Image from https://neurosciencenews.com/decision-making-brain-19504/

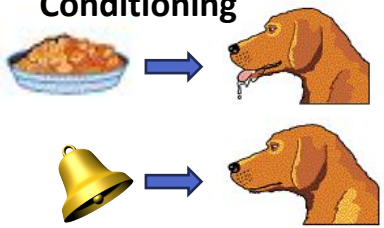
Section: Cohen et al. 2012

Background Recap:

Section W2

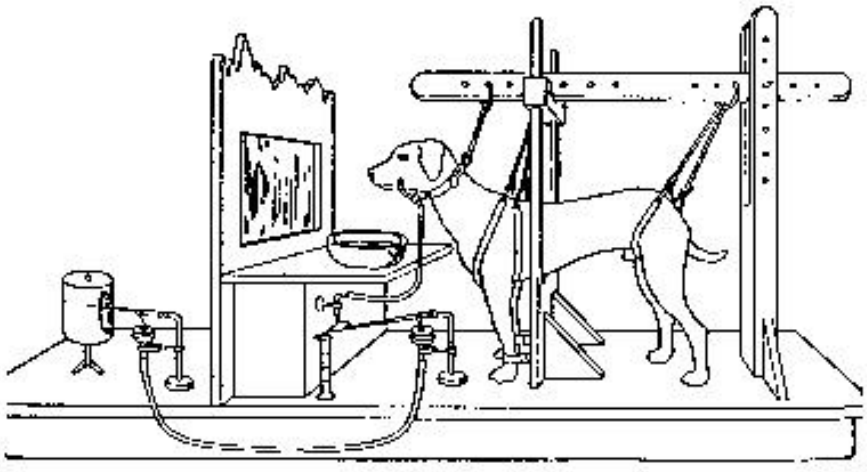
Pavlov's classical conditioning

Before Conditioning



During Conditioning

After Conditioning

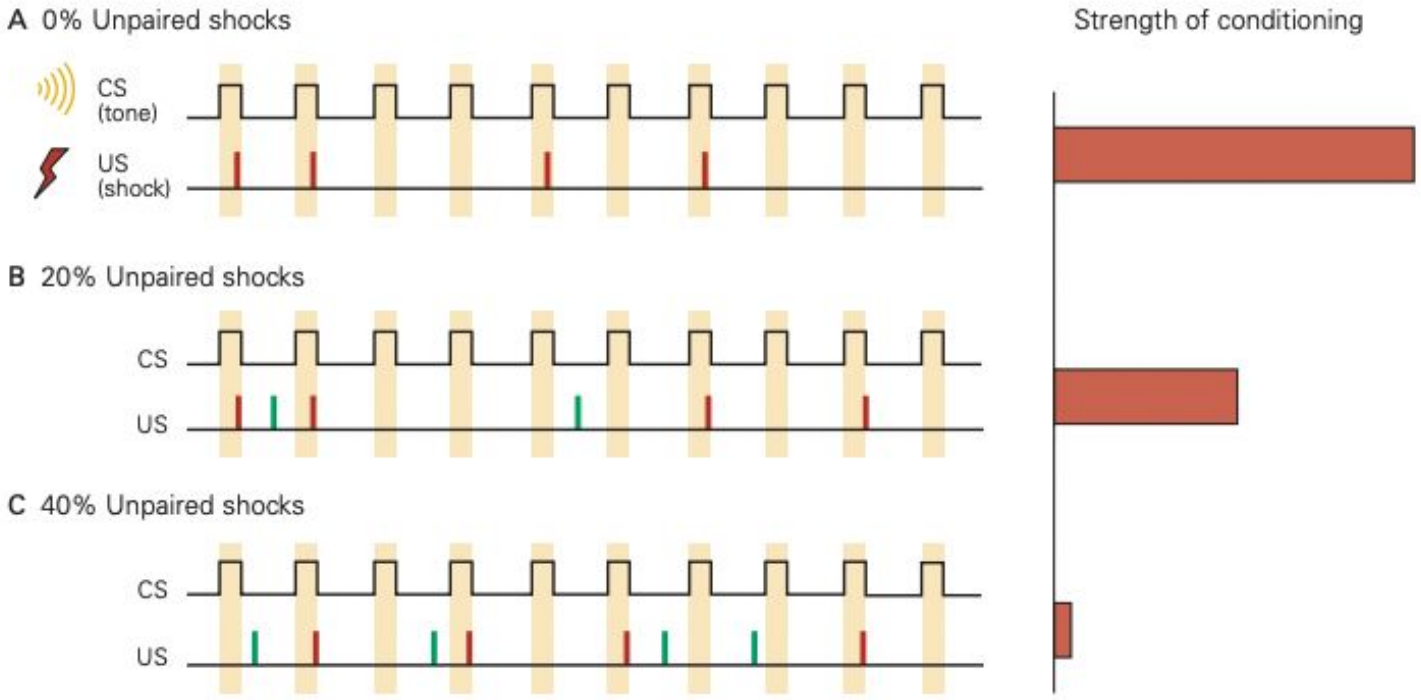


https://en.wikipedia.org/wiki/Classical_conditioning#/media/File:Ivan_Pavlov_research_on_dog's_reflex_setup.jpg

Ten of the more photogenic of Pavlov's dogs. Krasavietz (upper left), Beck, Milkah, Ikar, Joy, Tungus, Arleekin, Ruslan, Toi and Murashka (bottom right). The rest of Pavlov's dogs and their corresponding *Drosophila* memory mutants can be found on the author's webpage at www.cshl.org.

https://www.sciencedirect.com/science/article/pii/S0960982203000666

Classical conditioning depends on degree of stimulus-outcome correlation

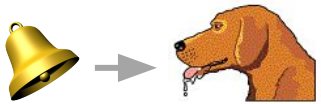


Kandel. Figure 65–12

Kamin's blocking experiment

1. Conditioning

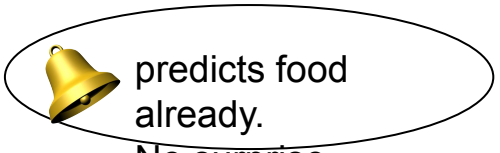
2. After conditioning



3. 2nd conditioning

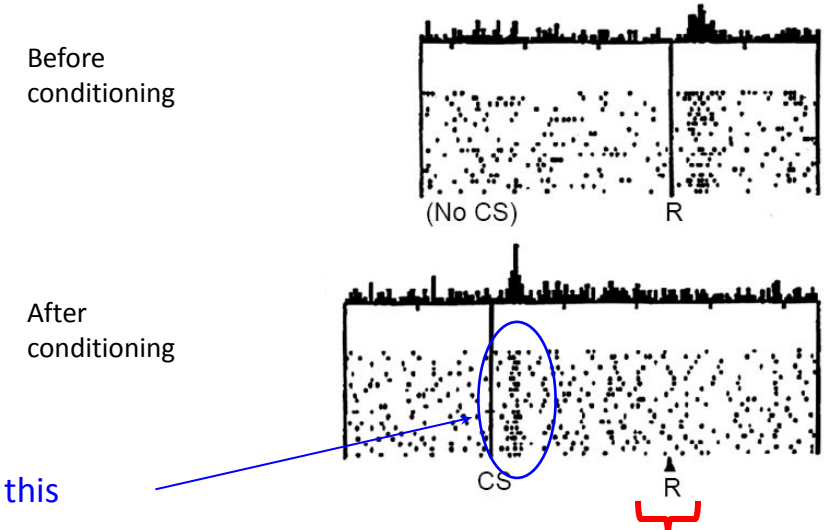
4. Test

Kamin, L. J. (1969). Predictability, Surprise, Attention, and Conditioning. In B. A. Campbell, & R. M. Church (Eds.), Punishment Aversive Behavior (pp. 279-296). New York: Appleton- Century-Crofts



