Inferring spikes from calcium imaging
in dopamine neurons

Weston Fleming[1]*, Sean Jewell[2]*, Ben Engelhard[1], Daniela M. Witten[2], Ilana B. Witten[1]

[1] Princeton Neuroscience Institute, Princeton University, Princeton NJ 08544

[2] Department of Statistics & Biostatistics, University of Washington

* these authors contributed equally

Abstract

Calcium imaging has led to discoveries about neural correlates of behavior in subcortical neurons, including dopamine (DA) neurons. However, spike inference methods have not been tested in most populations of subcortical neurons. To address this gap, we simultaneously performed calcium imaging and electrophysiology in DA neurons in brain slices, and applied a recently developed spike inference algorithm to the GCaMP fluorescence. This revealed that individual spikes can be inferred accurately in this population. Next, we inferred spikes *in vivo* from calcium imaging from these neurons during Pavlovian conditioning, as well as during navigation in virtual reality. In both cases, we quantitatively recapitulated previous *in vivo* electrophysiological observations. Our work provides a validated approach to infer spikes from calcium imaging in DA neurons, and implies that aspects of both tonic and phasic spike patterns can be recovered.

Introduction

The use of genetically encoded calcium indicators like GCaMP has made it possible to record neural activity from large populations of neurons with cell-type-specificity (1–3). These advances have spurred development of spike inference methods to estimate the underlying spikes from the imaged calcium (4–14), but to date their application has been largely limited to cortical and hippocampal neurons (15–21). Thus, it is unclear how well these algorithms perform on other cell-types, particularly those with tonic levels of baseline activity. One population of particular interest is midbrain dopamine (DA) neurons, given the shift in recent years to record from these neurons through calcium imaging (22–27), combined with the lack of established and validated methods to relate calcium and electrophysiology data in these neurons.

Two challenges arise when trying to infer spike rates from a new cell-type. The first is the acquisition of ground-truth data (simultaneous imaging and electrophysiology data from the same neurons), which is needed to validate a spike inference method in a new cell-type. The second is that spike inference algorithms rely on tuning parameters, and it is unclear how to select them, especially on a new cell-type. For example, DA neurons have distinct calcium currents (28–30) and buffering properties (31–33) compared to cortical neurons, which may affect the relationship between calcium fluorescence and spikes.

Because of these challenges, most imaging studies focused on subcortical neurons such as DA neurons have directly correlated calcium fluorescence to behavioral variables without accounting for the slow calcium dynamics (22–27,34,35). This approach risks introducing inaccuracies in characterizing the relationship between neural activity and behavior, and makes it difficult to relate findings from calcium imaging to those arising from electrophysiology.

Thus, to first address the lack of ground-truth data in the case of midbrain DA neurons, we collected simultaneous GCaMP imaging and cell-attached electrophysiology recordings *in vitro* (given the technical limitation of simultaneous *in vivo* measurements of these two signals). We then used these data to validate a recently-developed spike inference algorithm in these neurons (4).

Next, to address the issue of how to select the tuning parameters on new datasets when ground-truth spikes are not available (which is inevitably the case for subcortical *in vivo* imaging data), we made use of the fact that this spike inference algorithm depends on only two tuning parameters, both of which have a biophysical interpretation. The first corresponds to the decay rate of the calcium indicator, which we measured in our *in vitro* data. The second determines the average estimated firing rate of the recorded neurons, which we selected based on mean firing rates previously observed in midbrain DA neurons *in vivo* (36–39), and which can be measured by phototagging a new cell type of interest.

Using this straightforward approach to select the tuning parameters, we were able to quantitatively recapitulate multiple published properties of DA spiking from *in vivo* imaging data. During delivery of unexpected reward, inferred firing rates are of the appropriate magnitude and duration when compared to *in vivo* electrophysiology data. During omission of expected reward, pauses in firing in response to reward omission can be detected. Underscoring the utility of spike inference, inferred spikes, but not the calcium fluorescence itself, recapitulates the previously reported relationship between reward

delivery and reward omission. In addition to these fast response properties, slow upward and downward ramps in inferred tonic firing rate as animals approach rewards can also be detected. On the other hand, inference does not capture certain features of previously published data, including the reported variability in tonic firing rates across DA neurons. Thus, we describe strengths as well as limitations of a simple approach for inferring spikes in DA neurons. We expect this approach to generalize to other subpopulations of genetically identified subcortical neurons.

Results

Validation of spike inference using simultaneous calcium imaging and electrophysiology *in vitro*

To determine the efficacy of spike inference in ventral tegmental area (VTA) DA neurons, we generated a ground-truth dataset by performing simultaneous calcium imaging and cell-attached electrophysiological recordings of spontaneously active, burst-firing DA neurons expressing either GCaMP6f or GCaMP6m (**Fig. 1a, Supp. fig. 2**; GCaMP6f at 30°C, n=12 recordings from 9 total cells; GCaMP6f at 37°C, n=9 recordings from 5 total cells; GCaMP6m at 37°C, n=19 recordings from 9 total cells). We performed these simultaneous measurements *in vitro*, given that combining these approaches *in vivo* in the VTA, which is located deep in the brain, is not currently feasible. In addition to the data presented in **Fig. 1**, we also recorded from pacemaker cells which had faster GCaMP decay kinetics (**Supp. fig. 1**).

To estimate spike times on the basis of calcium fluorescence, and to compare to the corresponding ground-truth spike measurements, we leveraged a generative model of the calcium dynamics resulting from spikes (8,14) (**Fig. 1b**). In this model, the fluorescence trace y_t at the t th time step is a noisy observation of the underlying calcium concentration c_t, which is assumed to decay with rate $0 < \gamma < 1$ unless there is a spike. A spike causes an instantaneous increase in the calcium concentration (**Fig. 1b**). As previously discussed (5), we fit the model by minimizing the difference between the observed fluorescence and modeled calcium, while applying an L0 regularization penalty on the total number of inferred spikes. A recently-developed dynamic programming approach is used to efficiently solve this non-convex optimization problem (**Fig. 1b;** also see Methods).

This spike inference algorithm depends on only two parameters, which are both easily interpretable: i) the calcium decay rate γ, and ii) the tuning parameter λ from the regularization, which controls the total number of inferred spikes in a recording - in other words, the average inferred firing rate. The decay rate γ was selected based on fitting an exponential curve to the GCaMP fluorescence in segments without spikes (**Fig. 1c-e**; see Methods for curve fitting).

To assess the agreement between the inferred spikes and the true spikes, we calculated the van Rossum distance metric for the inferred and true spikes for each neuron across a range of values of the tuning parameter λ (**Fig. 1f**; see Methods) (40,41). This distance was smallest when the inferred average firing rate (computed as the number of inferred spikes divided by the length of the recording) equaled the true average firing rate, indicating that the true firing rate provides the optimal choice of λ. Based on this observation, to select λ for each neuron, we matched the inferred firing rate to the true firing rate of each neuron for our *in vitro* data.
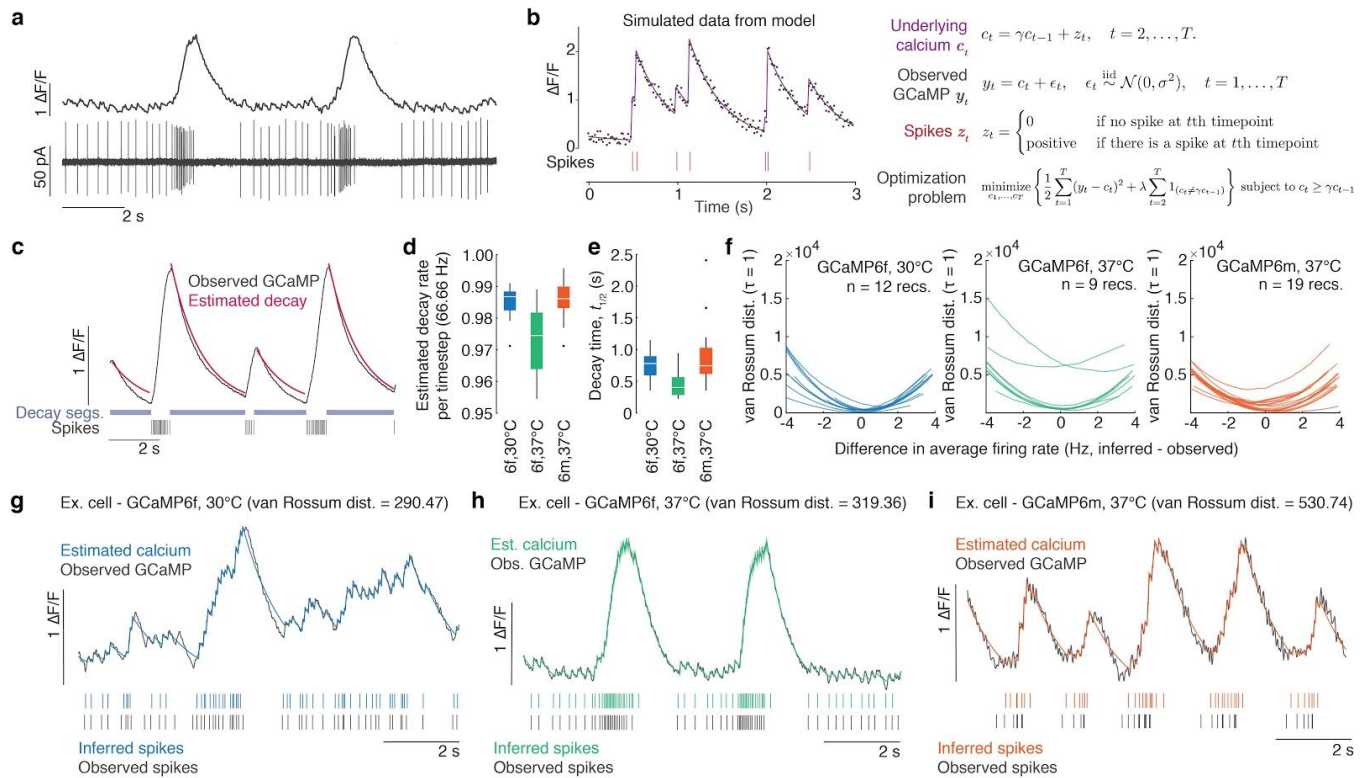
3

Figure 1. Spike inference applied to VTA DA neurons *in vitro*. a. Example data trace of simultaneous GCaMP6f (top) and raw cell-attached electrophysiology trace (bottom). **b.** Generative model of calcium dynamics. Simulated data (left) from this model (right) shows the underlying calcium c_t (purple line) decays at rate γ until at time s there is a spike, so that $z_s > 0$ (red vertical lines below). GCaMP observations y_t (black dots) are noisy realizations of the underlying calcium concentration. **c.** Example of decay estimation using multiple spike-free segments ("Decay segs.", light blue). Observed GCaMP (black) decays at the same rate between spikes (black vertical lines below). The rate of exponential decay, γ, was estimated by fitting an exponential decay model to the observed GCaMP in the spike-free segments. **d.** Estimated decay rate per timestep for GCaMP6f at 30°C (median = 0.987, Q1 = 0.982, Q3 = 0.988), GCaMP6f at 37°C (median = 0.975, Q1 = 0.964, Q3 = 0.982), and GCaMP6m at 37°C (median = 0.986, Q1 = 0.983, Q3 = 0.990). **e.** Decay time (half life, $t_{1/2}$) for GCaMP6f at 30°C (median = 0.780 s, Q1 = 0.593 s, Q3 = 0.894 s), GCaMP6f at 37°C (median = 0.403 s, Q1 = 0.283 s, Q3 = 0.565 s), and GCaMP6m at 37°C (median = 0.75 s, Q1 = 0.61 s, Q3 = 1.03 s). **f.** van Rossum distance for each experimental condition and for each neuron, using the median decay rate across all recordings for that type of GCaMP and temperature, as a function of the difference between the average inferred firing rate and average observed firing rate. The distance is minimized when the average firing rate between the inferred and observed spikes is similar. **g.** Example of correspondence between observed GCaMP (black line, top) and estimated calcium (blue line, top), and observed spikes (black vertical dashes, bottom) and inferred spikes (blue vertical dashes, bottom) in a single cell expressing GCaMP6f measured at 30°C. **h.** Example of correspondence between observed GCaMP (black line, top) and estimated calcium (green line, top), and observed spikes (black vertical dashes, bottom) and inferred spikes (green vertical dashes, bottom) in a single cell expressing GCaMP6f measured at 37°C. **i.** Example of correspondence between observed GCaMP (black line, top) and estimated calcium (orange line, top), and observed spikes (black vertical dashes, bottom) and inferred spikes (orange vertical dashes, bottom) in a single cell expressing GCaMP6m measured at 37°C.

Using this approach, inferred spike times tended to match true spike times in the *in vitro* dataset, during periods of tonic as well as burst activity (**Fig. 1g-i, Supp. figs. 3-5**). However, some patterns of activity were not well-recovered using this approach (**Supp. figs. 3-5**). There were occasional inaccuracies when estimating timing of the final spikes of a burst, or the first tonic spikes following a burst, when the GCaMP signal is still decaying (as seen in **Fig. 1i**). Additionally, inferred spikes cannot recreate the timing of recorded spikes at frequencies above double the sampling frequency (67 Hz; see Methods).

In summary, our spike inference method appears well-suited to estimate both tonic and phasic activity in DA neurons, with some limitations on estimating exact spike placement during and immediately after

4

burst activity, and with a limit to the maximum inferred firing rate. Given these successes, as well as limitations, in this *in vitro* data, we next asked whether our spike inference method could reproduce previously reported patterns of phasic and tonic activity in identified DA neurons *in vivo*.

Application of spike inference to *in vivo* imaging data during the presentation of unexpected reward

While unexpected reward is known to generate widespread phasic responses in DA neurons based on electrophysiology data (39,42–45), the extent to which spike inference can quantitatively recapitulate the magnitude and duration of the unexpected reward response has not been determined. Thus, we applied this spike inference algorithm to recently published *in vivo* two-photon calcium imaging data from VTA DA neurons during unexpected reward delivery (22) (**Fig. 2a**,**b**), and compared to the firing rates obtained from *in vivo* electrophysiology recordings from VTA DA neurons in a previously published dataset from other mice performing a different conditioning task (39). For these spike estimates, the decay γ was set to the median decay rate from the *in vitro* measurements at 37°C (**Fig. 1d**), whereas the tuning parameter λ was set to correspond to a mean firing rate of 6 Hz, consistent with previous *in vivo* measurements of average DA neuron firing rates (36,38,39,46).

