Ventral tegmental area neurons use an ensemble code for representing information about ongoing actions

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

Affiliations:

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

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

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

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

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

[6]Department of Neuroscience, University of Pittsburgh, A210 Langley Hall, Pittsburgh, PA 15260.

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

We observed that ensembles of VTA neurons encode information about ongoing behavior in real-time. As rats executed varying numbers of actions to obtain the same reward, the firing rates of different neurons were preferentially tuned to low, medium, or high numbered bins of actions, allowing information about ongoing behavior to be decoded from the ensemble activity. This ensemble code could be used to organize and sustain behavioral responses.

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 also depends on VTA dopamine input (Goldman-Rakic, 1988; Robbins & Arnsten, 2009; Salamone *et al.*, 2009). 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 during reward-motivated behavior (Aberman & Salamone, 1999; Ishiwari *et al.*, 2004; Mingote *et al.*, 2005). Recordings from VTA neurons, however, have primarily focused on homogenous responses to reward or aversion related events that require no or few actions (Schultz, 1998; Matsumoto *et al.*, 2016). Little is known about how VTA neurons encode information when multiple or varying number of actions are required to achieve a goal.

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 (Fig. 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 trials that were randomly reinforced, unpredictable and varying numbers of actions were required but, importantly, each action in a trial was equally likely to be reinforced.

Action response rates increased as reinforcement probability decreased (Fig. 1C, D), with the final RR10 sessions significantly higher than all other sessions (Fig. 1D; $F_{(6,24)} = 4.726$, $p = 0.003$). To investigate a potential relationship between behavioral performance and action number, we divided the data RR10 data into three bins: 1-7 actions (low), 8-14 actions

(medium), and 15-21 actions (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 (Fig. 1C). Within a trial, the inter-action interval significantly increased with increasing action number bins (Fig. 1E; $F_{(2,44)}$ = 11.020, $p < 0.001$). This indicated that behavior was sensitive to action number. The latency to retrieve the reward ($r_{(3468)}$ = -0.017, p = 0.309) or initiate responding in the next trial, however, was not correlated with binned action number ($r_{(3445)}$ = 0.015, p = 0.389). These data suggested the number of actions performed in a trial did not affect behavioral responses to reward or cues.

Neurophysiological recordings

VTA units (n=375, Fig. S1-S3) were recorded over 7 sessions from 10 rats. 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 as putative 'dopamine' (n = 155) or 'non-dopamine' neurons (n=220) based on validated criteria (Grace & Bunney, 1983; Schultz, 1998; Ungless & Grace, 2012). This permits comparison with previous work, with the caveat that some dopamine neurons co-release other neurotransmitters (Tritsch *et al.*, 2012), or potential inaccuracies (Margolis *et al.*, 2006) in these approaches. Analyses were also performed on reward-responsive dopamine neurons, as this subgroup may be a more conservative estimate of neuronal identity (Lak *et al.*, 2014; Eshel *et al.*, 2016).

Actions in a trial are accurately discriminated from VTA neuronal ensemble activity

In order to understand how VTA neurons encoded information during performance of serial actions, we examined how activity was modulated by actions in RR10 sessions. We observed preferential responding to low, medium, or high numbered actions amongst simultaneously recorded neurons (Fig. 2A-C). Serial actions significantly modulated the activity of 51% of the population, with no preferential effect on dopamine or non-dopamine neurons ($\chi^2_{(1)}$ = 0.179, p = 0.673). 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 (Fig. 3A). 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 (Fig. 3A, B). Tuning curve depth (maximum – minimum) did not differ significantly between putative dopamine and non-dopamine neurons (Fig. 3C; $t_{(154)} = 3.224$, $p = 0.084$). Averaging activity across neurons, a traditional approach to analyzing VTA neuronal data concealed this heterogeneity in either group of neurons (Fig. 3D, E). Time elapsed in a trial, and not actions, was a significant predictor of activity in only 7% of the neurons, suggesting action-evoked activity reflected behavior and not the passage of time.

The diversity of tuning curves that we observed in VTA 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), between all simultaneously recorded pairs of neurons. These elements of the correlation structure are reflective of functional connectivity between neurons and similarity in tuning curves, respectively (Cohen & Kohn, 2011). Correlations were examined between all possible pairings of VTA neurons (dopamine – dopamine, non-dopamine – non-dopamine, and dopamine – non-dopamine).

Noise correlations during actions were significantly lower in pairs of non-dopamine neurons than pairs containing a dopamine neuron (Fig. 4A; $F_{(2,858)} = 4.052$, $p = .018$). Noise correlations were decreased in low action number bins ($F_{(2,1716)} = 14.625$, $p < .001$), demonstrating that the correlation structure increases with subsequent actions in a trial. These differences are not likely due to differences in spike count, as the pairwise geometric mean spike count was not associated with noise correlation magnitude in any bin (low: $r = 0.008$, $p = .805$; med: $r = -0.015$, $p = .650$; high: $r = -0.048$, $p = .158$). Signal correlation strength did not

differ between pairings of VTA neurons (Fig. 3I; $F_{(2,858)} = 0.289$, $p = .749$). There was a strong association between signal correlations and noise correlations (Fig. 4B, Table 1), suggesting there is a higher degree of functional connectivity between pools of similarly tuned neurons. This in turn raises the possibility that distinct patterns of functional connectivity contribute to the diversity of tunings to actions.

Taken together, the tuning curve heterogeneity, absence of a clear population average signal for serial actions, and the association between action-evoked signal and noise correlations, suggested that information about ongoing sequences of behavior may be encoded by VTA ensembles. To quantify the accuracy of this encoding mechanism we discriminated low-action, medium-action and high-action number bins from population-averaged activity (Fig. 5A) or the collective activity of VTA ensembles (Fig. 5B). The same cross-validated linear discriminant analysis decoding algorithms were used for both the ensemble and population-average data. Ensemble activity was discriminated significantly more accurately than population-averaged activity (Fig. 5C; $\chi^2_{(1)} = 21.435$, $p < 0.001$). 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 ($\chi^2_{(1)} = 1.300$, $p = 0.254$). Ensemble activity was discriminated significantly more accurately than shuffled control data (Fig. 5C; permutation test, $p < 0.001$). In contrast, population averaged activity was decoded at chance levels (Fig. 5C; permutation test, $p = 0.240$). Taken together, 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.

