LETTER

doi:10.1038/nature10754

Neuron-type-specific signals for reward and punishment in the ventral tegmental area

Jeremiah Y. Cohen[1]*, Sebastian Haesler[1]*, Linh Vong[2], Bradford B. Lowell[2] & Naoshige Uchida[1]

Dopamine has a central role in motivation and reward. Dopaminergic neurons in the ventral tegmental area (VTA) signal the discrepancy between expected and actual rewards (that is, reward prediction error)[1–3], but how they compute such signals is unknown. We recorded the activity of VTA neurons while mice associated different odour cues with appetitive and aversive outcomes. We found three types of neuron based on responses to odours and outcomes: approximately half of the neurons (type I, 52%) showed phasic excitation after reward-predicting odours and rewards in a manner consistent with reward prediction error coding; the other half of neurons showed persistent activity during the delay between odour and outcome that was modulated positively (type II, 31%) or negatively (type III, 18%) by the value of outcomes. Whereas the activity of type I neurons was sensitive to actual outcomes (that is, when the reward was delivered as expected compared to when it was unexpectedly omitted), the activity of type II and type III neurons was determined predominantly by reward-predicting odours. We 'tagged' dopaminergic and GABAergic neurons with the light-sensitive protein channelrhodopsin-2 and identified them based on their responses to optical stimulation while recording. All identified dopaminergic neurons were of type I and all GABAergic neurons were of type II. These results show that VTA GABAergic neurons signal expected reward, a key variable for dopaminergic neurons to calculate reward prediction error.

Dopaminergic neurons fire phasically (100–500 ms) after unpredicted rewards or cues that predict reward[1–3]. Their response to reward is reduced when a reward is fully predicted. Furthermore, their activity is suppressed when a predicted reward is omitted. From these observations, previous studies hypothesized that dopaminergic neurons signal discrepancies between expected and actual rewards (that is, they compute reward prediction error (RPE)), but how dopaminergic neurons compute RPE is unknown.

Dopaminergic neurons make up about 55–65% of VTA neurons; the rest are mostly GABAergic inhibitory neurons[4–6]. Many addictive drugs inhibit VTA GABAergic neurons, which increases dopamine release (called disinhibition), a potential mechanism for reinforcing the effects of these drugs[7–12]. Despite the known role of VTA GABAergic neurons inhibiting dopaminergic neurons *in vitro*[13], little is known about their role in normal reward processing. One obstacle has been the difficulty of identifying different neuron types with extracellular recording techniques. Conventionally, spike waveforms and other firing properties have been used to identify presumed dopaminergic and GABAergic neurons[1,2,14,15], but this approach has been questioned recently[5,16]. We thus aimed to observe how dopaminergic and GABAergic neurons process information about rewards and punishments.

We classically conditioned mice with different odour cues that predicted appetitive or aversive outcomes. The possible outcomes were big reward, small reward, nothing, or punishment (a puff of air delivered to the animal's face). Each behavioural trial began with a conditioned stimulus (CS; an odour, 1 s), followed by a 1-s delay and an unconditioned stimulus (US; the outcome). Within the first two behavioural sessions, mice began licking towards the water-delivery tube in the delay before rewards arrived, indicating that they quickly learned the CS–US associations (Fig. 1). The lick rate was significantly higher preceding big rewards than small ones (paired *t*-tests between lick rates for big versus small rewards for each session, $P < 0.05$ for each mouse).

We recorded the activity of VTA neurons while mice performed the conditioning task. All 95 neurons showed task-related responses (analysis of variance (ANOVA), all $P < 0.001$), thus all recorded neurons were used in the following analyses. Observing the temporal profiles of responses in trials with rewards, we found neurons that showed firing patterns that resemble those of dopaminergic neurons found in non-human primates[1,2,15]. These neurons were excited phasically by reward-predicting stimuli or reward (Fig. 2a, top). We also found many neurons with firing patterns distinct from typical dopaminergic neurons. These neurons showed persistent excitation during the delay before rewards, in response to reward-predicting odours (Fig. 2a, middle). Other neurons showed persistent inhibition to reward-predicting odours (Fig. 2a, bottom). To characterize the responses of the population, we measured the temporal response profile of each neuron during big-reward trials by quantifying firing rate changes from baseline in 100-ms bins using a receiver operating characteristic (ROC) analysis (Fig. 2b and Supplementary Fig. 1). We calculated the area under the ROC curve (auROC) at each time bin. Values greater than 0.5 indicate increases in firing rate relative to baseline, whereas values less than 0.5 indicate decreases.

To classify these response profiles, we used principal component analysis (PCA) followed by unsupervised, hierarchical clustering. This yielded three clusters of neurons that were separated according to (1) the magnitude of activity during the delay between CS and US, and (2) the magnitude of responses to the CS or US (Fig. 2c). Forty-nine neurons (52%) were classified as type I, which showed phasic responses. Twenty-nine neurons (31%) were classified as type II, which showed sustained excitation to reward-predicting odours, whereas 17 neurons (18%) were classified as type III, which showed sustained inhibition (Fig. 2d).

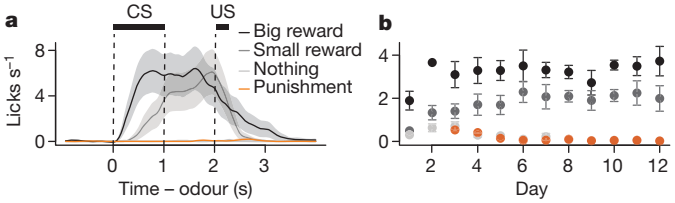


Figure 1 | Odour-outcome association task in mice. a, Licking behaviour from a representative experimental session. Black bars indicate CS and US delivery. Shaded regions around lick traces denote standard error of the mean (s.e.m.). **b**, Mean ± s.e.m. licks during the delay between CS and US as a function of days of the experiment across animals.

[1]Department of Molecular and Cellular Biology, Center for Brain Science, Harvard University, Cambridge, Massachusetts 02138, USA. [2]Division of Endocrinology, Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts 02215, USA.
*These authors contributed equally to this work.

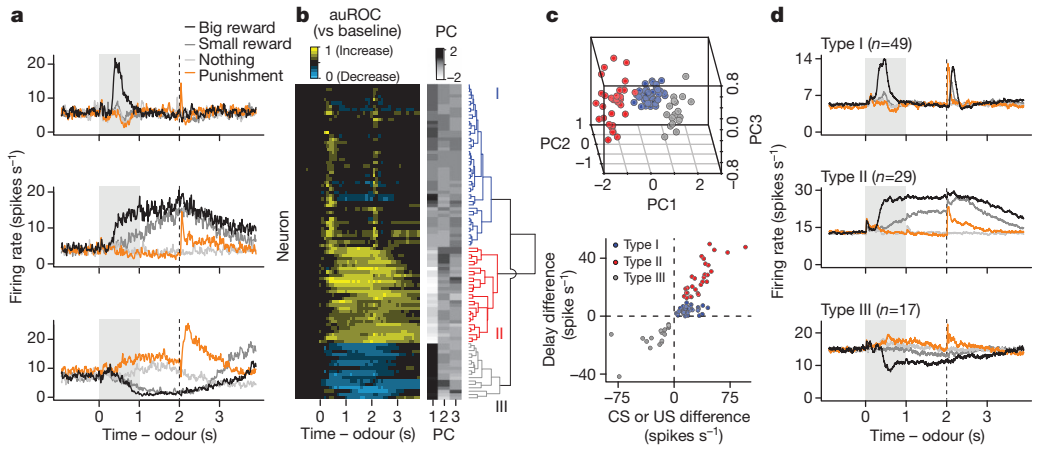
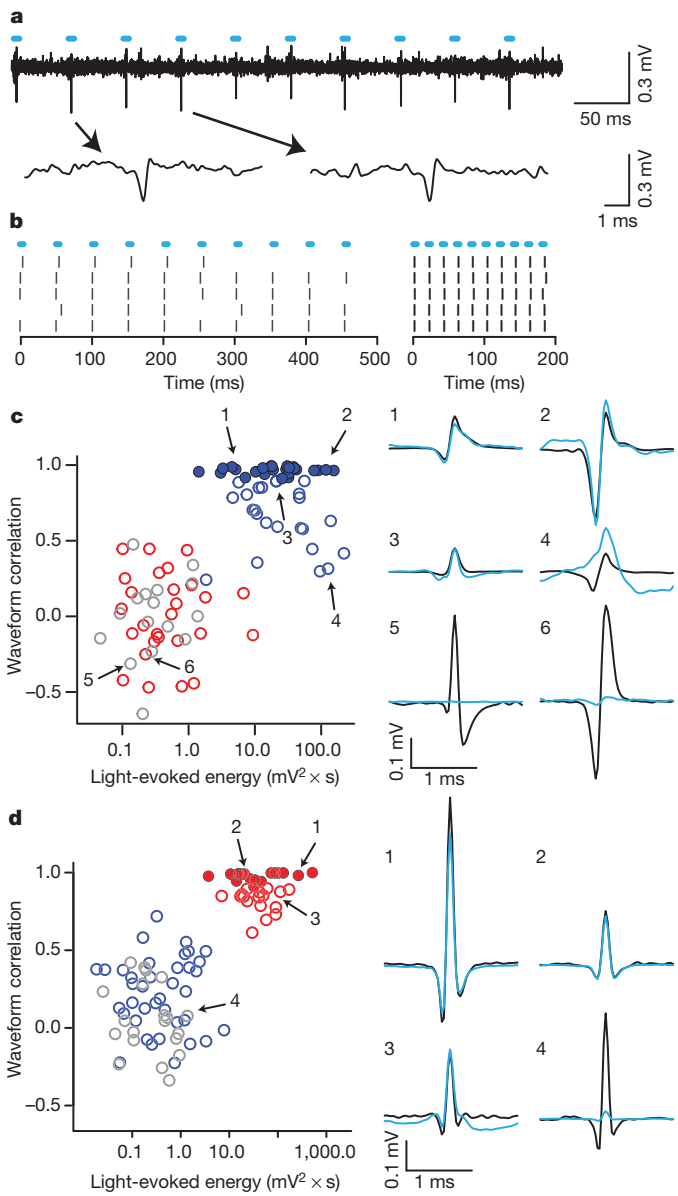


Figure 2 | VTA neurons show three distinct response types. **a**, Responses of example neurons. **b**, Responses of all neurons. Yellow, increase from baseline; cyan, decrease from baseline. Each row represents one neuron. The similarity order of the three main clusters is arranged to match the order presented in **a**. **c**, Top: the first three principal components of the auROC curves. Points are coloured based on hierarchical clustering from the dendrogram. Bottom: classification of neurons based on response differences between big-reward and no-reward trials during the delay versus during the CS or US. **d**, Average firing rates from type I–III neurons.



To identify dopaminergic neurons, we expressed channelrhodopsin-2 (ChR2), a light-gated cation channel[17,18], in dopaminergic neurons (see Methods). We confined expression to dopaminergic neurons by injecting adeno-associated virus containing FLEX-ChR2 (AAV-FLEX-ChR2)[19] into transgenic mice expressing Cre recombinase under the control of the promoter of the dopamine transporter (*DAT*; also called *Slc6a3*) gene (Supplementary Figs 2 and 3). For each neuron, we measured the response to light pulses and the shape of spontaneous spikes. We observed many neurons that fired after light pulses (Fig. 3a, b). We calculated the correlation between the spontaneous spike waveform and light-evoked voltage response and plotted it against the energy of light-evoked responses for each recording (Fig. 3c). This yielded two distinct clusters: one that showed significant responses to light pulses and one that did not. To identify dopaminergic neurons stringently, we applied the criterion that the light-evoked waveform must look almost identical to the spontaneous waveform (correlation coefficient >0.9). Twenty-six neurons met this criterion (filled blue points in Fig. 3c). Consistent with direct light activation rather than indirect, synaptic activation, all 26 neurons showed light-evoked spikes within a few milliseconds of light onset with small jitter, and followed high-frequency light stimulation of 50 Hz (Supplementary Fig. 4). These properties strongly indicate that these 26 neurons expressed ChR2. We therefore designate these 26 neurons as identified dopaminergic neurons. All identified dopaminergic neurons were of type I. Conversely, none of type II or III neurons was activated by light (red and grey points in Fig. 3c).

