

Temporal Coding of Reward Value in Monkey Ventral Striatal Tonicly Active Neurons

Rossella Falcone,^{1*} David B. Weintraub,^{2*} Tsuyoshi Setogawa,¹ John H. Wittig Jr.,² Gang Chen,³ and Barry J. Richmond¹

¹Laboratory of Neuropsychology, National Institute of Mental Health (NIMH), Bethesda, Maryland 20892, ²Surgical Neurology Branch, National Institute of Neurological Disorders and Stroke, Bethesda, Maryland 20824, and ³Scientific and Statistical Computing Core, NIMH/NIH/Department of Health and Human Services (DHHS), Bethesda, Maryland 20892

The rostromediovventral striatum is critical for behavior dependent on evaluating rewards. We asked what contribution tonically active neurons (TANs), the putative striatal cholinergic interneurons, make in coding reward value in this part of the striatum. Two female monkeys were given the option to accept or reject an offered reward in each trial, the value of which was signaled by a visual cue. Forty-five percent of the TANs use temporally modulated activity to encode information about discounted value. These responses were significantly better represented using principal component analysis than by just counting spikes. The temporal coding is straightforward: the spikes are distributed according to a sinusoidal envelope of activity that changes gain, ranging from positive to negative according to discounted value. Our results show that the information about the relative value of an offered reward is temporally encoded in neural spike trains of TANs. This temporal coding may allow well tuned, coordinated behavior to emerge.

Key words: monkey; striatum; tonically active neurons

Significance Statement

Ever since the discovery that neurons use trains of pulses to transmit information, it seemed self-evident that information would be encoded into the pattern of the spikes. However, there is not much evidence that spike patterns encode cognitive information. We find that a set of interneurons, the tonically active neurons (TANs) in monkeys' striatum, use temporal patterns of response to encode information about the discounted value of offered rewards. The code seems straightforward: a sinusoidal envelope that changes gain according to the discounted value of the offer, describes the rate of spiking across time. This temporal modulation may provide a means to synchronize these interneurons and the activity of other neural elements including principal output neurons.

Introduction

This laboratory has been studying the encoding of information about reward value in a group of interconnected cortical brain regions in monkeys, including the rhinal cortex and the orbito-frontal cortex for evaluating the reward value of external factors such as visual cues (Simmons et al., 2007; Bouret and Richmond,

2010; Clark et al., 2013; Eldridge et al., 2016). These cortical regions project strongly into the striatum, particularly to the rostral part of the mediovventral part of the striatum (Haber and McFarland, 1999; Averbeck et al., 2014). So, this part of the striatum is categorized as limbic striatum. The rostromediovventral striatum is critical for reward processing (Shidara et al., 1998; Simmons et al., 2007). Inactivation of this area, either through lesions or reversible chemogenetic methods, interferes with judgment of the values associated with familiar visual images (Nagai et al., 2016; Rothenhoefer et al., 2017).

To explore the role of the rostromediovventral striatum further, we recorded single neurons from this region. Neurons in this part of the striatum encode temporally discounted value (Cai et al., 2011). We were interested in characterizing how the responses of the tonically active neurons (TANs), the putative striatal cholinergic interneurons (Wilson et al., 1990; Aosaki et al., 1995), would change across different predicted values, where manipulating size and temporal discounting changed the reward's

Received April 16, 2019; revised July 15, 2019; accepted July 18, 2019.

Author contributions: D.B.W. and B.J.R. designed research; D.B.W. and J.H.W. performed research; R.F., T.S., and B.J.R. analyzed data; G.C. contributed unpublished reagents/analytic tools; R.F. and B.J.R. wrote the paper.

This work was supported (in part) by the Intramural Research Program of the NIMH and includes the relevant Annual Report number in the following format (ZIAMH002619). We thank Mark Eldridge and Janita Turchi for comments on the paper. The opinions expressed in this article are the authors' own and do not reflect the views of the NIH, the DHHS, or the United States government.

The authors declare no competing financial interests.

*R.F. and D.B.W. contributed equally to this work.

Correspondence should be addressed to Barry J. Richmond at bjr@ln.nimh.nih.gov.

<https://doi.org/10.1523/JNEUROSCI.0869-19.2019>

Copyright © 2019 the authors

discounted value. The TANs are especially interesting because they are the main source of acetylcholine in the striatum, and because, despite their rarity (they account for 1–2% of striatal neurons; Graveland and DiFiglia, 1985), each one projects to a wide area of the striatum, and each is thought to contact a large number of medium spiny neurons, the striatal output neurons (Izzo and Bolam, 1988; Contant et al., 1996). Their release of acetylcholine also seems to have a large effect on local striatal dopamine release (Cachope et al., 2012; Threlfell et al., 2012). Finally, cholinergic blockers alleviate some symptoms in patients with Parkinson's disease.

Most of what we know about the activity of TANs in monkeys comes from studies of the dorsal part of the striatum (Aosaki et al., 1994; Apicella et al., 1997, 2011; Ravel et al., 2006; Nougaret and Ravel, 2015). To date there seems to have been only one recording study of TANs in monkeys that specifically included ventral striatum (Marche et al., 2017). We determined how the discounted value of rewards is coded by TANs in the rostromediodorsal striatum, using a visually-cued multi-drop/multi-delay reward task. All of the TANs in our monkeys carried information about the discounted value of the reward. For half of the TANs, the neural responses were modulated in an unexpected manner; the pause and pulse scaled together as a function of the discounted value. That is, the envelope of the neuronal spiking looked sinusoidal, with the sinusoid changing gain according to the discounted value of the offered reward. Thus, the rostromediodorsal striatal TANs encode the value of the predicted reward, and they do so with one of two types of temporal coding: a temporally localized scaling of a pulse, or a temporally localized sinusoidal modulation of the number of spikes.

Materials and Methods

Subjects

Two female rhesus monkeys, both weighing 7.0 kg were used. The experimental procedures followed the Institute of Laboratory Animal Research *Guide for the Care and Use of Laboratory Animals* and were approved by the NIMH Animal Care and Use Committee.

Surgical procedures

Surgical procedures were performed using sterile technique under general anesthesia in a fully equipped operating room. An initial surgery was performed to place a head fixation post as well as scleral eye coil. After several weeks of recovery, a 3.0 T MR image was obtained to determine the location of the targeted area of the striatum and a second sterile surgical procedure was performed to place the recording well. Additional MR scans were obtained later, with tungsten electrodes at recording sites to define the recording area (see Fig. 3A).

Data collection

A noncommercial software package, REX (Hays et al., 1982; <https://datashare.nei.nih.gov/LsrDirectoryServlet>), was used to control stimuli presentation, reward delivery, bar touches, and to monitor the eye position. We monitored the eye position using a magnetic search coil (Judge et al., 1980). Neural activity of single units was recorded extracellularly with tungsten microelectrodes (impedance, 1.5 M Ω ; FHC or Microprobe). A guide tube was positioned using a stereotaxic plastic insert with holes 1 mm apart in a rectangular grid (Crist Instruments; <http://www.cristinstrument.com/>). The electrode was inserted through the guide tube. Electrical signals were captured, and single units were isolated online with a TDT system (Tucker-Davis Technologies). Data analyses were performed in the R statistical computing environment (R Foundation for Statistical Computing, Team, 2018) and in MATLAB (R2014b, MathWorks).

Experimental design and statistical analysis

Task. The monkey sat in a primate chair, with its head fixed, in front of a monitor (47 × 30 cm; 1024 × 768 pixel; 120 Hz) located 57 cm from the

front of the chair. A touch-sensitive bar was mounted on the chair at the level of the monkey's hands. Both monkeys used the right hand to grasp the bar, and the left hand was resting on top of the right hand. A metal tube was positioned in front of the monkey's lips to allow the monkey to drink when a reward was delivered.

Figure 1A shows the sequence of events in the multi-drop/multi-delay reward task used here. The trial started when the monkey touched the bar. After 500 ms a red square (0.40° × 0.40°) appeared at the center of the screen. When the monkey's eye positions were inside a virtual 20 × 20° window, 1 of 9 possible visual cues (9.2° × 7.3°) appeared for 1 of 5 durations between 750 and 2000 ms (step size 313 ms). Each visual cue indicated a particular offer, with the set of offers made by combining one of three reward sizes (2, 4, or 6 drops of liquid reward) with one of three delays to deliver the reward (1, 4, or 7 s). The monkey was required to continue to hold the bar and maintain its gaze within the virtual window until the red square changed color to purple or yellow with equal probability (Fig. 1A). Releasing the bar while the purple spot was present indicated that the monkey was accepting the offer. Releasing the bar when the yellow square was present indicated that the monkey was refusing the offer. The purple or yellow could appear first, and after 600 ms the other color, yellow or purple would replace it if the monkey had not released the bar during the initial purple or yellow period. In this design the monkey could not anticipate whether the acceptance or refusal period would be the first interval or the second, preventing the monkey from preprogramming the motor response in the cue period.

Upon accepting the offer, the purple square turned green (feedback signal) and the drops of reward were delivered after the delay indicated in the offer. Upon refusing the offer, the cue and fixation target disappeared, and a new trial would begin with a new offer.

