the problem of binding the animal’s action and its subsequent outcome might be resolved through eligibility trace (Fig. 1B, ET), as proposed in TD(λ) algorithm where the parameter λ controls the decay rate for the eligibility trace (Sutton and Barto 1998). In contrast to one-step temporal difference learning [or TD(0)] algorithm, memory traces for previous actions are maintained across many trials in the TD(λ) algorithm. In this regard, a recent study that combined unit recording and modeling suggested a process similar to TD(λ) algorithm (Pan et al. 2005). A main difference between the representation of serial compound stimulus and the eligibility trace is that the latter is explicitly related to the gating of learning process.

A potential biological substrate for either serial compound stimulus representation or eligibility might be provided by sustained neural activity in the VS. If memory signals in the present study indeed correspond to eligibility trace for previous actions, then it spans at least two trials in the current task. It would be important to test in a future study whether the duration of memory signals in the VS can be adjusted depending on the demands of a specific task. It should be noted, however, that some of the results from our study are not consistent with eligibility trace. Whereas eligibility trace for an action should sum over the same repeated choices (Sutton and Barto 1990, 1998), some neurons showed the opposite preference for the current and previous choices (e.g., Fig. 3B). Hence, memory signals in the VS might represent the serial compound stimulus related to the animal’s previous action or multiple processes including the serial compound stimulus as well as eligibility trace. It is also possible that VS memory signals may serve other unknown functions. Further work is required to clarify this issue.

Source of memory signals in the VS

Currently, the source of signals related to the animal’s previous action in the VS is not known. Nevertheless, cortico-striatal projections originating from the prefrontal cortex are likely to convey such information to the VS. Consistent with this possibility, neurons in the rat medial prefrontal cortex, which sends direct projections to the VS (see Vertes 2004 and references therein), often displayed signals more strongly related to the animal’s previous action than the animal’s future...
action in a spatial delayed alternation task, especially while the animal was still learning the task (Baeg et al. 2003). It is possible that memory signals for previous actions exist in the loop consisting of the prefrontal cortex, VS, ventral pallidum/substantia nigra pars reticulata, and thalamus. Therefore it would be important to test whether memory signals for previous actions also exist in the ventral pallidum/substantia nigra pars reticulata and the mediodorsal/ventromedial thalamic nuclei, which receive input projections from the VS and project back to the prefrontal cortex (Groenewegen et al. 1999). Neurons recorded from the dorsolateral prefrontal cortex (DLPFC) of monkeys that were trained to make stochastic choices in a simulated competitive game also displayed robust signals related to the animal’s choice in the previous trial (Barraclough et al. 2004). The activity of DLPFC was sometimes influenced by the animal’s choice two or three trials before the current trial (Seo et al. 2007). Whether signals related to previous actions also exist in the primate striatum is currently unknown. Nevertheless the presence of such signals in the DLPFC raises the possibility that memory signal for previous actions is also maintained in the cortico-basal ganglia loop of the primate brain. Although the DLPFC does not send direct projections to the VS (Haber and McFarland 1999), it projects heavily to the medial and orbital prefrontal cortices (Cavada et al. 2000), which in turn send direct projections to the VS (Haber et al. 1995, 2006). The signals related to previous actions in the DLPFC may be transmitted to the VS via such indirect projections. This possibility needs to be tested in future studies.

Role of the VS in action selection

Our results show that VS neurons do not carry signals on the animal’s choice of future action, which is at odds with the proposed role of the VS in action selection (e.g., Nicola 2007; Pennartz et al. 1994). One line of supporting evidence for this proposal is the finding that some neurons in rat VS selectively responded to a particular cue-response combination (Nicola et al. 2004; Setlow et al. 2003). In these studies, however, one of the sensory cues signaled an available reward, whereas the other cue signaled either no reward or an aversive stimulus. Thus it is not clear whether the observed pattern of activity represents stimulus-action association or stimulus-reward association (i.e., reward expectancy) (Knutson and Cooper 2005; O’Doherty 2004; Schultz 2006). In this regard, Daw (2003) has shown stimulus-evoked anticipatory responses of VS neurons that were independent on rat’s choice of action but dependent on the type of reward, supporting the latter possibility. Our results also indicate that VS neurons do not carry information about specific sensory cues or associated behavioral responses when correct choices always lead to the same amount of reward. Similarly, Chang et al. (2002) have shown that VS neuronal ensemble activity 1 s before lever press did not readily differentiate correct/error or left/right lever presses that led to the same amount of reward. In monkey VS, a considerable number of neurons responded to reward-predicting stimuli, whereas few neurons showed future action-selective activity in a go-no-go task (Schultz et al. 1992). The results from these studies therefore do not support the proposed role of the VS in action selection and are consistent with the behavioral studies suggesting the role of the VS in Pavlovian rather than instrumental conditioning (reviewed in Cardinal et al. 2002). They are also consistent with the human brain-imaging studies that did not find VS activation in association with action selection. Instead, activation of other brain regions, such as the dorsal striatum (DS), was correlated with action selection (reviewed in O’Doherty 2004), which is in line with numerous physiological studies demonstrating DS neural signals related to impending movement direction of the monkey (Alexander and Crutcher 1990; Hikosaka et al. 1989; Kobayashi et al. 2007; Pasqueulle et al. 2007; Pasupathy and Miller 2005; Samejima et al. 2005). Thus with the caveat that the VS may contribute to action selection under some circumstances (e.g., Nicola 2007; Redgrave et al. 1999), our results suggest that action selection generally takes place elsewhere in the brain.

Acknowledgments

We thank J. Kim for help on figure preparation. Present addresses: Y. B. Kim, Dept. of Neuroscience, University of Pittsburgh, Pittsburgh, PA 15260; and E. H. Baeg, Dept. of Psychiatry, University of Pittsburgh, Pittsburgh, PA 15213.

Grants

This work was supported by Korea Research Foundation grant KRF-2004-037-H00021 to Y. B. Kim, KRF-2005-015-E00032, Korea Science and Engineering Foundation Grant 2005-000×10199×0, the 21C Frontier Research Program. and the Cognitive Neuroscience Program of the Korea Ministry of Science and Technology to M. W. Jung.

References

Haber SN, Kim KS, Maillly P, Calzavara R. Reward-related cortical inputs define a large striatal region in primates that interface with associative...


