1 **Networks of VTA neurons encode real-time information about uncertain numbers of**

2 **actions executed to earn a reward**

3 **Running Title: VTA ensembles encode serial actions**

4 Jesse Wood[1], Nicholas W. Simon[2], Spencer Koerner[3,4], Robert E. Kass[3,4,5], Bita Moghaddam[6]

5

6

7 **Affiliations:**

8 [1]Department of Psychiatry, University of Pittsburgh, 450 Technology Drive Suite 223,
9 Pittsburgh, PA 15219.

10 [2]Department of Psychology, University of Memphis, Psychology Building 400 Innovation Drive,
11 Memphis, TN 38152

12 [3]Department of Statistics, Carnegie Mellon University, 132 Baker Hall, Pittsburgh, PA 15213.

13 [4]Center for the Neural Basis of Cognition, Carnegie Mellon University and the University of
14 Pittsburgh, 4400 5th Ave, Pittsburgh PA, 15213

15 [5]Machine Learning Department, Carnegie Mellon University, Gates Hillman Center 8203,
16 Pittsburgh, PA 15213

17 [6]Department of Behavioral Neuroscience, Oregon Health and Sciences University, Portland, OR
18 97239

19

20 *Correspondence to: Jesse Wood, email address: jtw22@pitt.edu

1 **ABSTRACT**

2 Multiple and unpredictable numbers of actions are often required to achieve a goal. In order to

3 organize behavior and allocate effort so that optimal behavioral policies can be selected, it is

4 necessary to continually monitor ongoing actions. Real-time processing of information related to

5 actions and outcomes is typically assigned to the prefrontal cortex and basal ganglia, but also

6 depends on midbrain regions, especially the ventral tegmental area (VTA). We were interested

7 in how individual VTA neurons, as well as networks within the VTA, encode salient events

8 when an unpredictable number of serial actions are required to obtain a reward. We recorded

9 from ensembles of putative dopamine and non-dopamine neurons in the VTA as animals

10 performed multiple cued trials in a recording session where, in each trial, serial actions were

11 randomly rewarded. While averaging population activity did not reveal a response pattern, we

12 observed that different neurons selectively tuned to low, medium, or high numbered actions in a

13 trial. This preferential tuning of putative dopamine and non-dopamine VTA neurons to different

14 subsets of actions in a trial allowed information about binned action number to be decoded from

15 the ensemble activity. At the network level, tuning curve similarity was positively associated

16 with action-evoked noise correlations, suggesting that action number selectivity reflects

17 functional connectivity within these networks. Analysis of phasic responses to cue and reward

18 revealed that the requirement to execute multiple and uncertain numbers of actions weakens both

19 cue-evoked responses and cue-reward response correlation. The functional connectivity and

20 ensemble coding scheme that we observe here may allow VTA neurons to cooperatively provide

21 a real-time account of ongoing behavior. These computations may be critical to cognitive and

22 motivational functions that have long been associated with VTA dopamine neurons.

23

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; Montague et al., 1996; Schultz et al., 1997; Schultz, 1998; Bayer and Glimcher, 2005; Morris et al., 2006; Roesch et al., 2007; Salamone et al., 2007; Dayan and Niv, 2008; Glimcher, 2011).

3

1   Thus, little is known about how VTA neurons encode information that could be used to organize

2   and sustain a series of actions.

3   To begin to understand how VTA neurons monitor a series of actions in real-time, we

4   recorded from VTA ensembles during a task where rats executed an unpredictable number of

5   self-paced actions to earn reward. We found that individual dopamine and non-dopamine

6   neurons fired selectively during different actions in a trial. Whereas population-averaging

7   concealed this mosaic of activity patterns, it was possible to decode low, medium, and high

8   numbered actions within a trial from the collective activity of all neurons. These data provide

9   novel evidence of ensemble encoding by VTA neurons and provide critical insight into how

10   VTA neurons may contribute to complex behaviors by encoding real-time information about

11   ongoing actions.

12   **MATERIALS AND METHODS**

13   **Subjects and Apparatus**

14   All procedures were conducted in accordance with the National Institutes of

15   Health's Guide to the Care and Use of Laboratory Animals, and approved by the University of

16   Pittsburgh's Institutional Animal Care and Use Committee. Behavioral and neuronal data were

17   collected from 10 adult, experimentally naïve, male Sprague-Dawley rats (Harlan, Frederick,

18   MD) singly housed on a 12-hour light cycle (lights on at 7pm). These sample sizes were chosen

19   because they are sufficient to detect behavioral and VTA neuronal effects in a variety of rat

20   operant procedures in our lab (Kim et al., 2010; Kim et al., 2012; Kim et al., 2016). At the time

21   of surgery, rats weighed approximately 350 grams (approximately 12 weeks of age). Under

22   isoflurane anesthesia, 8 or 16 channel, 50 µm stainless steel, Teflon insulated, microelectrode

23   arrays were implanted in left VTA (relative to Bregma: -5.30 mm posterior, 0.8 mm lateral, and

1   8.3 mm ventral). All experiments were conducted in a standard operant chamber (Coulbourn

2   Instruments, Allentown, PA). The operant chamber had a food trough on one wall and a single

3   nose poke port on the opposite wall. Both the food trough and nose poke port could be

4   illuminated and were equipped with an infrared beam, which detected the animal's entry. The

5   operant chamber system controller was configured to send the time of behavioral and

6   environmental events to the recording interface via standard TTL pulses and a digital interface.

7   **Behavior**

8   Each rat was given 7 days to recover from surgery and food restricted to 90% of their free

9   feeding body weights. Rats were habituated to handling for 5 minutes per day for 3 consecutive

10   days, before being habituated to being handled and connected to a headstage cable in the

11   procedure room for 2 additional days. Following habituation, rats were given a single 30 minute

12   magazine training session in the operant chamber, in which sugar pellets were delivered on a

13   variable time 75s reinforcement schedule. When each pellet was delivered, the pellet trough was

14   illuminated for 4s. The animal's behavior had no programmed consequences in the magazine

15   training session.

16   Following the magazine training session, each animal began instrumental conditioning.

17   During all instrumental conditioning sessions, each trial began with illumination of the nose poke

18   port (cue light onset). This served as a discriminative stimulus that reinforcing outcomes (sugar

19   pellets) were available (termed the 'response period'), contingent upon the animal executing

20   actions (nose pokes into the lit port). In each trial, actions were reinforced randomly, according

21   to a predetermined probability. When an action was executed, the behavioral system controller

22   randomly drew an outcome state (either reinforcement or no programmed consequence) with

23   replacement, according to the probability of reinforcement. Each action was reinforced randomly

5

1   and independently of the animal's action history within that trial or session. When an action was

2   reinforced, the cue light was immediately extinguished (cue light offset) and nose pokes had no

3   additional programmed consequences. A 0.500s delay between the final action and outcome

4   delivery was instituted to temporally separate these events, as done in previous work (Schultz et

5   al., 1993). Following this delay, the outcome was delivered to the animal and the food trough

6   was illuminated. Outcomes were delivered into the food trough from a standard pellet magazine,

7   via the operation of a smaller stepper motor and dispenser. The food trough remained illuminated

8   and the task did not progress until the animal retrieved the outcome. Once the animal retrieved

9   the outcome, a variable length intertrial interval (ITI) of 10-12s was initiated. In each session,

10   180 trials were administered.

11   In the first instrumental conditioning session, actions were reinforced with a probability

12   of 1 (each action was reinforced) equivalent to a fixed ratio 1 (FR01) reinforcement schedule. In

13   the second session, the probability that an action was reinforced was decreased across three

14   blocks of trials. In the first block of 60 trials, actions were reinforced with a probability of 1

15   (FR01). In the second block of 60 trials, each action had a 1 in 3 chance of being reinforced

16   (random ration 3, RR03). In the third block of 60 trials, the probability was further decreased to

17   0.2 (random ratio 5, RR05). In sessions 3 and 4, actions were reinforced with a 0.2 probability

18   for all trials (RR05). In sessions 5-7, actions were reinforced with a probability of 0.1 for all

19   trials (random ratio 10, RR10). In all trials but the FR01 trials, animals were required to execute

20   an unpredictable, varying, and randomly determined number of actions per trial. Random

21   reinforcement was utilized to limit the ability of the animal to correctly anticipate reward

22   delivery. Actions differed from each other mainly in terms of their location within the action

23   series in each trial (the action number within a trial, e.g. 1st action, 2nd action, 3rd action, etc.).

1 In each trial, each animal's action rate was calculated as the number of actions divided by the

2 duration of the response period. This served as a measure of behavioral conditioning and

3 performance. Changes in behavior across sessions were assessed with repeated measure analysis

4 of variance (ANOVA), and repeated measures contrasts were applied as appropriate (Kass et al.,

5 2014). The time interval between each action (inter-action interval) was measured for bins of

6 low, medium, and high numbered actions (actions 1-7, 8-14, and 15-21). Statistical differences

7 between binned inter-action intervals were assessed with repeated measures ANOVA. This

8 measured the animal's behavioral sensitivity to increasing action numbers throughout a trial. To

9 examine the effects of the number of actions performed in a trial on fatigue, motivation, or

10 attention, the number of actions performed in each trial was correlated with the latency to

11 retrieve the reward, or initiate the next trial.

12 **Histology**

13 Following the completion of experiments, animals were perfused with saline and brains

14 were extracted. Each brain was stored in a mixture of sucrose and formalin. The brains were then

15 frozen and sliced in 60 μm coronal sections on a cryostat, before being stained with cresyl-violet.

16 The location of each implant was histologically verified under light microscope according to

17 Swanson's brain atlas (Swanson, 2004). Animals were excluded if electrode location could not

18 be confirmed in VTA.

19 **Electrophysiology**

20 During experiments, animals were attached to a flexible headstage cable and motorized

21 commutator that allowed the animal to move freely about the operant chamber, with minimal

22 disruption of behavior (Plexon, Dallas, TX). Neural data were recorded via the PlexControl

23 software package, operating a 64-channel OmniPlex recording system (Plexon, Dallas, TX).