If the monkey (1) released the bar during the cue period, (2) did not maintain gaze in the virtual window, or (3) failed to release the bar during any of the go periods, an error was registered and an error cue (white rectangle with a red cross; 28.7° × 17.8°) appeared on the screen for 1000 ms. After an error, the next trial was a correction trial, during which the offer was repeated.

Each monkey was tested using two visual cue sets (Set 1 and Set 2 for Monkey 1 and Set 1 and Set 3 for Monkey 2; Fig. 1B). All the analyses were performed using data from both visual cue sets.

Neurophysiology. The neurons were classified as TANs based on their discharge rate and the interspike interval (ISI) distributions (Kimura et al., 1990; Apicella, 2002). ISI and discharge rates were calculated from all the data for each recorded neuron, without regard to individual behavioral events. All TANs had a discharge rate ranged from 2 to 8 spikes/s ($N = 50$; $M \pm SD$: 4.4 spikes/s \pm 1.5). The mean of the ISI distributions for the neural recordings ranged from 127 to 439 ms ($N = 50$; $M \pm SD$: 252 \pm 85 ms).

Neural activity was analyzed using data from, 75–750 ms after the cue appeared, labeled “Cue Period”. This period ended before the monkey received the instruction about how to respond, i.e., whether the purple (accept) or yellow (refuse) spot appeared.

To be included in the analyses the cell had to be recorded for at least 10 completed trials for each of the 9 offers.

Principal component extraction. For each cell, the response in the period from 75 to 750 ms after cue appearance for every trial was low-pass filtered with a Gaussian pulse ($\sigma = 50$ ms) and resampled at 25 ms intervals, giving 28 samples. Then, the Karhunen–Loeve transform was applied to the responses from all conditions, thereby extracting a set of 28 principal components for each neuron and the coefficients for each trial (Richmond and Optican, 1987).

The shapes of the principal components represent deviation through time from the average spike density. By construction the first principal component represents the best possible linear measure of the response pattern through time. The coefficient of the first principal component represents the gain of the pattern of activity represented by the first principal component for each trial. The coefficient of the first principal component was used as the dependent variable in our analyses. In the ANOVAs there was a coefficient for each trial or response.

To quantify the relation between the neural response and the discounted value, we compared the discounted values to the mean coefficient

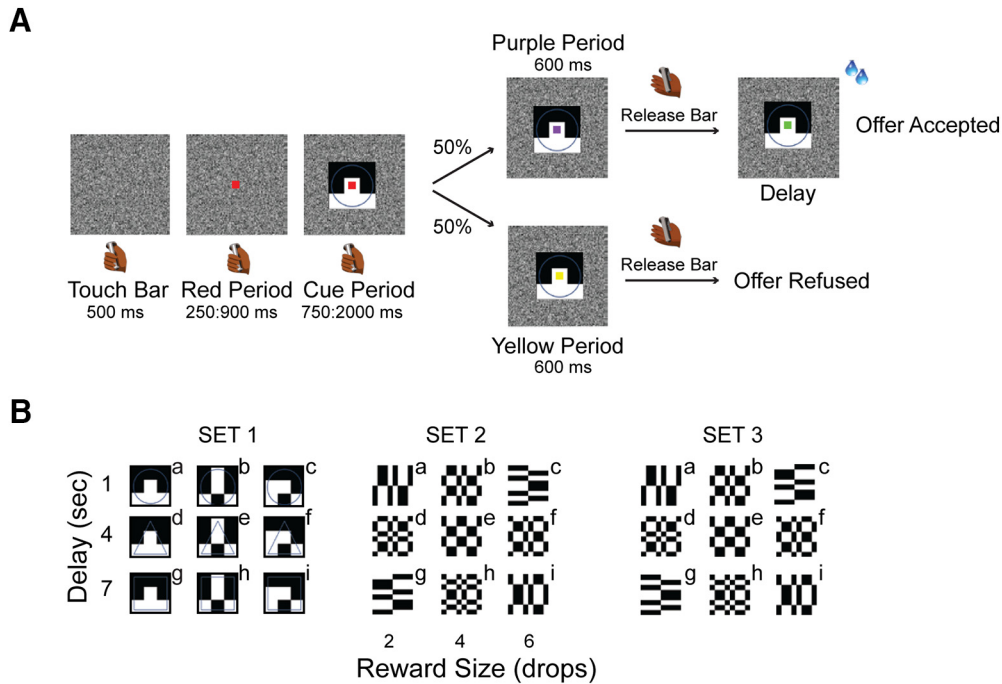


Figure 1. Task description. **A**, The monkey begins a trial by touching a bar. A red square appears first, followed by a visual cue behind the red square. The cue signals the type of offer, made by a combination of reward size (2, 4, or 6 drops) and delay (1, 4, or 7 s). Next, the red square turns purple or yellow, and serves as a go-signal. When the square is purple, the monkey can release the bar to accept the offer indicated by the cue. The purple square turns green and the reward indicated by the cue is delivered. When the square is yellow, the monkey can release the bar to refuse the offer and the screen goes blank and the monkey can start a new trial. If the animal does not release the bar, the square changes again to yellow or purple, and again, the monkey can respond according to the square color. **B**, Sets of visual cues used (Sets 1 and 2 for Monkey 1; Sets 1 and 3 for Monkey 2). The letters from “a” to “i” indicate the nine possible offers.

coefficients of the first principal component. It is usual to average these values for each individual experimental condition (here the 9 offers) and use the average of each as a measure of central tendency along with a representation of the uncertainty, here through the SE. The principal components were extracted using the “pca” function in MATLAB R2014b (MathWorks).

Linear discriminant analysis. To assess the population activity, we constructed and tested classifiers based on the linear discriminant analysis. 60% of the original data were randomly sampled and used as TRAINING set. The remaining 40% were used as TESTING set (“sample” in the “R” statistical programming language). The TRAINING set was used to train the classifier (“lda” in the R statistical programming language) and the TESTING set was used for making predictions (“predict” in the R statistical programming language). This random sampling of the data and testing the classifier was repeated 30 times.

Information theoretic analysis. For the convenience of the reader, the text describing the information theoretic analysis is modified from an earlier paper from this laboratory where the information theoretic analysis is explained in detail (Optican and Richmond, 1987).

The information theoretic analysis is based on the probabilistic relation between the responses of the neurons and the experimental conditions (here the cues). It quantifies how well a set R of responses r allows us to discriminate among the stimuli S, given by the following:

$$I(S;R) = \sum_s \left[P(s|r) \log_2 \frac{P(s|r)}{P(s)} \right] r$$

where S is the set of the visual stimuli s (9 visual cues) and R is the set of the responses r (here either PC1, or PC1 and PC2 simultaneously).

The brackets represent an average over the response distribution $P(r)$. $P(s)$ is the a priori probability of the stimulus s , and $P(s|r)$ is the conditional probability of stimulus s given that response r occurred.

To determine whether PC2 carries information that is not available from knowing PC1 we use PC1 and PC2 simultaneously. If the information carried by PC1 and PC2 were independent, then the information from the joint code would be equal to adding the information from PC1

and PC2 each measured independently. If the PC2 carried no additional information the amount of information carried in the joint code would equal that carried by PC1 alone.

We estimated the conditional probabilities $P(s|r)$ using a neural network with one hidden layer that was trained using backpropagation where r (either PC1 or PC1 and PC2 simultaneously) was the input. The learning rate η (set to 0.05) and the inertia α (set to 0.6) are used to control the speed/accuracy trade off of the learning.

Once we have trained the network so that its outputs $O_s(r)$ provide a good estimate of $P(s|r)$, we can substitute $O_s(r)$ for $P(s|r)$ in the expression above and average over a dataset r^μ to estimate the transmitted information:

$$I_{\text{est}}(S;R) = \frac{1}{n} \sum_{\mu,s} \left[O_s(r^\mu) \log_2 \frac{O_s(r^\mu)}{P(s)} \right]$$

with $P(s)$ estimated as $n^{-1} \sum_{\mu} O_s(r^\mu)$, where n is the total number of dimensions in the response (i.e., $n = 1$ for PC1, $n = 2$ for PC1 and PC2 simultaneously).

The neural network was trained using an early stopping procedure (Kjaer et al., 1994) to control the overfitting. We divided the data into training and testing segments. The network was tested on the test set while the training set was used to drive the backpropagation algorithm. Training was stopped when the test set error reached a minimum. The data were divided into training and test sets at least three different ways for each analysis. To obtain the final estimates, $I_{\text{est}}(S|R)$, the individual estimates for the three test sets were averaged for each neuron.

Results

Behavior

Two monkeys performed a task in which each of nine visual cues indicated a different value offer. Each offer was a combination of a reward size (2, 4, or 6 drops of liquid reward) and a discounting delay (1, 4, or 7 s after the choice). One offer was made in each trial, and the monkeys chose whether to accept or refuse the offer.