We first compared the magnitude of the phasic response to unexpected reward in inferred spikes with spikes previously measured via electrophysiology in DA neurons in another group of mice (electrophysiology data from (39)). The unexpected reward response was defined as the baseline subtracted mean firing rate over a 600 ms window following reward onset (baseline calculated over a 1s period before reward) (39). We found that inferred spikes showed a similar increase in firing rate following reward compared to the Eshel et al. electrophysiology data (inferred spikes median = 9.7 Hz, Q1 = 5.9 Hz, Q3 = 12.0 Hz; electrophysiology spikes median = 9.5 Hz, Q1 = 6.3 Hz, Q3 = 12.0 Hz ; **Fig. 2e-g**). The peak amplitude of the unexpected reward response based on inferred spikes was also similar to those in the electrophysiology dataset (inferred spikes median = 28.7 Hz, Q1 = 22.7 Hz, Q3 = 38.3 Hz; electrophysiology median = 30.5 Hz, Q1 = 22.6 Hz, Q3 = 40.9 Hz; **Fig. 2h**).
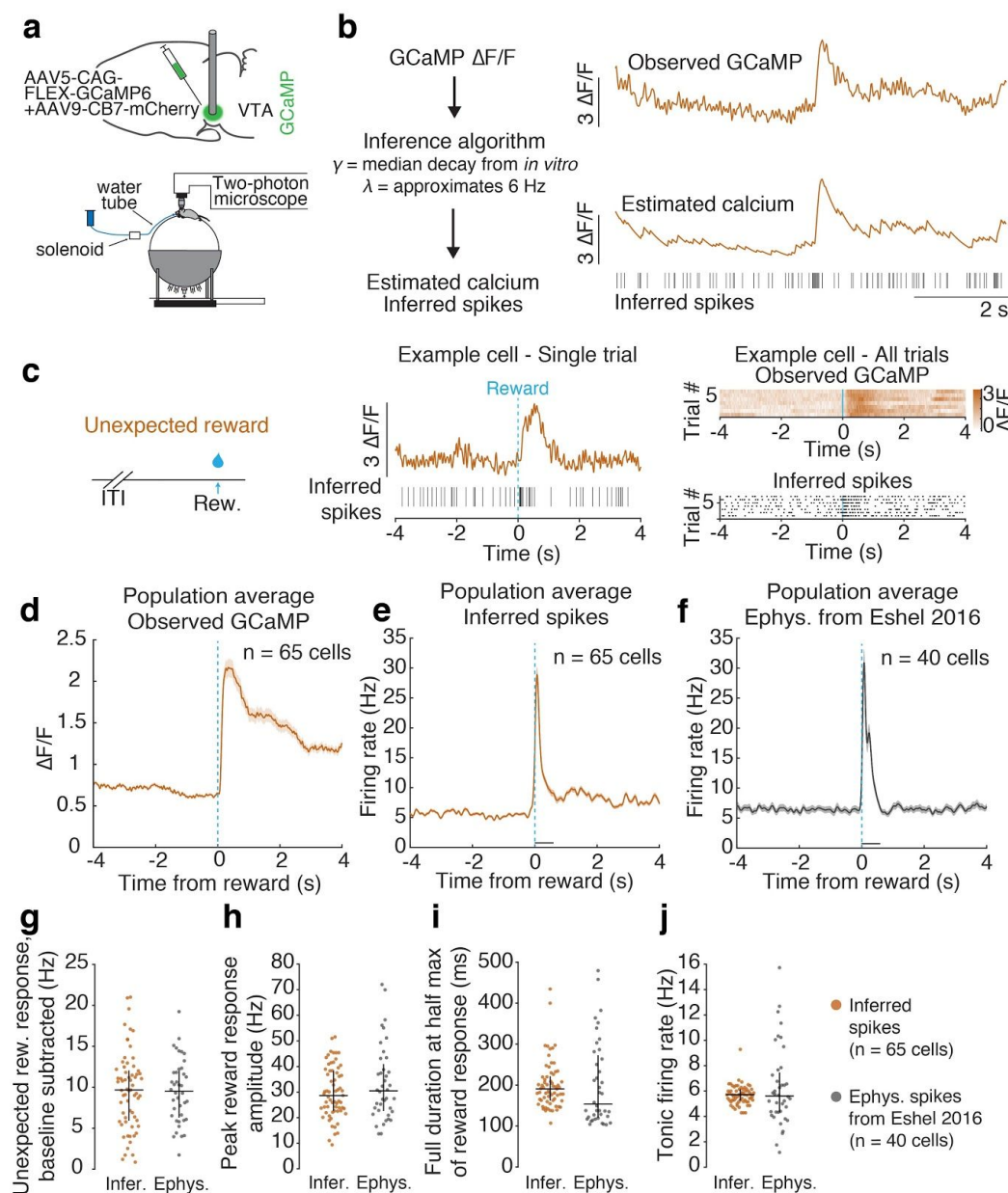
Figure 2. Spike inference applied to *in vivo* calcium imaging in VTA DA during presentation of unexpected reward generates firing rate modulations comparable to those previously reported in other animals via electrophysiology. a. Top: Schematic of the surgical strategy, where GCaMP6 is expressed in VTA DA neurons and a GRIN lens is implanted for imaging. Bottom: Schematic of recording setup, where the mouse is headfixed and VTA DA neurons are recorded via 2-photon calcium imaging. **b.** Schematic of spike estimation approach as applied to the *in vivo* data, where the observed ΔF/F is used to generate estimated calcium and inferred spikes using the decay parameter γ from *in vitro* experiment and a λ selected to target a 6 Hz average estimated firing rate. **c.** Left: Schematic showing unexpected reward is delivered after random inter-trial intervals. Middle: Example of an unexpected reward trial from a single cell, showing observed GCaMP (orange) and inferred spikes (black vertical lines). Right: All unexpected reward trials from an example cell, showing observed GCaMP (top) and inferred spikes (bottom). **d.** Mean observed GCaMP from population around presentation of unexpected reward (n = 65 cells). **e.** Mean population firing rate from inferred spikes from population around presentation of unexpected reward (n = 65 cells). **f.** Mean population firing rate from spikes recorded via electrophysiology (from Eshel et al., 2016; n = 40 cells) around presentation of unexpected reward. **g-j.** Comparison of inferred and electrophysiology spikes from Eshel et al., 2016. **g.** Unexpected reward response, where reward response is the mean firing rate over the first 600 ms following reward presentation, baseline subtracted using the mean firing rate over a 1 s period before reward presentation (inferred spikes over baseline median = 9.7 Hz, Q1 = 5.9 Hz, Q3 = 12.0 Hz; electrophysiology spikes over baseline median = 9.5 Hz, Q1 = 6.3 Hz, Q3 = 12.0 Hz). **h.** Peak reward response amplitude in inferred and electrophysiology spikes, where peak is maximum value of PSTH in the first 600 ms period following reward presentation (inferred spikes median = 28.7 Hz, Q1 = 22.7 Hz, Q3 = 38.3 Hz; electrophysiology median = 30.5 Hz, Q1 = 22.6 Hz, Q3 = 40.9 Hz). **i.** Full duration at half max of reward response peak in inferred and electrophysiology spikes (inferred spikes median full duration at half max= 190.6 ms, Q1 = 162.2 ms, Q3 = 221.4 ms; electrophysiology median full duration at half max = 153.5 ms, Q1 = 120.4 ms, Q3 = 272.5 ms). **j.** Mean tonic firing rates prior to presentations of unexpected reward, where tonic firing rates are calculated as the mean firing rate over a 1 s period prior to reward presentation (inferred spikes tonic firing rate median = 5.7 Hz, Q1 = 5.3 Hz, Q3 = 6.1 Hz; electrophysiology tonic firing rate median = 5.6 Hz, Q1 = 4.3 Hz, Q3 = 7.5 Hz). Vertical bars are interquartile range (Q1 and Q3).

In contrast to the prolonged ∆F/F modulation to unexpected reward (**Fig. 2c,d**), inferred spikes displayed a subsecond, phasic modulation (**Fig. 2c,e**). In fact, the duration of the unexpected reward response was similar between inferred spikes and electrophysiology spikes (inferred spikes median full duration at half max = 190.6 ms, Q1 = 162.2 ms, Q3 = 221.4 ms; electrophysiology median full duration at half max= 153.5 ms, Q1 = 120.4 ms, Q3 = 272.5 ms; **Fig. 2i**).

However, the electrophysiology population average firing rate displayed a biphasic response which was not evident in the inferred spikes (**Fig. 2f**). It is unclear if this difference reflects a limitation of spike inference to detect the time course of a biphasic burst, or is instead due to differences in the experimental design between the calcium imaging and the electrophysiology datasets. In the calcium imaging experiment, reward size was constant (reward size 8 µL), whereas for the electrophysiology dataset, reward size varied (ranging from 0.2 to 20 µL). Thus, it is possible that the second bump of the reward response is due to receiving a larger-than-expected reward (10 µL reward, the closest volume to our experiment) (43).

The median tonic firing rate of the inferred spike dataset was similar to that of the electrophysiology dataset (inferred spike tonic firing rate median = 5.7 Hz, Q1 = 5.3 Hz, Q3 = 6.1 Hz; electrophysiology tonic firing rate median = 5.6 Hz, Q1 = 4.3 Hz, Q3 = 7.5 Hz; tonic firing rate calculated as the mean firing rate over a 1 s window prior to reward; **Fig. 2g**). This implies that spike inference correctly assigns the total number of estimated spikes (which is dictated by λ) to phasic versus tonic firing. However, across neurons, the inferred spikes displayed less variability than the electrophysiology data in the tonic firing rates, presumably as a result of applying the same tuning parameter λ for all neurons, when in reality DA neurons exhibit a range of average *in vivo* firing rates between 1-10 Hz (36–38).

Taken together, our spike inference applied to *in vivo* DA imaging data closely recapitulates the median magnitude and duration of phasic activity to unexpected reward, as previously measured via electrophysiology, as well as the median level of tonic activity. However, the approach did not reproduce cross-neuron variability in tonic firing rates, and there may also be limitations to detecting a biphasic reward response.

Application of spike inference to *in vivo* imaging data during Pavlovian conditioning

We next inferred spikes from VTA DA calcium imaging data taken during Pavlovian conditioning (**Fig. 3a**; 2 s tone presentation followed immediately by reward delivery). This is one of the most commonly utilized behavioral paradigms to study VTA DA neuron activity (43,44,47–49). Our goal was to determine whether spike inference recapitulates the previously reported features of DA firing activity during Pavlovian conditioning, including the transient, phasic response to a cue that predicts reward (50), the pause resulting from the omission of an expected reward (39,44,51), and finally, the previously established relationships across neurons in response to expected, unexpected, and omitted rewards (39,51). Because we did not have electrophysiology data from the same Pavlovian task as our imaging data, we were not able to make direct comparisons to electrophysiology for these data.

Inferred spike rates displayed a transient phasic burst to the cue onset (median full duration at half max = 190.4 ms, Q1 = 170.7 ms, Q3 = 206.0 ms), consistent with previous electrophysiology reports (**Fig. 3a,c,d, Supp. fig. 6**) (50). This fast, phasic response contrasted with GCaMP fluorescence, which remained elevated throughout the 2 s duration of the cue that predicted reward (**Fig. 3a,b**).
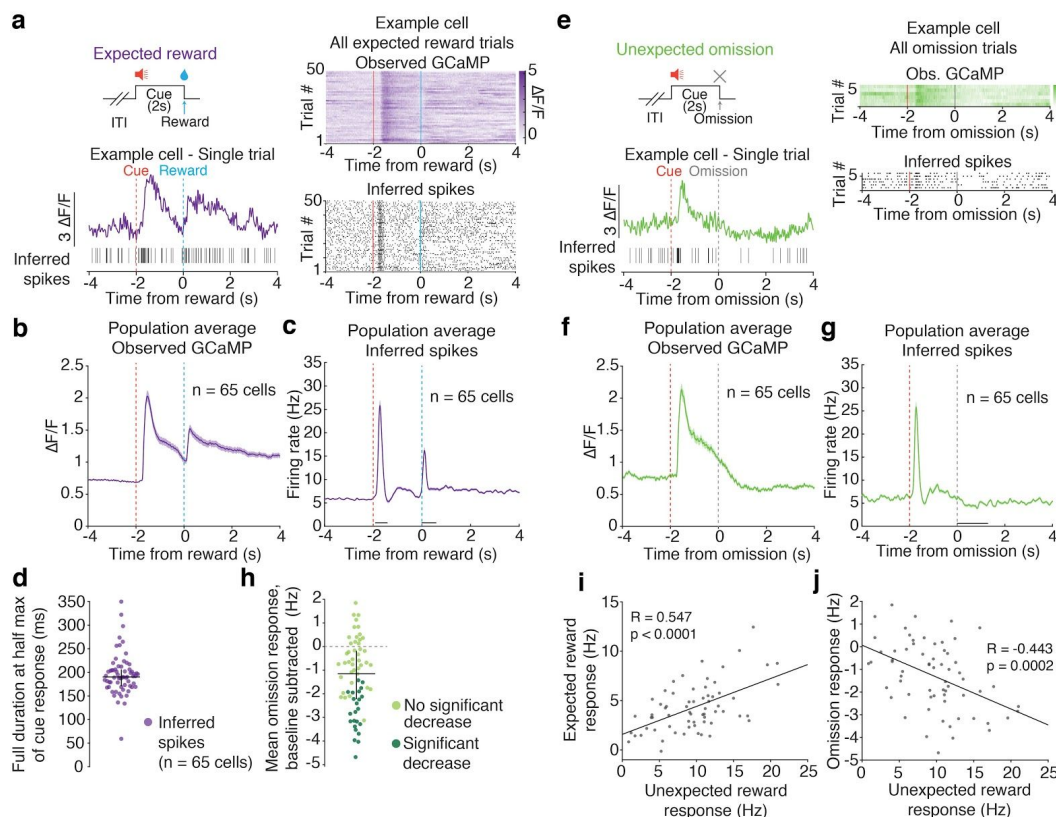
Figure 3. Spike inference applied to Pavlovian conditioning *in vivo* imaging data recapitulates transient cue responses, pauses following reward omission, and previously reported relationships between expected and unexpected reward responses. a. Top left: Schematic showing expected reward is delivered after a 2 s cue presentation. Bottom left: Example expected reward trial from a single cell, showing observed GCaMP (purple) and inferred spikes (black vertical lines). Right: All expected reward trials from the example cell, showing observed GCaMP (top) and inferred spikes (bottom). **b.** Mean observed GCaMP from population around presentation of expected reward. **c.** Mean inferred spikes from population around presentation of expected reward. **d.** Full duration at half max of cue response in inferred spikes (median = 190.4 ms, Q1 = 170.7 ms, Q3 = 206.0 ms). Vertical bars are interquartile range (Q1 and Q3). **e.** Top left: Schematic showing unexpected omission of reward, where reward is omitted after a 2 s cue presentation that previously predicted reward. Bottom left: Example unexpected omission trial from a single cell, showing observed GCaMP (green) and inferred spikes (black vertical lines). Right: All unexpected omission trials from example cell, showing observed GCaMP (top) and inferred spikes (bottom). **f.** Mean population observed GCaMP around unexpected omission of reward. **g.** Mean population firing rate from inferred spikes around unexpected omission of reward. Shaded areas are SEM. **h.** Mean omission response of inferred spikes, where omission response is mean firing rate over 1,300 ms following onset of reward omission, baseline subtracted using the mean firing rate over 1 s period before trial start (median = -1.2 Hz, Q1 = -2.2 Hz, Q3 = -0.2 Hz). Neurons that exhibited a significant decrease in firing following reward omission (23/65 neurons; 35.4% of population) are darker green. Vertical bars are interquartile range (Q1 and Q3). **i.** Scatterplot of expected reward response versus unexpected reward response, using inferred firing rate of each neuron, recapitulates correlations in Eshel et al. 2016. **j.** Scatterplot of omission response versus unexpected reward response, using inferred firing rate of each neuron, recapitulates correlations in Eshel et al. 2016. Responses in **i** and **j** are baseline subtracted.