1   Neural data were buffered by a unity gain headstage and then a preamplifier. The digitized

2   broadband signal was then band-pass filtered (100 Hz – 7 KHz). High-pass filtering can affect

3   spike waveform shapes and neuronal identification, but with freely moving animals it is

4   necessary to apply these filters to remove artifacts from the neuronal signal (Ungless and Grace,

5   2012). The filter pass bands that were utilized in the current manuscript are consistent with those

6   that have previously been used to record from dopamine containing brain regions (Schultz et al.,

7   1993; Fiorillo et al., 2003; Tobler et al., 2005). Data were digitized at 40 KHz and continuously

8   recorded to hard disk. Voltage thresholds were applied to the digitized spike data offline (Offline

9   Sorter, Plexon, Dallas, TX). Single units were sorted using standard techniques, and were

10 utilized only if they had a signal to noise ratio in excess of 2/1, and were clearly separated from

11 noise clusters and other single unit clusters. Examples of recordings and sorting are show in

12 supplemental material (Fig. S1).

13   A VTA neuron was classified as dopaminergic if it had broad action potentials, greater

14 than 1.4 ms in duration, and a mean baseline firing rate less than 10 Hz. These criteria are similar

15 to those used in previous studies (Hyland et al., 2002; Fiorillo et al., 2003; Anstrom and

16 Woodward, 2005; Pan et al., 2005; Tobler et al., 2005; Anstrom et al., 2007; Totah et al., 2013).

17 All remaining neurons were classified as non-dopaminergic. All analyses were conducted on the

18 entire population of neurons that were recorded to provide a classification-free examination of

19 the data. Our approach to VTA cell classification is consistent with classical techniques. It

20 should be pointed out, however, that indirect identification (using electrophysiological features

21 as opposed genotypic features) can lead to classification errors (Margolis et al., 2006; Lammel et

22 al., 2008; Cohen et al., 2012). In general, qualitatively similar responses were observed in both

23 groups of neurons. Units recorded in different sessions were considered separate units, as

1   methods to estimate neuronal identity between sessions are not widely used with VTA

2   recordings. In total, the following numbers of neurons were recorded in each of 7 sessions:

3   dopamine (reward responsive): 11(7), 19(10), 21(12), 29(19), 27(13), 21(9), 27(16); non-

4   dopamine (reward responsive): 31(17), 35(14), 39(27), 34(13), 25(12), 29(11), 27(9).

5   **Neuronal Data Analysis**

6   Each single unit's spike times were binned into spike counts (0.025s bins) within a trial.

7   Binned spike counts were aligned to all events (e.g. cue light onset, actions, time period between

8   cue light offset and outcome delivery, and outcome delivery). A stable, four s portion of the ITI

9   (5s to 1s prior to cue light onset) served as the neuronal activity baseline. Single unit firing rates

10   were Z-score normalized relative to baseline and zero-centered before activity was averaged

11   together. Each unit's normalized activity was examined in 0.250s windows around events (cue

12   onset: +0.050 – 0.300s, relative to cue onset; -0.125 – +0.125s, relative to the time of action

13   execution; time period between cue offset and outcome delivery: +0.150 – 0.400s, relative to

14   execution of the last action; outcome delivery: +0.050 – 0.300s, relative to delivery). To assess

15   between-session changes in population-level evoked activity, windowed activity was compared

16   with a between groups two way ANOVA, with session number and neuron type (dopamine or

17   non-dopamine) as grouping variables. In all cases, protected Fisher's least significant difference

18   tests were applied as appropriate.

19   A unit was classified as being activated or suppressed by an event if it met two criteria: 1)

20   a significant paired samples t-test comparing raw (non-normalized) baseline firing rates with raw

21   evoked firing rates, and 2) three or more consecutive bins of mean activity within the event-

22   window, that were in excess of a 95% confidence interval around the baseline mean. With

23   respect to a given task event, the proportions of units classified as being activated or suppressed

1 were calculated. Differences in the proportions of activated units between sessions were

2 compared with a Chi-squared test of independence. As expected, insufficient numbers of

3 suppressed neurons (in some sessions zero) were obtained to permit reliable statistical analyses

4 of this class of responses. Suppressed neurons are plotted for clarity, but not analyzed in detail.

5 The terminology "action-evoked" neuronal responses refers to activity around the time of

6 action execution, without assuming that the action is solely responsible for evoking this neuronal

7 response. Each unit's activity was examined as a function of action number (a unit's mean

8 response to each nth numbered action within a trial, across all trials). These analyses were

9 restricted to the RR10 sessions (sessions 5-7). RR10 sessions required larger numbers of actions

10 per trial, on average, and would ensure that there were a sufficient number of higher numbered

11 actions (e.g. actions 18, 19, 20, 21, etc.) for analysis. All action number analyses utilized actions

12 1 through 21. While even higher numbered actions occurred in some trials, these actions

13 occurred less frequently and were excluded from action number analyses, as there was

14 insufficient sample size for reliable statistical analysis. To remove any effects of impending

15 reward delivery on the action evoked neuronal responses, only unrewarded actions were used in

16 this analysis. Preliminary analyses suggested that including rewarded actions had little effect on

17 the results. Action evoked neuronal responses were binned into 3 bins of 7 consecutive actions

18 representing low, medium, and high numbered actions (actions 1-7, 8-14, and 15-21,

19 respectively).

20 We measured correlations in spike count between pairs of dopamine neurons (n = 172),

21 pairs of non-dopaminergic neurons (n = 276), and mixed pairs (a dopamine and a non-dopamine

22 neuron, n = 387) in RR10 sessions. Correlations were only calculated for pairs of simultaneously

10

1   recorded neurons. Pairwise neuronal correlations were assessed in the same .250 sec windows

2   used to quantify firing rates, using the standard Pearson's correlation,

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

3   where the correlation coefficient, r, is the summed product of the ith point's residual from the

4   sample mean, $\bar{x}$ or $\bar{y}$, normalized by the sample standard deviation, s. Noise correlations

5   represent the degree to which the joint activity of simultaneously recorded pairs of neurons

6   fluctuates from trial to trial. Noise correlations were calculated on the trial-by-trial spike counts

7   observed in each window. For actions, these evoked responses also examined according to action

8   number within a trial, or binned action (actions 1-7, 8-14, or 15-21). Signal correlations refer to

9   the correlation between pairs of tuning curves in simultaneously recoded neurons. We corrected

10   for chance levels of correlation, which were estimated as the mean correlation of 1000 pairs of

11   randomly shuffled data. Fisher's Z transformation

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

12   was used to transform correlation coefficients for statistical comparisons. The geometric mean

13   spike count was calculated to measure the average joint activity of each pair of neurons.

14   Individual neurons preferred different subsets of action numbers. To determine the extent

15   to which elapsed time since trial start or action number could account for action evoked neuronal

16   activity, we performed a Poisson regression of spike count onto these variables. The percentages

17   of neurons with significant coefficients are reported. To visualize the various tunings of VTA

18   neurons to action number, each neuron's mean activity as a function of action number is

19   displayed. Because evoked firing rates could span a large range of values, each tuning curve was

11

1   scaled so that the maximum evoked activity was equal to 1 and the minimum evoked activity

2   was equal to 0. Scaled tuning curves were used only for visualization. To compute the number of

3   maxima, each tuning curve was fit with a cubic smoothing spline (smoothing parameter 0.001)

4   and maxima were detected as points in the spline with derivatives equal to 0.

5   **Population Average Decoder**

6   A decoder classified binned action number according to the trial-averaged spike count,

7   also averaged across neurons, on an action-by-action basis. This is called the 'population average

8   decoder.' Binned action number, A, is defined as 3 bins of 7 consecutive actions (1-7, 8-14, or

9   15-21). The classifier maximizes

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

10 Here $p(\bar{R}|A)$ represents the probability of the population averaged response occurring, given a

11 particular action bin. The probability of an action belonging to an action bin is denoted by $p(A)$,

12 and is uniform across action bins in the current work. That is, the prior probability for each bin,

13 $p(A)$, is $\frac{1}{3}$. For these analyses, $p(\bar{R}|A)$ is calculated from the sample distributions of mean action

14 evoked population responses as follows. First, the population average of all neurons' trial-

15 averaged responses is calculated for all actions in all bins. Next, one action is selected as a test

16 observation. Then, $p(\bar{R}|A)$ is assumed to be the best fit Gaussian of unknown mean and variance,

17 fit from all trials in the bin A (excluding the test observation, if it lies in A). The result is 3

18 smooth distributions of mean activity, corresponding to the 3 bins of action numbers. The

19 average activity of the test observations is calculated, and classified as the most likely action bin

20 to have evoked this response, based on the relationship between population averaged activity and

12

1 action bins in the training dataset (the actions not tested). The classifier selects the action bin

2 most likely to have evoked the population response via

$$\widehat{A^*} = \underset{A}{\operatorname{argmax}}\ p(A \mid R^*) = \underset{A}{\operatorname{argmax}}\ p(R^* \mid A) \times p(A).$$

3 Here, R^* is the population average of the held out action number. $\hat{A}^*$ is the resulting estimated

4 action bin. Testing is repeated for each action number.

5 **Ensemble Decoder**

6   Assuming $\bar{R}$ is not a sufficient characterization of $\{R_i\}_{i=1}^N$, which we will call $\vec{R}$, $p(\bar{R}|A)$

7 does not thoroughly represent the information in $p(\vec{R}|A)$. Therefore, we instead assume relevant

8 information is captured by a higher-dimensional representation: $(W\vec{R}|A) \sim N(\mu_A, \Sigma)$. The

9 projection, W, estimated by principal components analysis, is necessary to ensure the

10 covariance, Σ , can be estimated by the usual pooled sample covariance, $\hat{\Sigma}$. This maximizes the

11 dimensionality of the space into which we project $\vec{R}$, for a given sample size. We then maximize

12 $p(A|W\vec{R})$. This maintains the varying firing rates in each unit, elicited by actions within

13 different bins. Note that the population average decoder is the special case of this approach

14 wherein W is a row vector of ones. The projection and classifier can be learned with each

15 neuron's set of evoked trial-averaged spike counts as a feature. Each action's set of observed

16 trial-averaged spike counts, the new test observations, $\overrightarrow{R *}$, are then projected via the weights

17 estimated from the training data and classified sequentially (that is, we estimate the set of

18 $\{\hat{\mu}_A\}_{A=1}^3$ on all the remaining action numbers, which is equivalent to training, and maximize the

19 analogous posterior).

