

Reward Unpredictability inside and outside of a Task Context as a Determinant of the Responses of Tonicly Active Neurons in the Monkey Striatum

Sabrina Ravel,¹ Pierangelo Sardo,² Eric Legallet,¹ and Paul Apicella¹

¹Laboratoire de Neurobiologie Cellulaire et Fonctionnelle, Centre National de la Recherche Scientifique, 13402 Marseille Cedex 20, France, and ²Istituto di Fisiologia Umana, Università di Palermo, 90134 Palermo, Italy

Tonicly active neurons (TANs) in the monkey striatum are involved in detecting motivationally relevant stimuli. We recently provided evidence that the timing of conditioned stimuli strongly influences the responsiveness of TANs, the source of which is likely to be the monkey's previous experience with particular temporal regularities in sequential task events. To extend these findings, we investigated the relationship of TAN responses to a primary liquid reward, the timing of which is more or less predictable to the monkey either outside of a task or during instrumental task performance. Reward predictability was indexed by the timing characteristics of the mouth movements. The responsiveness of TANs to reward increased with the range and variability of time periods before reward, notably when the liquid was delivered outside of a task. A change in the temporal order of events in a task context produced an increase

of response to reward, suggesting an influence of the predicted nature of the event in addition to its time of occurrence. By contrast, we observed no substantial changes in neuronal activity at the expected time of reward when this event failed to occur, suggesting that these neurons do not appear to carry information about an error in reward prediction. These results demonstrate that TANs constitute a neuronal system that is involved in detecting unpredicted reward events, irrespective of the specific behavioral situation in which such events occur. The responses influenced by stimulus prediction may constitute a neuronal basis for the notion that striatal processing is crucial for habit learning.

Key words: striatum; basal ganglia; TANs; prediction; reward; learning

Several lines of evidence suggest that the striatum, the main input structure of the basal ganglia, and the ascending dopamine (DA) system from the midbrain are involved in the acquisition and maintenance of reward-mediated behaviors (Graybiel, 1995; Houk et al., 1995; Robbins and Everitt, 1996; Schultz, 1998). Single-neuron recording studies in the striatum of behaving monkeys have demonstrated that a particular group of neurons, known as the tonicly active neurons (TANs), respond to stimuli that are conditioned by association with primary rewards (Kimura et al., 1984; Kimura, 1986; Apicella et al., 1991; Aosaki et al., 1994b; Raz et al., 1996) and to stimuli having inherent appetitive value (Apicella et al., 1997; Ravel et al., 1999). In this respect, TANs may provide a signal that reports rewarding properties of stimuli in a fashion very similar to DA neurons, suggesting that these two neuronal systems are involved in the motivational control of behavior (Aosaki et al., 1994b; Schultz, 1998; Sardo et al., 2000).

In recent work, we found that the majority of TANs respond to the unsignaled delivery of a liquid reward outside the context of a behavioral task, whereas these responses are reduced consider-

ably when the same reward is delivered on correct instrumental responding (Apicella et al., 1997). We interpret this context-dependent activity to be a facilitation of the responses of TANs when animals did not actively control the timing of reward through learned behavioral reactions. Using an instrumental task involving the same stimulus with the same behavioral response contingency, we found that TAN responses to a trigger stimulus for movement are reduced in the presence of an instruction cue given at a fixed time interval before the trigger (Apicella et al., 1998), and the number of neurons responding to the trigger increased when the usual duration of the instruction–trigger interval was changed (Sardo et al., 2000). Although these findings are consistent with the hypothesis that temporal predictability of a conditioned stimulus is a crucial determinant of the responses of TANs, it is possible that the responsiveness depends on the specific learning situation in which the stimulus occurs. In particular, it is equivocal whether the variations of neuronal responsiveness reflect temporal aspects of stimulus prediction or merely an influence of the behavior subjected to particular conditioning contingencies. One of the main purposes of the present experiment was to verify whether the prediction effect would remain present while primary liquid rewards were delivered in two different behavioral states. To this end, we tested the impact of temporal components of reward delivery on the activity of TANs inside and outside of a learned behavioral task to determine whether differences in neuronal responses were related to differences in the predictions about the timing of reward or to differences in the type of associative learning specific to the behavioral situations. Furthermore, because of a potential relationship between TANs and DA neurons of the midbrain, we wanted to

Received March 19, 2001; revised April 25, 2001; accepted May 10, 2001.

This research was supported by Centre National de la Recherche Scientifique, the European Human Capital and Mobility Program (Grant CHRX-CT94-0463), and the Biomed II Program of the European Commission (Grant BMH4-CT95-0608). We thank R. Massarino for expert mechanical work and C. Wirig for designing the electronic device.

Correspondence should be addressed to Paul Apicella, Laboratoire de Neurobiologie Cellulaire et Fonctionnelle, Centre National de la Recherche Scientifique, 31 chemin Joseph Aiguier, 13402 Marseille Cedex 20, France. E-mail: apicella@lncf.cnrs-mrs.fr.

Copyright © 2001 Society for Neuroscience 0270-6474/01/215730-10\$15.00/0

know whether TAN responses may encode errors in prediction of reward.

MATERIALS AND METHODS

Data were collected from two male macaque monkeys (*Macaca fascicularis*, monkeys A and B, 5–6 kg) that were trained in an instrumental task in which they were required to perform visually triggered arm-reaching movements to obtain a liquid reward. We also recorded from the striatum of a third monkey performing another version of the instrumental task that we used with monkeys A and B. The experimental protocol was performed according to the National Institutes of Health *Guide for the Care and Use of Laboratory Animals* and the French laws on animal experimentation.

The monkeys were seated in a restraining box that was described in an earlier publication (Apicella et al., 1997) and faced a panel placed ~30 cm in front of them. A red light-emitting diode (LED) and a contact-sensitive metal knob mounted 10 mm below the LED were located at the center of the panel, at arm's length and at eye level of the animal. A metal bar was situated centrally on the lower part of the panel at waist level of the animal. Depending on the testing condition used, the sliding door located at the front of the box could be opened or closed to allow or prevent manual access to the panel. A drinking tube was positioned directly in front of the monkey's mouth for the delivery of apple juice as a reward. Briefly, a trial started with the monkey keeping its hand on the bar. The animal was required to maintain this contact for a randomly varied period of 0.5–2 sec, after which a red light was illuminated. In response to the presentation of this stimulus, the monkey had to release the bar to reach and touch the knob below the illuminated LED. When the animal touched the target, the light was turned off and a solenoid valve dispensed a small amount of apple juice (0.3 ml) as a reward. After target acquisition, the monkey had to move back to the bar and wait for the total duration of the current trial (7 sec) to elapse before a new trial began. An error trial was recorded when monkeys took longer than 1 sec to initiate or execute the movement. The monkey did not know exactly when the trigger stimulus would occur, because both the length of the intertrial intervals and the delay between the start of a trial and the onset time of the trigger were varied. The solenoid valve was inside a sound-proof container that was located outside the experimental room so that the monkey was not able to hear the click sound emitted by the valve opening. During recording sessions, the monkey's face was continuously monitored using a video camera. The drinking tube was equipped with force transducers (strain gauges) with which the contact between the lips or tongue and the spout was recorded as a behavioral index of the monkey's ability to predict the moment of reward delivery. During the training and recording periods, the monkeys were deprived of water in their home cage and received apple juice during the experiments. Unlimited water access was allowed for at least 1 d each week.

