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], and **Naoshige Uchida**[1]

[1]Department of Molecular and Cellular Biology, Center for Brain Science, Harvard University, Cambridge, MA 02138

[2]Division of Endocrinology, Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, 02215

Abstract

Dopamine plays a key role in motivation and reward. Dopaminergic neurons in the ventral tegmental area (VTA) signal the discrepancy between expected and actual rewards (i.e., reward prediction error, RPE)[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 neurons 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 RPE 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*, 17%) by the value of outcomes. While the activity of *Type I* neurons was sensitive to actual outcomes (i.e., when the reward was delivered as expected vs. unexpectedly omitted), the activity of *Types II* and *III* neurons was determined predominantly by reward-predicting odours. We "tagged" dopaminergic and GABAergic neurons with the light-sensitive protein channelrhodopsin-2 (ChR2) 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 RPE.

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 (i.e., they compute 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 toward 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 (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, S1). 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, while 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, while 17 neurons (18%) were classified as *Type III*, which showed sustained inhibition (Fig. 2d).

To identify dopaminergic neurons, we expressed 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*) gene (Fig. S2, S3). 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 ms of light onset with small jitter, and followed high-frequency light stimulation of 50 Hz (Fig. S4). 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 *Types 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 *Types II* or *III* neurons. We recorded from 92 VTA neurons in mice expressing Cre recombinase under the control of the endogenous vesicular GABA transporter (*Vgat*) gene. These mice showed similar licking behaviour to *DAT*-Cre mice (Fig. S5). We applied the PCA parameters 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, S4). 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 (Fig. S6).

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] (Fig. S7), 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 17 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 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 (9/17 neurons $P < 0.05$, t-test; mean auROC = 0.407, $t_{16} = 2.56$, $P < 0.05$; Fig. S8a,b). Thus, the majority of 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 16/17 identified GABAergic neurons, Fig. S7a; regression slopes, Fig. S10i). 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, S8), in contrast to identified dopaminergic neurons. None of the identified GABAergic neurons, and only two 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 *Types II* and *III* neurons was modulated predominantly by reward-predicting cues but not actual reward.

Recent studies have revealed a diversity of dopaminergic neurons in their responses to aversive stimuli: some are excited, others inhibited[15]. To test whether this diversity exists in dopaminergic and GABAergic VTA neurons, we delivered airpuffs in some sessions. Identified dopaminergic neurons showed some diversity: while most significant responses were inhibition, some were excitation (Fig. 4h,i, S9). In contrast, most *Types II* and *III* neurons (and 13/14 identified GABAergic neurons) were excited by airpuffs.

Detecting the discrepancy between expected and actual outcomes plays 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 (Fig. S11a)[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] (Fig. S11b). VTA GABAergic neurons synapse preferentially onto dendrites of dopaminergic neurons[28], while other inhibitory inputs synapse 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 *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 fiber. 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. Following experiments, we performed immunohistochemistry to localize recording sites amid dopaminergic neurons.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. Science. 1997; 275:1593–1599. [PubMed: 9054347]

2. Bayer HM, Glimcher PW. Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron. 2005; 47:129–141. [PubMed: 15996553]

3. Schultz W. Behavioral theories and the neurophysiology of reward. Annu Rev Psychol. 2006; 57:87–115. [PubMed: 16318590]

4. Swanson LW. The projections of the ventral tegmental area and adjacent regions: a combined fluorescent retrograde tracer and immunofluorescence study in the rat. Brain Res Bull. 1982; 9:321–353. [PubMed: 6816390]

5. Margolis EB, Lock H, Hjelmstad GO, Fields HL. The ventral tegmental area revisited: is there an electrophysiological marker for dopaminergic neurons? J Physiol. 2006; 577:907–924. [PubMed: 16959856]

6. Nair-Roberts RG, et al. Stereological estimates of dopaminergic, GABAergic and glutamatergic neurons in the ventral tegmental area, substantia nigra and retrorubral field in the rat. Neuroscience. 2008; 152:1024–1031. [PubMed: 18355970]

7. Hyman SE, Malenka RC, Nestler EJ. Neural mechanisms of addiction: the role of reward-related learning and memory. Annu Rev Neurosci. 2006; 29:565–598. [PubMed: 16776597]

8. Lüscher C, Malenka RC. Drug-evoked synaptic plasticity in addiction: from molecular changes to circuit remodeling. Neuron. 2011; 69:650–663. [PubMed: 21338877]

