

# Dopamine neurons share common response function for reward prediction error

Neir Eshel<sup>1,2</sup>, Ju Tian<sup>1</sup>, Michael Bukwich<sup>1</sup> & Naoshige Uchida<sup>1</sup>

Dopamine neurons are thought to signal reward prediction error, or the difference between actual and predicted reward. How dopamine neurons jointly encode this information, however, remains unclear. One possibility is that different neurons specialize in different aspects of prediction error; another is that each neuron calculates prediction error in the same way. We recorded from optogenetically identified dopamine neurons in the lateral ventral tegmental area (VTA) while mice performed classical conditioning tasks. Our tasks allowed us to determine the full prediction error functions of dopamine neurons and compare them to each other. We found marked homogeneity among individual dopamine neurons: their responses to both unexpected and expected rewards followed the same function, just scaled up or down. As a result, we were able to describe both individual and population responses using just two parameters. Such uniformity ensures robust information coding, allowing each dopamine neuron to contribute fully to the prediction error signal.

Dopamine is critical for motivated behavior, enabling animals to learn what is rewarding and to pursue those rewards<sup>1–3</sup>. Although confined to a small portion of the midbrain, dopamine-producing neurons project diffusely throughout the brain. Dopamine neurons have large cell bodies, broad axonal arborization<sup>4</sup> and the ability to release dopamine both at synapses and into the extracellular space<sup>5</sup>. At the same time, they show relatively stereotyped electrophysiological properties<sup>6</sup>, can electrically couple with each other<sup>7</sup> and demonstrate high levels of correlation in their *in vivo* firing rates<sup>8–11</sup>. With these characteristics, dopamine neurons are in a prime position to broadcast a common signal to the rest of the brain<sup>12</sup>.

The identity of this broadcast signal, however, remains controversial. Although recent evidence<sup>3,13</sup> has suggested that there is diversity in dopamine neuron genetic profiles<sup>14</sup>, physiology<sup>15,16</sup>, connectivity<sup>17–19</sup> and responses to aversive stimuli<sup>20,21</sup>, the majority of dopamine neurons seem to follow a very simple pattern<sup>11</sup>. When a reward is delivered unexpectedly, dopamine neurons fire a burst of action potentials. If the reward is fully expected, dopamine neurons no longer respond to it. And if an expected reward is unexpectedly omitted, dopamine neurons dip from their baseline firing rate. Taken together, it appears that dopamine neurons signal reward prediction error, or the difference between the reward an animal expects to receive and the reward it actually receives. This signal has been demonstrated in monkeys<sup>22</sup>, rodents<sup>23,24</sup> and humans<sup>25</sup> and fits seamlessly into computational models of learning<sup>26,27</sup>. Specifically, dopamine prediction errors are thought to act as a teaching signal, strengthening actions and associations that lead to reward and weakening those that do not.

Although most dopamine neurons appear to encode prediction errors, their responses need not be identical. For example, different dopamine neurons may have different thresholds for responding to unexpected rewards<sup>20</sup> or different extents of suppression by reward expectation. Such diversity, however, would undercut each neuron's

ability to convey the same signal. An ideal broadcast signal would be similar enough from neuron to neuron that downstream targets could decode the same information, regardless of the subset of dopamine neurons that they contact. The quantitative homogeneity of dopamine reward prediction errors has never been assessed.

In particular, we recently found that dopamine neurons use subtraction to calculate prediction error, and that the signal they subtract arises from neighboring ventral tegmental area (VTA) GABA neurons<sup>28</sup>. This analysis, however, was carried out at the level of the population. It remains unclear whether individual dopamine neurons perform the same calculation or whether there is specialization that is overlooked when averaging across the population. Given the ability of individual dopamine neurons to influence diffuse downstream targets, as well as the debate over functional heterogeneity in the dopamine system, we performed a more fine-grained analysis of prediction error encoding at the single-neuron level.

We recorded from optogenetically identified dopamine neurons in the lateral VTA as mice performed classical conditioning tasks that varied reward size, expectation level or both. This allowed us to determine the full prediction error functions of individual dopamine neurons and assess how these functions related to each other. Rather than specialization, we found notable uniformity from neuron to neuron: when it comes to reward prediction errors, all dopamine neurons appeared to follow the same template, just scaled up or down. This scaled system greatly simplifies information coding, enabling prediction errors to be broadcasted robustly and consistently throughout the brain.

## RESULTS

### Dopamine neurons show unified reward response function

We recorded from the lateral VTA (Supplementary Fig. 1a) while mice performed a classical conditioning task with two interleaved trial types (Fig. 1a). On roughly half the trials, we delivered various sizes

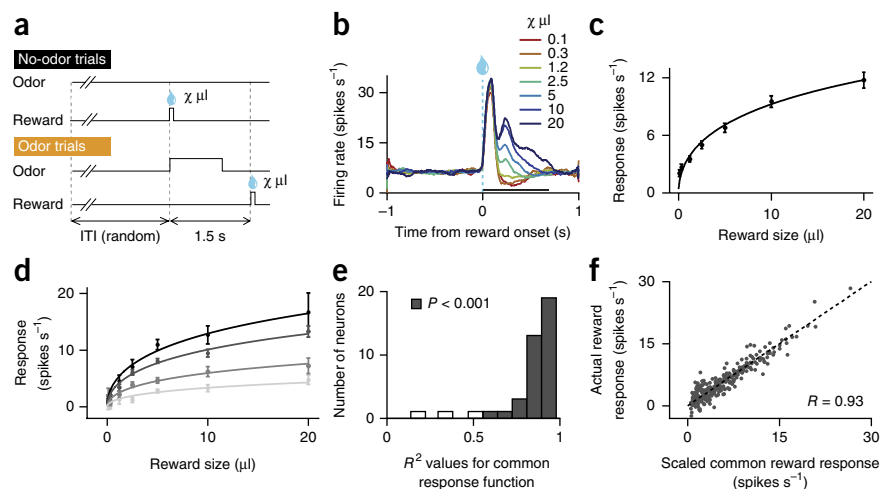
<sup>1</sup>Center for Brain Science, Department of Molecular and Cellular Biology, Harvard University, Cambridge, Massachusetts, USA. <sup>2</sup>MD-PhD Program, Harvard Medical School, Boston, Massachusetts, USA. Correspondence should be addressed to N.U. (uchida@mcb.harvard.edu).

Received 4 September 2015; accepted 7 January 2016; published online 8 February 2016; doi:10.1038/nn.4239

**Figure 1** Dopamine neurons share a common response function to unexpected rewards.

(a) Schematic of variable-reward task. Thirsty mice received various sizes of water reward ( $\chi$ ), ranging from 0.1 to 20  $\mu\text{l}$ . On no-odor trials, both the timing and size of reward were unpredictable. On odor trials, an odor cue predicted when reward would arrive, but the size remained unpredictable. Trials were randomly interleaved and inter-trial intervals (ITIs) were drawn from an exponential distribution, such that mice could not predict when the next trial would begin. (b) Firing rates of optogenetically identified dopamine neurons in response to different sizes of unexpected reward. (c) Average dopamine neuron responses (mean  $\pm$  s.e.m.) to different sizes of reward using a 600-ms window after reward onset. Line, best-fit Hill function (Online Methods).  $n = 40$  neurons. (d) Average responses of four example dopamine neurons (mean  $\pm$  s.e.m. across trials). Each neuron was fit by the same Hill function as in c, scaled by a single factor ( $\alpha$ ).

(e) Histogram of  $R^2$  values denoting how much of each neuron's variance in reward responses could be explained by a common response function, scaled by a single parameter. (f) Dopamine neuron responses predicted by a scaled common response function versus actual responses to each reward size. Each dot reflects the response of one identified dopamine neuron to one reward size. Dotted line, identity ( $y = x$ ).  $P = 4.6 \times 10^{-127}$ , Pearson's correlation.  $R$ , correlation coefficient.



of reward in the absence of any cue. On these trials, both the timing and size of reward were unexpected. On the other half of trials, an odor cue informed the mouse when reward would come, but the size was still unexpected. The light-gated ion channel rhodopsin (ChR2) was expressed selectively in dopamine neurons, which lets us identify neurons as dopaminergic on the basis of their responses to light<sup>23</sup>. A subset of this data was analyzed elsewhere<sup>28</sup>.

We recorded a total of 170 neurons, which clustered into three response types (Supplementary Fig. 2a–c). The largest cluster (84 neurons) showed phasic responses to reward; of these, 40 neurons were optogenetically identified as dopamine neurons (Supplementary Fig. 3a–g). We focused on these identified neurons, but our findings remained consistent when we included all 84 putative dopamine neurons (Supplementary Fig. 4).

We first examined how dopamine neurons respond to unexpected rewards of various sizes. Consistent with previous reports<sup>23,29</sup>, dopamine neurons monotonically increased their responses with increasing reward (Fig. 1b). After an initial peak unrelated to value, dopamine neurons showed a second peak that clearly distinguished reward size<sup>12</sup>. When averaged across the population, dopamine responses followed a simple saturating function (we used a Hill function for convenience, but see Online Methods for other potential fits; Fig. 1c). However, different neurons showed marked diversity in the size of their responses, ranging from 2 to almost 30 spikes  $\text{s}^{-1}$  above baseline for the largest reward. Are these neurons specialized for different reward sizes or do they all follow the same curve?