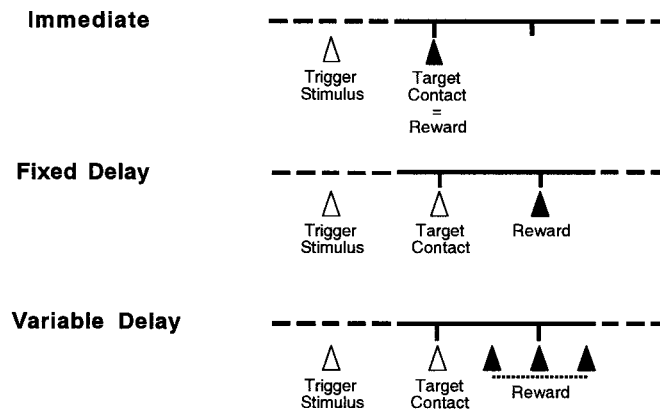
Behavioral procedures

In the present experiments, the monkey's ability to predict the time of reward delivery was manipulated under two behavioral situations: (1) an instrumental task condition in which the monkey performed an arm-reaching movement leading to delivery of reward, and (2) a free reward condition in which the same liquid was delivered outside of a task and in the absence of any external reward-predicting signal. Animals were informed of the change of situation by the experimenter entering the recording room to open (instrumental task condition) or close (free reward condition) the sliding door of the restraining box.

Instrumental task condition. In the instrumental task condition, the monkey obtained reward immediately after contact with the target (Fig. 1). This condition will be referred to as the "immediate condition." A variant of the task, called the "fixed delay condition," was designed to investigate the effect of delaying the delivery of reward after target contact by a fixed interval of 1 sec. In another version of the task, called the "variable delay condition," the time of reward was randomly varied relative to target contact (0.5, 1, and 1.5 sec). Both monkeys had achieved a consistent correct performance rate of >95% in the immediate condition before the neuronal recording started, and we used the delayed reinforcement procedures only during neuronal data collection. All three task conditions were run in separate blocks of 30–40 trials throughout the course of recording sessions, the order of the conditions being counterbalanced.

After recording from monkeys A and B, one additional monkey was trained and then tested in another version of the reaching task that we

Instrumental Task Condition



Free Reward Condition

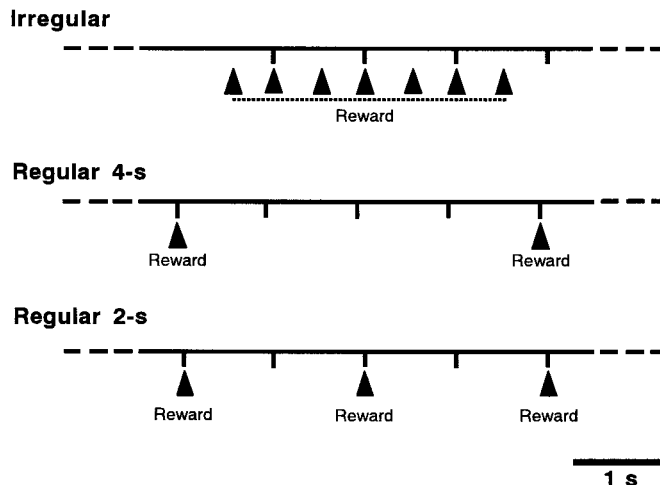


Figure 1. Description of the sequence of events in the testing conditions. **Instrumental Task Condition.** This is a task state in which the monkey awaited the visual stimulus with its hand on the bar. Illumination of the light signaled the monkey to release the bar and reach for the target, which it had to touch to receive a liquid reward. **Free Reward Condition.** This is a no-task state in which the monkey remained motionless and received the same liquid reward in the absence of any external predictive signal. The present study used three versions of each behavioral situation. In the task state, the delivery of reward occurred immediately after contact with the target (*Immediate*) or was delayed after target contact by a constant 1 sec interval (*Fixed Delay*) or a randomly varying interval of 0.5–1.5 sec (*Variable Delay*). In the no-task state, liquid was delivered at irregular time intervals of 5.5–8.5 sec (*Irregular*) or at a constant interval of 4 sec (*Regular 4-s*) or 2 sec (*Regular 2-s*). In each of these two behavioral states, the various conditions were tested in separate blocks of trials.

called the "surprising reward condition." For this condition, exactly the same procedure was used as that used for the immediate condition, except that both the specific times at which stimuli occurred and the nature of stimuli were variable from trial to trial. This condition was designed to compare neuronal activity for trials with reward normally occurring on correct target reaching, as compared with trials on which an unexpected premature reward was automatically delivered at the start of a trial. Each trial began with the monkey maintaining its hand on the bar for a random period of 0.5–2 sec, after which the trigger stimulus was presented 70% of the time (usual trials) and reward occurred 30% of the time (surprising trials) during the same session. Thus, even if one event

was more frequent than the other, the monkey could not be sure which event would occur first on any given trial. When the monkey received the surprising reward, it had to wait for a variable time period while resting its hand on the bar until the total duration of the trial had elapsed (7 sec). The surprising reward condition comprised 80 trials. In a variant of this condition, the monkey did not know exactly what would be the first event of each trial, but it could know when this event would occur because an external cue signaled its onset time at the end of a constant interval. A first visual stimulus (green light, 500 msec duration), serving as a temporal cue, appeared in the center of the screen at the location at which the trigger stimulus (two-colored LED) was presented. This cue began a waiting interval of 1.5 sec that lasted when the next event occurred. At the end of this interval, the trigger stimulus was presented 70% of the time, and reward occurred 30% of the time during the same session. This design combined cueing of the moment of the upcoming event and uncertainty about the nature of the event. This condition comprised 80 trials.

Free reward condition. Monkeys A and B were also subjected to a testing procedure in which the fruit juice was delivered without engaging the animal in any specific task. Delivery of reward was tested by administering the liquid in the absence of any external stimulus predicting its time of delivery (Fig. 1). The time interval between successive liquid deliveries was randomly varied from trial to trial and ranged from 5.5 to 8.5 sec to make it impossible for animals to have reliable temporal information during a run of trials. We refer to this situation as the “irregular condition.” To investigate the influence of the rate of reward delivery, we introduced regularities in inter-reward intervals so that each reward might serve the animal as a time marker from which it could predict the timing of the next reward. In the first condition, the liquid was delivered once every 4 sec. In the second condition, the constant interval was shortened to 2 sec. We refer to these as the “regular 4 sec condition” and the “regular 2 sec condition.” In each condition, the trials were run in separate blocks of 30–40 trials, and the order of the conditions was counterbalanced across recording sessions to ensure that the changes in neuronal responsiveness in these conditions were not caused by order effects.

Classically conditioned task. Monkeys A and B were also trained on a typical Pavlovian conditioning procedure in which a visual stimulus is repeatedly paired with a primary reward. In this condition, the sliding door of the restraining box was closed, and a red light was illuminated at unpredicted times as an external signal preceding the delivery of reward by a fixed interval of 1 sec. The visual stimulus duration was 0.3 sec, and its onset was initiated by the experimenter. Both monkeys received numerous training sessions with this same time interval, and on some occasions, the duration was prolonged to 2 sec to assess the effects of delivering reward beyond the time of its usual occurrence. The monkeys practiced the condition using the long signal–reward interval very infrequently.

Surgical procedures

After completing training on the immediate condition of the instrumental task, each monkey was surgically prepared for chronic single-neuron recording experiments. The surgery was performed in aseptic conditions and under pentobarbital sodium anesthesia (35 mg/kg, i.v.; Sanofi, Libourne, France). A stainless steel chamber (outer diameter, 25 mm) and a head-restraining device were implanted on the skull at the following coordinates according to the Szabo and Cowan (1984) atlas: anteroposterior 18, lateral 6. The position of the chamber was chosen to allow the search for neurons to be performed throughout the caudate nucleus and putamen. The dura mater was left intact. Prophylactic antibiotics (Ampicillin, 17 mg/kg for 12 hr; Bristol-Myers Squibb, Paris, France) were injected intramuscularly on the day of the surgery and for 5 d after the surgery. The recording chamber was filled with an antibiotic solution (flumequine; Sanofi) and sealed with a removable cap.

