



# Networks of VTA Neurons Encode Real-Time Information about Uncertain Numbers of Actions Executed to Earn a Reward

Jesse Wood<sup>1\*</sup>, Nicholas W. Simon<sup>2</sup>, F. Spencer Koerner<sup>3,4</sup>, Robert E. Kass<sup>3,4,5</sup> and Bitá Moghaddam<sup>6</sup>

<sup>1</sup> Department of Psychiatry, University of Pittsburgh, Pittsburgh, PA, United States, <sup>2</sup> Department of Psychology, University of Memphis, Memphis, TN, United States, <sup>3</sup> Department of Statistics, Carnegie Mellon University, Pittsburgh, PA, United States, <sup>4</sup> Center for the Neural Basis of Cognition, Carnegie Mellon University and the University of Pittsburgh, Pittsburgh, PA, United States, <sup>5</sup> Machine Learning Department, Carnegie Mellon University, Pittsburgh, PA, United States, <sup>6</sup> Department of Behavioral Neuroscience, Oregon Health and Sciences University, Portland, OR, United States

## OPEN ACCESS

### Edited by:

Bruno Poucet,  
Centre National de la Recherche  
Scientifique (CNRS), France

### Reviewed by:

Nandakumar Narayanan,  
University of Iowa, United States  
Frederic Ambroggi,  
University of California, San Francisco,  
United States

### \*Correspondence:

Jesse Wood  
jtw22@pitt.edu

**Received:** 19 April 2017

**Accepted:** 14 July 2017

**Published:** 08 August 2017

### Citation:

Wood J, Simon NW, Koerner FS,  
Kass RE and Moghaddam B (2017)  
Networks of VTA Neurons Encode  
Real-Time Information about  
Uncertain Numbers of Actions  
Executed to Earn a Reward.  
*Front. Behav. Neurosci.* 11:140.  
doi: 10.3389/fnbeh.2017.00140

Multiple and unpredictable numbers of actions are often required to achieve a goal. In order to organize behavior and allocate effort so that optimal behavioral policies can be selected, it is necessary to continually monitor ongoing actions. Real-time processing of information related to actions and outcomes is typically assigned to the prefrontal cortex and basal ganglia, but also depends on midbrain regions, especially the ventral tegmental area (VTA). We were interested in how individual VTA neurons, as well as networks within the VTA, encode salient events when an unpredictable number of serial actions are required to obtain a reward. We recorded from ensembles of putative dopamine and non-dopamine neurons in the VTA as animals performed multiple cued trials in a recording session where, in each trial, serial actions were randomly rewarded. While averaging population activity did not reveal a response pattern, we observed that different neurons were selectively tuned to low, medium, or high numbered actions in a trial. This preferential tuning of putative dopamine and non-dopamine VTA neurons to different subsets of actions in a trial allowed information about binned action number to be decoded from the ensemble activity. At the network level, tuning curve similarity was positively associated with action-evoked noise correlations, suggesting that action number selectivity reflects functional connectivity within these networks. Analysis of phasic responses to cue and reward revealed that the requirement to execute multiple and uncertain numbers of actions weakens both cue-evoked responses and cue-reward response correlation. The functional connectivity and ensemble coding scheme that we observe here may allow VTA neurons to cooperatively provide a real-time account of ongoing behavior. These computations may be critical to cognitive and motivational functions that have long been associated with VTA dopamine neurons.

**Keywords:** ventral tegmental area, VTA, dopamine, assemblies, ensemble coding, serial actions, instrumental conditioning, action sequence

## INTRODUCTION

Organization of goal-directed behavior requires real-time monitoring of ongoing actions. This function is typically assigned to the prefrontal cortex and striatum, but several lines of evidence demonstrate that organization of a series of behaviors also depends on VTA dopamine input to these regions (Goldman-Rakic, 1988; Robbins and Arnsten, 2009; Salamone et al., 2009). For instance, lesions of VTA dopamine projections to the striatum, or antagonists of dopamine receptors, impair the capacity to perform large numbers of actions without impacting the ability to complete one (or a few) action(s) (Aberman and Salamone, 1999; Ishiwari et al., 2004; Mingote et al., 2005). Additionally, VTA neurons are implicated in several cognitive constructs necessary for organizing an ongoing series of behaviors, such as working memory, attention, effort, and cognitive flexibility (Salamone and Correa, 2002; Seamans and Yang, 2004; Robbins and Arnsten, 2009; Fujisawa and Buzsaki, 2011). Thus, VTA neurons are critical for organizing, sustaining, and adapting an ongoing sequence of behaviors.

Additional evidence that VTA neurons participate in structuring and organizing a behavioral series is provided by studies showing that dopamine release in nucleus accumbens precedes a series of two lever presses (Wassum et al., 2012), and dopamine neurons encode the first and last actions in a fixed response eight (FR8) reinforcement schedule (Jin and Costa, 2010). These response patterns, however, do not reveal how VTA neurons encode information *during* a series of actions. The expansive literature on dopamine signaling has mostly used tasks that require subjects to execute a single instrumental action, or Pavlovian tasks with no explicit action requirement. These experiments have generally focused on neuronal responses before behavior is initiated or after outcomes are delivered (Ljungberg et al., 1992; Schultz et al., 1993, 1997; Montague et al., 1996; Schultz, 1998; Bayer and Glimcher, 2005; Morris et al., 2006; Roesch et al., 2007; Salamone et al., 2007; Dayan and Niv, 2008; Glimcher, 2011). Thus, little is known about how VTA neurons encode information that could be used to organize and sustain a series of actions.

To begin to understand how VTA neurons monitor a series of actions in real-time, we recorded from VTA ensembles during a task where rats executed an unpredictable number of self-paced actions to earn reward. We found that individual dopamine and non-dopamine neurons fired selectively during different actions in a trial. Whereas, population-averaging concealed this mosaic of activity patterns, it was possible to decode low, medium, and high numbered actions within a trial from the collective activity of all neurons. These data provide novel evidence of ensemble encoding by VTA neurons and provide critical insight into how VTA neurons may contribute to complex behaviors by encoding real-time information about ongoing actions.

## MATERIALS AND METHODS

### Subjects and Apparatus

All procedures were conducted in accordance with the National Institutes of Health's Guide to the Care and Use of Laboratory

Animals, and approved by the University of Pittsburgh's Institutional Animal Care and Use Committee. Behavioral and neuronal data were collected from 10 adult, experimentally naïve, male Sprague-Dawley rats (Harlan, Frederick, MD) singly housed on a 12-h light cycle (lights on at 7 p.m.). These sample sizes were chosen because they are sufficient to detect behavioral and VTA neuronal effects in a variety of rat operant procedures in our lab (Kim et al., 2010, 2012, 2016). At the time of surgery, rats weighed ~350 g (~12 weeks of age). Under isoflurane anesthesia, 8 or 16 channel, 50  $\mu$ m stainless steel, Teflon insulated, microelectrode arrays were implanted in left VTA (relative to Bregma:  $-5.30$  mm posterior, 0.8 mm lateral, and 8.3 mm ventral). All experiments were conducted in a standard operant chamber (Coulbourn Instruments, Allentown, PA). The operant chamber had a food trough on one wall and a single nose poke port on the opposite wall. Both the food trough and nose poke port could be illuminated and were equipped with an infrared beam, which detected the animal's entry. The operant chamber system controller was configured to send the time of behavioral and environmental events to the recording interface via standard TTL pulses and a digital interface.

### Behavior

Each rat was given 7 days to recover from surgery and food restricted to 90% of their free feeding body weights. Rats were habituated to handling for 5 min per day for 3 consecutive days, before being habituated to being handled and connected to a headstage cable in the procedure room for 2 additional days. Following habituation, rats were given a single 30 min magazine training session in the operant chamber, in which sugar pellets were delivered on a variable time 75 s reinforcement schedule. When each pellet was delivered, the pellet trough was illuminated for 4 s. The animal's behavior had no programmed consequences in the magazine training session.

Following the magazine training session, each animal began instrumental conditioning. During all instrumental conditioning sessions, each trial began with illumination of the nose poke port (cue light onset). This served as a discriminative stimulus that reinforcing outcomes (sugar pellets) were available (termed the "response period"), contingent upon the animal executing actions (nose pokes into the lit port). In each trial, actions were reinforced randomly, according to a predetermined probability. When an action was executed, the behavioral system controller randomly drew an outcome state (either reinforcement or no programmed consequence) with replacement, according to the probability of reinforcement. Each action was reinforced randomly and independently of the animal's action history within that trial or session. When an action was reinforced, the cue light was immediately extinguished (cue light offset) and nose pokes had no additional programmed consequences. A 0.500 s delay between the final action and outcome delivery was instituted to temporally separate these events, as done in previous work (Schultz et al., 1993). Following this delay, the outcome was delivered to the animal and the food trough was illuminated. Outcomes were delivered into the food trough from a standard pellet magazine, via the operation of a smaller stepper motor and dispenser. The food trough remained illuminated and the task

did not progress until the animal retrieved the outcome. Once the animal retrieved the outcome, a variable length intertrial interval (ITI) of 10–12 s was initiated. In each session, 180 trials were administered.

In the first instrumental conditioning session, actions were reinforced with a probability of 1 (each action was reinforced) equivalent to a fixed ratio 1 (FR01) reinforcement schedule. In the second session, the probability that an action was reinforced was decreased across three blocks of trials. In the first block of 60 trials, actions were reinforced with a probability of 1 (FR01). In the second block of 60 trials, each action had a 1 in 3 chance of being reinforced (random ratio 3, RR03). In the third block of 60 trials, the probability was further decreased to 0.2 (random ratio 5, RR05). In sessions 3 and 4, actions were reinforced with a 0.2 probability for all trials (RR05). In sessions 5–7, actions were reinforced with a probability of 0.1 for all trials (random ratio 10, RR10). In all trials but the FR01 trials, animals were required to execute an unpredictable, varying, and randomly determined number of actions per trial. Random reinforcement was utilized to limit the ability of the animal to correctly anticipate reward delivery. Actions differed from each other mainly in terms of their location within the action series in each trial (the action number within a trial, e.g., 1st action, 2nd action, 3rd action, etc.). In each trial, each animal's action rate was calculated as the number of actions divided by the duration of the response period. This served as a measure of behavioral conditioning and performance. Changes in behavior across sessions were assessed with repeated measure analysis of variance (ANOVA), and repeated measures contrasts were applied as appropriate (Kass et al., 2014). The time interval between each action (inter-action interval) was measured for bins of low, medium, and high numbered actions (actions 1–7, 8–14, and 15–21). Statistical differences between binned inter-action intervals were assessed with repeated measures ANOVA. This measured the animal's behavioral sensitivity to increasing action numbers throughout a trial. To examine the effects of the number of actions performed in a trial on fatigue, motivation, or attention, the number of actions performed in each trial was correlated with the latency to retrieve the reward, or initiate the next trial.

## Histology

Following the completion of experiments, animals were perfused with saline and brains were extracted. Each brain was stored in a mixture of sucrose and formalin. The brains were then frozen and sliced in 60  $\mu\text{m}$  coronal sections on a cryostat, before being stained with cresyl-violet. The location of each implant was histologically verified under light microscope according to Swanson's brain atlas (Swanson, 2004). Animals were excluded if electrode location could not be confirmed in VTA.