We next examined whether spike inference was able to recover the decrease in tonic firing that occurs following reward omission (sometimes called "pause", **Fig. 3e**) (39,44,51). Inferred spike rates in 35.4% of neurons (23/65) displayed a significant decrease in response to reward omission (**Fig. 3e** for example neuron; **Fig. 3g,h, Supp. fig. 7** for population summaries; median omission response = -1.2 Hz, Q1 = -2.2 Hz, Q3 = -0.2 Hz; n = 65 neurons). Though a direct comparison cannot be made, electrophysiology data from a Pavlovian experiment from Eshel et al. (2016) following odor-reward pairings revealed that 41.9% of neurons (13/31) exhibited significant decreases in firing rate in response to reward omission (median omission response of -1.5 Hz; data not shown; 7 trials per cell were analyzed in order to match the available trial number in the imaging dataset). Thus, spike inference is able to recover pauses to reward omission in a similar proportion to that observed with electrophysiology.

Finally, it has previously been reported that across DA neurons, responses to unexpected reward are positively correlated with responses to expected reward, while unexpected reward responses are negatively correlated with responses to reward omission (39,51). We recovered both of these relationships on the basis of inferred spikes (**Fig. 3i,j**). These relationships were not reliably recovered directly from GCaMP fluorescence without applying spike inference (**Supp. fig. 8**). This is because, as mentioned above, the fluorescence in response to the cue has not returned to baseline at the time of the reward or omission (**Fig. 3b,f**). Therefore, accounting for this baseline offset through spike inference avoids inaccuracies in response estimates at the time of reward.

Application of spike inference to _in vivo_ DA calcium imaging data collected during a task in virtual reality

In addition to the phasic modulation of DA activity in response to reward or reward-predicting cues, there are an increasing number of reports suggesting that activity may "ramp" as animals approach a reward in space (22,52–56). Most of these studies were based on calcium imaging (22,56), or on measures of downstream DA release (52,53,55), but more recently evidence has emerged from _in vivo_ electrophysiology of ramps in the tonic firing of VTA DA neurons during reward approach (54).

Thus, a key question is whether spike inference recovers the recently observed upward and downward ramps in tonic DA neuron firing. To our knowledge, spike inference has not previously been validated in the context of tonic firing rate modulations in any cell type.

To this end, we applied our spike inference procedure to a recent virtual reality T-maze study with 2-photon imaging of DA neurons (**Fig. 4a**) (22). Again, for these spike estimates, the decay γ was set to the median decay rate from the _in vitro_ measurements at 37°C (γ = 0.970 for GCaMP6f at 37°C; γ = 0.984 for GCaMP6m at 37°C at 60 Hz sampling frequency), and the tuning parameter λ was set to correspond to a mean firing rate of 6 Hz. Ramps in both GCaMP fluorescence and inferred spikes were often evident on single trials (**Fig. 4b-i**). Example neurons with prominent downward or upward ramps in GCaMP fluorescence showed corresponding ramps in inferred firing rates consistent with modulation of tonic activity (where tonic activity is defined as a firing rate between 1-10 Hz; **Fig. 4f-k**). To quantify how ramps in GCaMP fluorescence translated to changes in inferred firing rate, for each neuron we calculated the linear change in both fluorescence and inferred firing rate over the course of the maze (**Fig. 4l**). Across the population of 303 neurons, there was a strong correlation in the extent (and direction) of change in the fluorescence signal and the inferred spike rates across the maze (correlation coefficient R = 0.8328; p = 0.0077; **Fig. 4l**), with a change in ∆F/F of 0.25 from the beginning to the end of the maze associated with a change in firing rate of ~1.9 Hz (**Fig. 4l**). Inferred firing rates increased by a maximum of 6.8 Hz over the course of the maze in the case of upward ramps, and decreased by a maximum of 4.5 Hz for downward ramps (**Fig. 4l**).

These changes in tonic firing rates are consistent with a recent electrophysiology study that showed that many VTA DA neurons exhibit upward or downward ramps in tonic firing rate during reward approach in virtual reality (54). Thus, we show that spike inference can produce the subtle, slow-timescale changes in tonic firing rates that have been reported in DA neurons as animals approach rewards.
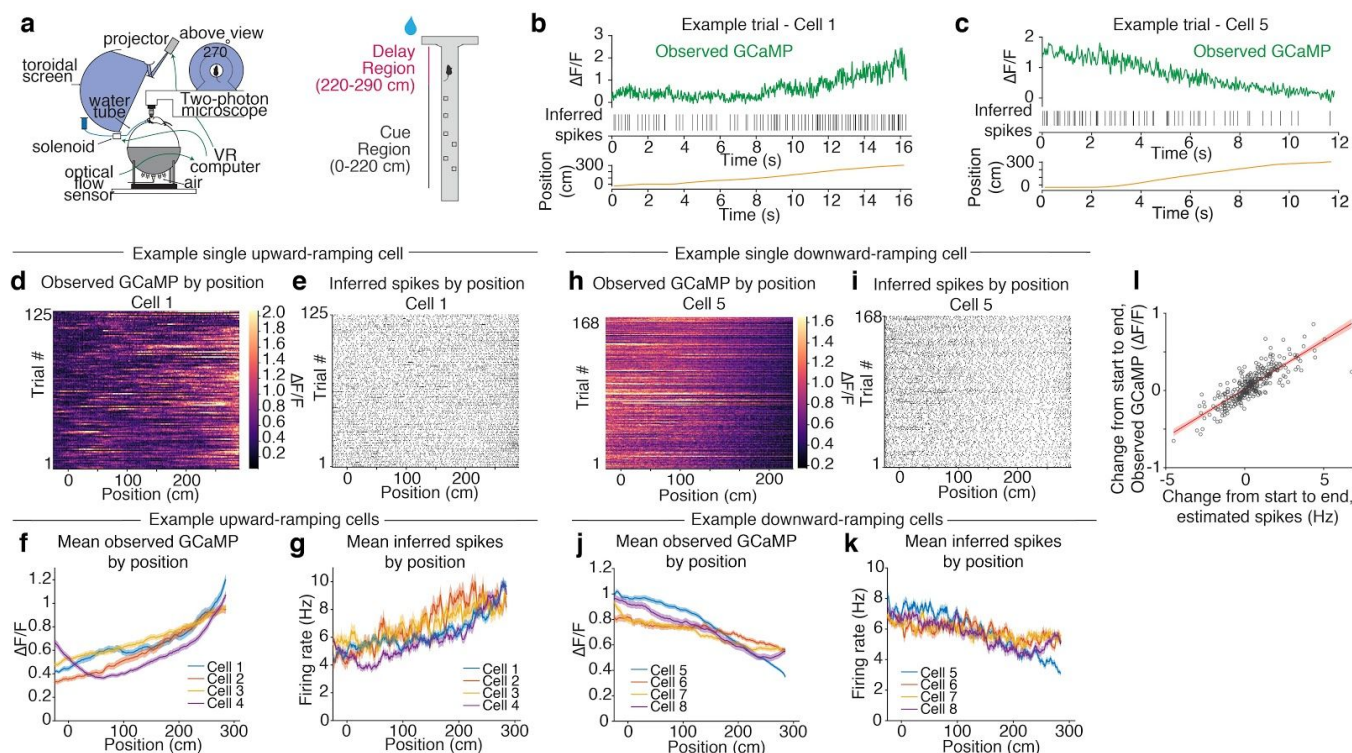
9

Figure 4. Upward and downward ramps in inferred spikes during reward approach in a virtual reality environment, in agreement with recent reports from electrophysiology. a. Left: Schematic of neural recording and behavioral setup. The mouse navigates a virtual reality (VR) environment and 2p microscope records neural data. Right: Schematic of VR T-maze paradigm. The mouse navigates a linear maze and at the end must turn to the side that featured more cues in the Cue Region to receive reward. **b.** Example of single trial data from an upward-ramping cell. Both observed GCaMP (green) and inferred spikes ramp upward over time as the mouse moves down the maze (position trace, yellow). **c.** Example of a single trial from a downward-ramping cell. Both observed GCaMP (green) and inferred spikes ramp downward over time as the mouse moves down the maze (position trace, yellow). **d**. All trials from the example upward-ramping cell in (**b**) of observed GCaMP by position. **e**. All trials from the example upward-ramping cell in (**b**) showing inferred spikes by position. **f**. Mean observed GCaMP by position for example upward-ramping cells. **g**. Mean inferred firing rate by position for example upward-ramping cells. **h**. All trials from the example downward-ramping cell in (**c**) of observed GCaMP by position. Heatmap color scales constrained to data between 1st and 99th percentile. **i.** All trials from the example downward-ramping cell in (**c**) showing inferred spikes by position. **j.** Mean observed GCaMP by position for example downward-ramping cells. **k**. Mean inferred firing rate by position for example downward-ramping cells. Shaded areas are SEM. **l**. Scatterplot showing how change in observed GCaMP from beginning to end of maze for each neuron relates to change in inferred firing rate for that neuron. Each data point represents a single neuron and its mean change in observed GCaMP and inferred firing rate. Red line is linear least-squares fit; shaded region is 95% confidence of the fit.

Discussion

Here, we establish and validate procedures to infer spikes from calcium imaging in DA neurons, making use of a recently developed spike inference algorithm (4) that depends on only two tuning parameters, which can be selected based on knowledge of basic biophysical properties of the cell-type of interest. We first validate this approach for spike inference on a new ground-truth dataset that we have collected, which consists of simultaneous calcium imaging and electrophysiology recordings in DA neurons *in vitro*. We then apply this same approach to recently published *in vivo* DA calcium imaging data during Pavlovian conditioning and virtual reality-based navigation. We find that spike inference successfully recapitulates various features of spiking in DA neurons that were previously reported

based on electrophysiological measurements, including the magnitude of the phasic unexpected reward response, average tonic firing rates, the relationship between expected and unexpected reward responses during Pavlovian conditioning, pauses during the omission of expected reward, and the recently reported magnitude of upward and downward ramping tonic activity during VR-based navigation.

Limitations of spike inference in VTA DA neurons

Despite these successes, we also noted limitations to DA spike inference *in vivo* and *in vitro*. *In vitro*, there were some missed or mis-timed spikes. *In vivo*, we did not recapitulate the observed variability in tonic firing rates, nor the smaller second peak in response to unexpected reward that has been previously reported. Since we did not have a true ground-truth dataset for the *in vivo* data, it is impossible to know whether the subtle differences we noted relative to previously published work should be attributed to shortcomings of spike inference, differences in behavioral paradigms in our imaging study versus the previous electrophysiology study, or some other factor.

One source of discrepancy between spike inference and true spikes is that we are limited to inferring spikes at double the imaging sampling frequency (60 Hz for *in vivo* data, 67 Hz for *in vitro* data; see Methods). While average DA neuron phasic firing rates in response to reward are typically <60 Hz (**Fig. 2h**), firing rates can momentarily exceed this value. This may be less relevant for our *in vivo* analyses, which involved averaging firing rates across trials for a given neuron.

Another source of discrepancy could be caused by limitations in the signal-to-noise of the imaging data. While single-spike events are often visually detectable in the fluorescence trace by eye in the *in vitro* data, *in vivo* recordings tend to have lower signal-to-noise ratios that may preclude reliable detection of single-spike events (57). Since we only considered trial-averaged properties, and since phasic DA neuron responses even to very small rewards consist of two or more spikes (39,58), these issues may be less relevant in our data.

Finally, the fact that we did not tune the two free parameters for the spike inference algorithm per neuron in the *in vivo* data could also produce discrepancies. This might be particularly relevant for λ , given that we assumed a target firing rate of 6 Hz, whereas in reality VTA DA neuron average firing rates may range between 1-10 Hz (36–39,46). Thus, examining robustness of any conclusions across a range of λ may be valuable.