Neuronal recordings

The experimental methods that were used to record single-neuron activity have been described elsewhere (Apicella et al., 1997). For a recording session, the monkey was placed in the restraining box with its head fixed. Single-neuron recording was performed with glass-coated tungsten electrodes (length of exposed tips, 9–12 μm ; diameter, 3 μm) that were passed inside a guide cannula (outer diameter, 0.6 mm) at the beginning of each session. After penetration of the dura, the electrode was advanced toward the striatum with a hydraulic microdrive (MO-95; Narishige, Tokyo, Japan) until the activity of one neuron was isolated.

Signals from neuronal activity were conventionally amplified, filtered (bandpass, 0.3–1.5 kHz), and converted to digital pulses through a window discriminator. Presentation of visual stimuli, collection of movement parameters, mouth movements and single-neuron activity, and the delivery of liquid reward were controlled by a computer.

In the present study, TANs were easily discriminated from other striatal neurons on the basis of spontaneous discharge rate and spike waveshape (Kimura et al., 1984; Alexander and DeLong, 1985; Kimura, 1986; Hikosaka et al., 1989; Aosaki et al., 1994b; Apicella et al., 1997). As previously mentioned, each of the testing conditions was performed in a block of trials. During the recording of any neuron, neuronal activity in the instrumental task conditions was generally studied first, and, if the isolation could be sustained for a sufficient period of testing, the tests were continued in the free reward conditions. A total of 106 neurons were recorded only in the free reward conditions.

Data analysis

Performance in the instrumental task condition was measured by using the reaction time, which was the time between the onset of the trigger stimulus and release of the bar, and the movement time, which was the time taken to move from the bar to the target. Upper limits of both behavioral parameters were specified (1 sec for each). Behavioral performance was assessed by calculating the median (50th percentile) of reaction and movement times of correct responses for each session. The data for each condition were taken from 30 and 20 sessions in monkeys A and B, respectively, with the exception of the variable condition in monkey A, which comprised 14 sessions. The Mann–Whitney *U* test was used for comparison between task conditions. Signals from the strain gauge circuit (mouth movements) were digitized at 100 Hz and stored into an analog file continuously during each block of trials. The timing characteristics of the mouth movements that animals performed in the different conditions were assessed off-line by single-trial analysis.

Significant changes in neuronal activity were detected on the basis of a Wilcoxon signed rank test (Apicella et al., 1997). Only neurons with statistically significant changes against control activity were counted as responsive. The baseline discharge rate was calculated from the mean firing frequency during the 500 msec before the presentation of the trigger stimulus in the instrumental task condition and before the delivery of reward in the free reward condition. A test window of 100 msec duration was moved in steps of 10 msec, starting at the onset of a particular stimulus event. The onset of a response was taken to be the beginning of the first of five consecutive steps showing a significant difference ($p < 0.05$) as against the average spike discharge rate value calculated during the 500 msec control period. The offset of a response was defined by the first of five consecutive steps in which activity returned to control levels. The magnitude of response was given by counting neuronal impulses during the period of modulation and expressed as percentage below control activity from each neuron showing a significant response. The results concerning neuronal data were analyzed by standard statistical methods described in the text. These included comparisons of fractions of responding neurons and response parameters among the testing conditions.

To give a description of the properties of the whole population of TANs, we summed activity of all neurons tested in the different conditions and made population histograms (Ljungberg et al., 1992; Aosaki et al., 1994b). These histograms were constructed from all TANs recorded in monkeys A and B in each testing condition, independent of individual neuronal responses. For each neuron, a normalized perievent time histogram was obtained by dividing the content of each bin by the number of trials. The population histogram was obtained by averaging all normalized histograms referenced to a particular event.

Histology

Toward the end of the experiment in monkeys A and B, neuronal recording sites were marked with small electrolytic lesions by passing negative currents through the microelectrode (20 μA for 15–20 sec). After the completion of all testing, the monkeys were killed with an overdose of sodium pentobarbital and perfused transcardially with 0.9% saline followed by a fixative (4% paraformaldehyde, pH 7.4 phosphate buffer). Frozen coronal sections (50 μm thick) were cut through the region of the recordings and stained with cresyl violet. The monkey striatum was divided into two territories based on topographic projections from different cortical regions: the associative striatum, which includes the caudate nucleus and anterior putamen, and the sensorimotor striatum, which includes more posterior regions of the putamen. These

two territories are innervated, respectively, by associative cortical areas and by the primary motor and sensory cortical areas (Künzle, 1975, 1977; Selemon and Goldman-Rakic, 1985; Parent, 1990). Differences in distributions of neuronal responses between these two distinct territories of the striatum were determined with the χ^2 test. The recording sites were not localized in the third monkey, which is still used in neurophysiological experiments.

RESULTS

Behavior

Monkeys A and B consistently showed >95% correct task performance in every condition. Reaction and movement time analyses failed to reveal significant differences between conditions in the two monkeys ($p > 0.01$, Mann–Whitney U test), indicating that arm movements were performed with equal speed, regardless of the moment of reward delivery after target contact. As illustrated in Figure 2, the timing characteristics of the mouth movements showed clear differences between conditions. In the instrumental task condition (Fig. 2, left), both monkeys systematically licked the spout immediately before the receipt of reward when a fixed or variable delay period was introduced between the target contact and reward, whereas the licking activity started on correct reaching when the reward event coincided with target contact. In the free reward condition (Fig. 2, right), both monkeys made licking movements continuously before the delivery of reward in the irregular and the regular 4 sec conditions. This differed from the regular 2 sec condition in which licking movements occurred over a short period centered around the time at which the reward was delivered. In this latter condition, visual inspection of the monkey's face by video monitoring showed rhythmic tongue protrusions that were well synchronized with the rate of reward delivery. The restriction of mouth movements to a narrow time range in the immediate condition and the regular 2 sec condition suggests that the temporal structure of these testing conditions provided more exact information as to the time at which the liquid would be delivered, thus making it more predictable.

Neuronal data

In monkeys A and B, we recorded the activity of 179 striatal neurons that were categorized as TANs on the basis of their relatively continuous spontaneous activity as well as their typical extracellularly recorded action potential waveform (see Materials and Methods). Of the 179 neurons studied, 152 responded to reward in at least one of our testing conditions. The responses of TANs consisted of a phasic depression of the tonic firing, followed in some instances by a transient increase in discharge rate. These response patterns agree with those previously reported and distinguish TANs from other neurons in the striatum (Kimura et al., 1984; Aosaki et al., 1994b; Apicella et al., 1997).

Influence of the timing of reward inside and outside of a task context

As it has been previously reported (Apicella et al., 1997), a number of TANs responded to reward that was delivered in a reaching task when the animal's hand contacted the target. In the present report, such responses were seen in 35 of 70 (50%) neurons studied in the immediate condition. All neurons were also tested when reward was dispensed 1 sec after target contact and, in 37 of them, when reward followed target contact with a random interval ranging from 0.5 to 1.5 sec. Responses to reward were observed in 43 of 70 (61%) and 25 of 37 (68%) neurons in the fixed delay and variable delay conditions, respectively. A somewhat higher proportion of neurons responded to reward that