We found that, despite the diversity in firing rates, almost every individual neuron was well-described by the same curve, multiplied by a parameter  $\alpha$  (Fig. 1d). To analyze the similarity between neurons, we first fit a curve to the entire population data set. Then we determined the value of  $\alpha$  that, when multiplied by the average curve, best fit each neuron. This scaled function accounted for a large fraction of response variability across light-identified dopamine neurons (mean coefficient of determination,  $R^2 = 0.84$ ,  $n = 40$  neurons; Fig. 1e). Furthermore, a scatter-plot comparing observed reward responses with those predicted by a scaled function showed a high degree of correlation (Fig. 1f). Indeed, adding a parameter for  $x$  intercept (that is, the reward size threshold for eliciting a response) failed to improve the fits (Online Methods). Thus, although dopamine neurons displayed

different response magnitudes, they shared the same shape. One neuron's response function could be converted to another's simply by multiplying the curve by a single value.

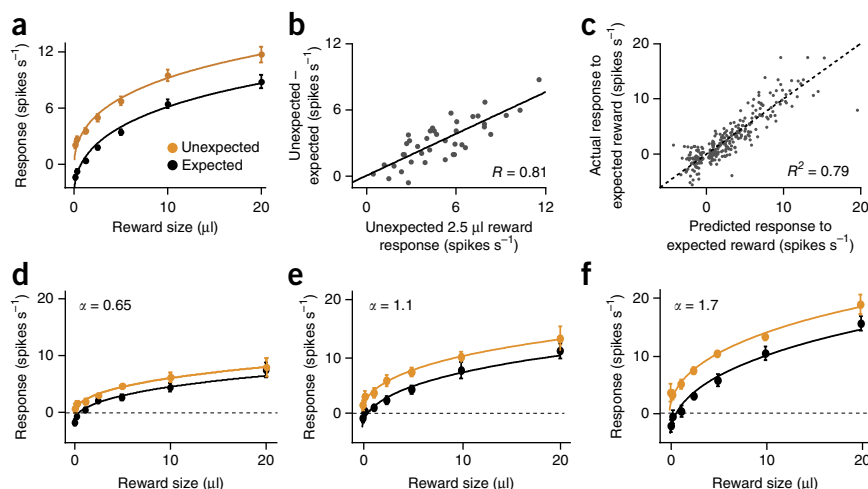
### Expectation causes scaled subtraction of dopamine response

How does expectation shift this curve? We designed our task so that the timing of reward was either expected (that is, after an odor cue) or unexpected (that is, in the absence of any cue). In both cases, reward size was varied randomly. In this way, we could assess how a single level of expectation modulated dopamine responses across a range of reward sizes. As we describe in more detail elsewhere<sup>28</sup>, expectation caused a subtractive shift in dopamine reward responses (Fig. 2a). Regardless of reward size, the odor cue caused a reduction of  $3.21 \pm 0.28$  spikes  $\text{s}^{-1}$  (mean  $\pm$  s.e.m. across neurons).

As with the reward response, however, there was considerable neuron-to-neuron variability in the amount of subtraction. Again, the question arises: are different dopamine neurons specialized for different levels of expectation or is there underlying homogeneity in their responses? We found that the variability in subtraction correlated with variability in reward response: those neurons that responded more to reward were also suppressed more by expectation (Fig. 2b). In other words, subtraction was scaled by a neuron's responsiveness. This correlation held for all reward sizes that we tested (Supplementary Fig. 5).

Together with the common response function (Fig. 1c), the existence of scaled subtraction suggests a notable feature of dopamine prediction error responses: they are simply scaled versions of each other. By averaging across the population, we derived two 'master' functions (Fig. 2a): one for unexpected reward and one for expected reward. The expected reward function is merely a subtractive transformation of the unexpected reward function. With these functions in hand, it is straightforward to recreate any individual neuron's activity. First, we used a neuron's responses to unexpected reward to determine  $\alpha$ , the factor that scales the neuron's responses to the average (Fig. 1d–f). Then we used the same  $\alpha$  to predict that neuron's response to expected reward. Specifically, we took the average expected reward function (Fig. 2a) and multiplied it by  $\alpha$ . Dopamine responses to expected reward, predicted in this way, corresponded well with the actual data (Fig. 2c–f). Thus, at least in a simple classical

**Figure 2** Prediction errors are calculated through scaled subtraction. (a) Average dopamine neuron responses (mean  $\pm$  s.e.m.) to different sizes of unexpected (orange circles) and expected (black circles) reward. Orange line, best-fit Hill function for unexpected reward (same as Fig. 1c). Black line, subtractive shift of the orange line.  $n = 40$  neurons. (b) Response to unexpected 2.5- $\mu$ l reward versus effect of expectation for this reward size. Line, best-fit linear regression.  $P = 3.5 \times 10^{-10}$ , Pearson's correlation.  $R$ , correlation coefficient. See **Supplementary Figure 5** for all other reward sizes. (c) Dopamine neuron responses predicted by a model of scaled subtraction versus actual responses to expected reward. Dotted line, identity ( $y = x$ ).  $P = 2.1 \times 10^{-96}$ , Pearson's correlation. (d–f) Average responses (mean  $\pm$  s.e.m.) of three example dopamine neurons with low (d), middle (e) and high (f) scaling factors ( $\alpha$ ). The same  $\alpha$  was used to produce the fits for both unexpected and expected reward.



conditioning task, there is marked uniformity in the way in which different dopamine neurons calculate reward prediction error.

### Linear suppression of responses across the population

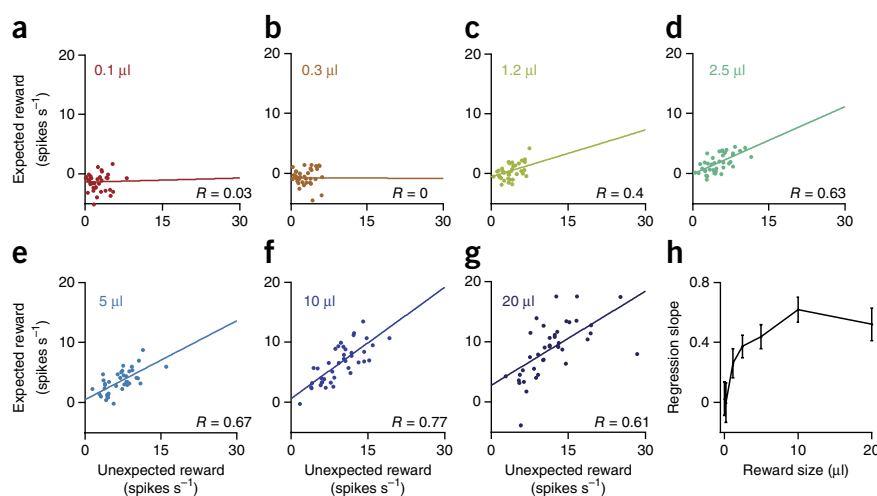
One consequence of this scaled system is that, for a given reward size, expectation should cause a linear suppression across neurons. Take, for example, the response to a 2.5- $\mu$ l reward. On average, expectation suppressed dopamine responses from 5 spikes  $s^{-1}$  to 1.8 spikes  $s^{-1}$ , roughly a 64% reduction. In the first example neuron (Fig. 2d), this corresponded to a decrease from 3.3 to 1.5 spikes  $s^{-1}$ ; the second neuron (Fig. 2e) decreased from 5.5 to 2 spikes  $s^{-1}$  and the third neuron (Fig. 2f) decreased from 7.5 to 3 spikes  $s^{-1}$ . The absolute number of spikes varied from neuron to neuron, but the proportional suppression stayed approximately the same. Thus, a scatter plot of unexpected reward versus expected reward showed a clear linear relationship (Fig. 3). We then focused on the slope of this relationship, which we believe has important implications for prediction error coding.

We defined parameter  $\beta$  as the proportional reduction of dopamine response when reward is expected. Unlike  $\alpha$ , which denotes an individual neuron's responsiveness and varies from neuron to neuron,  $\beta$  is the same across neurons. The value of  $\beta$ , however, changes with task conditions. For example, we saw that the slope changed with different reward sizes (Fig. 3h). What if we change the task?

To accurately encode reward prediction error signals, dopamine neurons should have larger proportional reductions with higher levels of expectation. In other words,  $\beta$  should scale with expectation. To test this hypothesis, we ran a separate experiment in which we held reward size constant, but varied expectation level (Fig. 4a). Mice performed a classical conditioning task with three odors: one predicting reward with 10% probability, the second predicting reward with 50% probability and the third predicting reward with 90% probability. Occasionally, we also delivered reward unexpectedly, in the absence of any odor. As before, we recorded in lateral VTA (Supplementary Fig. 1b) and used ChR2 to identify neurons as dopaminergic (Supplementary Fig. 3h–n).

Mice learned the task well. After the odor was presented, but before the reward was delivered, the mice displayed graded anticipatory licking, responding most for the 90% odor, less for 50% and even less for the 10% odor (Fig. 4b). Furthermore, higher reward probabilities caused reduced reward responses in dopamine neurons (Fig. 4c).

