Dopamine and value

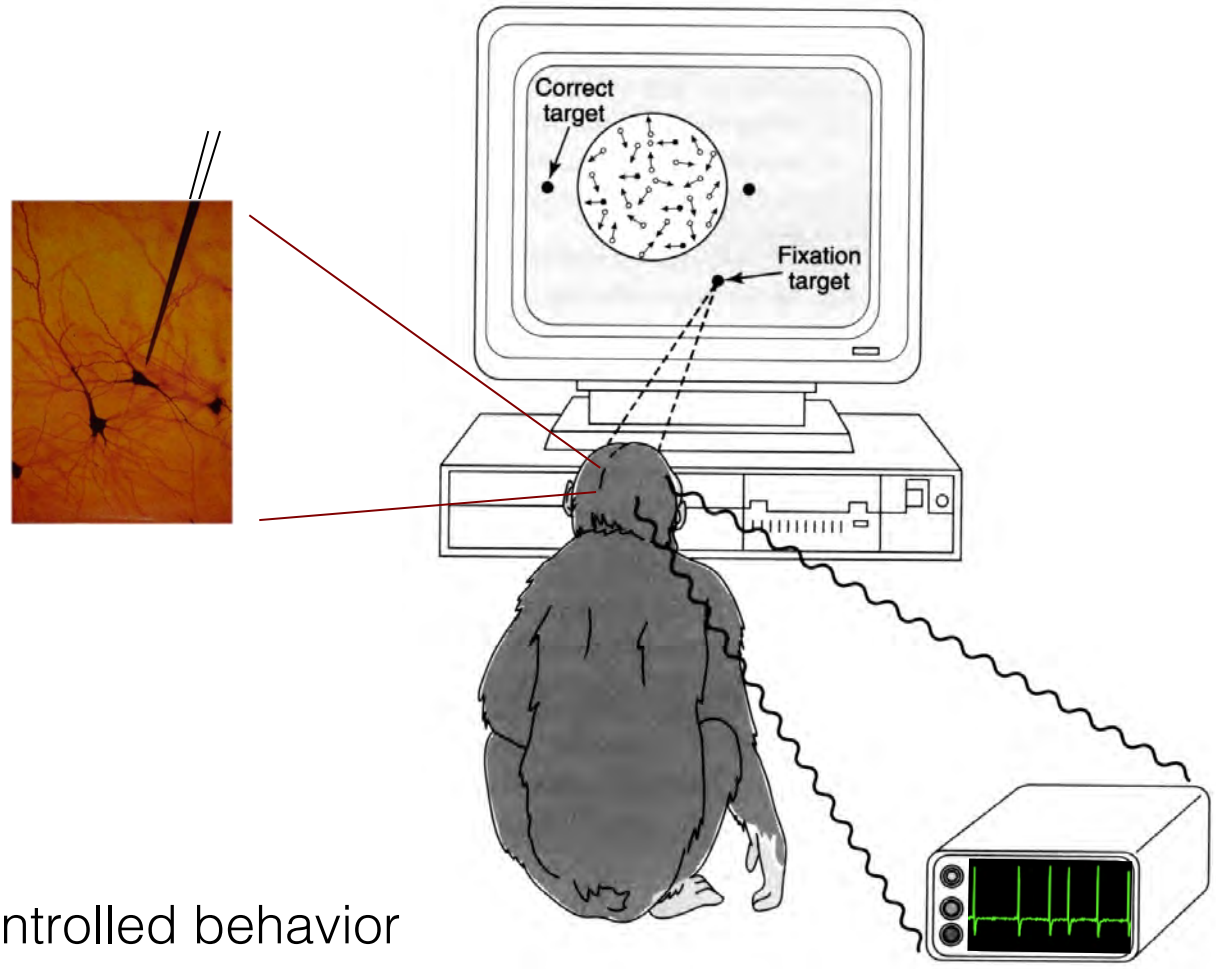
Naoshige Uchida
Center for Brain Science
Dept. of Molecular and Cellular Biology
Harvard University

