Discrete coding of stimulus value, reward expectation, and reward prediction error in the dorsal striatum

Kei Oyama,1,2 Yukina Tateyama,1 István Hernádi,3 Philippe N. Tobler,4 Toshio Iijima,1 and Ken-Ichiro Tsutsui1
1Division of Systems Neuroscience, Tohoku University Graduate School of Life Sciences, Sendai, Japan; 2Department of Physiology, Tohoku University School of Medicine, Sendai, Japan; 3Department of Experimental Zoology and Neurobiology and Szentagothai Research Center, University of Pécs, Pécs, Hungary; and 4Laboratory for Social and Neural Systems Research, Department of Economics, University of Zurich, Zurich, Switzerland

Submitted 29 January 2015; accepted in final form 13 August 2015

The striatum, an input stage of the basal ganglia, receives projections from almost all areas of the cerebral cortices (Bolam et al. 2000) as well as from dopamine neurons in the substantia nigra pars compacta (SNc; Anden et al. 1964). These diverse anatomic inputs make the striatum a structure ideal for integrating sensory and motor information from the cerebral cortex with reward information from the SNc. Previous studies have shown that striatal neurons are activated at various events within a trial in a behavioral task such as instruction cue, delay, and execution of the movement leading to reward delivery (Apicella et al. 1992; Hikosaka et al. 1989a,b; Kimura 1990; Rolls et al. 1983) and that such task-related activity is modulated by the likelihood of obtaining the reward (Cromwell and Schultz 2003; Hassani et al. 2001; Hollerman et al. 1998; Kawagoe et al. 1998; Nakamura et al. 2012). Furthermore, investigators using behavioral tasks in which the probability of obtaining reward by one or another action dynamically changed found striatal neurons to track the value of a specific action (Samejima et al. 2005) or that of the one actually chosen (Lau and Glimcher 2008).

A useful way to investigate how the striatum integrates information from the SNc and the cerebral cortex is to record single-unit activity in these structures while animals are performing the same task. The activity of dopamine neurons in the SNc has been widely investigated using a probabilistic Pavlovian conditioning task in which the association between the conditioning stimulus (CS) and subsequent reward (US) is varied parametrically across the full probability range (P = 0–1). Using such a task, Schultz and colleagues have found that dopamine neurons code reward prediction error in their phasic response to the stimulus and the reward and that they also code reward uncertainty in their tonic activity between the CS and the outcome (Fiorillo et al. 2003). We have recently recorded single-unit activity in the rat dorsal striatum and SNc during a probabilistic Pavlovian conditioning task and found that a group of neurons in the dorsal striatum codes reward prediction error information in the same manner as dopamine neurons in the SNc (Oyama et al. 2010). Whereas in that study we focused on the neurons coding reward prediction error, in this study we analyzed the same data set looking for any task-related variation of activity. Furthermore, we analyzed data from an additional experiment extending the delay duration.

MATERIALS AND METHODS

Subjects. Twenty-one male albino Wistar rats weighing 220–300 g were used as subjects. They were individually housed under a 12:12-h light-dark cycle with light onset at 8:00 PM. Throughout the experiments, they were treated in accordance with the National Institutes of Health (NIH) Guide for the Care and Use of Laboratory Animals, the Tohoku University Guidelines for Animal Care and Use, and the American Physiological Society (APS) Guiding Principles for the Care and Use of Vertebrate Animals in Research and Training. The experimental plan of the present study was approved and licensed (2015LSA-011) by the Institutional Animal Care and Use Committee of Tohoku University.

Apparatus. Experiments were conducted in a dimly lit, sound-attenuated room. Brief auditory CSs were generated by a personal computer and presented diagonally from two loudspeakers (ASP-701; ELECOM) 30 cm from the head of the rat. An infrared sensor system was used to detect conditioned and unconditioned spout-licking movement.

Behavioral procedure and task. Before behavioral training, a head-fixation device consisting of two metal tubes and a stainless steel screw as a grounded reference electrode were fixed to the skull with dental cement under anesthesia induced by a combination of ketamine (80.0 mg/kg) and xylazine (0.8 mg/kg). After recovery from the
surgery, each rat was habituated to an acrylic half-cylinder restraining device (diameter: 8.5 cm, length: 15 cm) that was combined with a stereotaxic head-fixation frame (SR-5R; Narishige). During the task training and single-unit recording sessions, the rat was placed in the restraining device with its head fixed firmly and painlessly in a stereotaxic device (Fig. 1A).

The rats were trained with a probabilistic classic conditioning procedure. Five different auditory stimuli with the same intensity but with different frequencies ranging from 1.2 to 14 kHz (1.2, 2, 5, 9, and 14 kHz) were used as CSs, each indicating a different reward probability \((P = 0, 0.25, 0.5, 0.75, \text{ or } 1.0)\). Combinations of tone frequencies and reward probabilities were varied between rats. To dissociate reward-probability-dependent neuronal activity from the auditory sensory response, the combinations of tone frequencies and reward probabilities were organized so that a reward-probability-dependent tuning of response amplitude would appear as multipeaked tuning when responses are plotted against log-aligned tone frequency. This allowed us to dissociate reward-probability-dependent activity from typical sensory responses that would be expressed as single-peaked tuning when responses are plotted against log-aligned tone frequency. This resulted in a situation where a reward probability of 0% would be associated with a single-peak tuning curve, and a reward probability of 100% would be associated with a single-peak tuning curve.

In each trial, a 1.5-s CS was followed by a 0.5-s delay. Whether reward occurred immediately after the delay was determined probabilistically depending on the CS, and in a rewarded trial a solenoid valve opened for 250 ms and delivered 50 μL of a sucrose solution through a spout in front of the rat’s mouth. The intertrial interval (ITI) was usually set to one of six durations, each consisting of a fixed 4 s plus an exponentially distributed interval with a mean of 5 s. The exception was when an unpredicted reward was given during the ITI. In that case, the time between the end of the previous trial and the unpredicted reward and the time between the unpredicted reward and the start of the next trial were both set to 1 of the above regular ITI durations. Trial sequence was predetermined by a personal computer so that each of the 5 CSs and the unpredicted reward appeared twice in a block of 10 trials. A daily session consisted of 600 trials.