- **Learning occurs only when expectation is violated!**
- *What is the neural basis of this?*

Dopamine neurons in the ventral tegmental area

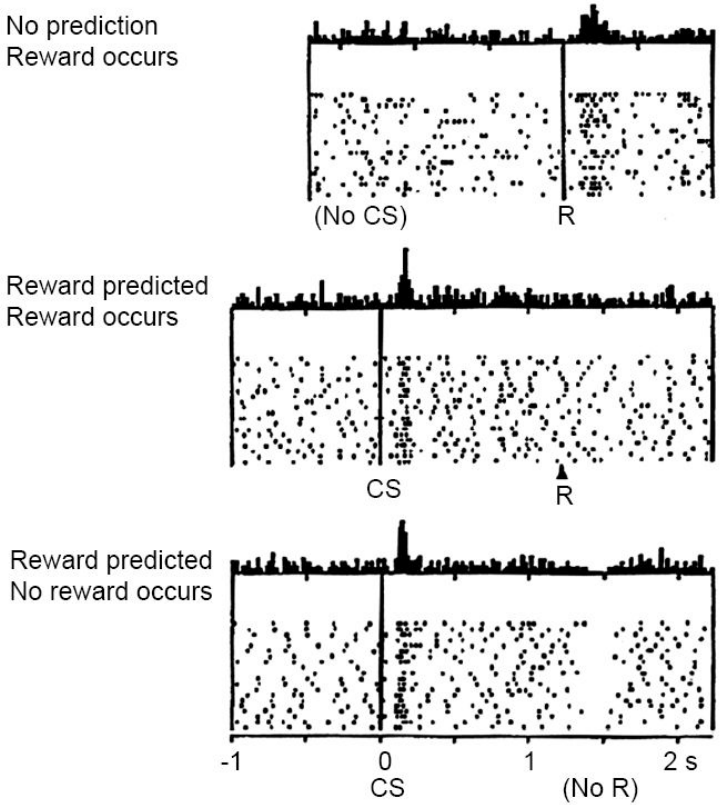


What is this for?

- Lack of reward responses when the reward was fully predicted

(Schultz, Dayan, Montague, 1997)

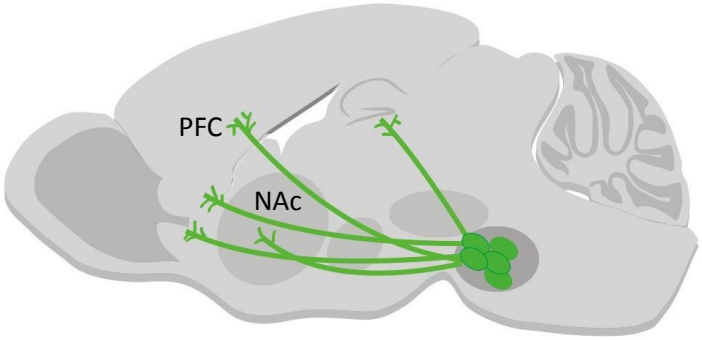
Dopamine as ~~reward~~ temporal difference (TD) error: reward prediction errors!

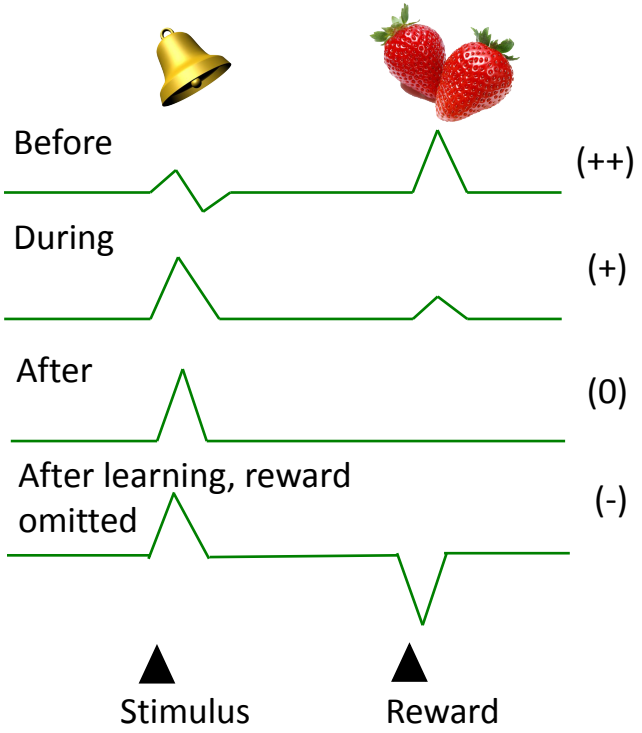


- Dopaminergic (DA) neurons fire phasically (100–500 ms) after unpredicted rewards or cues that predict reward.
- Their response to reward is reduced when a reward is fully predicted (the phasic firing happens at cue presentation).
- DA activity is suppressed when a predicted reward is omitted (negative prediction error).

(Schultz, Dayan, Montague, 1997)