Understanding the brain

- Computational theory / goal
- Algorithm
- Implementation

The three levels by David Marr

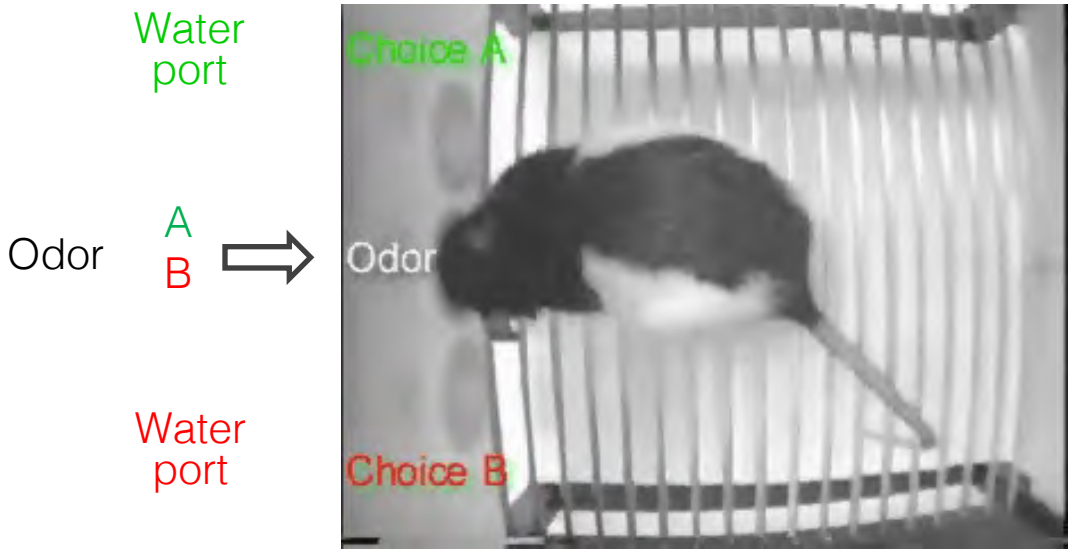
Single neuron recording in alert monkey



- Well-controlled behavior
- Recording in awake animal

(After Glimcher)