Next, we asked whether GABAergic neurons could be mapped to type II or III neurons. We recorded from 92 VTA neurons in mice expressing Cre recombinase under the control of the endogenous vesicular γ-aminobutyric acid (GABA) transporter (*Vgat*; also called *Slc32a1*) gene. These mice showed similar licking behaviour to *DAT*-Cre mice (Supplementary Fig. 5). We applied the PCA parameters

Figure 3 | Identifying dopaminergic and GABAergic neurons. **a**, Voltage trace from 10 pulses of 20 Hz light stimulation (cyan bars). Two light-triggered spikes are shown below. **b**, Response from this neuron to 20 Hz (left) and 50 Hz (right) stimulation. Ticks represent spikes. **c**, Quantification of light-evoked responses and mapping of response types in *DAT*-Cre mice. Blue, type I; red, type II; grey, type III neurons. Identified dopaminergic neurons are indicated by filled circles. Abscissa: energy (integral of the squared voltage values, $\int v^2 dt$) of the light-evoked response from each neuron. Ordinate: cross-correlation between the mean spontaneous spike and the light-evoked response. Example neurons are shown to the right (black, spontaneous spikes; cyan, light-evoked voltages). **d**, Light-evoked responses in *Vgat*-Cre mice. Conventions are the same as in **c**.

from the 95 neurons from *DAT*-Cre mice to the 92 neurons from *Vgat*-Cre mice. This yielded 38 type I neurons, 34 type II neurons and 20 type III neurons. Using the same criteria for GABAergic neurons as we used for dopaminergic neurons, we identified 17 GABAergic neurons (Fig. 3d and Supplementary Fig. 4). All 34 type II neurons fell in the upper cluster in Fig. 3d. We also found type I neurons that were inhibited by optical stimulation, consistent with local GABAergic stimulation (Supplementary Fig. 6).

Our data set of identified dopaminergic neurons allows us to characterize their diversity. We observed that some were excited by reward, some were excited by a reward-predicting CS, and some were excited by both (Fig. 4a–c). Although previous studies in non-human primates found similar variability[20,21] (Supplementary Fig. 7), this result may suggest that some dopaminergic neurons do not strictly follow canonical RPE coding. However, the US responses may be due to the delay between CS and US, known to increase the US response due to temporal uncertainty[20]. In addition, this diversity was correlated with the effect of training that occurred over several days across the population of dopaminergic neurons, even after animals had reached asymptotic behavioural performance (Fig. 1b). Soon after reaching a behavioural performance criterion, many dopaminergic neurons showed stronger responses to US over CS but the preference gradually shifted to CS over several days (Fig. 4d; Pearson correlation, $r = 0.42$, $P < 0.05$). This is consistent with a previous study in non-human primates that showed US responses gradually disappear over >1 month of training[21]. Thus, identified dopaminergic neurons appear to respond to CS and US similarly to those reported in non-human primate studies.

Another important response property that supports RPE coding in dopaminergic neurons is their decrease in firing rate when an expected reward is omitted[1,3]. We thus omitted reward unexpectedly on 10% of big-reward trials in some sessions. Fifteen of seventeen dopaminergic neurons showed a decrease in firing rate upon reward omission relative to reward delivery (Fig. 4f, g). The two dopaminergic neurons that were not modulated by reward omission were excited by big-reward CS, but fired close to 0 spikes s^{-1} otherwise; the low firing rate at the time of reward left little room to 'dip' further. We obtained similar results when we compared the firing rate upon reward omission to the baseline firing rate (9 of 17 neurons $P < 0.05$, *t*-test; mean auROC = 0.407, $t_{16} = 2.56$, $P < 0.05$; Supplementary Fig. 8a, b). Thus, most dopaminergic neurons coded RPE when expected reward was omitted.

GABAergic neurons showed persistent activity during the delay period, which parametrically encoded the value of upcoming outcomes (paired *t*-tests between no-, small- and big-reward trials, all $P < 0.001$ for 16 of 17 identified GABAergic neurons, Supplementary Fig. 7a; regression slopes, Supplementary Fig. 10i). This suggests that these neurons encode expectation about rewards. If this is the case, one prediction is that the activity of these neurons is not modulated by delivery or omission of reward. Indeed, GABAergic (and unidentified type II) and type III neurons were not significantly modulated by the presence or absence of reward itself (Fig. 4f, g and Supplementary Fig. 8), in contrast to identified dopaminergic neurons. None of the identified GABAergic neurons, and only 2 of 17 unidentified type II neurons, showed significant decreases in firing rate relative to when reward was delivered. None of the 11 type III neurons showed significant modulation by reward omission. Thus, the activity of type II and III neurons was modulated predominantly by reward-predicting cues but not actual reward.

Recent studies have revealed a diversity of responses of dopaminergic neurons to aversive stimuli: some are excited, others inhibited[15]. To test whether this diversity exists in dopaminergic and GABAergic VTA neurons, we delivered air puffs in some sessions. Identified dopaminergic neurons showed some diversity: although most significant responses were inhibitory, some were excitatory (Fig. 4h, i and Supplementary Fig. 9). In contrast, most type II and III neurons (and 13 of 14 identified GABAergic neurons) were excited by air puffs.

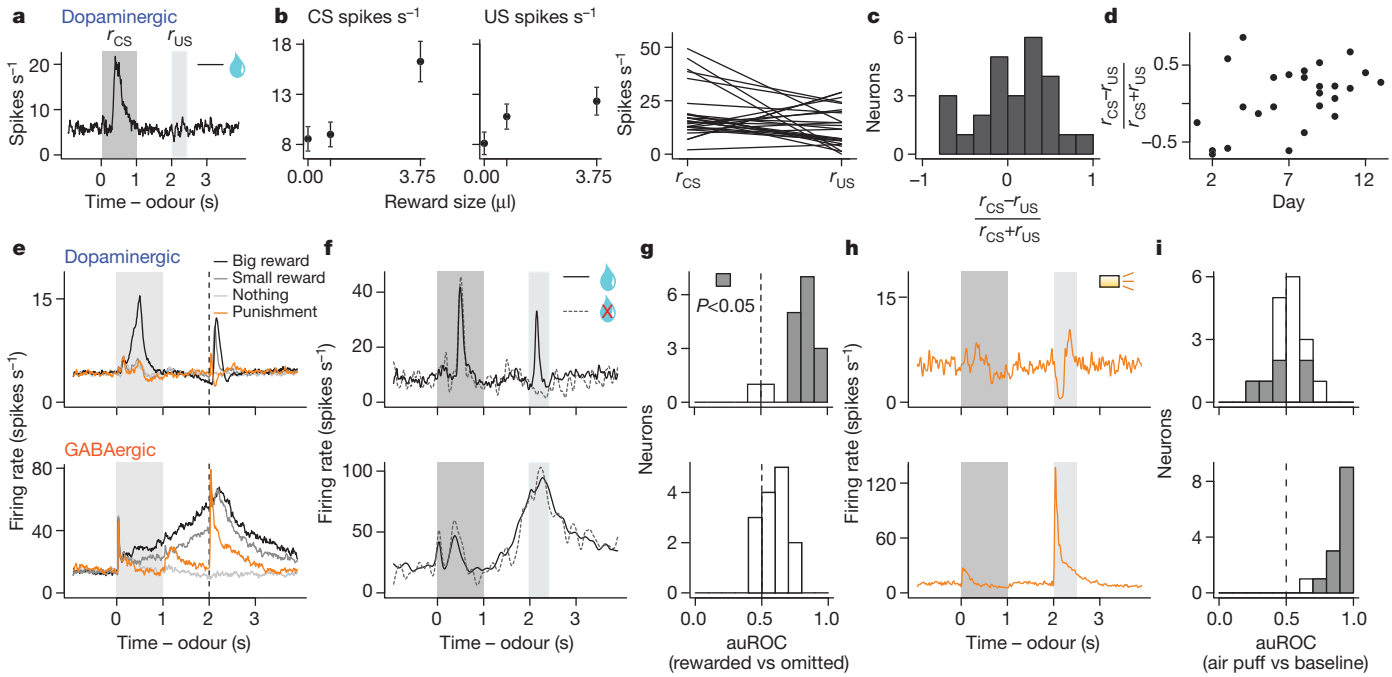


Figure 4 | Response variability based on CS–US preference, reward omission and air puffs. **a**, Response of a dopaminergic neuron during big-reward trials. **b**, Firing rate (mean ± s.e.m.) versus reward size (left) and in response to big-reward-predicting CS and big-reward US for each dopaminergic neuron (right). **c**, Histogram of CS–US index for dopaminergic neurons. **d**, CS–US index versus day after the behaviour was learned. **e**, Average responses of dopaminergic and GABAergic neurons. **f**, Responses of a dopaminergic and GABAergic neuron for reward present (solid) and unexpectedly absent (dashed) on big-reward trials. **g**, Histograms of differences in firing rates during the outcome period (2–2.5 s) between rewarded and reward-omitted trials for dopaminergic (top) and GABAergic (bottom) neurons. Values are represented using auROC (<0.5, rewarded < omitted; 0.5, no difference; >0.5, rewarded > omitted). Significant values are filled (*t*-test, $P < 0.05$). **h**, Responses of a dopaminergic and GABAergic neuron during punishment trials. **i**, Histograms of auROC values during the air puff (2–2.5 s) relative to baseline (<0.5, decrease; >0.5, increase from baseline).

Detecting the discrepancy between expected and actual outcomes has a critical role in optimal learning[1,22,23]. Although phasic firing of VTA dopaminergic neurons may act as such an error signal, how this is computed remains largely unknown. Models have postulated the existence of value-dependent, inhibitory input to dopaminergic neurons that persists during the delay between a CS and US (Supplementary Fig. 11a)[1,23]. Our data indicate that VTA GABAergic neurons provide such an inhibitory input that counteracts excitatory drive from primary reward when the reward is expected. In addition, these neurons were excited by aversive stimuli, potentially contributing to suppression of firing in some dopaminergic neurons in response to aversive events (Fig. 4). Previous work showed that VTA GABAergic neurons receive inputs from prefrontal cortex and subcortical areas that could provide reward-related signals[24–29]. Phasic excitation of VTA GABAergic neurons could be driven by inputs from lateral habenula neurons that are phasically excited by aversive stimuli[29]. These habenular neurons do not show sustained activity between CS and US, so it is unlikely that they provide reward expectation signals to VTA GABAergic neurons. Instead, these signals may come from the pedunculopontine nucleus[25] or orbitofrontal cortex[27] (Supplementary Fig. 11b). VTA GABAergic neurons form synapses preferentially onto dendrites of dopaminergic neurons[28], whereas other inhibitory inputs form synapses onto their somata[29]. Dendritic inhibition is thought to be weaker than somatic 'shunting' inhibition[28] but appears well suited for deriving graded outputs by 'arithmetically' combining excitatory and inhibitory inputs.

A major effect of drugs of addiction is inhibition of VTA GABAergic neurons[7,8]. If VTA GABAergic neurons are involved in computation of RPE, inhibition of GABAergic neurons by addictive drugs could lead to sustained RPE even after the learned effects of drug intake are well established, thereby resulting in sustained reinforcement of drug taking[30]. Understanding local circuits in VTA in the context of learning theory may thus provide crucial insights into normal as well as abnormal functions of reward circuits.

METHODS SUMMARY

All surgical and experimental procedures were in accordance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals and approved by the Harvard Institutional Animal Care and Use Committee. We injected male *DAT*-Cre and *Vgat*-Cre mice with adeno-associated virus carrying FLEX-ChR2 into the VTA and implanted a head plate and a microdrive containing six tetrodes and an optical fibre. While mice performed a classical conditioning task, we recorded spiking activity from VTA neurons. We delivered pulses of light to activate ChR2 and classified neurons as dopaminergic, GABAergic or unidentified. After experiments, we performed immunohistochemistry to localize recording sites amid dopaminergic neurons.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 9 May; accepted 2 December 2011.

Published online 18 January 2012.

1. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275,** 1593–1599 (1997).
2. Bayer, H. M. & Glimcher, P. W. Midbrain dopamine neurons encode a quantitative reward prediction error signal. *Neuron* **47,** 129–141 (2005).
3. Schultz, W. Behavioral theories and the neurophysiology of reward. *Annu. Rev. Psychol.* **57,** 87–115 (2006).
4. Swanson, L. W. The projections of the ventral tegmental area and adjacent regions: a combined fluorescent retrograde tracer and immunofluorescence study in the rat. *Brain Res. Bull.* **9,** 321–353 (1982).
5. 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.* **577,** 907–924 (2006).
6. Nair-Roberts, R. G. *et al.* Stereological estimates of dopaminergic, GABAergic and glutamatergic neurons in the ventral tegmental area, substantia nigra and retrorubral field in the rat. *Neuroscience* **152,** 1024–1031 (2008).
7. Hyman, S. E., Malenka, R. C. & Nestler, E. J. Neural mechanisms of addiction: the role of reward-related learning and memory. *Annu. Rev. Neurosci.* **29,** 565–598 (2006).
8. Lüscher, C. & Malenka, R. C. Drug-evoked synaptic plasticity in addiction: from molecular changes to circuit remodeling. *Neuron* **69,** 650–663 (2011).
9. Johnson, S. W. & North, R. A. Opioids excite dopamine neurons by hyperpolarization of local interneurons. *J. Neurosci.* **12,** 483–488 (1992).
10. Mansvelder, H. D., Keath, J. R. & McGehee, D. S. Synaptic mechanisms underlie nicotine-induced excitability of brain reward areas. *Neuron* **33,** 905–919 (2002).
11. Szabo, B., Siemes, S. & Wallmichrath, I. Inhibition of GABAergic neurotransmission in the ventral tegmental area by cannabinoids. *Eur. J. Neurosci.* **15,** 2057–2061 (2002).
12. Tan, K. R. *et al.* Neural bases for addictive properties of benzodiazepines. *Nature* **463,** 769–774 (2010).
13. Dobi, A., Margolis, E. B., Wang, H.-L., Harvey, B. K. & Morales, M. Glutamatergic and nonglutamatergic neurons of the ventral tegmental area establish local synaptic contacts with dopaminergic and nondopaminergic neurons. *J. Neurosci.* **30,** 218–229 (2010).
14. Steffensen, S. C., Svingos, A. L., Pickel, V. M. & Henriksen, S. J. Electrophysiological characterization of GABAergic neurons in the ventral tegmental area. *J. Neurosci.* **18,** 8003–8015 (1998).
15. Matsumoto, M. & Hikosaka, O. Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature* **459,** 837–841 (2009).
16. Lammel, S. *et al.* Unique properties of mesoprefrontal neurons within a dual mesocorticolimbic dopamine system. *Neuron* **57,** 760–773 (2008).
17. Nagel, G. *et al.* Channelrhodopsin-2, a directly light-gated cation-selective membrane channel. *Proc. Natl Acad. Sci. USA* **100,** 13940–13945 (2003).
18. Boyden, E. S., Zhang, F., Bamberg, E., Nagel, G. & Deisseroth, K. Millisecond-timescale, genetically targeted optical control of neural activity. *Nature Neurosci.* **8,** 1263–1268 (2005).
19. 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).
20. Fiorillo, C. D., Newsome, W. T. & Schultz, W. The temporal precision of reward prediction in dopamine neurons. *Nature Neurosci.* **11,** 966–973 (2008).
21. Takikawa, Y., Kawagoe, R. & Hikosaka, O. A possible role of midbrain dopamine neurons in short- and long-term adaptation of saccades to position-reward mapping. *J. Neurophysiol.* **92,** 2520–2529 (2004).
22. Rescorla, R. A. & Wagner, A. R. in *Classical Conditioning II: Current Research and Theory* (eds Black, A. H. & Wagner, A. R.) 64–99 (New York, 1972).
23. Houk, J. C., Adams, J. L. & Barto, A. G. in *Models of Information Processing in the Basal Ganglia* (eds Houk, J. C., Davis, J. L. & Beiser, D. G.) 249–270 (MIT Press, 1995).
24. Carr, D. B. & Sesack, S. R. Projections from the rat prefrontal cortex to the ventral tegmental area: target specificity in the synaptic associations with mesoaccumbens and mesocortical neurons. *J. Neurosci.* **20,** 3864–3873 (2000).
25. Okada, K., Toyama, K., Inoue, Y., Isa, T. & Kobayashi, Y. Different pedunculopontine tegmental neurons signal predicted and actual task rewards. *J. Neurosci.* **29,** 4858–4870 (2009).
26. Matsumoto, M. & Hikosaka, O. Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature* **447,** 1111–1115 (2007).
27. Takahashi, Y. K. *et al.* Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. *Nature Neurosci.* **14,** 1590–1597 (2011).
28. Omelchenko, N. & Sesack, S. R. Ultrastructural analysis of local collaterals of rat ventral tegmental area neurons: GABA phenotype and synapses onto dopamine and GABA cells. *Synapse* **63,** 895–906 (2009).
29. Jhou, T. C., Fields, H. L., Baxter, M. G., Saper, C. B. & Holland, P. C. The rostromedial tegmental nucleus (RMTg), a GABAergic afferent to midbrain dopamine neurons, encodes aversive stimuli and inhibits motor responses. *Neuron* **61,** 786–800 (2009).
30. Redish, A. D. Addiction as a computational process gone awry. *Science* **306,** 1944–1947 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Meister, V. N. Murthy, J. D. Schall and R. P. Heitz for comments, C. Dulac for sharing resources, C. I. Moore, J. Ritt and J. Siegle for advice about microdrives, K. Deisseroth for the AAV-FLEX-ChR2 construct, and E. Soucy and J. Greenwood for technical support. This work was supported by a Howard Hughes Medical Institute Fellowship from the Helen Hay Whitney Foundation (J.Y.C.); the Human Frontiers Science Program (S.H.); a Howard Hughes Medical Institute Collaborative Innovation Award, a Smith Family New Investigator Award, the Alfred Sloan Foundation, the Milton Fund (N.U.); F32 DK078478, P30 DK046200 (L.V.); and R01 DK075632, R01 DK089044, P30 DK046200, P30 DK057521 (B.B.L.).

Author Contributions J.Y.C. and S.H. collected and analysed data. J.Y.C., S.H. and N.U. designed experiments and wrote the paper. L.V. and B.B.L. generated *Vgat*-Cre mice.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to N.U. (uchida@mcb.harvard.edu).

METHODS

Animals. We used six adult male mice, backcrossed with C57/BL6J mice, heterozygous for Cre recombinase under the control of the *DAT* gene (B6.SJL-Slc6a3$^{tm1.1(cre)Bkmn}$/J, The Jackson Laboratory)[31] and six adult male mice, backcrossed with C57/BL6J mice, heterozygous for Cre recombinase under the control of the *Vgat* gene (*Vgat*-ires-Cre)[32]. Animals were housed on a 12-h dark/12-h light cycle (dark from 06:00 to 18:00) and each performed the conditioning task at the same time of day, between 07:00 and 18:00.

Surgery and viral injections. Mice were surgically implanted with a head plate and a microdrive containing electrodes and an optical fibre. During surgery, we injected 200–500 nl adeno-associated virus (AAV), serotype 5, carrying an inverted ChR2 (H134R)-EYFP flanked by double *loxP* sites[18,19,33] into the VTA stereotactically (from bregma: 3.1 mm posterior, 0.8 mm lateral, 4–4.5 mm ventral). Our expression was highly selective (<1% of ChR2-positive neurons were negative for the dopaminergic neuron marker tyrosine hydroxylase, TH) and efficient (>90% of TH-expressing neurons co-expressed ChR2; Supplementary Figs 2 and 3).

For each unidentified type I neuron in *DAT*-Cre mice, we were unable to identify it as dopaminergic using our stringent criteria not because there was no response on the electrode after light pulses, but because the response did not match the shape of the spontaneous spike waveforms (open blue points in Fig. 3c). This may have been from changing the spike shape by opening many ChR2 or due to 'spillover' from nearby neurons, contaminating the signal.

After injections in *Vgat*-Cre mice, we confirmed that our ChR2 expression was selective (<1% of ChR2-positive neurons were negative for the GABAergic neuron marker glutamic acid decarboxylase, GAD65/67) and efficient (but less efficient than in *DAT*-Cre mice; >60% of GAD65/67-expressing neurons co-expressed ChR2) (Supplementary Fig. 2).

For fluorogold tracer injections (Fluorochrome), we injected two animals each with 10–25 nl into the following areas: prefrontal cortex (from bregma: 2.1 mm anterior, 0.25 mm lateral, 1.6 mm ventral), medial striatum (1.4 mm anterior, 1.0 mm lateral, 4.0 mm ventral) and basolateral amygdala (1.4 mm posterior, 2.8 mm lateral, 4.3 mm ventral).

All surgery was performed under aseptic conditions with animals under ketamine/medetomidine anaesthesia (60/0.5 mg kg^{-1}, intraperitoneal, respectively). Analgesia (ketoprofen, 5 mg kg^{-1}, intraperitoneal) was administered postoperatively.

Behavioural task. After >1 week of recovery, mice were water-deprived in their home cage. Weight was maintained above 90% of their full body weight. Animals were head-restrained using a custom-made metal plate and habituated for 1–2 d while head-restrained before training on the task. Odours were delivered with a custom-made olfactometer[34]. Each odour was dissolved in paraffin oil at 1/100 dilution. Thirty microlitres of diluted odour was placed inside a filter-paper housing. Odours were isoamyl acetate, 1-butanol, *N*-butyl acetate, citral, eugenol, (+) limonene and (−) carvone, and differed for different animals. Odourized air was further diluted with filtered air by 1:10 to produce a 500 ml min^{-1} total flow rate. Licks were detected by breaks of an infrared beam placed in front of the water tube.

We delivered one of four odours, selected pseudorandomly, for 1 s, followed by a delay of 1 s and an outcome. Each odour predicted a different outcome: a large drop of water (3.75 µl; valve open for 100 ms), a small drop of water (0.75 µl; valve open for 20 ms), no outcome, or an air puff delivered to the animal's face. Air puff trials were included for 17 identified dopaminergic neurons, 14 identified GABAergic neurons, and 15 type III neurons. Intertrial intervals (ITIs) were drawn from an exponential distribution with a rate parameter of 10 (that is, $ITI = 10\exp(-10x)$). This resulted in a flat ITI hazard function, ensuring that expectation about the start of the next trial did not increase over time. Data from *DAT*-Cre mice were obtained from 85 sessions (9–23 sessions per animal, 14 ± 4.8 sessions, mean ± s.d.). For 17 identified dopaminergic neurons, 14 identified GABAergic neurons, and 11 type III neurons, we omitted rewards during 10% of big-reward trials. Animals performed between 400 and 1,000 trials per day (533 ± 120 trials, mean ± s.d.). Data from *Vgat*-Cre mice were obtained from 71 sessions.

Electrophysiology. We recorded extracellularly from multiple neurons simultaneously using a custom-built 200-µm-fibreoptic-coupled screw-driven microdrive with six implanted tetrodes (four wires wound together). For three *Vgat*-Cre mice, we used a 105-µm fibre-optic. Tetrodes were glued to the fibre optic with epoxy. The ends of the tetrodes were 350–600 µm from the end of the fibre optic. Neural signals and time stamps for behaviour were recorded using a DigiLynx recording system (Neuralynx). Broadband signals from each wire filtered between 0.1 and 9,000 Hz were recorded continuously at 32 kHz. To extract the timing of spikes, signals were band-pass-filtered between 300–6,000 Hz and sorted online and off-line using custom software.

To verify that our recordings targeted dopaminergic or GABAergic neurons, we used ChR2 to observe stimulation-locked spikes[35]. The optical fibre was coupled with a diode-pumped solid-state laser with analogue amplitude modulation (Laserglow Technologies). Before and after each behavioural session, we delivered trains of 10 light pulses, each 5-ms long, at 1, 5, 10, 20 and 50 Hz at 473 nm at 5–20 mW mm^{-2}. We used frequencies of 1, 10, 20, 50 and 100 Hz in *Vgat*-Cre mice. Higher intensities typically resulted in light-evoked spike waveforms that did not match spontaneous ones (open blue points in Fig. 3c and open red points in Fig. 3d). Therefore, we adjusted the light intensity after observing the responses. The increasing latency of light-evoked spiking as a function of stimulation frequency indicates that the response was not due to photochemical artefact (Supplementary Fig. 4f, j). Spike shape was measured using a broadband signal (0.1– 9,000 Hz) sampled at 32 kHz. This ensured that particular features of the spike waveform were not missed. A previous study found decreasing probability of activation with increasing frequency[33]. This discrepancy may be explained by the low levels of light we used to identify neurons and the difference in experimental preparation, *in vitro* versus *in vivo*[36].

We used two criteria to include a neuron in our data set. First, the neuron must have been well isolated (signal-to-noise ratio of >5 dB). Second, the neuron must have been recorded within 500 µm of an identified dopaminergic neuron (or a type I neuron in *Vgat*-Cre mice), to ensure that all neurons came from VTA. Using distance cutoffs of 400, 300 or 200 µm yielded similar estimates of the proportion of type I neurons as the full data set (49 of 91, 49 of 86 and 49 of 82, respectively, in *DAT*-Cre mice).

Recording sites were further verified histologically with electrolytic lesions using 15–20 s of 100 µA direct current and from the optical fibre track (Supplementary Fig. 12).

Data analysis. To measure firing rates, peristimulus time histograms (PSTHs) were constructed using 10-ms bins. To calculate spike density functions, PSTHs were convolved with a function resembling a postsynaptic potential[37]. To determine whether a neuron showed a significant task-related response, we calculated an ANOVA on the trial-by-trial firing rates (spikes s^{-1}) during the baseline period (1 s before odour onset), CS period (from odour onset to odour offset), delay (from odour offset to outcome onset), and US period (from outcome onset to 500 ms after outcome onset). The factors were task epoch (that is, baseline, CS, delay, or US) and outcome type. All neurons showed a main effect of time; 183 of 187 neurons showed a main effect of outcome type; 184 of 187 neurons showed a time–outcome interaction ($P < 0.001$).