## Electrophysiology

During experiments, animals were attached to a flexible headstage cable and motorized commutator that allowed the animal to move freely about the operant chamber, with minimal disruption of behavior (Plexon, Dallas, TX). Neural data were recorded via the PlexControl software package, operating a 64-channel OmniPlex recording system (Plexon, Dallas, TX).

Neural data were buffered by a unity gain headstage and then a preamplifier. The digitized broadband signal was then band-pass filtered (100 Hz–7 KHz). High-pass filtering can affect spike waveform shapes and neuronal identification, but with freely moving animals it is necessary to apply these filters to remove artifacts from the neuronal signal (Ungless and Grace, 2012). The filter pass bands that were utilized in the current manuscript are consistent with those that have previously been used to record from dopamine containing brain regions (Schultz et al., 1993; Fiorillo et al., 2003; Tobler et al., 2005). Data were digitized at 40 KHz and continuously recorded to hard disk. Voltage thresholds were applied to the digitized spike data offline (Offline Sorter, Plexon, Dallas, TX). Single units were sorted using standard techniques, and were utilized only if they had a signal to noise ratio in excess of 2/1, and were clearly separated from noise clusters and other single unit clusters. Examples of recordings and sorting are show in Supplemental Material (**Figure S1**).

A VTA neuron was classified as dopaminergic if it had broad action potentials, > 1.4 ms in duration, and a mean baseline firing rate < 10 Hz. These criteria are similar to those used in previous studies (Hyland et al., 2002; Fiorillo et al., 2003; Anstrom and Woodward, 2005; Pan et al., 2005; Tobler et al., 2005; Anstrom et al., 2007; Totah et al., 2013). All remaining neurons were classified as non-dopaminergic. All analyses were conducted on the entire population of neurons that were recorded to provide a classification-free examination of the data. Our approach to VTA cell classification is consistent with classical techniques. It should be pointed out, however, that indirect identification (using electrophysiological features as opposed genotypic features) can lead to classification errors (Margolis et al., 2006; Lammel et al., 2008; Cohen et al., 2012). In general, qualitatively similar responses were observed in both groups of neurons. Units recorded in different sessions were considered separate units, as methods to estimate neuronal identity between sessions are not widely used with VTA recordings. In total, the following numbers of neurons were recorded in each of 7 sessions: dopamine (reward responsive): 11(7), 19(10), 21(12), 29(19), 27(13), 21(9), 27(16); non-dopamine (reward responsive): 31(17), 35(14), 39(27), 34(13), 25(12), 29(11), 27(9).

## Neuronal Data Analysis

Each single unit's spike times were binned into spike counts (0.025 s bins) within a trial. Binned spike counts were aligned to all events (e.g., cue light onset, actions, time period between cue light offset and outcome delivery, and outcome delivery). A stable, four s portion of the ITI (5 s to 1 s prior to cue light onset) served as the neuronal activity baseline. Single unit firing rates were Z-score normalized relative to baseline and zero-centered before activity was averaged together. Each unit's normalized activity was examined in 0.250 s windows around events (cue onset: +0.050 to 0.300 s, relative to cue onset; -0.125 to +0.125 s, relative to the time of action execution; time period between cue offset and outcome delivery: +0.150 to 0.400 s, relative to execution of the last action; outcome delivery: +0.050 to 0.300 s, relative to delivery). To assess between-session changes in population-level evoked activity, windowed activity was compared with a between groups two way ANOVA,

with session number and neuron type (dopamine or non-dopamine) as grouping variables. In all cases, protected Fisher's least significant difference tests were applied as appropriate.

A unit was classified as being activated or suppressed by an event if it met two criteria: (1) a significant paired samples *t*-test comparing raw (non-normalized) baseline firing rates with raw evoked firing rates, and (2) three or more consecutive bins of mean activity within the event-window, that were in excess of a 95% confidence interval around the baseline mean. With respect to a given task event, the proportions of units classified as being activated or suppressed were calculated. Differences in the proportions of activated units between sessions were compared with a Chi-squared test of independence. As expected, insufficient numbers of suppressed neurons (in some sessions zero) were obtained to permit reliable statistical analyses of this class of responses. Suppressed neurons are plotted for clarity, but not analyzed in detail.

The terminology "action-evoked" neuronal responses refers to activity around the time of action execution, without assuming that the action is solely responsible for evoking this neuronal response. Each unit's activity was examined as a function of action number (a unit's mean response to each *n*th numbered action within a trial, across all trials). These analyses were restricted to the RR10 sessions (sessions 5–7). RR10 sessions required larger numbers of actions per trial, on average, and would ensure that there were a sufficient number of higher numbered actions (e.g., actions 18, 19, 20, 21, etc.) for analysis. All action number analyses utilized actions 1 through 21. While even higher numbered actions occurred in some trials, these actions occurred less frequently and were excluded from action number analyses, as there was insufficient sample size for reliable statistical analysis. To remove any effects of impending reward delivery on the action evoked neuronal responses, only unrewarded actions were used in this analysis. Preliminary analyses suggested that including rewarded actions had little effect on the results. Action evoked neuronal responses were binned into 3 bins of 7 consecutive actions representing low, medium, and high numbered actions (actions 1–7, 8–14, and 15–21, respectively).

We measured correlations in spike count between pairs of dopamine neurons ( $n = 172$ ), pairs of non-dopaminergic neurons ( $n = 276$ ), and mixed pairs (a dopamine and a non-dopamine neuron,  $n = 387$ ) in RR10 sessions. Correlations were only calculated for pairs of simultaneously recorded neurons. Pairwise neuronal correlations were assessed in the same 0.250 s windows used to quantify firing rates, using the standard Pearson's correlation,

$$r = \frac{1}{P-1} \sum_{i=1}^P \frac{(x_i - \bar{x})(y_i - \bar{y})}{s_x s_y}$$

where the correlation coefficient,  $r$ , is the summed product of the *i*th point's residual from the sample mean,  $\bar{x}$  or  $\bar{y}$ , normalized by the sample standard deviation,  $s$ . Noise correlations represent the degree to which the joint activity of simultaneously recorded pairs of neurons fluctuates from trial to trial. Noise correlations were calculated on the trial-by-trial spike counts observed in each window. For actions, these evoked responses also examined

according to action number within a trial, or binned action (actions 1–7, 8–14, or 15–21). Signal correlations refer to the correlation between pairs of tuning curves in simultaneously recorded neurons. We corrected for chance levels of correlation, which were estimated as the mean correlation of 1,000 pairs of randomly shuffled data. Fisher's *Z* transformation

$$\frac{1}{2} * \ln\left(\frac{1+r}{1-r}\right)$$

was used to transform correlation coefficients for statistical comparisons. The geometric mean spike count was calculated to measure the average joint activity of each pair of neurons.

Individual neurons preferred different subsets of action numbers. To determine the extent to which elapsed time since trial start or action number could account for action evoked neuronal activity, we performed a Poisson regression of spike count onto these variables. The percentages of neurons with significant coefficients are reported. To visualize the various tunings of VTA neurons to action number, each neuron's mean activity as a function of action number is displayed. Because evoked firing rates could span a large range of values, each tuning curve was scaled so that the maximum evoked activity was equal to 1 and the minimum evoked activity was equal to 0. Scaled tuning curves were used only for visualization. To compute the number of maxima, each tuning curve was fit with a cubic smoothing spline (smoothing parameter 0.001) and maxima were detected as points in the estimated curve with derivatives equal to 0.

## Population Average Decoder

A decoder classified binned action number according to the trial-averaged spike count, also averaged across neurons, on an action-by-action basis. This is called the "population average decoder." Binned action number,  $A$ , is defined as 3 bins of 7 consecutive actions (1–7, 8–14, or 15–21). The classifier maximizes

$$p(A|\bar{R}) \propto p(\bar{R}|A) \times p(A).$$

Here  $p(\bar{R}|A)$  represents the probability of the population averaged response occurring, given a particular action bin. The probability of an action belonging to an action bin is denoted by  $p(A)$ , and is uniform across action bins in the current work. That is, the prior probability for each bin,  $p(A)$ , is  $\frac{1}{3}$ . For these analyses,  $p(\bar{R}|A)$  is calculated from the sample distributions of mean action evoked population responses as follows. First, the population average of all neurons' trial-averaged responses is calculated for all actions in all bins. Next, one action is selected as a test observation. Then,  $p(\bar{R}|A)$  is assumed to be the best fit Gaussian of unknown mean and variance, fit from all trials in the bin  $A$  (excluding the test observation, if it lies in  $A$ ). The result is 3 smooth distributions of mean activity, corresponding to the 3 bins of action numbers. The average activity of the test observations is calculated, and classified as the most likely action bin to have evoked this response, based on the relationship between population averaged activity and action bins in the

training dataset (the actions not tested). The classifier selects the action bin most likely to have evoked the population response via

$$\hat{A}^* = \operatorname{argmax}_A p(A | R^*) = \operatorname{argmax}_A p(R^* | A) \times p(A).$$

Here,  $R^*$  is the population average of the held out action number.  $\hat{A}^*$  is the resulting estimated action bin. Testing is repeated for each action number.

## Ensemble Decoder

Assuming  $\bar{R}$  is not a sufficient characterization of  $\{R_i\}_{i=1}^N$ , which we will call  $\bar{R}$ ,  $p(\bar{R}|A)$  does not thoroughly represent the information in  $p(\bar{R}|A)$ . Therefore, we instead assume relevant information is captured by a higher-dimensional representation:  $(W\bar{R}|A) \sim N(\mu_A, \Sigma)$ . The projection,  $W$ , estimated by principal components analysis, is necessary to ensure the covariance,  $\Sigma$ , can be estimated by the usual pooled sample covariance,  $\hat{\Sigma}$ . This maximizes the dimensionality of the space into which we project  $\bar{R}$ , for a given sample size. We then maximize  $p(A|W\bar{R})$ . This maintains the varying firing rates in each unit, elicited by actions within different bins. Note that the population average decoder is the special case of this approach wherein  $W$  is a row vector of ones. The projection and classifier can be learned with each neuron's set of evoked trial-averaged spike counts as a feature. Each action's set of observed trial-averaged spike counts, the new test observations,  $\vec{R}^*$ , are then projected via the weights estimated from the training data and classified sequentially (that is, we estimate the set of  $\{\hat{\mu}_A\}_{A=1}^3$  on all the remaining action numbers, which is equivalent to training, and maximize the analogous posterior).

## Statistical Testing of the Decoders

To test whether the two decoders produce equal or different classification accuracy, we fit a binomial generalized linear model and controlled for action number, session, and unit type (dopaminergic vs. non-dopaminergic). For each trial of action bin  $A$ , tested on decoder  $d$ , the log odds of the probability of correct classification was assumed to be linear in its covariates as given by the logistic regression function

$$\logit \left( \mathbb{E} \left[ I_A(\hat{A}) \mid (I_d, I_A, I_S, I_{DA}) \right] \right) = \beta_0 + \beta_d \cdot I_d + \beta_A \cdot I_A + \beta_S \cdot I_S + \beta_{DA} \cdot I_{DA},$$

where  $I_A(\hat{A})$  is an indicator signifying correct or incorrect classification, and  $I_d, I_A, I_S, I_{DA}$  are indicators (or a set of indicators) for the decoder, action bin, session, and dopaminergic unit, respectively. Likelihood ratio tests were used to test the effect of each group of factors, corresponding to each feature. Additionally, statistical significance of decoder performance (vs. chance levels of correct classification) at each action number was determined via permutation test. Approximations to the exact  $p$ -values can be determined by assigning classes to each action according to a uniform prior distribution, and calculating the proportion of sets of resulting classifications that perform better than the classification rates observed using the previously discussed methods.