Reward delivery evoked responses

Reward delivery evoked greater activity in dopamine neurons than in non-dopamine neurons (Fig. 6A, B; $F_{(1,361)} = 9.159$, $p = .003$). This pattern did not differ across sessions (Fig. 6C; $F_{(6,361)} = 1.352$, $p = .233$). The proportion of dopamine neurons responding to reward delivery did not vary significantly across sessions (Fig. 6D; 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 learning progressed, cue offset evoked great activity levels (Fig. 6E; $F_{(6,361)} = 4.776$, $p < .001$), and activated a greater proportion of neurons (Fig. 6F; 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 classically attributed to dopamine neurons (Montague *et al.*, 1996; Schultz, 1998).

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. 7A), which is consistent with previous observations (Schultz, 1998; Roesch *et al.*, 2007). Non-dopamine neurons followed a similar pattern (Fig. 7B). During random ratio sessions, the population response to the cue decreased significantly as response requirement increased (Fig. 7C; session, $F_{(6,361)} = 2.667$, $p = .015$). This pattern of responding did not differ between dopamine and non-dopamine neurons (Figs. 3D, E; $F_{(1,361)} = 1.543$, $p = .215$). We found the same effects when analysis was restricted to neurons which were significantly activated by reward delivery (session, $F_{(6,175)} = 4.961$, $p < .001$; 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 (Fig. 7D; all VTA, $X^2_{(6)} = 23.844$, $p = .001$; dopamine, $X^2_{(6)} = 20.109$, $p = .003$; non-dopamine, $X^2_{(6)} = 10.636$, $p = .100$).

Cue evoked responses decouple from reward evoked responses with learning

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 S4). This correlation decreased in RR05 sessions (Table S4), and was no longer significant by RR10 sessions (Table S4). A similar pattern was observed in dopamine neurons that were activated by reward delivery (Table S4). Taken together, this suggests that, though cue and reward responses are positively correlated in Pavlovian conditioning tasks (Eshel *et al.*, 2016), when serial actions are reinforced randomly cue-evoked responses become decoupled from reward responses.

Modulation of noise correlations by cue onset and reward delivery

We calculated the signal and noise correlation (Cohen & Kohn, 2011) 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 (Fig. 8; $F_{(4,2496)} = 9.482$, $p < .001$). There was no difference in baseline noise correlations between any pairing of VTA neurons (Fig. 8; $F_{(2,832)} = 0.855$, $p = .426$). Noise correlations evoked by cue onset were significantly greater in pairs containing a dopamine neuron than pairs of non-dopamine neurons (Fig. 8; $F_{(2,832)} = 5.881$, $p = .003$). There was no association between cue evoked noise correlations and pairwise cue evoked geometric mean spike counts ($r = 0.008$, $p = .817$), 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 (Fig. 8; $F_{(2,832)} = 2.794$, $p = .062$). 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 a series of nose pokes to earn rewards. Both dopamine and non-dopamine VTA neurons were preferentially tuned to unique subsets of actions, and pairs of neurons with similar tuning curves had higher action-evoked noise correlations. This diversity of tunings allowed serial actions to be accurately discriminated from the collective activity of VTA ensembles. In line with previous work suggesting that dopamine neurons are activated by conditioned stimuli predictive of reward delivery (Schultz, 1998), population responses to cue offset following action completion (predicting impending rewards) increased as learning occurred. In contrast, cue onset preceding a series of behaviors did not evoke VTA neuronal responses, and the correlation between cue onset and reward delivery responses was diminished in sessions requiring serial actions, suggesting that VTA responses were sensitive to the contingency and contiguity of stimuli with reward.

Encoding information during an action series

Previous reports indicate dopamine neurons encode errors in the predicted value of rewards and encode action value (Montague *et al.*, 1996; Schultz, 1998; Morris *et al.*, 2006; Roesch *et al.*, 2007). Dopamine neurons also phasically fire during the first and last action in a sequence, which may be critical to initiating and terminating behavior (Jin & Costa, 2010), and dopamine release in a short sequence of actions occurs during the first action (Wassum *et al.*, 2012). In addition to these roles, we propose that the collective activity of ensembles of VTA neurons has the capacity to encode information about ongoing behavior in real-time. This code could be used to organize and sustain responding, which is a fundamental feature of

dopamine's role in cognition (Aberman & Salamone, 1999; Ishiwari *et al.*, 2004; Salamone *et al.*, 2009). Thus, our data provide new insight into how VTA neurons encode a signal that could be critical for supporting the unique motivational and cognitive demands of trials requiring large numbers of actions (Goldman-Rakic, 1998; Aberman & Salamone, 1999; Salamone & Correa, 2002; Seamans & Yang, 2004; Robbins & Roberts, 2007; Salamone *et al.*, 2007; Robbins & Arnsten, 2009).

VTA dopamine neurons project to several networks that may utilize VTA signals during serial actions to organize behavior. For instance, dopaminergic 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 & Salamone, 1999; Ishiwari *et al.*, 2004; Mingote *et al.*, 2005). A second line of evidence suggests that prefrontal cortex processes which require real-time information about ongoing behavior, such as behavioral flexibility and executive functions, also require innervation from VTA dopamine neurons (Goldman-Rakic, 1998; Seamans & Yang, 2004; Robbins & Roberts, 2007; Robbins & Arnsten, 2009). Dopamine release occurring exclusively at the beginning of a trial cannot encode information that emerges during performance of an action series, and therefore, could not be used to organize or sustain behavior in real-time. Thus, it may be necessary for VTA to continually encode information throughout execution of a lengthy series of actions, so that striatal and prefrontal networks can track effort expenditure and progress toward a goal.

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 & 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 have more similar connection properties. VTA receives inputs from an expansive set of afferent

structures (Geisler & Zahm, 2005; Geisler *et al.*, 2007; Watabe-Uchida *et al.*, 2012), 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 suggests that VTA network properties are critical to understanding how information about serial behaviors 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.

Both dopamine and non-dopamine neurons had similarly 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 & 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.