9. Johnson SW, North RA. Opioids excite dopamine neurons by hyperpolarization of local interneurons. J Neurosci. 1992; 12:483–488. [PubMed: 1346804]

10. Mansvelder HD, Keath JR, McGehee DS. Synaptic mechanisms underlie nicotine-induced excitability of brain reward areas. Neuron. 2002; 33:905–919. [PubMed: 11906697]

11. Szabo B, Siemes S, Wallmichrath I. Inhibition of GABAergic neurotransmission in the ventral tegmental area by cannabinoids. Eur J Neurosci. 2002; 15:2057–2061. [PubMed: 12099913]

12. Tan KR, et al. Neural bases for addictive properties of benzodiazepines. Nature. 2010; 463:769–774. [PubMed: 20148031]

13. Dobi A, Margolis EB, Wang H-L, Harvey BK, Morales M. Glutamatergic and nonglutamatergic neurons of the ventral tegmental area establish local synaptic contacts with dopaminergic and nondopaminergic neurons. J Neurosci. 2010; 30:218–229. [PubMed: 20053904]

14. Steffensen SC, Svingos AL, Pickel VM, Henriksen SJ. Electrophysiological characterization of GABAergic neurons in the ventral tegmental area. J Neurosci. 1998; 18:8003–8015. [PubMed: 9742167]

15. Matsumoto M, Hikosaka O. Two types of dopamine neuron distinctly convey positive and negative motivational signals. Nature. 2009; 459:837–841. [PubMed: 19448610]

16. Lammel S, et al. Unique properties of mesoprefrontal neurons within a dual mesocorticolimbic dopamine system. Neuron. 2008; 57:760–773. [PubMed: 18341995]

17. Nagel G, et al. Channelrhodopsin-2, a directly light-gated cation-selective membrane channel. Proc Natl Acad Sci U S A. 2003; 100:13940–13945. [PubMed: 14615590]

18. Boyden ES, Zhang F, Bamberg E, Nagel G, Deisseroth K. Millisecond-timescale, genetically targeted optical control of neural activity. Nat Neurosci. 2005; 8:1263–1268. [PubMed: 16116447]

19. Atasoy D, Aponte Y, Su HH, Sternson SM. A FLEX switch targets Channelrhodopsin-2 to multiple cell types for imaging and long-range circuit mapping. J Neurosci. 2008; 28:7025–7030. [PubMed: 18614669]

20. Fiorillo CD, Newsome WT, Schultz W. The temporal precision of reward prediction in dopamine neurons. Nat Neurosci. 2008; 11:966–973. [PubMed: 18660807]

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. 2004; 92:2520–2529. [PubMed: 15163669]

22. Rescorla, RA.; Wagner, AR. Classical Conditioning II: Current Research and Theory. Black, AH.; Wagner, AR., editors. New York: Appleton-Century-Crofts; 1972. p. 64-99.

23. Houk, JC.; Adams, JL.; Barto, AG. Models of Information Processing in the Basal Ganglia. Houk, JC.; Davis, JL.; Beiser, DG., editors. MIT Press; 1995. p. 249-270.

24. Carr DB, Sesack SR. Projections from the rat prefrontal cortex to the ventral tegmental area: target specificity in the synaptic associations with mesoaccumbens and mesocortical neurons. J Neurosci. 2000; 20:3864–3873. [PubMed: 10804226]

25. Okada K, Toyama K, Inoue Y, Isa T, Kobayashi Y. Different pedunculopontine tegmental neurons signal predicted and actual task rewards. J Neurosci. 2009; 29:4858–4870. [PubMed: 19369554]

26. Masumoto M, Hikosaka O. Lateral habenula as a source of negative reward signals in dopamine neurons. Nature. 2007; 447:1111–1115. [PubMed: 17522629]

27. Takahashi YK, et al. Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. Nat Neurosci. in press. 10.1038/nn.2957

28. Omelchenko N, Sesack SR. Ultrastructural analysis of local collaterals of rat ventral tegmental area neurons: GABA phenotype and synapses onto dopamine and GABA cells. Synapse. 2009; 63:895–906. [PubMed: 19582784]

29. Jhou TC, Fields HL, Baxter MG, Saper CB, Holland PC. The rostromedial tegmental nucleus (RMTg), a GABAergic afferent to midbrain dopamine neurons, encodes aversive stimuli and inhibits motor responses. Neuron. 2009; 61:786–800. [PubMed: 19285474]