## RESULTS

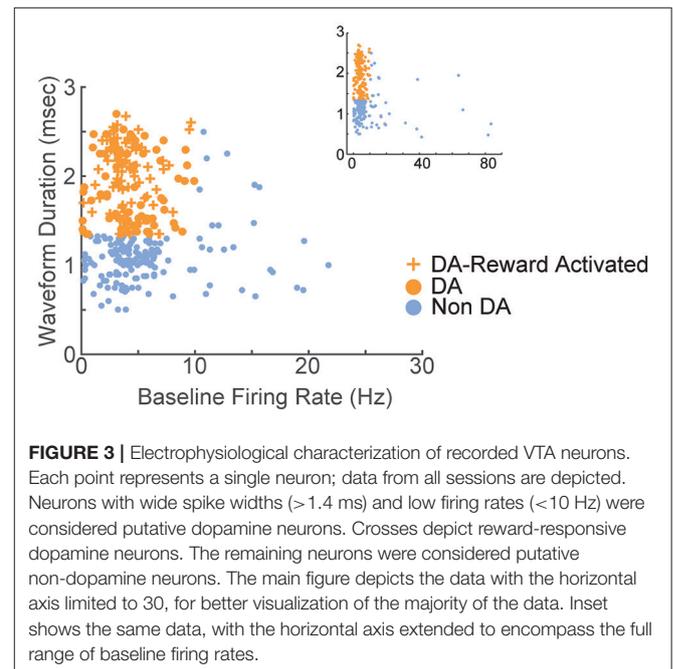
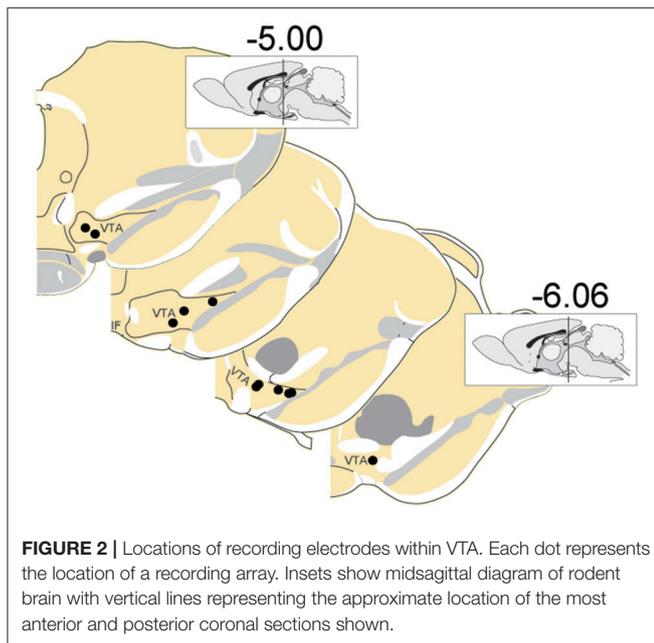
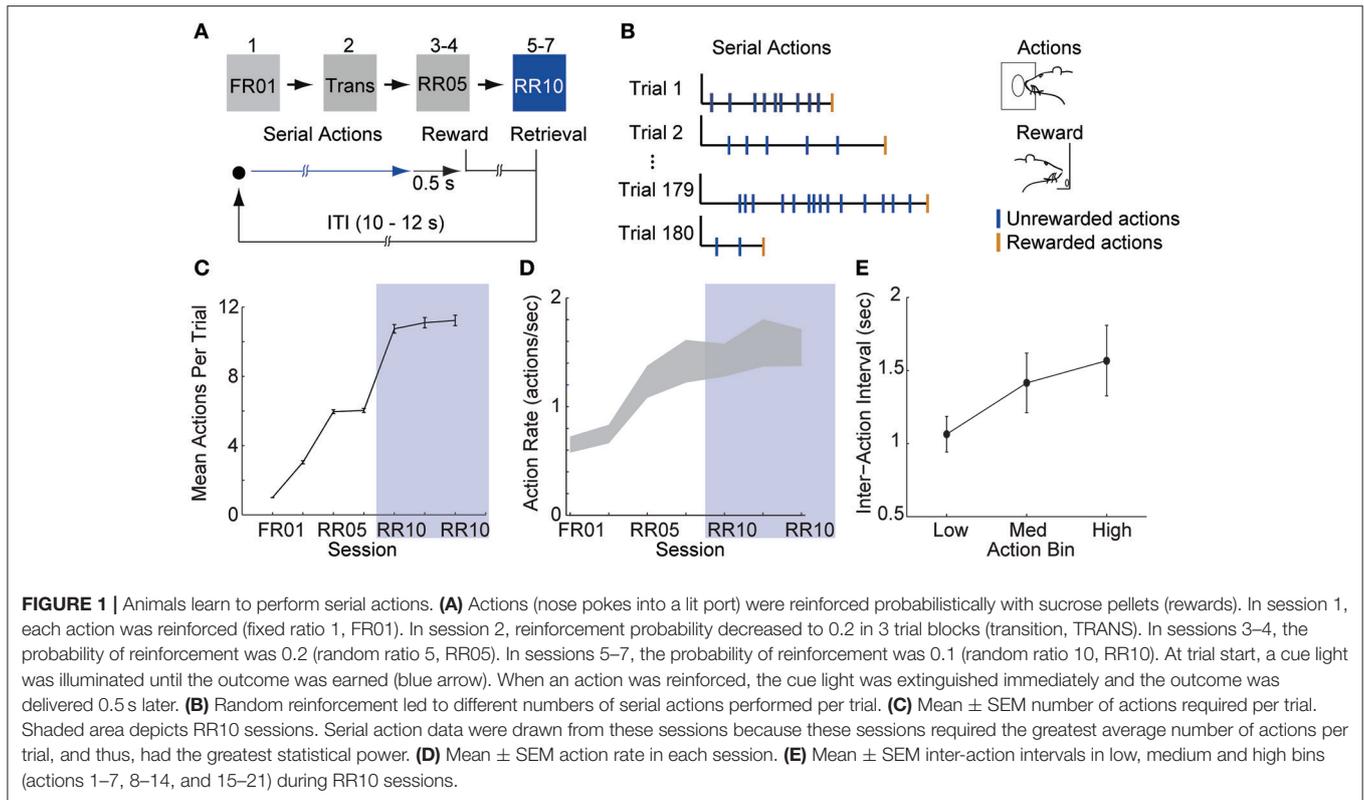
### Animals Learn to Execute Serial Actions for Reward

VTA activity was recorded while rats learned to execute multiple actions for random reinforcement with one sugar pellet (Figures 1A,B). In the first session, each action was rewarded (FR01). In session 2, the reward probability was decreased from 1 to 0.2 across 3 blocks of trials. In sessions 3 and 4, actions were reinforced at a probability of 0.2 (RR05). In sessions 5–7, actions were reinforced at a probability of 0.1 (RR10). In all randomly reinforced trials, unpredictable and varying numbers of actions were required, but each action was equally likely to be reinforced.

Action response rates increased as reinforcement probability decreased (Figures 1C,D), and the response rate in the final RR10 sessions was significantly higher than all other sessions [Figure 1D;  $F_{(6, 24)} = 4.726$ ,  $p = 0.003$ ]. To investigate a potential relationship between behavioral performance and action number, we divided the RR10 data into three bins: actions 1–7 (low), actions 8–14 (medium), and actions 15–21 (high). We analyzed serial actions in trials reinforced with a probability of 0.1 (RR10 sessions, sessions 5–7), because this reinforcement schedule produced the greatest average number of actions performed per trial (Figure 1C). Within a trial, the interaction interval significantly increased with increasing action number bins [Figure 1E;  $F_{(2, 44)} = 11.020$ ,  $p < 0.001$ ]. This indicated that behavior was sensitive to action number. The latency to retrieve the reward [ $r_{(3, 468)} = -0.017$ ,  $p = 0.309$ ] or initiate responding in the next trial, however, was not correlated with binned action number [ $r_{(3, 445)} = 0.015$ ,  $p = 0.389$ ]. These data suggest that the number of actions performed in a trial did not affect behavioral responses to reward or cues.

### Neurophysiological Recordings

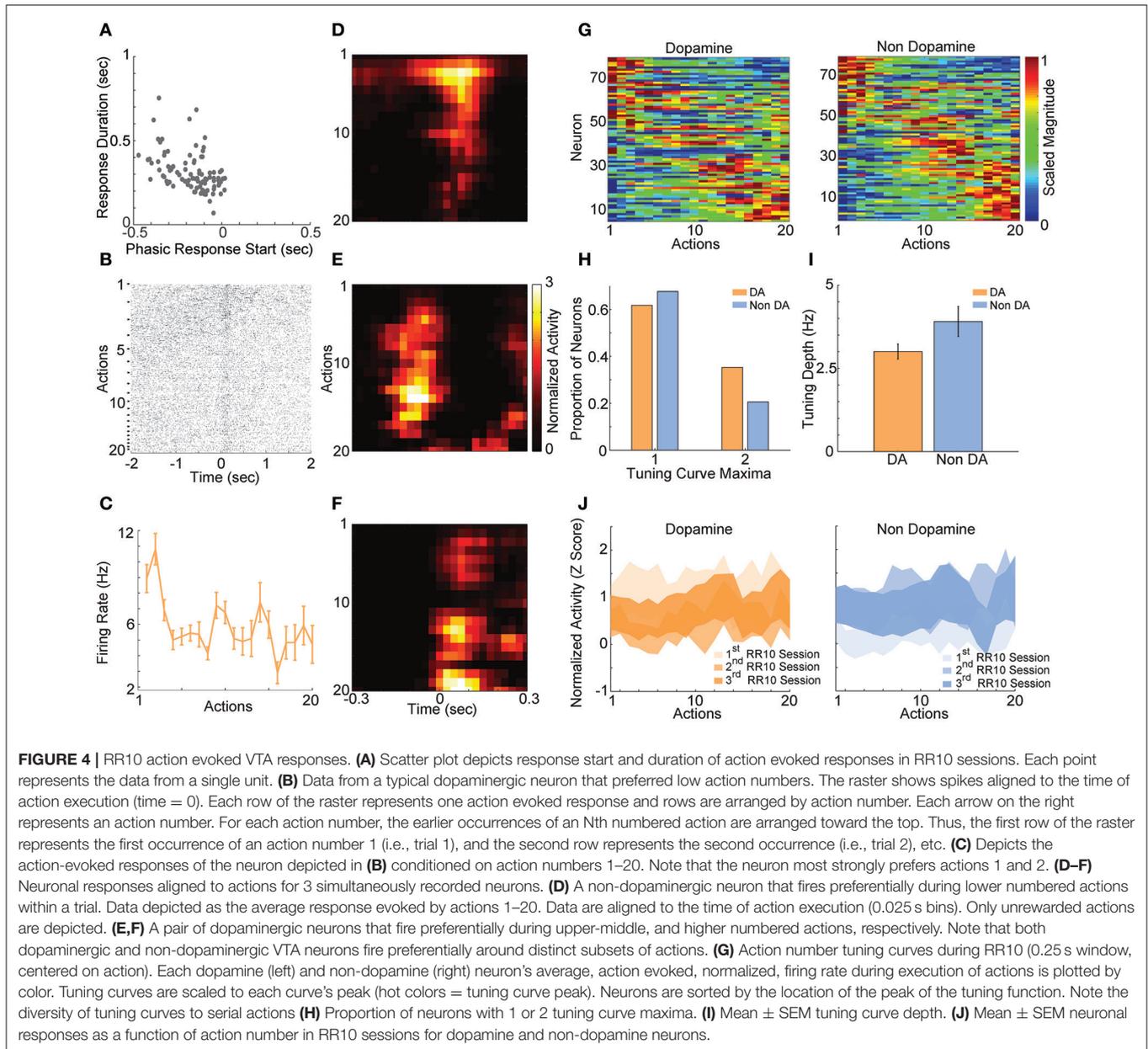
A total of 375 units (in 7 sessions and from 10 rats) were recorded from electrodes histologically localized to VTA (Figure 2). All neurons were included in these analyses because we were interested in understanding the full diversity of activity patterns in the VTA. Neurons were classified (Figure 3) as putative “dopamine” ( $n = 155$ ) or “non-dopamine” neurons ( $n = 220$ ) based on electrophysiological criteria. This classification approach permits comparison with previous work, despite potential inaccuracies (Margolis et al., 2006) and with the caveat that some dopamine neurons co-release other neurotransmitters (Tritsch et al., 2012). We did not observe strong clustering in the electrophysiological profiles of these neurons. It is important to underscore that neurons were collected with no attempt to select for particular electrophysiological characteristics. When VTA neurons are recorded in an unbiased manner, as done here, weak clustering may be expected (Kiyatkin and Rebec, 1998). In addition to the above classification, analyses were performed on reward-responsive dopamine neurons, as this subgroup may be a more conservative estimate of dopamine neuron identity (Lak et al., 2014; Eshel et al., 2016).