20 **Statistical Testing of the Decoders**

13

1   To test whether the two decoders produce equal or different classification accuracy, we

2   fit a binomial generalized linear model and controlled for action number, session, and unit type

3   (dopaminergic versus non-dopaminergic). For each trial of action bin A, tested on decoder d, the

4   log odds of the probability of correct classification was assumed to be linear in its covariates as

5   given by the logistic regression function

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

6   where $I_A(\hat{A})$ is an indicator signifying correct or incorrect classification, and I_d, I_A, I_S, I_{DA} are

7   indicators (or a set of indicators) for the decoder, action bin, session, and dopaminergic unit,

8   respectively. Likelihood ratio tests were used to test the effect of each group of factors,

9   corresponding to each feature. Additionally, statistical significance of decoder performance

10   (versus chance levels of correct classification) at each action number was determined via

11   permutation test. Approximations to the exact p-values can be determined by assigning classes to

12   each action according to a uniform prior distribution, and calculating the proportion of sets of

13   resulting classifications that perform better than the classification rates observed using the

14   previously discussed methods.

15   **RESULTS**

16   **Animals learn to execute serial actions for reward**

17   VTA activity was recorded while rats learned to execute multiple actions for random

18   reinforcement with one sugar pellet (Fig. 1A, B). In the first session, each action was rewarded

19   (FR01). In session 2, the reward probability was decreased from 1 to 0.2 across 3 blocks of trials.

20   In sessions 3 and 4, actions were reinforced at a probability of 0.2 (RR05). In sessions $5 - 7$,

21   actions were reinforced at a probability of 0.1 (RR10). In all randomly reinforced trials,

14

1   unpredictable and varying numbers of actions were required, but each action was equally likely

2   to be reinforced.

3   Action response rates increased as reinforcement probability decreased (Fig. 1C, D), and

4   the response rate in the final RR10 sessions was significantly higher than all other sessions (Fig.

5   1D; $F_{(6,24)}$ = 4.726, p = 0.003). To investigate a potential relationship between behavioral

6   performance and action number, we divided the RR10 data into three bins: actions 1-7 (low),

7   actions 8-14 (medium), and actions 15-21 (high). We analyzed serial actions in trials reinforced

8   with a probability of 0.1 (RR10 sessions, sessions 5-7), because this reinforcement schedule

9   produced the greatest average number of actions performed per trial (Fig. 1C). Within a trial, the

10   inter-action interval significantly increased with increasing action number bins (Fig. 1E; $F_{(2,44)}$ =

11   11.020, p < 0.001). This indicated that behavior was sensitive to action number. The latency to

12   retrieve the reward ($r_{(3468)}$ = -0.017, p = 0.309) or initiate responding in the next trial, however,

13   was not correlated with binned action number ($r_{(3445)}$ = 0.015, p = 0.389). These data suggest that

14   the number of actions performed in a trial did not affect behavioral responses to reward or cues.

15   **Neurophysiological recordings**

16   A total of 375 units (in 7 sessions and from 10 rats) were recorded from electrodes

17   histologically localized to VTA (Fig. 2). All neurons were included in these analyses because we

18   were interested in understanding the full diversity of activity patterns in the VTA. Neurons were

19   classified (Fig. 3) as putative 'dopamine' (n = 155) or 'non-dopamine' neurons (n=220) based on

20   electrophysiological criteria. This classification approach permits comparison with previous

21   work, despite potential inaccuracies (Margolis et al., 2006) and with the caveat that some

22   dopamine neurons co-release other neurotransmitters (Tritsch et al., 2012). We did not observe

23   strong clustering in the electrophysiological profiles of these neurons. It is important to

15

1   underscore that neurons were collected with no attempt to select for particular

2   electrophysiological characteristics. When VTA neurons are recorded in an unbiased manner, as

3   done here, weak clustering may be expected (Kiyatkin and Rebec, 1998). In addition to the

4   above classification, analyses were performed on reward-responsive dopamine neurons, as this

5   subgroup may be a more conservative estimate of dopamine neuron identity (Lak et al., 2014;

6   Eshel et al., 2016).

7   **VTA neurons are diversely tuned to serial actions**

8   In order to understand how VTA neurons encode information during serial actions, we

9   examined how activity was modulated by actions in RR10 sessions. Action evoked responses

10   initiated -0.173±0.013 sec prior to actions and had an average duration of 0.313±0.012 sec (Fig.

11   4A). Thus, the majority of neurons had response durations several hundred ms faster than the

12   majority of inter-action intervals, and therefore, action-evoked responding was well separated

13   between successive actions in a trial. When RR10 actions were examined according to their

14   number within a trial (e.g. 1st, 2nd, 3rd action in a trial), we observed that individual neurons fired

15   selectively for subsets of actions (Fig. 4B, C). We observed preferential responding to low,

16   medium, or high numbered actions within the population of simultaneously recorded neurons

17   (Fig. 4D-F). To determine how VTA neurons responded to serial actions, we calculated each

18   neuron's tuning curve as a function of action number within a trial (Fig. 4G). Action tuning

19   curves mostly had a single peak (local maxima, $\chi^2_{(1)} = 9.921$, p = 0.002; neuron type, $\chi^2_{(1)} =$

20   1.267, p = 0.260), and different neurons preferred different subsets of actions (Fig. 4G,H).

21   Tuning curve depth (maximum – minimum) did not differ significantly between putative

22   dopamine and non-dopamine neurons (Fig. 4I; $t_{(154)} = 3.224$, p = 0.084). Averaging activity

16

1   across neurons, a traditional approach to analyzing VTA neuronal data, concealed this

2   heterogeneity in both groups of neurons (Fig. 4J).

3   The diversity of the tuning curves of VTA neurons suggested that network properties

4   may be critical to information processing during serial actions. To gain insight into VTA

5   network-based information processing, we calculated the trial-by-trial correlations in spike

6   counts (noise correlations), and correlations between tuning curves (signal correlations) for all

7   simultaneously recorded pairs of neurons. Noise correlations reflect functional connectivity

8   between neurons while signal correlations reflect similarity in tuning curves (Cohen and Kohn,

9   2011). Correlations were examined between all possible pairings of VTA neurons (dopamine –

10   dopamine, non-dopamine – non-dopamine, and dopamine – non-dopamine).

11   We found that noise correlations during actions were significantly lower in pairs of non-

12   dopamine neurons than pairs containing a dopamine neuron (Fig. 5A; $F_{(2,858)} = 4.052$, p = .018).

13   The magnitude of these correlations was decreased during low numbered actions compared to

14   medium or high numbered actions ($F_{(2,1716)} = 14.625$, p < .001). It is unlikely that these

15   differences were due to differences in spike count because the pairwise geometric mean spike

16   count was not associated with noise correlation magnitude in any bin (low: r = 0.008, p = .805;

17   med: r = -0.015, p = .650; high: r = -0.048, p = .158). In contrast, signal correlation strength did

18   not differ between different pairings of VTA neurons (Fig. 5B; $F_{(2,858)} = 0.289$, p = .749). There

19   was a strong association between signal correlations and noise correlations (Fig. 5B, Table 1),

20   suggesting that there is a high degree of functional connectivity between pools of similarly tuned

21   neurons. This result raises the possibility that different patterns of functional connectivity

22   contribute to the diversity of tunings to actions.

23

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

Table 1. Table depicts the correlation between action number signal correlations and action evoked noise correlations. 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).

1

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 (Fig. 6A) or the collective activity of VTA ensembles (Fig. 6B). The same cross-validated linear discriminant analysis decoding algorithms were used for both the ensemble and population-average data (see Methods). Ensemble activity was discriminated more accurately than shuffled control data (Fig. 6C; 85% accuracy versus chance levels of correct decoding [33% accuracy], permutation test, $p < 0.001$). In contrast, population averaged activity was decoded at chance levels (Fig. 6C; permutation test, $p = 0.240$). Binned action number was discriminated from ensemble activity significantly more accurately than population-averaged activity (Fig. 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

1    $(\chi^2_{(1)} = 1.300$, p = 0.254). These data suggest that real-time information about ongoing behavior

2   is encoded by the collective activity of VTA neurons, but not the averaged activity of VTA

3   dopamine or non-dopamine neurons. Thus, VTA ensemble activity accurately predicted the

4   current action bin, and could be used to organize, pace, and sustain effortful behaviors.

5   To determine if serial actions or the passage of time preferentially modulated neuronal

6   activity, we examined the relationship between action evoked spike counts and action number, or

7   the corresponding time within a trial. Action number predicted spike counts of a significantly

8   larger proportion of neurons than time elapsed in a trial ($\chi^2_{(1)} = 20.643$, p < 0.001; 29% versus

9   7% of neurons, respectively). Actions did not preferentially modulate the activity of either

10   dopamine or non-dopamine neurons ($\chi^2_{(1)} = 0.179$, p = 0.673). These data suggest action-evoked

11   activity reflected ongoing behavior and do not support a role for encoding the passage of time in

12   this task.

13   Neuronal responses may have represented reward anticipation (Bromberg-Martin et al.,

14   2010; Totah et al., 2013) in a manner which changed throughout a series of actions. If this were

15   the case, the number of actions performed in a trial would be correlated with neuronal activity in

16   a consistent direction. Thus, we examined the correlation between neuronal responses and the

17   number of actions in each trial. Only a small proportion of dopamine and non-dopamine neurons

18   had significant correlations between the response evoked by the first action in a trial and the

19   number of actions in the previous trial (Table 2). Similar proportions of these neurons were

20   positively or negatively correlated, indicating that this relationship was inconsistent between

21   neurons (Table 2). Likewise, very few neurons had responses to the last action in each trial that

22   were significantly correlated with the number of actions in that trial; there was no significant

1 difference in the proportion of dopamine and non-dopamine neurons with positive or negative

2 correlations (Table 2).