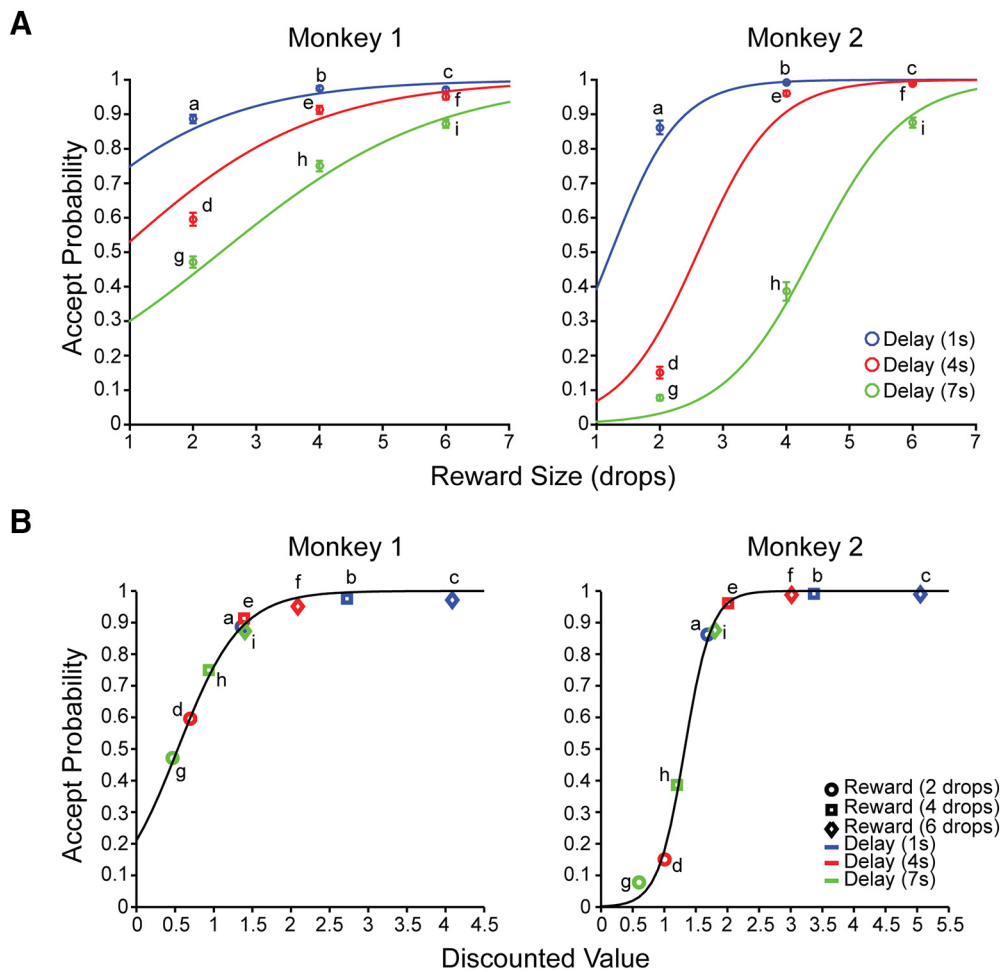


Figure 2. Behavioral performance. **A**, Mean accept probabilities (\pm SEM) as a function of reward size (drops), and delay to reward. The colored lines indicate the best fits using a logistic regression model with reward size and delay as independent variables and accept rate as the dependent variable. The probability of accepting the offer is highest for the largest reward with the shortest delay and becomes progressively smaller as the reward size becomes smaller and/or the delay becomes longer. **B**, Accept probability (colored points) estimated from hyperbolic discounting for Monkey 1 and exponential discounting for Monkey 2 (see Results). The black curves show accept probability fit using a sigmoidal function (see Results). Letters from “a” to “i” indicate the nine possible offers.

If the offer was accepted, the cue-defined reward was delivered after the corresponding delay (Fig. 1A; see Materials and Methods). If the offer was rejected, the monkey could immediately begin a new trial.

For the behavioral analyses, we used the data from all sessions in which neural recording was conducted (79 for Monkey 1; 68 for Monkey 2). The probability of accepting an offer, calculated independently for each of the nine offers, was modeled as a binary variable using logistic regression with reward size and delay as independent continuous variables,

$$\text{Accept Rate} = \frac{e^{(\beta_0 + \beta_1(\text{reward size}) + \beta_2(\text{delay}) + \beta_3(\text{reward size} \times \text{delay}))}}{1 + (e^{(\beta_0 + \beta_1(\text{reward size}) + \beta_2(\text{delay}) + \beta_3(\text{reward size} \times \text{delay}))})}$$

The monkeys' behavior was influenced by both reward size and delay, such that the probability of accepting the offer was highest for the largest reward with the shortest delay, and became progressively smaller as the reward became smaller and/or the delay became longer (Fig. 2A, Monkey 1: intercept: $z = 5.4$, $p = 7.9 \times 10^{-8}$, reward size: $z = 17.1$, $p < 2 \times 10^{-16}$, delay: $z = -13$, $p < 2 \times 10^{-16}$, reward size \times delay: $z = -2.7$, $p = 8.0 \times 10^{-3}$; Monkey 2: intercept: $z = -9.1$, $p < 2 \times 10^{-16}$, reward size: $z = 25.8$, $p < 2 \times 10^{-16}$, delay: $z = -19.3$, $p < 2 \times 10^{-16}$, reward size \times delay: $z = -6.5$, $p = 7.5 \times 10^{-11}$).

Certain combinations of reward sizes and delays, that is, different cues, were treated as equivalent. For example, Monkey 1 treated small reward/short delay, medium reward/medium delay, and large reward/long delay, as equivalent, that is, the reward size and delay interact such that longer delays can be compensated by larger rewards (Fig. 2A, a, e, i, respectively). The interaction between reward size and delay can be described using simple reward discounting models with the performance directly related to reward size and inversely to the delay, as shown in the following formulae:

$$\text{Exponential discounting model: } DV = \frac{R}{e^{kD}}$$

$$\text{Hyperbolic discounting model: } DV = \frac{R}{1 + kD}$$

DV is the discounted value expressed as the probability of accepting the offer, R is the reward size in drops, k is the discount factor fit separately to the data for each monkey, and D is the delay until reward delivery.

The discounted value was better estimated by the hyperbolic discounting model in Monkey 1 [Akaike information criterion (AIC): Exponential model vs Hyperbolic model = -35.4 vs