## VTA Neurons Are Diversely Tuned to Serial Actions

In order to understand how VTA neurons encode information during serial actions, we examined how activity was modulated by actions in RR10 sessions. Action evoked responses initiated

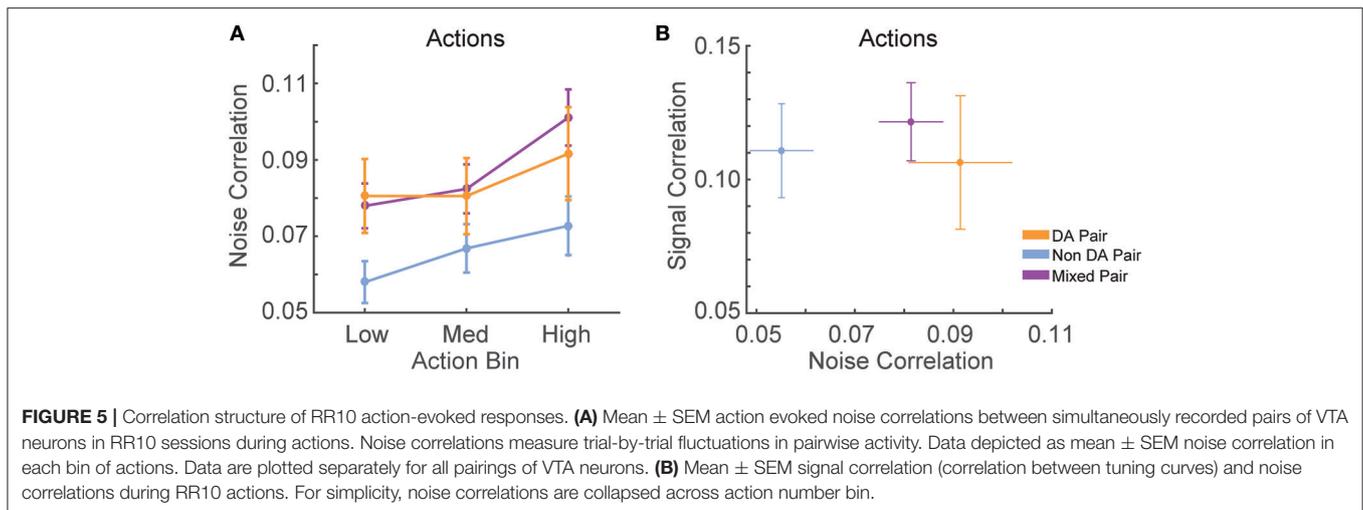
$-0.173 \pm 0.013$  s prior to actions and had an average duration of  $0.313 \pm 0.012$  s (**Figure 4A**). Thus, the majority of neurons had response durations several 100 ms faster than the majority of inter-action intervals, and therefore, action-evoked responding



was well-separated between successive actions in a trial. When RR10 actions were examined according to their number within a trial (e.g., 1st, 2nd, 3rd action in a trial), we observed that individual neurons fired selectively for subsets of actions (**Figures 4B,C**). We observed preferential responding to low, medium, or high numbered actions within the population of simultaneously recorded neurons (**Figures 4D–F**). To determine how VTA neurons responded to serial actions, we calculated each neuron's tuning curve as a function of action number within a trial (**Figure 4G**). Action tuning curves mostly had a single peak [local maxima,  $\chi^2_{(1)} = 9.921$ ,  $p = 0.002$ ; neuron type,  $\chi^2_{(1)} = 1.267$ ,  $p = 0.260$ ], and different neurons preferred different subsets of actions (**Figures 4G,H**). Tuning curve depth

(maximum–minimum) did not differ significantly between putative dopamine and non-dopamine neurons [**Figure 4I**;  $t_{(154)} = 3.224$ ,  $p = 0.084$ ]. Averaging activity across neurons, a traditional approach to analyzing VTA neuronal data, concealed this heterogeneity in both groups of neurons (**Figure 4J**).

The diversity of the tuning curves of VTA neurons suggested that network properties may be critical to information processing during serial actions. To gain insight into VTA network-based information processing, we calculated the trial-by-trial correlations in spike counts (noise correlations), and correlations between tuning curves (signal correlations) for all simultaneously recorded pairs of neurons. Noise correlations reflect functional connectivity between neurons while signal



correlations reflect similarity in tuning curves (Cohen and Kohn, 2011). Correlations were examined between all possible pairings of VTA neurons (dopamine–dopamine, non-dopamine–non-dopamine, and dopamine–non-dopamine).

We found that noise correlations during actions were significantly lower in pairs of non-dopamine neurons than pairs containing a dopamine neuron [Figure 5A;  $F_{(2, 858)} = 4.052$ ,  $p = 0.018$ ]. The magnitude of these correlations was decreased during low numbered actions compared to medium or high numbered actions [ $F_{(2, 1,716)} = 14.625$ ,  $p < 0.001$ ]. It is unlikely that these differences were due to differences in spike count because the pairwise geometric mean spike count was not associated with noise correlation magnitude in any bin (low:  $r = 0.008$ ,  $p = 0.805$ ; med:  $r = -0.015$ ,  $p = 0.650$ ; high:  $r = -0.048$ ,  $p = 0.158$ ). In contrast, signal correlation strength did not differ between different pairings of VTA neurons [Figure 5B;  $F_{(2, 858)} = 0.289$ ,  $p = 0.749$ ]. There was a strong association between signal correlations and noise correlations (Figure 5B, Table 1), suggesting that there is a high degree of functional connectivity between pools of similarly tuned neurons. This result raises the possibility that different patterns of functional connectivity contribute to the diversity of tunings to actions.

## Actions in a Trial Are Accurately Discriminated from VTA Neuronal Ensemble Activity

The combination of tuning curve heterogeneity, lack of population average signal for serial actions, and the association between action-evoked signal and noise correlations suggested that information about ongoing sequences of actions may be encoded by VTA ensembles. To quantify the accuracy of this encoding mechanism, we differentiated low, medium, and high action number bins from either population-averaged activity (Figure 6A) or the collective activity of VTA ensembles (Figure 6B). The same cross-validated linear discriminant analysis decoding algorithms were used for both the ensemble and population-average data (see Methods). Ensemble activity

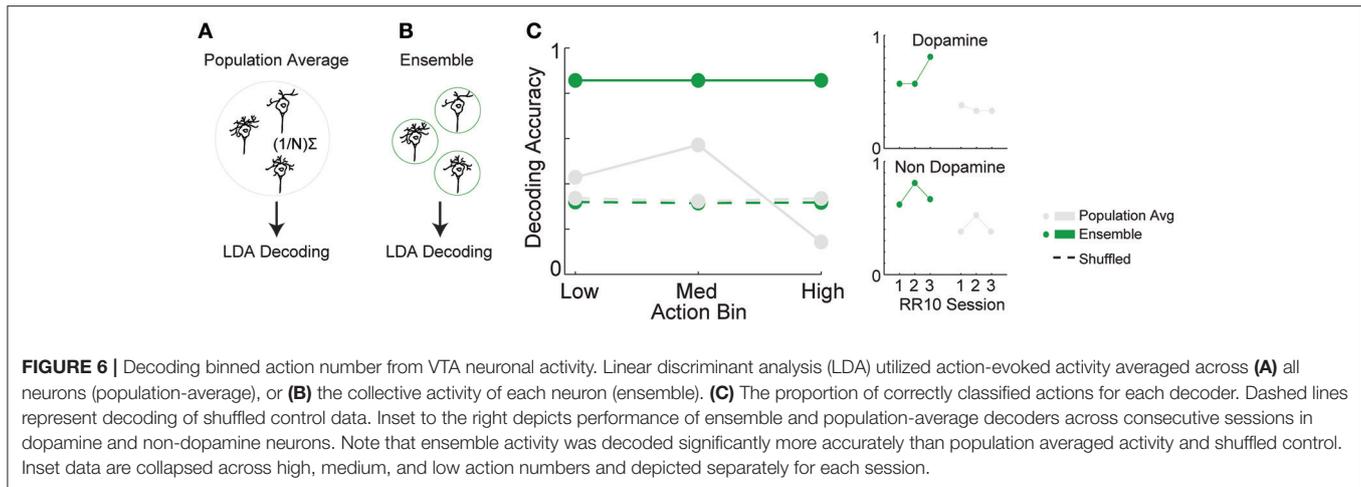
**TABLE 1 |** Table depicts the correlation between action number signal correlations and action evoked noise correlations.