Light-evoked spikes were detected during the 10 ms after light onset. If <20% of light pulses had a spike (defined as a waveform that matched that of the isolated unit) during the 10 ms after light onset (lower left points in Fig. 3c, d), the maximum voltage during that interval was used as the light-evoked 'response'. Cross-correlations between spontaneous and light-evoked spike waveforms were calculated by aligning the positive peak of each waveform, averaging separately, and aligning the peaks of the averages. The cross-correlation coefficients reported in Fig. 3 are from a lag of 0 ms. The correlation was calculated using the full duration of the spontaneous spike (spike duration was measured as the first time until the last time at which the voltage was significantly different from baseline using Wilcoxon rank sum tests). The energy of the light-evoked waveform is defined as the integral of the squared voltage values ($\int v^2 dt$). To ensure that measurements of the fidelity of light-evoked responses were not biased by poorly isolated neurons, we plotted the probability of firing at 50 Hz light stimulation against L-ratio, a measure of isolation quality[38] (Supplementary Fig. 13). L-ratio approximates the fraction of 'contaminated' spikes; smaller L-ratios indicate better isolation. One hundred and eighty of 187 neurons in the data set had L-ratios <0.05.

CS–US indices were calculated as (CS − US)/(CS + US), where CS is the difference between the peak firing rate (maximum value of the PSTH) in the 1 s after odour onset and the baseline firing rate, and US is the difference between the peak firing rate in the 0.5 s after reward onset and the baseline firing rate. The baseline firing rate was calculated as the mean of the PSTH in the 0.5 s before odour onset.

ROC curves were calculated by comparing the distribution of firing rates across trials in 100-ms bins to the distribution of baseline firing rates (900 ms before odour onset; Supplementary Fig. 1). PCA was calculated using the singular value decomposition of the auROC. Learning was measured using the lick responses (mean licks s^{-1}) across consecutive days of behavioural sessions. Curves were fit with logistic functions, $k/(1+\exp(-t))$. The task was considered fully learned if the learning rate was within 95% of the carrying capacity of the logistic (k). All 12 animals learned the task within the first 3 days.

Hierarchical clustering was done using the first three PCs of the auROC curves using a Euclidean distance metric and complete agglomeration method.

All statistical tests were done with corrections for multiple comparisons (Bonferroni or Tukey). Analyses were done with R (http://www.r-project.org/).

Immunohistochemistry. After recording, which lasted between 9 and 23 days, mice were given an overdose of ketamine/medetomidine, exsanguinated with saline, perfused with paraformaldehyde, and brains were cut in 50 µm coronal or horizontal sections. Sections were immunostained with antibodies to TH and

secondary antibodies labelled with Cy3 (Jackson Immunoresearch). Sections were further stained with 4′,6-diamidino-2-phenylindole (DAPI) or TO-PRO-3 (Invitrogen) to visualize nuclei. Recording sites were identified and verified to be amid EYFP staining and TH staining in VTA. Brain slices from additional virus-injected *DAT*-Cre and *Vgat*-Cre animals were stained for TH and GAD-65/67 (Millipore) with secondary antibodies labelled with Cy3 or Alexa 594 (Invitrogen). Cellular fluorescence intensities of ChR2-EYFP were measured using Volocity Image Analysis Software (Perkin Elmer).

31. 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).
32. Vong, L. *et al.* Leptin action on GABAergic neurons prevents obesity and reduces inhibitory tone to POMC neurons. *Neuron* **71,** 142–154 (2011).
33. Tsai, H. C. *et al.* Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science* **324,** 1080–1084 (2009).
34. Uchida, N. & Mainen, Z. F. Speed and accuracy of olfactory discrimination in the rat. *Nature Neurosci.* **6,** 1224–1229 (2003).
35. 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).
36. Zhao, S. *et al.* Cell-type specific channelrhodopsin-2 transgenic mice for optogenetic dissection of neural circuitry function. *Nature Methods* **8,** 745–752 (2011).
37. Thompson, K. G., Hanes, D. P., Bichot, N. P. & Schall, J. D. Perceptual and motor processing stages identified in the activity of macaque frontal eye field neurons during visual search. *J. Neurophysiol.* **76,** 4040–4055 (1996).
38. 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).

doi:10.1038/nature10754

Supplementary note 1

To compare our data with previous studies, we measured the width of spikes from identified dopaminergic neurons and unidentified neurons from *DAT*-Cre mice. Previous studies used a criterion that dopaminergic neurons should have wide spikes. We found significant diversity in spike waveforms from identified dopaminergic neurons (Fig. S4). There was no significant difference across neuron types in spike duration (ANOVA $F_{2,89} = 0.15$, $P > 0.7$). We found that *Type I* neurons (and identified dopaminergic neurons, when considered separately) had lower baseline firing rates than *Types II* or *III* neurons (*Type II* - *Type I* mean ± 95% CI, 6.73 ± 6.72 spikes/s, Tukey HSD test, $P < 0.05$; *Type III* - *Type I*, 11.11 ± 8.84 spikes/s, $P < 0.01$), although many *Types II* and *III* neurons had low firing rates (<10 spikes/s). Thus, many identified dopaminergic neurons in the present study would have been missed using previous criteria.

Supplementary note 2

Models of RPE typically assume two excitatory inputs, those from reward-predicting cues (CS) and those from reward (US), in addition to inhibitory inputs reflecting reward expectation discussed above (Fig. S12a). Where do these excitatory inputs come from? That some dopaminergic neurons respond preferentially to reward (and weakly to CS) raises the possibility that these US-preferring neurons may provide other CS-preferring dopaminergic neurons with the reward signals. If this is the case, these neurons should meet the following predictions. First, excitation of these neurons should cause phasic excitation of other dopaminergic neurons (the ones calculating RPE). Second, these neurons should not receive inhibitory inputs encoding reward expectation. First, we observed that the excitation of dopaminergic neurons (using ChR2 stimulation in *DAT*-Cre mice) did not cause synaptic excitation of dopaminergic neurons (Fig. S6d). This result does not support the view that a class of "reward-coding" dopaminergic neurons excites other "RPE-coding" dopaminergic neurons. Second, the magnitude of inhibition during reward omission was as large in US-preferring dopaminergic

neurons as in CS-preferring dopaminergic neurons (Figs. 4, S8, S9). This does not support the view that there is a class of dopaminergic neurons that "purely" encodes reward signals. These results suggest that reward-encoding neurons, as proposed in the models, are located outside the VTA. Similarly, although CS-preferring dopaminergic neurons can, in principle, provide excitatory inputs to other dopaminergic neurons, our results do not support this view, suggesting that excitatory inputs for CS also come from outside the VTA.

Supplementary note 3

Our experiments allow us to map a neuron's function onto the transmitter it releases. The percentage of *Type I* neurons (49/95, or 52%) is close to the estimate of the percentage of dopaminergic neurons in rat VTA (55-65%)[6-8]. Furthermore, the unidentified *Type I* neurons showed similar task-related activity to identified dopaminergic neurons (Fig. S11a,b). Together with the observation that *Type I* neurons, but not *Types II* and *III* neurons, responded to light stimulation in *DAT*-Cre mice, this suggests that unidentified *Type I* neurons were also dopaminergic (although we cannot rule out the possibility that some *Type I* neurons were non-dopaminergic). The second experiment showed that identified GABAergic neurons were of *Type II*. The identity of *Type III* neurons remains to be determined. They may release a different neurotransmitter, such as glutamate[14], but we cannot rule out the possibility that *Type III* neurons are GABAergic or dopaminergic (although the latter seems unlikely, given the high efficiency of ChR2 expression in dopaminergic neurons throughout VTA, regardless of projection target; see Figs. S2, S3).

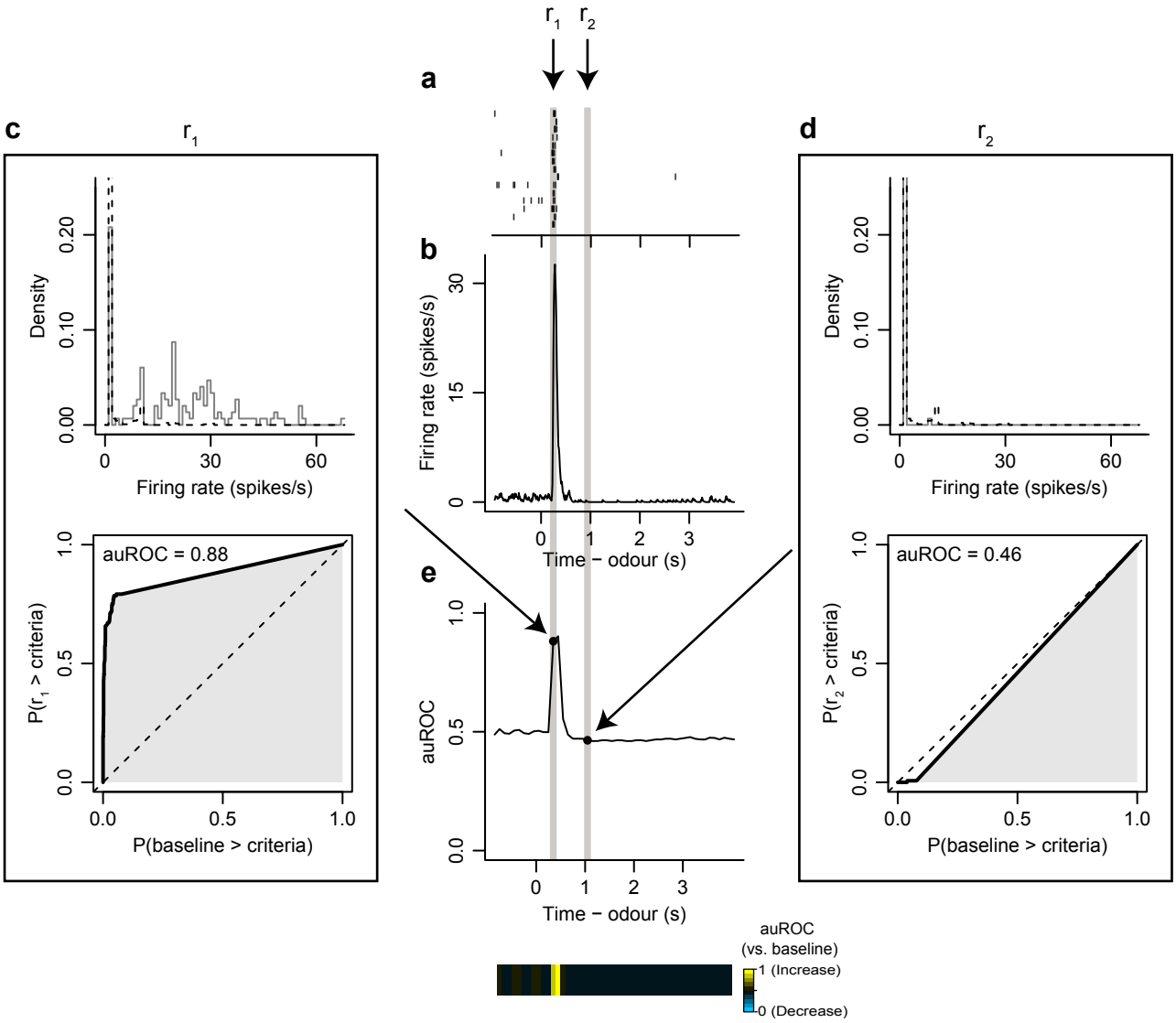


Figure S1. ROC analysis. a, Raster plot from 15 trials of 149 big-reward trials from a dopaminergic neuron. r_1 and r_2 correspond to two example 100-ms bins. **b**, Average firing rate of this neuron. **c**, Area under the receiver operating characteristic curve (auROC) for r_1, in which the neuron increased its firing rate relative to baseline. We compared the histogram of spike counts during the baseline period (dashed line) to that during a given bin (solid line) by moving a criterion from zero to the maximum firing rate (in this example, 68 spikes/s). We then plotted the probability that the activity during r_1 was greater than the criteria against the probability that the baseline activity was greater than the criteria. The area under this curve quantifies the degree of overlap between the two spike count distributions (i.e., the discriminability of the two). The histogram in the top panel is truncated at an ordinate value of 0.2 for display purposes. **d**, Similar to (**c**), but for r_2, which corresponds to an auROC value close to 0.5 (i.e., activity close to baseline). **e**, auROC response profile for the full duration of the task, with r_1 and r_2 indicated by arrows. Shown below are the heat map values, as depicted in Fig. 2.