Dopamine circuitry of the brain



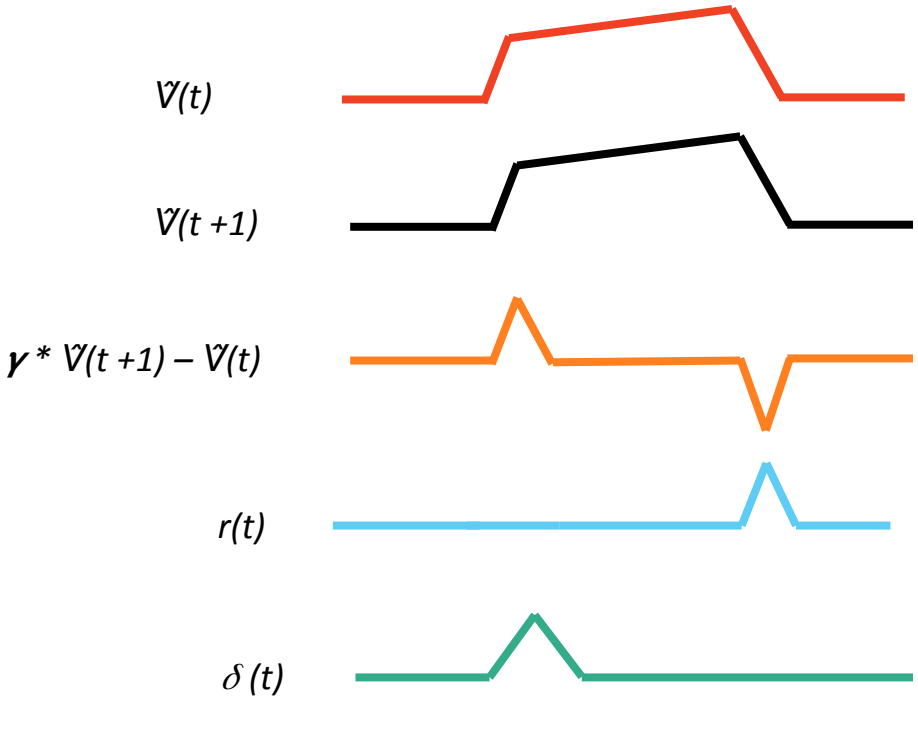


- Dopaminergic neurons are ~55–65% of VTA neurons
- The rest are mostly GABAergic inhibitory neurons or Glutamatergic neurons

How could a system encode a temporal difference (TD) error

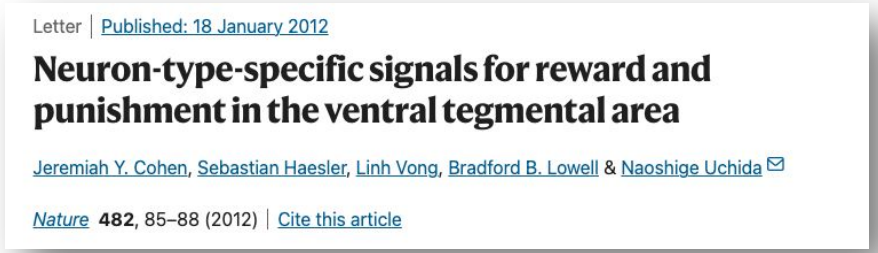
TD error as a derivative-like computation:
(neurally doable!)

$\delta(t) = r(t) + \gamma * V(t+1) - V(t)$



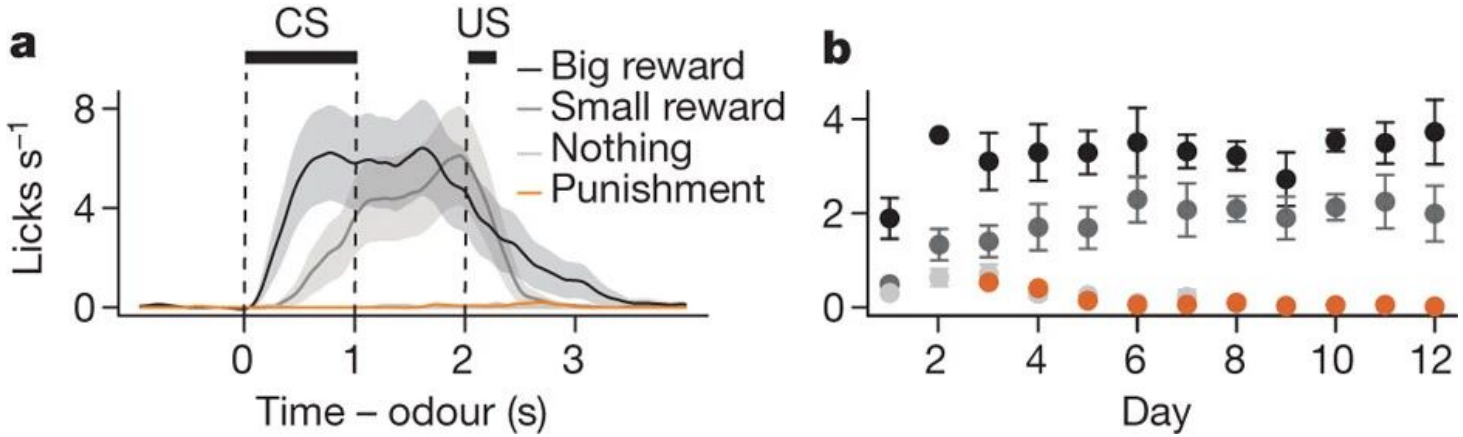
t = time
r = reward
$V(t)$ = value
γ = discount factor
δ = prediction error

Section Paper:



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)[1,2,3], 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.**

Figure 1: Odour-outcome association task in mice

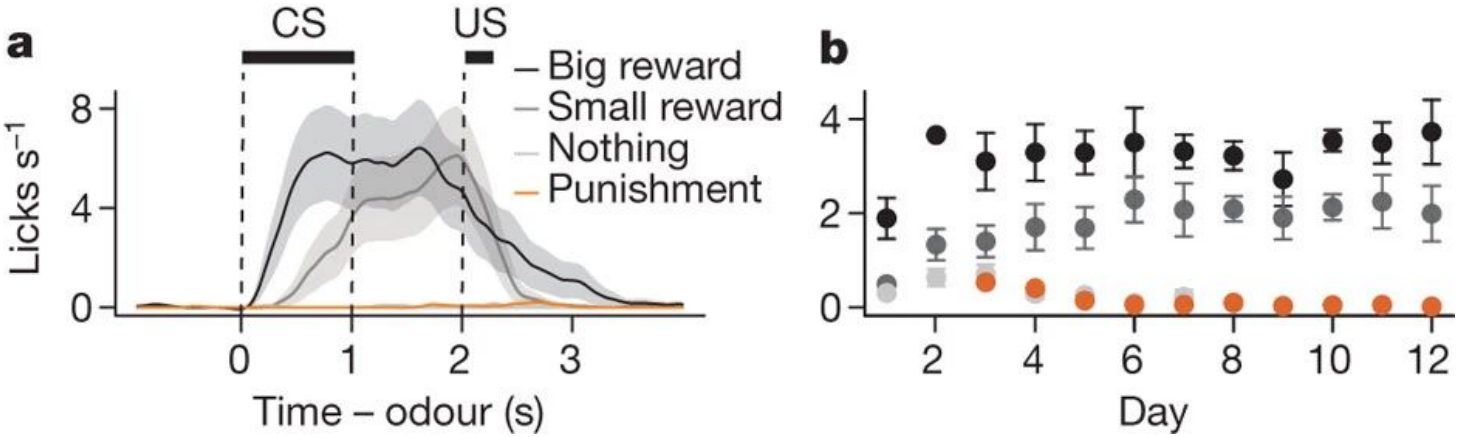


a, Licking behaviour from a representative experimental session. Black bars indicate CS and US delivery. Shaded regions around lick traces denote standard error of the mean (s.e.m.).
b, Mean ± s.e.m. licks during the delay between CS and US as a function of days of the experiment across animals.

Describe the behavioral task. What are the CS & US for each condition?