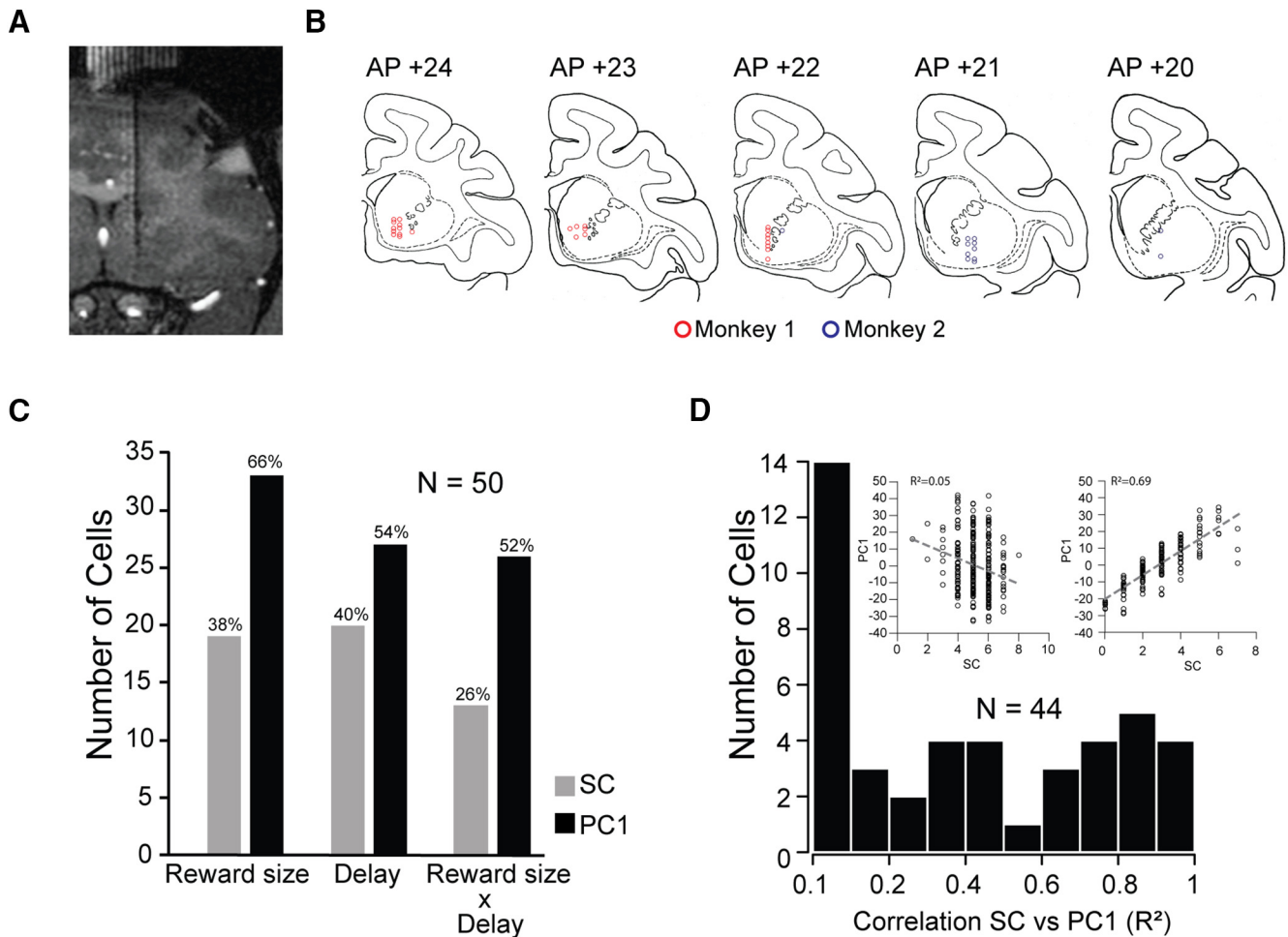


Figure 3. Recording sites and number of cells modulated by the main factors. **A**, Frontal T1-weighted MR scan with one electrode in place, in one monkey. **B**, Locations of recorded TANs (red dots, Monkey 1; blue dots, Monkey 2), reconstructed based on MRI and reported on the *Rhesus Macaque “RED” Symmetrical Brain Atlas* (Saunders and Phil, 2006). **C**, Cells modulated by reward size and/or delay in the cue period using spike count (gray bars) or the coefficient of PC1 (black bars) as dependent variable in the two-way ANOVA. **D**, Distribution of the linear correlation coefficients between spike count (SC) and the coefficient of the PC1 of the significant cells from the two-way ANOVA. Left inset, Scatter plot between SC and PC1 for one cell with low correlation (R^2) between SC and PC1. Right inset, Scatter plot between SC and PC1 for one cell with high correlation (R^2) between SC and PC1.

–41.9] and the exponential model in Monkey 2 (AIC: Exponential model vs Hyperbolic model = –34.3 vs –32.8).

The relationship between the accept probability and discounted value (Fig. 2B, black curves) was represented by a two-parameter logistic function:

$$\text{Accept Probability} = \frac{1}{1 + e^{(-\alpha DV - \beta)}}$$

where parameter α controlled the steepness of the sigmoidal curve, and parameter β determined the degree of horizontal translation. A general-purpose optimization function (“optim” in the R statistical programming language) was used to find the best fits for α and β (α and β for each monkey).

Single neuron activity

Neural activity is modulated by the information represented in the visual cue

While monkeys performed the task, we recorded 50 TANs from the rostromediovventral striatum (34 in Monkey 1; 16 in Monkey 2), located between the pyramidal tract and the nucleus accumbens (see Materials and Methods; Fig. 3A,B).

The neural activity expressed as total spikes count in the cue period significantly modulated by reward and/or delay in 31/50

TANs (two-way ANOVA, $p < 0.05$; Fig. 3C, gray bars). On examination of the spike raster plots, it appeared that for at least some of the other 19 TANs the distribution of spikes changed as a function of the offer (Fig. 4A, compare response to 6 drops, 1 s delay vs 2 drops, 7 s).

To represent the modulation of the pause-pulse relation, as in Figure 4B, a principal component analysis of the responses was performed for each of the 50 neurons (see Materials and Methods).

When the coefficients of the first principal component (PC1) were taken as the dependent variable, 44 cells showed significant modulation to reward size and/or delay (two-way ANOVA, $p < 0.05$; Fig. 3C, black bars); 13 more than reported when looking at spike count alone.

Across the population of neurons, the correlations of spike count versus PC1 went from nearly zero (Fig. 3D, left inset) to almost 1 (Fig. 3D, right inset).

For 20/44 neurons (10 from Monkey 1; 10 from Monkey 2), the correlation between the coefficient of PC1 and spike count was < 0.3 (Fig. 3D, and left inset). For these neurons, PC1 explained more variance ($M \pm SD$: $19.4 \pm 14.1\%$) than was explained by the spike count ($M \pm SD$: $5.8 \pm 2.8\%$; $N = 20$; two-sample t test: mean of the differences $\pm SD = 13.5 \pm 3.2$,

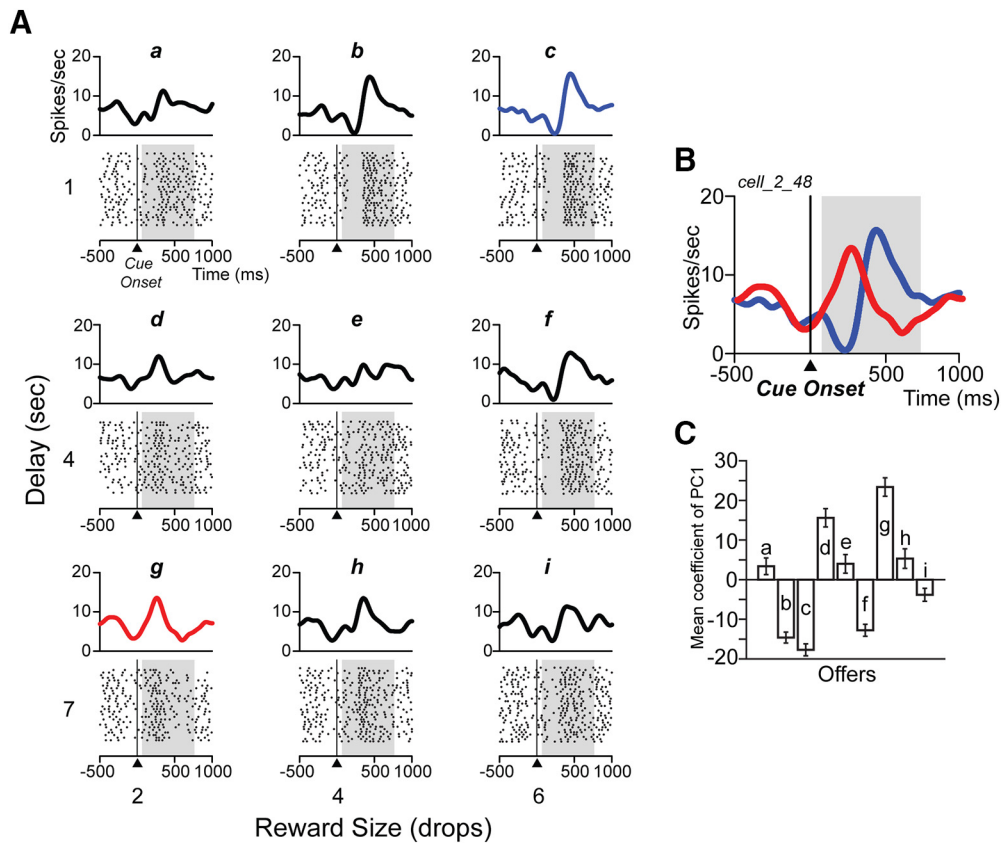


Figure 4. Spike raster plots for one sinusoidal neuron. **A**, Raster displays from one neuron for the nine offers (indicated from “a” to “i” and related to Fig. 1B), aligned to cue onset (time 0). The spike density is shown above each raster. Shading shows the 75–750 ms analysis interval. **B**, Spike density curves of two offers (largest value, 6 drops/1 s delay, blue curve; top right in **A**, and smallest value, 2 drops/7 s delay, red curve; bottom left in **A**), illustrating the inversion of the pause-pulse phase according to reward value. **C**, Mean coefficient of PC1 (\pm SE) for each of nine offers (indicated from “a” to “i”) for the neuron in **A**.

$t_{(38)} = 4.21$, $p = 0.00015$). The shape of PC1 for each of the neurons with poor correlation between spike count and PC1 appeared sinusoidal (Fig. 5A).

For the neuron in Figure 4A the mean coefficient of the first principal component for three offers (indicated as b, c, and f) looked similar in gain and all three were negative; similarly, the responses to the offers indicated as d and g were similar in gain and both positive (Fig. 4C). The responses to the other offers (a, e, h, i) appeared all had low gain.

The first principal component provides a convenient and compact means to quantify the time varying response across offers. When the PC1 (Fig. 5A, black curve) weighted by the coefficient, that is the gain (Fig. 4C), has been added to the average response, good reconstructions of the experimental data emerged (Fig. 5D).

For 24/44 neurons the correlation between the coefficient of PC1 and spike count was >0.3 (Fig. 3D, and right inset). For these neurons, the percentage of variance explained by PC1 ($M \pm SD$: $16.2 \pm 9.6\%$) was not significantly more than that explained by spike count ($M \pm SD$: $12.7 \pm 7.4\%$; $N = 24$; two-sample t test: mean of the differences \pm SD = 3.5 ± 2.5 , $t_{(46)} = 1.40$, $p = 0.17$). The shape of PC1 for each of these 24 neurons was unimodal (18/24, 14 from Monkey 1 and 4 from Monkey 2; Fig. 5B) or sustained (6/24, 5 from Monkey 1 and 1 from Monkey 2; Fig. 5C). For this unimodal group of neurons, the variability of the response among the offers was localized mostly only at the pulse phase (Fig. 6A, in detail in B).

This means that for this group of neurons, counting spikes especially during the pulse phase is a good representation of the

coding. We chose to use the coefficient of the first principal component so we did not have to examine neuron-by-neuron to adjust placement of the best coding window in time. There was no difference in the anatomical localization of these two populations of neurons. For each individual neuron, we estimated how much information about the visual cue was present in the neuronal response code formed using PC1 as the response measure (Richmond and Optican, 1987).