An additional experiment was conducted to identify the task events to which the activity of the recorded neurons was time-locked. In this experiment, 6 rats were used as subjects. The length of the delay period was extended in a stepwise fashion, and three different auditory stimuli were used as CSs indicating reward probabilities of 0, 50, and 100%. Again, an unpredicted reward was occasionally given during the ITI. The CS indicating a reward probability of 50% appeared twice as often as the CSs indicating reward probabilities of 0 or 100%. As a consequence, the number of rewarded trials in the 50% condition was the same as that in the 100% condition. Each recording session for a delay duration consisted of 60–90 trials, and the initial 20–30 trials after delay extension were excluded from analysis (allowing rats to adapt to the new timing of the reward delivery). After a neuron was isolated, the neuronal activity was first recorded with a 0.5-s delay, and then the delay duration (i.e., the time without an explicit timing cue) was extended to 1.5 s. The delay was then extended from 1.5 to 3.5 s and, finally, set back to 0.5 s.

**Single-unit recording.** The recording session began after the rat’s anticipatory licking responses discriminated between probabilities during the CS and delay period. Chronic access to the brain was provided by using a second surgical procedure to open a hole in the skull and attach a recording chamber over it. The position of the hole (anteriorposterior = +2.0 to −1.5 mm from bregma; lambda = 1.0–4.5 mm from the midline) was determined according to the standard stereotaxic atlas (Paxinos and Watson 2005). After recovery from surgery, the activity of single neurons was recorded extracellularly, using tungsten microelectrodes with a platinized tip (1–3 MΩ measured at 1 kHz, 0.125-mm-diameter shaft; FHC), while the rat performed the Pavlovian conditioning task. The electrode was attached to a hydraulic microdrive (MO-15; Narishige) so that it could be advanced into the brain. Electrophysiological signals were amplified (10,000 times) and band-pass filtered (low-cut: 100 Hz; high-cut: 10,000 Hz) with a standard biophysical amplifier (BioAmp A2-v6; Supertech) and were displayed on an oscilloscope (CS-4125A; Kenwood). The amplified signals were also rendered audible and presented to the experimenter through headphones. The action potentials of isolated neurons were sorted by a window discriminator (DDIS-1; Bak Electronics) and displayed on a digital storage oscilloscope (DOS-7040; Kenwood). The recorded electrophysiological signals were digitized at 25 kHz by using an analog-to-digital conversion interface (Power1401; CED) and then stored on a hard disk of a personal computer (X100; IBM). The times of the detected action potentials, licking movements, and task events were also stored on the hard disk. Rasters and histograms showing the neuronal activity recorded under each probability condition and the response to the unpredicted reward were displayed online on a liquid-crystal display (LCD) video screen. If visual inspection suggested that the neural activity was related to one or more task events (CS, delay, and/or reward) or to the unpredicted reward given during the ITI, we stored the recorded data on the computer for offline analysis. The data set included the activity recorded in at least 7 trials for each probability...
condition, and in this study the activity was subjected to further analysis.

Analysis of neuronal activity. When analyzing the activity of each neuron, we focused on the activity during 4 time periods within the trial: stimulus-related activity from CS onset until 750 ms after the CS onset (1st half of the CS presentation), stimulus-related activity from 750 ms before the CS offset to CS offset (2nd half of the CS presentation), reward expectation activity from 1,000 ms before reward delivery until reward delivery, and reward activity (and prediction error response) from reward onset until 500 ms after reward delivery.

We classified neurons into 3 groups according to which of the 3 prereward time windows was the 1 in which the neuron showed the greatest change in activity (in this analysis, we excluded the time window after reward delivery, as we wanted to classify neurons based on the activity before receipt of the reward). We calculated the $t$-value by comparing the baseline activity (500 ms before CS onset) and the activity in each time window. The time window that yielded the largest absolute $t$-value was considered the 1 with the greatest change in activity.

Neurons showing the greatest change during the 750 ms after the CS onset (1st half of the CS presentation), during the 750 ms before the CS offset (2nd half of the CS presentation), and during the 1,000 ms before reward delivery were classified as CS phasic, CS tonic, and US buildup neurons, respectively.

To analyze the information content coded by each neuron in each time window, we conducted a multiple linear regression analysis on a trial-by-trial basis with the reward probability and uncertainty as regressors ($P < 0.05$, uncorrected for multiple regressors). Uncertainty was quantified as relative variance, which was calculated as the difference between the baseline activity and the activity in each time window. The time window that yielded the largest absolute $t$-value was considered the 1 with the greatest change in activity.

Neurons showing the greatest change during the 750 ms after the CS onset (1st half of the CS presentation), during the 750 ms before the CS offset (2nd half of the CS presentation), and during the 1,000 ms before reward delivery were classified as CS phasic, CS tonic, and US buildup neurons, respectively.

To analyze the information content coded by each neuron in each time window, we conducted a multiple linear regression analysis on a trial-by-trial basis with the reward probability and uncertainty as regressors ($P < 0.05$, uncorrected for multiple regressors). Uncertainty was quantified as relative variance, which was calculated as the difference between the baseline activity and the activity in each time window. The time window that yielded the largest absolute $t$-value was considered the 1 with the greatest change in activity.

Neurons showing the greatest change during the 750 ms after the CS onset (1st half of the CS presentation), during the 750 ms before the CS offset (2nd half of the CS presentation), and during the 1,000 ms before reward delivery were classified as CS phasic, CS tonic, and US buildup neurons, respectively.

To analyze the information content coded by each neuron in each time window, we conducted a multiple linear regression analysis on a trial-by-trial basis with the reward probability and uncertainty as regressors ($P < 0.05$, uncorrected for multiple regressors). Uncertainty was quantified as relative variance, which was calculated as the difference between the baseline activity and the activity in each time window. The time window that yielded the largest absolute $t$-value was considered the 1 with the greatest change in activity.

Neurons showing the greatest change during the 750 ms after the CS onset (1st half of the CS presentation), during the 750 ms before the CS offset (2nd half of the CS presentation), and during the 1,000 ms before reward delivery were classified as CS phasic, CS tonic, and US buildup neurons, respectively.