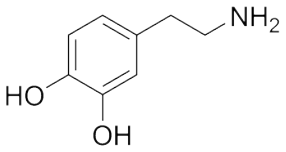
Our approach: rodent models

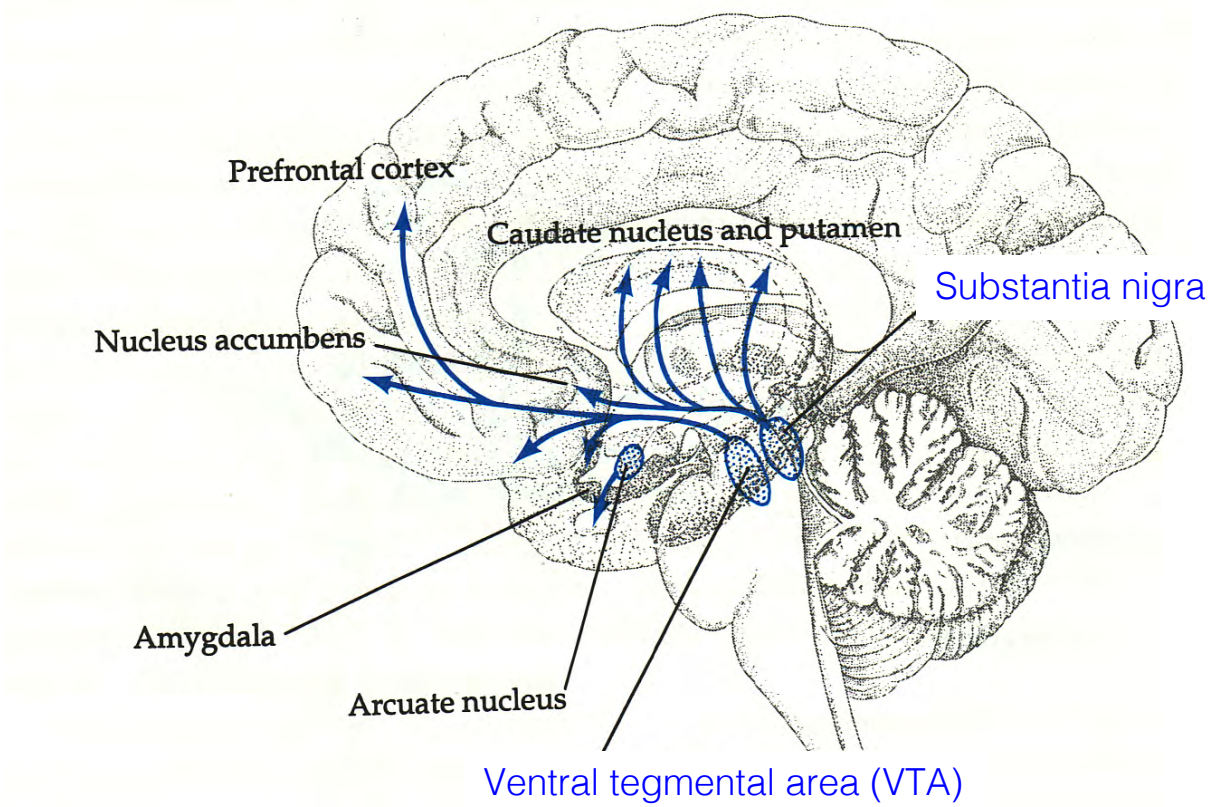


(Uchida and Mainen, 2003)

- Well-controlled behavior
- Electrophysiology
- Molecular/genetic tools

Dopamine





A10: VTA
A9: SNc
A8: retrorubral field

(From Neuron to Brain)

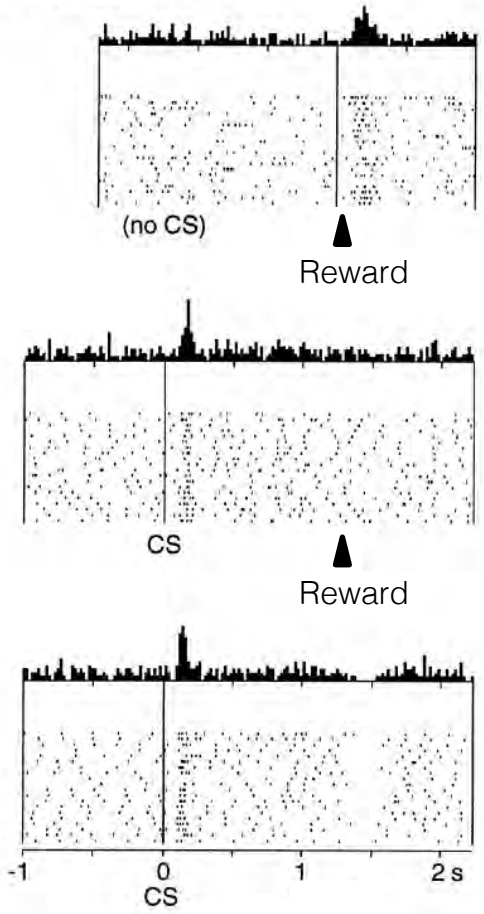
Firing of putative dopamine neurons

No prediction
Reward occurs

CS predicts reward
Reward occurs

(CS: conditioned stimulus)