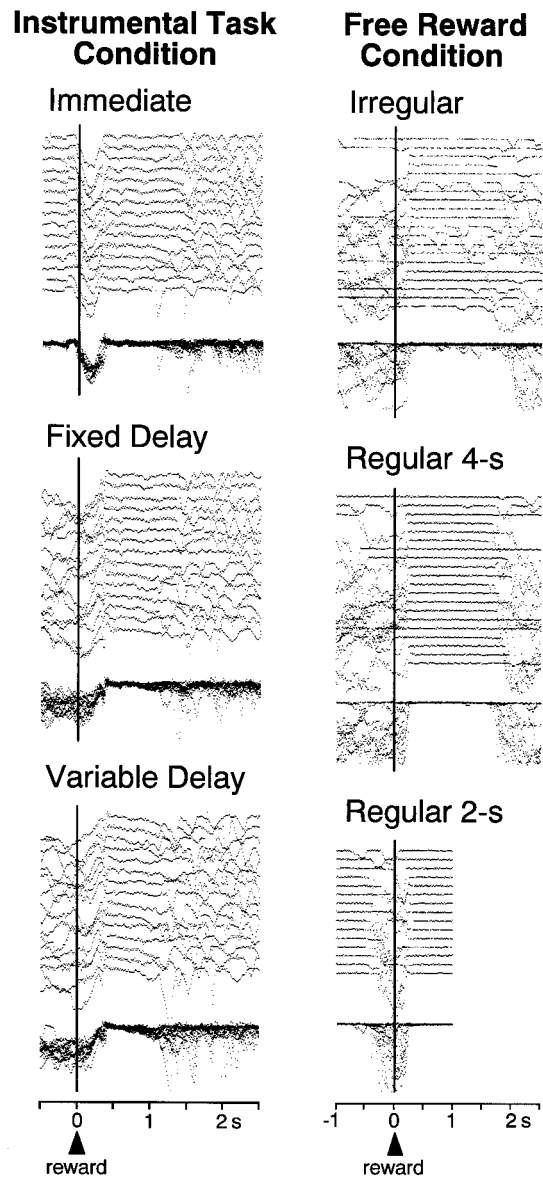


Figure 2. Oral behavior in the testing conditions. Mouth contacts with the spout were detected by means of force transducers (strain gauges). Data from the different variants of each testing condition are ordered from *top* to *bottom*. For each condition are shown consecutive and superimposed traces of mouth movement records aligned on the delivery of reward. Data in each condition are shown from 15–20 consecutive trials, the chronological order of trials being preserved from *top* to *bottom*. The monkey started to lick the spout on the delivery of reward in the immediate condition and maintained oral activity thereafter, whereas licking movements began before the reception of liquid in the two delayed reinforcement procedures. Mouth movements appeared similar in the irregular and regular 4 sec conditions and consisted of frequent, randomly timed licking movements before the delivery of reward. By contrast, in the regular 2 sec condition, licking movements were restricted to a short time period centered on the receipt of reward so that the movements were performed in synchrony with reward delivery. These mouth movements were recorded simultaneously with neuronal activity.

was given at the end of the 1 sec delay after target contact, as compared with the immediate condition, but this difference was not significant ($\chi^2 = 1.85$; $df = 1$; $p > 0.05$). There was also a tendency for the fraction of neurons responding to reward to increase as the length of delay varied, but again this was not significant ($\chi^2 = 3.03$; $df = 1$; $p > 0.05$). Although none of these

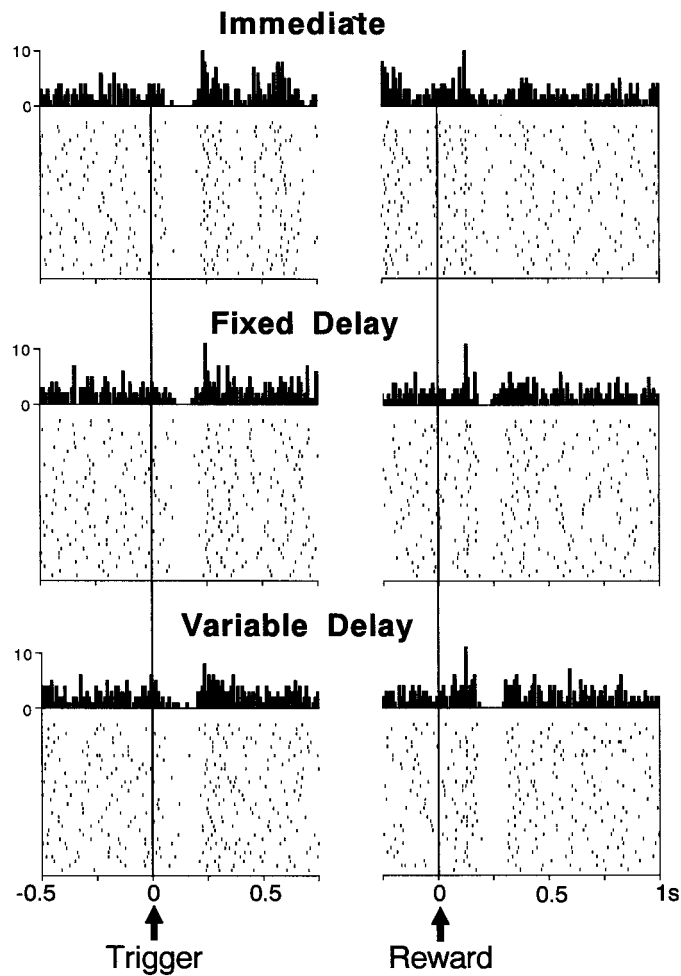


Figure 3. Influence of delaying the delivery of reward in the instrumental task condition in one tonically active neuron. Each dot in the raster displays represents one neuronal impulse, and each line of dots represents the neuronal activity occurring during a single trial. Dot displays and perievent time histograms are aligned on the onset of the trigger stimulus or the delivery of reward, which are marked by vertical lines. The sequence of trials is shown chronologically from top to bottom in each raster display. The change of condition occurred over three successive blocks of trials. This neuron showed a weak response to reward in the immediate condition but had a clear-cut response to the same event in both the fixed delay and variable delay conditions. This neuron also showed a response to the trigger stimulus that persisted in all instances. Vertical calibration on histograms is in impulses per bin. Bin width for histograms is 10 msec.

increases in frequencies of responding neurons reached statistical significance, we observed in some cases marked differences in the responsiveness to reward among the three task conditions. The neuron for which activity is shown in Figure 3 exhibited a weak depression in its activity as the reward was delivered on target contact, although reward responses were enhanced when a fixed or variable delay separated the target contact from the delivery of reward. The same neuron responded to the trigger stimulus, and the responses were unaffected by the relative proximity to the delivery of reward. There were no significant differences in the proportions of neurons showing trigger responses among the three task conditions ($p > 0.05$; χ^2 test).

A high percentage of TANs responded to the delivery of liquid at irregular time intervals outside of a task, consistent with previous findings (Apicella et al., 1997). Of 137 neurons, 105 (77%) showed responses to successive reward deliveries sepa-

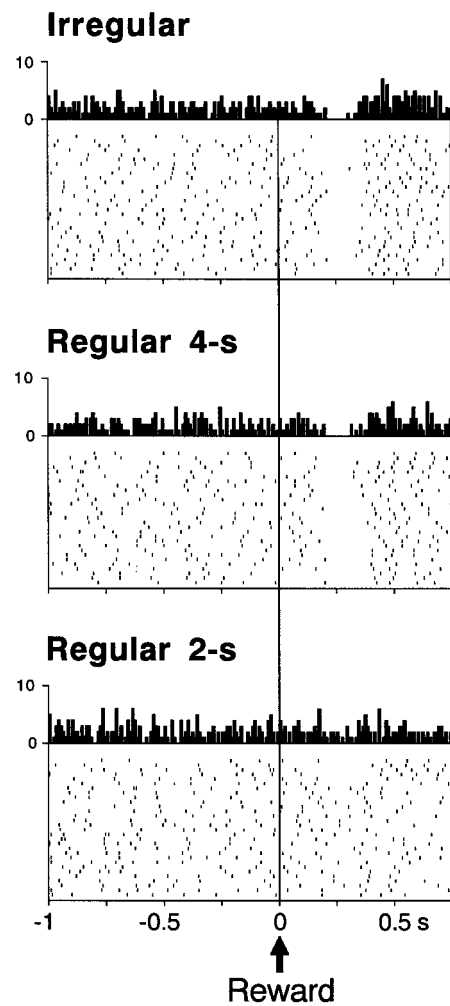


Figure 4. Influence of shortening the interval between consecutive deliveries of liquid in the free reward condition in one tonically active neuron. Conventions are similar to those in Figure 3. The change of condition occurred over three successive blocks of trials. Response to reward occurred in both the irregular and regular 4 sec conditions and disappeared almost immediately when the monkey was switched from the regular 4 sec condition to the regular 2 sec condition.