	Rho	p-value
DA–DA Pair	0.571	<0.001
DA–NDA Pair	0.569	<0.001
NDA–NDA Pair	0.528	<0.001

Data are taken from RR10 sessions. Data depicted separately for each pair type (DA-DA, dopamine pair; DA-NDA, dopamine and non-dopamine pair; NDA-NDA, non-dopamine pair).

was discriminated more accurately than shuffled control data (Figure 6C; 85% accuracy vs. chance levels of correct decoding [33% accuracy], permutation test,  $p < 0.001$ ). In contrast, population averaged activity was decoded at chance levels (Figure 6C; permutation test,  $p = 0.240$ ). Binned action number was discriminated from ensemble activity significantly more accurately than population-averaged activity [Figure 6C;  $\chi^2_{(1)} = 21.435$ ,  $p < 0.001$ ]. Discrimination was stable across sessions [ $\chi^2_{(2)} = 1.097$ ,  $p = 0.578$ ], and between low, medium, and high action number bins [ $\chi^2_{(2)} = 3.614$ ,  $p = 0.164$ ]. There was no difference in accuracy between dopamine and non-dopamine neurons [ $\chi^2_{(1)} = 1.133$ ,  $p = 0.287$ ]. When reward responsiveness was also included in the model, dopamine and non-dopamine neurons were decoded at similar accuracy [ $\chi^2_{(1)} = 1.300$ ,  $p = 0.254$ ]. These data suggest that real-time information about ongoing behavior is encoded by the collective activity of VTA neurons, but not the averaged activity of VTA dopamine or non-dopamine neurons. Thus, VTA ensemble activity accurately predicted the current action bin, and could be used to organize, pace, and sustain effortful behaviors.

To determine if serial actions or the passage of time preferentially modulated neuronal activity, we examined the relationship between action evoked spike counts and action number, or the corresponding time within a trial. Action number predicted spike counts of a significantly larger proportion of neurons than time elapsed in a trial [ $\chi^2_{(1)} = 20.643$ ,  $p < 0.001$ ;



29 vs. 7% of neurons, respectively]. Actions did not preferentially modulate the activity of either dopamine or non-dopamine neurons [ $\chi^2_{(1)} = 0.179$ ,  $p = 0.673$ ]. These data suggest action-evoked activity reflected ongoing behavior and do not support a role for encoding the passage of time in this task.

Neuronal responses may have represented reward anticipation (Bromberg-Martin et al., 2010; Totah et al., 2013) in a manner which changed throughout a series of actions. If this were the case, the number of actions performed in a trial would be correlated with neuronal activity in a consistent direction. Thus, we examined the correlation between neuronal responses and the number of actions in each trial. Only a small proportion of dopamine and non-dopamine neurons had significant correlations between the response evoked by the first action in a trial and the number of actions in the previous trial (**Table 2**). Similar proportions of these neurons were positively or negatively correlated, indicating that this relationship was inconsistent between neurons (**Table 2**). Likewise, very few neurons had responses to the last action in each trial that were significantly correlated with the number of actions in that trial; there was no significant difference in the proportion of dopamine and non-dopamine neurons with positive or negative correlations (**Table 2**).

We further investigated if reward anticipation modulated neuronal responses by examining the relationship between responses evoked by the cue light turning off, reward delivery, or reward retrieval, and the number of actions performed in each trial. We found no significant differences in the proportion of dopamine or non-dopamine neurons with positive or negative correlations between responses evoked by cue onset and the number of actions in the previous trial or between cue light offset, reward delivery, or reward retrieval responses and the number of actions in the current trial (**Table 2**). Taken together, these results indicate that performance of serial actions did not influence encoding of reward anticipation, supporting the notion that VTA neurons encoded information about ongoing behaviors as opposed to information about time or reward anticipation in the current task.

**TABLE 2** | Table depicts the percentage of neurons with significantly correlated responses between task events and the number of actions in each trial.

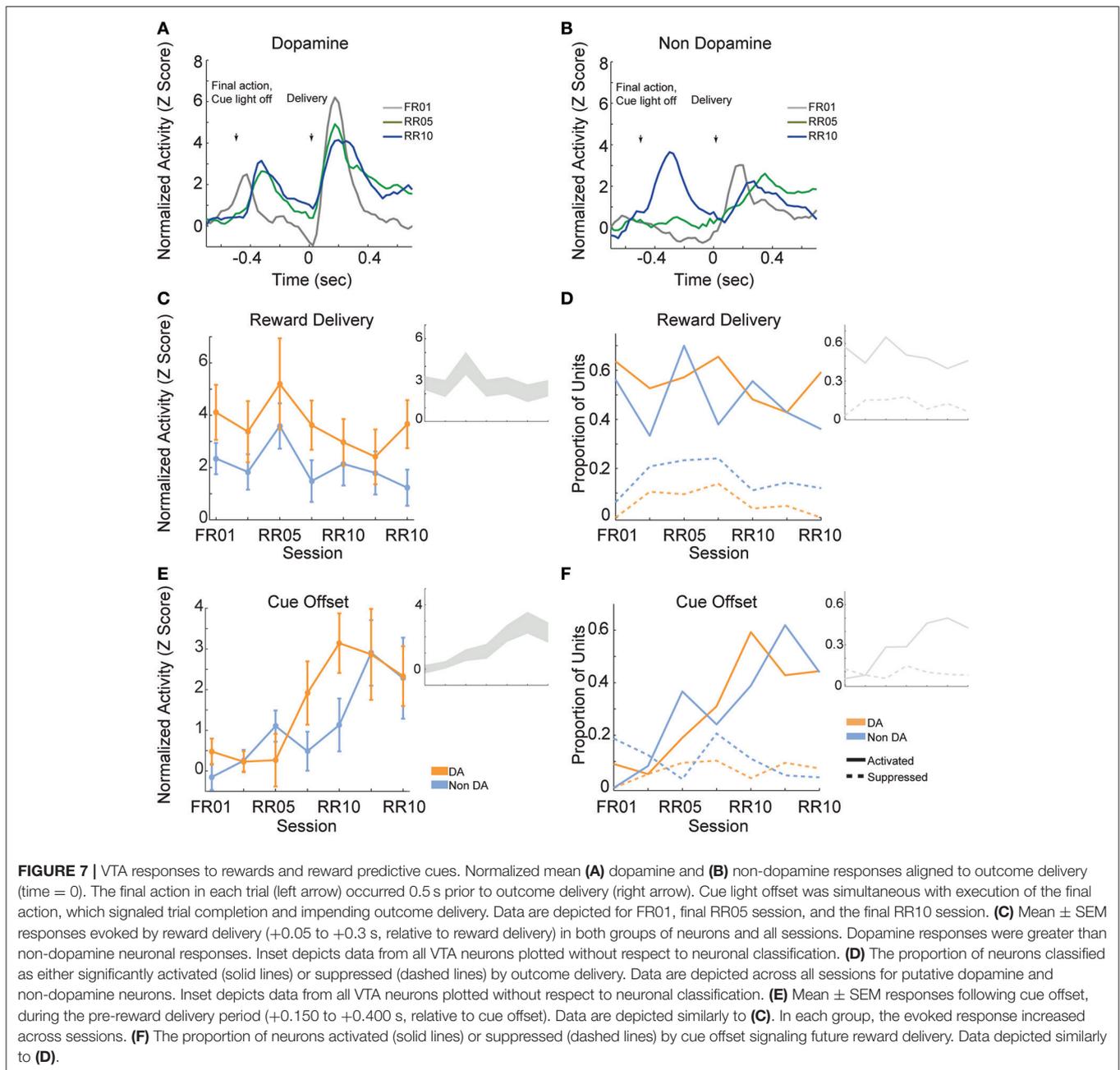
	DA positive corr. (%)	DA negative corr. (%)	Non-DA positive corr. (%)	Non-DA negative corr. (%)	Fisher's exact test
Response to first action and previous trial's action number	4	1.33	3.70	1.23	$P = 0.71$
Response to last action and current trial's action number	2.67	1.33	4.94	4.94	$P = 0.77$
Response to cue on and previous trial's action number	0.00	2.67	2.47	1.23	$P = 0.25$
Response to cue off and current trial's action number	6.67	2.67	11.11	4.94	$P = 0.81$
Response to delivery and current trial's action number	5.33	5.33	3.70	2.47	$P = 0.80$
Response to retrieval and current trial's action number	1.33	1.33	1.23	1.23	$P = 0.67$

Data are taken from RR10 sessions. Pearson's correlations were calculated on spike counts taken from 0.25 s windows around events. The far right column depicts the p-value corresponding to a Fisher's exact test of these proportions

## Reward Delivery Evoked Responses

Reward delivery evoked greater activity in dopamine neurons than in non-dopamine neurons [**Figures 7A,B**;  $F_{(1, 361)} = 9.159$ ,  $p = 0.003$ ]. This pattern did not change across sessions [**Figure 7C**;  $F_{(6, 361)} = 1.352$ ,  $p = 0.233$ ]. The proportion of dopamine neurons responding to reward delivery did not vary significantly across sessions [**Figure 7D**; all neurons,  $\chi^2_{(6)} = 9.289$ ,  $p = 0.158$ ; non dopamine,  $\chi^2_{(6)} = 13.234$ ,  $p = 0.039$ ].

After the final action in a trial, the cue light was immediately extinguished and reward was delivered 0.5 s later. We hypothesized that cue light offset would evoke VTA