CS predicts reward
No reward occurs

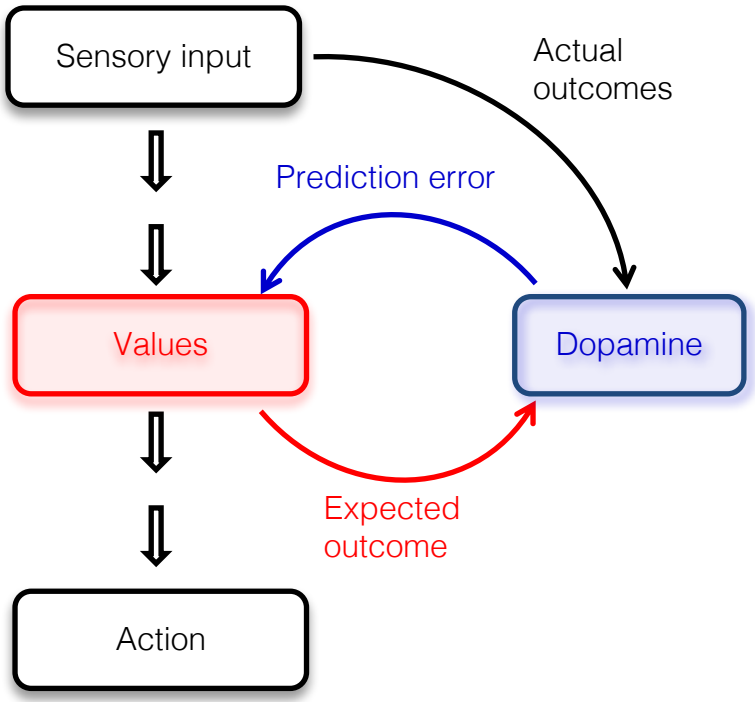


$$\delta(t) = r(t) + \gamma \hat{V}(t+1) - \hat{V}(t)$$

(Schultz, Dayan, Montague, 1997)

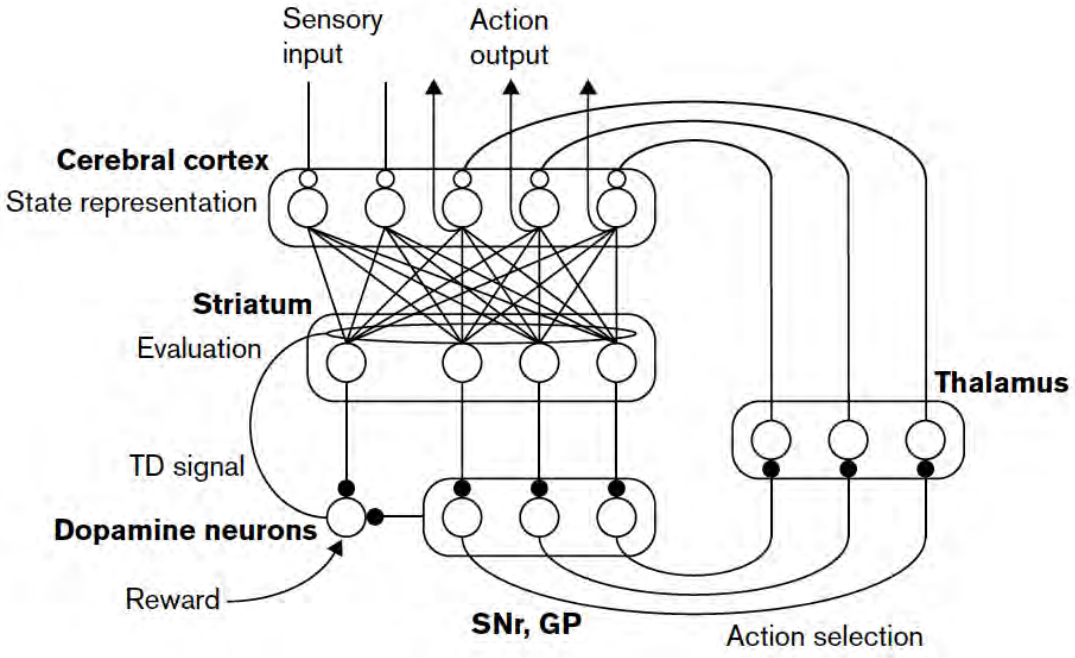
Reinforcement learning models

(1)



(After Sugrue et al.,2005)

(2)