Odor A cue - big reward
Odor B cue - small reward
Odor C cue - no reward
Odor D cue - punishment – air puff.

Figure 1: Odour-outcome association task in mice

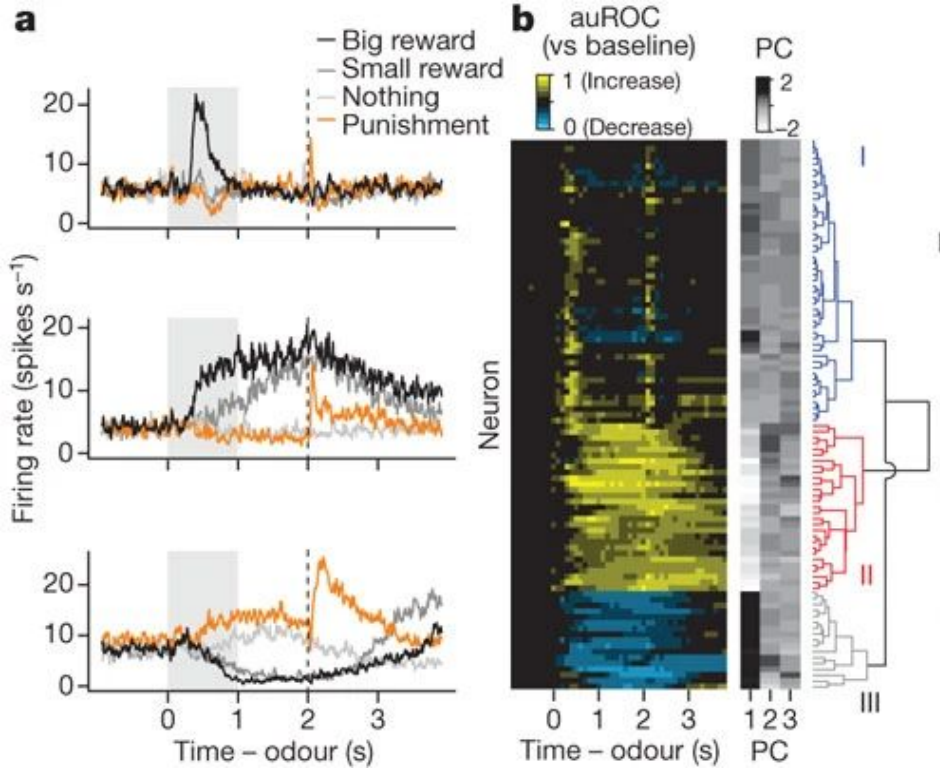


a, Licking behaviour from a representative experimental session. Black bars indicate CS and US delivery. Shaded regions around lick traces denote standard error of the mean (s.e.m.).
b, Mean ± s.e.m. licks during the delay between CS and US as a function of days of the experiment across animals.

What are the mice learning on day 1?
When do they hit asymptotic performance?

Mice began licking towards the water-delivery tube in the delay before rewards arrived → they quickly learned the CS–US associations. On day 6 they hit asymptotic performance for small reward and punishment, and on day 2 for big reward.

Figure 2: VTA neurons show three distinct response types



What is the baseline?

Baseline is the neural activity in a window of time of 1s before odor onset.

How was auROC calculated?

See next slide.