To analyze the information content coded by each neuron in each time window, we conducted a multiple linear regression analysis on a trial-by-trial basis with the reward probability and uncertainty as regressors ($P < 0.05$, uncorrected for multiple regressors). Uncertainty was quantified as relative variance, which was calculated as the difference between the baseline activity and the activity in each time window. The time window that yielded the largest absolute $t$-value was considered the 1 with the greatest change in activity.

Neurons showing the greatest change during the 750 ms after the CS onset (1st half of the CS presentation), during the 750 ms before the CS offset (2nd half of the CS presentation), and during the 1,000 ms before reward delivery were classified as CS phasic, CS tonic, and US buildup neurons, respectively.

To analyze the information content coded by each neuron in each time window, we conducted a multiple linear regression analysis on a trial-by-trial basis with the reward probability and uncertainty as regressors ($P < 0.05$, uncorrected for multiple regressors). Uncertainty was quantified as relative variance, which was calculated as the difference between the baseline activity and the activity in each time window. The time window that yielded the largest absolute $t$-value was considered the 1 with the greatest change in activity.

Neurons showing the greatest change during the 750 ms after the CS onset (1st half of the CS presentation), during the 750 ms before the CS offset (2nd half of the CS presentation), and during the 1,000 ms before reward delivery were classified as CS phasic, CS tonic, and US buildup neurons, respectively.

To analyze the information content coded by each neuron in each time window, we conducted a multiple linear regression analysis on a trial-by-trial basis with the reward probability and uncertainty as regressors ($P < 0.05$, uncorrected for multiple regressors). Uncertainty was quantified as relative variance, which was calculated as the difference between the baseline activity and the activity in each time window. The time window that yielded the largest absolute $t$-value was considered the 1 with the greatest change in activity.

Neurons showing the greatest change during the 750 ms after the CS onset (1st half of the CS presentation), during the 750 ms before the CS offset (2nd half of the CS presentation), and during the 1,000 ms before reward delivery were classified as CS phasic, CS tonic, and US buildup neurons, respectively.

To analyze the information content coded by each neuron in each time window, we conducted a multiple linear regression analysis on a trial-by-trial basis with the reward probability and uncertainty as regressors ($P < 0.05$, uncorrected for multiple regressors). Uncertainty was quantified as relative variance, which was calculated as the difference between the baseline activity and the activity in each time window. The time window that yielded the largest absolute $t$-value was considered the 1 with the greatest change in activity.

Neurons showing the greatest change during the 750 ms after the CS onset (1st half of the CS presentation), during the 750 ms before the CS offset (2nd half of the CS presentation), and during the 1,000 ms before reward delivery were classified as CS phasic, CS tonic, and US buildup neurons, respectively.
to reward probability or reward uncertainty, which can be quantified as relative variance (1 at $P = 0.5$ and 0 at $P = 0$ and 1). In this paper, we have used the term uncertainty according to the definition by Fiorillo et al. (2003), although recent studies have also referred to uncertainty as “risk” (e.g., Burke and Tobler 2011). We included uncertainty in the regression model because tonic activity of midbrain dopamine neurons is correlated with reward uncertainty (Fiorillo et al. 2003). Table 1 summarizes the results of the multiple linear regression analysis. The proportions of neurons that showed a positive correlation between the activity and reward probability were highest in the 3 time windows before the reward delivery, especially
During the 1st half of the CS presentation, the majority (53%, 40/76) showed a positive correlation between CS response and reward probability. On the other hand, half of the CS phasic neurons (50%, 38/76) showed a negative correlation between reward response and reward probability. In total, 34% (26/76) showed a positive correlation between the CS response and reward probability and showed a negative correlation between the reward response and reward probability.

Figures 3 and 4 show the activity of a representative CS phasic neuron and the average histograms of the activity of 26 CS phasic neurons that showed CS and reward responses positively and negatively related to the reward probability. They showed a phasic response both to the CS and to the reward. The response during the 1st half of the CS was positively related to reward probability ($r = 0.55$, $P < 0.0001$), whereas the reward response was negatively related to reward probability ($r = -0.44$, $P < 0.0001$) and was highest for the unpredicted reward. These neurons were considered to code a reward prediction error: the discrepancy between the prediction and occurrence of a reward. These neurons were almost the same population as the one we have previously reported as reward prediction error-coding neurons (Oyama et al. 2010).

Table 1. Summary of the relationship between the activity of each type of neuron and reward probability or uncertainty

<table>
<thead>
<tr>
<th>Type</th>
<th>Window (ms)</th>
<th>Probability</th>
<th>Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
<td>Negative</td>
<td>Positive</td>
</tr>
<tr>
<td>CS phasic, $n = 76$</td>
<td>CS 1st half (0–750)</td>
<td>40 (53*)</td>
<td>2 (3)</td>
</tr>
<tr>
<td></td>
<td>CS 2nd half (750–1,500)</td>
<td>16 (21*)</td>
<td>8 (11*)</td>
</tr>
<tr>
<td></td>
<td>Before reward (1,000–2,000)</td>
<td>17 (22*)</td>
<td>7 (10*)</td>
</tr>
<tr>
<td></td>
<td>After reward (2,000–2,500)</td>
<td>1 (1)</td>
<td>38 (50*)</td>
</tr>
<tr>
<td>CS tonic, $n = 23$</td>
<td>CS 1st half (0–750)</td>
<td>12 (52*)</td>
<td>0 (0)</td>
</tr>
<tr>
<td></td>
<td>CS 2nd half (750–1,500)</td>
<td>13 (57*)</td>
<td>1 (4)</td>
</tr>
<tr>
<td></td>
<td>Before reward (1,000–2,000)</td>
<td>12 (52*)</td>
<td>1 (4)</td>
</tr>
<tr>
<td>US buildup, $n = 95$</td>
<td>After reward (2,000–2,500)</td>
<td>0 (0)</td>
<td>2 (9*)</td>
</tr>
<tr>
<td></td>
<td>CS 1st half (0–750)</td>
<td>26 (27*)</td>
<td>3 (3)</td>
</tr>
<tr>
<td></td>
<td>CS 2nd half (750–1,500)</td>
<td>46 (48*)</td>
<td>1 (1)</td>
</tr>
<tr>
<td></td>
<td>Before reward (1,000–2,000)</td>
<td>65 (68*)</td>
<td>1 (1)</td>
</tr>
<tr>
<td></td>
<td>After reward (2,000–2,500)</td>
<td>28 (29*)</td>
<td>10 (11*)</td>
</tr>
</tbody>
</table>