Relationship to recent validation of spike inference in cortical neurons

Until recently, systematic comparisons of neural correlates derived from electrophysiology data and imaging data were lacking. Recent work has begun to address this gap, by comparing neural tuning properties derived from calcium imaging versus electrophysiology data in cortex (59,60).

Siegle et al.(59) reported discrepancies in visual cortex between neural correlates based on calcium imaging versus electrophysiology, with a larger fraction of responsive neurons from electrophysiology and higher stimulus selectivity in calcium imaging, even after spike inference (using the same spike inference algorithm we apply in this paper). A forward spike-to-calcium model could not account for

11

these discrepancies, suggesting that it was not due to the efficacy of spike inference. Instead, the discrepancy seemed to be caused by a recording bias with electrophysiology towards highly active neurons, and also inaccuracies of spike sorting. This conclusion was based on the fact that by biasing the imaging data towards highly active neurons, and also by using stricter cell sorting criteria, the discrepancies could be largely eliminated. By contrast, Wei et al.(60) found that the discrepancies they observed in tuning properties in anterolateral motor cortex between imaging and electrophysiology could only be partially accounted for by spike inference, but could be accounted for by a forward spike-to-calcium model. This may suggest limitations in the application of their spike inference algorithms to their data.

In our case of VTA DA neuron data, as discussed above, we only noticed minor discrepancies between inferred spikes from *in vivo* imaging and previously reported electrophysiology data. The conclusions from the Siegle paper that discrepancies may be due primarily to recording biases and poorly isolated units in electrophysiology data may explain our success, as these issues may be less applicable for our data. In particular, our focus on a genetically identified cell-type would help eliminate recording biases, while the phototagging used for the electrophysiology data collection requires well-isolated units. Finally, the challenge of correctly selecting the tuning parameters for spike inference, which was likely an issue for Wei et al., is less likely a problem with our spike inference algorithm, which has only two free parameters that are likely to be relatively consistent within a genetically-identified population.

In summary, we present and validate a simple and successful method to infer spikes from VTA DA neurons. We expect this straightforward approach to generalize to other subpopulations of genetically identified subcortical neurons.

METHODS

Animals and surgery

All experimental procedures were conducted in accordance with the National Institutes of Health guidelines and were reviewed by the Princeton University Institutional Animal Care and Use Committee (IACUC). A total of 39 mice were used across all experiments. We used either DAT::IRES-Cre mice (n=17, The Jackson Laboratory strain 006660; extensively characterized in (61)) or mice resulting from the cross of DAT^{IREScre} mice and the GCaMP6f reporter line Ai148 mice (62) (n=22, Ai148xDAT::cre, The Jackson Laboratory strain 030328).

For the slice recording experiments, we used male and female Ai148xDAT::cre mice (n=8) and DAT::IRES-Cre mice (n=3). For the Pavlovian conditioning experiments, male and female Ai148xDAT::cre mice (n=8) were used. For the virtual reality experiments, male DAT::IRES-Cre mice (n=14) and male Ai148xDAT::cre mice (n=6) were used.

Mice were maintained on a 12-hour light on – 12-hour light off schedule. All procedures were conducted during their light off period. Mice were 2-6 months old.

For Pavlovian and virtual reality experiments, mice between 8-12 weeks underwent sterile stereotaxic

surgery under isoflurane anesthesia (3-4% for induction, 0.75-1.5% for maintenance). The skull was exposed and the periosteum removed using a delicate bone scraper (Fine Science Tools). The edges of the skin were affixed to the skull using a small amount of Vetbond (3M). We injected 800 nL of a viral combination of AAV5-CAG-FLEX-GCaMP6m-WPRE-SV40 (n=12) or AAV5-CAG-FLEX-GCaMP6f-WPRE-SV40 (n=2; UPenn Vector Core) with $1.6x10^{12}$/mL titer and AAV9-CB7-CI-mCherry-WPRE-rBG (UPenn Vector Core) with $2.3x10^{12}$/mL titer. Two such injections were made at stereotactic coordinates: 0.5 mm lateral, 2.6 or 3.8 mm posterior, 4.7 mm in depth (relative to bregma). After the injections, we implanted a 0.6 mm diameter GRIN lens (GLP-0673, Inscopix or NEM-060-25-10-920-S-1.5p, GrinTech) in the VTA using a 3D printed custom lens holder. After implantation, a small amount of diluted metabond cement (Parkell) was applied to affix the lens to the skull using a 1 ml syringe and 18 gauge needle. After 20 minutes, the lens holder grip on the lens was loosened while the lens was observed through the microscope used for surgery to ascertain there was no movement of the lens. Then, a previously described titanium headplate was positioned over the skull using a custom tool and aligned parallel to the stereotax using an angle meter (2). The headplate was then affixed to the skull using metabond. A titanium ring was then glued to the headplate using dental cement blackened with carbon.

For ex vivo recordings of GCaMP6m, mice between 6-7 weeks underwent stereotaxic surgery under isoflurane anesthesia (3-4% for induction, .75-1.5% for maintenance). The skull was exposed and the periosteum removed using a delicate bone scraper (Fine Science Tools). The edges of the skin were affixed to the skull using a small amount of Vetbond (3M). We injected 500 nL of AAV5-CAG-FLEX-GCaMP6m-WPRE-SV40 (n=3) with $1.6x10^{12}$/mL titer. Injections were made at stereotactic coordinates: 0.5 mm lateral, 3.2 mm posterior, 4.75 mm in depth (relative to bregma).

Virtual-reality behavioural system

To enable a navigation-based decision-making task under head-fixed conditions, we used a virtual-reality system similar to that previously described (22,63,64) (**Fig. 4a**). Mice were held head-fixed under a two-photon microscope using two custom headplate holders and ran on an air-supported, Styrofoam spherical treadmill that was 8 inches (20.3 cm) in diameter. We found that the precise alignment of the mouse on top of the sphere was important for maintaining good behavioural performance; therefore, we used a custom alignment tool for this purpose. The sphere's movement were measured using an optical flow sensor (ADNS3080) located underneath the sphere and controlled by an Arduino Due; this information was sent to the virtual-reality computer, running the ViRMEn software engine (65) (https://pni.princeton.edu/pni-software-tools/virmen) under MATLAB, which displayed and controlled the virtual-reality environment.

The display was projected using a DLP projector (Mitsubishi HD4000) running at 85 Hz onto a custom toroidal screen with a 270° horizontal field of view. Reward delivery was accomplished by sending by a TTL pulse from the virtual-reality computer to a solenoid valve (NResearch), which released a drop of a water to a lick tube located slightly in front and below the mice's mouth. The tone signifying trial failure was played through conventional computer speakers (Logitech). The setup was enclosed in a custom-designed cabinet built from optical rails (Thorlabs) and lined with sound-absorbing foam sheeting (McMaster-Carr).

In vivo optical imaging and data acquisition

Imaging was performed using a custom- built, virtual-reality-compatible two-photon microscope (64).

13

The microscope was equipped with a pulsed Ti:sapphire laser (Chameleon Vision, Coherent) tuned to 920 nm. The scanning unit used a 5-mm galvanometer and an 8-kHz resonant scanning mirror (Cambridge Technologies). The collected photons were split into two channels by a dichroic mirror (FF562-Di03, Semrock). The light for the respective green and red channels was filtered using bandpass filters (FF01-520/60 and FF01-607/70, Semrock), and then detected using GaAsP photomultiplier tubes (1077PA-40, Hamamatsu). The signal from the photomultiplier tubes was amplified using a high-speed current amplifier (59-179, Edmund). Black rubber tubing was attached to the objective (Zeiss 20×, 0.5 NA) as a light shield covering the space from the objective to the titanium ring surrounding the GRIN lens. Double-distilled water was used as the immersion medium. The microscope could be rotated along the mediolateral axis of the mice, allowing alignment of the optical axes of the microscope objective and GRIN lens as previously described for microprism imaging (64). Control of the microscope and image acquisition was performed using the ScanImage software (Vidrio Technologies37) that was run on a separate (scanning) computer. Images were acquired at 30 Hz at a resolution of 512 × 512 pixels. Average beam power measured at the front of the objective was 40–60 mW. Synchronization between the behavioural logs and acquired images was achieved by sending behavioural information each time the virtual-reality environment was refreshed from the virtual-reality computer to the scanning computer via an I2C serial bus; behavioural information was then stored in the header of the image files.

Behavioral training

Seven days after the surgery, mice were started on a water restriction protocol, with a daily allotment of water of 1–1.5 ml. Mice were monitored for signs of dehydration or reductions in body mass below 80% of the initial value. If any of these conditions occurred, mice were given ad libitum access to water until recovery. The animals were handled daily from the start of water restriction. Five days after starting water restriction and handling, mice began training in the behavioural setup. Training consisted of a shaping procedure with nine levels of T-mazes with progressively longer stem length and cognitive difficulty. After shaping concluded, in each session the first few trials (5–30) were warm-up trials drawn from mazes 5–8, and then trials from the final maze (#9) were used for the remainder of the session. Warm-up trials were excluded from all analyses presented in the paper. The mice typically received their daily allotment of water during task performance; if not, the remainder was provided to them at the end of the day.

Details of the virtual reality behavioral task

At the beginning of each trial, mice were presented with the start of a virtual T-maze. After 30 cm (start region; -30 cm to 0 cm in **Fig. 4**) the cue region began, in which cues randomly appeared on either side of the corridor. The portion of the maze where cues were presented (cue region) was 220 cm long (0 cm to 220 cm in **Fig. 4**), and after it the stem of the T-maze continued for another 80 cm in which no cues were presented (delay region; 220 cm to 300 cm in **Fig. 4**). At the end of the T-maze, the mouse had to enter one of the arms.Turning into the correct (more cues) side would elicit a water reward (6.4 µl), whereas an incorrect choice elicited a tone (pulsing 6–12-kHz tone for 1 s). At the time of reward or tone delivery, the visual environment froze for 1 s, and then disappeared for 2 s (after a successful trial) or 5 s (after a failed trial) before another trial was started. Cue distribution is described in detail in our previous methods (22).

Pavlovian conditioning and unexpected reward behavioral paradigm

After water restriction and handling, mice were habituated to head fixation for 2–3 sessions. Training consisted of 5 sessions (1 session per day); each session consisted of 50 reward deliveries (8 µL of water reward). During training, each reward was preceded by a 2 s tone that ended at the time of reward delivery. The time between a reward and the next tone delivery was sampled from an exponential distribution with a mean of 40 s. The tone consisted of a sum of multiple sine waves with frequencies of 2, 4, 6, 8 and 16 kHz, and an amplitude of 70 dB. All of the mice exhibited anticipatory licking by the end of the five days (increase in lick rate after tone presentation but before reward delivery). Some of the mice were previously trained for several days in a similar protocol in which the tone amplitude was 60 dB and the time between reward and subsequent tone was sampled from a uniform distribution between 5 and 15 s; these mice did not exhibit anticipatory licking until trained in the final protocol. After training, mice underwent a single test session that consisted of 64 trials; 50 of those trials were identical to the training trials (tone followed by reward), 7 trials were unexpected reward trials (reward delivery with no preceding tone) and 7 trials were unexpected omissions (tone not followed by reward). In all cases, the intertrial interval was sampled from an exponential distribution with a mean of 40 s. Trial identity was sampled randomly with the following exceptions: (1) the first five trials were standard trials (tone plus reward); (2) the first two non-standard trials were unexpected reward trials.

Unexpected reward paradigm from Eshel et al. 2016
We used previously collected in vivo electrophysiology data as a comparison to our inferred spike data during unexpected delivery of reward (39). After >1 week of recovery from stereotactic surgery involving implantation of a headplate and microdrive containing optrodes, mice were water-restricted. Weight was maintained above 90% of baseline body weight. Animals were head-restrained and habituated for 1-2 d before training. Mice underwent a variable-reward task, where mice received either water reward that was either predicted by odor delivery ("trial type 1"; 45% of all trials), or where water rewards of various sizes were delivered without any odor ("trial type 2"; 45% of all trials). For both experiments, reward sizes were chosen pseudorandomly from the following set: 0.1, 0.3, 1.2, 2.5, 5, 10, or 20 µL. For our comparison, we only looked at unexpected delivery of 10 µL rewards. We used these trials because they most closely matched the conditions of our unexpected reward trials (where unexpected 8 µL water rewards were delivered in a similar head-fixed protocol). We used electrophysiology recorded extracellularly from VTA during this task. Electrophysiology methods were previously described in detail (39).

Motion correction procedure for in vivo fluorescence data
Deep brain imaging can be associated with spatially non-uniform fast motion (frame to frame), as well as spatially non-uniform slow drift of the field of view (over several minutes). To perform accurate motion correction despite the spatial non-uniformity, we divided the video into small regions ("patches") that had relatively uniform motion, and separately corrected the motion within each patch, as described below. Motion correction was performed on the red channel of the recording when available, otherwise it was performed on the green channel (n = 9).

As described previously (22), before dividing the video into patches, we first performed rigid motion correction using a standard normalized cross-correlation method, to eliminate any spatially uniform motion ('matchTemplate' function in the openCV package in Python). This correction was performed on non-overlapping 50 s video clips to eliminate concerns that slow drift over the course of minutes would degrade performance. The template for the cross-correlation was calculated by dividing each clip into

non-overlapping sections of 100 frames, calculating the mean image of each section and obtaining the median of the mean images. Before these motion correction steps, the video was pre-processed as follows: (i) thresholded by subtracting a constant number and setting negative values to 0, such that the lower roughly 50% of pixels was 0; (ii) using the openCV function 'erode' (with a scalar '1' kernel); and (iii) convolved with a Gaussian (s.d. = 2 pixels). Motion correction and template calculation were performed iteratively ten times or until all absolute shifts were less than one pixel in both axes. Finally, the 50 s clips had to be aligned to each other. This required generating a 'master template' for the entire video, and then using the same normalized cross-correlation procedure as before ('matchTemplate' function). The master template was calculated by taking the median of the templates of all clips.