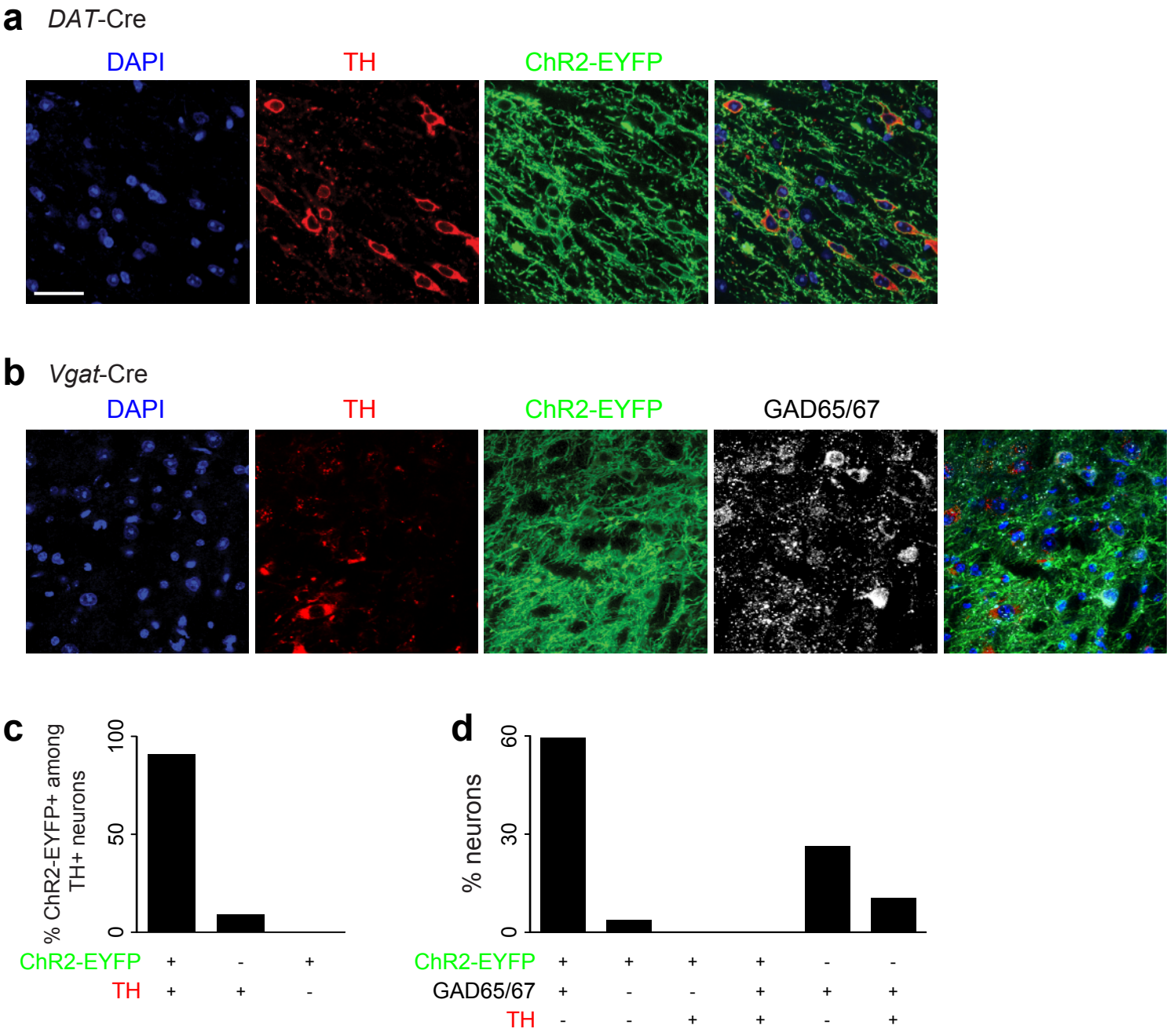


Figure S2. Efficient and specific expression of ChR2-EYFP in dopaminergic and GABAergic neurons. **a**, Blue: DAPI (nuclear marker). Red: immunostaining for TH (dopaminergic neurons). Green: ChR2-EYFP. Scale bar is 50 μm. **b**, Colors as in **a**; White: immunostaining for GAD65/67. Slices are from *DAT*-Cre (**a**) or *Vgat*-Cre (**b**) mice injected with rAAV5-ChR2-EYFP. **c**, Percentage of neurons labeled for EYFP-ChR2 and TH (n = 666 neurons in 2 mice). **d**, Percentage of neurons labeled for EYFP-ChR2, TH and GAD65/67 (n = 433 neurons in 2 mice).

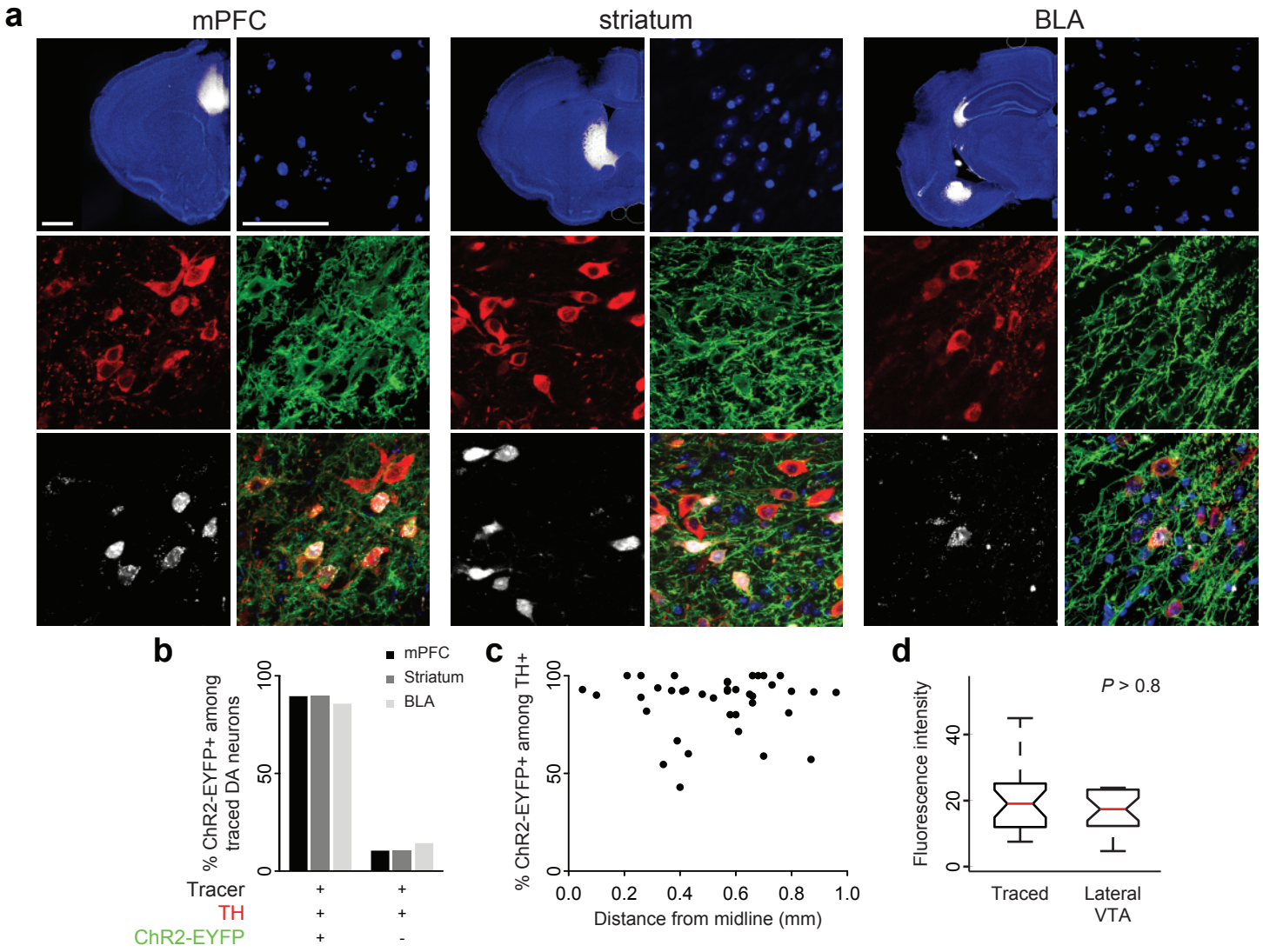


Figure S3. ChR2 expression is uniform across dopaminergic neurons with different projection targets. A recent study found lower levels of DAT in dopaminergic neurons projecting to prefrontal cortex, nucleus accumbens core, and basolateral amygdala versus those projecting to lateral nucleus accumbens shell and dorsolateral striatum[16]. Because Cre expression in the *DAT*-Cre mouse is under the control of the *DAT* gene promoter, this could have biased ChR2 expression toward cells with high *DAT* promoter activity. To confirm that our recordings were not biased toward dopaminergic neurons with a particular projection target, we injected fluorogold, a retrograde tracer, into the prefrontal cortex, striatum, or basolateral amygdala in AAV-FLEX-ChR2-injected *DAT*-Cre mice, and counted the proportion of TH-expressing neurons that also expressed fluorogold and ChR2. **a**, Fluorogold tracing from medial prefrontal cortex (mPFC, left), striatum (middle), and basolateral amygdala (BLA, right) with stains for cell nuclei (To-Pro-3, blue), TH (red), ChR2-EYFP (green), fluorogold (white), and their overlay. Scale bars: 1 mm (left) and 50 μm (right). **b**, The infection efficiency of dopaminergic neurons with identified projection to mPFC, striatum and BLA was similar to the overall infection efficiency measured throughout VTA (Fig. S2). The percentage of VTA neurons triple-labeled by tracer, TH, and ChR2-EYFP was similar across projection target sites, suggesting that virus-mediated ChR2 expression did not preferentially target a subpopulation of dopaminergic neurons (mPFC, n = 67, striatum, n = 145, BLA, n = 35). **c**, Dopaminergic neurons with high *DAT* expression are mainly found in the lateral extent of VTA[21], while low *DAT* expressing neurons are distributed throughout the medial to lateral extent of VTA. Consistent with the finding in (**b**), the viral infection efficiency was similar along the medial-lateral axis of VTA (n = 39). **d**, To evaluate possible differences in cellular ChR2 expression levels, we measured the relative ChR2-EYFP fluorescence intensity over background in dopaminergic neurons projecting to mPFC, BLA or striatum and dopaminergic neurons in the lateral VTA (parabrachial pigmented nucleus). ChR2-EYFP expression was similar in all analysed dopaminergic neurons (n = 25 for each condition).