Numbers in parentheses below the title of the time window show the time from the conditioning stimulus (CS) onset. Numbers in parentheses in each cell show percentages of the neurons. Asterisks indicate that the proportion is greater than chance level (permutation test, $P < 0.05$).
Figure 5 shows the distributions of the standardized $\beta$-coefficients for reward probability and uncertainty of all CS phasic neurons in each time window. In the three time windows before the reward delivery, the distribution for reward probability showed a positive deviation from 0 ($P < 0.05$, Wilcoxon signed-rank test; Fig. 5A). On the other hand, the distribution of the reward responses for reward probability showed a negative deviation ($P < 0.05$, Wilcoxon signed-rank test; Fig. 5A). The distribution for uncertainty did not show any deviation during the first half of the CS presentation ($P > 0.05$, Wilcoxon signed-rank test), that during the second half of the CS presentation and that before reward showed positive deviations from 0 ($P < 0.05$, Wilcoxon signed-rank test; Fig. 5B).

To compare the effect size of each regressor, we compared the absolute values of the standardized $\beta$-coefficients for reward probability and uncertainty. In every time window, the absolute value of the standardized $\beta$-coefficient for reward probability was greater than that of the standardized $\beta$-coefficient for uncertainty ($P < 0.05$, paired $t$-test). We also confirmed this tendency in the proportion of neurons that showed a statistically significant activity change during the 1st half of the CS presentation and after the reward ($\chi^2$-test, $P < 0.05$ with Bonferroni correction).

**Neurons with tonic CS response (CS tonic neurons).** Of the 194 recorded neurons, 12% (23/194) were classified as CS tonic neurons. They typically showed a tonic response during the CS presentation, and 30% of those neurons (7/23) showed a significant response to the reward. The proportions of neu-
rons that showed a positive correlation between activity and reward probability were highest in the 3 time windows before the reward delivery, especially during the 2nd half of the CS presentation (Table 1). In that time window, the majority (57%, 13/23) showed a positive correlation between CS response and reward probability.

Figures 6 and 7 show the activity of a representative CS tonic neuron and the average histograms of the activity of 13 CS tonic neurons that showed a CS response positively related to the reward probability. The neurons were tonically active during the presentation of the CS and showed no response to reward delivery. The CS response was positively related to reward probability ($r = 0.78, P < 0.0001$).

Figure 8 shows the distribution of the standardized β-coefficients for reward probability and uncertainty of all CS tonic neurons in each time window. In the three time windows before the reward delivery, the distribution for reward probability showed a positive deviation from 0 ($P < 0.05$, Wilcoxon...
The distribution of reward responses for reward probability did not show any deviation ($P > 0.05$, Wilcoxon signed-rank test; Fig. 8A). In every time window, the distribution for uncertainty did not show any deviation ($P > 0.05$, Wilcoxon signed-rank test; Fig. 8B). In the three time windows before the reward delivery, the absolute value of the standardized $\beta$-coefficient for reward probability was greater than that of the standardized $\beta$-coefficient for uncertainty ($P < 0.05$, paired $t$-test). We also confirmed this tendency in the proportion of neurons that showed a statistically significant activity change during the 2nd half of the CS presentation ($\chi^2$-test, $P < 0.05$ with Bonferroni correction).

**Fig. 8.** Distributions of the standardized $\beta$-coefficients of CS tonic neurons. Conventions are the same as in Fig. 5.

**Neurons with pre-US buildup activity (US buildup neurons).** Of the 194 recorded neurons, 49% (95/194) were classified as US buildup neurons. They typically showed gradually increasing activity toward the time of reward delivery. Of them, 47% (45/95) also showed a significant response to the reward, whereas 53% (50/95) did not. The proportions of neurons that showed a positive correlation between the activity and reward probability were highest in every time window, especially during the 1,000 ms before the reward delivery (Table 1). In that time window, the majority (68%, 65/95) showed a positive correlation between CS response and reward probability.

**Fig. 9.** Activity of 2 representative US buildup neurons, 1 with a reward response (A) and the other without a reward response (B). Conventions are the same as in Fig. 3.
Fig. 10. Population activity of US buildup neurons. A: peak-normalized and baseline activity-subtracted population histograms of the US buildup neurons that showed prereward activity positively related to reward probability with a reward response \((n = 31)\). B: peak-normalized and baseline activity-subtracted population histograms of the US buildup neurons that showed prereward activity positively related to reward probability without a reward response \((n = 34)\). Conventions are the same as in Fig. 4.

Figures 9A and 10A show the activity of a representative US buildup neuron and the average histograms of the activity of 31 US buildup neurons that showed prereward activity positively related to reward probability with a reward response \((n = 31)\). The activity of these neurons gradually increased toward the time of reward delivery and also was high after the reward. The activity during the 1,000 ms before the reward delivery was positively related to reward probability \((r = 0.68, P < 0.0001)\), whereas the activity after the time of reward delivery was not \((P > 0.1)\).

Figures 9B and 10B show the activity of a representative US buildup neuron and the average histograms of the activity of the 34 US buildup neurons that showed prereward activity positively related to reward probability without a reward response. As in the reward-responsive class of US buildup neurons, activity gradually increased toward the time of the reward delivery. The activity of these neurons, however, steeply decreased after reward delivery. The activity during the 1,000 ms before the reward delivery was positively and linearly related to reward probability \((r = 0.70, P < 0.0001)\). The activity after the time of the reward delivery showed a positive correlation with reward probability \((r = 0.35, P < 0.0001)\), but this may simply reflect the preceding probability-dependent buildup activity. The activity after the usual time of reward was higher in unrewarded trials than in rewarded trials (comparison in intermediate reward-probability conditions; 75, 50, and 25%; \(P < 0.0001\), 2-way ANOVA).