3   We further investigated if reward anticipation modulated neuronal responses by

4 examining the relationship between responses evoked by the cue light turning off, reward

5 delivery, or reward retrieval, and the number of actions performed in each trial. We found no

6 significant differences in the proportion of dopamine or non-dopamine neurons with positive or

7 negative correlations between responses evoked by cue onset and the number of actions in the

8 previous trial or between cue light offset, reward delivery, or reward retrieval responses and the

9 number of actions in the current trial (Table 2). Taken together, these results indicate that

10 performance of serial actions did not influence encoding of reward anticipation, supporting the

11 notion that VTA neurons encoded information about ongoing behaviors as opposed to

12 information about time or reward anticipation in the current task.

	DA Positive Corr	DA Negative Corr	Non-DA Positive Corr	Non-DA Negative Corr	Fisher's Exact Test
Response to First Action & Previous Trial's Action Number	4 %	1.33 %	3.70 %	1.23 %	P = 0.71
Response to Last Action & Current Trial's Action Number	2.67 %	1.33 %	4.94 %	4.94 %	P = 0.77
Response to Cue On & Previous Trial's Action Number	0.00 %	2.67 %	2.47 %	1.23 %	P = 0.25
Response to Cue Off & Current Trial's Action Number	6.67 %	2.67 %	11.11 %	4.94 %	P = 0.81
Response to Delivery & Current Trial's Action Number	5.33 %	5.33 %	3.70 %	2.47 %	P = 0.80
Response to Retrieval & Current Trial's Action Number	1.33 %	1.33 %	1.23 %	1.23 %	P = 0.67

Table 2. Table depicts the percentage of neurons with significantly correlated responses between task events and the number of actions in each trial. Data are taken from RR10 sessions. Pearson's correlations were calculated on spike counts taken from 0.25s 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 (Fig. 7A, B; $F_{(1,361)} = 9.159$, p = .003). This pattern did not change across sessions (Fig. 7C; $F_{(6,361)} = 1.352$, p = .233). The proportion of dopamine neurons responding to reward delivery did not vary significantly across sessions (Fig. 7D; all neurons, $X^2_{(6)} = 9.289$, p = .158; non dopamine, $X^2_{(6)} = 13.234$, p = .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 (Fig. 7E; $F_{(6,361)} = 4.776$, p < .001), and activated a greater proportion of neurons (Fig. 7F; all neurons, $X^2_{(6)} = 45.949$, p < .001; dopamine neurons, $X^2_{(6)} = 22.093$, p = .001; non dopamine neurons, $X^2_{(6)} = 29.744$, p < .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; Cohen et al., 2012; Kim 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 (Fig. 8A), which is consistent with previous observations (Schultz, 1998; Roesch et al., 2007). Non-dopamine neurons followed a similar pattern (Fig. 8B). During random ratio sessions, the population response to the cue decreased significantly as response requirement increased (Fig. 8C; session, $F_{(6,361)} = 2.667$, p = .015). This pattern of responding did not differ between dopamine and non-dopamine neurons (Fig. 8C;

1 $F_{(1,361)} = 1.543$, p = .215). We found the same effects when analysis was restricted to neurons that

2 were significantly activated by reward delivery (session, $F_{(6,175)} = 4.961$, p < .001; neuron type,

3 $F_{(1,175)} = 3.912$, p = 0.050). The data among neurons not responsive to reward delivery were

4 inconclusive, but followed the same trend (session, $F_{(6,172)} = 2.096$, p = 0.056; neuron type,

5 $F_{(1,172)} = 0.824$, p = 0.365). Likewise, significantly fewer neurons were modulated by cue light

6 onset in random ratio sessions (Fig. 8D; all VTA, $X^2_{(6)} = 23.844$, p = .001; dopamine, $X^2_{(6)} =$

7 20.109, p = .003; non-dopamine, $X^2_{(6)} = 10.636$, p = .100).

8 **Cue and reward evoked responses are decoupled when performing serial actions**

9 We assessed the correlation between each neuron's responses to cue onset and reward

10 delivery in order to understand the relationship between encoding of these events. During FR01

11 sessions, responses evoked by these events were significantly correlated in both dopamine and

12 non-dopamine neurons (Table 3). This correlation decreased in RR05 sessions (Table 3), and

13 was no longer significant by RR10 sessions (Table 3). A similar pattern was observed in

14 dopamine neurons that were activated by reward delivery (Table 3). This suggests that, although

15 cue and reward responses are positively correlated in Pavlovian conditioning tasks (Eshel et al.,

16 2016), cue-evoked responses become decoupled from reward responses during an instrumental

17 random reinforcement schedule. Figure 9 depicts two example neurons in the FR01 session (Fig.

18 9A) and the final RR10 session (Fig. 9B). Note that this example neuron has similar responses to

19 cue onset and reward delivery in the FR01 session (Fig. 9A), while the example neuron (same

20 animal) recorded during the RR10 session does not respond to cue onset, but is activated by

21 reward delivery (Fig. 9B).

22

23

	FR01	RR05	RR10
Dopamine	r = 0.653 p = .029	r = 0.321 p = .023	r = 0.096 p = .411
Reward responsive dopamine	r = 0.496 p = .014	r = 0.280 p = 0.018	r = 0.231 p = 0.054
Non-dopamine	r = 0.517 p = .003	r = 0.584 p < .001	r = 0.002 p = .981

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

1 **Modulation of noise correlations by cue onset and reward delivery in serial action trials**

2   We calculated the signal and noise correlation between pairs of VTA neurons during cues

3 and reward delivery in RR10 sessions. We found a significant interaction between pair type

4 (dopamine pairs, non-dopamine pairs, and mixed pairs) and event (Fig. 10; $F_{(4,2496)} = 9.482$, p <

5 .001). There was no difference in baseline noise correlations between any pairing of VTA

6 neurons (Fig. 10; $F_{(2,832)} = 0.855$, p = .426). Noise correlations evoked by cue onset were

7 significantly greater in pairs containing a dopamine neuron than pairs of non-dopamine neurons

8 (Fig. 10; $F_{(2,832)} = 5.881$, p = .003). There was no association between cue evoked noise

9 correlations and pairwise cue evoked geometric mean spike counts (r = 0.008, p = .817),

10 suggesting that mean activity level did not account for this difference. Noise correlations were

11 strongest during reward delivery overall. Noise correlations during reward delivery did not differ

12 significantly between pairings of neurons (Fig. 10; $F_{(2,832)} = 2.794$, p = .062). These data confirm

13 previous findings that correlated activity in VTA circuits emerges when rewarding outcomes can

14 be earned (Joshua et al., 2009; Kim et al., 2012; Eshel et al., 2016).

15

16

1 **DISCUSSION**

2   We investigated how VTA neurons encode information when animals execute an

3 uncertain numbers of actions to earn a reward. Dopamine and non-dopamine VTA neurons were

4 preferentially tuned to different subsets of binned action number in a manner that allowed serial

5 actions to be accurately discriminated from the collective activity of VTA ensembles. In

6 addition, pairs of neurons with similar action tuning curves had higher action-evoked noise

7 correlations suggesting that VTA network properties may be critical to information processing

8 during serial actions. Our analyses further showed that this pattern of tuning is distinct from

9 reward prediction related activity or the passage of time. We also demonstrated that cues

10 preceding a series of actions, unlike a single rewarded action, do not evoke neuronal responses

11 and are decoupled from reward, suggesting that cue-related information is utilized differently

12 when a single versus an unpredictable series of actions are required to obtain reward.

13 **Encoding information during an action series**

14   Previous reports indicate that dopamine neurons encode errors in the predicted value of

15 rewards as well as action values (Montague et al., 1996; Schultz, 1998; Morris et al., 2006;

16 Roesch et al., 2007). Further, dopamine neurons phasically fire during the first and last action in

17 a sequence, which may be critical to initiating and terminating behavior (Jin and Costa, 2010).

18 Our data indicate that the collective activity of ensembles of VTA neurons also has the capacity

19 to encode information about ongoing behavior in real-time.

20   The ensemble code revealed here may be used to organize and sustain actions, which is a

21 fundamental feature of dopamine's role in cognition (Aberman and Salamone, 1999; Ishiwari et

22 al., 2004; Salamone et al., 2009). In particular, this signal could be critical for supporting the

23 unique motivational and cognitive demands when large numbers of actions are required

24

1   (Goldman-Rakic, 1998; Aberman and Salamone, 1999; Salamone and Correa, 2002; Seamans

2   and Yang, 2004; Robbins and Roberts, 2007; Salamone et al., 2007; Robbins and Arnsten, 2009).

3   VTA dopamine neurons project to several networks that may utilize this signal to organize

4   behavior. For instance, dopamine innervation of the ventral striatum is selectively required for

5   completing a lengthy series of instrumental actions, but not for executing small numbers of

6   actions for reward (Aberman and Salamone, 1999; Ishiwari et al., 2004; Mingote et al., 2005).

7   Dopamine projections to the prefrontal cortex are important for working memory and related

8   constructs that require real-time information about ongoing behavior (Goldman-Rakic, 1998;

9   Seamans and Yang, 2004; Robbins and Roberts, 2007; Robbins and Arnsten, 2009). VTA

10   neurons could contribute to these functions by continually encoding information throughout

11   execution of a lengthy series of actions, allowing striatal and prefrontal networks to track goal-

12   directed effort expenditure and progress.

13   Our analyses suggest that VTA responses do not represent the passage of time in the

14   current task. Although dopamine responses are sensitive to the timing of reward delivery and

15   necessary for perception of time and interval timing (Lake and Meck, 2013; Bermudez and

16   Schultz, 2014; Parker et al., 2015; Soares et al., 2016), neuronal responses were not consistently

17   modulated by time elapsed in a trial. Instead, activity during a series of actions was most strongly

18   modulated by action execution. This is likely because reward delivery was not contingent upon

19   elapsed time, and VTA neurons are sensitive to the contingency between action execution and

20   outcome delivery. Experimental designs with an explicit contingency between elapsed time and

21   reward delivery, such as a fixed interval reward schedule, may reveal a more complex

22   relationship between action execution and within trial timing in the responses of VTA neurons.

1   We also found that the number of actions required in RR10 trials did not modulate VTA

2   responses to the rewarded action or the first action in the next trial. This suggests that recent trial

3   history does not affect action evoked responses at the beginning or end of a series of actions.