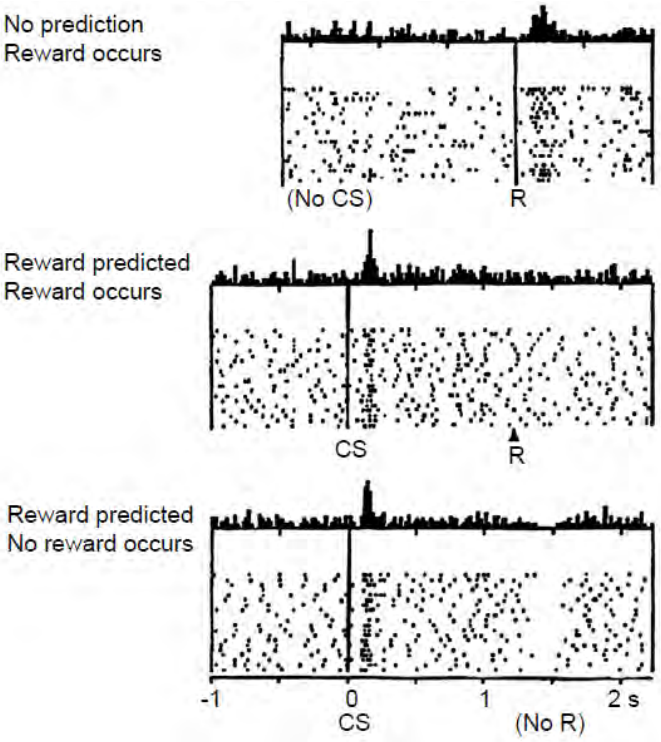
(Doya, 2000)

SCIENCE • VOL. 275 • 14 MARCH 1997

A Neural Substrate of Prediction and Reward

Wolfram Schultz, Peter Dayan, P. Read Montague*

- Neurobiology
  Phasic dopamine

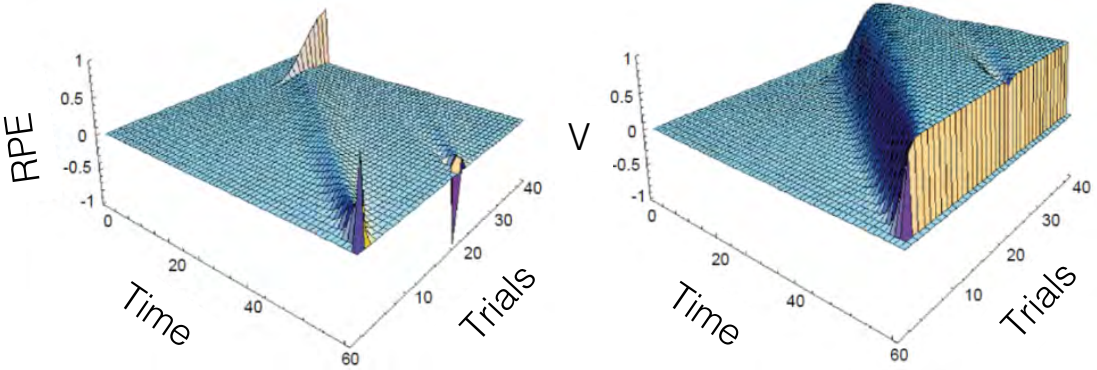


- Reinforcement learning
  Temporal difference (TD) learning theory

$$V(t) = E[\gamma^0 r(t) + \gamma^1 r(t+1) + \gamma^2 r(t+2) + \cdots] \quad (1)$$