To determine whether more information about the reward size and the delay carried by the visual cues that was not available from PC1 could instead be available using the coefficient of the second principal component, PC2, we used the information theoretic analysis (Optican and Richmond, 1987) on the response code formed either by PC1 alone or from an extended code using both PC1 and PC2 (see Materials and Methods). The information carried by the extended code ($M \pm SD$: $0.18 \text{ bits} \pm 0.12$) was not significantly different from the information carried by PC1 alone ($M \pm SD$: $0.15 \text{ bits} \pm 0.12$; $N = 44$, two sample t test: mean of the differences \pm SD = -0.04 ± 0.03 , $t_{(86)} = -1.51$, $p = 0.13$). Thus, PC1 alone was sufficient to describe the stimulus-dependent information carried by the temporally modulated neural responses of all the neurons.

The neural activity was also analyzed in the reward period for each of the 50 cells. The reward-related activity was defined as the activity during the 400 ms after the onset of reward delivery. A two-way ANOVA was computed for each neuron with factors reward size and delay, and the coefficient of principal component (PC1) as the dependent variable, to identify units modulated according to these factors. During the reward period a smaller

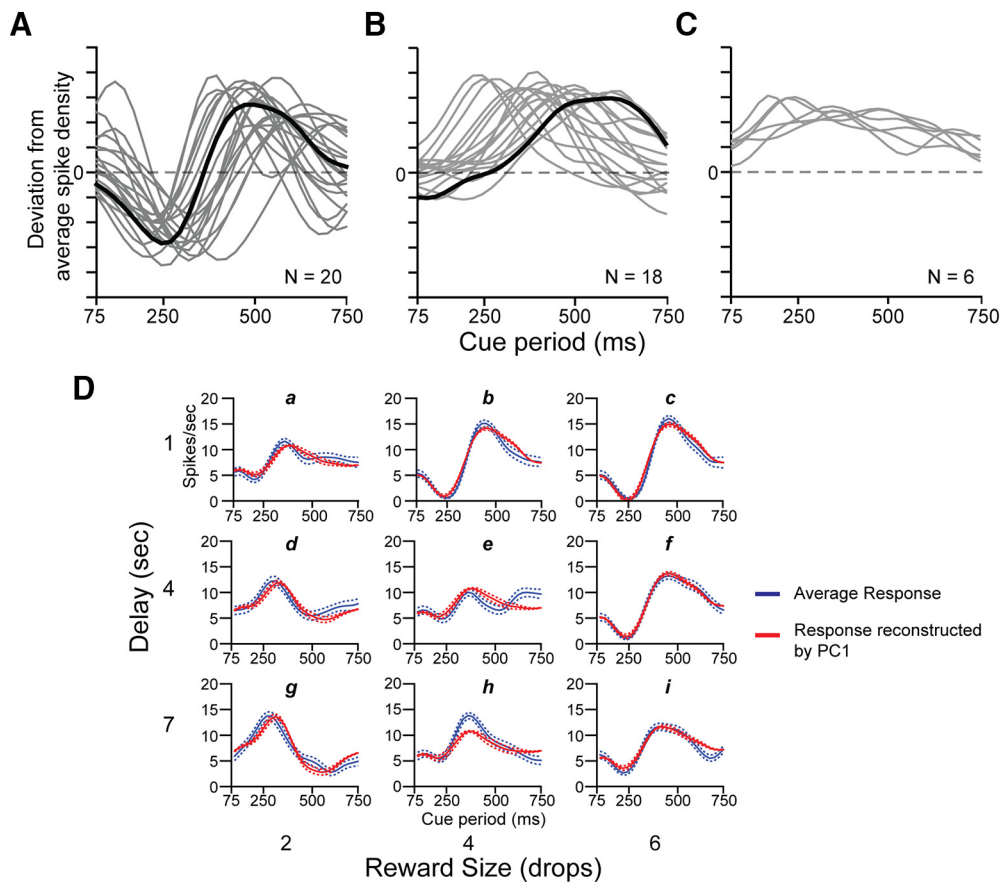


Figure 5. First principal component by shape. **A**, Sinusoidal first principal components for 20 neurons showing reversed pause-pulse response among offers (gray and black curves). The black curve corresponds to the sinusoidal shape of the first principal component of the neuron in Figure 4A. The two peaks, one positive and the other negative for PC1, show that the most variance in the overall response is localized in two periods of the cue period. In the neural response of the neuron in Figure 4A, these two periods roughly coincide with those where the pause becomes pulse or vice versa and the pulse became pause or vice versa (shown in detail in Fig. 4B). **B**, Unimodal first principal component for 18 neurons (gray and black curves). The black curve corresponds to unimodal shape of the first principal component of the neuron in Figure 6A. This curve shows that the variance across the overall response is mostly localized in one period of the whole cue period. This period roughly coincides with the starting of the pulse phase of the neural response and with the starting of the variability of the response among the offers (shown in detail in Fig. 6B). **C**, First principal component for six neurons with sustained firing (gray curves). **D**, Average response (blue line) and average response reconstructed by adding PC1 weighted by coefficient of PC1 to the average response (red line) for the neuron in Figure 4A, for each offer (indicated from “a” to “i”). Dashed lines represent SE.

proportion of TANs have responses related to the reward size and/or delay (25/50) than during the cue period ($\chi^2_{(1, N=50)} = 16.88, p = 0.00004$).

Few cells (5/25) showed a sinusoidal shape of the first principal component in the reward period. Of these five, two cells showed a sinusoidal shape of the first principal component, both in cue and reward period. All other neurons (20/25) showed unimodal or sustained shape of PC1 (11 unimodal and 9 sustained). Of 11 cells with unimodal shape of PC1, 3 showed sinusoidal PC1 and 2 showed sustained PC1 in cue period. Of nine cells with sustained shape of PC1, two showed also sustained shape in the cue period. This result shows that the response features to the value seemed to be restricted to the cue period.

The neural response scales with the discounted value of the visual cue

For 17/44 (40%) neurons, there was a significant linear correlation between the mean coefficient of the first principal component of each offer and the discounted value calculated from the behavior (Fig. 7A, D). The linear correlation with the discounted value was seen for both sinusoidal and unimodal neurons. Of these 17 neurons 8 had unimodal PC1 (6 unimodal and 2 sustained) and 9 were sinusoidal.

The responses to some stimuli appeared similar to each other. To determine how many groups of stimulus-elicited responses might legitimately be distinguished, a cluster analysis was performed followed by a set of ANOVAs, i.e., were there really nine distinguishable responses, or was the number likely to be fewer?

We performed divisive clustering for each cell (“hclust” function in R with agglomeration method “ward.D2”). Using the cluster dendrogram as a guide the responses for each cell were treated as though there were 2, 3, 4, or 9 distinguishable levels (Fig. 7B).

For each cell, four separate one-way ANOVAs were performed with the number of levels as the independent variable. These models were compared (“anova” function in R using test “ChiSq”). The model with largest number of levels that was significantly different from the next larger number of levels was taken to signify the most complex model (largest number of levels) that was justifiable. This procedure showed that the visual cue-elicited responses could be grouped into 2 levels (36/44 cells), 3 levels (7/44 cells), or 4 levels (1/44 cells). Thus, overall the responses of these neurons fell mostly into two groups (Fig. 7E).

For the 17 neurons that showed the modulation linearly related to the discounted value, the neural activity appeared clustered at two levels, high and low reward value (Fig. 7C, dashed

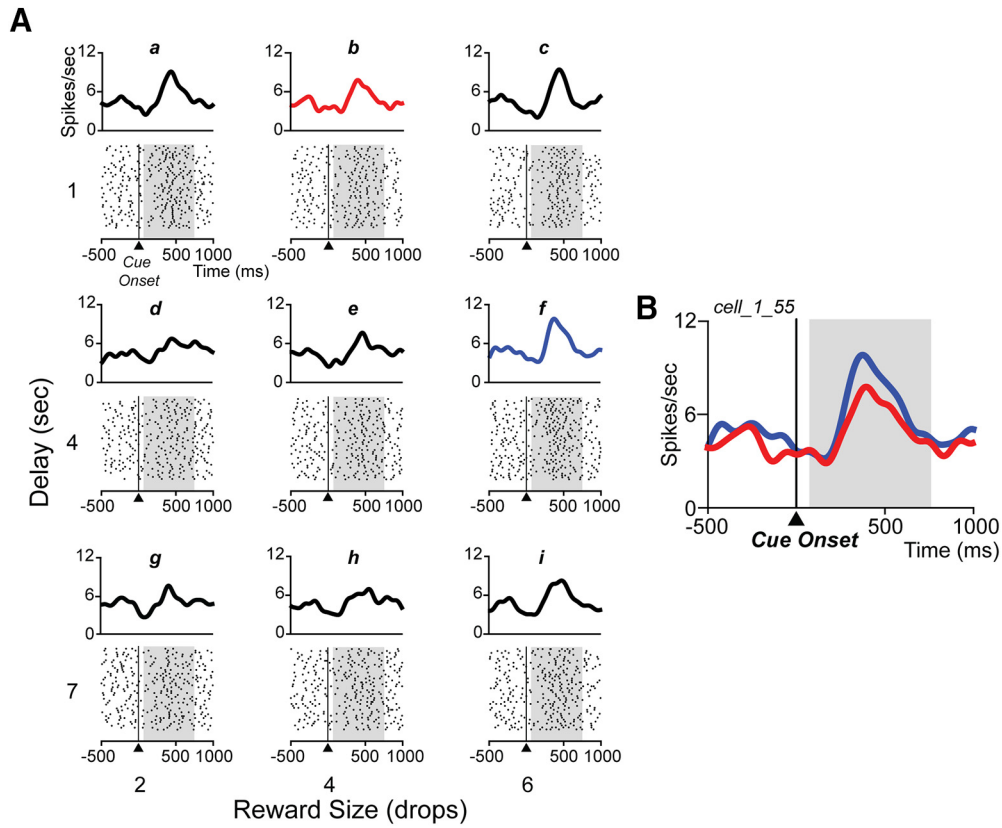


Figure 6. Spike raster plots for one unimodal neuron. **A**, Raster displays of the response of one neuron among the nine offers (indicated from “a” to “i” and related to Fig. 1B), aligned to cue onset (time 0). The spike density is shown above each raster. Shading shows the 75–750 ms analysis interval. **B**, Spike density curves of two offers (6 drops/4 s as delay, blue curve; middle, right in **A**) and 4 drops/1 s as delay (red curve; middle, top in **A**), illustrating the modulation of the pulse phase according to the offer.