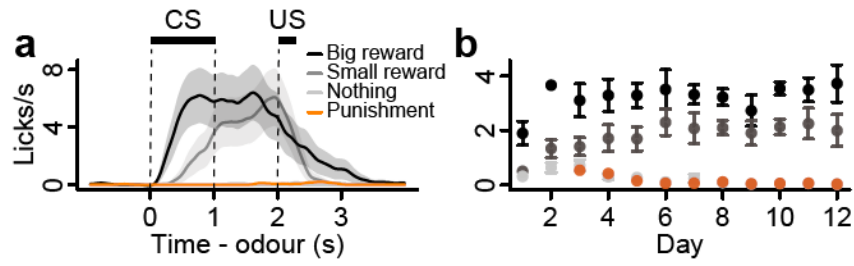30. Redish AD. Addiction as a computational process gone awry. Science. 2004; 306:1944–1947. [PubMed: 15591205]

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 SEM. **b**, Mean ± SEM licks during the delay between CS and US as a function of days of the experiment across animals.
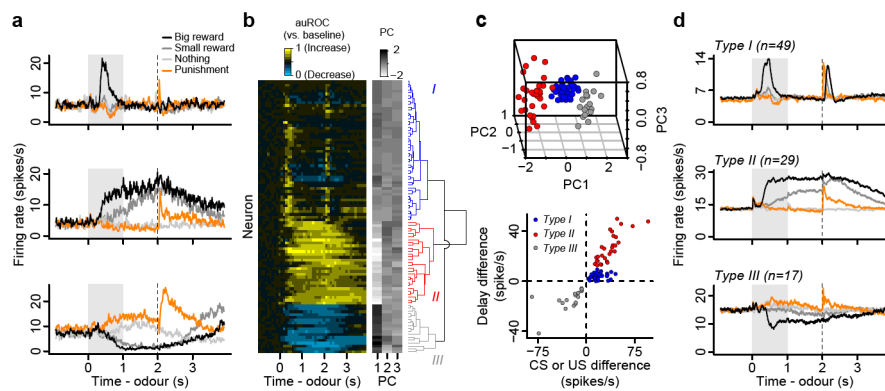
Figure 2. VTA neurons show three distinct response types