Conclusions

Our work elucidates a novel mechanism for encoding information about a series of actions from start to finish. The ensemble signal we observed represents one of the first observations of such a code in the dopamine system. We propose that encoding real-time information about ongoing behavior may be a critical component of dopamine function. Accordingly, behavioral organization pathologies may arise from compromised integrity and lack of cooperation in VTA networks (Moghaddam & Wood, 2014). Dysregulated behavioral organization in psychiatric disorders may arise from compromised integrity and lack of cooperation in VTA networks. Moving forward, it is critical to incorporate ensemble encoding into our theories of healthy and unhealthy VTA function.

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-hour light cycle (lights on at 7pm). 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; Kim *et al.*, 2012; Kim *et al.*, 2016). At the time of surgery, rats weighed approximately 350 grams (approximately 12 weeks of age). Under isoflurane anesthesia, 8 or 16 channel, 50 µm 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). A subset of rats also was implanted in left prelimbic cortex (data not presented). 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 minutes 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 minute magazine training session in the operant chamber, in which sugar pellets were delivered on a variable time 75s reinforcement schedule. When each pellet was delivered, the pellet trough was illuminated for 4s. 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 0.500s 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 10-12s 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 ration 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 µm 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 & 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.

A VTA neuron was classified as dopaminergic if it had broad action potentials, greater than 1.4 ms in duration, and a mean baseline firing rate less than 10 Hz. These criteria are similar to those used in previous studies (Hyland *et al.*, 2002; Fiorillo *et al.*, 2003; Anstrom & 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. Should neurons have been misclassified, it is unlikely that this strongly affects the conclusions of the current work. 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 (0.025s 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 (5s to 1s 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 0.250s windows around events (cue onset: +0.050 – 0.300s, relative to cue onset; -0.125 – +0.125s, relative to the time of action execution; time period between cue offset and outcome delivery: +0.150 – 0.400s, relative to execution of the last action; outcome delivery: +0.050 – 0.300s, 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 .250 sec windows used to quantify firing rates, using the standard Pearson's correlation,

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

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 recoded neurons. We corrected for chance levels of correlation, which were estimated as the mean correlation of 1000 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 proportions of neurons with significant coefficients for either predictor variable 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 (Fig. 3A). 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 spline 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

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

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 $\vec{R}$, $p(\bar{R}|A)$ does not thoroughly represent the information in $p(\vec{R}|A)$. Therefore, we instead assume relevant information is captured by a higher-dimensional representation: $(W\vec{R}|A) \sim N(\mu_A, \Sigma)$. The projection, W, estimated by principal components analysis, is necessary to ensure the covariance, Σ , can be estimated by the usual pooled sample covariance, $\hat{\Sigma}$. This maximizes the dimensionality of the space into which we project $\vec{R}$, for a given sample size. We then maximize $p(A|W\vec{R})$. 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, $\overrightarrow{R*}$, 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 versus 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

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

where $I_A\left(\hat{A}\right)$ 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 (versus 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.

ACKNOWLEDGEMENTS

Funding was obtained from the NIMH MH048404 (BM), MH064537 (SK), MH064537 (REK), MH016804 (JW); NIDA DA031111 (JW), DA035050 (NWS), DA022762 (SK); RK Mellon Foundation (SK); Andrew Mellon Fellowship (JW). JW designed experiments, performed experiments, analyzed data, and wrote the manuscript. NWS designed experiments and edited the manuscript. SK analyzed the data. REK analyzed data and edited the manuscript. BM designed experiments and wrote the manuscript. The authors have no conflicts of interest related to this work.

REFERENCES

Aberman, J.E. & 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.

Anstrom, K.K., Cromwell, H.C. & Woodward, D.J. (2007) Effects of restraint and haloperidol on sensory gating in the midbrain of awake rats. *Neuroscience*, **146**, 515-524.

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

Barker, D.J., Root, D.H., Zhang, S. & Morales, M. (2016) Multiplexed neurochemical signaling by neurons of the ventral tegmental area. *J Chem Neuroanat*, **73**, 33-42.

Cohen, M.R. & Kohn, A. (2011) Measuring and interpreting neuronal correlations. *Nat Neurosci*, **14**, 811-819.

Eshel, N., Tian, J., Bukwich, M. & Uchida, N. (2016) Dopamine neurons share common response function for reward prediction error. *Nat Neurosci*, **19**, 479-486.

Fiorillo, C.D., Tobler, P.N. & Schultz, W. (2003) Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, **299**, 1898-1902.

Geisler, S., Derst, C., Veh, R.W. & Zahm, D.S. (2007) Glutamatergic afferents of the ventral tegmental area in the rat. *J Neurosci*, **27**, 5730-5743.

Geisler, S. & Zahm, D.S. (2005) Afferents of the ventral tegmental area in the rat-anatomical substratum for integrative functions. *J Comp Neurol*, **490**, 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.

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.

Grace, A.A. & Bunney, B.S. (1983) Intracellular and extracellular electrophysiology of nigral dopaminergic neurons--1. Identification and characterization. *Neuroscience*, **10**, 301-315.

Hyland, B.I., Reynolds, J.N., Hay, J., Perk, C.G. & Miller, R. (2002) Firing modes of midbrain dopamine cells in the freely moving rat. *Neuroscience*, **114**, 475-492.

Ishiwari, K., Weber, S.M., Mingote, S., Correa, M. & 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.

Jin, X. & Costa, R.M. (2010) Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature*, **466**, 457-462.

Joshua, M., Adler, A., Prut, Y., Vaadia, E., Wickens, J.R. & Bergman, H. (2009) Synchronization of midbrain dopaminergic neurons is enhanced by rewarding events. *Neuron*, **62**, 695-704.

Kass, R.E., Eden, U.T. & Brown, E.N. (2014) *Analysis of neural data*.

Kim, Y., Simon, N.W., Wood, J. & Moghaddam, B. (2016) Reward Anticipation Is Encoded Differently by Adolescent Ventral Tegmental Area Neurons. *Biol Psychiatry*, **79**, 878-886.

Kim, Y., Wood, J. & Moghaddam, B. (2012) Coordinated activity of ventral tegmental neurons adapts to appetitive and aversive learning. *PLoS One*, **7**, e29766.

Kim, Y.B., Matthews, M. & 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.

Lak, A., Stauffer, W.R. & Schultz, W. (2014) Dopamine prediction error responses integrate subjective value from different reward dimensions. *Proc Natl Acad Sci U S A*, **111**, 2343-2348.