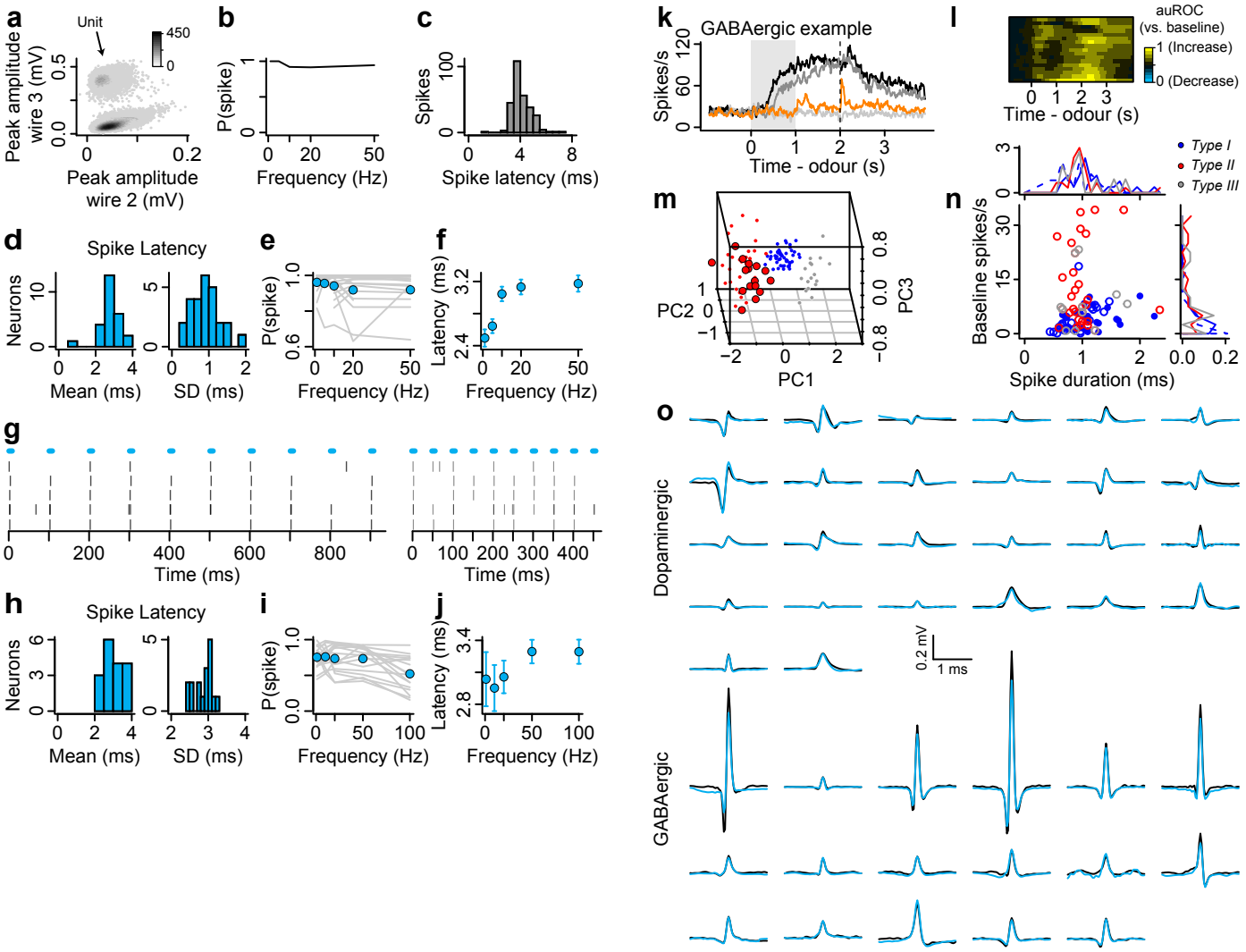


Figure S4. Identification of light-responsive neurons. **a**, Isolation of the neuron in Fig. 3a. Binned scatterplot shows peak amplitude of spike waveforms from two wires of the tetrode. Isolated unit is the upper cluster. The lower cluster is noise. **b**, Probability of a spike (for the second through tenth pulses in each train) as a function of stimulation frequency for this neuron. **c**, Spike latency relative to light onset for this neuron. **d**, Histogram of mean (left) and SD (right) spike latency to light stimulation for 26 identified dopaminergic neurons. **e**, Probability of a spike as a function of stimulation frequency for each dopaminergic neuron (grey) and the mean across dopaminergic neurons (cyan). **f**, Mean ± SEM spike latency as a function of stimulation frequency. **g**, Response from a GABAergic neuron to 5 repetitions of 10 Hz (left) or 20 Hz (right) light stimulation (cyan bars). Ticks represent spikes. **h**, Histogram of mean (left) and SD (right) spike latency to light stimulation for 17 identified GABAergic neurons. **i**, Probability of a spike as a function of stimulation frequency for each GABAergic neuron (grey) and the mean across GABAergic neurons (cyan). **j**, Mean ± SEM spike latency as a function of stimulation frequency. **k**, Response pattern of an example identified GABAergic neuron. **l**, Temporal response profiles of identified GABAergic neurons. Conventions are as in Fig. 2b. **m**, The first three principal components of the auROC curves of all neurons from Fig. 2c (small points) and from identified GABAergic neurons (large points), using the model fit from Fig. 2c. Each GABAergic neuron fell within 95% confidence intervals of the *Type II* cluster defined in *DAT*-Cre mice. **n,** Baseline firing rate vs. spike duration for neurons of each type with density histograms in the margins. Spike duration was calculated as the time at which the voltage was significantly different from baseline (1 ms of pre-spike voltage). **o**, Mean spontaneous (black) and light-evoked (cyan) spike waveforms from 26 identified dopaminergic neurons and 17 identified GABAergic neurons.

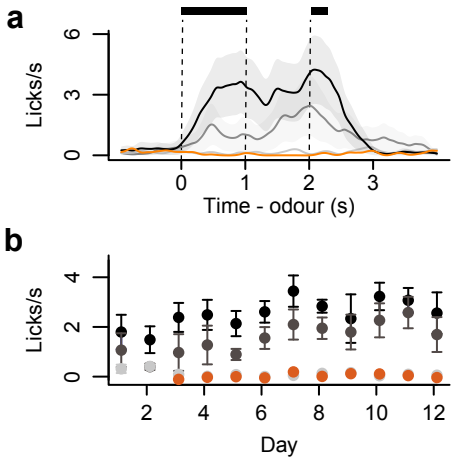


Figure S5. **a**, Licking behaviour from a representative experimental session from a *Vgat*-Cre mouse. Black bars indicate CS and US delivery. Shaded regions around lick traces denote SEM. **b**, mean ± SEM licks during the delay between CS and US as a function of days of the experiment across *Vgat*-Cre animals.

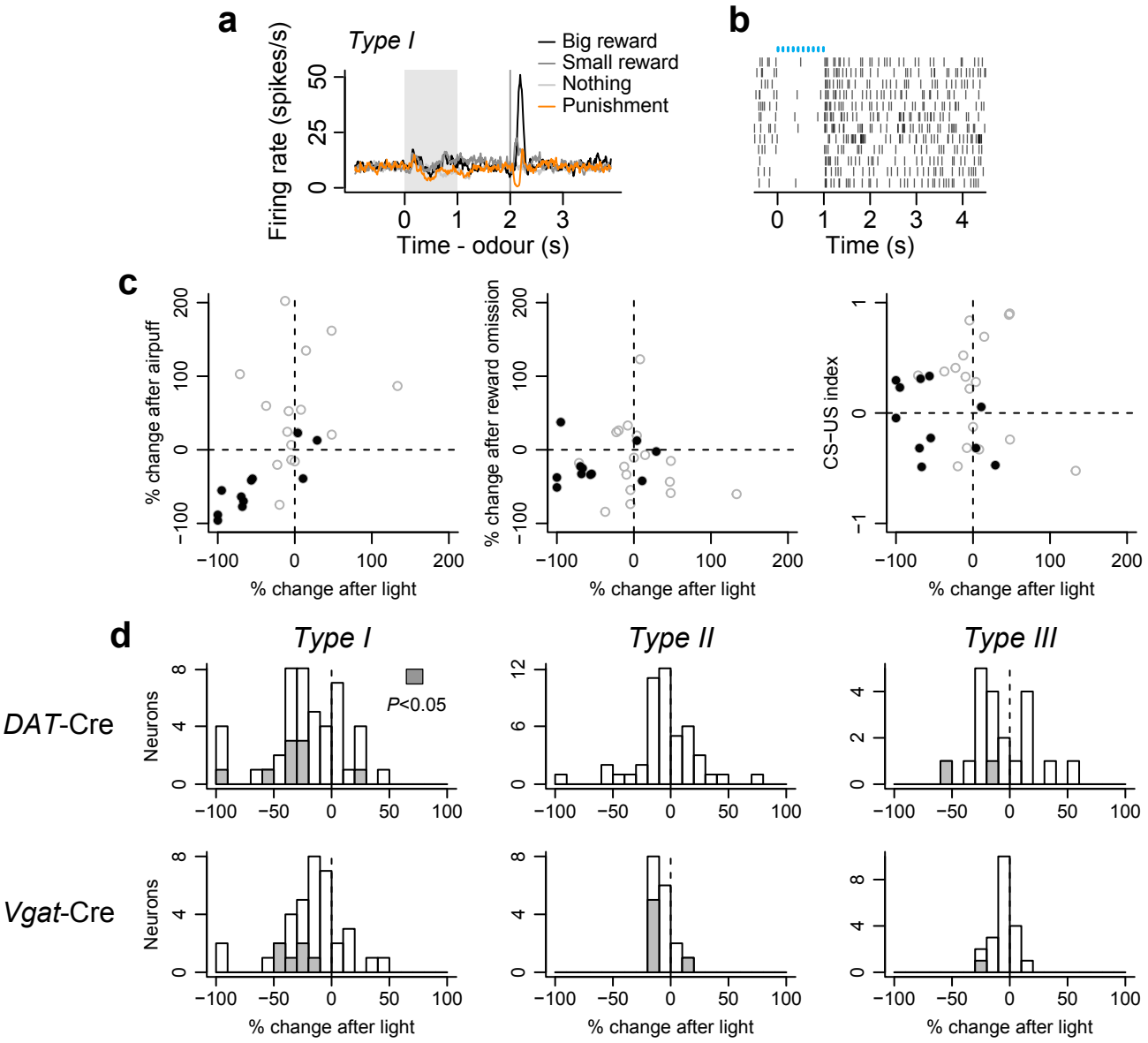


Figure S6. Putative synaptic effects of ChR2 stimulation. **a**, Firing rates from a *Type I* neuron in a *Vgat*-Cre mouse. **b**, Response of the neuron to 12 trials of 10 Hz light pulses (cyan bars). The neuron was inhibited by GABAergic stimulation. **c**, Scatter plots of % change (maximal magnitude of change in the 5-20 ms following light pulses) in firing rate after airpuff, reward omission, and CS-US index vs. % change (maximal magnitude of change) in firing rate after light stimulation in 28 *Type I* neurons from *Vgat*-Cre mice. The 11 neurons for which light stimulation had a significant effect are shown in black (Wilcoxon rank sum test, $P < 0.05$). There was a significant correlation between % change in firing rate after airpuff and % change in firing rate after light stimulation across the population ($r = 0.59$, $P < 0.001$) and for the 11 neurons with a significant response to light stimulation ($r = 0.86$, $P < 0.001$). **d**, % firing rate changes from baseline in the 5-20 ms following all light pulses in *DAT*-Cre (top row) and *Vgat*-Cre (bottom row) mice for each neuron type.

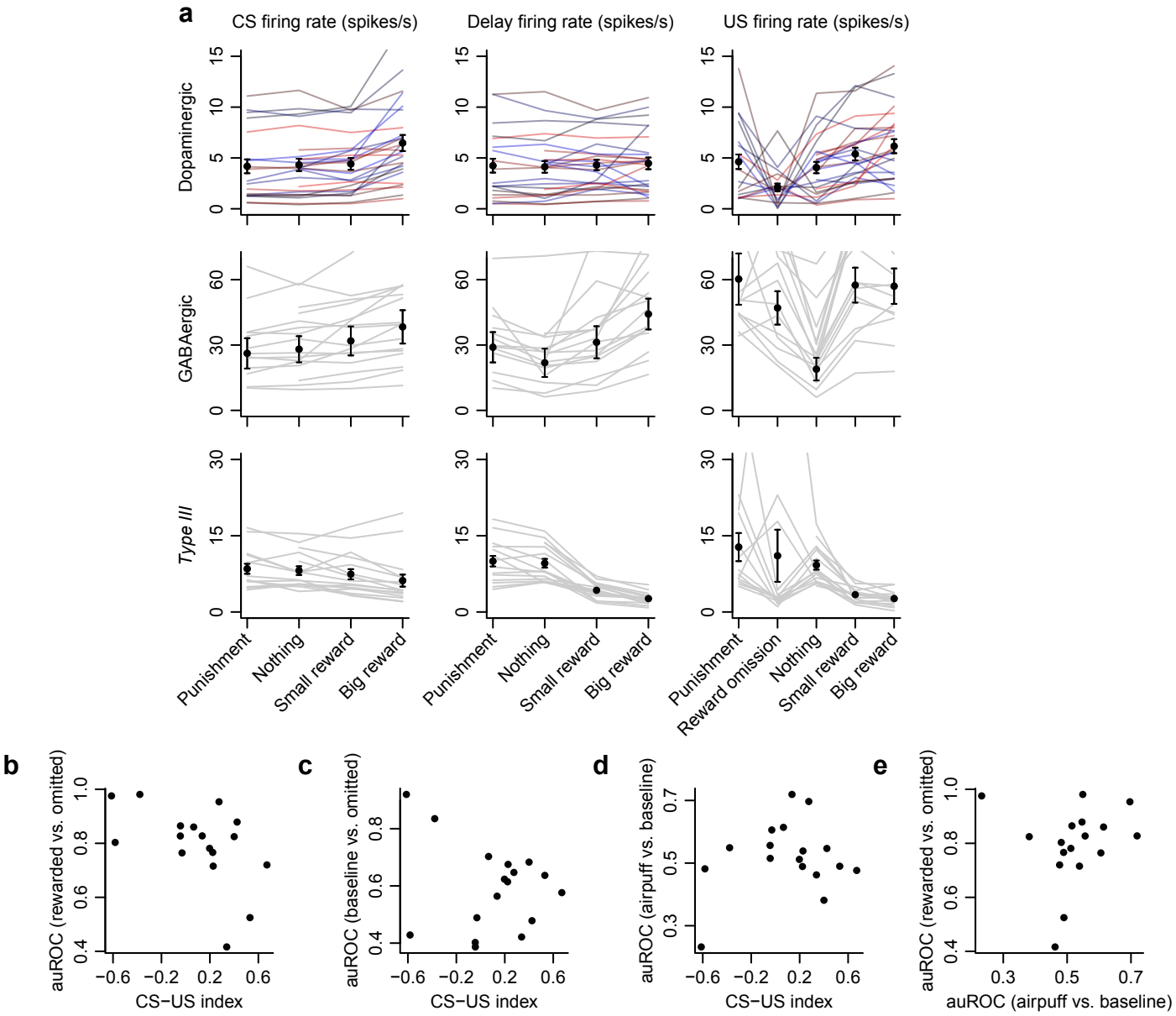


Figure S7. **a**, Mean firing rates for identified dopaminergic (top row), GABAergic (middle row), and *Type III* (bottom row) neurons during CS (left column), delay between CS and US (middle column), and 500 ms after US onset (right column). Each neuron is plotted in grey with the mean and SEM overlaid. Individual dopaminergic neurons are plotted as a function of their CS-US index, with red indicating large CS-US index (i.e., higher firing rate for CS vs. US), blue indicating small CS-US index. **b**, auROC values for reward present compared to reward absent versus CS-US index. There was a significant negative correlation (r = -0.51, P < 0.05) that disappeared with a leverage analysis, indicating that the correlation was driven mostly by one or two neurons. **c**, auROC values for baseline compared to reward absent versus CS-US index. There was no significant correlation. **d**, auROC values for reward present compared to reward absent versus auROC values for airpuff compared to baseline. There was no significant correlation. **e**, auROC values for airpuff compared to baseline versus CS-US index. There was no significant correlation.