Reminder, how to calculated an auROC

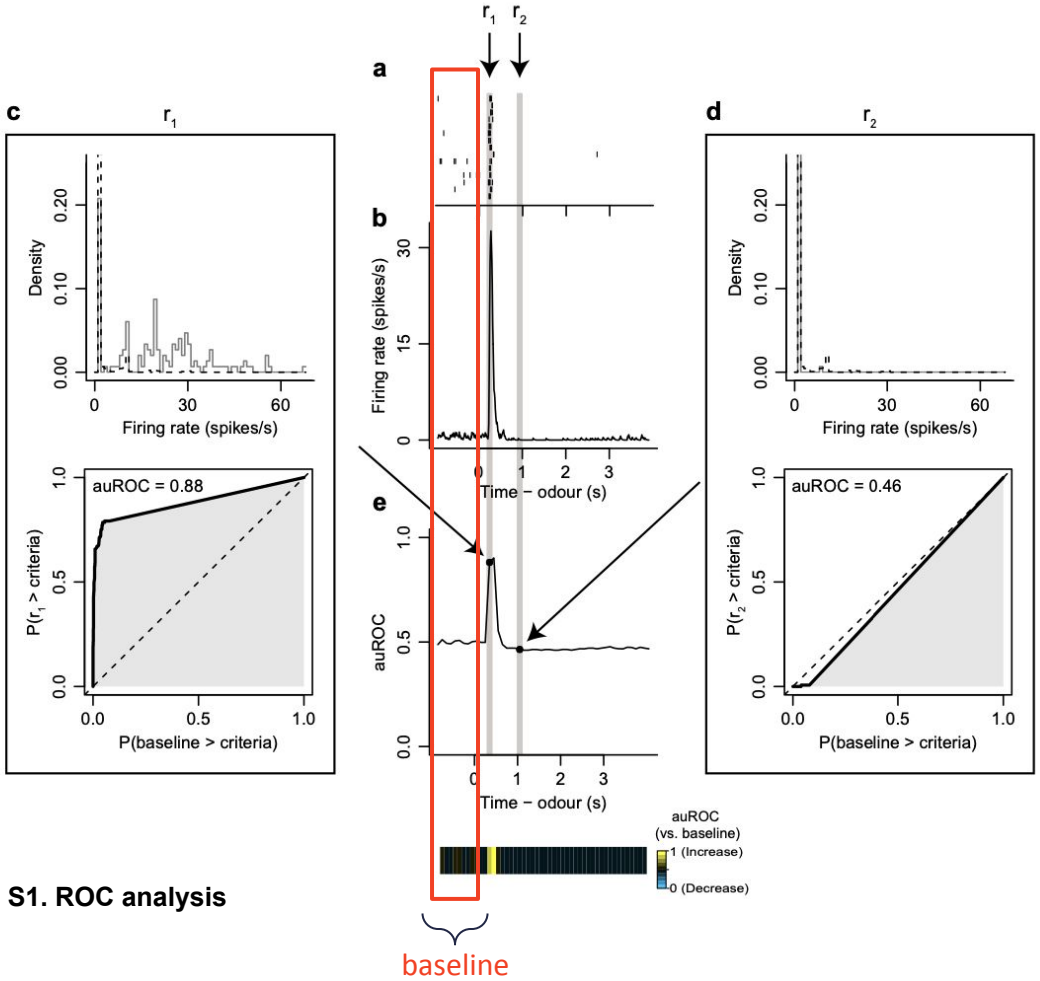


Figure S1. ROC analysis

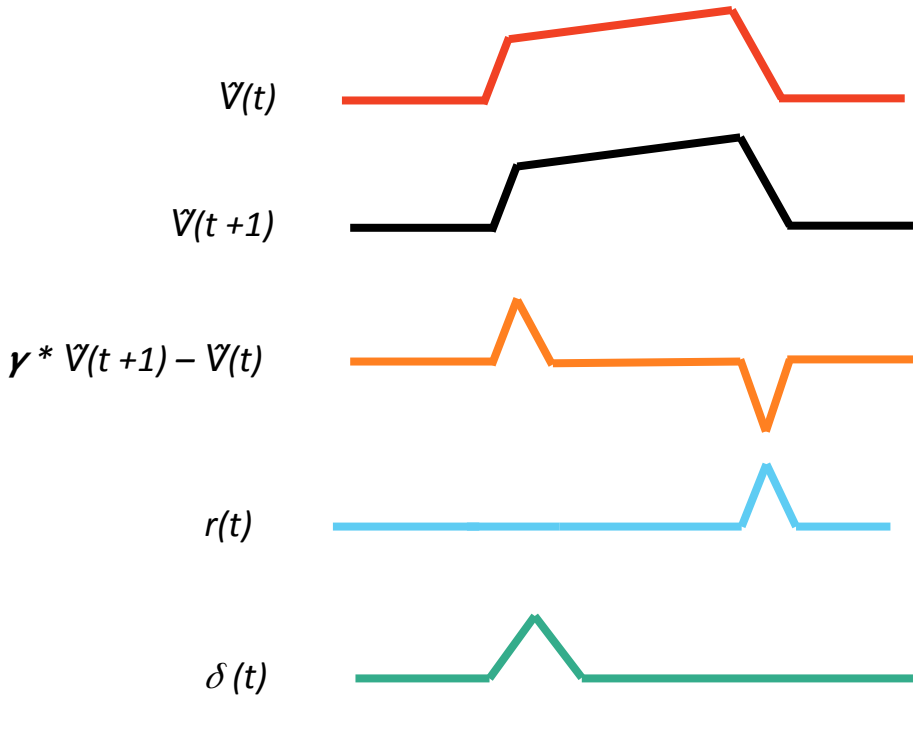
- *The auROC was calculated by comparing the histogram of spike counts during the baseline period with the histogram during a specific bin.*
- *The larger the auROC area above 0.5 is, the more spiking that occurred during the specific bin.*

Let's explore this more in the coding hands-on!

How could a system encode a temporal difference (TD) error

TD error as a derivative-like computation:
(neurally doable!)

$\delta(t) = r(t) + \gamma * V(t+1) - V(t)$

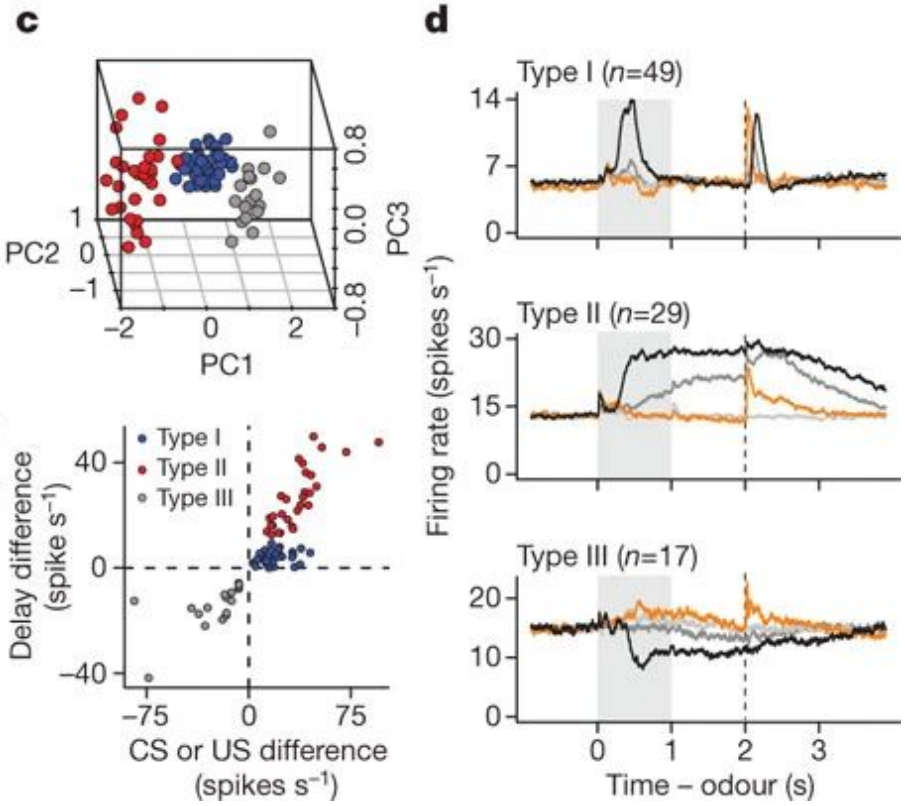


(derivative like)

Dopamine neurons

t = time
r = reward
$V(t)$ = value
γ = discount factor
δ = prediction error

Figure 2: VTA neurons show three distinct response types



What do you think these neurons types are encoding? You can also check (a).

Type 1: RPE
Type 2: Value encoding
Type 3: Ambiguous, could be a reversed value encoding

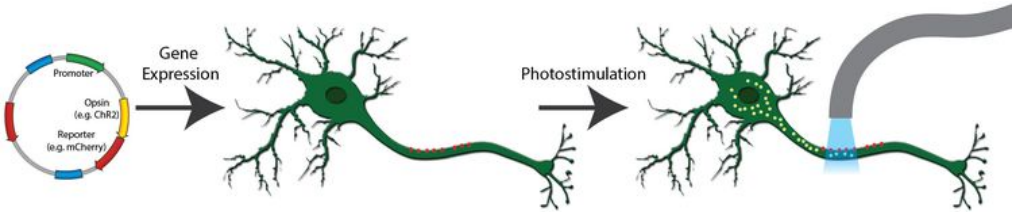
Do the Type 1 neurons match the canonical RPE coding? Explain.

Canonical RPE coding: shouldn't see phasic activity at the reward delivery (CS) only at US.

- *Mice didn't overtrained on it so still uncertainty, see D-RL in course.*
- *Time between US and CS (2sec) is long → play a role in certainty and conditioning.*