Margolis, E.B., Lock, H., Hjelmstad, G.O. & Fields, H.L. (2006) The ventral tegmental area revisited: is there an electrophysiological marker for dopaminergic neurons? *J Physiol*, **577**, 907-924.

Matsumoto, H., Tian, J., Uchida, N. & Watabe-Uchida, M. (2016) Midbrain dopamine neurons signal aversion in a reward-context-dependent manner. *eLife*, **5**.

Mingote, S., Weber, S.M., Ishiwari, K., Correa, M. & 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.

Moghaddam, B. & Wood, J. (2014) Team work matters: coordinated neuronal activity in brain systems relevant to psychiatric disorders. *JAMA psychiarty*.

Montague, P.R., Dayan, P. & 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. & Bergman, H. (2006) Midbrain dopamine neurons encode decisions for future action. *Nat Neurosci*, **9**, 1057-1063.

Pan, W.X., Schmidt, R., Wickens, J.R. & 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.

Robbins, T.W. & Arnsten, A.F. (2009) The neuropsychopharmacology of fronto-executive function: monoaminergic modulation. *Annu Rev Neurosci*, **32**, 267-287.

Robbins, T.W. & Roberts, A.C. (2007) Differential regulation of fronto-executive function by the monoamines and acetylcholine. *Cerebral cortex*, **17 Suppl 1**, i151-160.

Roesch, M.R., Calu, D.J. & Schoenbaum, G. (2007) Dopamine neurons encode the better option in rats deciding between differently delayed or sized rewards. *Nat Neurosci*, **10**, 1615-1624.

Salamone, J.D. & Correa, M. (2002) Motivational views of reinforcement: implications for understanding the behavioral functions of nucleus accumbens dopamine. *Behav Brain Res*, **137**, 3-25.

Salamone, J.D., Correa, M., Farrar, A. & Mingote, S.M. (2007) Effort-related functions of nucleus accumbens dopamine and associated forebrain circuits. *Psychopharmacology (Berl)*, **191**, 461-482.

Salamone, J.D., Correa, M., Farrar, A.M., Nunes, E.J. & Pardo, M. (2009) Dopamine, behavioral economics, and effort. *Front Behav Neurosci*, **3**, 13.

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

Schultz, W. (2010) Dopamine signals for reward value and risk: basic and recent data. *Behav Brain Funct*, **6**, 24.

Schultz, W., Apicella, P. & 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**, 900-913.

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

Swanson, L.W. (2004) *Brain maps III : structure of the rat brain : an atlas with printed and electronic templates for data, models, and schematics*. Elsevier, Academic Press, Amsterdam ; Boston.

Tobler, P.N., Fiorillo, C.D. & Schultz, W. (2005) Adaptive coding of reward value by dopamine neurons. *Science*, **307**, 1642-1645.

Totah, N.K., Kim, Y. & Moghaddam, B. (2013) Distinct prestimulus and poststimulus activation of VTA neurons correlates with stimulus detection. *J Neurophysiol*, **110**, 75-85.

Tritsch, N.X., Ding, J.B. & Sabatini, B.L. (2012) Dopaminergic neurons inhibit striatal output through non-canonical release of GABA. *Nature*, **490**, 262-266.

Ungless, M.A. & Grace, A.A. (2012) Are you or aren't you? Challenges associated with physiologically identifying dopamine neurons. *Trends Neurosci*, **35**, 422-430.

Wassum, K.M., Ostlund, S.B. & Maidment, N.T. (2012) Phasic mesolimbic dopamine signaling precedes and predicts performance of a self-initiated action sequence task. *Biol Psychiatry*, **71**, 846-854.

Watabe-Uchida, M., Zhu, L., Ogawa, S.K., Vamanrao, A. & Uchida, N. (2012) Whole-brain mapping of direct inputs to midbrain dopamine neurons. *Neuron*, **74**, 858-873.

Fig. 1

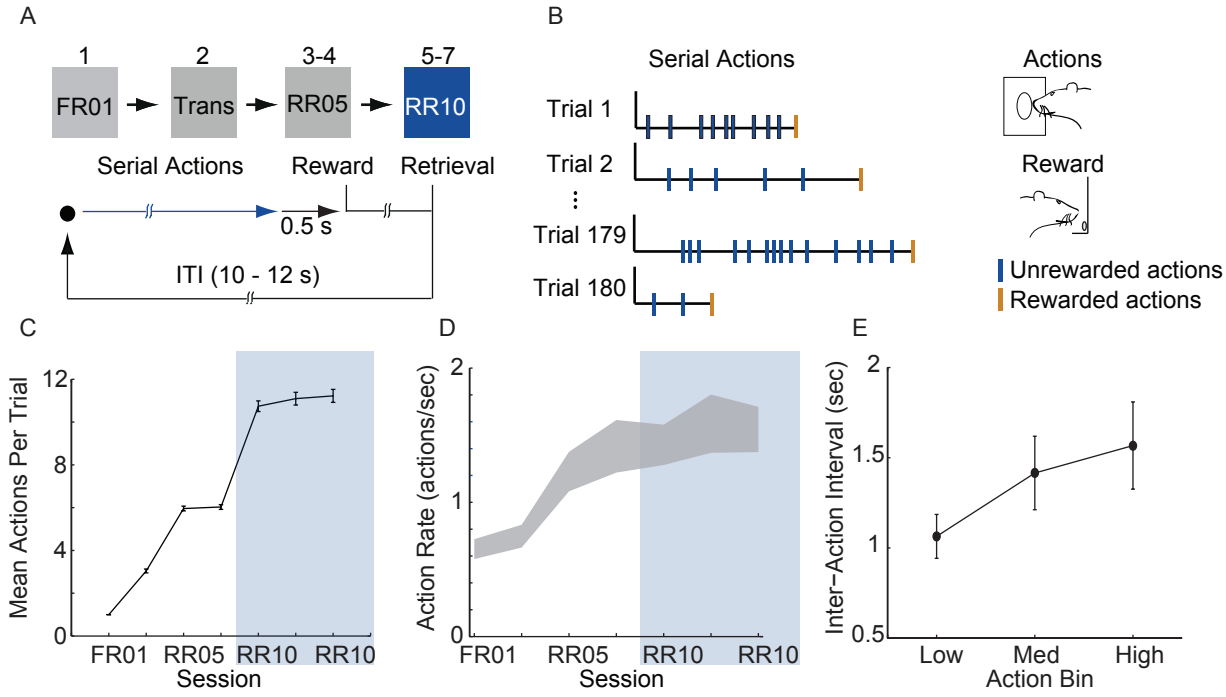

Fig. 1. (A) Actions (nose pokes into a lit port) were reinforced probabilistically with sucrose pellets (rewards). In session 1, each action was reinforced (fixed ratio 1, FR01). In session 2, reinforcement probability decreased to 0.2 in 3 trial blocks (transition, TRANS). In sessions 3-4, the probability 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