Figure 11 shows the distribution of the standardized \(\beta\)-coefficients for reward probability and uncertainty of all US buildup neurons in each time window. In every time window, the distribution for reward probability showed a positive deviation from 0 \((P < 0.05\), Wilcoxon signed-rank test; Fig. 11A). The distribution for uncertainty did not show any deviation during the first half of the CS presentation \((P > 0.05\), Wilcoxon signed-rank test), that during the second half of the CS presentation and that before reward showed a positive deviation from 0 \((P < 0.05\), Wilcoxon signed-rank test), and that after the reward delivery did not show any deviation from 0.
In every time window, the absolute value of the standardized β-coefficient for reward probability was greater than that of the standardized β-coefficient for uncertainty ($P < 0.05$, paired $t$-test). We also confirmed this finding in the proportion of neurons that showed a statistically significant activity difference during the 1,000 ms before the reward delivery ($\chi^2$-test, $P < 0.05$ with Bonferroni correction).

**Impact of delay extensions.** In classifying the neurons as above, we assumed that the activity between the CS onset and the reward onset of CS phasic and CS tonic neurons was time-locked to the CS onset and that of US buildup neurons was time-locked to the reward onset. To test this assumption, we conducted an additional experiment in which we recorded the activity of neurons while we changed the length of the delay from 0.5 to 1.5 s, then to 3.5 s, and finally back to 0.5 s. The activity of representative CS phasic, CS tonic, and US buildup neurons in this delay manipulation is shown in Fig. 12, A–C. We calculated the onset and peak latency of event-related activity in each neuron (Fig. 13A). In CS phasic neurons, the onset and the peak latency of activity related to the CS onset remained unchanged ($P > 0.05$, Kruskal-Wallis test). The responses related to the reward onset occurred similarly, irrespective of delay duration. In CS tonic neurons, the onset and the peak latency of responses related to the CS remained unchanged ($P > 0.05$, Kruskal-Wallis test). Thus we confirmed that the prereward activity of CS phasic and CS tonic neurons was time-locked to the CS onset, whereas the postreward activity of CS phasic neurons was time-locked to the reward onset. In strong contrast to the prereward responses of CS phasic and CS tonic neurons, the onset times of prereward activity in US buildup neurons shifted away from the CS onset when the delay to reward increased ($P < 0.0001$, Kruskal-Wallis test). The peak response times with respect to the reward onset remained unchanged ($P > 0.05$, Kruskal-Wallis test). Thus in terms of peak response time, the activity of US buildup neurons was time-locked to the reward onset.

We also examined whether the magnitude of the event-related activity changed with the delay extension (Fig. 13B). On average, CS phasic neurons showed weaker CS responses and stronger reward responses with delay extension, CS tonic neurons showed weaker CS responses, and US buildup neurons showed reduced activity during the last 1 s before reward time ($P < 0.05$, 1-way ANOVA). However, it appeared that the activity level changed not only abruptly with the shift of the delay duration, but also gradually as the daily session progressed. It is possible that the gradual decrease of the neuronal firing reflects the decrease of the animal’s motivation. We used as an indicator of motivational level the number of licking movements after reward delivery. To dissociate abrupt change with the shift of the delay duration and gradual change with the motivational level, we applied multiple linear regression analysis on an individual-neuron basis with the delay length as 1 factor and with the number of licking movements after reward delivery as another factor ($P < 0.05$, uncorrected). In this analysis, the number of licking movements was averaged for rewarded trials in 10 trials, and the averaged data were applied to all of 10 trials. The results are summarized in Table 2. Of all the CS phasic neurons we found ($n = 14$), 6 neurons showed reduced CS responses only with delay extension, 1 showed reduced response only with the decrease of licking movement, and 1 showed both effects. Moreover, 6 neurons showed stronger reward responses with delay extension, and 1 showed stronger responses with the decrease of licking movement. Of all the CS tonic neurons we found ($n = 6$), 1 showed reduced CS responses only with delay extension, 1 showed reduced response only with the decrease of licking movement, and 3 showed both effects. Of all the US buildup neurons we found ($n = 19$), 4 showed reduced CS responses only with delay extension, 6 showed reduced response only with the decrease of licking movement, and 4 showed both effects. Thus in about half of the neurons, the analysis of individual neuron data confirmed the main findings from the analysis of average data. We found that a considerable number of striatal neurons were affected by the delay factor, whereas some of them and some others were affected by the motivational factor.

**Relationship between licking movement and neuronal activity.** As it is known that the striatum is involved in motor functions and that many neurons fire in relation to movement (Hikosaka et al. 1989a; Rolls et al. 1983), it is possible that the
observed neuronal activity is related to movement rather than reward probability. As we have previously shown (Oyama et al. 2010), anticipatory licking movement was often positively correlated with reward probability. In such cases, it is impossible to include these two factors as regressors in a multiple linear regression model because of the multicollinearity between them. Therefore, to evaluate whether the reward probability or the licking movement was more suitable to explain the change of neuronal activity, we first divided recorded neurons into two groups: those recorded when probability-dependent anticipatory licking movement was not observed and those recorded when it was observed. For the former group, we applied a model in which reward probability, uncertainty, and the number of licking movements were included as regressors. For the latter group, we applied two models of multiple linear regression analysis independently: one that included reward probability and uncertainty as regressors, and one in which reward probability was replaced by the number of licking movements. For neurons that showed correlations with both reward probability and licking movement, we compared the $r^2$ value for each model to find out which factor affected the neuronal activity more. The results are summarized in Tables 3 and 4.

Of all CS phasic neurons, the activity of 60 was recorded when probability-dependent anticipatory licking movement was not observed. Of these neurons, 27 showed a correlation only with reward probability, 3 showed a correlation only with licking movement, and 5 showed correlations with both factors during the 1st half of the CS presentation. Of 16 neurons for which activity was recorded when probability-dependent anticipatory licking movement was observed, 8 showed a correlation only with reward probability, none showed a correlation only with licking movement, and 2 showed correlations with both factors. Both of those neurons had a higher $r^2$ value for the model including licking movement as a regressor.

Of all CS tonic neurons, the activity of 17 was recorded when probability-dependent anticipatory licking movement was not observed. Of those neurons, 10 showed a correlation only with reward probability, none showed a correlation only with licking movement, and 1 showed correlations with both factors during the 2nd half of the CS presentation. Of 6 neurons for which activity was recorded when probability-dependent anticipatory licking movement was observed, 3 showed a correlation only with reward probability, 2 showed a correlation only with licking movement, and none showed correlations with both factors.