$$V(t) = E[r(t) + \gamma V(t+1)] \quad (2)$$

$$\delta(t) = r(t) + \gamma \hat{V}(t+1) - \hat{V}(t) \quad (3)$$

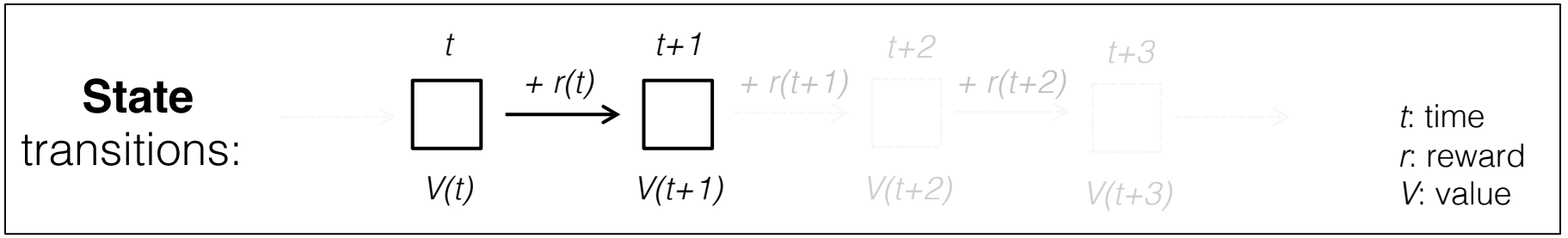


- Animal learning theory
  Learning is driven by prediction errors

Kamin, Rescorla, Wagner

Dopamine as temporal difference (TD) error

Dopamine as temporal difference (TD) error



- **Value function**: the sum of all future reward

$$V(t) = r(t) + \underbrace{r(t+1) + r(t+2) + \cdots}_{V(t+1)}$$

$$\hat{V}(t) = r(t) + \hat{V}(t+1)$$

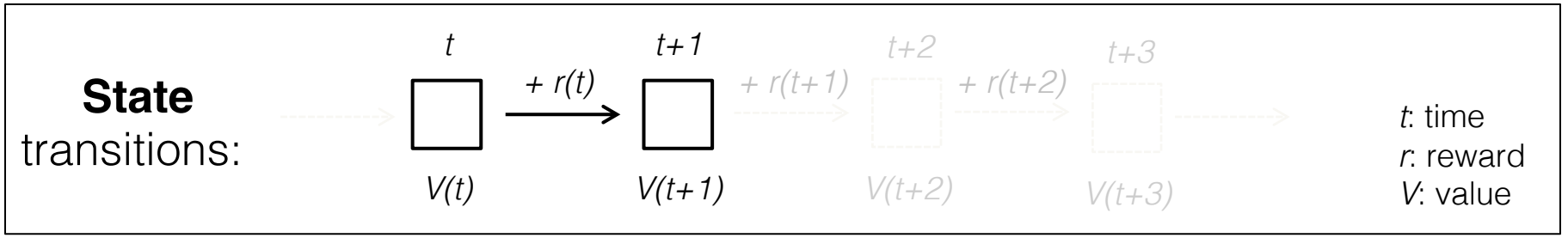
- **Temporal difference (TD) error**

$$\delta = r(t) + \hat{V}(t+1) - \hat{V}(t) \quad \Rightarrow \quad \text{Update } V(t)$$

$$\hat{V}(t) \leftarrow \hat{V}(t) + \alpha \cdot \delta \quad (\alpha\text{: learning rate})$$

(Sutton, 1988; Sutton & Barto, 1998; Montague et al., 1996; Schultz et al., 1997)

Dopamine as temporal difference (TD) error



- **Value function**: the sum of all future reward

$$V(t) = r(t) + \underbrace{\gamma \cdot r(t+1) + \gamma^2 \cdot r(t+2) + \cdots}_{\gamma \cdot V(t+1)}$$

$$\hat{V}(t) = r(t) + \gamma \cdot \hat{V}(t+1)$$

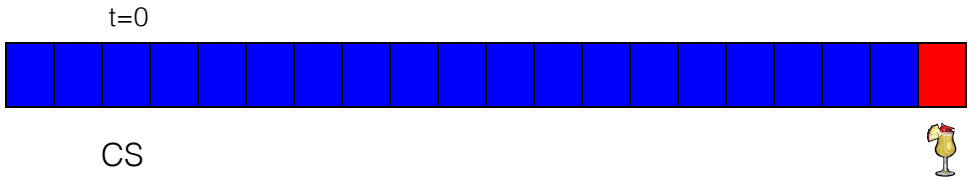
- **Temporal difference (TD) error**

$$\delta = r(t) + \gamma \cdot \hat{V}(t+1) - \hat{V}(t) \quad \Rightarrow \quad \text{Update } V(t)$$