4   Similarly, VTA responses to stimuli associated with reward were not consistently correlated with

5   the number of actions that an animal performed, suggesting that VTA correlates of reward

6   anticipation did not increase or decrease according to action number. This result was expected, as

7   each action was reinforced with equal probability.

8   **VTA ensemble encoding**

9   Neurons with similar action number tuning curves had higher action-evoked noise

10   correlations in spike count, which reflect shared connectivity between neurons (Cohen and Kohn,

11   2011). This suggests that different action number tunings may arise from unique inputs or

12   differing connection strengths between inputs and VTA neurons with similar tunings may share

13   connection properties. VTA receives inputs from an expansive set of afferent structures (Geisler

14   and Zahm, 2005; Geisler et al., 2007), and the diversity of inputs that converge in VTA may

15   contribute to preferential firing in different subsets of serial actions.

16   The association between noise correlations and action selectivity indicates that VTA

17   network properties are critical to understanding how information about serial behavior is

18   encoded. Networks can encode information redundantly or through diverse activity patterns and

19   heterogeneous tunings naturally lead to the capacity to encode information as ensembles.

20   Accordingly, serial actions were accurately discriminated from ensemble activity, but not from

21   population-averaged activity. Though previous studies suggest VTA activity is highly redundant

22   (Joshua et al., 2009; Schultz, 2010; Glimcher, 2011; Kim et al., 2012; Eshel et al., 2016), this

23   work was limited to tasks requiring very few actions. The diversity of VTA tuning curves may

1   increase to match the expanded behavioral state space of serial actions compared with the limited

2   state space of single action trials (Eshel et al., 2016). Thus, ensemble encoding of information

3   may occur when the complexity of a task increases, such as during ongoing serial actions.

4   Units classified as either putative dopamine or non-dopamine neurons had comparable

5   mixed tunings to serial actions, suggesting that dopamine neurons can cooperate with non-

6   dopamine VTA neurons to encode information about serial actions. These ensemble signals

7   consisting of multiple neurotransmitters could allow information to be decoded through multiple

8   signaling mechanisms, which may diversify the spatiotemporal properties of this signal

9   (Seamans and Yang, 2004; Kim et al., 2010; Barker et al., 2016). The activity of different types

10 of VTA neurons must ultimately be coordinated and unified to represent information, and the

11 present work demonstrates how different types of VTA neurons could collectively encode

12 information about an ongoing series of actions.

13 **Cue and Reward Evoked Responses**

14   In the first session, animals executed a single action to earn rewards (FR01). Similar to

15 previous studies (Ljungberg et al., 1992; Schultz et al., 1993; Morris et al., 2006; Roesch et al.,

16 2007), cue light onset at the start of the trial evoked homogenous, phasic population activation in

17 VTA neurons. There was also a significant correlation between cue and reward delivery evoked

18 responses. These observations are consistent with a convincing body of literature suggesting that

19 dopamine neurons encode action value predicted by cues in the environment (Montague et al.,

20 1996; Dayan and Abbott, 2001). On the other hand, when animals were required to execute

21 multiple and random numbers of actions (random reinforcement schedules) to earn rewards we

22 observed weakening of two critical neuronal responses by VTA neurons: (1)-loss of cue-evoked

23 responses and (2) decorrelation of cue and reward evoked responses. This suggests that VTA

27

1   neuronal responses to reward predictive cues are modulated by the contingency and/or contiguity

2   between cues, action, and outcomes. When cues or discriminative stimuli are presented in FR01

3   trials, there is often a close temporal association between cue and reward delivery, as animals

4   tend to execute actions within several seconds to earn rewards. As only one action is required in

5   this setting, there may be a more accurate predictive relationship between cue onset and reward

6   delivery. Thus, the unpredictable nature of a random reinforcement schedule may prevent the cue

7   from evoking a reward prediction error. This finding is important because while previous studies

8   have suggested that cue-evoke reward prediction errors can be used to guide action selection in a

9   wide variety of settings, these data are drawn largely from Pavlovian and FR01-based tasks.

10   Cue offset in the current task is similar to a traditional Pavlovian conditioned stimulus,

11   as offset was completely predictive of subsequent reward delivery with high temporal contiguity

12   (0.5 sec between offset and delivery). Consistent with previous studies in which Pavlovian

13   associations evoke VTA responses after associations are formed (Schultz, 1998), responses

14   evoked by cue offset grew larger in successive sessions of the task. Our experiment was not

15   designed to dissociate the effects of learning from the effects of different reinforcement

16   schedules on information processing, and future experiments can directly address this topic.

17   Taken together, these data suggest VTA neurons encode real-time information about ongoing

18   actions in conjunction with information about cues with strong contiguity and contingency with

19   reward delivery.

20   **Conclusions**

21   Our date reveal a novel form of information processing by VTA neurons. The unique

22   ensemble coding scheme that we observed allows heterogeneous groups of VTA neurons to

23   provide a real-time account of ongoing behavior. This coding scheme may subserve the well-

1   established role of dopamine in goal-directed actions, decision making, and the behavioral

2   disorganization and amotivation associated with illnesses such as ADHD and schizophrenia.

3   **ACKNOWLEDGEMENTS**

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(2)**,** 545-552.

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(2)**,** 515-524. doi: S0306-4522(07)00091-7 [pii] 10.1016/j.neuroscience.2007.01.060.

Anstrom, K.K., and Woodward, D.J. (2005). Restraint increases dopaminergic burst firing in awake rats. *Neuropsychopharmacology* 30(10)**,** 1832-1840.

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(1)**,** 129-141. doi: S0896-6273(05)00467-8 [pii] 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(1637)**,** 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(5)**,** 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(7383)**,** 85-88. doi: 10.1038/nature10754.

Cohen, M.R., and Kohn, A. (2011). Measuring and interpreting neuronal correlations. *Nat Neurosci* 14(7)**,** 811-819. doi: 10.1038/nn.2842nn.2842 [pii].

Dayan, P., and Abbott, L.F. (2001). *Theoretical neuroscience : computational and mathematical modeling of neural systems.* Cambridge, Mass.: Massachusetts Institute of Technology Press.

Dayan, P., and Niv, Y. (2008). Reinforcement learning: the good, the bad and the ugly. *Curr Opin Neurobiol* 18(2)**,** 185-196. doi: S0959-4388(08)00076-7 [pii] 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(3)**,** 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(5614)**,** 1898-1902.

Fujisawa, S., and Buzsaki, G. (2011). A 4 Hz Oscillation Adaptively Synchronizes Prefrontal, VTA, and Hippocampal Activities. *Neuron* 72(1)**,** 153-165. doi: S0896-6273(11)00739-2 [pii]10.1016/j.neuron.2011.08.018.

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(21)**,** 5730-5743.

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(3)**,** 270-294.

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.10142691081014269108 [pii].

Goldman-Rakic, P. (1998). The cortical dopamine system: Role in memory and cognition. *Advances in Pharmacology* 42**,** 707-711.

Goldman-Rakic, P.S. (1988). Topography of cognition: parallel distributed networks in primate association cortex. *Annual Reviews in Neurosciences* 11**,** 137-156.

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(2)**,** 475-492. doi: S0306452202002671 [pii].

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(1-2)**,** 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(7305)**,** 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(5)**,** 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.*

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(11)**,** 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(1)**,** e29766. doi: 10.1371/journal.pone.0029766.

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. *The European journal of neuroscience* 32(9)**,** 1564-1572. doi: 10.1111/j.1460-9568.2010.07371.x.

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(4)**,** 1285-1309.

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(6)**,** 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(2)**,** 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(5)**,** 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(1)**,** 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.

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(6)**,** 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(5)**,** 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(8)**,** 1057-1063.

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(26)**,** 6235-6242.

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. *Cerebral cortex* 17 Suppl 1**,** i151-160. 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(12)**,** 1615-1624.

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(1-2)**,** 3-25.

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(3)**,** 461-482. doi: 10.1007/s00213-006-0668-9.

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.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology* 80(1)**,** 1-27.

Schultz, W. (2010). Dopamine signals for reward value and risk: basic and recent data. *Behav Brain Funct* 6**,** 24. doi: 1744-9081-6-24 [pii] 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. *Journal of Neuroscience* 13(3)**,** 900-913.

Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. *Science* 275(5306)**,** 1593-1599.

Seamans, J.K., and Yang, C.R. (2004). The principal features and mechanisms of dopamine modulation in the prefrontal cortex. *Prog Neurobiol* 74(1)**,** 1-58.