rated by a randomly varied interval of 5.5–8.5 sec. Among this sample, many neurons were also tested with a constant interval between liquid deliveries. Responses to reward were seen in 97 of 131 (74%) and 74 of 128 (58%) neurons in the regular 4 sec and regular 2 sec conditions, respectively. In contrast to the instrumental task, frequencies of neuronal responses to reward changed significantly over the free reward conditions. The fraction of neurons showing reward responses was significantly higher in both the irregular ($\chi^2 = 10.70$; $df = 1$; $p < 0.01$) and the regular 4 sec ($\chi^2 = 7.60$; $df = 1$; $p < 0.01$) conditions, as compared with the regular 2 sec condition, but not between the irregular and the regular 4 sec conditions ($\chi^2 = 0.24$; $df = 1$; $p > 0.05$). One example of a neuron tested with liquid delivered at different intervals is shown in Figure 4. This neuron gave a response when the liquid came at irregular intervals or at the same 4 sec intervals, although the response disappeared almost immediately with a constant interval of 2 sec.

Table 1 shows the response parameters that were recorded in the various testing conditions. No significant differences were observed in the latency, duration, and magnitude of reward re-

Table 1. Latencies, durations, and magnitudes of reward responses in the two behavioral situations

Behavioral situation	<i>N</i>	Latency (msec)	Duration (msec)	Magnitude (% decrease)
Instrumental task condition				
Immediate	35	171 ± 83	160 ± 74	50 ± 17
Fixed delay	40	188 ± 68	143 ± 74	55 ± 16
Variable delay	23	193 ± 59	126 ± 86	53 ± 20
Free reward condition				
Irregular	92	169 ± 42	147 ± 56	63 ± 18
Regular 4 sec	89	179 ± 37	143 ± 57	60 ± 16
Regular 2 sec	70	175 ± 51	143 ± 64	53 ± 16

Values are means ± SD. *N*, numbers of TANs. Magnitudes are indicated as decreases in percentage below baseline activity during the control period (see Materials and Methods). For quantitative analysis, some recorded neurons were discarded if noise artifact caused by jerky movements by the monkeys contaminated the data, thus explaining the lower number of neurons used for the response parameters analysis, compared with the size of the samples of responding neurons.

sponses among instrumental task conditions (one-way ANOVA followed by Fisher's test, $p > 0.05$). Responses to reward varied insignificantly ($p > 0.05$) in all three free reward conditions in terms of latency and duration, whereas the magnitude of change was significantly greater for both the irregular ($F_{(1,157)} = 6.79$; $p < 0.01$) and regular 4 sec ($F_{(1,155)} = 6.84$; $p < 0.01$) conditions than for the regular 2 sec condition. Thus, both the incidence of responsive neurons and the magnitude of the response decreased when repetitive deliveries of liquid occurred at the same 2 sec intervals.

To determine whether the responsiveness of TANs to reward was influenced by the particular behavioral state in which the liquid was delivered, we compared responses under the variable delay condition to responses under the irregular and regular 4 sec conditions, namely conditions that were supposed to minimize the temporal predictability of reward. No significant differences were observed in the fraction of neurons responding to reward between the variable delay condition and the irregular and regular 4 sec conditions ($p > 0.05$; χ^2 test). On the other hand, magnitudes of reward responses were significantly higher both in the irregular ($F_{(1,111)} = 10.11$; $p < 0.01$) and the regular 4 sec ($F_{(1,108)} = 7.70$; $p < 0.05$) conditions, as compared with the variable delay condition. This indicates that changes in the timing of reward in the context of the instrumental task were somewhat less effective than rewards given outside of a task in terms of response magnitude. This may be because the delivery of reward is more unpredicted in the free reward condition than in the instrumental task condition.

To give an overview of the response properties of the whole population of TANs tested in the two distinct behavioral states, population responses are illustrated in Figure 5. In the instrumental task condition, qualitative inspection of the data revealed that the population average showed a progressive enhancement of the response to reward when passing from the immediate to the fixed delay and variable delay conditions. This shows that reward responses in the delayed reinforcement procedures were sufficiently strong to result in a response of the whole population of TANs recorded in each condition, even if differences between the fraction of neurons showing reward responses and response magnitudes determined individually for each responding neuron in the three conditions did not reach statistical significance. On the other hand, the average response to the trigger stimulus was about the same, regardless of the timing of reward. The relationship between neuronal activity and the liquid delivered in the free reward condition was assessed in a similar manner (Fig. 5). The

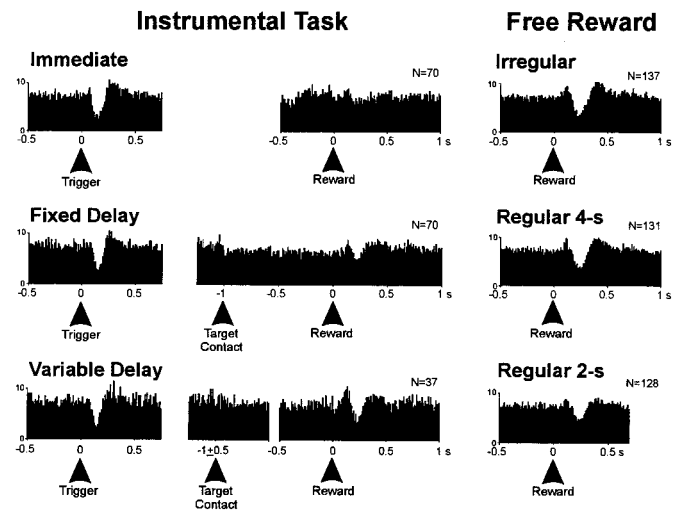


Figure 5. Population responses of tonically active neurons to events occurring in the two behavioral situations. Histograms of the activity of TANs for a sample included responsive and nonresponsive neurons recorded. Because of randomly varying intervals between target contact and reward (1 ± 0.5 sec), the time base in the variable delay condition was split, and histograms were separately referenced to target contact and reward. *N*, Number of neurons included for each population histogram.

population histograms showed higher magnitudes of responses to reward in the irregular and regular 4 sec conditions, as compared with those in the regular 2 sec condition, although neurons responded with equal vigor in the irregular and regular 4 sec conditions. As already pointed out, analyses of the proportions of responding neurons and the response magnitudes substantiated these differences.

Influence of the temporal order of task events

In the experiments reported above, we have tried to impair the monkey's ability to predict the timing of reward during task performance by delaying reward after target contact. However, the monkey could predict that an upcoming reward would follow correct instrumental responses in any case at some point of time because the temporal order of task events remained the same on every trial and every condition. It is possible that the test of the influence of temporal predictability of reward would be made more stringent by changing the event sequence itself. To this aim, we used, for a third monkey, the surprising reward condition in which reward was delivered unexpectedly soon at the beginning of