Figure 3: Identifying dopaminergic and GABAergic neurons

Cre recombinase under the control of DAT or VGAT gene



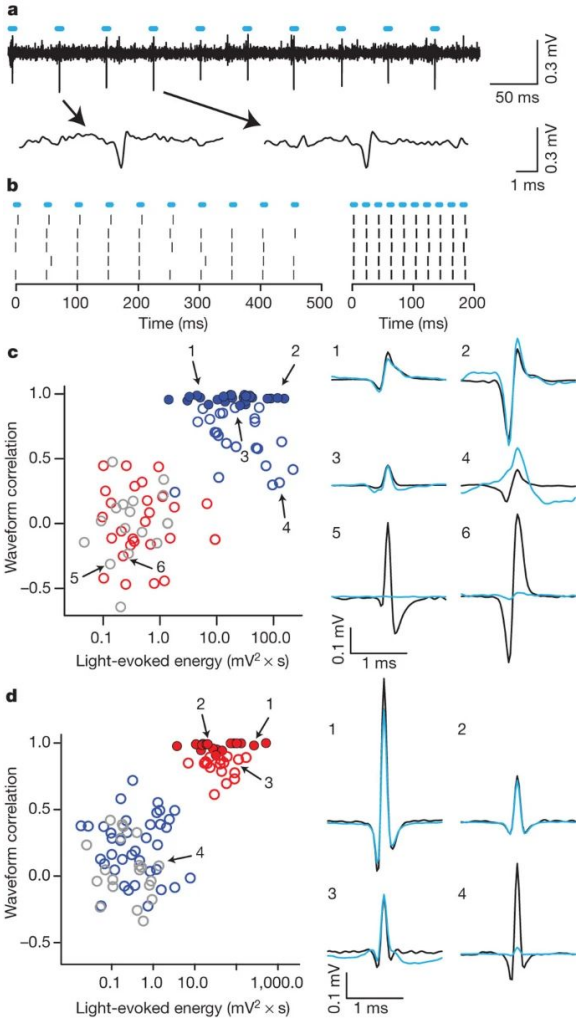


Figure 3: Identifying dopaminergic and GABAergic neurons

Explain how dopaminergic and GABAergic neurons were identified.

AAV injection (Channelrhodopsin-2) into mice expressing Cre recombinase under the control of the DAT or VGAT promoter → selective expression of ChR2 in neurons.

What were the criteria that was used in cell type categorization?

Researchers used waveform correlation between light-evoked and endogenous spikes. They also measured the spike amplitude (though the light-evoked energy calculation.

What type of cells give type 1 responses? What type give type 2?

Type 1: DA
Type 2: GABA

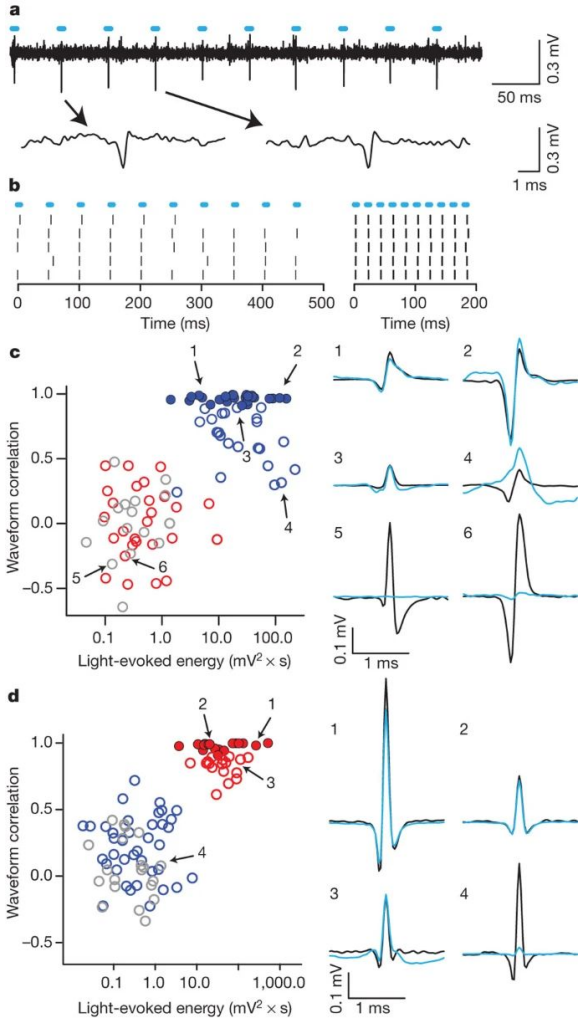
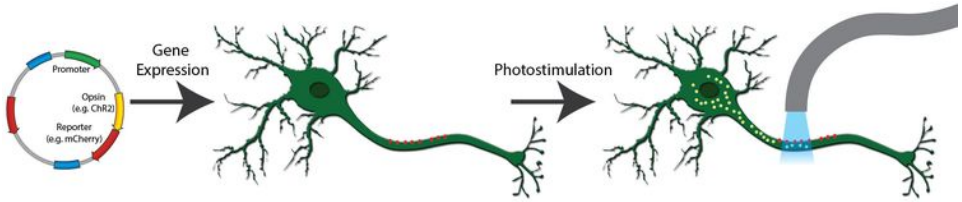


Figure 3: How dopaminergic neurons were identified?

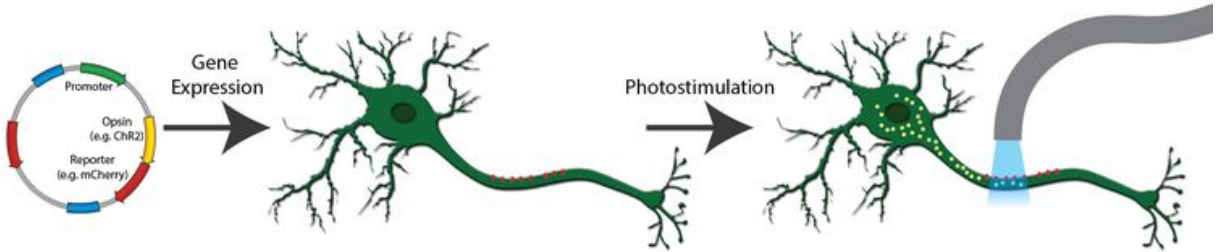
- *Channelrhodopsin-2 (ChR2): It is a protein that acts as a light-gated cation channel (recap from Week 1).*
- *Adeno-associated virus (AAV) with FLEX-ChR2: To introduce ChR2 into dopaminergic neurons, scientists use a modified virus called an adeno-associated virus (AAV). This virus is engineered to carry the gene for ChR2, but it's designed in a way (using a system called FLEX) that the gene can only be activated in specific circumstances (recap from Week 1).*
- *Cre recombinase and FLEX system: This is a genetic engineering technique used to control where and when a particular gene is expressed. In this case, they use transgenic mice that have been genetically modified to produce an enzyme called Cre recombinase specifically in dopaminergic neurons. The Cre enzyme can activate the ChR2 gene carried by the AAV, but only in cells where Cre is present. The FLEX system ensures that the ChR2 gene is only expressed in the presence of Cre recombinase.*
- *Dopamine Transporter (DAT) Promoter: The specificity for dopaminergic neurons is achieved by using the promoter of the DAT gene (a promoter is a region of DNA that initiates the transcription (activation) of a particular gene). The DAT gene is active specifically in dopaminergic neurons, so by using its promoter, the Cre recombinase enzyme is produced only in these neurons. This means that the ChR2 will also be expressed only in these neurons, thanks to the FLEX system.*



This leads to the selective expression of ChR2 in dopamin/GABAergic neurons. With ChR2 present, these neurons can now be controlled by light!