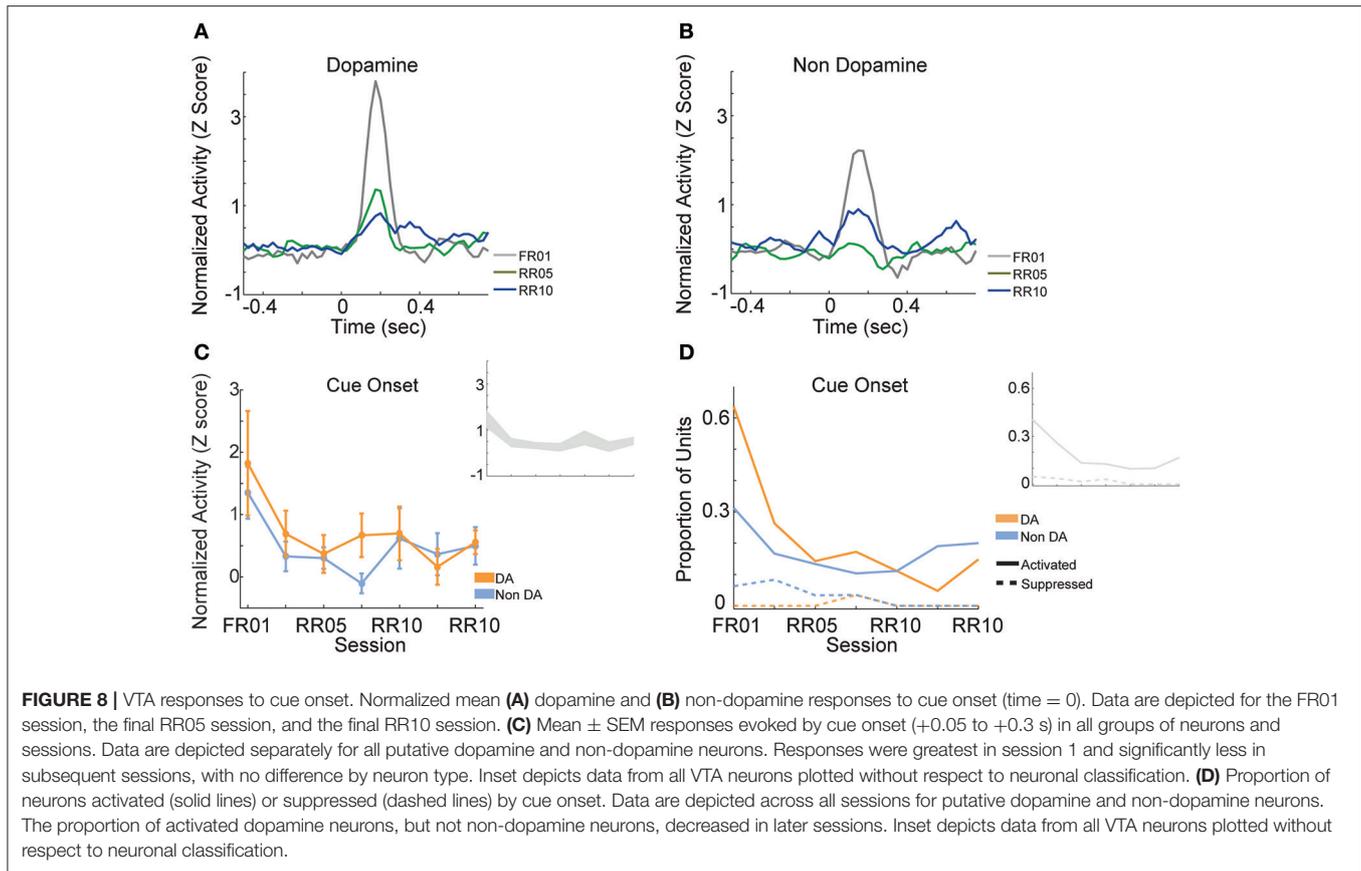


responses, similar to a reward prediction error, because this event signaled new information about unpredictable outcome delivery (Montague et al., 1996; Schultz, 1998). As training progressed, cue offset evoked great activity levels [Figure 7E;  $F_{(6, 361)} = 4.776$ ,  $p < 0.001$ ], and activated a greater proportion of neurons [Figure 7F; all neurons,  $\chi^2_{(6)} = 45.949$ ,  $p < 0.001$ ; dopamine neurons,  $\chi^2_{(6)} = 22.093$ ,  $p = 0.001$ ; non dopamine neurons,  $\chi^2_{(6)} = 29.744$ ,  $p < 0.001$ ]. This finding confirms that we observed responses to cues and rewards that are commonly attributed to dopamine neurons (Montague et al., 1996; Schultz, 1998), but

also occur in non-dopaminergic cells in the VTA (Kim et al., 2010, 2012; Cohen et al., 2012).

## Cue Evoked Responses

When animals were required to perform a single action to earn rewards (FR01), the majority of dopamine neurons encoded cue onset (Figure 8A), which is consistent with previous observations (Schultz, 1998; Roesch et al., 2007). Non-dopamine neurons followed a similar pattern (Figure 8B). During random ratio sessions, the population response to the



cue decreased significantly as response requirement increased [Figure 8C; session,  $F_{(6, 361)} = 2.667, p = 0.015$ ]. This pattern of responding did not differ between dopamine and non-dopamine neurons [Figure 8C;  $F_{(1, 361)} = 1.543, p = 0.215$ ]. We found the same effects when analysis was restricted to neurons that were significantly activated by reward delivery [session,  $F_{(6, 175)} = 4.961, p < 0.001$ ; neuron type,  $F_{(1, 175)} = 3.912, p = 0.050$ ]. The data among neurons not responsive to reward delivery were inconclusive, but followed the same trend [session,  $F_{(6, 172)} = 2.096, p = 0.056$ ; neuron type,  $F_{(1, 172)} = 0.824, p = 0.365$ ]. Likewise, significantly fewer neurons were modulated by cue light onset in random ratio sessions [Figure 8D; all VTA,  $\chi^2_{(6)} = 23.844, p = 0.001$ ; dopamine,  $\chi^2_{(6)} = 20.109, p = 0.003$ ; non-dopamine,  $\chi^2_{(6)} = 10.636, p = 0.100$ ].

## Cue and Reward Evoked Responses Are Decoupled When Performing Serial Actions

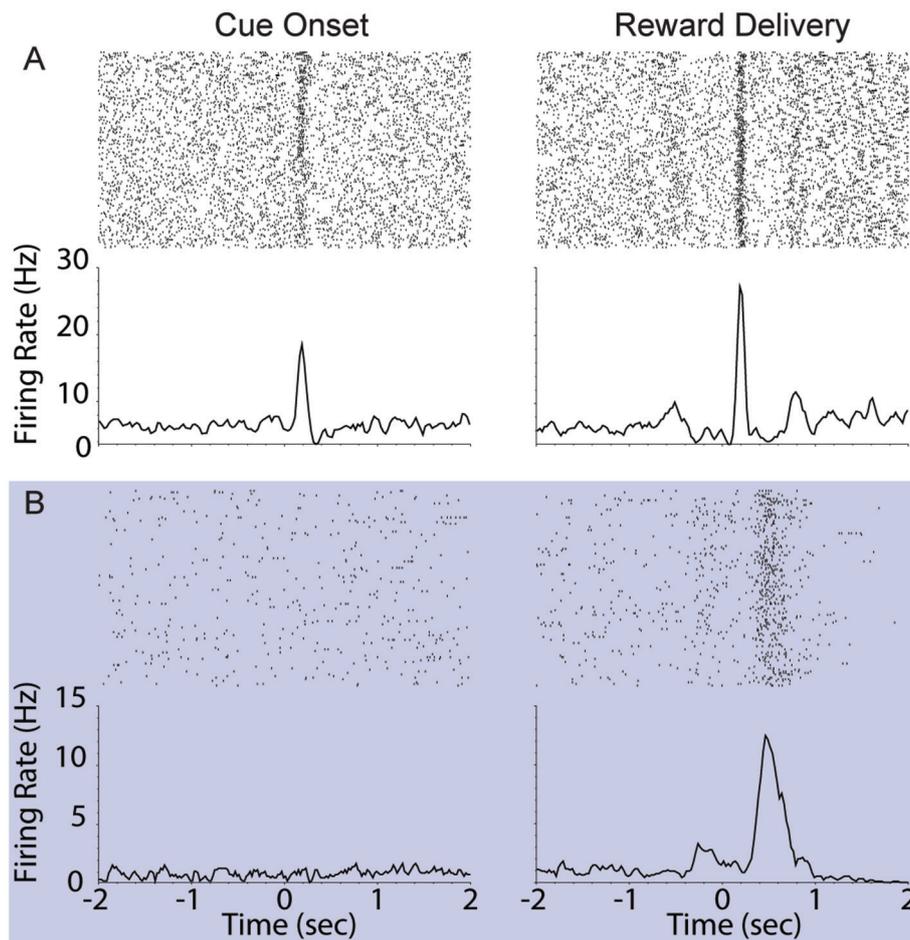
We assessed the correlation between each neuron's responses to cue onset and reward delivery in order to understand the relationship between encoding of these events. During FR01 sessions, responses evoked by these events were significantly correlated in both dopamine and non-dopamine neurons (Table 3). This correlation decreased in RR05 sessions (Table 3), and was no longer significant by RR10 sessions (Table 3). A similar pattern was observed in dopamine neurons that were activated by reward delivery (Table 3). This suggests that,

**TABLE 3 |** Table depicts the correlation between cue onset and reward delivery evoked responses.

	FR01	RR05	RR10
Dopamine	$r = 0.653$ $p = 0.029$	$r = 0.321$ $p = 0.023$	$r = 0.096$ $p = 0.411$
Reward responsive dopamine	$r = 0.496$ $p = 0.014$	$r = 0.280$ $p = 0.018$	$r = 0.231$ $p = 0.054$
Non-dopamine	$r = 0.517$ $p = 0.003$	$r = 0.584$ $p < 0.001$	$r = 0.002$ $p = 0.981$

Data depicted separately for dopamine neurons, reward responsive dopamine neurons, and non-dopamine neurons, in FR01, RR05, and RR10 sessions. Note that the correlation between cue and reward delivery evoked response magnitudes decreases from FR01 to RR10 sessions.

although cue and reward responses are positively correlated in Pavlovian conditioning tasks (Eshel et al., 2016), cue-evoked responses become decoupled from reward responses during an instrumental random reinforcement schedule. Figure 9 depicts two example neurons in the FR01 session (Figure 9A) and the final RR10 session (Figure 9B). Note that this example neuron has similar responses to cue onset and reward delivery in the FR01 session (Figure 9A), while the example neuron (same animal) recorded during the RR10 session does not respond to cue onset, but is activated by reward delivery (Figure 9B).



**FIGURE 9** | Example VTA responses to cue onset and reward delivery in first session (FR01) and final RR10 session. **(A)** Top left panel depicts representative example raster of cue evoked response (cue onset at time = 0; trials arranged in chronological order starting at top). Each dash represents one spike. Smoothed mean response is depicted below raster. The same neuron's response to reward delivery is depicted in the upper right panel. Note the phasic response to both cue onset and reward delivery in the FR01 session (top row). **(B)** Representative example neuronal responses recorded in final RR10 session [data recorded from same animal in **(A,B)**]. Data depicted similarly to **(A)** with different Y axis on bottom panels. Note that this neuron responds to reward delivery, but not cue onset.

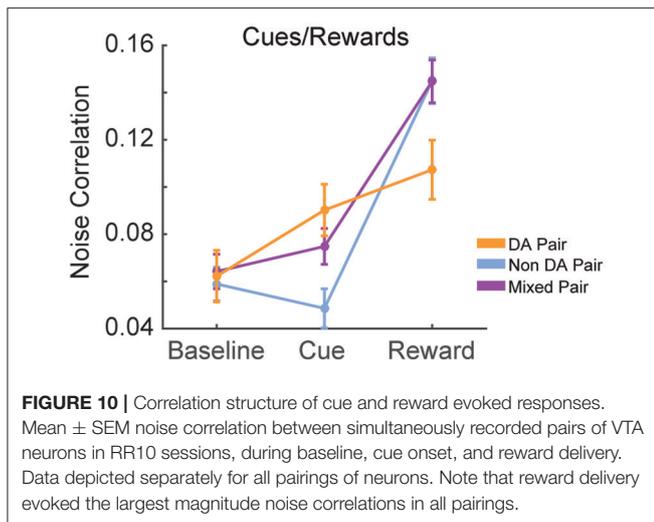
## Modulation of Noise Correlations by Cue Onset and Reward Delivery in Serial Action Trials

We calculated the signal and noise correlation between pairs of VTA neurons during cues and reward delivery in RR10 sessions. We found a significant interaction between pair type (dopamine pairs, non-dopamine pairs, and mixed pairs) and event [**Figure 10**;  $F_{(4, 2,496)} = 9.482, p < 0.001$ ]. There was no difference in baseline noise correlations between any pairing of VTA neurons [**Figure 10**;  $F_{(2, 832)} = 0.855, p = 0.426$ ]. Noise correlations evoked by cue onset were significantly greater in pairs containing a dopamine neuron than pairs of non-dopamine neurons [**Figure 10**;  $F_{(2, 832)} = 5.881, p = 0.003$ ]. There was no association between cue evoked noise correlations and pairwise cue evoked geometric mean spike counts ( $r = 0.008, p = 0.817$ ), suggesting that mean activity level did not account for this difference. Noise correlations were strongest during reward

delivery overall. Noise correlations during reward delivery did not differ significantly between pairings of neurons [**Figure 10**;  $F_{(2, 832)} = 2.794, p = 0.062$ ]. These data confirm previous findings that correlated activity in VTA circuits emerges when rewarding outcomes can be earned (Joshua et al., 2009; Kim et al., 2012; Eshel et al., 2016).