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 *Types I - III* neurons.
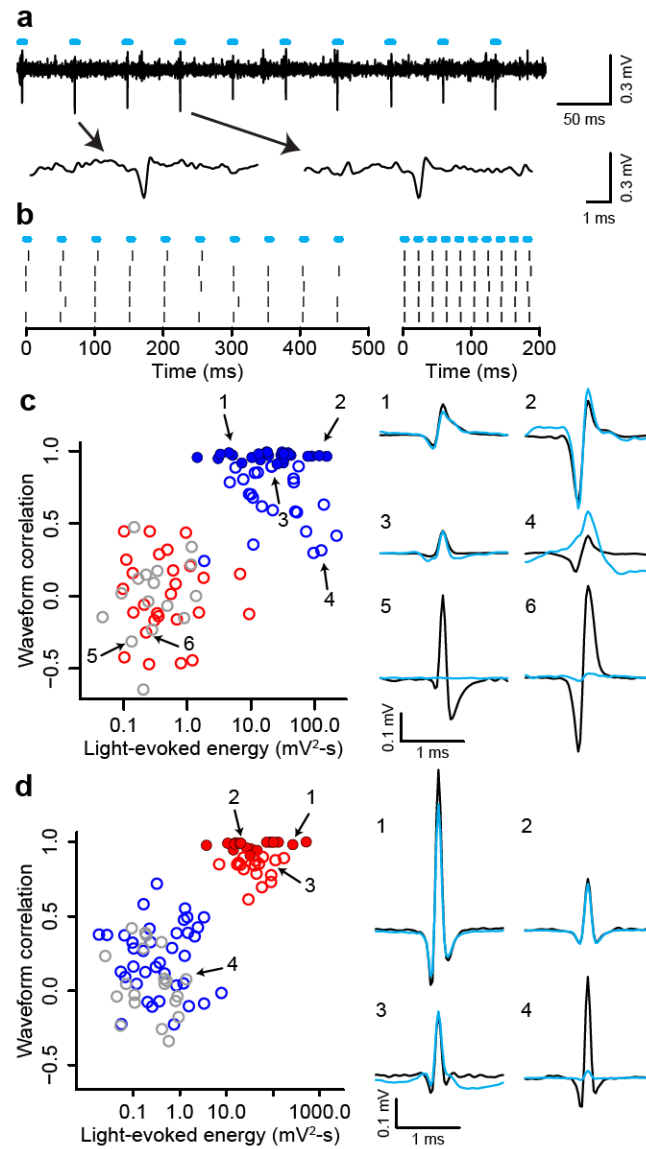
Figure 3. Identifying dopaminergic and GABAergic neurons
a, Voltage trace from 10 pulses of 20 Hz light stimulation (cyan bars). Two light-triggered spikes are shown below. **b**, Response from this neuron to 20 Hz (left) and 50 Hz (right) stimulation. Ticks represent spikes. **c**, Quantification of 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**).
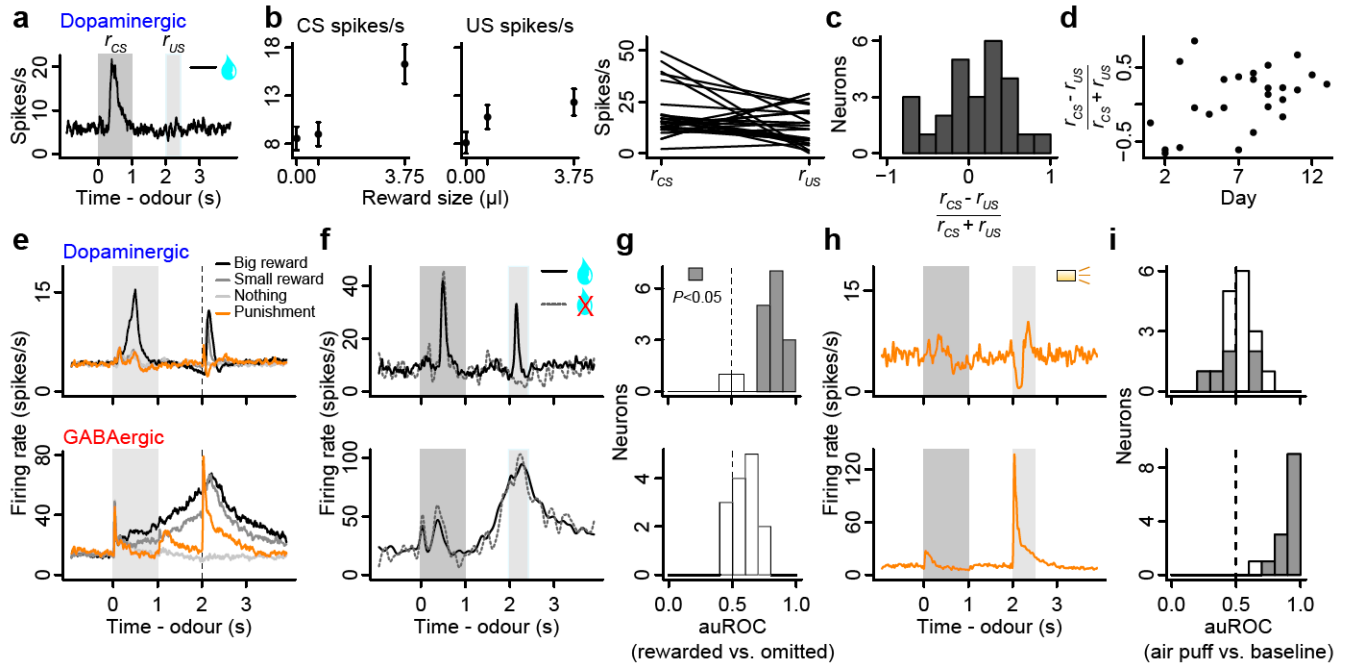
Figure 4. Response variability based on CS-US preference, reward omission and air puffs
a, Response of a dopaminergic neuron during big-reward trials. **b**, Firing rate (mean ± SEM) vs. reward size (left) and in response to big-reward-predicting CS and big-reward US for each dopaminergic neuron (right). **c**, Histogram of CS-US index for dopaminergic neurons. **d**, CS-US index vs. 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 airpuff (2-2.5 s) relative to baseline (<0.5, decrease; >0.5, increase from baseline).