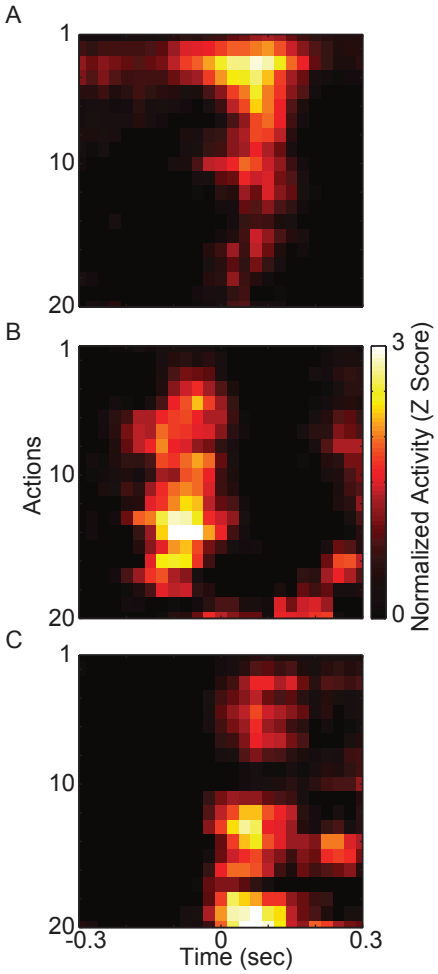

Fig. 2. Representative neuronal responses aligned to actions for 3 simultaneously recorded neurons. **(A)** A non-dopaminergic neuron that fires preferentially during lower numbered actions within a trial. Data depicted as the average response evoked by actions 1-20. Data are aligned to the time of action execution (0.025 sec bins). Only unrewarded actions are depicted. **(B-C)** 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.

Fig. 3

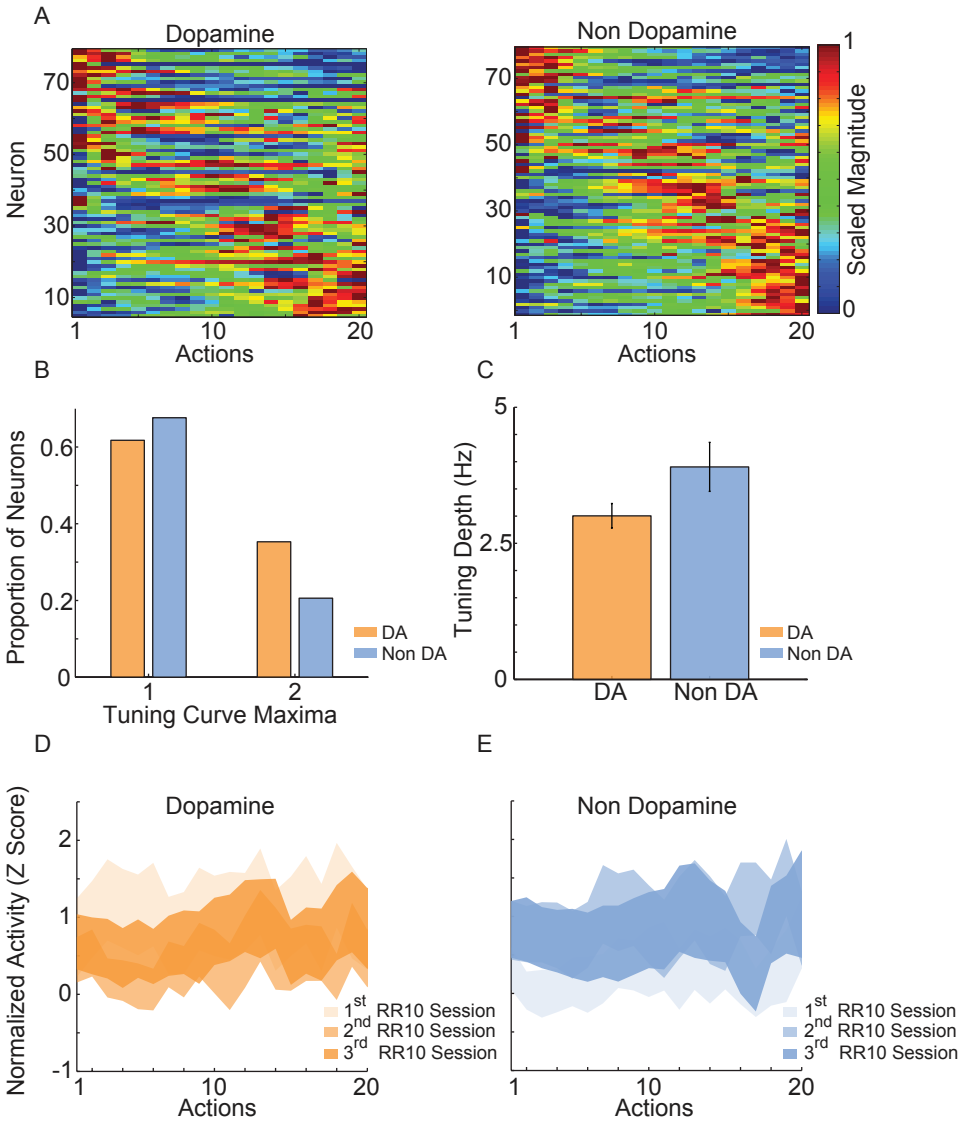

Fig. 3. (A) Action number tuning curves during RR10 (0.25 sec window, centered on action). Each dopamine (left) and non-dopamine (right) neuron's average, action evoked, normalized, firing rate during execution of actions is plotted by color (hot colors = tuning curve peak). Neurons are sorted by the location of the peak of the tuning function. Note the diversity of tuning curves to serial actions **(B)** Proportion neurons with 1 or 2 tuning curve maxima. **(C)** Mean ± SEM tuning curve depth. **(D)** Mean ± SEM neuronal responses as a function of action number in RR10 sessions for dopamine and **(E)** non-dopamine neurons.