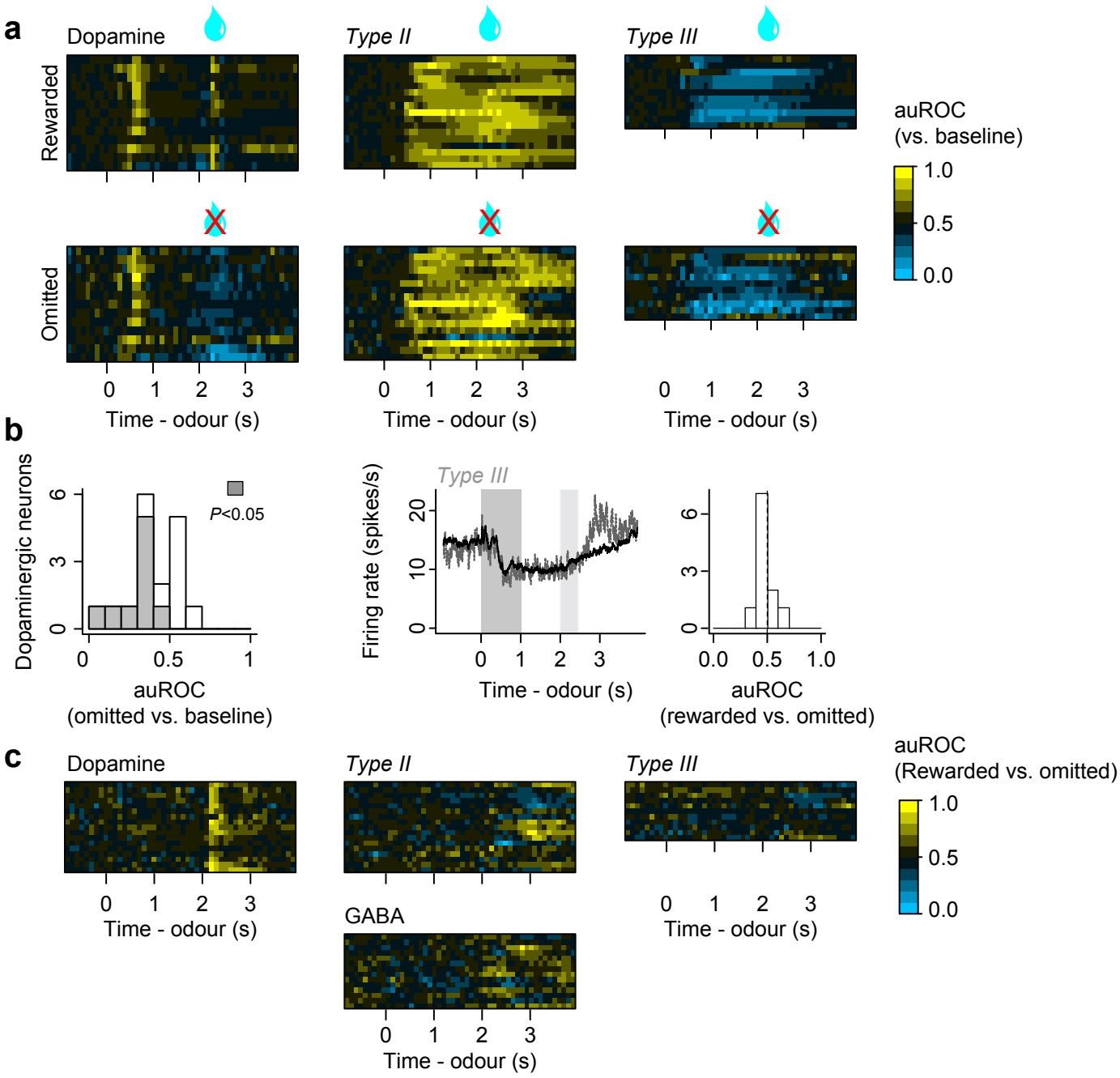


Figure S8. Response profiles of all neurons during reward-omitted trials. **a,** Firing patterns during big-reward trials in which reward was delivered (top) and omitted (bottom). Firing rate changes from baseline were quantified using the auROC curve. Values were computed between each neuron's activity across time and its baseline activity. Yellow: increase from baseline, cyan: decrease from baseline. **b,** Histogram of auROC values during the reward omission period relative to baseline for dopaminergic neurons (left) and example *Type III neuron* and histogram of auROC values for rewarded versus omitted trials. **c,** Difference between rewarded and reward-omitted trials for big-reward trials for identified dopaminergic neurons, unidentified *DAT*-Cre *Types II* and *III* neurons, and identified GABAergic neurons. auROC values were computed between reward-present and reward-absent trials. Each row represents one neuron. Yellow: rewarded > omitted, cyan: rewarded < omitted.

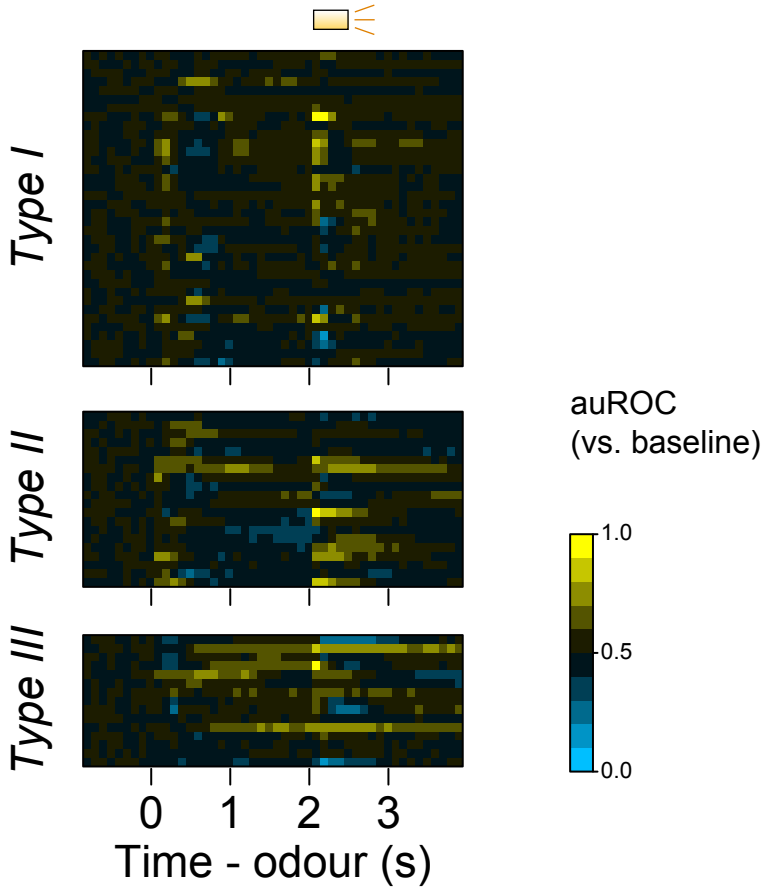


Figure S9. Response profiles of all neurons during airpuff trials. auROC curves for punishment trials relative to baseline activity.

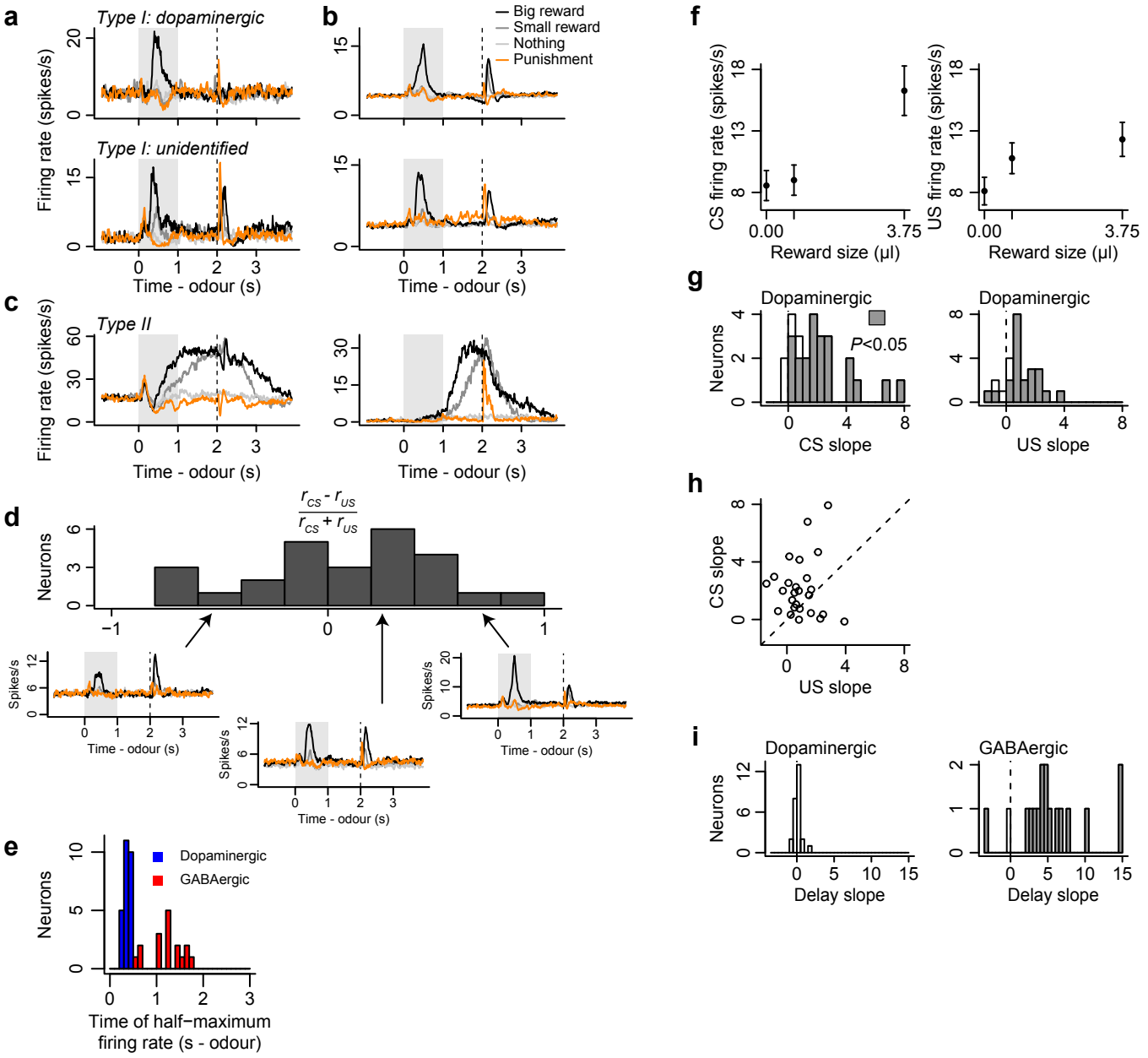


Figure S10. Firing patterns of more example neurons and CS and US responses in dopaminergic neurons. **a**, Examples of firing rates from identified and unidentified *Type I* neurons. **b**, Average firing rates from 26 identified dopaminergic neurons and 23 unidentified *Type I* neurons. **c**, Examples of firing rates from *Type II* neurons. **d**, Average firing rates from the lowest, middle, and upper third of the CS-US index histogram in Fig. 4c (reproduced here). **e**, The time of half-maximum firing rate during big-reward trials from odour onset to reward onset. Identified GABAergic neuron firing rates slowly rose to a peak, while identified dopaminergic neuron firing rates peaked during the CS. **f**, Firing rates for dopaminergic neurons during CS and US. **g**, Regression slopes of the firing rates versus reward size during CS and US for dopaminergic neurons during CS and. **h**, Slope values for each dopaminergic neuron during CS and US. **i**, Slope values during the delay between CS and US. While the delay activity of GABAergic neurons was parametrically modulated by the value of reward, none of identified dopaminergic neurons showed such modulation. Unidentified *Type II* neurons showed a similar modulation by reward value as GABAergic neurons (43/47 unidentified *Type II* neurons showed significant delay-period slope values and 40/47 showed significant differences between no-, small- and big-reward trials, $P < 0.001$).