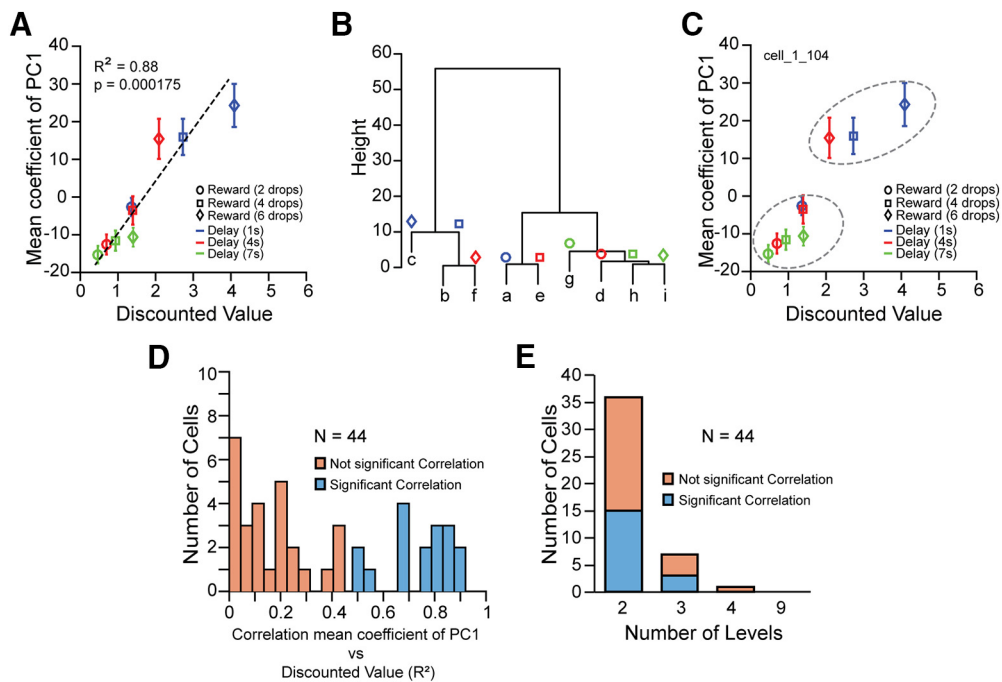


Figure 7. Neural response related to discounted value. **A**, Linear correlation between the mean coefficient of PC1 for each offer and the discounted value calculated from the behavior for one neuron. Error bars represent SE. Letters from “a” to “i” correspond to the nine offers (related to Fig. 1B). **B**, Cluster dendrogram for the same neuron as in **A**, using the mean coefficient of PC1. Based on the cluster dendrogram the responses were treated as if they were distinguishable among the nine levels corresponding to the nine offers (from “a” to “i”), or as if they were distinguishable in progressively fewer levels; that is, in four levels: “c”, “b/f”, “a/e”, “g/d/h/i”, in three levels: “c/b/f”, “a/e”, “g/d/h/i”, or in two levels: “c/b/f”, “a/e/g/d/h/i”, where the response in each level was indistinguishable. **C**, The two dashed ovals represent the two major levels identified after the model selection analysis, where the hypothesis is that neural responses to the offers c, b, and f are indistinguishable from each other, and distinguishable from the neural responses to the offers a, e, g, d, h, and i, which are each indistinguishable from each other. Error bars represent SE. **D**, Distribution of the linear correlation coefficients (R^2) between the mean coefficient of PC1 and discounted value for each neuron. **E**, Cells with not significant (orange bars) and significant (cyan bars) linear correlation between the mean coefficient of PC1 and discounted value, in which the neural response to the offers was distinguishable among 2, 3, 4, or 9 levels.

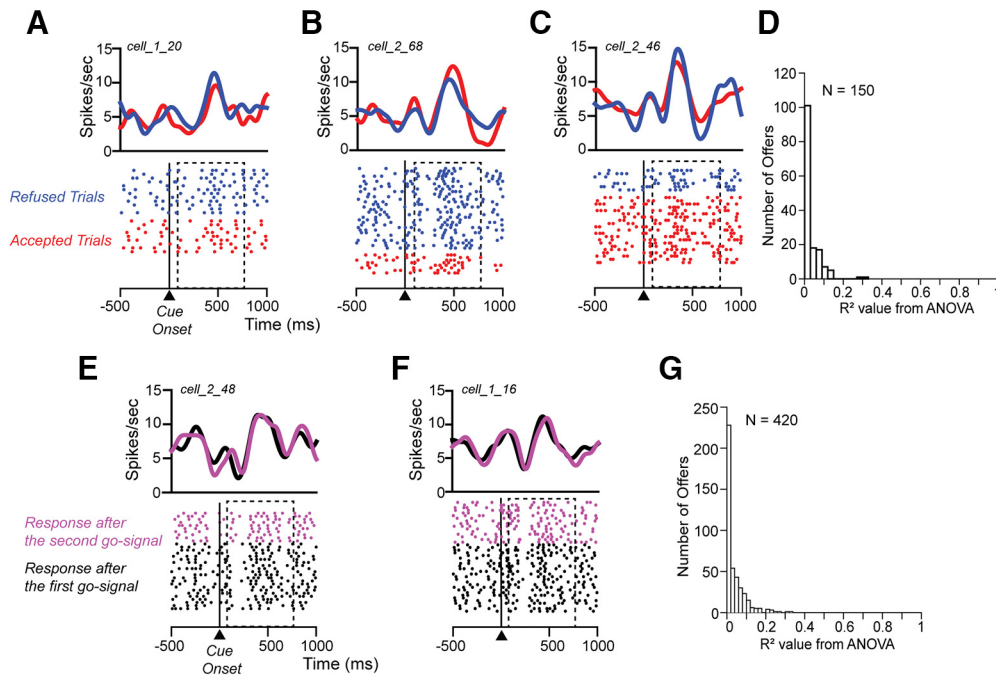


Figure 8. Neural response is not related to predicted choice and to motor response. Raster displays with spike densities on top of three different neurons when visual cue indicating 2 drops/long delay (**A**), 4 drops/long delay (**B**), and 6 drops/long delay (**C**) are shown to the monkey. The activity is aligned to the cue onset (time 0); dashed rectangles show the 75–750 ms analysis interval. Trials are sorted based on the choice, accept (red) or refuse (blue). **D**, Distribution of R^2 values extracted from the ANOVA for each offer. Raster displays with spike densities on top of two different neurons when visual cue indicating 6 drops/long delay (**E**), 6 drops/short delay (**F**), are shown to the monkey. The activity is aligned to the cue onset (time 0); dashed rectangles show the 75–750 ms analysis interval. Trials are sorted based on whether the response is made after the first go-signal (black) or after the second go-signal (pink). **G**, Distribution of R^2 values extracted from the ANOVA for each offer.

ovals) or in three levels of high, medium and low value; that is, when the cues represented outcomes with similar value to the monkeys (Fig. 7C; stimuli labeled f, b, and c elicit indistinguishable responses), the neural activity failed to distinguish among cue elicited responses within the cluster. Overall, the cue-elicited responses are related to the value associated with the cue.

To determine whether the offers were distinguishable by the population activity we constructed and tested decoding classifiers based on linear discriminant analysis (see Materials and Methods). The average decoding accuracy estimated as mean across 30 iterations was $11.6 \pm 0.44\%$ ($M \pm SD$).

To exclude the hypothesis that the modulation of the response in the cue period was related to the intention of the monkey to accept or refuse the offer, a one-way ANOVA ($p < 0.05$) was performed on each offer where the animal responded with at least 5 accepted and 5 refused trials ($N = 150$; 119 offers for Monkey 1 and 31 for Monkey 2). For 95% of the offers the neural response when the monkey accepted was indistinguishable from that when the monkey refused (Fig. 8A,B,C). The R^2 values extracted from the ANOVA for all 150 offers were small (Fig. 8D). The 7/150 that were significant were just about exactly what would be expected by chance at the 5% level.