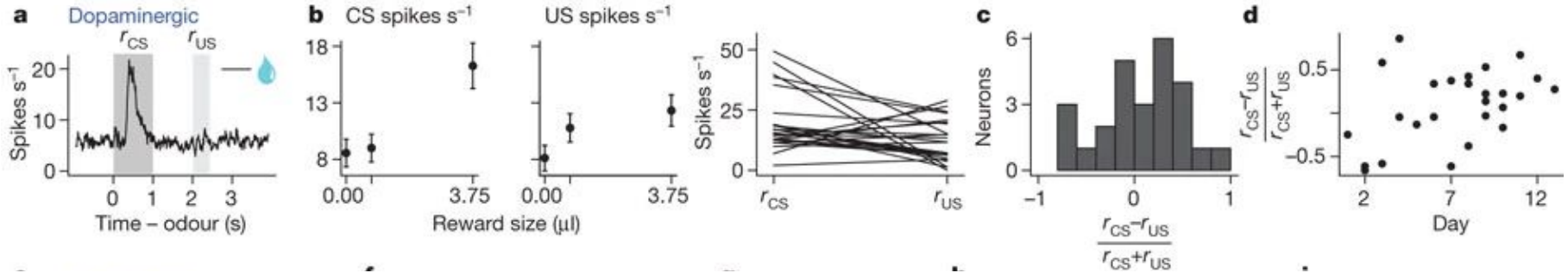
Figure 3: How dopaminergic neurons were identified?

- **ChR2 (Channelrhodopsin-2):** Light-gated cation channel (recap Week 1).
- **AAV-FLEX-ChR2:**
  - AAV delivers **ChR2 gene**.
  - **FLEX system** ensures activation only in specific conditions.
- **Cre Recombinase & FLEX System:**
  - **Cre enzyme** expressed in dopaminergic neurons (via transgenic mice).
  - **FLEX system** ensures ChR2 activation only in Cre+ cells.
- **DAT Promoter:**
  - Controls **Cre recombinase** expression in dopaminergic neurons.
  - Ensures **ChR2 is only in these neurons**.



This leads to the selective expression of ChR2 in dopamin/GABAergic neurons. With ChR2 present, these neurons can now be controlled by light!

Figure 4: Response variability based on CS–US preference, reward omission and air puffs



How do you expect firing of a Dopaminergic neuron to change across learning?

The response should change, from peaking at the US, to peaking at the CS.

Do you think the neuron in Fig. 4a is from before or after learning?

Neuron peaks after odor onset (CS) so after learning.

Figure 4: Response variability based on CS–US preference, reward omission and air puffs

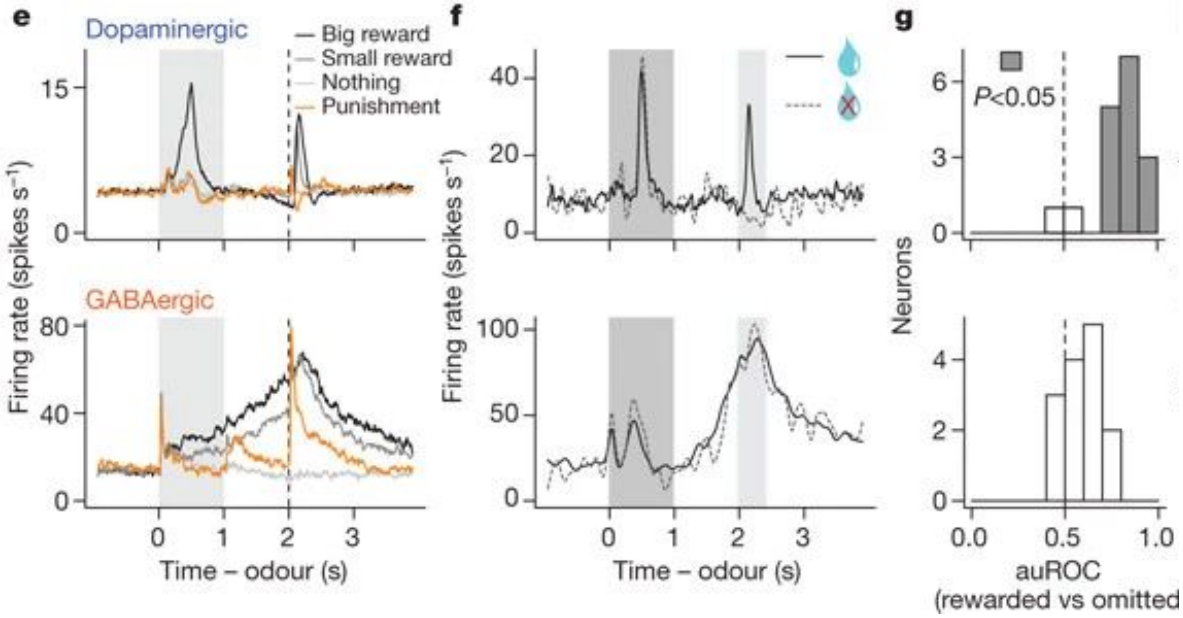
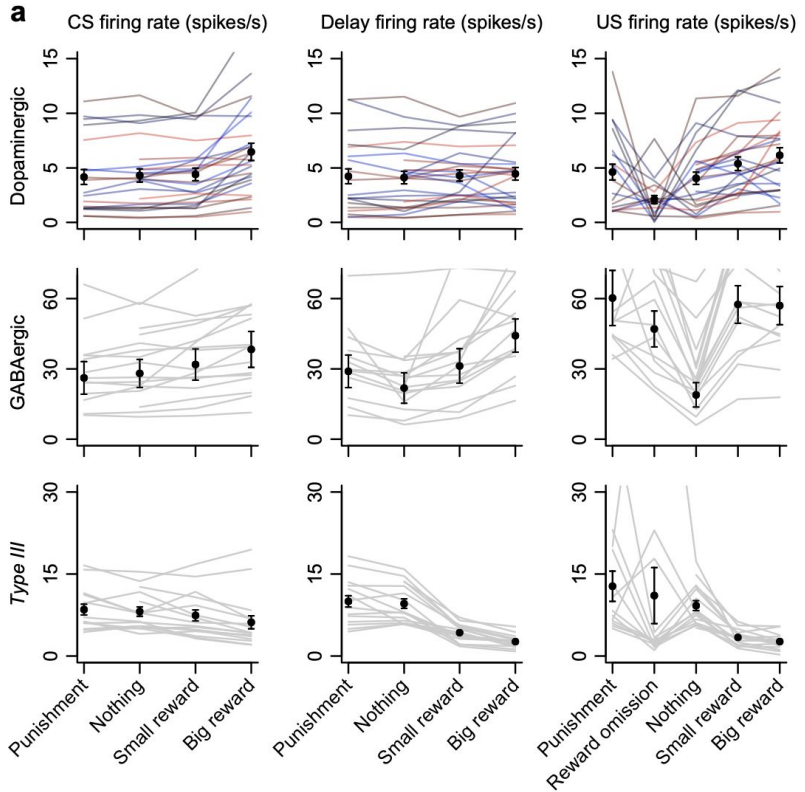


Fig. 4f-g: What is the response of dopaminergic neurons when the expected reward was omitted? Does it support RPE coding and why?