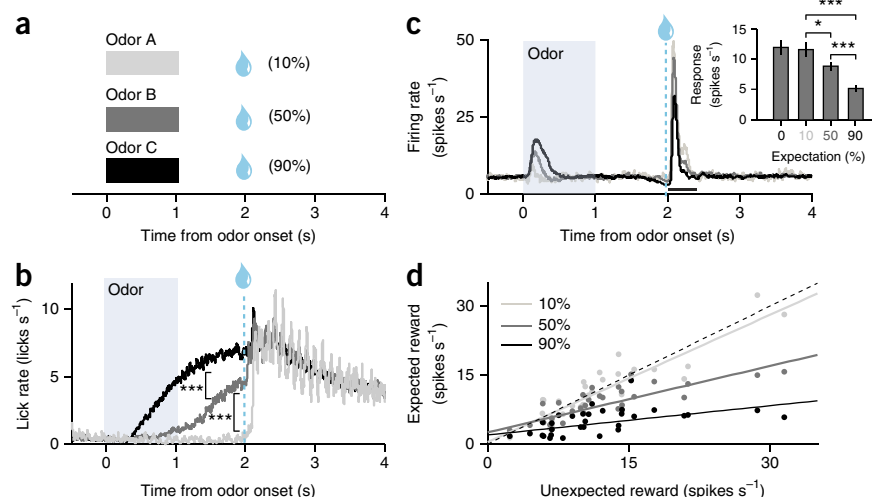
As predicted by our model of scaled subtraction, there was a systematic, linear effect of expectation across the population of dopamine neurons (Fig. 4d). Compared with completely unexpected reward, 10% expectation reduced firing by about 8% (Pearson's  $R = 0.89$ ,  $P = 6.9 \times 10^{-12}$ ), 50% expectation reduced firing by about 54% ( $R = 0.82$ ,  $P = 1.8 \times 10^{-8}$ ) and 90% expectation reduced firing by about 80% ( $R = 0.51$ ,  $P = 0.004$ ). Thus, each level of expectation suppressed responses by a particular proportion ( $\beta$ ), which stayed the same from neuron to neuron. This linear suppression is consistent with a common 'template', or response function, for prediction errors. The only difference between neurons is how much the template is scaled.



**Figure 3** Scaled subtraction produces linear suppression across the population. (a–g) Dopamine neuron responses to unexpected reward versus expected reward of various sizes ( $n = 40$  neurons). Pearson's correlation,  $P$  values: 0.83, 0.98, 0.01,  $1.4 \times 10^{-5}$ ,  $2.6 \times 10^{-6}$ ,  $7.8 \times 10^{-9}$  and  $2.7 \times 10^{-5}$  for a–g, respectively.  $R$ , correlation coefficients. (h) The regression slope ( $\pm$ s.e.) for each of the reward sizes plotted in a–g.

**Figure 4** Expectation level determines proportional suppression of dopamine responses. **(a)** Schematic of variable-expectation task. Odors predicted reward at 10%, 50% or 90% probability. The size and timing of reward never varied. **(b)** Average licking during the three trial types ( $n = 21$  behavioral sessions). After odor onset, but before reward delivery, mice displayed anticipatory licking corresponding to the level of expectation. \*\*\* $t(20) = 9.89$ ,  $P = 3.8 \times 10^{-9}$ ; and  $t(20) = 13.0$ ,  $P = 3.4 \times 10^{-11}$ ,  $t$  test.

**(c)** Average firing rates of identified dopamine neurons to different levels of expectation. Inset, average dopamine neuron responses in a 400-ms window after reward delivery (mean  $\pm$  s.e.m. across 31 neurons). \* $t(30) = 2.4$ ,  $P = 0.03$ ; \*\*\* $t(30) = 4.1$ ,  $P = 2.5 \times 10^{-4}$ ; and  $t(30) = 6.4$ ,  $P = 4.5 \times 10^{-7}$ ,  $t$  test. **(d)** Unexpected reward responses versus expected reward responses across the population of identified dopamine neurons ( $n = 31$ ). Reward probability differed from trial to trial, but reward size remained the same. Dotted line, identity ( $y = x$ ). Solid lines, best-fit regressions for each level of expectation. The slope of each line is  $1 - \beta$  (see text).



We have described two parameters,  $\alpha$  and  $\beta$ , which are orthogonal to each other. Each neuron has a different value of  $\alpha$  depending on its overall level of responsiveness. In fact, at least in these simple tasks, the value of  $\alpha$  is the only difference between neurons; otherwise their response functions are identical. Conversely, for a given task condition, all neurons share the same value of  $\beta$ . But as task conditions change (that is, by varying reward amounts or levels of expectation), then  $\beta$  changes accordingly. In other words,  $\alpha$  reflects the neuron and  $\beta$  reflects the task. Taken together, these two parameters fully describe the responses of individual dopamine neurons to simple conditioning tasks (for a mathematical description, please see **Supplementary Note**).

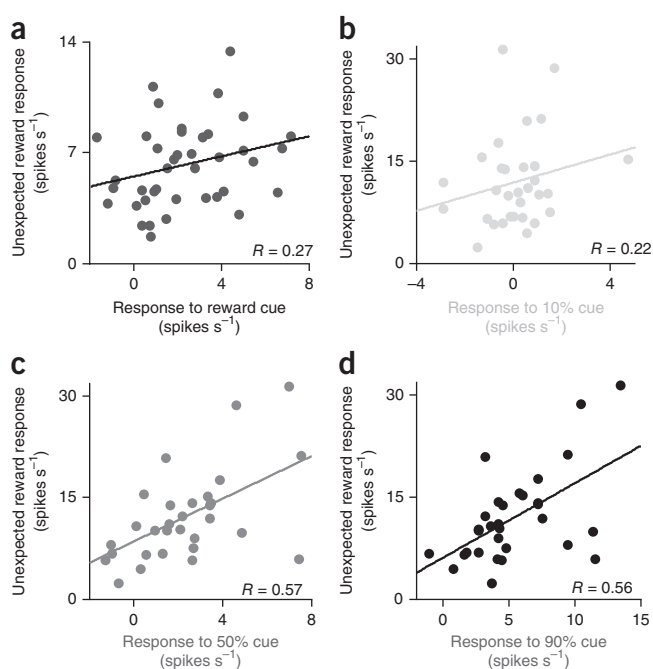
### Cue responses also scale with responsiveness

To develop our model for prediction error coding, we focused mostly on the reward response, comparing trials in which reward is expected to trials in which reward is unexpected. However, we also analyzed

dopamine neuron responses to conditioned stimuli. In our variable-reward experiment, dopamine neurons responded more to an odor associated with reward than to an odor associated with nothing ( $2.31 \pm 0.35$  spikes  $s^{-1}$  versus  $-0.38 \pm 0.16$  spikes  $s^{-1}$ , mean  $\pm$  s.e.m.;  $P = 3.1 \times 10^{-7}$ ,  $t$  test). Similarly, in the variable-expectation experiment, dopamine neurons responded most to the cue associated with 90% reward ( $5.38 \pm 0.61$  spikes  $s^{-1}$ ), less to the 50% cue ( $2.18 \pm 0.43$  spikes  $s^{-1}$ ) and least to the 10% cue ( $0.10 \pm 0.25$  spikes  $s^{-1}$ ;  $P = 6.5 \times 10^{-5}$ ,  $1.7 \times 10^{-8}$  and  $4.0 \times 10^{-9}$ , respectively,  $t$  test for all pairs). These results are consistent with previous findings that dopamine neurons are finely tuned to stimuli associated with reward, in addition to reward itself<sup>27,29</sup>. Notably, dopamine cue responses appeared to follow the scaled system that we described above, although the correlations were not substantial for cues predicting small reward. In general, those neurons with large cue responses also showed large reward responses (for the variable-reward experiment,  $P = 0.096$ ; for the variable-expectation experiment,  $P = 0.001$ ,  $7.4 \times 10^{-4}$ , and 0.24 for the 90%, 50%, and 10% cues, respectively; **Fig. 5**). Thus, prediction errors were scaled both for strong conditioned stimuli and for unconditioned stimuli, consistent with the notion that each neuron broadcasts the same signal.

### Dopamine neurons show high noise correlations

Up to this point, we averaged over trials to examine how dopamine neurons respond to different sizes of reward and levels of expectation. To understand how this information might be decoded by downstream structures in real time, it is also important to examine trial-by-trial correlations in activity. To do so, we took pairs of simultaneously recorded dopamine neurons and examined how they responded to repeated presentations of the same stimulus (that is, noise correlations). Low noise correlations imply that trial-by-trial changes in



**Figure 5** Cue response scales with reward response. **(a)** Results from the variable-reward experiment ( $n = 40$  dopamine neurons). Response to reward-predicting odor versus response to unexpected reward (averaged across all sizes).  $P = 0.096$ , Pearson's correlation.  $R$ , correlation coefficient. **(b–d)** Results from the variable-expectation experiment ( $n = 31$  dopamine neurons). Response to odors predicting 10% **(b)**, 50% **(c)** or 90% reward **(d)** versus response to unexpected reward.  $P = 0.24$ ,  $7 \times 10^{-4}$  and 0.001 for **b–d**, respectively, Pearson's correlation.