Of all US buildup neurons, the activity of 69 was recorded when probability-dependent anticipatory licking movement was not observed. Of those neurons, 29 showed a correlation only with reward probability, 6 showed a correlation only with licking movement, and 19 neurons showed correlations with both factors during the 1,000 ms before the reward delivery. Of 26 neurons for which activity was recorded when probability-dependent anticipatory licking movement was observed, 14 showed a correlation only with reward probability, 1 showed a correlation only with licking movement, and 4 showed correlations with both factors. All 4 of the neurons that showed correlations with both factors had a higher $r^2$ value for the model including licking movement as a regressor.

Thus a few neurons that were originally considered to be nondifferential were found to show licking-related activity by
movement. Types of neurons decreased with delay extension or decrease of licking extension or decrease of licking movement, whereas prereward activity of all US buildup, CS tonic, CS phasic (reward), reward delivery each type of neuron and delay length or licking movement after animals’ intrinsic expectation level, it is unclear buildup neurons, showed licking-related activity. However, as observed, only a small number were found more likely to be licking-movement-dependent rather than reward probability. The numbers of those neurons are listed in parentheses of Table 4 under the heading Both. These results suggest that the activity of striatal neurons recorded under the Pavlovian conditioning paradigm was related more to reward probability than to reward onset for the reward response of CS phasic neurons, 750 ms before CS offset for CS tonic neurons, and 1,000 ms before reward onset for the reward response of US buildup neurons.

Fig. 13. Effects of delay extension on the onset and peak latency (A) and the magnitude of event-related activity (B) in each type of neuron. A: onset and peak latencies of CS phasic neurons related to the CS onset (top left) and to the reward onset (top right), CS tonic neurons related to the CS onset (bottom left), and US buildup neurons related to the reward onset (Rew. res.; bottom right). Filled circles and open triangles represent the onset and peak latency of each neuron, respectively. B: magnitude of the CS (top left) and reward response (top right) of CS phasic neurons, CS response of CS tonic neurons (bottom left), and prereward activity of US buildup neurons (bottom right). Filled circles, filled squares, and open triangles represent the activity in the 100%, 50%, and 0% conditions, respectively. The time window for determining the peak normalized response was 750 ms from CS onset for the CS response of CS phasic neurons, 500 ms from reward onset for the reward response of CS phasic neurons, 750 ms before CS offset for CS tonic neurons, and 1,000 ms before reward onset for US buildup neurons.

reanalyzing the data with a model including the number of licking movements as a factor. In Table 3, the numbers of those neurons are listed under the heading Licking Only. On the other hand, several neurons that were originally considered to be probability-dependent were found to also show licking-related activity. The numbers of those neurons are listed in Table 3 under the heading Both. Of the neurons that were recorded while probability-dependent licking movement was observed, only a small number were found more likely to be licking-movement-dependent rather than reward probability. The numbers of those neurons are listed in parentheses of Table 4 under the heading Both. These results suggest that the activity of striatal neurons recorded under the Pavlovian conditioning paradigm was related more to reward probability than licking movement, although some neurons, especially some US buildup neurons, showed licking-related activity. However, as anticipatory spout-licking behavior may emerge with the increase of animals’ intrinsic expectation level, it is unclear whether the observed licking-related activity was directly related to motor function or was related only indirectly through such motivational factors.

Relationship between tone frequency and neuronal activity. As it is known that the striatum receives sensory inputs from various cortices, we tested whether the observed activity of striatal neurons can be explained by the typical sensory responses that would appear as single-peak tuning when activity in auditory-related areas is plotted against tone frequency (Bordi and LeDoux 1992; Doron et al. 2002; Sally and Kelly 1988; Sutter and Schreiner 1991). In designing the task, we determined the combination of the tone frequency and the reward probability so that the reward-probability-dependent response would not appear as single-peak tone tuning. We tested for single-peeked tuning by Gaussian curve-fitting to the response magnitude of each type of neuron against the logarithmically scaled tone frequency (in this curve-fitting, the time windows for the responses were for CS phasic neurons the 1st half of the CS presentation, for CS tonic neurons the 2nd half of the CS presentation, and for US buildup neurons 1,000 ms before the reward delivery). Since we found good fitting only for 1 US buildup neuron, we think there is little possibility that the activity of recorded striatal neurons was an artifact of the simple auditory tone tuning.

Table 2. Summary of the relationship between the activity of each type of neuron and delay length or licking movement after reward delivery

<table>
<thead>
<tr>
<th>Type</th>
<th>Delay Length Only</th>
<th>Licking Only</th>
<th>Both</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS phasic (CS), n = 14</td>
<td>6</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>CS phasic (reward), n = 14</td>
<td>5</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>CS tonic, n = 6</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>US buildup, n = 19</td>
<td>4</td>
<td>6</td>
<td>4</td>
</tr>
</tbody>
</table>

Note that reward responses of CS phasic neurons increased with delay extension or decrease of licking movement, whereas prereward activity of all types of neurons decreased with delay extension or decrease of licking movement.

Table 3. Summary of the relationship between the activity of each type of neuron and reward probability or licking movement recorded without probability-dependent licking movement

<table>
<thead>
<tr>
<th>Type</th>
<th>Probability Only</th>
<th>Licking Only</th>
<th>Both</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS phasic, n = 60</td>
<td>27</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>CS tonic, n = 17</td>
<td>10</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>US buildup, n = 69</td>
<td>6</td>
<td>29</td>
<td>19</td>
</tr>
</tbody>
</table>
US buildup, CS tonic, and recorded with probability-dependent licking movement each type of neuron and reward probability or licking movement.

Comparison of firing property and waveforms between neuron types. We compared the baseline firing rates of the above 3 types of neurons. The baseline firing rates of CS phasic, CS tonic, and US buildup neurons were $3.1 \pm 2.5$, $3.0 \pm 1.9$, and $3.3 \pm 3.1$ (means $\pm$ SD) spikes per second, respectively, and did not differ between neuron types ($P > 0.05$, 1-way ANOVA). We also compared the duration of the waveforms of action potentials (width of negative component at half-maximum). The durations for CS phasic, CS tonic, and US buildup neurons were $205 \pm 31, 223 \pm 45$, and $212 \pm 30$ (means $\pm$ SD) $\mu$s, respectively, and did not differ significantly between neuron types ($P > 0.05$, 1-way ANOVA).