*Dopaminergic neurons decrease their firing rate when an anticipated reward is omitted (**negative prediction error**) which is aligned with canonical RPE coding.*

Figure 4: Response variability based on CS–US preference, reward omission and air puffs



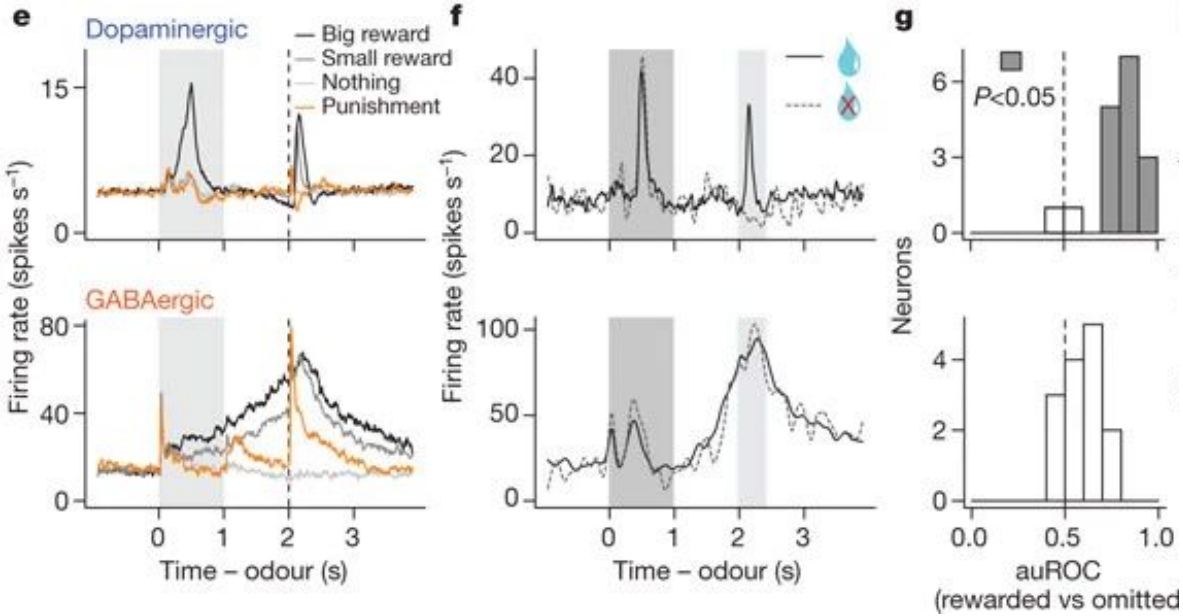
Supplementary Figure 7a

Based on the figure S7a what can we say about fairing patterns of GABAergic neurons during the delay period? What does it tell us about their encoding properties?

GABAergic neurons shows persistent activity during the delay and encode the value of upcoming potential outcomes (no-, small- and big-reward).

This suggests that these neurons encode expectation about rewards.

Figure 4: Response variability based on CS–US preference, reward omission and air puffs



Look at the GABAergic neurons response to presence / omission of reward. Does this support the hypothesis of value encoding?

GABAergic neurons maintain a consistent level of activity regardless of whether a reward is delivered or omitted. This supports GABAergic neurons encoding prediction of the reward rather than presence or absence of actual reward.

Conclusions

Our dataset 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 variability20,21 (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 uncertainty20. 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 training21. Thus, identified dopaminergic neurons appear to respond to CS and US similarly to those reported in non-human primate studies.

Paper round-up

- They identify 3 types of neurons in the ventral tegmental area.
- They differentiate dopaminergic and GABAergic neurons using optogenetic tools.
- They characterize dopaminergic neurons diversity (excited by either reward, reward-predicting CS or both) which seems to be related to the effect of training.
- They show that some dopaminergic neurons might not strictly follow canonical RPE coding.
- They show that GABAergic neurons parametrically encoded the value of upcoming outcomes.

What did we learn? What questions do we have?

- **What points do they make in the discussion?**

- **Is anything unclear?**

- **What would you do next if you had to design an experiment?**

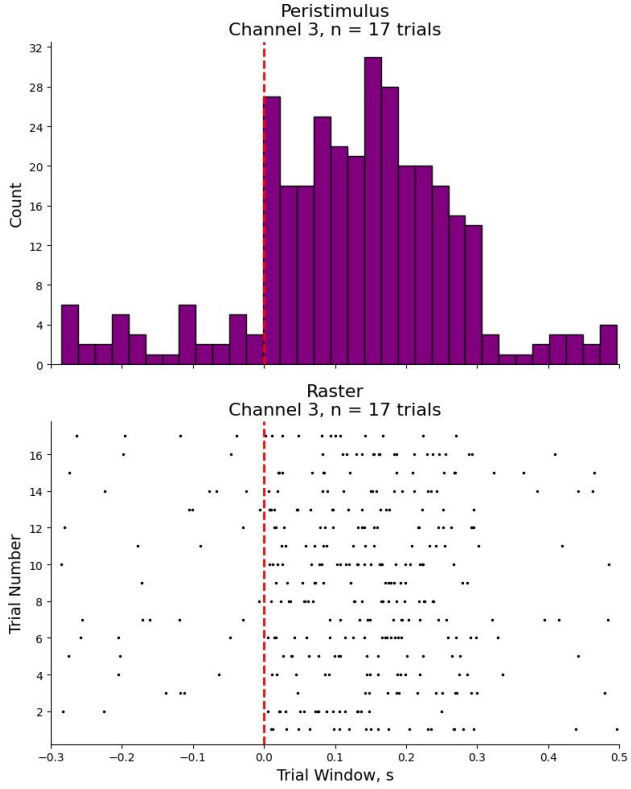
- *Author's next work: serotonergic neurons on the same task (+ dopaminergic): https://elifesciences.org/articles/6346.*
- *Investigate the provenance of the input signals to GABAergic neurons in VTA between phasic excitation (aversive stimuli) and sustained activity between US and CS (value encoding): https://www.sciencedirect.com/science/article/pii/S0896627312002814?via%3Dihub.*
- *How is value encoded in the brain more specifically (why type 2 and 3)? Distributed RL: https://www.nature.com/articles/s41586-019-1924-6.*

Key concept: peri-stimulus time histogram

The Peri-Stimulus Time Histogram (PSTH) plots the average firing rate of a neuron over time relative to the onset of a stimulus. Here's how it's typically calculated:

1. Define a time window around the onset of the stimulus.
2. Divide this time window into small bins.
3. Count the number of spikes (action potentials) that occur within each bin across multiple trials.
4. Average the spike counts across trials for each bin.
5. Plot the average spike count (firing rate) for each bin as a function of time.

https://github.com/MMathisLab/Nx-435_EPFL



https://colab.research.google.com/github/MMathisLab/Nx-435_EPFL/blob/main/Notebooks/Demo_PSTH.ipynb