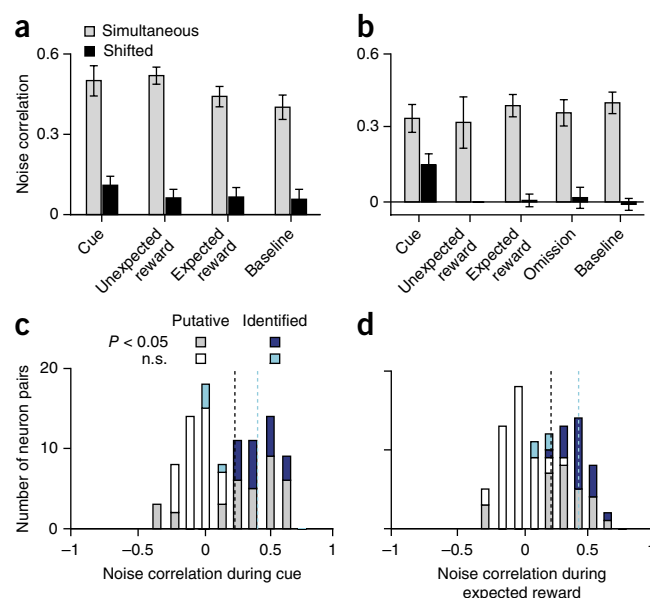


**Figure 6** Dopamine neurons show high noise correlations in every task epoch. (a,b) Noise correlations (mean  $\pm$  s.e.m.) between pairs of simultaneously recorded light-identified dopamine neurons in the variable-reward experiment (a,  $n = 11$  pairs) and the variable-expectation experiment (b,  $n = 12$  pairs). Correlations were calculated by examining trial-by-trial variations in spiking during different task epochs (Online Methods). Gray bars, correlations on simultaneous trials. Black bars, correlations in which one neuron's data was shifted by one trial. (c,d) Histograms of noise correlations between pairs of simultaneously recorded neurons that were either light-identified (cyan or dark blue) or putative (gray or white) dopamine neurons ( $n = 114$  pairs in total). Data are combined from both the variable-reward and variable-expectation experiments, and reflect correlations during the reward-predicting cue (c) and during delivery of expected reward (d). Gray or dark blue, significant noise correlation ( $P < 0.05$ , Pearson's correlation). White or cyan, n.s. ( $P > 0.05$ ). Black dotted lines, mean noise correlation for putative dopamine neuron pairs. Cyan dotted lines, mean noise correlation for light-identified dopamine neuron pairs.

firing rate are independent between neurons; thus, by pooling many dopamine neurons, a downstream neuron could cancel out the noise and decode the signal more effectively<sup>30,31</sup>. In contrast, high noise correlations imply that averaging over many neurons may not increase information content, as the noise cannot be canceled.

In the variable-reward experiment (Fig. 1), we collected 11 pairs of simultaneously recorded light-identified dopamine neurons. In the variable-expectation experiment (Fig. 4), we collected 12 pairs of light-identified neurons. To increase these numbers, we used an unsupervised clustering approach to identify putative dopamine neurons (Supplementary Fig. 2) and examined all pairs of simultaneously recorded putative dopamine neurons (31 pairs in the variable-reward experiment and 60 pairs in the variable-expectation experiment).

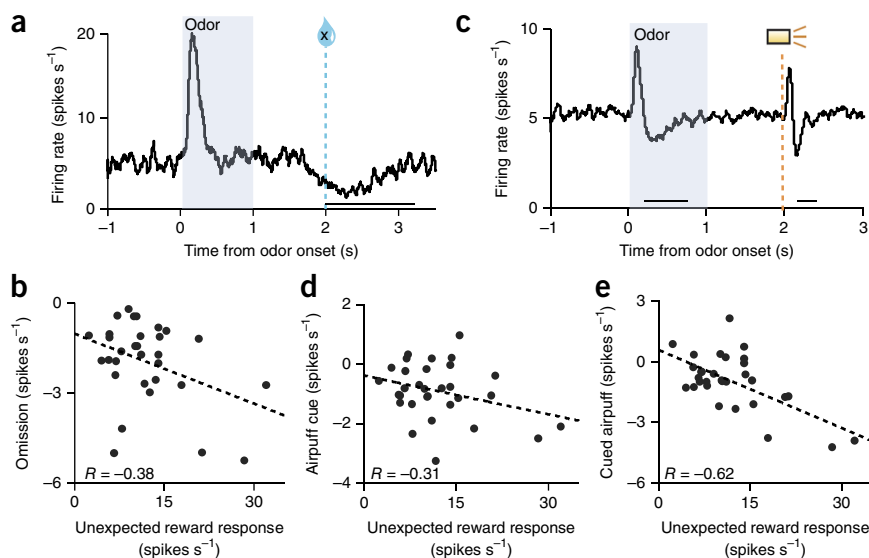
Consistent with previous results in putative dopamine neurons in substantia nigra<sup>8</sup> and a mixture of neuron types in VTA<sup>10</sup>, we found a high level of noise correlations in identified dopamine neurons (Fig. 6). These correlations were similar for both the variable-reward experiment (Fig. 6a) and the variable-expectation experiment (Fig. 6b), and for every experimental condition, including the conditioned stimulus, unexpected reward delivery, expected reward delivery and reward omission. Note that noise correlations were high for both responses above baseline (for example, the reward) and responses below baseline (for example, reward omission). Indeed, there were also high noise correlations in spontaneous baseline activity. Such correlations are



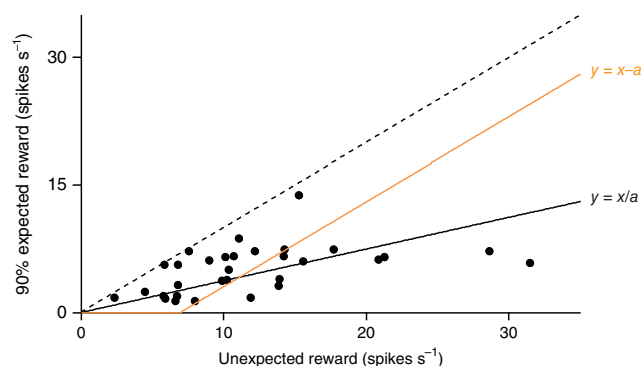
particularly impressive given low baseline firing rates, which can mask noise correlations<sup>31</sup>. Furthermore, the noise correlations essentially disappeared when we shifted the trials of one member of each pair, so that the pairs were offset by one trial (Fig. 6a,b). This implies that the correlations were indeed a result of task events, rather than long-timescale changes such as firing rate drift over the course of a session. The results also held when we included all putative dopamine neurons (Fig. 6c,d), and when we only included pairs of neurons on different tetrodes to eliminate any bias from spike sorting (data not shown). In contrast, pairs of putative dopamine and GABA neurons did not show substantial noise correlations (Supplementary Fig. 6). Our findings suggest that, in addition to homogeneity in mean firing rates, dopamine neurons are highly correlated even in their trial-by-trial fluctuations. This implies that downstream neurons would not benefit by averaging over many dopamine neurons to receive the prediction error signal.

### Responses to aversive events

Some of the dopamine neurons that we recorded showed weak responses to our stimuli; although they shared the same response



**Figure 7** Reward response correlates with response to aversive events. (a) Average firing of identified dopamine neurons ( $n = 31$  neurons) in the variable-expectation task during trials in which reward was omitted despite 90% expectation. (b) Response to unexpected reward versus response to reward omission. The omission response was averaged over a 1,300-ms window from the time of expected reward to include the entire duration of the effect.  $P = 0.04$ , Pearson's correlation.  $R$ , correlation coefficient. (c) Average firing of identified dopamine neurons during airpuff trials. An odor cue predicted airpuff with 90% probability. (d) Response to unexpected rewards versus response to airpuff cues, averaged from 200–800 ms after cue onset to include the full duration of inhibition, without the initial excitatory response.  $P = 0.09$ , Pearson's correlation. (e) Response to unexpected rewards versus response to cued airpuffs, averaged from 100–400 ms after airpuff onset to include the full duration of inhibition.  $P = 1.5 \times 10^{-4}$ , Pearson's correlation.



**Figure 8** Common response function allows even weakly responsive neurons to contribute. For identified dopamine neurons in the variable-expectation experiment, response to unexpected rewards versus 90% expected rewards (re-plotted from **Fig. 4d**). Black line, linear regression, which assumes that each neuron's response is divided by the same value. The slope of this line corresponds to  $1 - \beta$  (see text). Orange line, hypothetical scenario in which each neuron's response is subtracted by the same value, rather than divided. Dotted line, identity ( $y = x$ ).

function, their absolute firing rates were low. We wondered whether these neurons might be specialized for other stimuli, such as aversive stimuli. Thus, we examined the correlation between reward responses and responses to three presumably aversive events: omission of expected reward (**Fig. 7a,b**), prediction of an airpuff to the face (**Fig. 7c,d**) and airpuff itself (**Fig. 7c,e**). Note that, for the airpuff cue and airpuff itself, there was a biphasic response consisting of an initial excitatory response followed by a dip. The excitation may be a result of response generalization, as the sound of the airpuff valve resembled the sound of the reward valve<sup>12</sup>. We focused on the inhibitory phase, which encoded a more genuine value response. We found that neurons that were highly responsive to unexpected rewards tended to also be highly responsive to aversive events: they showed greater levels of suppression below baseline. This tendency is consistent with a recent study of putative dopamine neurons in monkeys that also used reward omission and airpuff, as well as bitter liquid<sup>20</sup>. Thus, it is not the case that some neurons specialize for aversive events; rather, highly responsive neurons in one setting appear to be highly responsive in other settings as well. Conversely, weakly reward-responsive units were likely to have weak suppressions or even net-positive reactions to airpuff cues and airpuff itself. The latter neurons, which responded positively to both rewarding and punishing events, resemble the 'salience-coding' neurons described previously<sup>21</sup>, although their lack of excitation to reward omission (**Fig. 7b**) may argue against this interpretation<sup>20</sup>.