Recording sites. The recording sites of each neuron were reconstructed histologically and superimposed onto coronal sections of the left hemisphere of the standard rat brain atlas (Paxinos and Watson 2005). Figure 14 shows the recording sites of each neuron type. We found that all three neuron types were widely distributed within the dorsal striatum without any specific topographical clustering for any of the three types.

DISCUSSION

In this study, we recorded single-unit activity in the dorsal striatum of head-fixed rats that had been pretrained in a probabilistic Pavlovian conditioning task using auditory cues. The neurons recorded in rats performing this task could be categorized into three types based on their firing patterns. CS phasic neurons showed a phasic response to the CS onset, and the magnitude of this response was positively related to reward probability. The majority of these neurons also showed a phasic reward response for which magnitude was negatively related to reward probability (Fig. 4). Thus many CS phasic neurons showed greater phasic responses to CSs that predicted higher reward probability and showed greater phasic responses to less probable rewards. These firing properties correspond to the firing properties of midbrain dopamine neurons and indicate that a subset of CS phasic neurons code a reward prediction error at both the CS onset and the reward onset. In our previous study (Oyama et al. 2010), we compared this type of neuron with midbrain dopamine neurons and concluded that they have highly similar firing properties. CS tonic neurons showed a tonic response during the CS presentation without a reward response, and the magnitude of this response was positively related to reward probability (Fig. 7). These neurons can be considered to code the value of the stimulus. US buildup neurons showed gradually increasing activity toward the time of reward delivery, and the magnitude of the prereward activity was positively related to reward probability. The firing of this type of neuron may reflect the animal’s internal expectation about the upcoming reward (Fig. 10).

Table 4. Summary of the relationship between the activity of each type of neuron and reward probability or licking movement recorded with probability-dependent licking movement

<table>
<thead>
<tr>
<th>Type</th>
<th>Probability Only</th>
<th>Licking Only</th>
<th>Both</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS phasic, $n = 16$</td>
<td>8</td>
<td>0</td>
<td>2 (2)</td>
</tr>
<tr>
<td>CS tonic, $n = 6$</td>
<td>3</td>
<td>2</td>
<td>0 (0)</td>
</tr>
<tr>
<td>US buildup, $n = 26$</td>
<td>14</td>
<td>1</td>
<td>4 (4)</td>
</tr>
</tbody>
</table>

Numbers in parentheses show the number of neurons that had a higher $r^2$ value for the model including licking movement as a regressor.

For all three types of neurons, we examined how activity changed as the delay duration between CS offset and reward onset was prolonged. As expected, the phasic responses in CS phasic neurons were time-locked to the CS and reward onset, and the tonic response in CS tonic neurons was time-locked to the CS onset. The buildup activity in US buildup neuron continued to peak at the reward onset regardless of the delay extension, but the buildup activity itself was prolonged as the delay was extended. These results confirmed the validity of the interpretation of the function of each neuron type. In general, extension of the time interval between the CS and the reward leads to a decrease of the value of the CS or the value of the reward when measured at the time of the CS, which in behavioral economics is known as “temporal discounting” (Ainslie 1975). It has been reported that value-coding neurons show this devaluing effect induced by delay prolongation (Cai et al. 2011; Kobayashi and Schultz 2008). We found that for CS phasic and CS tonic neurons, the CS responses in the...
high-probability conditions decreased as the delay was prolonged. Given that the reward value predicted by the CS decreases with longer delays, the prediction error that occurs at the actual delivery of reward increases. In addition, the timing of reward is less precise with longer CS-US intervals, which also increases the prediction error (Fiorillo et al. 2008). In accordance with this, we also observed that for CS phasic neurons the reward responses under the high-probability conditions were stronger when the delay was prolonged. When the delay was again set to the original duration, the activity level did not return all the way to the original level: the peak of the CS response in CS phasic and CS tonic neurons and the peak of the buildup activity in US buildup neurons were lower than in the initial session. As the whole procedure of extending the delay duration in two steps and then bringing the delay duration back to the original length took a long time, the motivational level of the animal could have been reduced considerably during the procedure. It is possible that the data reflect not only the effect of time discounting, but also an overall decrease of the motivational level over time.

In the present task, the activity of the majority of the recorded neurons depended on reward probability as indicated by different CSs. This dependency on reward probability implies that the activity was the product of associative learning during the extensive training phase. Previous studies on the synaptic mechanisms in the striatum have shown that long-term potentiation can occur at corticostriatal synapses when the striatal neuron receives both cortical and dopaminergic inputs (Canales et al. 2002; Reynolds et al. 2001; Wickens et al. 1996). In addition, dopamine neurons, which send dense projections to the striatum, fire in response to unexpected rewards, i.e., when a positive reward prediction error occurs (Schultz 1998). These results suggest that when a reward is given after the presentation of a CS, the corticostriatal synapses that transmit the sensory information of the CS would be strengthened. During extensive training, the synapses that transmit the information of a CS indicating higher reward probability, which is more frequently followed by reward, would be strengthened further. As a result of this process, the presentation of a CS indicating higher reward probability would elicit greater striatal activation. This may be the mechanism by which information about stimulus value is acquired. Similarly, the probability-dependent phasic CS response of the reward prediction error-coding neurons may be formed through this process. Moreover, it is conceivable that the firing of neurons coding stimulus value may differentially change the animals’ internal motivational state, the elevation of which would be reflected in the buildup activity of reward expectation-coding neurons toward the time of reward delivery. The reward expectation signal would lead to preparation for appropriately acquiring the reward such as directing attention to the reward and preparing to execute the reward-acquiring action. At the timing of the reward delivery, the reward expectation signal would be used to calculate the reward prediction error signal, which is represented in the phasic response to the reward of the reward prediction error-coding neurons.