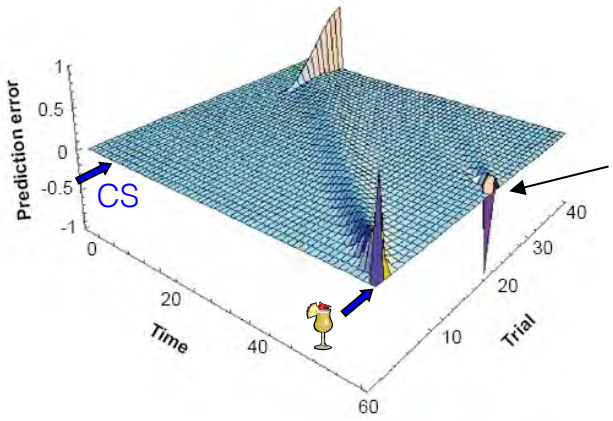
$$\hat{V}(t) \leftarrow \hat{V}(t) + \alpha \cdot \delta \quad (\alpha\text{: learning rate})$$

(Sutton, 1988; Sutton & Barto, 1998; Montague et al., 1996; Schultz et al., 1997)

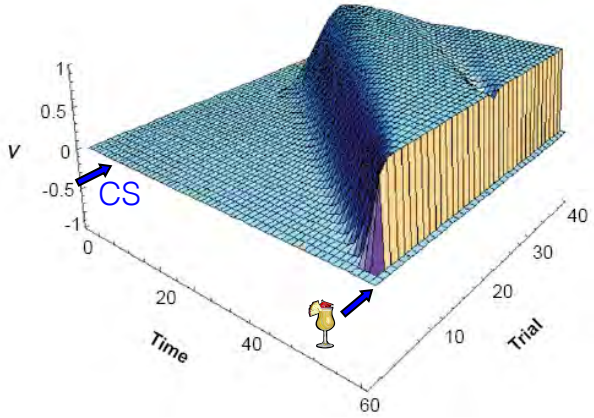
Dopamine as temporal difference (TD) error



TD-error signal



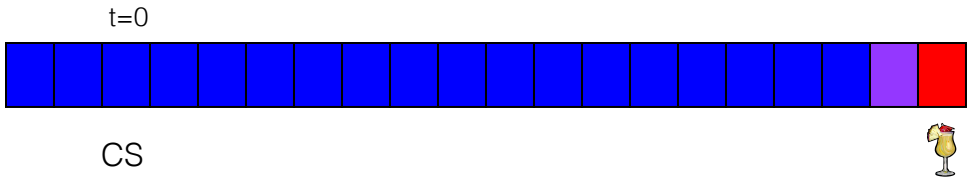
Value



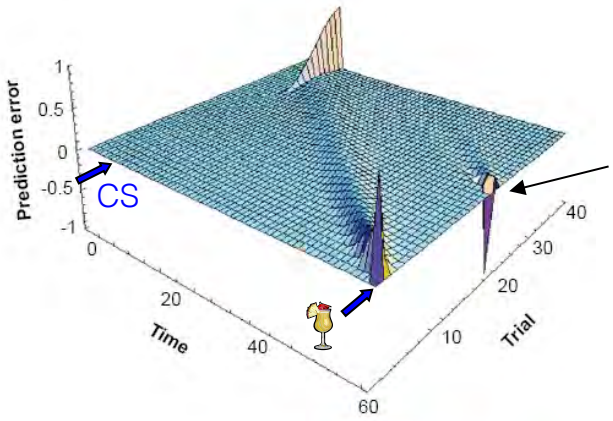
- Learning to predict future rewards

(Schultz, Dayan, Montague, 1997)

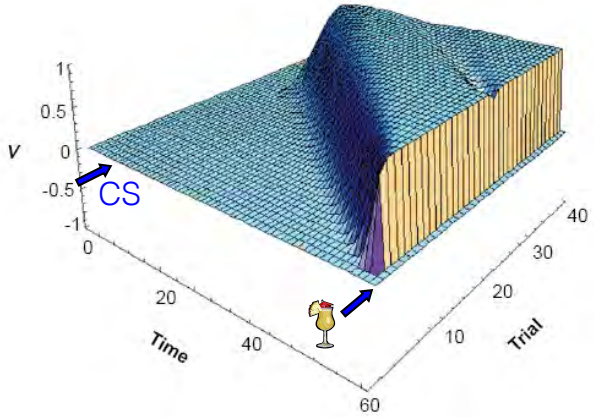
Dopamine as temporal difference (TD) error