The next step of motion correction involved compensating for spatially non-uniform, slow drift by estimating the drift in local patches. Patches were defined manually around neurons of interest to contain objects that drifted coherently (patch width 80–160 pixels). To estimate the drift of each patch over time, we used a non-rigid image registration algorithm (demons algorithm, 'imregdemons' function in MATLAB). This algorithm outputs a pixel-by-pixel correction. However, direct application of this correction risks distorting the shape of the neurons or the amplitude of the signals. Therefore, we applied a uniform correction for each patch, based on the average shift of all pixels in the patch (based on the demons output). We implemented the demons algorithm on the templates from the 50 s clips described in the previous paragraph, again using the median of these templates as the master template. The registration and master templates were computed iteratively 20 times, or until the increase in the average correlation between each corrected template and the overall template was less than the s.e.m. of these correlations. We found that the performance of the non-rigid registration improved if the templates were first processed through a local normalization procedure (66).

Finally, we performed standard rigid motion correction using the normalized cross-correlation method on each patch and clip. We then repeated the rigid motion correction after taking a rolling mean of every two frames and downsampling the video by a factor of two. This increased the signal strength; we used this downsampled video for subsequent analysis. After correcting for motion within clips, we had to correct across clips. To this end, we performed rigid motion correction on the clip templates. The motion correction code can be found in: https://github.com/benengx/Deep-Brain-Motion-Corr.

***In vitro* recordings to compare GCaMP6 fluorescence with electrophysiology in DA neurons**
In order to compare GCaMP6f and GCaMP6m fluorescence with spike times in DA neurons, we performed *ex vivo* slice imaging and electrophysiological recordings in Ai148xDAT::Cre mice or DAT::IRES-Cre mice virally expressing GCaMP6m. Mice were anesthetized with an i.p. injection of Euthasol (0.06ml/30g) and decapitated. After extraction, the brain was immersed in ice-cold carbogenated NMDG ACSF (92 mM NMDG, 2.5 mM KCl, 1.25 mM NaH2PO4, 30 mM NaHCO3, 20 mM HEPES, 25 mM glucose, 2 mM thiourea, 5 mM Na-ascorbate, 3 mM Na-pyruvate, 0.5 mM CaCl2·4H2O, 10 mM MgSO4·7H2O, and 12 mM N-Acetyl-L-cysteine) for 2 minutes. The pH was adjusted to 7.3-7.4. Afterwards coronal slices (300um) were sectioned using a vibratome (VT1200s, Leica) and then incubated in NMDG ACSF at 34°C for 15 minutes. Slices were then transferred into a holding solution of HEPES ACSF (92 mM NaCl, 2.5 mM KCl, 1.25 mM NaH2PO4, 30 mM NaHCO3, 20 mM HEPES, 25 mM glucose, 2 mM thiourea, 5 mM Na-ascorbate, 3 mM Na-pyruvate, 2 mM

CaCl2·4H2O, 2 mM MgSO4·7H2O and 12 mM N-Acetyl-l-cysteine, bubbled at room temperature with 95% 02/ 5% CO2) for at least 45 mins until recordings were performed.

During cell-attached recordings, slices were perfused with a recording ACSF solution (120 mM NaCl, 3.5 mM KCl, 1.25 mM NaH2PO4, 26 mM NaHCO3, 1.3 mM MgCl2, 2 mM CaCl2 and 11 mM D-(+)-glucose, continuously bubbled with 95% O2/5% CO2) held at either 30°C or 37°C. Picrotoxin (100 µM) was added to the recording solution to block tonic inhibition and promote spontaneous activity. Cell-attached recordings were performed using a Multiclamp 700B (Molecular Devices, Sunnyvale, CA) using pipettes with a resistance of 4-6 MOhm filled with a solution identical to the recording ACSF. Infrared differential interference contrast–enhanced visual guidance was used to select neurons that were 3–4 cell layers below the surface of the slices, which were held at room temperature while the recording solution was delivered to slices via superfusion driven by a peristaltic pump. Cell-attached recordings were collected once a seal (>50 MOhm to ~1 GOhm) between the recording pipette and the cell membrane was obtained. Action potential currents were recorded in voltage-clamp mode with voltage clamped at 0 mV. Cell-attached currents were low-pass filtered at 1 kHz and digitized and stored at 10 kHz (Clampex 9; MDS Analytical Technologies). All experiments were completed within 4 hours after slicing the brain. Fluorescence was imaged using a CMOS camera (ORCA-Flash 2.8, Hamamatsu) at 33.333 Hz (30 ms exposure windows) using a GFP filter cube set (exciter ET470/40x, dichroic T495LP, emitter ET525/50m).

Motion correction procedure for *in vitro* data
On our *in vitro* imaging data, we performed rigid motion correction using a standard normalized cross-correlation method, to eliminate any spatially uniform motion ('matchTemplate' function in the openCV package in Python). The template for the cross-correlation was hand-selected as a frame near the middle of the recording when cell fluorescence was relatively bright. Since these recordings were usually around 2 mins in length, we performed motion correction on the entire recording rather than sectioning the recording into individual "patches."

ROI selection and calculation of raw GCaMP for *in vivo* and *in vitro* data
As previously described (22), to determine ROIs for fluorescence measurement, we used the mean (*in vivo* only) or S.D. (*in vivo* and *in vitro*) projection of the recording to generate a clear outline of the cell body. ROIs were defined manually using this S.D. projection. An initial automatic annulus was generated by enlarging the borders of the ROI twice (by 5 µm and 10 µm); the annulus was the shape contained between the two enlarged borders, where we expect that observed activity would be due to the neuropil but not the cell itself. Next, we manually reshaped the annulus region to avoid any visible dendrites, processes or cell bodies, while approximately maintaining its original area.

To correct for neuropil contamination in the *in vivo* and *in vitro* imaging data, we subtracted a scaled version of the annulus fluorescence from the raw trace: $F_{corr}(t) = F_{raw}(t) - Y*F_{annulus}(t)$, in which $F_{raw}(t)$ is the mean fluorescence in the ROI of each neuron at time t; $F_{annulus}(t)$ is the mean fluorescence in the corresponding annulus ROI at time t; and Y is the correction factor (1,67). The correction factor is intended to reflect the fraction of the *z*-section that is generated by the neuropil versus the cell that is being imaged. The correction factor used was 0.58 for the *in vivo* data and 1 for the *in vitro* data. For the *in vitro* data, spike time estimates and decay rate estimates were similar whether a correction factor of 1 or 0.7 was used.

17

In vitro data organization

The in-vitro data consist of 12 recordings of GCaMP6f at 30°C; 9 recordings of GCaMP6f at 37°C; and 21 recordings of GCaMP6m at 37°C. Each recording consists of a time series of fluorescence intensities obtained via calcium imaging at 33.33 Hz as well as a set of ground-truth spike times obtained via an electrode. We defined bursting as a series of at least 3 spikes with a maximum inter-spike interval of 100 ms. Because we anticipated burst activity in the *in vivo* experiments, we focused our analysis on recordings where the fraction of spikes within a burst was >=10%. Recordings where the fraction of spikes within a burst was <10% were deemed pacemaker cells and were excluded from analysis.

In vitro data preprocessing

We performed data preprocessing to calculate ΔF/F and to further adjust for changes in baseline fluorescence.

First, we shifted and scaled the fluorescence F_t by the baseline fluorescence F_t^0 to produce a normalized trace, $G_t = \Delta F_t / \hat{F}_t^0 = (F_t - \hat{F}_t^0)) / \hat{F}_t^0$, where $\hat{F}_t^0$ is an estimate of the baseline fluorescence at t. We estimated $\hat{F}_t^0$ by the 8% percentile of $F_{t-L}, F_{t-(L-1)}, \ldots, F_t$ where $L = 60s \times 33.33Hz$, corresponding to a 60 second lag.

To further remove linear trends that may occur over the course of an entire fluorescence recording, we computed the residuals from a robust linear regression of G_t onto time with Gaussian mixture errors, as in Theis et al. (2016). Then, we subtracted the 1st percentile from these residuals, and divided by the difference between the 80th and 1st percentiles. We let $\tilde{G}_t$ denote the resulting scaled and shifted residuals.

Next, to allow for the possibility that more than one spike occurs within a (1/33.33)s window, we resampled the data $\tilde{G}_t$ from 33.33 Hz to 66.66 Hz using the scipy.signal.resample function in python. We call this time series the fluorescence trace, or the GCaMP ΔF/F, and denote it by y_t for $t = 1, \ldots, T$ time steps, where now the length of each time step is 1/66.66 seconds.

In vitro data analysis

Statistical model

In order to infer spike times on the basis of the calcium imaging data, we consider a model for calcium dynamics proposed by (14) and (8). We view the fluorescence trace y_t at the t th time step as a noisy observation of the underlying calcium concentration c_t, which is assumed to decay with rate $0 < \gamma < 1$, unless there is a spike. A spike causes an instantaneous increase in the calcium concentration. The full model is

$$y_t = c_t + \epsilon_t, \quad \epsilon_t \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2), \quad t = 1, \ldots, T,$$

18

$$c_t = \gamma c_{t-1} + z_t, \quad t = 2, \ldots, T,$$

where $z_t = 0$ unless there is a spike at the tth timepoint, in which case $z_t > 0$.

As previously discussed (4,5), a natural estimate of the spike times, i.e. the times s such that $z_s > 0$, is obtained by solving the non-convex optimization problem

$$\underset{c_1,\ldots,c_T}{\text{minimize}} \left\{ \frac{1}{2} \sum_{t=1}^{T} (y_t - c_t)^2 + \lambda \sum_{t=2}^{T} 1_{(c_t \neq \gamma c_{t-1})} \right\} \text{ subject to } c_t \geq \gamma c_{t-1}.$$

Here, λ is a tuning parameter that controls the trade-off between the goodness of fit to the data--as measured by the residual sum of squares between estimated calcium and the observed fluorescence--and the number of inferred spikes. This optimization problem is non-convex, but can be solved for the global optimum using a dynamic programming strategy laid out in Jewell et al. (2019). We say that we have inferred a spike if $\hat{z}_s = \hat{c}_s - \gamma \hat{c}_{s-1} > 0$. Smaller values of λ result in more inferred spikes and thus a higher inferred firing rate (number of inferred spikes / length of recording), whereas larger values of λ result in fewer inferred spikes, and thus a smaller inferred firing rate.

This model assumes that a spike leads to an instantaneous increase in the calcium; in reality, the calcium increase due to a spike is not instantaneous. Therefore, after fitting the model, we shift the inferred spike times by 4 time steps ($\approx 0.06s$).

To fit this model, we must specify the decay rate γ and a value of the tuning parameter λ.

Estimating the decay rate γ

According to the statistical model for calcium decay, the rate of calcium decay γ is constant across multiple inter-spike regions. Thus, given knowledge of the spiking events from the ground-truth data, we can estimate the decay rate from multiple inter-spike segments.

More precisely, given n spike indices (i.e., the times where there is at least one spike) $t_1, t_2, \ldots, t_n$, we defined $\overline{m}(i)$ as the index between t_i and t_{i+1} with maximal fluorescence value, $\overline{m}(i) = \underset{t_i \leq j \leq t_{i+1}}{\text{argmax}} \{y_j\}$. Similarly, we let $\underline{m}(i) = \underset{t_i \leq j \leq t_{i+1}}{\text{argmin}} \{y_j\}$ be the index with minimal fluorescence value. These indices segment the fluorescence trace from peak-to-trough between spikes. Of these, we considered segments whose peak-to-trough distance is at least $K = 10$ time steps,

$$\mathcal{I} = \bigcup_{\{i=1,\ldots,n-1 \mid \overline{m}(i) - \underline{m}(i) > K\}} (\overline{m}(i), \underline{m}(i)).$$

Then we estimated the decay rate by finding the value of γ that minimizes the residual sum of squares of the best exponential curve to each of these segments,

$$\hat{\gamma} \in \operatorname*{argmin}_{\gamma} \left\{ \sum_{(a,b)\in\mathcal{I}} \min_{c_0} \left\{ \sum_{t=a}^{b} (y_t - \gamma^{t-a} c_0)^2 \right\} \right\}.$$

This process is illustrated in **Fig. 1d**. Here, values of the fluorescence trace whose indices are contained in $\mathcal{I}$ are colored blue. Using these points, $\hat{\gamma}$ is the value that gives the lowest total residual sum of squares obtained by fitting exponential curves to each set of points. For each segment, the best exponential fit with decay rate $\hat{\gamma}$ is plotted in red.

Estimating the tuning parameter λ

The tuning parameter λ controls the tradeoff between goodness of fit to the data and the number of inferred spikes. A smaller value of λ leads to a higher inferred firing rate, and a larger value leads to a lower inferred firing rate.

On the *in vitro* data, we set λ so that the inferred firing rate equals the true firing rate. In greater detail, we defined $\lambda^\star$ to be the value that minimizes the difference between the number of inferred spikes and the number of time bins (of duration $(1 / 66.66)$s) with at least one spike,

$$\lambda^\star \in \operatorname*{argmin}_{\lambda} \left\{ |\#\{s | \hat{z}_s(\lambda) > 0\} - n| \right\},$$

where the inferred spikes $\{s | \hat{z}_s(\lambda) > 0\}$ were obtained by solving the optimization problem with the tuning parameter λ and setting $\hat{z}_s(\lambda) = \hat{c}_s - \gamma \hat{c}_{s-1}$, and where n is the number of time bins with at least one ground-truth spike. We efficiently computed $\lambda^\star$ using binary search.

Assessing model performance

To assess whether the inferred spikes agree with the ground-truth spikes, we used the van Rossum distance metric (40,41). This metric measures the distance between two discrete vectors of spike times $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{v} = (v_1, \ldots, v_{n'})$ by first mapping each to a continuous space and then using the usual Euclidean distance between two functions. Specifically, a discrete vector $\mathbf{u}$ is mapped to a continuous function, $f(t; \mathbf{u}) = \sum_{i=1}^{n} h(t - u_i)$, where h is the kernel

$$h(t) = \begin{cases} 0, & t < 0 \\ e^{-t/\tau}, & t \geq 0 \end{cases}$$

and where $\tau > 0$ is a time-scale parameter. The van Rossum distance $d(\mathbf{u}, \mathbf{v}; \tau)$ between two spike trains $\mathbf{u}$ and $\mathbf{v}$ is given as the distance between the induced functions $f(t; \mathbf{u})$ and $f(t; \mathbf{v})$,

$$d(\mathbf{u}, \mathbf{v}; \tau)^2 \quad = \tfrac{1}{\tau} \int_0^\infty (f(t; \mathbf{u}) - f(t; \mathbf{v}))^2 \mathrm{d}t.$$

In vivo **data analysis**

Fluorescence data preprocessing

The same data preprocessing pipeline described above in the **In vitro data preprocessing** section was used to de-trend using a robust linear regression, subtract the 1st percentile, scale by the difference between the 80th and 1st percentiles, and then upsample from 30 Hz to 60 Hz.