## DISCUSSION

Here we sought to define the computations that dopamine neurons make during a simple behavior. By recording from optogenetically identified dopamine neurons in lateral VTA during two classical conditioning tasks, we discovered a common template for prediction errors. Almost every dopamine neuron that we recorded appeared to follow this template, responding in the same way to different sizes of reward and different levels of expectation. The only difference between neurons was in the scale of their responses. These results suggest that, at least in a classical conditioning task, most lateral VTA dopamine neurons encode the same information, giving them the capacity to broadcast a single important value: prediction error.

We previously found that, at the population level, VTA dopamine neurons calculate prediction error through subtraction<sup>28</sup>. Here we

extended that analysis by examining individual dopamine neuron responses and assessing the extent of specialization versus homogeneity. We found a notable amount of the latter. Just two parameters could accurately describe prediction errors at both the individual and population levels (**Supplementary Note**). We use  $\alpha$  to denote a neuron's reward responsiveness: the higher a neuron's  $\alpha$ , the more it is excited by rewards and suppressed by reward expectation (**Fig. 2**). This parameter is likely intrinsic to the neuron and is the only factor that distinguishes one neuron from another. In contrast,  $\beta$  (**Figs. 2b** and **3**) is an orthogonal variable that determines the proportion of a dopamine neuron's response that will be suppressed by expectation. This parameter is the same for every neuron and depends on the behavioral context, including reward expectation (**Fig. 4d**) and reward size (**Fig. 3**). In other settings, it is likely that  $\beta$  would depend on other factors that affect expectation, such as delay time or learning. Knowing a neuron's  $\alpha$ , as well as the  $\beta$  determined by task conditions, makes it possible to predict that neuron's firing at most times in the task.

## Population coding of prediction errors

Such a compact system, in which each neuron is a scaled version of each other, may increase the robustness of prediction error coding. For example, when we presented an odor that predicted reward with 90 percent probability, dopamine neuron reward responses were suppressed by about 80 percent (**Fig. 8**). Averaging across the population, this corresponded to a decrease of about 7 spikes  $s^{-1}$ . Most neurons, however, did not subtract exactly 7 spikes  $s^{-1}$ . If each neuron had subtracted that amount, the result would resemble the orange line in **Figure 8**. In this case, many neurons could not contribute: since they fired less than 7 spikes  $s^{-1}$  for unexpected reward, subtracting that amount would make them hit a floor. By scaling the subtracted amount to a neuron's responsiveness, this system allows even weakly responsive units to contribute to prediction errors. Indeed, note that each neuron shown fell below the identity line (**Fig. 8**), implying that even neurons that responded minimally to unexpected reward were suppressed when reward was expected. In other words, each neuron calculated prediction error, even if its responses were weak.

The scaling that we discovered here ensures that each neuron's possible output is matched with the input it receives. This may be an efficient way to ensure consistent prediction errors. For example, matching inputs to outputs is probably less energy-intensive than enforcing identical firing rates in every neuron. Furthermore, the presence of weakly responding neurons is important: when expectation is large, strongly responding neurons can eventually hit a floor, whereas weakly responding neurons have more space to be suppressed, as a given expectation level suppresses only a small fraction of their spontaneous firing rate.

The redundancy in this system may also be adaptive, given the proposed importance of dopamine prediction errors for both learning and response vigor<sup>32</sup>. No single dopamine neuron can innervate all distal targets; likewise, no target neuron can receive inputs from all dopamine neurons. Redundancy allows the downstream neuron to decode a signal approximating the 'true' dopamine signal, despite inputs from only one or a small number of dopamine neurons. This reduces the requirement for precise connectivity, at least for the interpretation of reward prediction error.

It is thought that noise correlations arise from common inputs and can have detrimental effects on the ability of an ensemble of neurons to jointly convey their signals (population coding)<sup>33</sup>. We found that dopamine neurons exhibited extremely high noise correlations. Our data is consistent with the idea that different dopamine neurons

receive similar, common inputs. These results suggest that dopamine neurons need not jointly signal their outputs; rather, each neuron encodes a more or less complete prediction error signal. This mode of signaling may be particularly sensible if dopamine-recipient synapses rely on a small number of dopamine inputs, or even only one dopamine input<sup>34,35</sup>, rather than relying on many dopamine inputs.

### Homeostatic balance of excitation and inhibition

It is interesting to speculate how a system of scaled inhibition might arise. Previous studies, which focused mostly on cortex and hippocampus, have demonstrated an exquisite balance between excitation and inhibition<sup>36,37</sup>. Over time, neurons that are more active begin to receive more inhibition, thereby ensuring overall network stability. Such homeostasis is achieved through multiple mechanisms, including changes in synaptic strength, intrinsic excitability and synapse stabilization. Our results complement this literature, as we found that those dopamine neurons with higher spontaneous and reward-evoked firing rates also exhibited more suppression by expectation. In other words, inhibition was matched to excitation. We believe this result could naturally emerge from homeostatic processes that vary inhibitory synapse number or GABA receptor expression on the dopamine neurons, among other circuit features. Notably, despite these differences in excitatory or inhibitory strength (reflected by our parameter  $\alpha$ ), each neuron seemed to respond in the same way when task demands changed (reflected by our parameter  $\beta$ ), implying commonality of inputs.

Although homeostatic processes likely have a role over a developmental timescale, the balance between excitation and inhibition may also be vital over the course of a single trial<sup>38,39</sup>. For example, a recent study found hints of such a balance in the multiphasic responses of putative dopamine neurons<sup>38</sup>. By analyzing rebound excitation and inhibition, the researchers argued that dopamine neurons undergo tonic, homeostatic inhibition that varies with the magnitude of excitatory responses. Our results extend this argument to explain how such a balance might ensure fidelity of downstream signaling.

Our findings are also consistent with recent discoveries of one-dimensional dynamics in cortical circuits<sup>40,41</sup>. In the lateral intraparietal area, for example, neurons differed in their responses to three task epochs, but in a stereotyped way<sup>40</sup>. Knowing a neuron's response to one epoch allowed the experimenters to predict the responses to the other two. In other words, just as in our data, different neurons were scaled versions of each other. Such a one-dimensional system has been proposed to facilitate downstream decoding, thereby enhancing behavioral reliability.

### Diversity may come in other forms

Our finding of homogeneity among dopamine neurons is consistent with recent work showing a continuum of dopamine responses to both positive and negative stimuli, without any clear clusters<sup>20</sup>. But our results do not rule out diversity in the system<sup>3,13</sup>. First, we targeted a circumscribed region of the lateral VTA (Supplementary Fig. 1); it is possible that neurons in the substantia nigra or medial VTA would show different responses. Note, however, that in our recording region, we did not find any aspect of neural activity that varied with location. Second, our tasks focused primarily on responses to unexpected and expected water rewards. Dopamine neurons might show more diverse responses to different tasks, for example, involving working memory<sup>42</sup> or a broader array of aversive or salient stimuli<sup>20,21,43</sup>. The homogeneity may simply be explained by the computational simplicity of the classical conditioning task. Indeed, with increasing task complexity, more dopamine neurons might be released from tonic inhibition and

come on-line, changing the overall population response<sup>44</sup>. Third, our observation that responses of individual dopamine neurons follow a common template is consistent with recent anatomical observations that dopamine neurons projecting to different targets receive relatively similar sets of monosynaptic inputs<sup>45–47</sup>. However, a larger difference in responses might have arisen if we had targeted our analysis at dopamine neurons projecting to the tail of the striatum, which show greater differences in patterns of input<sup>45</sup>. Finally, even if the dopamine neuron firing rates are homogeneous, downstream neurons may read out this signal in diverging ways. For example, there may be different time courses of dopamine release<sup>48</sup>, different expression patterns of dopamine receptors and transporters<sup>19</sup>, or different co-released transmitters<sup>49,50</sup>, depending on the target region.

Dopamine neurons, in other words, may encode more than just reward prediction error. But when they do encode reward prediction error, they do so with marked consistency.

### METHODS

Methods and any associated references are available in the [online version of the paper](#).

*Note: Any Supplementary Information and Source Data files are available in the online version of the paper.*

### ACKNOWLEDGMENTS

We thank J. Fitzgerald for assistance with analysis, J. Assad, R. Born, J. Maunsell, R. Wilson and members of the Uchida laboratory for discussions, C. Dulac for sharing resources, and K. Deisseroth (Stanford University) for the AAV-FLEX-ChR2 construct. This work was supported by a Sackler Fellowship in Psychobiology (N.E.) and US National Institutes of Health grants T32GM007753 (N.E.), F30MH100729 (N.E.), 2T32MH020017-16 (M.B.), 5T32MH020017-17 (M.B.), R01MH095953 (N.U.) and R01MH101207 (N.U.).

### AUTHOR CONTRIBUTIONS

N.E., J.T. and N.U. designed the experiments. N.E. and M.B. collected data for the variable-reward task. J.T. collected data for the variable-expectation task. N.E. analyzed data and wrote the manuscript, with comments from J.T. and N.U.

### COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

1. Wise, R.A. Dopamine, learning and motivation. *Nat. Rev. Neurosci.* **5**, 483–494 (2004).
2. Salamone, J.D. & Correa, M. The mysterious motivational functions of mesolimbic dopamine. *Neuron* **76**, 470–485 (2012).
3. Bromberg-Martin, E.S., Matsumoto, M. & Hikosaka, O. Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron* **68**, 815–834 (2010).
4. Matsuda, W. *et al.* Single nigrostriatal dopaminergic neurons form widely spread and highly dense axonal arborizations in the neostriatum. *J. Neurosci.* **29**, 444–453 (2009).
5. Fuxe, K. *et al.* The discovery of central monoamine neurons gave volume transmission to the wired brain. *Prog. Neurobiol.* **90**, 82–100 (2010).
6. Grace, A.A. & Bunney, B.S. Intracellular and extracellular electrophysiology of nigral dopaminergic neurons. 1. Identification and characterization. *Neuroscience* **10**, 301–315 (1983).
7. Vandecasteele, M., Glowinski, J. & Venance, L. Electrical synapses between dopaminergic neurons of the substantia nigra pars compacta. *J. Neurosci.* **25**, 291–298 (2005).
8. Joshua, M. *et al.* Synchronization of midbrain dopaminergic neurons is enhanced by rewarding events. *Neuron* **62**, 695–704 (2009).
9. Morris, G., Arkadir, D., Nevet, A., Vaadia, E. & Bergman, H. Coincident but distinct messages of midbrain dopamine and striatal tonically active neurons. *Neuron* **43**, 133–143 (2004).
10. Kim, Y., Wood, J. & Moghaddam, B. Coordinated activity of ventral tegmental neurons adapts to appetitive and aversive learning. *PLoS One* **7**, e29766 (2012).
11. Schultz, W. Predictive reward signal of dopamine neurons. *J. Neurophysiol.* **80**, 1–27 (1998).
12. Schultz, W. Updating dopamine reward signals. *Curr. Opin. Neurobiol.* **23**, 229–238 (2013).
13. Roeper, J. Dissecting the diversity of midbrain dopamine neurons. *Trends Neurosci.* **36**, 336–342 (2013).

14. Blaess, S. *et al.* Temporal-spatial changes in Sonic Hedgehog expression and signaling reveal different potentials of ventral mesencephalic progenitors to populate distinct ventral midbrain nuclei. *Neural Dev.* **6**, 29 (2011).
15. Margolis, E.B., Lock, H., Hjelmstad, G.O. & Fields, H.L. The ventral tegmental area revisited: is there an electrophysiological marker for dopaminergic neurons? *J. Physiol. (Lond.)* **577**, 907–924 (2006).
16. Neuhoff, H., Neu, A., Liss, B. & Roeper, J. I(h) channels contribute to the different functional properties of identified dopaminergic subpopulations in the midbrain. *J. Neurosci.* **22**, 1290–1302 (2002).
17. Lammel, S. *et al.* Input-specific control of reward and aversion in the ventral tegmental area. *Nature* **491**, 212–217 (2012).
18. Watabe-Uchida, M., Zhu, L., Ogawa, S.K., Vamanrao, A. & Uchida, N. Whole-brain mapping of direct inputs to midbrain dopamine neurons. *Neuron* **74**, 858–873 (2012).
19. Lammel, S. *et al.* Unique properties of mesoprefrontal neurons within a dual mesocorticolimbic dopamine system. *Neuron* **57**, 760–773 (2008).
20. Fiorillo, C.D., Yun, S.R. & Song, M.R. Diversity and homogeneity in responses of midbrain dopamine neurons. *J. Neurosci.* **33**, 4693–4709 (2013).
21. Matsumoto, M. & Hikosaka, O. Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* **459**, 837–841 (2009).
22. Bayer, H.M. & Glimcher, P.W. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47**, 129–141 (2005).
23. Cohen, J.Y., Haesler, S., Vogt, L., Lowell, B.B. & Uchida, N. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* **482**, 85–88 (2012).
24. Pan, W.-X., Schmidt, R., Wickens, J.R. & Hyland, B.I. Dopamine cells respond to predicted events during classical conditioning: evidence for eligibility traces in the reward-learning network. *J. Neurosci.* **25**, 6235–6242 (2005).
25. D'Ardenne, K., McClure, S.M., Nystrom, L.E. & Cohen, J.D. BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science* **319**, 1264–1267 (2008).
26. Waelti, P., Dickinson, A. & Schultz, W. Dopamine responses comply with basic assumptions of formal learning theory. *Nature* **412**, 43–48 (2001).
27. Schultz, W., Dayan, P. & Montague, P.R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
28. Eshel, N. *et al.* Arithmetic and local circuitry underlying dopamine prediction errors. *Nature* **525**, 243–246 (2015).
29. Tobler, P.N., Fiorillo, C.D. & Schultz, W. Adaptive coding of reward value by dopamine neurons. *Science* **307**, 1642–1645 (2005).
30. Averbeck, B.B., Latham, P.E. & Pouget, A. Neural correlations, population coding and computation. *Nat. Rev. Neurosci.* **7**, 358–366 (2006).
31. Cohen, M.R. & Kohn, A. Measuring and interpreting neuronal correlations. *Nat. Neurosci.* **14**, 811–819 (2011).
32. Dayan, P. Twenty-five lessons from computational neuromodulation. *Neuron* **76**, 240–256 (2012).
33. Abbott, L.F. & Dayan, P. The effect of correlated variability on the accuracy of a population code. *Neural Comput.* **11**, 91–101 (1999).
34. Freund, T.F., Powell, J.F. & Smith, A.D. Tyrosine hydroxylase-immunoreactive boutons in synaptic contact with identified striatonigral neurons, with particular reference to dendritic spines. *Neuroscience* **13**, 1189–1215 (1984).
35. Zahm, D.S. An electron microscopic morphometric comparison of tyrosine hydroxylase immunoreactive innervation in the neostriatum and the nucleus accumbens core and shell. *Brain Res.* **575**, 341–346 (1992).
36. Turrigiano, G. Homeostatic signaling: the positive side of negative feedback. *Curr. Opin. Neurobiol.* **17**, 318–324 (2007).
37. Davis, G.W. Homeostatic control of neural activity: from phenomenology to molecular design. *Annu. Rev. Neurosci.* **29**, 307–323 (2006).
38. Fiorillo, C.D., Song, M.R. & Yun, S.R. Multiphasic temporal dynamics in responses of midbrain dopamine neurons to appetitive and aversive stimuli. *J. Neurosci.* **33**, 4710–4725 (2013).
39. Fiorillo, C.D. Towards a general theory of neural computation based on prediction by single neurons. *PLoS One* **3**, e3298 (2008).
40. Ganguli, S. *et al.* One-dimensional dynamics of attention and decision making in LIP. *Neuron* **58**, 15–25 (2008).
41. Fitzgerald, J.K. *et al.* Biased associative representations in parietal cortex. *Neuron* **77**, 180–191 (2013).
42. Matsumoto, M. & Takada, M. Distinct representations of cognitive and motivational signals in midbrain dopamine neurons. *Neuron* **79**, 1011–1024 (2013).
43. Zweifel, L.S. *et al.* Activation of dopamine neurons is critical for aversive conditioning and prevention of generalized anxiety. *Nat. Neurosci.* **14**, 620–626 (2011).
44. Grace, A.A., Floresco, S.B., Goto, Y. & Lodge, D.J. Regulation of firing of dopaminergic neurons and control of goal-directed behaviors. *Trends Neurosci.* **30**, 220–227 (2007).
45. Menegas, W. *et al.* Dopamine neurons projecting to the posterior striatum form an anatomically distinct subclass. *eLife* **4**, e10032 (2015).
46. Lerner, T.N. *et al.* Intact-brain analyses reveal distinct information carried by SNc dopamine subcircuits. *Cell* **162**, 635–647 (2015).
47. Beier, K.T. *et al.* Circuit architecture of VTA dopamine neurons revealed by systematic input-output mapping. *Cell* **162**, 622–634 (2015).
48. Zhang, L., Doyon, W.M., Clark, J.J., Phillips, P.E.M. & Dani, J.A. Controls of tonic and phasic dopamine transmission in the dorsal and ventral striatum. *Mol. Pharmacol.* **76**, 396–404 (2009).
49. Tritsch, N.X., Ding, J.B. & Sabatini, B.L. Dopaminergic neurons inhibit striatal output through non-canonical release of GABA. *Nature* **490**, 262–266 (2012).
50. Stuber, G.D., Hnasko, T.S., Britt, J.P., Edwards, R.H. & Bonci, A. Dopaminergic terminals in the nucleus accumbens but not the dorsal striatum corelease glutamate. *J. Neurosci.* **30**, 8229–8233 (2010).



## ONLINE METHODS

**Animals.** We used ten adult male mice, backcrossed for >5 generations with C57/BL6J mice, that were heterozygous for Cre recombinase under the control of the DAT gene (B6.SJL-*Slc6a3<sup>tm1.1(cre)Bkmn</sup>/J*, The Jackson Laboratory)<sup>51</sup>. Five animals were used in the variable-reward task (Fig. 1a) and five in the variable-expectation task (Fig. 4a). Animals were housed on a 12-h dark/12-h light cycle (dark from 07:00 to 19:00), one to a cage, and performed the task at the same time each day. All experiments were performed in accordance with the US National Institutes of Health Guide for the Care and Use of Laboratory Animals and approved by the Harvard Institutional Animal Care and Use Committee.