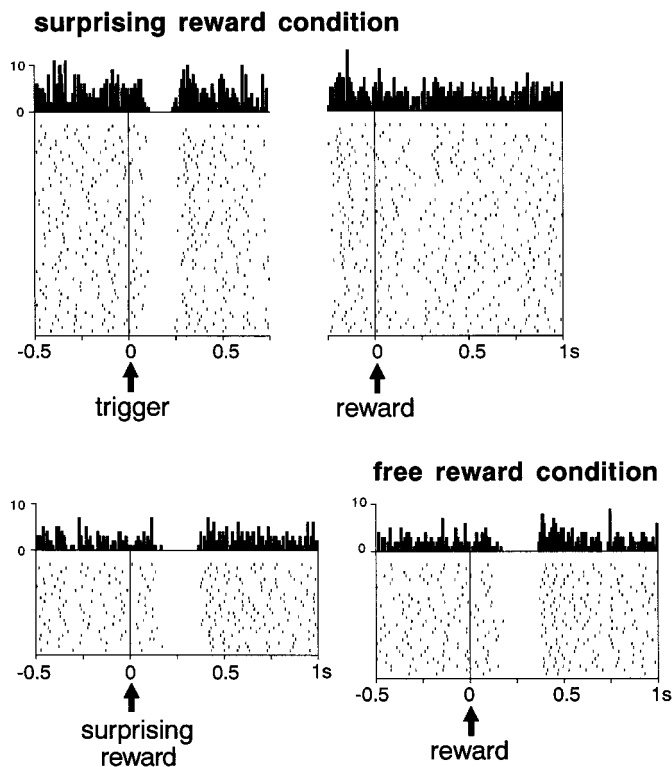


Figure 6. Influence of delivering reward unexpectedly early in the instrumental task condition in one tonically active neuron. Same conventions as in Figure 3. Data were collected during the same test session in which reward was delivered normally after target reaching (usual trials) or unexpectedly soon at the beginning of a trial (surprising trials). The two types of trials occurred pseudorandomly within the same block of trials and were separated off-line for analysis. The raster and histogram display in the lower right panel is for a separate block of trials in which the liquid was given outside of the task at unpredictable intervals (*free reward condition*). This neuron responded in a similar way to reward, both in the surprising and free reward trials.

some trials. This situation was quite different from the immediate condition in that the monkey was not able to predict precisely which event would be presented first on any given trial. Consistent with the notion of a reduced stimulus prediction, behavioral data showed that reaction times in the usual trials of the surprising reward condition were significantly longer (293 ± 41 msec, mean \pm SD) than those in the immediate condition (268 ± 52 msec) ($p < 0.01$, Mann–Whitney U test). Of the 52 neurons tested in this condition, 20 (38%) responded to reward that normally occurred on target contact, whereas 39 (75%) responded to the surprising reward. An example of a neuron showing a strong response to surprising rewards with no response to usual rewards is shown in Figure 6. As can be seen, this same neuron was also tested in the free reward condition with irregular inter-reward intervals, and the response was strikingly similar to the response elicited by the surprising reward. A sample of 12 neurons was tested in the surprising reward condition and the free reward condition, all of them being responsive to reward in both conditions. Magnitudes of reward responses were $-79 \pm 24\%$ and $-74 \pm 26\%$ in the surprising reward and free reward conditions, respectively ($p > 0.05$), demonstrating that depressions in activity related to surprising rewards were quantitatively equal to those that could be elicited by delivering the reward irregularly outside of a task.

Because the surprising reward condition combined the effect of

time uncertainty with the effect of event uncertainty, we added another condition in which a temporal cue pointed to either the trigger stimulus or a reward occurring 1.5 sec later. Among the 52 neurons tested in the surprising reward condition, 12 were tested in the cued one. All of these neurons responded to the surprising reward, regardless of the condition. Quantitative analysis in these neurons revealed that magnitudes of reward responses were significantly higher ($p < 0.01$) in the uncued surprising reward condition ($-86 \pm 12\%$) than in the cued one ($-65 \pm 15\%$), suggesting that the effect of reward uncertainty was weaker in a condition in which an external signal served as a temporal cue allowing for prediction of the time of occurrence of the surprising reward.

Influence of the expected time of reward delivery

Because the responses of TANs appeared to depend critically on the temporal predictability of reward, we were interested in the possibility that these neurons could carry information about the time of the usual occurrence of reward. In such a case, the response properties of TANs would be similar to those reported for the midbrain DA neurons, which are thought to provide a signal reflecting the omission of an expected reward at a particular point in time (Hollerman and Schultz, 1998). As far as could be seen visually by the averaged histograms (Fig. 5), there were no activity changes detectable on target contact when the moment of reward delivery did not match the monkey's prediction. To test specifically the ability of TANs to convey information about the expected time of reward delivery and to exclude possible confounding factors that were linked to target reaching, we used a testing condition lacking arm movement reactions. In a Pavlovian conditioning procedure, a visual signal preceded the delivery of liquid by 1 sec, and the monkeys were trained extensively with that same interval. As reported in an earlier study (Apicella et al., 1997), assessment of licks in this condition revealed that the presentation of the visual signal reliably elicited mouth movements throughout most of the time preceding the delivery of reward. It was reasoned that if the TANs are sensitive to the absence of expected rewards, one would observe a change in firing at the accustomed time of reward. Of the 20 neurons studied with the usual 1 sec delay, 6 (30%) responded to reward. Of these 20 neurons, 17 (85%) demonstrated a response to reward when the delay was prolonged to 2 sec, and only 2 of these neurons had detectable changes in their firing rates at the end of the 1 sec interval, consisting of a moderate yet statistically significant decrease in activity. The neuron shown in Figure 7 was depressed by reward delivered 2 sec after the visual signal, but exhibited no change in firing rate at the usual time of reward, i.e., 1 sec after the visual signal. This demonstrated that TANs are particularly sensitive to changes in the usual duration of the signal–reward interval, although not showing sufficiently differentiated changes in activity to be engaged in the detection of an error in the prediction of reward.

Locations of neurons with reward responses

Histological reconstructions of recording sites of all neurons tested in monkeys A and B are shown in Figure 8. Most of the neurons sampled were located in the dorsomedial regions of the striatum and distributed throughout the mediolateral extent of both the caudate nucleus and putamen. According to anatomical criteria for defining functionally distinct territories (see Materials and Methods), 99 neurons were located in the associative striatum and 73 in the sensorimotor striatum. Only a small number of

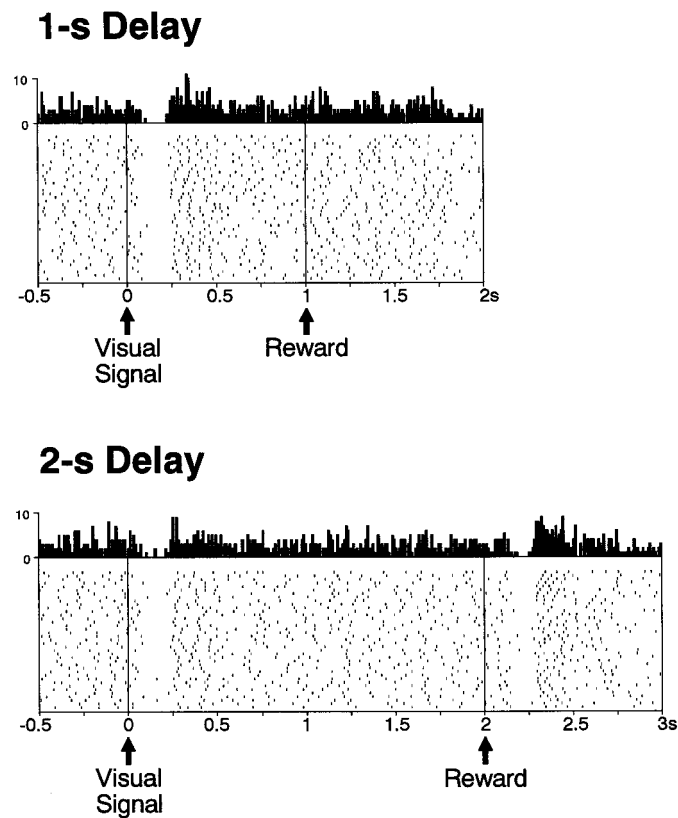


Figure 7. Influence of changing the time of reward delivery in the Pavlovian conditioning procedure in one tonically active neuron. Same conventions as in Figure 3. The neuron was first tested with the usual interval of 1 sec between the visual signal and reward. Then, the neuron was tested with the 2 sec delay. Delivering the liquid at the end of the well practiced 1 sec interval after the visual signal did not elicit a reward response, although such a response appeared when liquid delivery was delayed. Note that no change in firing rate was detectable at the usual time of reward delivery, i.e., 1 sec after the visual signal, when the event occurred later than expected.

neurons ($n = 7$) were recorded in the ventral striatum, which is composed of the nucleus accumbens and adjacent ventromedial caudate and ventral putamen. Reward responses were found in 86 neurons of the associative striatum (87%) and 59 neurons of the sensorimotor striatum (81%). Incidences of reward responses did not vary significantly between these two striatal territories ($p > 0.05$; χ^2 test), confirming previous results (Apicella et al., 1997). Responses that were dependent on the timing of reward were found in 55 neurons of the associative striatum (56%) and 34 neurons of the sensorimotor striatum (47%). The findings go further to provide evidence that response modulations related to reward predictability were not significantly different for the two territories of the striatum that were investigated ($p > 0.05$; χ^2 test).