We evaluated whether neural activity in cue period was related to the subsequent motor response; that is, does motor preparation influence the cue-related response? Each offer was split in trials where the animal released the bar after the first go-signal and in trials where the animal released the bar after the second go-signal (see Materials and Methods). A one-way ANOVA ($p < 0.05$) was performed on each offer where the animal responded for at least five trials after the first go-signal and five trials after the second go-signal ($N = 420$; 279 offers for Monkey 1 and 141 for Monkey 2). For 96% of the offers the neural response in the cue

period when the animal released the bar after the first go signal, the cue-related response was indistinguishable from that when the monkey released the bar after the second go signal (Fig. 8E,F). The R^2 values extracted from the ANOVA for these 420 offers were small. The 18/420 offers that were significant is just about exactly the number of them that would be expected to appear significant by chance at the 5% level (Fig. 8G). Thus, we were unable to identify an effect on the responses that could be attributed to motor anticipation.

Discussion

We studied how TANs in the rostromediodorsal striatum encode information about predicted rewards. In each trial the monkey was offered a single reward that could be accepted or refused. The reward value, signaled through its association with a visual cue, was constructed by combining one of three reward sizes with one of three delays. The monkeys modulated their behavior by accepting or refusing the offer according to the value the animal attributed to the predicted reward; they were increasingly likely to accept offers as the reward became larger and the delay became shorter. We observed that the values were very well described by a simple reinforcement learning model for the discounted value of the rewards. For both monkeys the ordering of values was the same, although the two monkeys showed different overall sensitivity to the discounted values.

All of the TANs we recorded modulated their activity in relation to the discounted value of the reward predicted by the cue. The response was localized in time. Past studies using Pavlovian conditioning tasks have shown that TANs respond to reward delivery and to conditioned stimuli that predict appetitive or aversive outcomes with a stereotyped response, a pause followed by a pulse (Kimura et al., 1990; Aosaki et al., 1994; Ravel et al.,

1999, 2001). The pause-pulse nature of TANs responses has become almost diagnostic for their identification (Marche et al., 2017) and is frequently used as one of the criteria to identify TANs. For at least some of the reward offerings each of our TANs showed the pause followed by an excitatory rebound response that has been taken as characteristic of TANs (Kimura et al., 1984, 1990; Apicella, 2002, 2017).

For approximately half of the neurons we recorded the intensity of the pulse was modulated by the predicted reward in a manner similar to that reported recently (Nougaret and Ravel, 2015). For the other half, the pattern of the response surprised us. The modulation involved the entire pause-pulse pattern. The pause and pulse were modulated simultaneously like a see-saw going up and down. For some cues there was a complete reversal of the pause-pulse pair so that the response became pulse-pause, that is, the temporal shape of the response depended on which the cue appeared. The see-saw modulation of these neurons interferes with detection of the response modulation when the spikes are counted over the whole interval because the pause and pulse are balanced.

To analyze this temporally modulated signal we used principal component analysis. The principal component analysis revealed what was obvious from inspection: the two phases of the response were modulated together. The first principal component looked like a sinusoid that was modulated across the different valued cues. The positive and negative components of sinusoidal principal component were equal, thereby showing why spike counting across the interval did not reveal neural response modulation.

There seems to be only one previous report of recordings taken from rat ventral striatum that shows a phase reversing modulation in TANs to unexpected reward delivery and omission during reward-based learning (Atallah et al., 2014). At this point it is not clear whether the phase reversing responses we have seen are characteristic of TANs throughout the striatum, and they can only be seen when there is a wide range of values to encode, or whether they are found only in this limbic striatum. In monkeys the one study that compares TANs in the dorsal, intermediate, and ventral striatum shows that the dorsal and ventral TANs have differences in response latency and pause width in response to rewarding events (Marche et al., 2017), so it is at least plausible to think this could be a characteristic found only in limbic striatum, but it is possible that the reversing modulation we have seen in ventral TANs is also a feature of dorsal TANs response, if they were to be recorded under the same experimental circumstances.

At this time, it is still not clear how the two responses phases, pause and pulse arise (Matsumoto et al., 2001; Nanda et al., 2009; Aosaki et al., 2010; Nakajima et al., 2017). For the neurons we term unimodal where the pulse is modulated according to the reward value and the pause seems unaffected, the origins of pause and pulse appear independent (Schulz and Reynolds, 2013). For the sinusoidal TANs with the sinusoidal first principal component, the inputs contributing to the pulse are equal in power to the inputs contributing to the pause.

Thus, we might have seen two classes of TANs characterized by different input patterns, balanced versus unbalanced excitatory and inhibitory inputs. The changes in TAN firing might well arise from the multitude of different afferents that TANs receive, mostly glutamatergic from the cortex and the thalamus (Matsumoto et al., 2001; Doig et al., 2014), and other afferents from GABAergic collaterals of striatal projection neurons as seen in monkeys (Gonzales et al., 2013), including from the ventral tegmental area that, as seen in mice, provide direct inhibition of the ventral striatal cholinergic interneurons (Brown et al., 2012).

The sinusoidal temporal modulation of the TANs firing we have reported here was unexpected. It is intriguing and opens the question of what inputs are driving it and what has contributed to it and whether there is any functional and/or cellular mechanism different from the mechanisms driving the unimodal responses.

Striatal cholinergic interneurons play an important role in behavioral flexibility, both for reversal learning (Brown et al., 2010) and set-shifting (Aoki et al., 2015). The temporal modulation could serve as a neural mechanism relating the animal's ability to react flexibly according to the contingency in the current trial.

Taking the results of Atallah et al. (2014) and Marche et al. (2017) into account with our results describes the sensitivity of the response to predicted reward value in ventral striatal TANs. It appears that the relatively rich repertoire of temporally modulated responses we found in these individual TANs arose as a consequence of having presented a relatively rich repertoire of reward values, and, perhaps, the need to compute value across two-dimensions, here temporal delay and reward size.

Approximately half of the neurons taken from each of the unimodal and the sinusoidal groups showed linear scaling of the coefficient of the first principal component against the value estimated from the reinforcement model of the behavior. In addition, responses were clustered in such a way that several visual cues of equivalent value elicited responses that were indistinguishable. This latter finding makes it unlikely that the neural responses were modulated by the identity of the visual cue but rather the responses are related to the value encoded.

Each TAN we recorded had its response localized in time with respect to the onset of the cue, with the response to each TAN occupying its own characteristic time window as seen by examining the positions of the first principal components in time. The total time tiled by the population of TANs is relatively wide, order 300 ms (compare Fig. 5, compare A, B, C). The time-locking of each TAN in time combined with the temporal coding makes it appear that the population of TANs has a particular signature for each stimulus. With this coding mechanism the message being carried by the population is sustained over this relatively extended time interval, presumably making the message available for a relatively prolonged period as might be needed by a decision-making process evaluating a set of options.

There is evidence for synchronization, seen as peaks in cross-correlograms, that might be related to this time-locking of TAN responses (Raz et al., 1996). Based on now classic experiments (Bryant and Segundo, 1976; Mainen and Sejnowski, 1995), we know that continuously changing rate, like the responses we have seen here, is a powerful mechanism for entraining the membrane of a target neuron so that the timing of spikes at the target exhibits a replicable pattern on each iteration of the same input. The coordinated activity across TANs could lead to the TANs having a coordinating effect on dopamine release (Cachope et al., 2012; Threlfell et al., 2012; Cachope and Cheer, 2014) and on projection neurons in a manner similar to what others previously reported. In this scenario TANs would influence the membrane potential of projection neurons, thereby coordinating the activity of these neural elements, through regulating their membrane up-states so that these targeted neural elements become coordinated (Graybiel et al., 1994; Stern et al., 1998; Carrillo-Reid et al., 2009; Kreitzer, 2009). Despite many efforts, it has been difficult to show that temporally encoded information plays a role in cognitive function (Ahissar et al., 1992; Haalman and Vaadia, 1998; Grün et al., 2002). Here, we observed that in an operant task rostromediodorsal striatal TANs in monkeys use temporally modulated

responses to encode cognitive information from a rich set of reward values. We speculate that this temporal coding is important for regulating appropriate value-based choices.