1 Soares, Atallah, and Paton (2016). Midbrain dopamine neurons control judgement of time.
2   *Science* 354(6317)**,** 1273-1277.
3 Swanson, L.W. (2004). *Brain maps III : structure of the rat brain : an atlas with printed and*
4   *electronic templates for data, models, and schematics.* Amsterdam ; Boston: Elsevier,
5   Academic Press.
6 Tobler, P.N., Fiorillo, C.D., and Schultz, W. (2005). Adaptive coding of reward value by
7   dopamine neurons. *Science* 307(5715)**,** 1642-1645. doi: 307/5715/1642 [pii]
8   10.1126/science.1105370.
9 Totah, N.K., Kim, Y., and Moghaddam, B. (2013). Distinct prestimulus and poststimulus
10   activation of VTA neurons correlates with stimulus detection. *J Neurophysiol* 110(1)**,** 75-
11   85. doi: 10.1152/jn.00784.2012.
12 Tritsch, N.X., Ding, J.B., and Sabatini, B.L. (2012). Dopaminergic neurons inhibit striatal output
13   through non-canonical release of GABA. *Nature* 490(7419)**,** 262-266. doi:
14   10.1038/nature11466.
15 Ungless, M.A., and Grace, A.A. (2012). Are you or aren't you? Challenges associated with
16   physiologically identifying dopamine neurons. *Trends Neurosci* 35(7)**,** 422-430. doi:
17   10.1016/j.tins.2012.02.003.
18 Wassum, K.M., Ostlund, S.B., and Maidment, N.T. (2012). Phasic mesolimbic dopamine
19   signaling precedes and predicts performance of a self-initiated action sequence task. *Biol*
20   *Psychiatry* 71(10)**,** 846-854. doi: 10.1016/j.biopsych.2011.12.019.
21
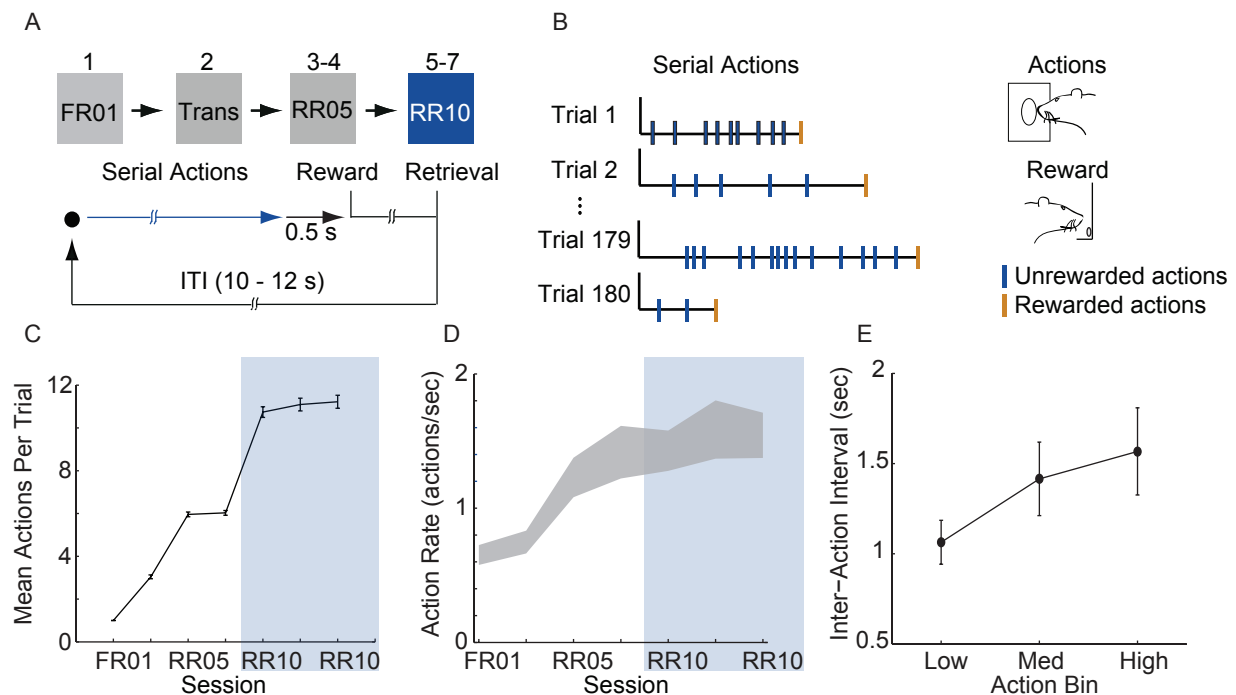
Fig. 1

Fig. 1. Animals learn to perform serial actions. (A) Actions (nose pokes into a lit port) were reinforced probabilistically with sucrose pellets (rewards). In session 1, each action was reinforced (fixed ratio 1, FR01). In session 2, reinforcement probability decreased to 0.2 in 3 trial blocks (transition, TRANS). In sessions 3-4, the probability of reinforcement was 0.2 (random ratio 5, RR05). In sessions 5-7, the probability of reinforcement was 0.1 (random ratio 10, RR10). At trial start, a cue light was illuminated until the outcome was earned (blue arrow). When an action was reinforced, the cue light was extinguished immediately and the outcome was delivered 0.5 sec later. (B) Random reinforcement led to different numbers of serial actions performed per trial. (C) Mean ± SEM number of actions required per trial. Shaded area depicts RR10 sessions. Serial action data were drawn from these sessions because these sessions required the greatest average number of actions per trial, and thus, had the greatest statistical power. (D) Mean ± SEM action rate in each session. (E) Mean ± SEM inter-action intervals in low, medium and high bins (actions 1-7, 8-14, and 15-21) during RR10 sessions.

Fig. 2
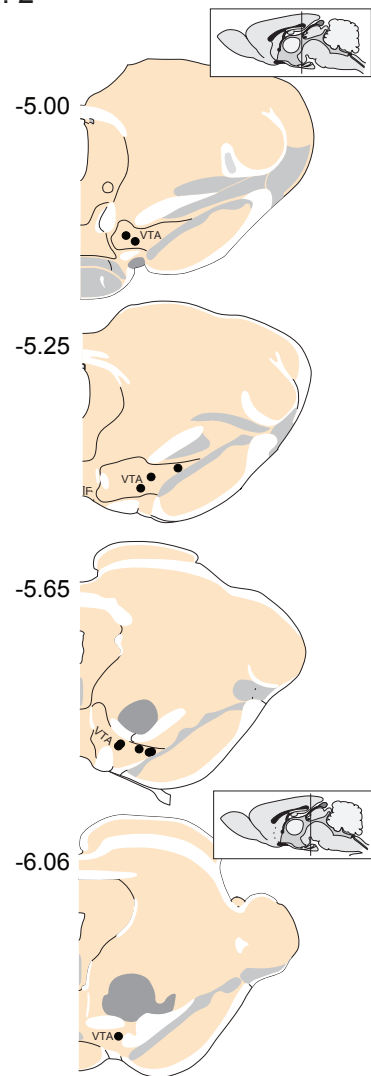
Fig. 2. Locations of recording electrodes within VTA. Each dot represents the location of a recording array. Insets show midsagittal diagram of rodent brain with vertical lines representing the approximate location of the most anterior and posterior coronal sections shown.
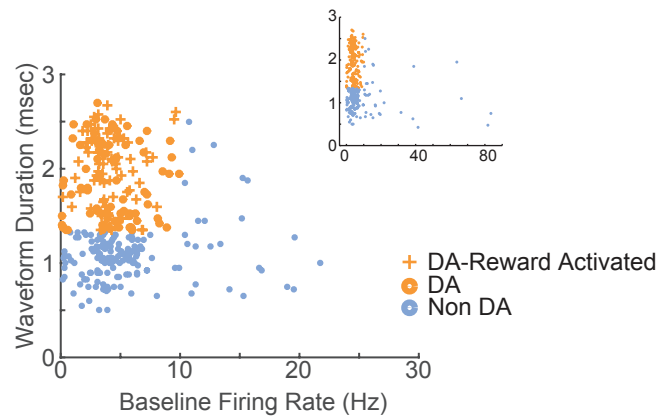
Fig. 3

Fig. 3. Electrophysiological characterization of recorded VTA neurons. Each point represents a single neuron; data from all sessions are depicted. Neurons with wide spike widths (>1.4 ms) and low firing rates (< 10 Hz) were considered putative dopamine neurons. Crosses depict reward-responsive dopamine neurons. The remaining neurons were considered putative non-dopamine neurons. The main figure depicts the data with the horizontal axis limited to 30, for better visualization of the majority of the data. Inset shows the same data, with the horizontal axis extended to encompass the full range of baseline firing rates.

Fig. 4
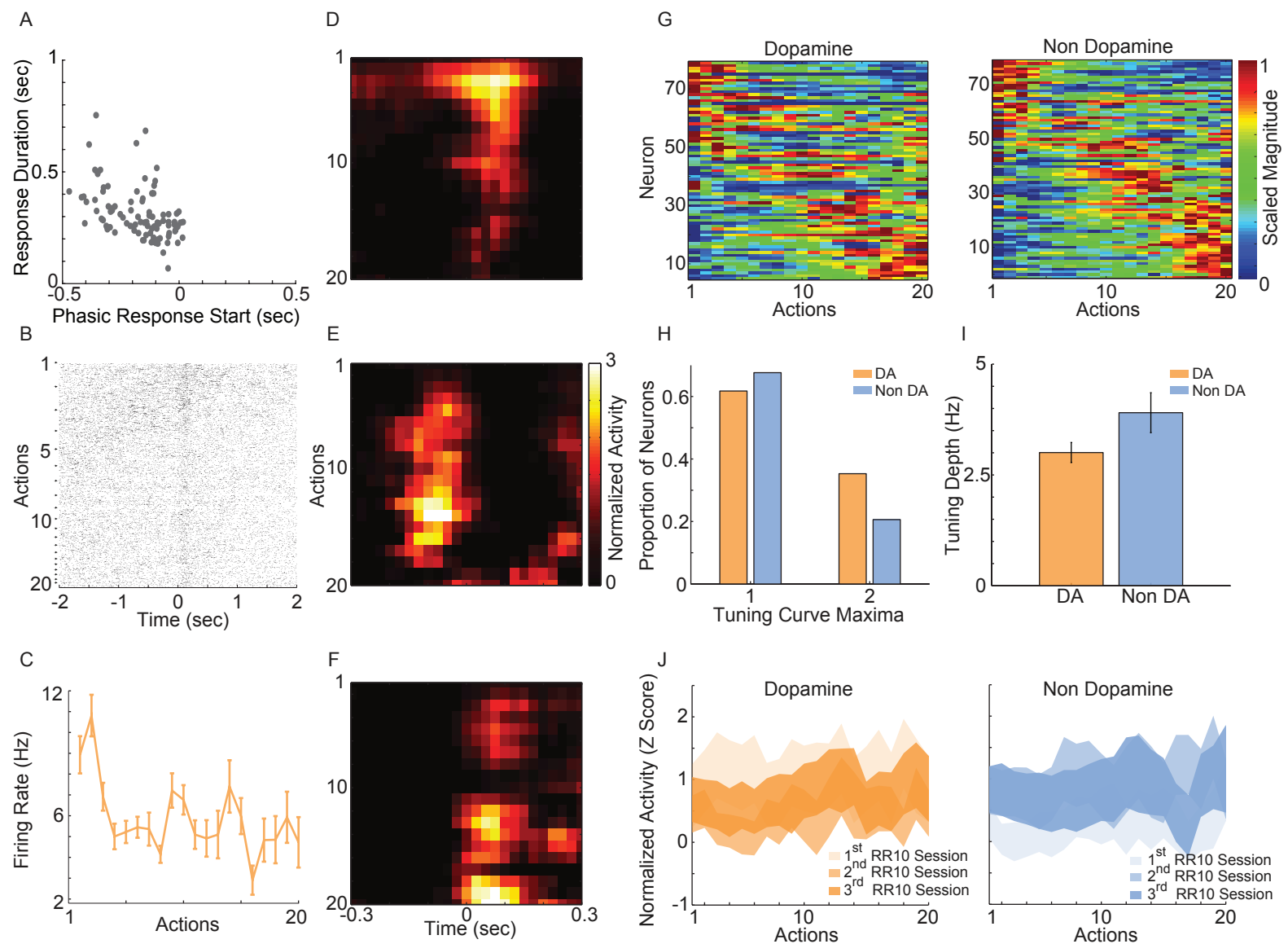
Fig. 4. RR10 action evoked VTA responses. (A) Scatter plot depicts response start and duration of action evoked responses in RR10 sessions. Each point represents the data from a single unit. (B) Data from a typical dopaminergic neuron that preferred low action numbers. The raster shows spikes aligned to the time of action execution (time = 0). Each row of the raster represents one action evoked response and rows are arranged by action number. Each arrow on the right represents an action number. For each action number, the earlier occurrences of an Nth numbered action are arranged toward the top. Thus, the first row of the raster represents the first occurrence of an action number 1 (i.e. trial 1), and the second row represents the second occurrence (i.e. trial 2), etc. (C) Depicts the action-evoked responses of the neuron depicted in (B) conditioned on action numbers 1-20. Note that the neuron most strongly prefers actions 1 and 2. (D-F) Neuronal responses aligned to actions for 3 simultaneously recorded neurons. (D) A non-dopaminergic neuron that fires preferentially during lower numbered actions within a trial. Data depicted as the average response evoked by actions 1-20. Data are aligned to the time of action execution (0.025 sec bins). Only unrewarded actions are depicted. (E-F) A pair of dopaminergic neurons that fire preferentially during upper-middle, and higher numbered actions, respectively. Note that both dopaminergic and non-dopaminergic VTA neurons fire preferentially around distinct subsets of actions.(G) Action number tuning curves during RR10 (0.25 sec window, centered on action). Each dopamine (left) and non-dopamine (right) neuron's average, action evoked, normalized, firing rate during execution of actions is plotted by color. Tuning curves are scaled to each curve's peak (hot colors = tuning curve peak). Neurons are sorted by the location of the peak of the tuning function. Note the diversity of tuning curves to serial actions (H) Proportion neurons with 1 or 2 tuning curve maxima. (I) Mean ± SEM tuning curve depth. (J) Mean ± SEM neuronal responses as a function of action number in RR10 sessions for dopamine and non-dopamine neurons.