Fig. 4

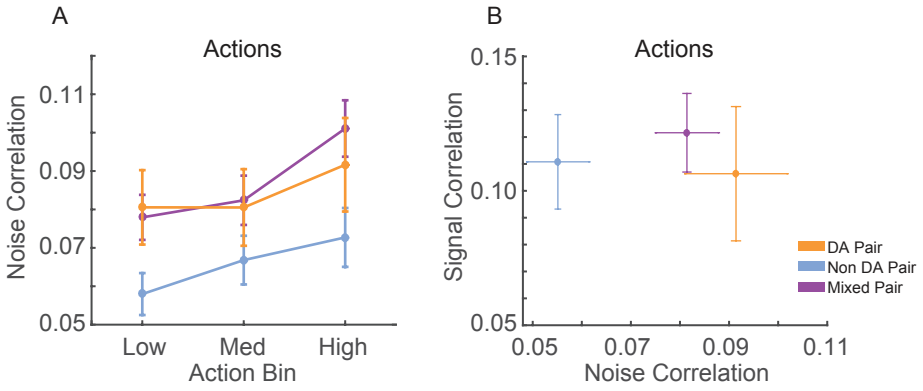

Fig. 4. (A) Mean ± SEM action evoked noise correlations between simultaneously recorded pairs of VTA neurons in RR10 sessions during actions. Noise correlations measure trial-by-trial fluctuations in pairwise activity. Data depicted as mean ± SEM noise correlation in each bin of actions. Data are plotted 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.

	Rho	p value
Dopamine – Dopamine Pair	0.571	< 0.001
Dopamine – Non-Dopamine Pair	0.569	< 0.001
Non-Dopamine – Non-Dopamine Pair	0.528	< 0.001

Table 1. Correlation between action number signal and noise correlations.

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.

Fig. 5

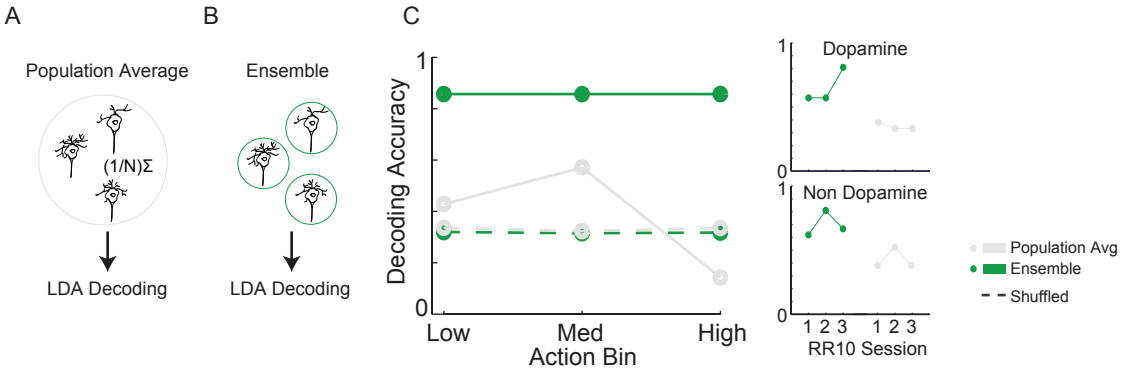

Fig. 5. Linear discriminant analysis (LDA) utilized action-evoked activity averaged across **(A)** all neurons (population-average), or **(B)** the collective activity of each neuron (ensemble). **(C)** The proportion of correctly classified actions for each decoder. Dashed lines represent decoding of 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. 6