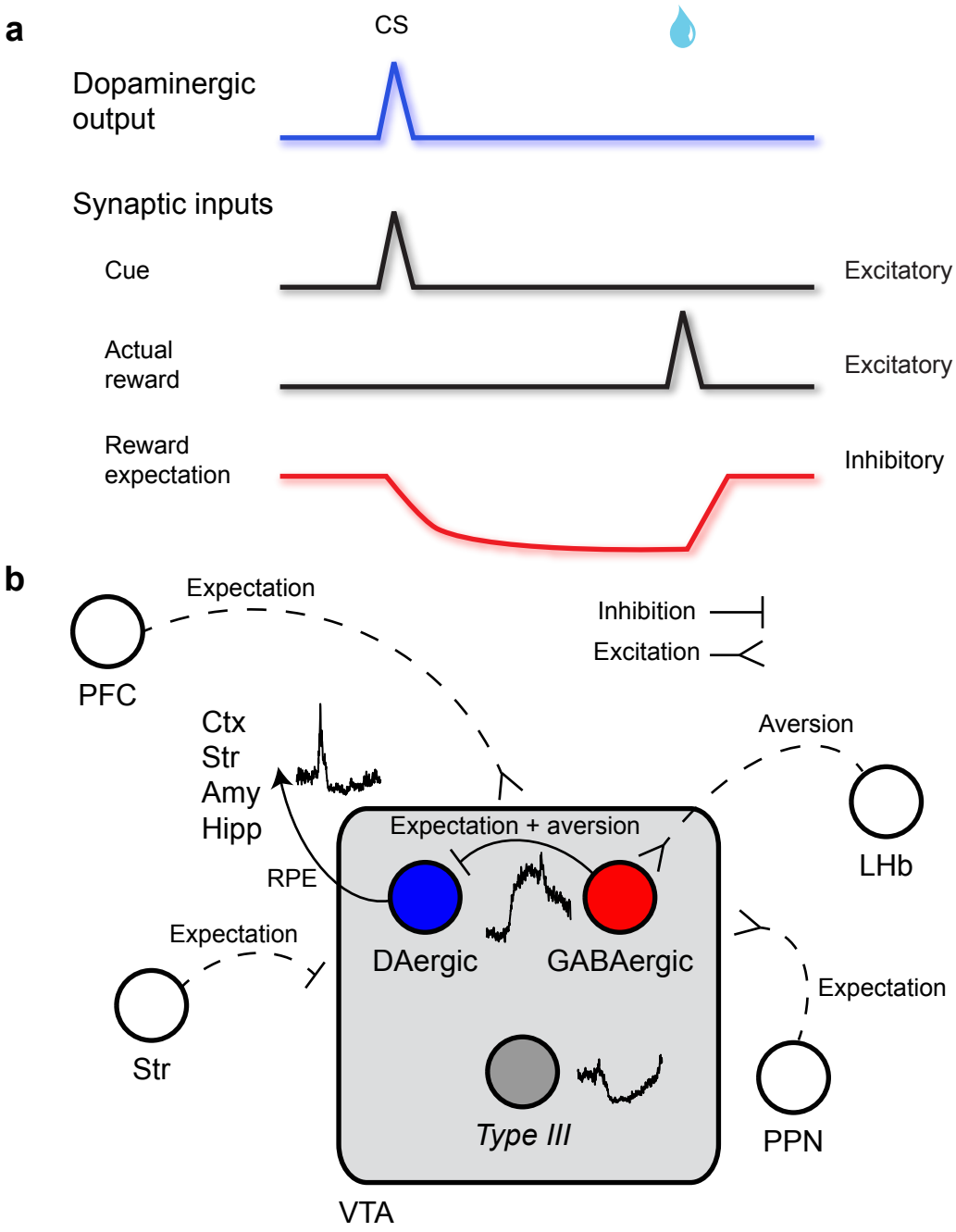


Figure S11. Proposed RPE model and circuit diagram. **a**, A synaptic model of RPE calculation (after Houk et al., 1995). Dopaminergic spikes are a result of excitatory inputs about reward-predicting cues and actual reward, and inhibitory inputs about reward expectation. **b**, Schematic interaction of VTA dopaminergic (DAergic) and GABAergic neurons. VTA receives reward-expectation input from prefrontal cortex (PFC, including orbitofrontal cortex), striatum (str) and pedunculopontine nucleus (PPN). VTA receives inputs about aversive stimuli from lateral habenula (LHb). VTA GABAergic neurons integrate information about reward expectation and aversive stimuli and dopaminergic neurons project RPE signals to cortex (ctx), striatum, amygdala (amy), hippocampus (hipp), and elsewhere. Example firing rates from each type of neuron we recorded are shown. We have omitted many other areas known to be important in this circuit (e.g., rostromedial tegmental nucleus). See text for details.

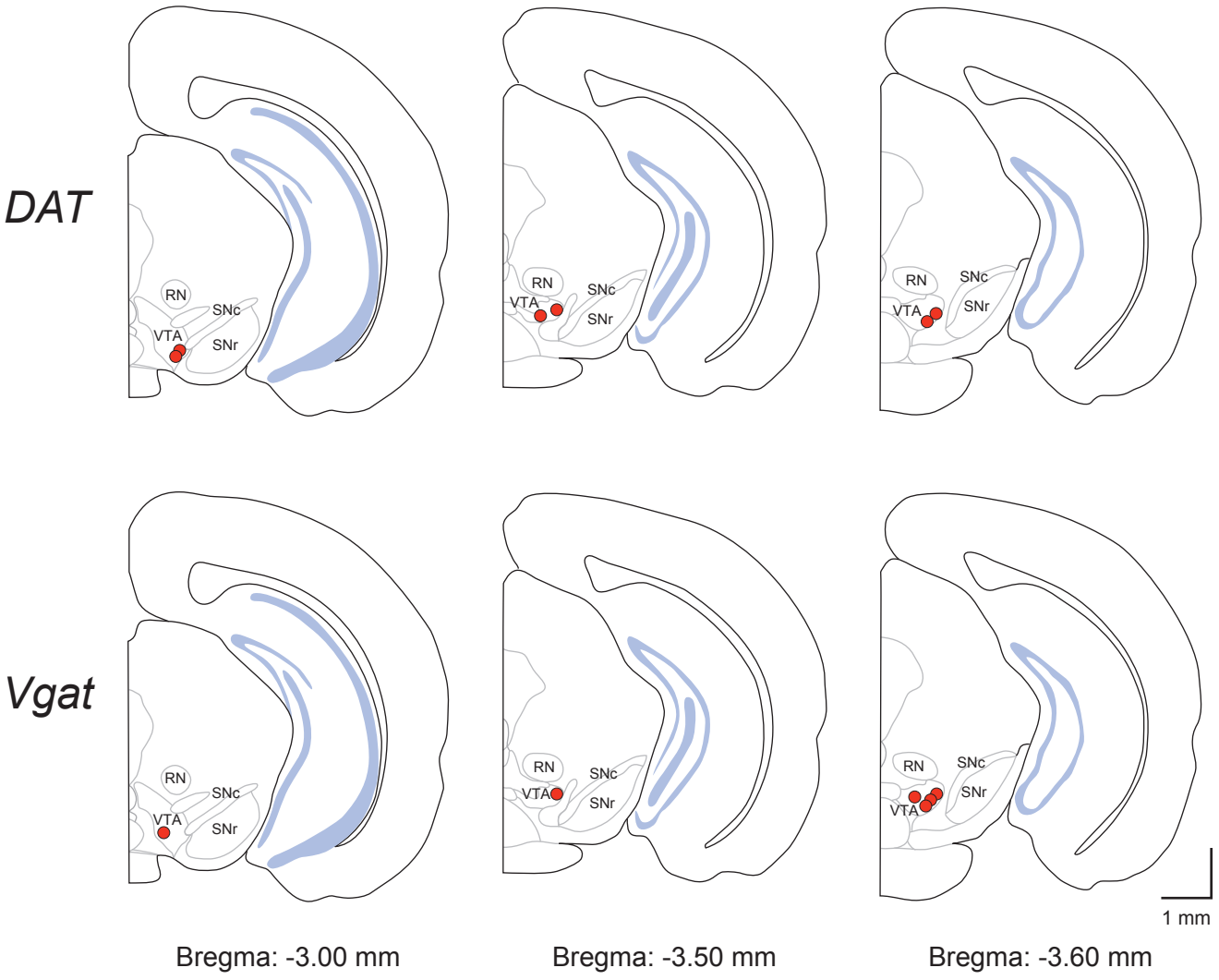


Figure S12. Histological reconstruction of recording sites (red circles). Labeled structures: red nucleus (RN), substantia nigra pars compacta (SNc), substantia nigra pars reticulata (SNr), VTA.

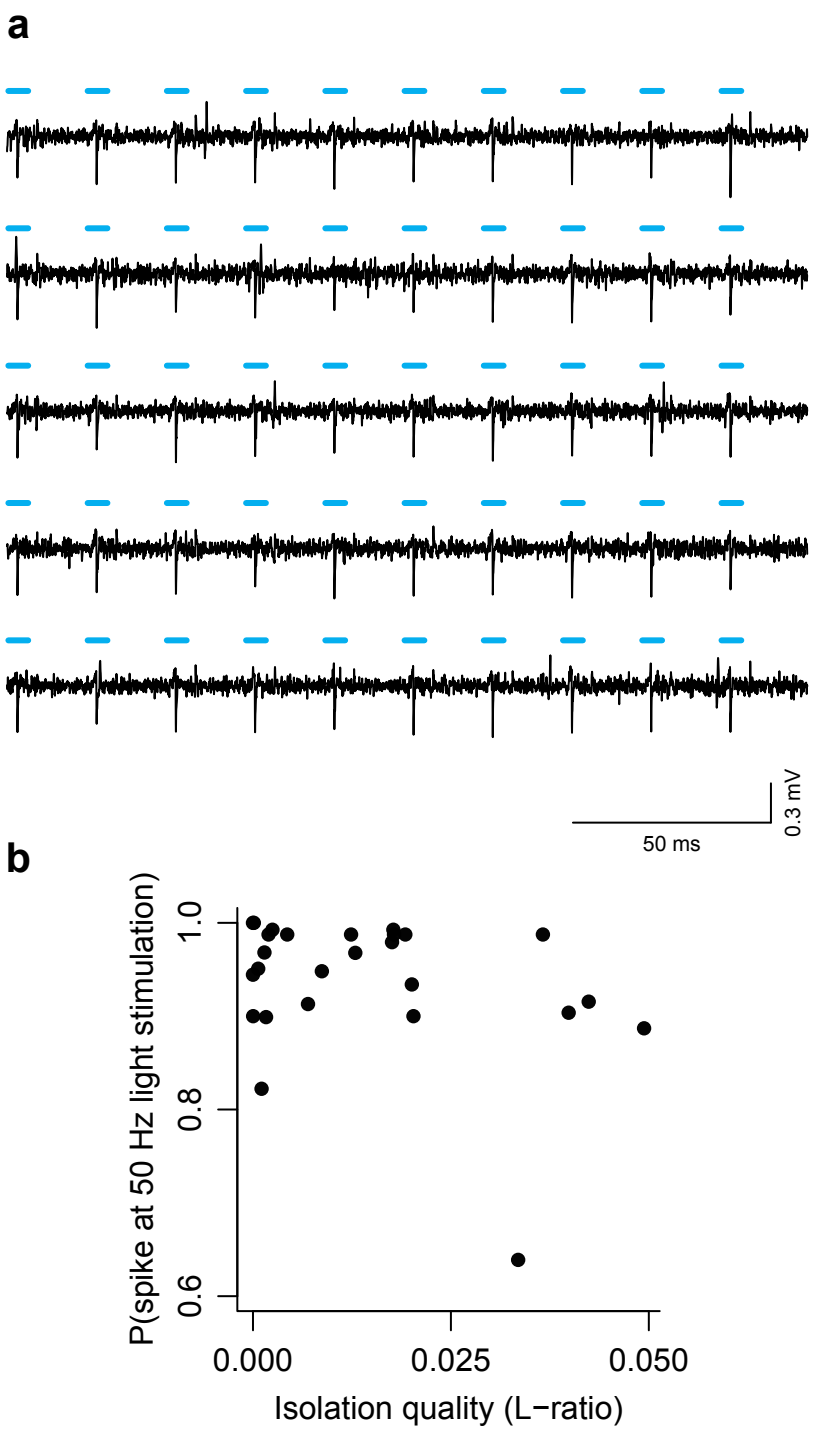


Figure S13. Dopaminergic neurons follow high-frequency light stimulation. **a**, Five consecutive trials of 50 Hz light stimulation from a dopaminergic neuron. **b**, Probability of following 50 Hz stimulation as a function of isolation quality, measured using the L-ratio. Smaller values indicate better isolation. The L-ratio for the example neuron is 3.6×10^{-8}.