**Surgery and viral injections.** All surgeries were performed under aseptic conditions with animals under either ketamine/medetomidine (60 and 0.5 mg kg<sup>-1</sup>, intraperitoneal, respectively) or isoflurane (1–2% at 0.5–1.0 l min<sup>-1</sup>) anesthesia. Analgesia (ketoprofen, 5 mg kg<sup>-1</sup> intraperitoneal; buprenorphine, 0.1 mg kg<sup>-1</sup>, intraperitoneal) was administered postoperatively. Mice underwent two surgeries, both stereotactically targeting left VTA (from bregma: 3.0 mm posterior, 0.8 mm lateral, 4–5 mm ventral). In the first surgery, we injected 200–500 nl adeno-associated virus (AAV, serotype 5) carrying an inverted ChR2 (H134R) fused to the fluorescent reporter eYFP and flanked by double *loxP* sites<sup>52,53</sup>. We previously showed that expression of this virus in dopamine neurons is highly selective and efficient<sup>23</sup>. After 2–4 weeks, we performed a second surgery to implant a head plate and custom-built microdrive containing 6–8 tetrodes and an optical fiber, as described<sup>23</sup>. Recording sites are displayed in **Supplementary Figure 1**. There was no systematic difference in dopamine responses as a function of recording location.

**Behavioral procedure.** After more than 1 week of recovery, mice were water-restricted in their cages. Weight was maintained above 90% of baseline body weight. Animals were head-restrained and habituated for 1–2 d before training. Odors were delivered with a custom-made olfactometer<sup>54</sup>. Each odor was dissolved in mineral oil at 1:10 or 1:100 dilution. 30 µl of diluted odor was placed inside a filter-paper housing, and then further diluted with filtered air by 1:20 to produce a 1,000 ml min<sup>-1</sup> total flow rate. Odors included isoamyl acetate, (+)-carvone, 1-hexanol, p-cymene, ethyl butyrate and 1-butanol, and differed for different animals. Licks were detected by breaks of an infrared beam placed in front of the water tube.

Each trial began with 1-s odor delivery, followed by a short delay (1 s in the variable-expectation task and 0.5 s in the variable-reward task) and an outcome (3.75 µl water for the variable-expectation task, or an amount ranging from 0.1–20 µl in the variable-reward task). Inter-trial intervals were drawn from an exponential distribution (mean: 7.6 s), resulting in a flat hazard function such that mice had constant expectation of when the next trial would begin. The tasks were purely classical conditioning: the licking behavior of the mice had no effect on whether water was delivered.

The variable-reward task (Fig. 1a) included three trial types, randomly inter-mixed. In trial type 1 (45% of all trials), an odor was delivered for 1 s, followed by a 0.5-s delay and a reward chosen pseudorandomly from the following set: 0.1, 0.3, 1.2, 2.5, 5, 10 or 20 µl. The frequency of each reward size was chosen to make the average reward approximately 5 µl. Reward sizes were determined by the length of time the water valve remained open: 4, 12, 25, 45, 75, 140 or 250 ms, respectively. In trial type 2 (45% of all trials), rewards of various sizes were delivered without any preceding odor. The reward sizes were identical to trial type 1. In these trials, the reward itself was considered the start of the trial, to ensure a flat hazard function. Comparing trial types 1 and 2 allowed us to determine how a constant level of expectation modulated responses to different sizes of reward. In trial type 3 (10% of all trials), a different odor was delivered, which was followed by no outcome. This trial type was included to ensure that the animals learned the task: they began to lick after the odor in trial type 1 but not after the odor in trial type 3 (data not shown). Animals performed between 300 and 500 trials per session.

In the variable-expectation task (Fig. 4), each trial began with one of four odors, selected pseudorandomly. One odor predicted reward with 10% probability, another predicted water reward with 50% probability, a third predicted water reward with 90% probability, and a fourth predicted an air puff to the

animal's face with 90% probability. In addition to the odor trials, fully unexpected rewards (in the absence of any odor) were also delivered on about 5% of trials in each session. Animals performed between 400 and 700 trials per day. The full data set for this experiment was also analyzed for another manuscript<sup>55</sup>.

**Electrophysiology.** Recording techniques were based on a previous study<sup>23</sup>. Briefly, we recorded extracellularly from VTA using a custom-built, screw-driven microdrive containing six or eight tetrodes (Sandvik) glued to a 200-µm optic fiber (ThorLabs). Tetrodes were affixed to the fiber so that their tips extended 350–600 µm from the end of the fiber. Neural and behavioral signals were recorded with a DigiLynx recording system (Neuralynx) or a custom-built system using a multi-channel amplifier chip (RHA2116, Intan Technologies LLC) and data acquisition device (PCIe-6351, National Instruments). Broadband signals from each wire were filtered between 0.1 and 9,000 Hz and recorded continuously at 32 kHz. To extract spike timing, signals were band-pass-filtered between 300 and 6000 Hz and sorted offline using SpikeSort3D (Neuralynx) or MClust-3.5 (A.D. Redish). At the end of each session, the fiber and tetrodes were lowered by 40–80 µm to record new units the next day.

To be included in the data set, a neuron had to be well-isolated (L ratio<sup>56</sup> < 0.05) and recorded within 0.5 mm of a light-identified dopamine neuron (see below), to ensure that it was recorded in VTA. Recording sites were also verified histologically with electrolytic lesions using 10–15 s of 30-µA direct current.

To identify neurons as dopaminergic, we used ChR2 to observe laser-triggered spikes<sup>23,57,58</sup>. The optical fiber was coupled with a diode-pumped solid-state laser with analog amplitude modulation (Laserglow Technologies). At the beginning and end of each recording session, we delivered trains of ten blue (473 nm) light pulses, each 5 ms long, at 1, 10, 20 and 50 Hz, with an intensity of 5–20 mW mm<sup>-2</sup> at the tip of the fiber. Spike shape was measured using a broadband signal (0.1–9,000 Hz) sampled at 32 kHz.

**Data analysis.** Peristimulus time histograms (PSTHs) were constructed using 1-ms bins and then convolved with a function resembling a postsynaptic potential:  $(1 - e^{-t}) * e^{-t/20}$  for time  $t$  in ms. Average firing rates in response to reward were calculated using a 600-ms window after reward onset for the variable-reward experiment and a 400-ms window after reward onset for the variable-expectation experiment. These windows were chosen to reflect the full duration of the neural response to reward, as determined by inspection of the average PSTH. Window sizes ranging from 300–1,000 ms were attempted and gave qualitatively similar results (data not shown). The reward response was longer for the variable-reward experiment likely because the reward itself took longer to deliver (that is, the valve was open for longer). Similarly, responses to reward-predicting odors were calculated using a 500-ms window, which was the duration of the population response. When we wished to determine the correlation between excitatory and inhibitory responses (Fig. 7), we did not include the initial excitatory peak in response to aversive stimuli; rather, we chose windows that reflected the full duration of the inhibition: 200–800 ms after the onset of an airpuff-predicting cue, and 100–400 ms after the onset of an airpuff itself. All responses were baseline-subtracted, with baseline defined as the 1 s before trial onset.

To identify neurons as dopaminergic, we used the Stimulus-Associated spike Latency Test (SALT<sup>58</sup>) to determine whether light pulses significantly changed a neuron's spike timing (Supplementary Fig. 3). We used a significance value of  $P < 0.001$ . To ensure that spike sorting was not contaminated by light artifacts, we also calculated waveform correlations between spontaneous and light-evoked spikes, as described<sup>23</sup>. All light-identified dopamine neurons had Pearson's correlation coefficients > 0.9.

In the variable-reward experiment, we identified putative dopamine neurons based on their firing patterns through an unsupervised clustering approach (Supplementary Fig. 2), similar to a previous study<sup>23</sup>. Briefly, receiver-operating characteristic (ROC) curves for each neuron were calculated by comparing the distribution of firing rates across trials in 100-ms bins (starting 1 s before expected reward and ending 1 s after expected reward) to the distribution of baseline firing rates (1 s before trial onset). PCA was calculated using the singular value decomposition of the area under the ROC. Hierarchical clustering was then done using the first three principal components of the auROC using a Euclidean distance metric and complete agglomeration method. As described<sup>23</sup>,

this method produced three clusters: one with phasic excitation to reward (type 1), one with sustained excitation to reward expectation (type 2), and one with sustained suppression to reward expectation (type 3). Type 1 neurons were classified as putatively dopaminergic and included in the paper. 40 out of 43 light-identified dopamine neurons fell into this cluster. The other three light-identified dopamine neurons showed phasic suppression to reward and were clustered as type 3. These three dopamine neurons showed qualitatively different responses than the others, and were not included in the data set. In previous histology experiments, we found very high specificity, that is, all ChR2-expressing neurons also expressed the dopamine marker tyrosine hydroxylase<sup>23</sup>. Thus, we expect that these three neurons were indeed dopaminergic, but might have a different function. However, we cannot completely rule out the possibility that these neurons were not dopaminergic.

As we describe in more detail elsewhere<sup>23,28</sup>, type 2 neurons are GABA-ergic and appear to provide a reward expectation signal that dopamine neurons use to calculate prediction error. Type 3 neurons are rarer, and their cell type and function remain unclear. Their responses are essentially mirror images of type 2 neurons, with greater suppressions for cues predicting larger rewards<sup>23</sup>, implying that they may also relay an expectation signal. Further work is needed to determine whether the relative dynamics of type 2 and type 3 neurons contribute to dopamine prediction error responses.

We focus our analysis on light-identified dopamine neurons ( $n = 40$  for the variable-reward task;  $n = 31$  for the variable-expectation task). However, we also examined responses of putative dopamine neurons ( $n = 44$  for the variable-reward task;  $n = 28$  for the variable-expectation task). Results were qualitatively similar for both light-identified and putative dopamine neurons (for example, see **Supplementary Fig. 4**).