Decay rate determination

We used the decay rates estimated from the *in vitro* experiment (**Fig. 1d**) for the *in vivo* analysis. In particular, we calculated a median decay rate for each GCaMP indicator and temperature based on the *in vitro* data. We adjusted these median decay rates to account for the difference in recording frequency between the *in vitro* and *in vivo* data (the latter was recorded at 60 Hz),

$$\text{decay rate at 60 Hz} = 1 - (66.66 / 60.00) * (1 - \text{decay rate at 66.66 Hz}).$$

The adjusted median decay rate for the appropriate GCaMP indicator (γ = 0.970 for GCaMP6f at 37°C; γ = 0.984 for GCaMP6m at 37°C) is used for all *in vivo* analysis.

Tuning parameter determination

We use binary search to identify a value of the tuning parameter λ that results in a set of inferred spikes with a target average firing rate r. In particular, we solve

$$\lambda^\star \in \underset{\lambda}{\text{argmin}} \left\{ |\#\{s | \hat{z}_s(\lambda) > 0\} - r \cdot T/60| \right\}$$

where T denotes the length of the recording in time steps. For the *in vivo* analysis, our target firing rate was 6 Hz (36–39,46).

Analysis of firing rate data in unexpected reward and Pavlovian conditioning experiment

For the unexpected reward and Pavlovian conditioning data (**Figs. 2c, 3e, 3h**), peristimulus time histograms (PSTHs) were constructed using electrophysiology and inferred spike times. For a given cell and trial type, we found all spikes that occurred within some window (-4 s to 4 s) of the behavioral event. Electrophysiology spike times were organized into 1 ms time bins, whereas inferred spike times were rounded to the nearest 1 ms. For each neuron, and for each trial type, the mean instantaneous firing rate at each time bin was calculated by summing the number of spikes occurring within that 1 ms time bin across trials, and dividing by the product of the number of trials and the bin width (1/1000 ms). We then convolved this mean firing rate with a Gaussian kernel (width = 200 ms, SD = 40 ms; normalized such that the sum of the kernel equaled 1), producing a smooth PSTH for each neuron. The population firing rate PSTHs (**Figs. 2e-f, 3c,g**), was calculated as the average of the individual cells' smooth PSTHs.

For a given cell, mean firing rates for a given period were calculated by taking the total number of spikes in some time window, and dividing by the product of the width of that window and the number of relevant trials. We calculated the baseline firing rate over a 1 s period immediately before reward delivery (for unexpected reward trials) or cue onset (for expected reward and reward omission trials). Cue, reward, and omission responses were baseline-subtracted using the baseline firing rates calculated across those respective trials. The time windows for cue, reward, and reward omission responses were selected for consistency with the Eshel et al. experiments (39). The responses to

21

unexpected and expected reward were calculated by finding the mean firing rate over a 600 ms window after reward onset. The response to cue presentation was calculated by finding the mean firing rate a 500 ms window beginning 150 ms after cue onset. The response to reward omission was calculated by finding the mean firing rate over a 1,300 ms window after reward omission onset. Peak amplitudes of cue and reward responses were calculated as the largest peak of that cell's smooth PSTH in the respective response window. Full duration at half max of cue and reward responses were calculated using the width of the largest peak, as measured at half the max value, of the smooth PSTH in the response window.

To detect significant changes in response to reward omission for a given neuron, for each trial we calculated the firing rate in a 1 s window prior to cue onset and the firing rate in a 1,300 ms window following omission onset. We performed a paired t-test to determine whether the firing rates during the omission window were significantly different from baseline, and categorized a neuron as having a significant omission response if the t-test rejected the null hypothesis at level $p = 0.05$ and the mean omission response was < 0 Hz.

Analyzing GCaMP fluorescence and inferred spike rates by maze position in the VR task
In Figure 4, we plotted GCaMP fluorescence and inferred spike rates as a function of position. To smooth this data for visualization (**Fig. 4g,k**), for a given cell we calculated moving averages of GCaMP fluorescence and inferred spike rate using a 10 cm window moved at 1 cm intervals along the maze. To compute the inferred spiking rate by position, we divided the number of inferred spikes within a given 10 cm window by the amount of time spent in that window. We then averaged GCaMP and inferred spike rates across trials to generate cell averages (**Fig. 4g,k**).

Quantifying ramps in VR task using GCaMP fluorescence and inferred spikes
To describe linear ramps by position in each neuron (**Fig. 4l**), we performed a linear regression of the observed GCaMP fluorescence or the inferred spike rates onto the mouse's position. We calculated the mean inferred spike rates and GCaMP fluorescence over maze positions using the same procedure for **Fig 4 g,k**, except that we used non-overlapping 10 cm position bins (with edges at maze positions -30 and 290 cm), where the positional value was the center of that bin. For **Fig. 4l**, we calculated the linear change in GCaMP or inferred spike rate for each neuron by multiplying the position coefficient of its respective linear model by the length of the maze (290 cm).

Code availability
The code for calcium imaging data preprocessing, decay rate estimation, and spike inference is available on github (https://github.com/jewellsean/spike_tools).
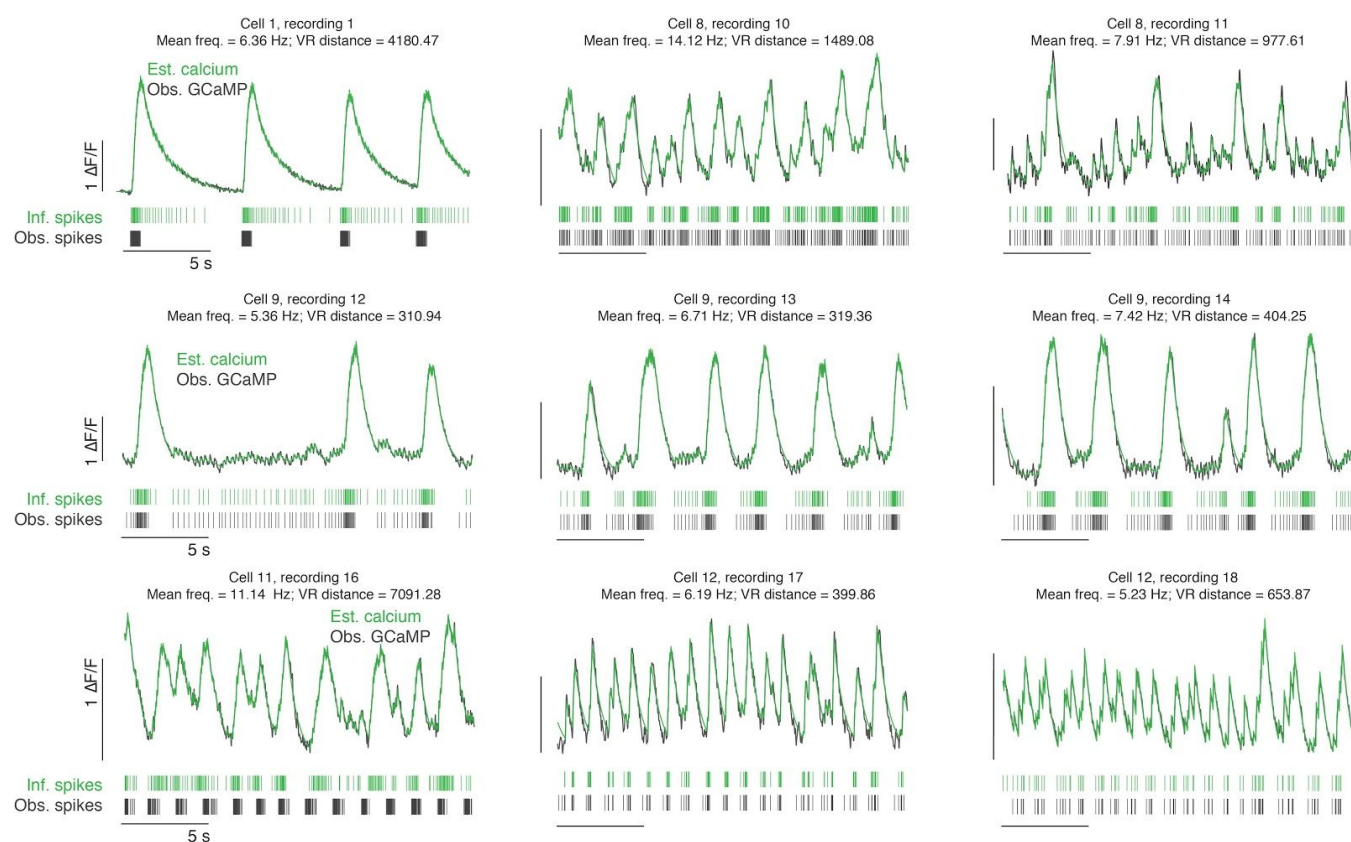
Supplementary figures
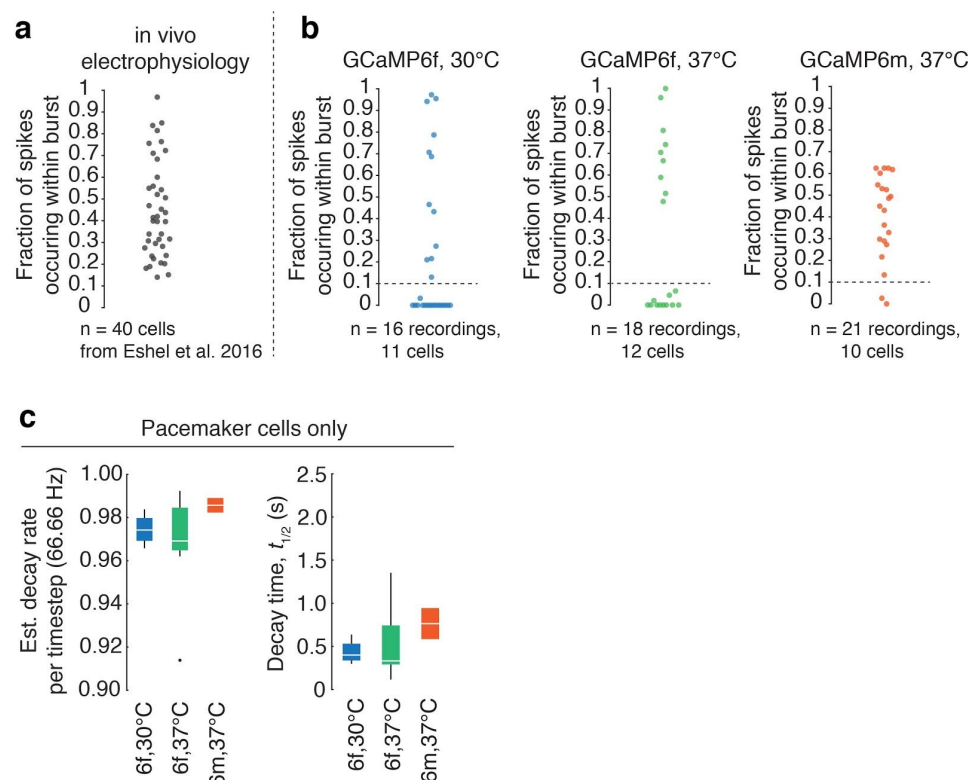
Supplementary figure 1. Identification of pacemaker cells by absence of phasic activity, and decay statistics for pacemaker cells. **a.** Percent of spikes that occurred within a burst in a previously collected in vivo electrophysiology dataset (Eshel et al., 2016). A burst was defined by a series of at least 3 spikes with <= 100 ms between each spike. All in vivo electrophysiology cells had >10% of spikes within a burst. This was chosen as the cut-off for inclusion in in vitro analysis. **b.** Percent of spikes that occured within a burst for all recordings in the GCaMP6f at 30°C, GCaMP6f at 37°C, and GCaMP6m at 37°C groups. Cells with <10% of spikes occurring within a burst were excluded from group analysis, since they featured little or no phasic activity. **c.** Left: Estimated decay rates for pacemaker cells for GCaMP6f at 30°C (n = 4 recordings across 2 cells; median decay rate = 0.974, Q1 = 0.969, Q3 = 0.980), GCaMP6f at 37°C (n=9 recordings across 7 total cells; median decay rate = 0.969, Q1 = 0.965, Q3 = 0.985), and GCaMP6m at 37°C (n=2 recordings from 1 cell; median decay rate = 0.986, Q1 = 0.982, Q3 = 0.989). Right: Half decay time for pacemaker cells for GCaMP6f at 30°C (median = 0.400 s, Q1 = 0.337 s, Q3 = 0.530 s), GCaMP6f at 37°C (median = 0.332 s, Q1 = 0.291 s, Q3 = 0.742 s) and GCaMP6m at 37°C (median = 0.764 s, Q1 = 0.583 s, Q3 = 0.944 s).
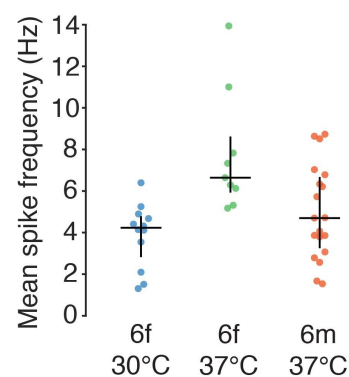
Supplementary figure 2. Mean firing frequency in bursty cells during in vitro recordings. Mean spike frequency during the simultaneous calcium imaging and electrophysiology in vitro experiment (GCaMP6f at 30°C, n=12 recordings from 9 total cells, median firing rate = 4.2 Hz, Q1 = 2.8 Hz, Q3 = 4.8 Hz; GCaMP6f at 37°C, n=9 recordings from 5 total cells, median firing rate = 6.5 Hz, Q1 = 5.9 Hz, Q3 = 8.6 Hz; GCaMP6m at 37°C, n=19 recordings from 9 total cells, median firing rate = 4.7 Hz, Q1 = 3.3 Hz, Q3 = 6.7 Hz). Vertical bars are interquartile range (Q1 and Q3).