Fig. 5

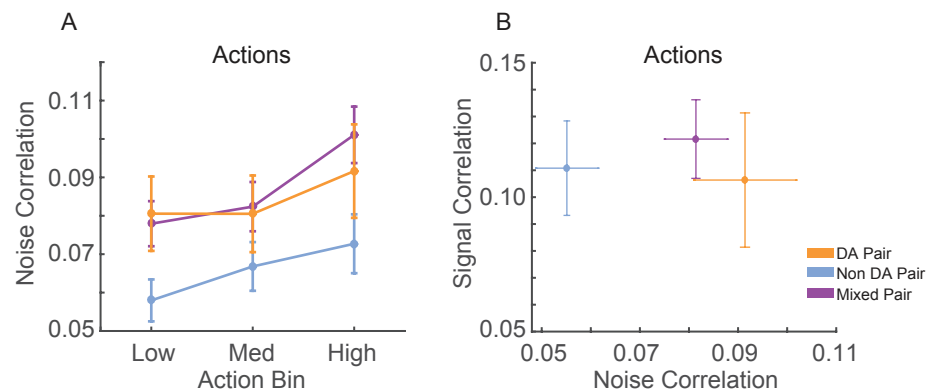

Fig. 5. Correlation structure of RR10 action-evoked responses. (A) Mean ± SEM action evoked noise correlations between simultaneously recorded pairs of VTA neurons in RR10 sessions during actions. Noise correlations measure trial-by-trial fluctuations in pairwise activity. Data depicted as mean ± SEM noise correlation in each bin of actions. Data are plotted separately for all pairings of VTA neurons. (B) Mean ± SEM signal correlation (correlation between tuning curves) and noise correlations during RR10 actions. For simplicity, noise correlations are collapsed across action number bin.
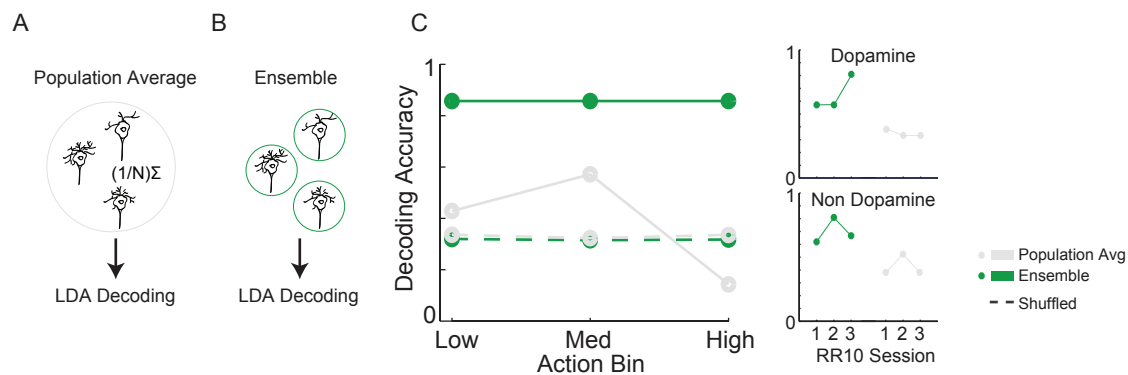
Fig. 6. Decoding binned action number from VTA neuronal activity. Linear discriminant analysis (LDA) utilized action-evoked activity averaged across (A) all neurons (population-average), or (B) the collective activity of each neuron (ensemble). (C) The proportion of correctly classified actions for each decoder. Dashed lines represent decoding of shuffled control data. Inset to the right depicts performance of ensemble and population-average decoders across consecutive sessions in dopamine and non-dopamine neurons. Note that ensemble activity was decoded significantly more accurately than population averaged activity and shuffled control. Inset data are collapsed across high, medium, and low action numbers and depicted separately for each session.
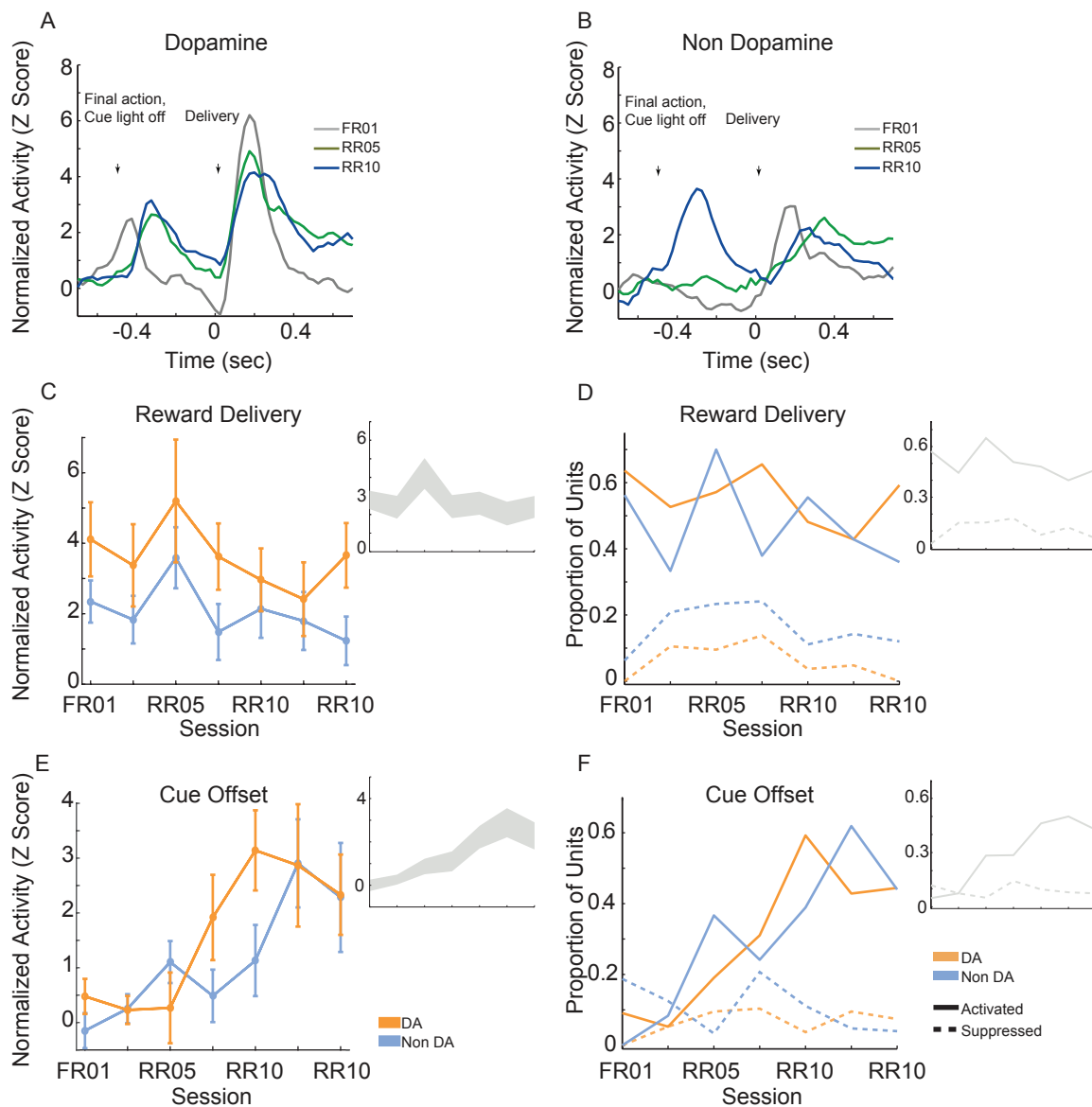
Fig. 7