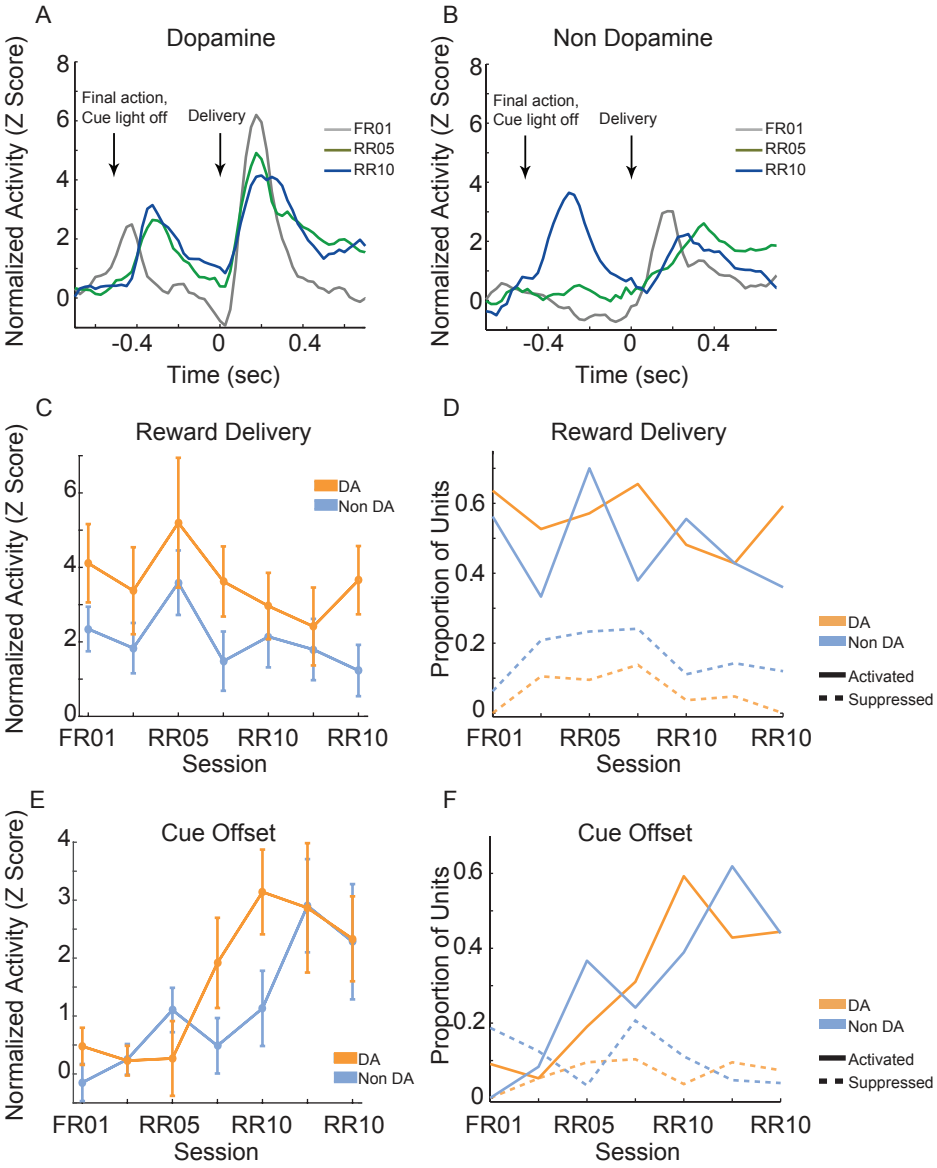

Fig. 6. Normalized mean **(A)** dopamine and **(B)** non-dopamine responses aligned to outcome delivery (time = 0). The final action in each trial (left arrow) occurred 0.5s prior to outcome delivery (right arrow). Cue light offset was simultaneous with execution of the final action, which signaled trial 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. **(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. **(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. 7

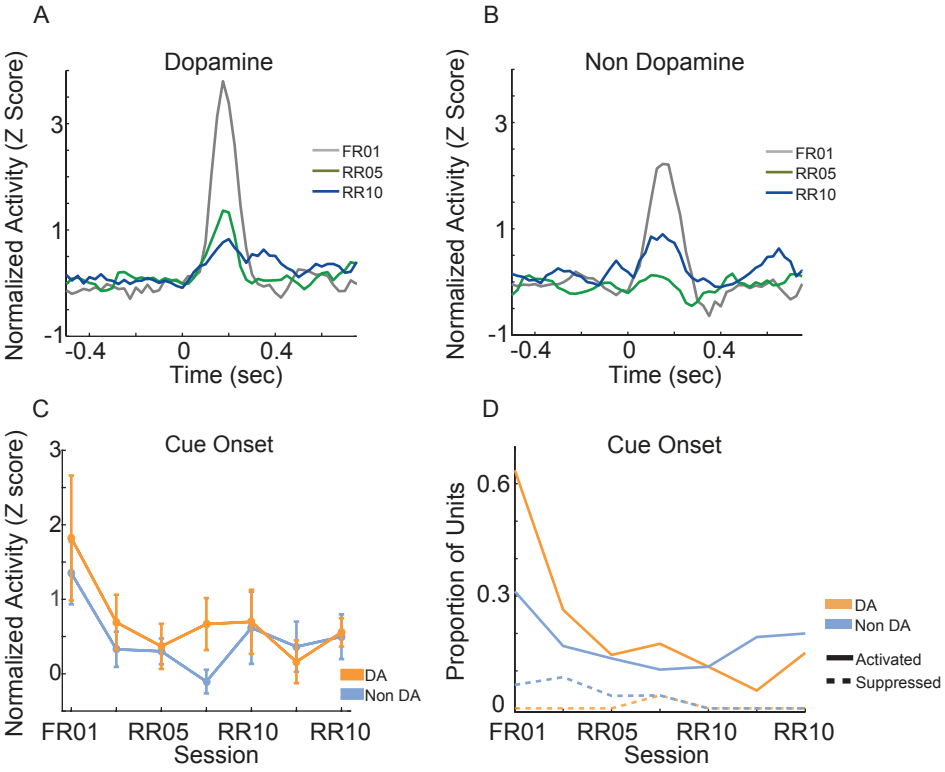

Fig. 7. Normalized mean **(A)** dopamine and **(B)** non-dopamine responses to cue onset (time = 0). Data are depicted for the FR01 session, the final RR05 session, and the final RR10 session. **(C)** Mean ± SEM responses evoked by cue onset (+0.05 - +0.3s) in all groups of neurons and sessions. Data 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. **(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.

Fig. 8

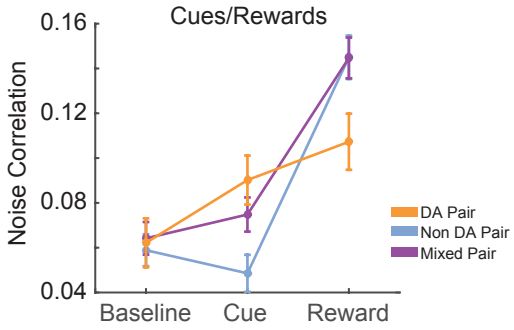

Fig. 8. Mean ± SEM noise correlation between simultaneously recorded pairs of VTA neurons in RR10 sessions, during baseline, cue, and reward delivery. Data depicted separately for all pairings of neurons. Note that reward delivery evoked the largest magnitude noise correlations in all pairings.

Fig. S1

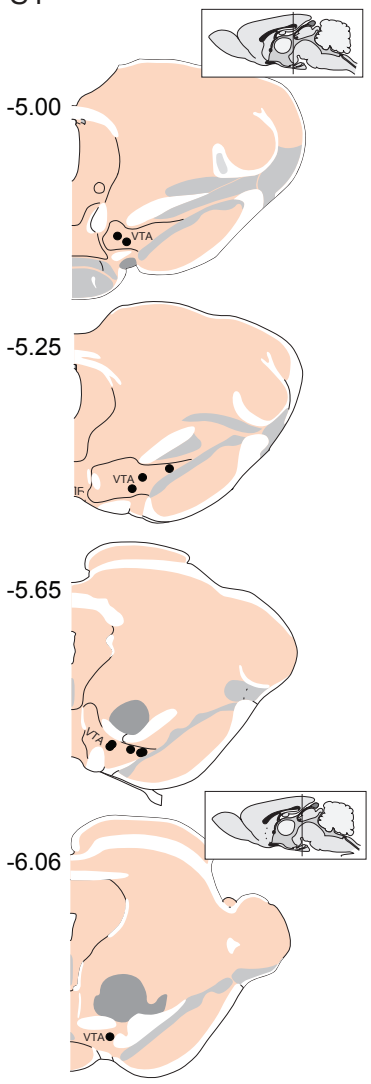

Fig. S1. Locations of recording electrodes within VTA. Each dot represents the location of a recording array. Insets show midsagittal diagram of rodent brain with vertical lines representing the approximate location of the most anterior and posterior coronal sections shown.