DISCUSSION

Our initial study had shown that the responses of TANs to reward were reduced or absent in an instrumental task condition in which reinforcement followed a movement, although stronger responding was observed in relation to the same reward given without external cues (Apicella et al., 1997). The present study brings the importance of temporal factors into greater focus by investigating the impact of changes in the timing of reward either outside of a task or during instrumental task performance. The data support

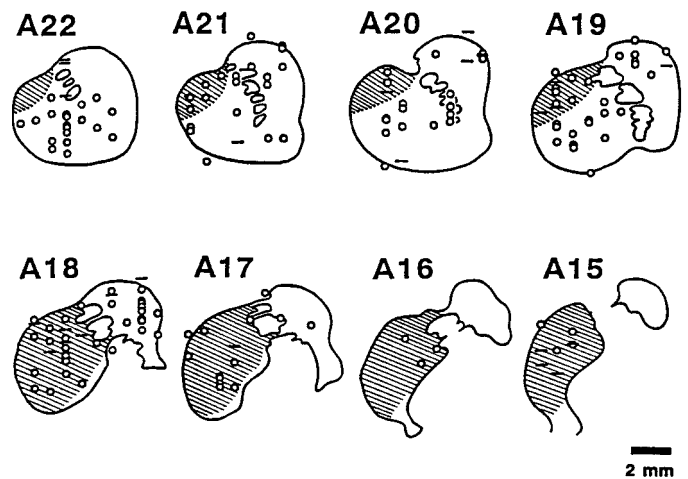


Figure 8. Locations of tonically active neurons with reward responses. Data from monkeys A and B were drawn at corresponding positions on coronal sections of the striatum. Anteroposterior levels are shown according to the atlas of Cowan and Szabo (1984). Striped areas represent approximate extent of the sensorimotor striatum, whereas white areas correspond to the associative striatum (see Materials and Methods). Open circles represent neurons responding to the delivery of reward; small horizontal lines represent neurons not responding to reward. There were no significant differences in the fractions of responsive neurons in the sensorimotor striatum as compared with the associative striatum.

the notion that TANs were really interested in reward events occurring at unpredictable times, the prediction effect on neuronal responsiveness being not constrained by the conditional structure of the testing situation in which reward is delivered. These findings complement and extend previous research in our laboratory indicating that the sensitivity of TANs to a trigger stimulus for movement depends critically on the temporal predictability of this event (Apicella et al., 1998; Sardo et al., 2000). The crucial information determining the response of TANs appears to be the affective value of stimuli, irrespective of the particular behavioral processes associated with the stimuli, such as initiation of learned movements or detection of primary rewards. We further demonstrate that temporal variations in stimulus occurrence are found to modulate the responsiveness of these neurons to appetitive motivating stimuli.

Reward unpredictability in the context of an instrumental task

Although the proportions of neurons responding to reward tended to increase when monkeys were exposed to delayed reinforcement, as compared with the condition in which reward delivery immediately followed behavioral responses, none of these increases reached statistical significance. However, it appears that the average response of all neurons that were recorded in the delayed reinforcement procedures was sufficiently strong to result in a net population response to reward, notably when the timing of reward was randomly varied relative to target reaching. The finding that TANs are relatively poorly related to delayed reinforcement, in terms of fraction of neurons responding, is probably attributable to the fact that performance levels and circumstances remain well organized and may allow for a more general level of reward predictability, at least with the range of target–reward intervals used in this report. This interpretation is strengthened by the finding that TANs were particularly responsive to reward in the framework of the instrumental task for trials in which reward was delivered sooner than expected. In the

surprising reward condition, we attempted to specify whether the increase in responsiveness reflects a lack of expectation for a probable event or a probable time. We found that the surprising presence of rewards produced an increase of the responses to reward, even when the monkey was cued to attend to the upcoming event at a particular time. This finding expands on our earlier findings by demonstrating that the timing of reward alone may not be the sole basis for modulation of TAN responses. It remains an open question as to how the effects of time uncertainty and event uncertainty might act together.

Another explanation for the effects of the timing of reward is that a difference in the monkey's level of attention to the relation between successive events may bring about the difference in the neuronal responsiveness. Because attention and prediction are difficult to untangle in the present experiments, further experimental work is required to examine the influence of each of the single factors separately.

Influence of the range of time intervals between reward events

Considerations of temporal range may provide an important clue for explaining the changes in the responsiveness of TANs to reward; the closer the stimuli are in time during testing, the weaker the responses to these stimuli are presumed to be. We found that the responses to reward delivered at a constant rate outside of a task were much more prevalent and of greater magnitude when the liquid was delivered at an interval of 4 sec than when it was given at a 2 sec interval on successive trials. As evidenced by the monkey's overt oral behavior, animals could only synchronize mouth movements with the rate of reward delivery in the regular 2 sec condition, suggesting that they appeared to take advantage of this short interval to predict the moment of the expected reception of reward. In this condition, the monkey used the delivery of reward as a temporal cue indicating that another reward would be delivered at a particular time point, whereas the animal was not able to retain this information to influence the timing of mouth movements when the inter-reward interval was longer than 2 sec. Analysis of the effect of the duration of constant intervals between a predictive cue and a movement-triggering stimulus on reaction time performance and the responses of TANs to this stimulus in our previous experiments (Sardo et al., 2000) has shown that the prediction effect was apparently confined to a certain range of time intervals and TANs did not discriminate among these intervals very well. In our study, it remains to be established with what degree stimulus unpredictability may depend on time perception processes to estimate intervals between successive events and on the monkey's experience with a particular pattern of events at fixed time intervals.

It must be emphasized that the responsiveness of TANs to primary liquid rewards has varied between different reports, according to the nature of the experimental designs used and to the monkey's experience with temporal regularities in sequential events. At first glance, our results seem to be contradictory to the results of Kimura and coworkers, claiming that TANs do not respond to a liquid given outside of a task (Kimura et al., 1990; Aosaki et al., 1994b). The apparent failure to find a response in these studies may relate to the fact that, under a temporally constant condition, the monkey probably would be able to find some cues in the context of the testing situation that would provide a basis for predicting the timing of reward.

Influence of the time of expected reward

Another new finding reported here is that TANs showed little or no differences in activity at the moment when the reward was expected, suggesting that they do not emit a signal that reports errors in prediction of reward. Although there is evidence of a role of DA inputs in the expression of the TAN responses to reward-related stimuli (Aosaki et al., 1994a; Raz et al., 1996; Watanabe and Kimura, 1998), the lack of firing rate modulation at the usual time of reward suggests that the information processing of TANs differs from that of midbrain DA neurons. As demonstrated by Schultz and coworkers, DA neurons have a specific capacity for the generation of an error signal when the time of reward delivery does not match the monkey's prediction (Hollerman and Schultz, 1998), and this response property may be crucial for reinforcement learning (Montague et al., 1996; Schultz et al., 1997; Schultz, 1998). Our results indicate that TANs do not emit a reward prediction error signal similar to DA neurons. However, because our analysis was confined to firing rate modulations of individual neurons, we cannot exclude that activity synchronization between simultaneously recorded TANs (Raz et al., 1996) can be used by the system for coding of prediction errors.