Supplementary figure 3. Example traces from all bursty recordings in GCaMP6f 30°C group. Examples of correspondence between observed GCaMP (black line, top) and estimated calcium (blue line, top), and observed spikes (black dashes, bottom) and inferred spikes (blue dashes, bottom) for each recording in GCaMP6f 30°C group. Traces are 20 s long. Scale bars are 1 ΔF/F and 5 s.

Supplementary figure 4. Example traces from all bursty recordings in GCaMP6f 37°C group. Examples of correspondence between observed GCaMP (black line, top) and estimated calcium (green line, top), and observed spikes (black dashes, bottom) and inferred spikes (green dashes, bottom) for each recording in GCaMP6f 37°C group. Traces are 20 s long. Scale bars are 1 ΔF/F and 5 s.
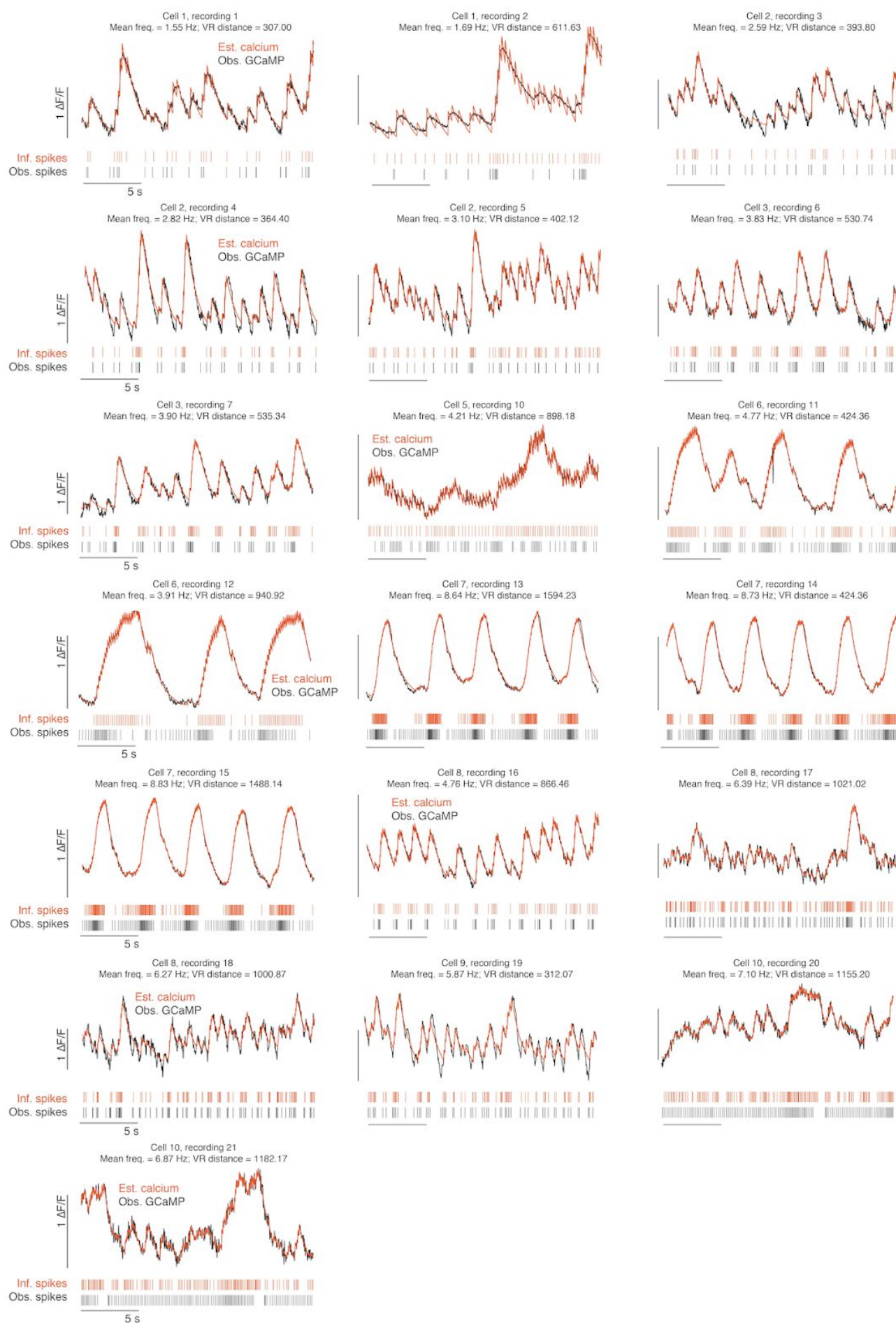
Supplementary figure 5. Example traces from all bursty recordings in GCaMP6m 37°C group. Examples of correspondence between observed GCaMP (black line, top) and estimated calcium (orange line, top), and observed spikes (black dashes, bottom) and inferred spikes (orange dashes, bottom) for each recording in the GCaMP6m 37°C group. Traces are 20 s long. Scale bars are 1 ΔF/F and 5 s.
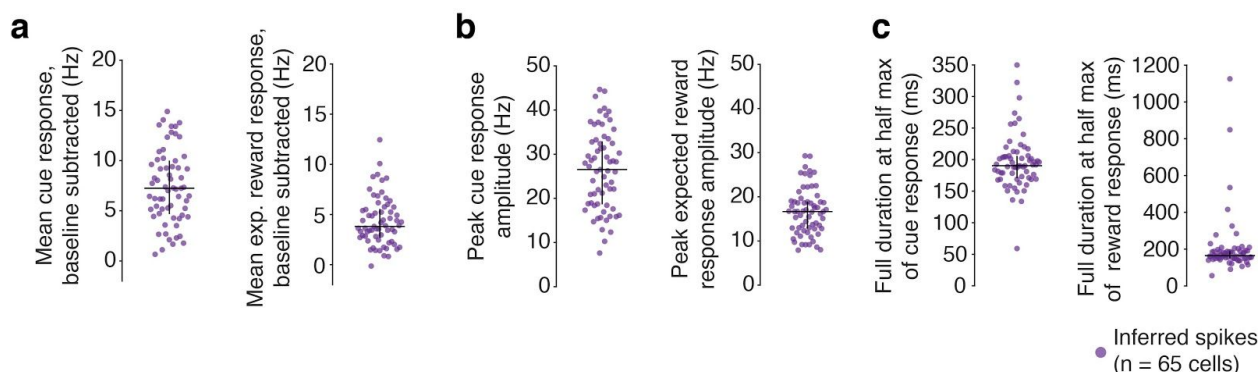
Supplementary figure 6. Mean cue and expected reward responses in inferred spikes. a. Left: Mean cue response, where cue response is the mean firing rate in a 500 ms period beginning 150 ms after cue onset, baseline subtracted using the mean firing rate over a 1 s period before cue (median = 7.3 Hz, Q1 = 4.7 Hz, Q3 = 10.0 Hz). Right: Mean expected reward response, where reward response is the mean firing rate over the first 600 ms following reward presentation, baseline subtracted using the mean firing rate over a 1 s period before cue (median = 3.8 Hz, Q1 = 2.7 Hz, Q3 = 5.6 Hz). **b.** Left: Peak cue response amplitude in inferred spikes, where peak is maximum value of PSTH in the 500 ms cue period (median = 26.5 Hz, Q1 = 18.6 Hz, Q3 = 32.9 Hz). Right: Peak expected reward response amplitude in inferred spikes, where peak is maximum value of PSTH in the first 600 ms period following reward presentation (median = 16.7 Hz, Q1 = 12.8 Hz, Q3 = 18.9 Hz). **c.** Left: Full duration at half max of cue response in inferred spikes (median = 190.4 ms, Q1 = 170.7 ms, Q3 = 206.0 ms). Right: Full duration at half max of expected reward response in inferred spikes (median = 164.9 ms, Q1 = 149.8 ms, Q3 = 189.5 ms). Vertical bars are interquartile range (Q1 and Q3).
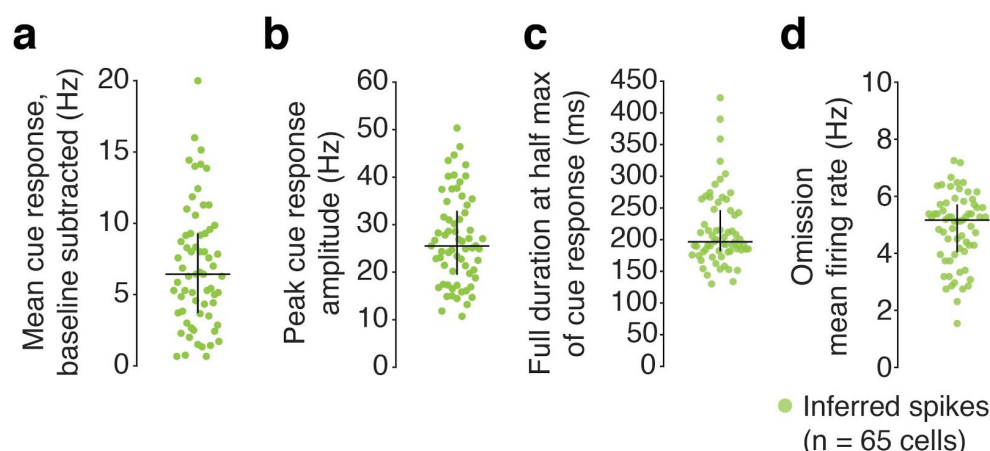
Supplementary figure 7. Mean cue and omission responses in inferred spikes during omission trials. a. Mean cue response during omission trials, where cue response is the mean firing rate over first 500 ms following cue presentation, baseline subtracted using the mean firing rate over a 1 s period before cue (median = 6.4 Hz, Q1 = 3.7 Hz, Q3 = 9.3 Hz). **b.** Peak cue response amplitude in inferred spikes during omission trials, where peak is maximum value of PSTH in the first 500 ms period following cue presentation (median = 25.5 Hz, Q1 = 19.5 Hz, Q3 = 32.9 Hz).**c.** Full duration at half max of cue response in inferred spikes during omission trials (median = 192.6 ms, Q1 = 181.6 ms, Q3 = 246.4 ms). d. Mean firing rate in inferred spikes over 1,300 ms period following onset of reward omission (median = 5.2 Hz, Q1 = 4.0 Hz, Q3 = 5.7 Hz). Vertical bars are interquartile range (Q1 and Q3).
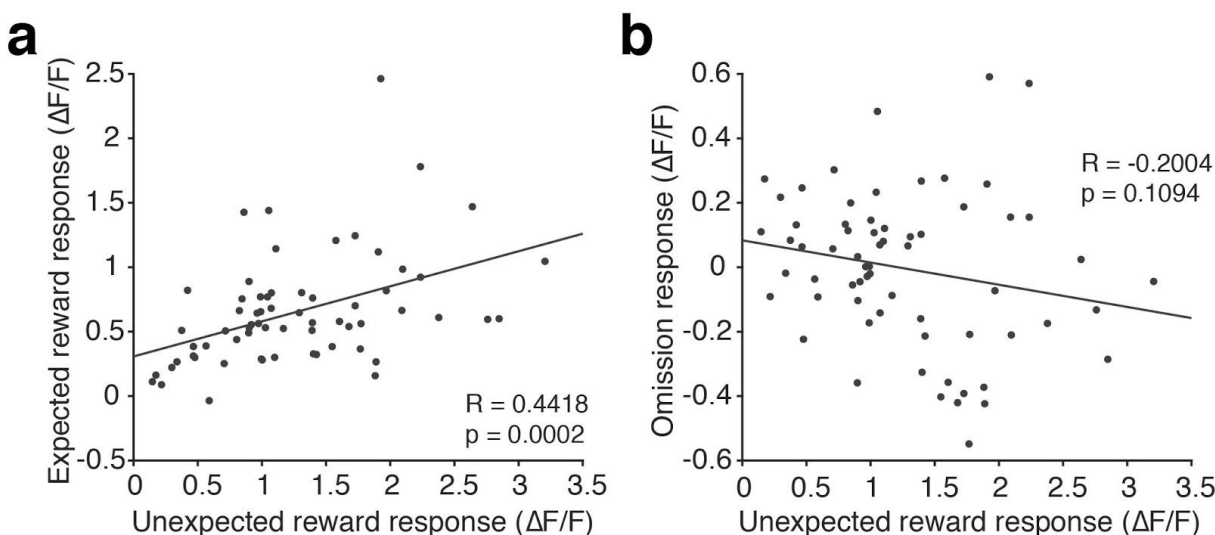
Supplementary figure 8. Reward and omission response correlations using GCaMP. a. Scatterplot of expected reward response versus unexpected reward response. Responses are calculated by finding mean ΔF/F value over a 600 ms window following reward presentation and baseline subtracting the mean ΔF/F value over a 1 s window before trial start (1 s before cue for expected reward; 1 s before reward for unexpected reward). **b.** Scatterplot of omission response versus unexpected reward response. Omission response is calculated by finding the mean ΔF/F value over a 1,300 ms window following reward omission and baseline subtracting the mean ΔF/F value over a 1 s window before cue.

REFERENCES

1. Chen T-W, Wardill TJ, Sun Y, Pulver SR, Renninger SL, Baohan A, et al. Ultrasensitive fluorescent proteins for imaging neuronal activity. Nature. 2013 Jul 18;499(7458):295–300.

2. Dombeck DA, Khabbaz AN, Collman F, Adelman TL, Tank DW. Imaging large-scale neural activity with cellular resolution in awake, mobile mice. Neuron. 2007 Oct 4;56(1):43–57.

3. de Vries SEJ, Lecoq JA, Buice MA, Groblewski PA, Ocker GK, Oliver M, et al. A large-scale standardized physiological survey reveals functional organization of the mouse visual cortex. Nat Neurosci. 2020 Jan;23(1):138–51.