Fig. 7. VTA responses to rewards and reward predictive cues. Normalized mean (A) dopamine and (B) non-dopamine responses aligned to outcome delivery (time = 0). The final action in each trial (left arrow) occurred 0.5s prior to outcome delivery (right arrow). Cue light offset was simultaneous with execution of the final action, which signaled trial completion and impending outcome delivery. Data are depicted for FR01, final RR05 session, and the final RR10 session. (C) Mean ± SEM responses evoked by reward delivery (+0.05 - +0.3s, relative to reward delivery) in both groups of neurons and all sessions. Dopamine responses were greater than non-dopamine neuronal responses. Inset depicts data from all VTA neurons plotted without respect to neuronal classification. (D) The proportion of neurons classified as either significantly activated (solid lines) or suppressed (dashed lines) by outcome delivery. Data are depicted across all sessions for putative dopamine and non-dopamine neurons. Inset depicts data from all VTA neurons plotted without respect to neuronal classification. (E) Mean ± SEM responses following cue offset, during the pre-reward delivery period (+0.150 - +0.400s, relative to cue offset). Data are depicted similarly to (C). In each group, the evoked response increased across sessions. (F) The proportion of neurons activated (solid lines) or suppressed (dashed lines) by cue offset signaling future reward delivery. Data depicted similarly to (D).
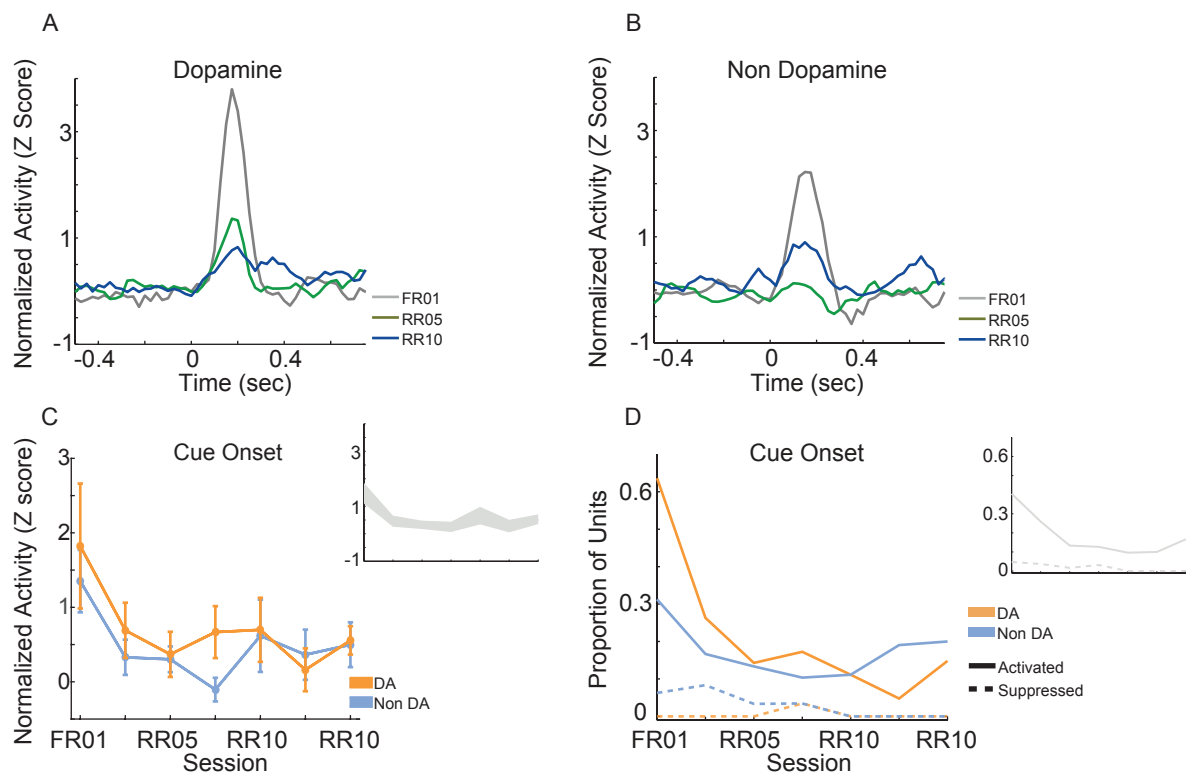
Fig. 8

Fig. 8. VTA responses to cue onset. Normalized mean (A) dopamine and (B) non-dopamine responses to cue onset (time = 0). Data are depicted for the FR01 session, the final RR05 session, and the final RR10 session. (C) Mean ± SEM responses evoked by cue onset (+0.05 - +0.3s) in all groups of neurons and sessions. Data are depicted separately for all putative dopamine and non-dopamine neurons. Responses were greatest in session 1 and significantly less in subsequent sessions, with no difference by neuron type. Inset depicts data from all VTA neurons plotted without respect to neuronal classification. (D) Proportion of neurons activated (solid lines) or suppressed (dashed lines) by cue onset. Data are depicted across all sessions for putative dopamine and non-dopamine neurons. The proportion of activated dopamine neurons, but not non-dopamine neurons, decreased in later sessions. Inset depicts data from all VTA neurons plotted without respect to neuronal classification.
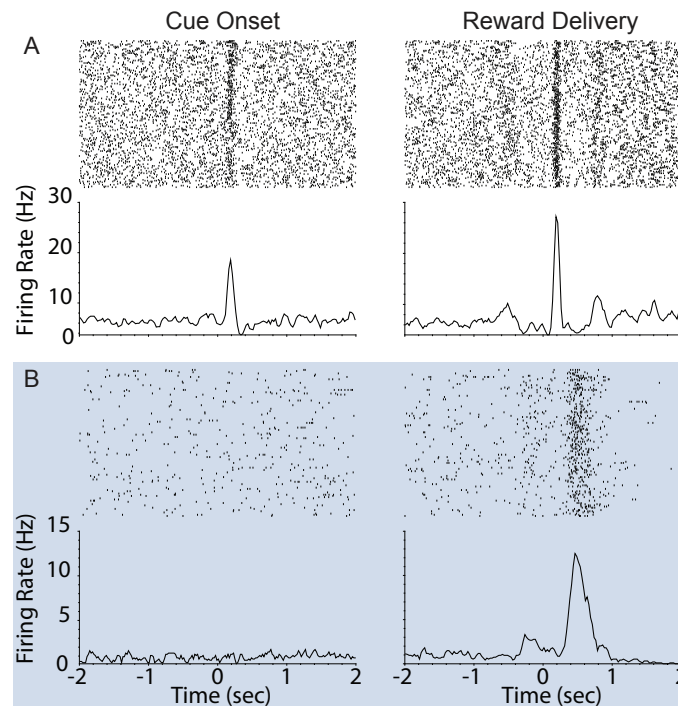
Fig. 9

Fig. 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 and 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 panels (A) and (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.

Fig. 10

Fig. 10. Correlation structure of cue and reward evoked responses. Mean ± SEM noise correlation between simultaneously recorded pairs of VTA neurons in RR10 sessions, during baseline, cue onset, and reward delivery. Data depicted separately for all pairings of neurons. Note that reward delivery evoked the largest magnitude noise correlations in all pairings.
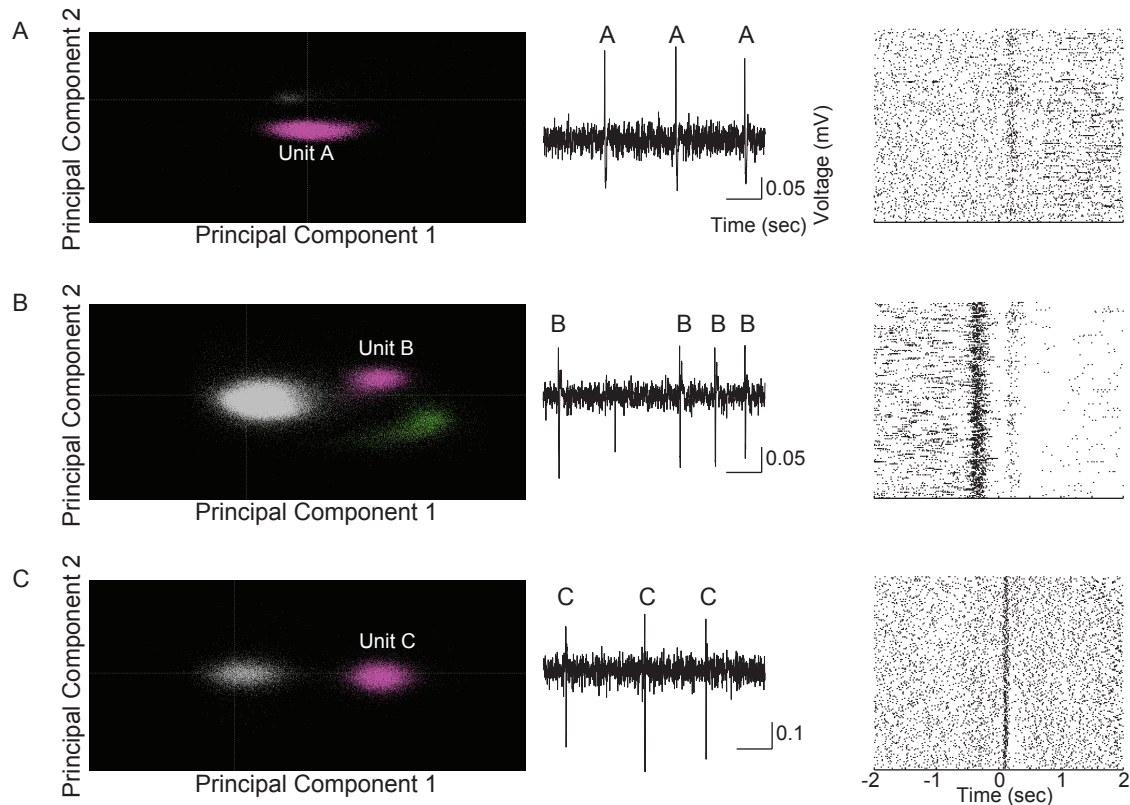
Fig. S1

Fig. 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 – 0s), 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.