To summarize, we have shown that the responsiveness of TANs recorded in wide areas of the primate striatum is dependent on the temporal structure of a series of events, including primary rewards. Moreover, TANs do not appear to encode an error in reward prediction, suggesting that their response properties are not similar to those described on DA neurons. There has been considerable interest in possible roles for TANs in the acquisition and maintenance of reward-mediated behaviors (Graybiel, 1995). We suggest that this physiologically homogeneous population of striatal neurons could be specialized for learning about the temporal relationship among external cues and events. Because TANs are thought to correspond to cholinergic local circuit neurons (Kimura et al., 1990; Wilson et al., 1990; Bennett and Wilson, 1999), they may modulate the impact of events on the activity of projection neurons in the striatum depending on their predictability in time. This provides a basis for a general principle that could underlie striatal processing that is crucial for the performance of goal-directed behaviors executed in an automatic fashion, with a minimum of attention to the temporal relationship between successive events that occur in a predictable manner. This consideration is especially important in light of the hypothesized role of the striatum in processes underlying the learning of skills and habits (Mishkin et al., 1984; Graybiel, 1995; Salmon and Butters, 1995; Knowlton et al., 1996; Teng et al., 2000). Although recent evidence suggests that DA (Aosaki et al., 1994a; Raz et al., 1996; Watanabe and Kimura, 1998) and thalamic (Matsumoto et al., 2001) influences on TANs are necessary in order for these neurons to express their typical pause response to motivationally relevant stimuli, further research is needed to characterize the neuronal systems providing input to the TANs that are responsible for monitoring the current temporal context in which such stimuli occur.

REFERENCES

- Alexander GE, DeLong MR (1985) Microstimulation of the primate neostriatum. II. Somatotopic organization of striatal microexcitable zones and their relation to neuronal response properties. *J Neurophysiol* 53:1417–1430.
- Aosaki T, Graybiel AM, Kimura M (1994a) Effect of the nigrostriatal dopamine system on acquired neural responses in the striatum of behaving monkeys. *Science* 265:412–415.
- Aosaki T, Tsubokawa H, Ishida A, Watanabe K, Graybiel AM, Kimura M

- (1994b) Responses of tonically active neurons in the primate's striatum undergo systematic changes during behavioral sensorimotor conditioning. *J Neurosci* 14:3969–3984.
- Apicella P, Schultz W, Scarnati E (1991) Tonicly discharging neurons of monkey striatum respond to preparatory and rewarding stimuli. *Exp Brain Res* 84:672–675.
- Apicella P, Legallet E, Trouche E (1997) Responses of tonically discharging neurons in the monkey striatum to primary rewards delivered during different behavioral states. *Exp Brain Res* 116:456–466.
- Apicella P, Ravel S, Sardo P, Legallet E (1998) Influence of predictive information on responses of tonically active neurons in the monkey striatum. *J Neurophysiol* 80:3341–3344.
- Bennett BD, Wilson CJ (1999) Spontaneous activity of neostriatal cholinergic interneurons *in vitro*. *J Neurosci* 19:5586–5596.
- Graybiel AM (1995) Building action repertoires: memory and learning functions of the basal ganglia. *Curr Opin Neurobiol* 5:733–741.
- Hikosaka O, Sakamoto M, Usui S (1989) Functional properties of monkey caudate neurons. I. Activities related to saccadic eye movements. *J Neurophysiol* 61:780–798.
- Hollerman JR, Schultz W (1998) Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat Neurosci* 1:304–309.
- Houk JC, Adams JL, Barto AG (1995) A model of how the basal ganglia generate and use neural signals that predict reinforcement. In: *Models of information processing in the basal ganglia* (Houk JC, Davis JL, Beiser DG, eds), pp 249–270. Cambridge, MA: MIT.
- Kimura M (1986) The role of primate putamen neurons in the association of sensory stimuli with movement. *Neurosci Res* 3:436–443.
- Kimura M, Rajkowski J, Evarts EV (1984) Tonicly discharging putamen neurons exhibit set dependent responses. *Proc Natl Acad Sci USA* 81:4998–5001.
- Kimura M, Kato M, Shimazaki H (1990) Physiological properties of projection neurons in the monkey striatum to the globus pallidus. *Exp Brain Res* 82:672–676.
- Knowlton BJ, Mangels JA, Squire LR (1996) A neostriatal habit learning system in humans. *Science* 273:1399–1402.
- Künzle H (1975) Bilateral projections from precentral motor cortex to the putamen and other parts of the basal ganglia. An autoradiographic study in *Macaca fascicularis*. *Brain Res* 88:195–209.
- Künzle H (1977) Projections from the primary somatosensory cortex to basal ganglia and thalamus in the monkey. *Exp Brain Res* 30:481–492.
- Ljungberg T, Apicella P, Schultz W (1992) Responses of monkey dopamine neurons during learning of behavioral reactions. *J Neurophysiol* 67:145–163.
- Matsumoto N, Minamimoto T, Graybiel AM, Kimura M (2001) Neurons in the thalamic CM-Pf complex supply striatal neurons with information about behaviorally significant sensory events. *J Neurophysiol* 85:960–976.
- Mishkin M, Malamut B, Bachevalier J (1984) Memories and habits: two neural systems. In: *Neurobiology of learning and memory* (Lynch G, McGaugh JL, Weinberger NM, eds), pp 65–77. New York: Guilford.
- Montague PR, Dayan P, Sejnowski TJ (1996) A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci* 16:1936–1947.
- Parent A (1990) Extrinsic connections of the basal ganglia. *Trends Neurosci* 13:254–258.
- Ravel S, Legallet E, Apicella P (1999) Tonicly active neurons in the monkey striatum do not preferentially respond to appetitive stimuli. *Exp Brain Res* 128:531–534.
- Raz A, Feingold A, Zelanskaya V, Vaadia E, Bergman H (1996) Neuronal synchronization of tonically active neurons in the striatum of normal and parkinsonian primates. *J Neurophysiol* 76:2083–2088.
- Robbins TW, Everitt BJ (1996) Neurobehavioral mechanisms of reward and motivation. *Curr Opin Neurobiol* 6:228–236.
- Salmon DP, Butters N (1995) Neurobiology of skill and habit learning. *Curr Opin Neurobiol* 5:184–190.
- Sardo P, Ravel S, Legallet E, Apicella P (2000) Influence of the predicted time of stimuli eliciting movements on responses of tonically active neurons in the monkey striatum. *Eur J Neurosci* 12:1801–1816.
- Schultz W (1998) Predictive reward signal of dopamine neurons. *J Neurophysiol* 80:1–27.
- Schultz W, Dayan P, Montague PR (1997) A neural substrate of prediction and reward. *Science* 275:1593–1599.
- Selemon LD, Goldman-Rakic PS (1985) Longitudinal topography and interdigitation of corticostriatal projections in the rhesus monkey. *J Neurosci* 5:776–794.
- Szabo J, Cowan WM (1984) A stereotaxic atlas of the brain of the cynomolgus monkey (*Macaca fascicularis*). *J Comp Neurol* 222:265–300.
- Teng E, Stefanacci L, Squire LR, Zola SM (2000) Contrasting effects on discrimination learning after hippocampal lesions and conjoint hippocampal–caudate lesions in monkeys. *J Neurosci* 20:3853–3863.
- Watanabe K, Kimura M (1998) Dopamine receptor-mediated mechanisms involved in the expression of learned activity of primate striatal neurons. *J Neurophysiol* 79:2568–2580.
- Wilson CJ, Chang HT, Kitai ST (1990) Firing patterns and synaptic potentials of identified giant aspiny interneurons in the rat neostriatum. *J Neurosci* 10:508–519.