4. Jewell SW, Hocking TD, Fearnhead P, Witten DM. Fast nonconvex deconvolution of calcium imaging data. Biostatistics [Internet]. 2019 Feb 8; Available from: http://dx.doi.org/10.1093/biostatistics/kxy083

5. Jewell S, Witten D. EXACT SPIKE TRAIN INFERENCE VIA ℓ0 OPTIMIZATION. Ann Appl Stat. 2018 Dec;12(4):2457–82.

6. Berens P, Freeman J, Deneux T, Chenkov N, McColgan T, Speiser A, et al. Community-based benchmarking improves spike rate inference from two-photon calcium imaging data. PLoS Comput Biol. 2018 May;14(5):e1006157.

7. Pachitariu M, Stringer C, Harris KD. Robustness of Spike Deconvolution for Neuronal Calcium Imaging. J Neurosci. 2018 Sep 12;38(37):7976–85.

8. Friedrich J, Zhou P, Paninski L. Fast online deconvolution of calcium imaging data. PLoS Comput Biol. 2017 Mar;13(3):e1005423.

9. Dyer EL, Studer C, Robinson JT, Baraniuk RG. A robust and efficient method to recover neural events from noisy and corrupted data. In: 2013 6th International IEEE/EMBS Conference on Neural Engineering (NER). 2013. p. 593–6.

10. Dyer EL, Duarte MF, Johnson DH, Baraniuk RG. Recovering Spikes from Noisy Neuronal Calcium Signals via Structured Sparse Approximation. In: Latent Variable Analysis and Signal Separation. Springer Berlin Heidelberg; 2010. p. 604–11.

11. Pnevmatikakis EA, Merel J, Pakman A, Paninski L. Bayesian spike inference from calcium imaging data. In: 2013 Asilomar Conference on Signals, Systems and Computers. 2013. p. 349–53.

12. Pnevmatikakis EA, Soudry D, Gao Y, Machado TA, Merel J, Pfau D, et al. Simultaneous Denoising, Deconvolution, and Demixing of Calcium Imaging Data. Neuron. 2016 Jan 20;89(2):285–99.

13. Vogelstein JT, Watson BO, Packer AM, Yuste R, Jedynak B, Paninski L. Spike inference from calcium imaging using sequential Monte Carlo methods. Biophys J. 2009 Jul 22;97(2):636–55.

14. Vogelstein JT, Packer AM, Machado TA, Sippy T, Babadi B, Yuste R, et al. Fast nonnegative deconvolution for spike train inference from population calcium imaging. J Neurophysiol. 2010

Dec;104(6):3691–704.

15. Yaksi E, Friedrich RW. Reconstruction of firing rate changes across neuronal populations by temporally deconvolved Ca2+ imaging. Nat Methods. 2006 May;3(5):377–83.

16. Sasaki T, Takahashi N, Matsuki N, Ikegaya Y. Fast and accurate detection of action potentials from somatic calcium fluctuations. J Neurophysiol. 2008 Sep;100(3):1668–76.

17. Deneux T, Kaszas A, Szalay G, Katona G, Lakner T, Grinvald A, et al. Accurate spike estimation from noisy calcium signals for ultrafast three-dimensional imaging of large neuronal populations in vivo. Nat Commun. 2016 Jul 19;7:12190.

18. Theis L, Berens P, Froudarakis E, Reimer J, Román Rosón M, Baden T, et al. Benchmarking Spike Rate Inference in Population Calcium Imaging. Neuron. 2016 May 4;90(3):471–82.

19. Greenberg DS, Houweling AR, Kerr JND. Population imaging of ongoing neuronal activity in the visual cortex of awake rats. Nat Neurosci. 2008 Jul;11(7):749–51.

20. Grewe BF, Langer D, Kasper H, Kampa BM, Helmchen F. High-speed in vivo calcium imaging reveals neuronal network activity with near-millisecond precision. Nat Methods. 2010 May;7(5):399–405.

21. Rupprecht P, Carta S, Hoffmann A, Echizen M. A deep learning toolbox for noise-optimized, generalized spike inference from calcium imaging data. bioRxiv [Internet]. 2020; Available from: https://www.biorxiv.org/content/10.1101/2020.08.31.272450v1.abstract

22. Engelhard B, Finkelstein J, Cox J, Fleming W, Jang HJ, Ornelas S, et al. Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. Nature. 2019 Jun;570(7762):509–13.

23. da Silva JA, Tecuapetla F, Paixão V, Costa RM. Dopamine neuron activity before action initiation gates and invigorates future movements. Nature. 2018 Feb 8;554(7691):244–8.

24. Gore BB, Soden ME, Zweifel LS. Visualization of plasticity in fear-evoked calcium signals in midbrain dopamine neurons. Learn Mem [Internet]. 2014; Available from: http://learnmem.cshlp.org/content/21/11/575.short

25. Howe MW, Dombeck DA. Rapid signalling in distinct dopaminergic axons during locomotion and reward. Nature. 2016 Jul 28;535(7613):505–10.

26. Cho JR, Chen X, Kahan A, Elliott Robinson J, Wagenaar DA, Gradinaru V. Dorsal raphe dopamine neurons signal motivational salience dependent on internal and external states [Internet]. 2020 [cited 2020 Aug 7]. p. 2020.07.27.222729. Available from: https://www.biorxiv.org/content/10.1101/2020.07.27.222729v1

27. Howe M, Ridouh I, Allegra Mascaro AL, Larios A, Azcorra M, Dombeck DA. Coordination of rapid cholinergic and dopaminergic signaling in striatum during spontaneous movement. Elife [Internet]. 2019 Mar 28;8. Available from: http://dx.doi.org/10.7554/eLife.44903

28. Foehring RC, Zhang XF, Lee JCF, Callaway JC. Endogenous calcium buffering capacity of substantia nigral dopamine neurons. J Neurophysiol. 2009 Oct;102(4):2326–33.

29. Puopolo M, Raviola E, Bean BP. Roles of subthreshold calcium current and sodium current in

spontaneous firing of mouse midbrain dopamine neurons. J Neurosci. 2007 Jan 17;27(3):645–56.

30. Wilson CJ, Callaway JC. Coupled oscillator model of the dopaminergic neuron of the substantia nigra. J Neurophysiol. 2000 May;83(5):3084–100.

31. Pan P-Y, Ryan TA. Calbindin controls release probability in ventral tegmental area dopamine neurons. Nat Neurosci. 2012 Jun;15(6):813–5.

32. Schmidt H. Three functional facets of calbindin D-28k. Front Mol Neurosci. 2012 Mar 15;5:25.

33. Schwaller B. Cytosolic Ca2+ Buffers Are Inherently Ca2+ Signal Modulators. Cold Spring Harb Perspect Biol [Internet]. 2020 Jan 2;12(1). Available from: http://dx.doi.org/10.1101/cshperspect.a035543

34. Donahue CH, Liu M, Kreitzer AC. Distinct value encoding in striatal direct and indirect pathways during adaptive learning [Internet]. Available from: http://dx.doi.org/10.1101/277855

35. Gritton HJ, Howe WM, Romano MF, DiFeliceantonio AG, Kramer MA, Saligrama V, et al. Unique contributions of parvalbumin and cholinergic interneurons in organizing striatal networks during movement. Nat Neurosci. 2019 Apr;22(4):586–97.

36. Mameli-Engvall M, Evrard A, Pons S, Maskos U, Svensson TH, Changeux J-P, et al. Hierarchical control of dopamine neuron-firing patterns by nicotinic receptors. Neuron. 2006 Jun 15;50(6):911–21.

37. Dahan L, Astier B, Vautrelle N, Urbain N, Kocsis B, Chouvet G. Prominent burst firing of dopaminergic neurons in the ventral tegmental area during paradoxical sleep. Neuropsychopharmacology. 2007 Jun;32(6):1232–41.

38. Floresco SB, West AR, Ash B, Moore H, Grace AA. Afferent modulation of dopamine neuron firing differentially regulates tonic and phasic dopamine transmission. Nat Neurosci. 2003 Sep;6(9):968–73.

39. Eshel N, Tian J, Bukwich M, Uchida N. Dopamine neurons share common response function for reward prediction error. Nat Neurosci. 2016 Mar;19(3):479–86.

40. van Rossum MC. A novel spike distance. Neural Comput. 2001 Apr;13(4):751–63.

41. Houghton C, Kreuz T. On the efficient calculation of van Rossum distances. Network. 2012 May 8;23(1-2):48–58.

42. Eshel N, Bukwich M, Rao V, Hemmelder V, Tian J, Uchida N. Arithmetic and local circuitry underlying dopamine prediction errors [Internet]. Vol. 525, Nature. 2015. p. 243–6. Available from: http://dx.doi.org/10.1038/nature14855

43. Tobler PN, Fiorillo CD, Schultz W. Adaptive coding of reward value by dopamine neurons. Science. 2005 Mar 11;307(5715):1642–5.

44. Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. Science. 1997 Mar 14;275(5306):1593–9.

45. Bayer HM, Glimcher PW. Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron. 2005 Jul 7;47(1):129–41.

46. Grace AA, Bunney BS. The control of firing pattern in nigral dopamine neurons: single spike firing. J Neurosci. 1984 Nov;4(11):2866–76.

47. Cohen JY, Haesler S, Vong L, Lowell BB, Uchida N. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. Nature. 2012 Jan 18;482(7383):85–8.

48. Starkweather CK, Babayan BM, Uchida N, Gershman SJ. Dopamine reward prediction errors reflect hidden-state inference across time. Nat Neurosci. 2017 Apr;20(4):581–9.

49. Flagel SB, Clark JJ, Robinson TE, Mayo L, Czuj A, Willuhn I, et al. A selective role for dopamine in stimulus–reward learning. Nature. 2011 Jan 1;469(7328):53–7.

50. Fiorillo CD, Newsome WT, Schultz W. The temporal precision of reward prediction in dopamine neurons. Nat Neurosci. 2008 Aug;11(8):966–73.

51. Fiorillo CD, Yun SR, Song MR. Diversity and homogeneity in responses of midbrain dopamine neurons. J Neurosci. 2013 Mar 13;33(11):4693–709.

52. Howe MW, Tierney PL, Sandberg SG, Phillips PEM, Graybiel AM. Prolonged dopamine signalling in striatum signals proximity and value of distant rewards. Nature. 2013 Aug 29;500(7464):575–9.

53. Carelli RM. Nucleus accumbens cell firing and rapid dopamine signaling during goal-directed behaviors in rats. Neuropharmacology. 2004;47 Suppl 1:180–9.

54. Kim HR, Malik AN, Mikhael JG, Bech P, Tsutsui-Kimura I, Sun F, et al. A unified framework for dopamine signals across timescales [Internet]. bioRxiv. 2019 [cited 2020 Jul 23]. p. 803437. Available from: https://www.biorxiv.org/content/10.1101/803437v1.full-text

55. Hamid AA, Pettibone JR, Mabrouk OS, Hetrick VL, Schmidt R, Vander Weele CM, et al. Mesolimbic dopamine signals the value of work. Nat Neurosci. 2016 Jan;19(1):117–26.

56. Guru A, Seo C, Post RJ, Kullakanda DS, Schaffer JA. Ramping activity in midbrain dopamine neurons signifies the use of a cognitive map. bioRxiv [Internet]. 2020; Available from: https://www.biorxiv.org/content/10.1101/2020.05.21.108886v1.abstract

57. Ledochowitsch P, Huang L, Knoblich U, Oliver M, Lecoq J, Reid C, et al. On the correspondence of electrical and optical physiology in in vivo population-scale two-photon calcium imaging [Internet]. 2019 [cited 2020 Oct 15]. p. 800102. Available from: https://www.biorxiv.org/content/10.1101/800102v1.full-text

58. Grace AA, Bunney BS. The control of firing pattern in nigral dopamine neurons: burst firing [Internet]. Vol. 4, The Journal of Neuroscience. 1984. p. 2877–90. Available from: http://dx.doi.org/10.1523/jneurosci.04-11-02877.1984

59. Siegle JH, Ledochowitsch P, Jia X, Millman D. Reconciling functional differences in populations of neurons recorded with two-photon imaging and electrophysiology. bioRxiv [Internet]. 2020; Available from: https://www.biorxiv.org/content/10.1101/2020.08.10.244723v1.abstract

60. Wei Z, Lin B-J, Chen T-W, Daie K, Svoboda K, Druckmann S. A comparison of neuronal population dynamics measured with calcium imaging and electrophysiology [Internet]. PLOS Computational Biology. 2020. Available from: https://doi.org/10.1371/journal.pcbi.1008198

61. Lammel S, Steinberg EE, Földy C, Wall NR, Beier K, Luo L, et al. Diversity of transgenic mouse

models for selective targeting of midbrain dopamine neurons. Neuron. 2015 Jan 21;85(2):429–38.

62. Daigle TL, Madisen L, Hage TA, Valley MT, Knoblich U, Larsen RS, et al. A Suite of Transgenic Driver and Reporter Mouse Lines with Enhanced Brain-Cell-Type Targeting and Functionality. Cell. 2018 Jul 12;174(2):465–80.e22.

63. Harvey CD, Coen P, Tank DW. Choice-specific sequences in parietal cortex during a virtual-navigation decision task. Nature. 2012 Mar 14;484(7392):62–8.

64. Low RJ, Gu Y, Tank DW. Cellular resolution optical access to brain regions in fissures: imaging medial prefrontal cortex and grid cells in entorhinal cortex. Proc Natl Acad Sci U S A. 2014 Dec 30;111(52):18739–44.

65. Aronov D, Tank DW. Engagement of neural circuits underlying 2D spatial navigation in a rodent virtual reality system. Neuron. 2014 Oct 22;84(2):442–56.

66. Sage D, Unser M. Teaching image-processing programming in Java. IEEE Signal Process Mag. 2003 Nov;20(6):43–52.

67. Cox J, Pinto L, Dan Y. Calcium imaging of sleep–wake related neuronal activity in the dorsal pons. Nat Commun [Internet]. 2016; Available from: https://www.nature.com/articles/ncomms10763?origin=ppub