References

- Ahissar E, Vaadia E, Ahissar M, Bergman H, Arieli A, Abeles M (1992) Dependence of cortical plasticity on correlated activity of single neurons and on behavioral context. *Science* 257:1412–1415.
- Aoki S, Liu AW, Zucca A, Zucca S, Wickens JR (2015) Role of striatal cholinergic interneurons in set-shifting in the rat. *J Neurosci* 35:9424–9431.
- Aosaki T, Tsubokawa H, Ishida A, Watanabe K, Graybiel AM, Kimura M (1994) Responses of tonically active neurons in the primate's striatum undergo systematic changes during behavioral sensorimotor conditioning. *J Neurosci* 14:3969–3984.
- Aosaki T, Kimura M, Graybiel AM (1995) Temporal and spatial characteristics of tonically active neurons of the primate's striatum. *J Neurophysiol* 73:1234–1252.
- Aosaki T, Miura M, Suzuki T, Nishimura K, Masuda M (2010) Acetylcholine-dopamine balance hypothesis in the striatum: an update. *Geriatr Gerontol Int* 10:S148–S157.
- Apicella P (2002) Tonically active neurons in the primate striatum and their role in the processing of information about motivationally relevant events. *Eur J Neurosci* 16:2017–2026.
- Apicella P (2017) The role of the intrinsic cholinergic system of the striatum: what have we learned from TAN recordings in behaving animals? *Neuroscience* 360:81–94.
- 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, Deffains M, Legallet E (2011) The role of striatal tonically active neurons in reward prediction error signaling during instrumental task performance. *J Neurosci* 31:1507–1515.
- Atallah HE, McCool AD, Howe MW, Graybiel AM (2014) Neurons in the ventral striatum exhibit cell-type-specific representations of outcome during learning. *Neuron* 82:1145–1156.
- Averbeck BB, Lehman J, Jacobson M, Haber SN (2014) Estimates of projection overlap and zones of convergence within frontal-striatal circuits. *J Neurosci* 34:9497–9505.
- Bouret S, Richmond BJ (2010) Ventromedial and orbital prefrontal neurons differentially encode internally and externally driven motivational values in monkeys. *J Neurosci* 30:8591–8601.
- Brown HD, Baker PM, Ragazzino ME (2010) The parafascicular thalamic nucleus concomitantly influences behavioral flexibility and dorsomedial striatal acetylcholine output in rats. *J Neurosci* 30:14390–14398.
- Brown MT, Tan KR, O'Connor EC, Nikonenko I, Muller D, Lüscher C (2012) Ventral tegmental area GABA projections pause accumbal cholinergic interneurons to enhance associative learning. *Nature* 492:452–456.
- Bryant HL, Segundo JP (1976) Spike initiation by transmembrane current: a white-noise analysis. *J Physiol* 260:279–314.
- Cachope R, Cheer JF (2014) Local control of striatal dopamine release. *Front Behav Neurosci* 8:188.
- Cachope R, Mateo Y, Mathur BN, Irving J, Wang HL, Morales M, Lovinger DM, Cheer JF (2012) Selective activation of cholinergic interneurons enhances accumbal phasic dopamine release: setting the tone for reward processing. *Cell Rep* 2:33–41.
- Cai X, Kim S, Lee D (2011) Heterogeneous coding of temporally discounted values in the dorsal and ventral striatum during intertemporal choice. *Neuron* 69:170–182.
- Carrillo-Reid L, Tecuapetla F, Vautrelle N, Hernández A, Vergara R, Galaraga E, Bargas J (2009) Muscarinic enhancement of persistent sodium current synchronizes striatal medium spiny neurons. *J Neurophysiol* 102:682–690.
- Clark AM, Bouret S, Young AM, Murray EA, Richmond BJ (2013) Interaction between orbital prefrontal and rhinal cortex is required for normal estimates of expected value. *J Neurosci* 33:1833–1845.
- Contant C, Umbriaco D, Garcia S, Watkins KC, Descarries L (1996) Ultrastructural characterization of the acetylcholine innervation in adult rat neostriatum. *Neuroscience* 71:937–947.
- Doig NM, Magill PJ, Apicella P, Bolam JP, Sharott A (2014) Cortical and thalamic excitation mediate the multiphasic responses of striatal cholinergic interneurons to motivationally salient stimuli. *J Neurosci* 34:3101–3117.
- Eldridge MA, Lerchner W, Saunders RC, Kaneko H, Krausz KW, Gonzalez FJ, Ji B, Higuchi M, Minamimoto T, Richmond BJ (2016) Chemogenetic disconnection of monkey orbitofrontal and rhinal cortex reversibly disrupts reward value. *Nat Neurosci* 19:37–39.
- Gonzales KK, Pare JF, Wichmann T, Smith Y (2013) GABAergic inputs from direct and indirect striatal projection neurons onto cholinergic interneurons in the primate putamen. *J Comp Neurol* 521:2502–2522.
- Graveland GA, DiFiglia M (1985) The frequency and distribution of medium-sized neurons with indented nuclei in the primate and rodent neostriatum. *Brian Res* 327:307–311.
- Graybiel AM, Aosaki T, Flaherty AW, Kimura M (1994) The basal ganglia and adaptive motor control. *Science* 265:1826–1831.
- Grün S, Diesmann M, Aertsen A (2002) Unitary events in multiple single-neuron spiking activity: I. Detection and significance. *Neural Comput* 14:43–80.
- Haalman I, Vaadia E (1998) Emergence of spatio-temporal patterns in neuronal activity. *Z Naturforsch C* 53:657–669.
- Haber SN, McFarland NR (1999) The concept of the ventral striatum in nonhuman primates. *Ann N Y Acad Sci* 877:33–48.
- Hays AV Jr, Richmond BJ, Optican LM (1982) A UNIX-based multiple process system for real-time data acquisition and control. WESCON Conference Proceeding, pp 2/1–1 to 2/1–10. OSTI ID: 5213621.
- Izzo PN, Bolam JP (1988) Cholinergic synaptic input to different parts of spiny striatonigral neurons in the rat. *J Comp Neurol* 269:219–234.
- Judge SJ, Richmond BJ, Chu FC (1980) Implantation of magnetic search coils for measurement of eye position: an improved method. *Vision Res* 20:535–538.
- Kimura M, Rajkowski J, Evarts E (1984) Tonicly discharging putamen neurons exhibit set-dependent responses. *Proc Natl Acad Sci U S A* 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.
- Kjaer TW, Hertz JA, Richmond BJ (1994) Decoding cortical neuronal signals: network models, information estimation and spatial tuning. *J Comput Neurosci* 1:109–139.
- Kreitzer AC (2009) Physiology and pharmacology of striatal neurons. *Annu Rev Neurosci* 32:127–147.
- Mainen ZF, Sejnowski TJ (1995) Reliability of spike timing in neocortical neurons. *Science* 268:1503–1506.
- Marche K, Martel AC, Apicella P (2017) Differences between dorsal and ventral striatum in the sensitivity of tonically active neurons to rewarding events. *Front Syst Neurosci* 11:52.
- 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.
- Nagai Y, Kikuchi E, Lerchner W, Inoue KI, Ji B, Eldridge MA, Kaneko H, Kimura Y, Oh-Nishi A, Hori Y, Kato Y, Hirabayashi T, Fujimoto A, Kumata K, Zhang MR, Aoki I, Suhara T, Higuchi M, Takada M, Richmond BJ, Minamimoto T (2016) PET imaging-guided chemogenetic silencing reveals a critical role of primate rostromedial caudate in reward evaluation. *Nat Commun* 7:13605.
- Nakajima A, Shimo Y, Uka T, Hattori N (2017) Subthalamic nucleus and globus pallidus interna influence firing of tonically active neurons in the primate striatum through different mechanisms. *Eur J Neurosci* 46:2662–2673.
- Nanda B, Galvan A, Smith Y, Wichmann T (2009) Effects of stimulation of the centromedian nucleus of the thalamus on the activity of striatal cells in awake rhesus monkeys. *Eur J Neurosci* 29:588–598.
- Nougaret S, Ravel S (2015) Modulation of tonically active neurons of the monkey striatum by events carrying different force and reward information. *J Neurosci* 35:15214–15226.
- Optican LM, Richmond BJ (1987) Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex: III. Information theoretic analysis. *J Neurophysiol* 57:162–178.
- 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.

- Ravel S, Sardo P, Legallet E, Apicella P (2001) Reward unpredictability inside and outside of a task context as a determinant of the responses of tonically active neurons in the monkey striatum. *J Neurosci* 21:5730–5739.
- Ravel S, Sardo P, Legallet E, Apicella P (2006) Influence of spatial information on responses of tonically active neurons in the monkey striatum. *J Neurophysiol* 95:2975–2986.
- 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.
- Richmond BJ, Optican LM (1987) Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex: II. Quantification of response waveform. *J Neurophysiol* 57:147–161.
- Rothenhoefer KM, Costa VD, Bartolo R, Vicario-Feliciano R, Murray EA, Averbach BB (2017) Effects of ventral striatum lesions on stimulus-based versus action-based reinforcement learning. *J Neurosci* 37:6902–6914.
- Saunders RC, Phil D (2006) Rhesus macaque “red” symmetrical brain atlas. Bethesda, MD: Laboratory of Neuropsychology, NIMH, NIH.
- Schulz JM, Reynolds JN (2013) Pause and rebound: sensory control of cholinergic signaling in the striatum. *Trends Neurosci* 36:41–50.
- Shidara M, Aigner TG, Richmond BJ (1998) Neuronal signals in the monkey ventral striatum related to progress through a predictable series of trials. *J Neurosci* 18:2613–2625.
- Simmons JM, Ravel S, Shidara M, Richmond BJ (2007) A comparison of reward-contingent neuronal activity in monkey orbitofrontal cortex and ventral striatum: guiding actions toward rewards. *Ann N Y Acad Sci* 1121:376–394.
- Stern EA, Jaeger D, Wilson CJ (1998) Membrane potential synchrony of simultaneously recorded striatal spiny neurons *in vivo*. *Nature* 394:475–478.
- Threlfell S, Lalic T, Platt NJ, Jennings KA, Deisseroth K, Cragg SJ (2012) Striatal dopamine release is triggered by synchronized activity in cholinergic interneurons. *Neuron* 75:58–64.
- 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.