## DISCUSSION

We investigated how VTA neurons encode information when animals execute an uncertain numbers of actions to earn a reward. Dopamine and non-dopamine VTA neurons were preferentially tuned to different subsets of binned action number in a manner that allowed serial actions to be accurately discriminated from the collective activity of VTA ensembles. In addition, pairs of neurons with similar action tuning curves had higher action-evoked noise correlations suggesting that VTA



network properties may be critical to information processing during serial actions. Our analyses further showed that this pattern of tuning is distinct from reward prediction related activity or the passage of time. We also demonstrated that cues preceding a series of actions, unlike a single rewarded action, do not evoke neuronal responses and are decoupled from reward, suggesting that cue-related information is utilized differently when a single vs. an unpredictable series of actions are required to obtain reward.

## Encoding Information during an Action Series

Previous reports indicate that dopamine neurons encode errors in the predicted value of rewards as well as action values (Montague et al., 1996; Schultz, 1998; Morris et al., 2006; Roesch et al., 2007). Further, dopamine neurons phasically fire during the first and last action in a sequence, which may be critical to initiating and terminating behavior (Jin and Costa, 2010). Our data indicate that the collective activity of ensembles of VTA neurons also has the capacity to encode information about ongoing behavior in real-time.

The ensemble code revealed here may be used to organize and sustain actions, which is a fundamental feature of dopamine's role in cognition (Aberman and Salamone, 1999; Ishiwari et al., 2004; Salamone et al., 2009). In particular, this signal could be critical for supporting the unique motivational and cognitive demands when large numbers of actions are required (Goldman-Rakic, 1998; Aberman and Salamone, 1999; Salamone and Correa, 2002; Seamans and Yang, 2004; Robbins and Roberts, 2007; Salamone et al., 2007; Robbins and Arnsten, 2009). VTA dopamine neurons project to several networks that may utilize this signal to organize behavior. For instance, dopamine innervation of the ventral striatum is selectively required for completing a lengthy series of instrumental actions, but not for executing small numbers of actions for reward (Aberman and Salamone, 1999; Ishiwari et al., 2004; Mingote et al., 2005). Dopamine projections to the prefrontal cortex are important for working memory and

related constructs that require real-time information about ongoing behavior (Goldman-Rakic, 1998; Seamans and Yang, 2004; Robbins and Roberts, 2007; Robbins and Arnsten, 2009). VTA neurons could contribute to these functions by continually encoding information throughout execution of a lengthy series of actions, allowing striatal and prefrontal networks to track goal-directed effort expenditure and progress.

Our analyses suggest that VTA responses do not represent the passage of time in the current task. Although dopamine responses are sensitive to the timing of reward delivery and necessary for perception of time and interval timing (Lake and Meck, 2013; Bermudez and Schultz, 2014; Parker et al., 2015; Soares et al., 2016), neuronal responses were not consistently modulated by time elapsed in a trial. Instead, activity during a series of actions was most strongly modulated by action execution. This is likely because reward delivery was not contingent upon elapsed time, and VTA neurons are sensitive to the contingency between action execution and outcome delivery. Experimental designs with an explicit contingency between elapsed time and reward delivery, such as a fixed interval reward schedule, may reveal a more complex relationship between action execution and within trial timing in the responses of VTA neurons.

We also found that the number of actions required in RR10 trials did not modulate VTA responses to the rewarded action or the first action in the next trial. This suggests that recent trial history does not affect action evoked responses at the beginning or end of a series of actions. Similarly, VTA responses to stimuli associated with reward were not consistently correlated with the number of actions that an animal performed, suggesting that VTA correlates of reward anticipation did not increase or decrease according to action number. This result was expected, as each action was reinforced with equal probability.

## VTA Ensemble Encoding

Neurons with similar action number tuning curves had higher action-evoked noise correlations in spike count, which reflect shared connectivity between neurons (Cohen and Kohn, 2011). This suggests that different action number tunings may arise from unique inputs or differing connection strengths between inputs and VTA neurons with similar tunings may share connection properties. VTA receives inputs from an expansive set of afferent structures (Geisler and Zahm, 2005; Geisler et al., 2007), and the diversity of inputs that converge in VTA may contribute to preferential firing in different subsets of serial actions.

The association between noise correlations and action selectivity indicates that VTA network properties are critical to understanding how information about serial behavior is encoded. Networks can encode information redundantly or through diverse activity patterns and heterogeneous tunings naturally lead to the capacity to encode information as ensembles. Accordingly, serial actions were accurately discriminated from ensemble activity, but not from population-averaged activity. Though previous studies suggest VTA activity is highly redundant (Joshua et al., 2009; Schultz, 2010; Glimcher, 2011; Kim et al., 2012; Eshel et al., 2016), this work was limited to tasks requiring very few actions. The diversity of VTA tuning curves may increase

to match the expanded behavioral state space of serial actions compared with the limited state space of single action trials (Eshel et al., 2016). Thus, ensemble encoding of information may occur when the complexity of a task increases, such as during ongoing serial actions.

Units classified as either putative dopamine or non-dopamine neurons had comparable mixed tunings to serial actions, suggesting that dopamine neurons can cooperate with non-dopamine VTA neurons to encode information about serial actions. These ensemble signals consisting of multiple neurotransmitters could allow information to be decoded through multiple signaling mechanisms, which may diversify the spatiotemporal properties of this signal (Seamans and Yang, 2004; Kim et al., 2010; Barker et al., 2016). The activity of different types of VTA neurons must ultimately be coordinated and unified to represent information, and the present work demonstrates how different types of VTA neurons could collectively encode information about an ongoing series of actions.

## Cue and Reward Evoked Responses

In the first session, animals executed a single action to earn rewards (FR01). Similar to previous studies (Ljungberg et al., 1992; Schultz et al., 1993; Morris et al., 2006; Roesch et al., 2007), cue light onset at the start of the trial evoked homogenous, phasic population activation in VTA neurons. There was also a significant correlation between cue and reward delivery evoked responses. These observations are consistent with a convincing body of literature suggesting that dopamine neurons encode action value predicted by cues in the environment (Montague et al., 1996; Dayan and Abbott, 2001). On the other hand, when animals were required to execute multiple and random numbers of actions (random reinforcement schedules) to earn rewards we observed weakening of two critical neuronal responses by VTA neurons: (1)-loss of cue-evoked responses and (2) decorrelation of cue and reward evoked responses. This suggests that VTA neuronal responses to reward predictive cues are modulated by the contingency and/or contiguity between cues, action, and outcomes. When cues or discriminative stimuli are presented in FR01 trials, there is often a close temporal association between cue and reward delivery, as animals tend to execute actions within several seconds to earn rewards. As only one action is required in this setting, there may be a more accurate predictive relationship between cue onset and reward delivery. Thus, the unpredictable nature of a random reinforcement schedule may prevent the cue from evoking a reward prediction error. This finding is important because while previous studies have suggested that cue- evoke reward prediction errors can be used to guide action selection in a wide variety of settings, these data are drawn largely from Pavlovian and FR01-based tasks.

Cue offset in the current task is similar to a traditional Pavlovian conditioned stimulus, as offset was completely predictive of subsequent reward delivery with high temporal contiguity (0.5 s between offset and delivery). Consistent with previous studies in which Pavlovian associations evoke VTA responses after associations are formed (Schultz, 1998), responses evoked by cue offset grew larger in successive

sessions of the task. Our experiment was not designed to dissociate the effects of learning from the effects of different reinforcement schedules on information processing, and future experiments can directly address this topic. Taken together, these data suggest VTA neurons encode real-time information about ongoing actions in conjunction with information about cues with strong contiguity and contingency with reward delivery.

## CONCLUSION

Our data reveal a novel form of information processing by VTA neurons. The unique ensemble coding scheme that we observed allows heterogeneous groups of VTA neurons to provide a real-time account of ongoing behavior. This coding scheme may subserve the well-established role of dopamine in goal-directed actions, decision making, and the behavioral disorganization and amotivation associated with illnesses such as ADHD and schizophrenia.

## AUTHOR CONTRIBUTIONS

JW designed experiments, performed experiments, analyzed data, and wrote the manuscript. NS designed experiments and edited the manuscript. FK analyzed the data. RK analyzed data and edited the manuscript. BM designed experiments and wrote the manuscript.

## ACKNOWLEDGMENTS

Funding was obtained from the NIMH MH048404 (BM), MH064537 (FK), MH064537 (RK), MH016804 (JW); NIDA DA031111 (JW), DA035050 (NS), DA022762 (FK); RK Mellon Foundation (FK); Andrew Mellon Fellowship (JW).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fnbeh.2017.00140/full#supplementary-material>

**Figure S1** | Representative examples of neuronal spike sorting, raw data and neuronal responses. **(A)** Spikes from unit A, a dopamine neuron, sorted according to the first 2 principal components of all threshold crossing waveforms (left). Purple points represent spikes that were assigned to unit A, and gray points represent noise that was not sorted into single unit spikes. A raw voltage trace (band pass filtered between 100 Hz and 7 KHz) corresponding to the same unit is depicted in the middle column. Examples of spikes belonging to unit A are notated in the trace. Unit A represents a typical cue-responsive unit. Raster plot depicts the unit's response aligned to cue onset (right). Each dash represents a single spike, and each row represents a single trial (first trial in the top row). Note the increased spike density just after cue onset (time 0) across all trials. **(B)** Representative cue-offset responsive non-dopaminergic neuron. Data plotted with the same conventions as **(A)**. In this example recording, two units were simultaneously recorded (left). Raster (right) depicts data during the time period between cue offset and reward delivery (−0.5 to 0 s), with spikes aligned to the time of reward delivery. Note the consistent increase in spike density after cue offset and preceding outcome delivery. **(C)** Data from a representative non-dopaminergic, outcome delivery responsive neuron is plotted with similar conventions as **(A)**. Raster (right) depicts neuronal activity aligned to the time of outcome delivery. Note the consistent delivery evoked response.