To examine the dose-response of dopamine neurons (**Fig. 1c**), we based our analysis on a previous study<sup>59</sup>. We first fit a hyperbolic ratio function (Hill function) to the unexpected reward data, where  $r$  denotes reward size

$$f(r) = f_{\max} \left( \frac{r^{0.5}}{r^{0.5} + \sigma^{0.5}} \right) \quad (1)$$

The function had two free parameters:  $f_{\max}$ , the saturating firing rate; and  $\sigma$ , the reward size that elicits half-maximum firing rate. We chose an exponent of 0.5 after fitting the data with exponents ranging from 0.1 to 2.0 (in steps of 0.1), and finding the exponent with the lowest mean squared error. Note that the Hill function is not the only possible function that could fit our data. For example, the power function  $f(r) = ar^k$ , where  $a = 3.73$  and  $k = 0.39$ , also did an excellent job, but this function does not saturate, and therefore we thought it was less likely to represent neuronal responses. Two-parameter exponential and log functions did not fit the data well, and adding a third parameter for  $x$  intercept did not noticeably improve the fit for any of the functions we attempted (mean-squared error of 10.31 for the two-parameter Hill function versus 10.20 for the three-parameter Hill function; 10.24 for the two-parameter power function versus 10.21 for the three-parameter power function;  $P = 0.09$  and  $0.38$ , respectively, using bootstrapping to compare the models). The conclusions of this manuscript do not depend on the exact function chosen to fit the data. In particular, for each function, subtraction was the best fit.

To explore how individual neurons differ from this average function (**Fig. 1d**), we kept the  $f_{\max}$  and  $\sigma$  that we fit to the population and found the parameter  $\alpha_i$  that best scaled this function to each neuron's reward responses (using the least-squares method).

$$f_i(r) = \alpha_i \times f_{\max} \left( \frac{r^{0.5}}{r^{0.5} + \sigma^{0.5}} \right) \quad (2)$$

We used this scaled function to predict how each neuron would respond to each size of reward, and then compared the predictions to the actual data (**Fig. 1e,f**). The high degree of accuracy supports the idea that dopamine neurons were following a common template in their responses to unexpected reward, that is, their response functions all took the same shape. In addition, this multiplicative scaling (equation (2)) was better able to fit individual dopamine neurons than three other simple models (equations (3–5)). Multiplicative scaling

had the lowest mean-squared error overall, and in any pairwise comparison, multiplicative scaling was a better fit in the majority of identified and putative dopamine neurons.

$$f_i(r) = \alpha_i + f_{\max} \left( \frac{r^{0.5}}{r^{0.5} + \sigma^{0.5}} \right) \quad (3)$$

$$f_i(r) = f_{\max} \left( \frac{(r \times \alpha_i)^{0.5}}{(r \times \alpha_i)^{0.5} + \sigma^{0.5}} \right) \quad (4)$$

$$f_i(r) = f_{\max} \left( \frac{(r + \alpha_i)^{0.5}}{(r + \alpha_i)^{0.5} + \sigma^{0.5}} \right) \quad (5)$$

After fitting the unexpected reward data, we explored what transformation could best mimic the effect of expectation. As we describe in detail elsewhere<sup>28</sup>, we found that subtraction provided the best fit (**Fig. 2a**).

$$f(r) = f_{\max} \left( \frac{r^{0.5}}{r^{0.5} + \sigma^{0.5}} \right) - E \quad (6)$$

In this equation,  $f_{\max}$  and  $\sigma$  were determined by the unexpected reward data; the only new parameter was the expectation factor  $E$ .

We noticed that the value of  $E$  varied from neuron to neuron, and that it correlated strongly with a neuron's responsiveness to unexpected reward (**Fig. 2b**). We wondered if this correlation would allow us to fit expected reward responses into the same common framework we applied to unexpected reward responses. In other words, could one type of response predict the other? To find out, we derived  $\alpha_i$  for each neuron using equation (2). This parameter is based solely on the unexpected reward data; it is the parameter that best scaled a neuron's responses to the average. Next, we multiplied this parameter by the population response to expected reward (**Fig. 2a**). This gave us the predicted response to expected reward for each neuron and each reward size (**Fig. 2d–f**). We compared these predicted values with the actual values, and found a high degree of similarity (**Fig. 2c**). This implies that dopamine neurons calculate prediction error in the same way, just scaled up or down.

To measure noise correlations (**Fig. 6** and **Supplementary Fig. 6**), we first calculated the trial-by-trial firing rates of each neuron during each task condition (reward-predicting cue, expected reward delivery, unexpected reward delivery, reward omission, and baseline). We used 1-s windows because this ensured that we included the full duration of response for each trial, and because short windows have been shown to hide noise correlations<sup>31</sup>. However, 500-ms windows gave us similar results (data not shown). We then transformed these firing rates into Z scores by subtracting the mean response to a given stimulus and dividing by the s.d. This gave us a vector of Z scores for each neuron and each condition; the length of the vector was the number of trials. For each pair of simultaneously-recorded neurons, we then calculated the Pearson's correlation between these vectors. To determine if the noise correlations were due to processes happening over long time scales (for example, increased satiety causing all VTA neurons to slowly decrease their firing rate), we also calculated the correlation after shifting one neuron's vector by one trial. Thus, the firing rate for trial 1 in neuron 1 was compared to the firing rate for trial 2 in neuron 2, and so on.

We first analyzed the two experiments separately (**Fig. 6a,b**), but to plot the histogram of noise correlations (**Fig. 6c,d**), we combined the two experiments. Specifically, to analyze the cue period, we combined the reward-predicting cue in the variable-reward task with the 90% reward cue in the variable-expectation class. To analyze the reward period, we combined the expected reward in the variable-reward task with the 90% expected reward in the variable-expectation task. For the variable-reward task, we calculated noise correlations separately for each reward size, but given that the values were consistent, we report the average over reward sizes.

Comparisons were performed with paired  $t$  tests or Wilcoxon rank-sum tests, with corrections for multiple comparisons (Bonferroni or Tukey). Correlations were done with Pearson's rho.  $P$  values less than 0.05 were considered significant, unless otherwise noted. Bootstrapping model comparisons were done in the following fashion: we resampled the data 1,000 times and determined

for each resample the mean squared error for both two-parameter and three-parameter functions. We calculated the  $P$  value by counting the number of resamples when the mean squared error was better for the two-parameter function (for example, if 1 resample out of 1,000 preferred the two-parameter function,  $P = 0.001$ ). Analyses were done using custom scripts in Matlab (Mathworks). Code is available on request. For an expected correlation coefficient greater than 0.6, the sample size needed for  $\alpha$  of 0.05 and  $\beta$  of 0.2 is approximately 19 neurons, which we exceeded in each experiment. Randomization and blinding were not employed.

**Immunohistochemistry.** After recording for 4–8 weeks, mice were given an overdose of ketamine/medetomidine, exsanguinated with saline, and perfused with 4% paraformaldehyde (vol/vol). Brains were cut in 100- $\mu$ m coronal sections on a vibrotome and immunostained with antibodies to tyrosine hydroxylase (AB152, 1:1,000, Millipore) to visualize dopamine neurons and 49,6-diamidino-2-phenylindole (DAPI, Vectashield) to visualize nuclei. Virus expression was determined through eYFP fluorescence. Slides were examined to verify that the optic fiber track was among VTA dopamine neurons and in a region expressing the virus.

A **Supplementary Methods Checklist** is available.

51. Bäckman, C.M. *et al.* Characterization of a mouse strain expressing Cre recombinase from the 3' untranslated region of the dopamine transporter locus. *Genesis* **44**, 383–390 (2006).
52. Boyden, E.S., Zhang, F., Bamberg, E., Nagel, G. & Deisseroth, K. Millisecond-timescale, genetically targeted optical control of neural activity. *Nat. Neurosci.* **8**, 1263–1268 (2005).
53. Atasoy, D., Aponte, Y., Su, H.H. & Sternson, S.M. A FLEX switch targets Channelrhodopsin-2 to multiple cell types for imaging and long-range circuit mapping. *J. Neurosci.* **28**, 7025–7030 (2008).
54. Uchida, N. & Mainen, Z.F. Speed and accuracy of olfactory discrimination in the rat. *Nat. Neurosci.* **6**, 1224–1229 (2003).
55. Tian, J. & Uchida, N. Habenula lesions reveal that multiple mechanisms underlie dopamine prediction errors. *Neuron* **87**, 1304–1316 (2015).
56. Schmitzer-Torbert, N. & Redish, A.D. Neuronal activity in the rodent dorsal striatum in sequential navigation: separation of spatial and reward responses on the multiple T task. *J. Neurophysiol.* **91**, 2259–2272 (2004).
57. Lima, S.Q., Hromádka, T., Znamenskiy, P. & Zador, A.M. PINP: a new method of tagging neuronal populations for identification during in vivo electrophysiological recording. *PLoS One* **4**, e6099 (2009).
58. Kvitsiani, D. *et al.* Distinct behavioural and network correlates of two interneuron types in prefrontal cortex. *Nature* **498**, 363–366 (2013).
59. Olsen, S.R., Bhandawat, V. & Wilson, R.I. Divisive normalization in olfactory population codes. *Neuron* **66**, 287–299 (2010).