Fig. S2

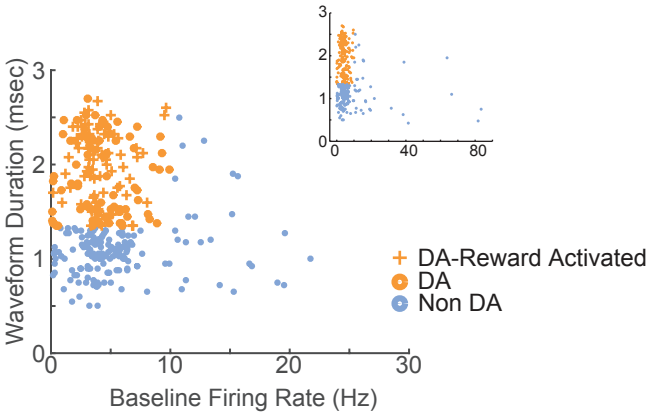

Fig. S2. Electrophysiological characterization of recorded VTA neurons. Each point represents a single neuron; data from all sessions are depicted. Neurons with wide spike widths (>1.4 ms) and low firing rates (< 10 Hz) were considered putative dopamine neurons. Crosses depict reward-responsive dopamine 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.

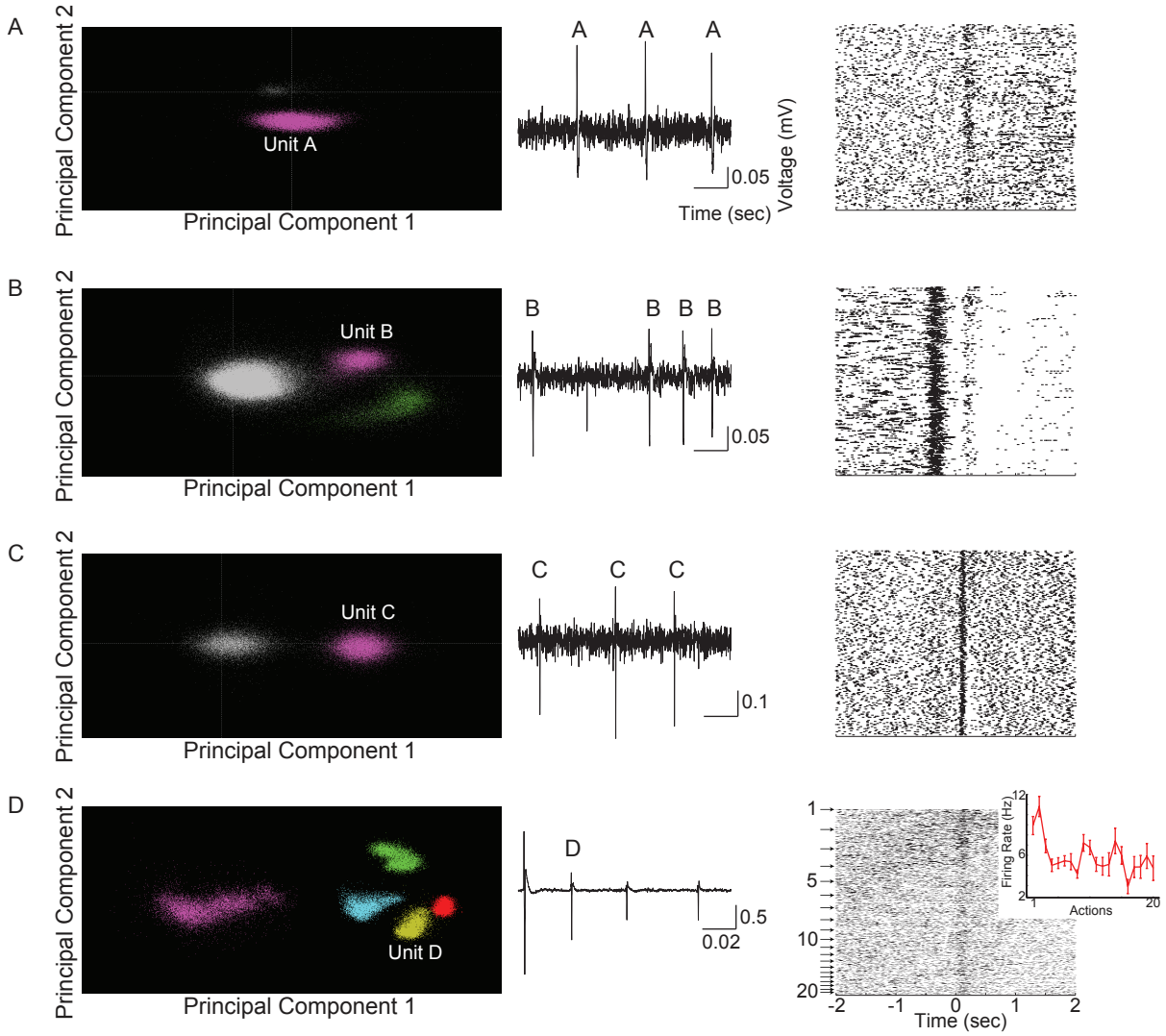

Fig. S3. Representative examples of neuronal spike sorting, raw data and neuronal responses. **(A)** Spikes from unit A, a dopamine neuron, sorted according to the first 2 principal components of all threshold crossing waveforms (left). Purple points represent spikes that were assigned to unit A, and 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. **(D)** Data from a typical dopaminergic neuron that preferred low action numbers. Spike sorting (left) plotted with similar conventions as (A). Several units were simultaneously recorded and the example voltage trace contains spikes from multiple units (middle). Only the spike corresponding to unit D (yellow) is notated. The raster (right) 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, and the second row represents the second occurrence, etc. The inset depicts the tuning curve across action numbers 1-20. Note that the neuron most strongly prefers actions 1 and 2, which is reflected in the tuning curve (inset) and the spike density in the raster.

	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 S4. The correlation between cue and reward delivery evoked responses.

Table S4. 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.