According to the firing properties of neurons and the waveform of action potentials (Oyama et al. 2010), it is most likely that we recorded from medium spiny projection neurons, which constitute the vast majority of striatal neurons (Apicella 2007; Oorschot 1996). Our results indicate that within the striatal medium spiny neurons there are discrete functional subtypes that code different aspects of reward. It is known that there are different subpopulations of medium spiny neurons such as neurons belonging to the direct pathway or indirect pathway and neurons located in the patch or matrix. Recent studies using transgenic animals and molecular biological techniques have found that neurons belonging to the direct and indirect pathway have different motor functions (Kravitz et al. 2010) and cognitive functions such as learning (Hikida et al. 2010; Kravitz et al. 2012). These results suggest that striatal neurons with different histochemical properties code different information. However, to understand how a neuron relates to a larger neuronal network and how it functions and interacts with other neurons, we need to investigate the precise morphological and histochemical background of the neuron, including its type, which receptors it expresses, and which other neurons it projects to. Staining a single neuron after having recorded from it (Oyama et al. 2013) during a behavioral paradigm will allow for such histochemical and morphological investigations and may reveal the functions and relationships of discrete subtypes of striatal neurons that code different reward-related information.

In our recorded neurons, only a small population showed activity related to reward uncertainty, which is maximal at a reward probability of 50% and gradually decreases as reward probability becomes smaller or larger [although we found in our previous study (Oyama et al. 2010) that none of the reward prediction error-coding neurons, a subset of CS phasic neurons of this study, showed activity related to uncertainty, that may have been a consequence of our having underestimated the number of neurons that code uncertainty because the statistical method we used was not as powerful as the 1 used in this study]. This suggests that the striatum is preferentially involved in coding parametric reward value rather than reward uncertainty. Such a conclusion is consistent with human imaging findings (Tobler et al. 2008) indicating that striatal activation is dependent on reward probability but not on reward uncertainty in a very similar probabilistic Pavlovian conditioning task. Furthermore, we found that only a small number of neurons showed negative correlations between CS-related activity and reward probability even though positive correlations between CS-related activity and reward probability were substantial and numerous. This suggests that in a probabilistic Pavlovian conditioning paradigm, negative reward value-coding is not common in dorsal striatal neurons. This contrasts with previous studies investigating striatal value representation in monkeys involved in an instrumental task in which 30–60% of task-related neurons coded value negatively (Cromwell and Schultz 2003; Samejima et al. 2005). In addition, our striatal neurons showed neither an increase nor decrease of activity at the time of an unexpected reward omission on neither the population nor single-neuron level. This suggests that negative reward prediction error was not coded by striatal neurons in our task and contrasts with what is known about dopamine neurons, which are known to code both positive and negative prediction errors in monkeys (Schultz 1998) and in rodents (Oyama et al. 2010). On the other hand, a recent study recording from monkeys performing an instrumental conditioning task demonstrated that both positive and negative prediction errors were represented in presumed medium spiny neurons primarily by increases in firing rates (Asaad and
Eskandar 2011). These inconsistencies with previous studies may be attributable to task or species differences.

In a behavioral task in which both reward and punishment were used as unconditioned stimuli, Matsumoto and Hikosaka (2009) claimed that a subpopulation of dopamine neurons encodes general motivational salience rather than value (but see Fiorillo et al. 2013). Their argument raises the possibility that striatal neurons showing probability-dependent activity may reflect the motivational salience but not the value of stimuli and outcomes. The behavioral task we used in this study, however, cannot dissociate value from motivational salience (Kahnt et al. 2014). Therefore, we cannot rule out the possibility that some neurons recorded in this study code motivational salience rather than value. To determine whether the activity of a neuron reflects reward, punishment, or motivational salience, we need to record the activity in a behavioral paradigm in which both reward and punishment are used as unconditioned stimuli.

It has been suggested that the dorsomedial striatum mediates action-outcome learning or goal-directed behavior and that the dorsolateral striatum mediates stimulus-response learning or habitual behavior (Barnes et al. 2005; Jog et al. 1999; Packard and Knowlton 2002; Yin et al. 2004, 2005). In this study, we recorded from both medial and lateral areas and found many active neurons, although we used a Pavlovian paradigm in which the animals were not required to perform any action. Our data suggest that the dorsal striatum is involved not only in goal-directed or habitual behavior, but also in more in general associative learning, including probabilistic Pavlovian conditioning.

GRANTS
This study was funded by Grants-in-Aid for Scientific Research (KAKENHI) #24223004, #24243067, and #19673002 to K.-I. Tsutsui. K. Oyama was supported by Japan Society for the Promotion of Science as a Research Fellow #24223004, #24243067, and #19673002 to K.-I. Tsutsui. K. Oyama was supported by Japan Society for the Promotion of Science as a Research Fellow #24223004, #24243067, and #19673002 to K.-I. Tsutsui. K. Oyama was supported by Japan Society for the Promotion of Science as a Research Fellow #24223004, #24243067, and #19673002 to K.-I. Tsutsui. K. Oyama was supported by Japan Society for the Promotion of Science as a Research Fellow #24223004, #24243067, and #19673002 to K.-I. Tsutsui. K. Oyama was supported by Japan Society for the Promotion of Science as a Research Fellow #24223004, #24243067, and #19673002 to K.-I. Tsutsui. K. Oyama was supported by Japan Society for the Promotion of Science as a Research Fellow #24223004, #24243067, and #19673002 to K.-I. Tsutsui. K. Oyama was supported by Japan Society for the Promotion of Science as a Research Fellow #24223004, #24243067, and #19673002 to K.-I. Tsutsui. K. Oyama was supported by Japan Society for the Promotion of Science as a Research Fellow #24223004, #24243067, and #19673002 to K.-I. Tsutsui.

DISCLOSURES
No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS
K.-I.T. conception and design of research; K.O. performed experiments; K.O. and Y.T. analyzed data; K.O., I.H., P.N.T., and K.-I.T. interpreted results of experiments; K.O. prepared figures; K.O. and K.-I.T. drafted manuscript; I.H., P.N.T., and T.I. edited and revised manuscript; K.-I.T. approved final version of manuscript.

REFERENCES
Asaad WF, Eskandar EN. Encoding of both positive and negative reward prediction errors by neurons of the primate lateral prefrontal cortex and caudate nucleus. J Neurosci 31: 17772–17787, 2011.

Burke CJ, Tobler PN. Coding of reward probability and risk by single neurons in animals. Front Neurosci 5: 121, 2011.
Oorschot DE. Total number of neurons in the neostriatal, pallidal, subthalamie, and substantia nigral nuclei of the rat basal ganglia: a stereological study using the Cavalieri and optical disector methods. J Comp Neurol 366: 580–599, 1996.