## REFERENCES

- Aberman, J. E., and Salamone, J. D. (1999). Nucleus accumbens dopamine depletions make rats more sensitive to high ratio requirements but do not impair primary food reinforcement. *Neuroscience* 92, 545–552. doi: 10.1016/S0306-4522(99)00004-4
- Anstrom, K. K., and Woodward, D. J. (2005). Restraint increases dopaminergic burst firing in awake rats. *Neuropsychopharmacology* 30, 1832–1840. doi: 10.1038/sj.npp.1300730
- Anstrom, K. K., Cromwell, H. C., and Woodward, D. J. (2007). Effects of restraint and haloperidol on sensory gating in the midbrain of awake rats. *Neuroscience* 146, 515–524. doi: 10.1016/j.neuroscience.2007.01.060
- Barker, D. J., Root, D. H., Zhang, S., and Morales, M. (2016). Multiplexed neurochemical signaling by neurons of the ventral tegmental area. *J. Chem. Neuroanat.* 73, 33–42. doi: 10.1016/j.jchemneu.2015.12.016
- Bayer, H. M., and Glimcher, P. W. (2005). Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* 47, 129–141. doi: 10.1016/j.neuron.2005.05.020
- Bermudez, M. A., and Schultz, W. (2014). Timing in reward and decision processes. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 369:20120468. doi: 10.1098/rstb.2012.0468
- Bromberg-Martin, E. S., Matsumoto, M., and Hikosaka, O. (2010). Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron* 68, 815–834. doi: 10.1016/j.neuron.2010.11.022
- Cohen, J. Y., Haesler, S., Vong, L., Lowell, B. B., and Uchida, N. (2012). Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* 482, 85–88. doi: 10.1038/nature10754
- Cohen, M. R., and Kohn, A. (2011). Measuring and interpreting neuronal correlations. *Nat. Neurosci.* 14, 811–819. doi: 10.1038/nn.2842
- Dayan, P., and Abbott, L. F. (2001). *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. Cambridge, MA: Massachusetts Institute of Technology Press.
- Dayan, P., and Niv, Y. (2008). Reinforcement learning: the good, the bad and the ugly. *Curr. Opin. Neurobiol.* 18, 185–196. doi: 10.1016/j.conb.2008.08.003
- Eshel, N., Tian, J., Bukwich, M., and Uchida, N. (2016). Dopamine neurons share common response function for reward prediction error. *Nat. Neurosci.* 19, 479–486. doi: 10.1038/nn.4239
- Fiorillo, C. D., Tobler, P. N., and Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science* 299, 1898–1902. doi: 10.1126/science.1077349
- Fujisawa, S., and Buzsaki, G. (2011). A 4 Hz oscillation adaptively synchronizes prefrontal, V. T. A., and hippocampal activities. *Neuron* 72, 153–165. doi: 10.1016/j.neuron.2011.08.018
- Geisler, S., and Zahm, D. S. (2005). Afferents of the ventral tegmental area in the rat-anatomical substratum for integrative functions. *J. Comp. Neurol.* 490, 270–294. doi: 10.1002/cne.20668
- Geisler, S., Derst, C., Veh, R. W., and Zahm, D. S. (2007). Glutamatergic afferents of the ventral tegmental area in the rat. *J. Neurosci.* 27, 5730–5743. doi: 10.1523/JNEUROSCI.0012-07.2007
- Glimcher, P. W. (2011). Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proc. Natl. Acad. Sci. U.S.A.* 108(Suppl. 3), 15647–15654. doi: 10.1073/pnas.1014269108
- Goldman-Rakic, P. (1998). The cortical dopamine system: role in memory and cognition. *Adv. Pharmacol.* 42, 707–711. doi: 10.1016/S1054-3589(08)60846-7
- Goldman-Rakic, P. S. (1988). Topography of cognition: parallel distributed networks in primate association cortex. *Ann. Rev. Neurosci.* 11, 137–156. doi: 10.1146/annurev.ne.11.030188.001033
- Hyland, B. I., Reynolds, J. N., Hay, J., Perk, C. G., and Miller, R. (2002). Firing modes of midbrain dopamine cells in the freely moving rat. *Neuroscience* 114, 475–492. doi: 10.1016/S0306-4522(02)00267-1
- Ishiwari, K., Weber, S. M., Mingote, S., Correa, M., and Salamone, J. D. (2004). Accumbens dopamine and the regulation of effort in food-seeking behavior: modulation of work output by different ratio or force requirements. *Behav. Brain Res.* 151, 83–91. doi: 10.1016/j.bbr.2003.08.007
- Jin, X., and Costa, R. M. (2010). Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature* 466, 457–462. doi: 10.1038/nature09263
- Joshua, M., Adler, A., Prut, Y., Vaadia, E., Wickens, J. R., and Bergman, H. (2009). Synchronization of midbrain dopaminergic neurons is enhanced by rewarding events. *Neuron* 62, 695–704. doi: 10.1016/j.neuron.2009.04.026
- Kass, R. E., Eden, U. T., and Brown, E. N. (2014). *Analysis of Neural Data*. New York, NY: Springer-Verlag.
- Kim, Y. B., Matthews, M., and Moghaddam, B. (2010). Putative gamma-aminobutyric acid neurons in the ventral tegmental area have a similar pattern of plasticity as dopamine neurons during appetitive and aversive learning. *Eur. J. Neurosci.* 32, 1564–1572. doi: 10.1111/j.1460-9568.2010.07371.x
- Kim, Y., Simon, N. W., Wood, J., and Moghaddam, B. (2016). Reward anticipation is encoded differently by adolescent ventral tegmental area neurons. *Biol. Psychiatry* 79, 878–886. doi: 10.1016/j.biopsych.2015.04.026
- Kim, Y., Wood, J., and Moghaddam, B. (2012). Coordinated activity of ventral tegmental neurons adapts to appetitive and aversive learning. *PLoS ONE* 7:e29766. doi: 10.1371/journal.pone.0029766
- Kiyatkin, E. A., and Rebec, G. V. (1998). Heterogeneity of ventral tegmental area neurons: single-unit recording and iontophoresis in awake, unrestrained rats. *Neuroscience* 85, 1285–1309. doi: 10.1016/S0306-4522(98)00054-2
- Lak, A., Stauffer, W. R., and Schultz, W. (2014). Dopamine prediction error responses integrate subjective value from different reward dimensions. *Proc. Natl. Acad. Sci. U.S.A.* 111, 2343–2348. doi: 10.1073/pnas.1321596111
- Lake, J. I., and Meck, W. H. (2013). Differential effects of amphetamine and haloperidol on temporal reproduction: dopaminergic regulation of attention and clock speed. *Neuropsychologia* 51, 284–292. doi: 10.1016/j.neuropsychologia.2012.09.014
- Lammel, S., Hetzel, A., Hackel, O., Jones, I., Liss, B., and Roeper, J. (2008). Unique properties of mesoprefrontal neurons within a dual mesocorticolimbic dopamine system. *Neuron* 57, 760–773. doi: 10.1016/j.neuron.2008.01.022
- Ljungberg, T., Apicella, P., and Schultz, W. (1992). Responses of monkey dopamine neurons during learning of behavioral reactions. *J. Neurophysiol.* 67, 145–163.
- Margolis, E. B., Lock, H., Hjelmstad, G. O., and Fields, H. L. (2006). The ventral tegmental area revisited: is there an electrophysiological marker for dopaminergic neurons? *J. Physiol.* 577(Pt 3), 907–924. doi: 10.1113/jphysiol.2006.117069
- Mingote, S., Weber, S. M., Ishiwari, K., Correa, M., and Salamone, J. D. (2005). Ratio and time requirements on operant schedules: effort-related effects of nucleus accumbens dopamine depletions. *Eur. J. Neurosci.* 21, 1749–1757. doi: 10.1111/j.1460-9568.2005.03972.x
- Montague, P. R., Dayan, P., and Sejnowski, T. J. (1996). A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* 16, 1936–1947.
- Morris, G., Nevet, A., Arkadir, D., Vaadia, E., and Bergman, H. (2006). Midbrain dopamine neurons encode decisions for future action. *Nat. Neurosci.* 9, 1057–1063. doi: 10.1038/nn1743
- Pan, W. X., Schmidt, R., Wickens, J. R., and Hyland, B. I. (2005). Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward-learning network. *J. Neurosci.* 25, 6235–6242. doi: 10.1523/JNEUROSCI.1478-05.2005
- Parker, K. L., Ruggiero, R. N., and Narayanan, N. S. (2015). Infusion of D1 dopamine receptor agonist into medial frontal cortex disrupts neural correlates of interval timing. *Front. Behav. Neurosci.* 9:294. doi: 10.3389/fnbeh.2015.00294
- Robbins, T. W., and Arnsten, A. F. (2009). The neuropsychopharmacology of fronto-executive function: monoaminergic modulation. *Annu. Rev. Neurosci.* 32, 267–287. doi: 10.1146/annurev.neuro.051508.135535
- Robbins, T. W., and Roberts, A. C. (2007). Differential regulation of fronto-executive function by the monoamines and acetylcholine. *Cereb. Cortex* 17(Suppl. 1), i151–i160. doi: 10.1093/cercor/bhm066
- Roesch, M. R., Calu, D. J., and Schoenbaum, G. (2007). Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat. Neurosci.* 10, 1615–1624. doi: 10.1038/nn2013
- Salamone, J. D., and Correa, M. (2002). Motivational views of reinforcement: implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behav. Brain Res.* 137, 3–25. doi: 10.1016/S0166-4328(02)00282-6
- Salamone, J. D., Correa, M., Farrar, A. M., Nunes, E. J., and Pardo, M. (2009). Dopamine, behavioral economics, and effort. *Front. Behav. Neurosci.* 3:13. doi: 10.3389/neuro.08.013.2009

- Salamone, J. D., Correa, M., Farrar, A., and Mingote, S. M. (2007). Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology (Berl)* 191, 461–482. doi: 10.1007/s00213-006-0668-9
- Schultz, W. (1998). Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80, 1–27.
- Schultz, W. (2010). Dopamine signals for reward value and risk: basic and recent data. *Behav. Brain Funct.* 6:24. doi: 10.1186/1744-9081-6-24
- Schultz, W., Apicella, P., and Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *J. Neurosci.* 13, 900–913.
- Schultz, W., Dayan, P., and Montague, P. R. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599. doi: 10.1126/science.275.5306.1593
- Seamans, J. K., and Yang, C. R. (2004). The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Prog. Neurobiol.* 74, 1–58. doi: 10.1016/j.pneurobio.2004.05.006
- Soares, S. Atallah, B. V., and Paton, J. J. (2016). Midbrain dopamine neurons control judgement of time. *Science* 354, 1273–1277. doi: 10.1126/science.aah5234
- Swanson, L. W. (2004). *Brain Maps III: Structure of the Rat Brain: An Atlas with Printed and Electronic Templates for Data, Models, and Schematics*. Amsterdam; Boston, MA: Elsevier, Academic Press.
- Tobler, P. N., Fiorillo, C. D., and Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science* 307, 1642–1645. doi: 10.1126/science.1105370
- Totah, N. K., Kim, Y., and Moghaddam, B. (2013). Distinct prestimulus and poststimulus activation of VTA neurons correlates with stimulus detection. *J. Neurophysiol.* 110, 75–85. doi: 10.1152/jn.00784.2012
- Tritsch, N. X., Ding, J. B., and Sabatini, B. L. (2012). Dopaminergic neurons inhibit striatal output through non-canonical release of GABA. *Nature* 490, 262–266. doi: 10.1038/nature11466
- Ungless, M. A., and Grace, A. A. (2012). Are you or aren't you? Challenges associated with physiologically identifying dopamine neurons. *Trends Neurosci.* 35, 422–430. doi: 10.1016/j.tins.2012.02.003
- Wassum, K. M., Ostlund, S. B., and Maimment, N. T. (2012). Phasic mesolimbic dopamine signaling precedes and predicts performance of a self-initiated action sequence task. *Biol. Psychiatry* 71, 846–854. doi: 10.1016/j.biopsych.2011.12.019

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Wood, Simon, Koerner, Kass and Moghaddam. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.