TD-error signal



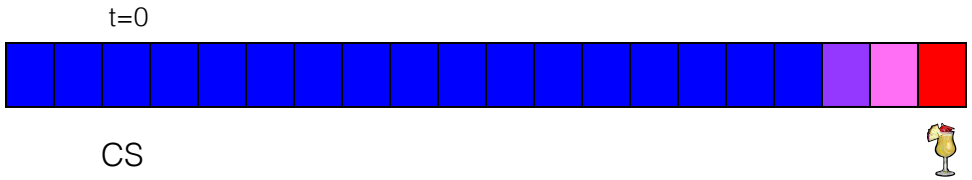
Value



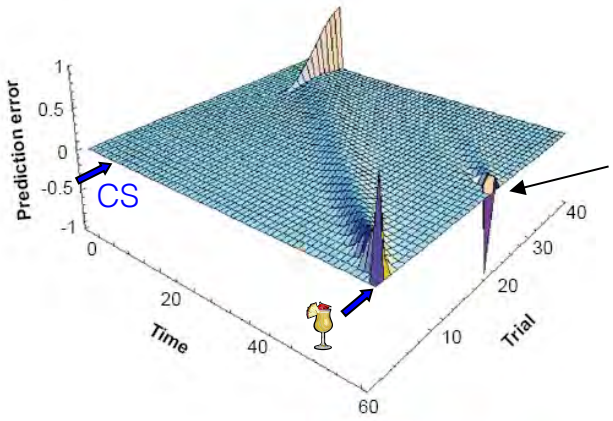
- Learning to predict future rewards

(Schultz, Dayan, Montague, 1997)

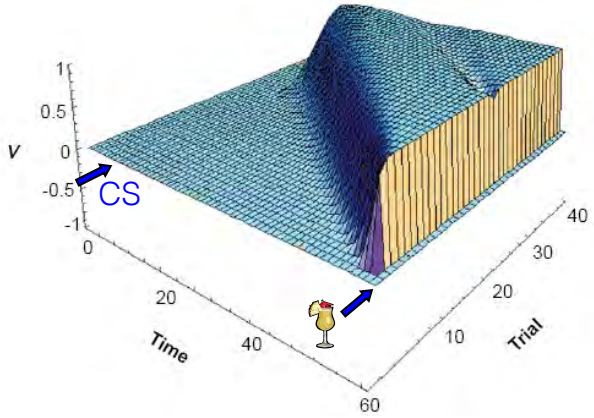
Dopamine as temporal difference (TD) error



TD-error signal



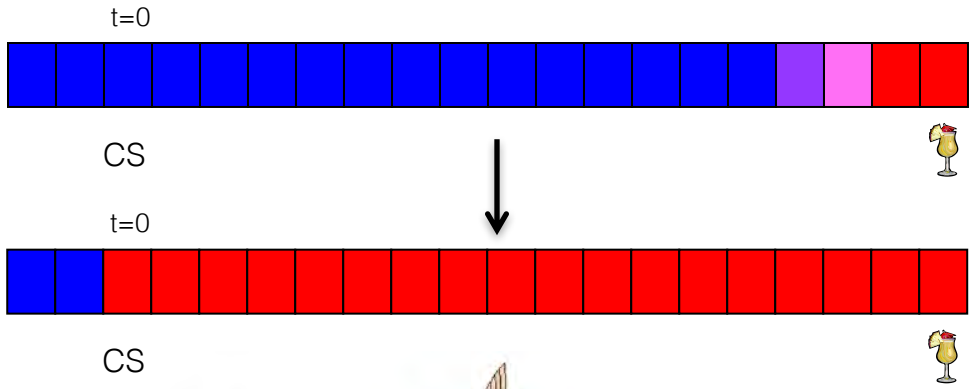
Value



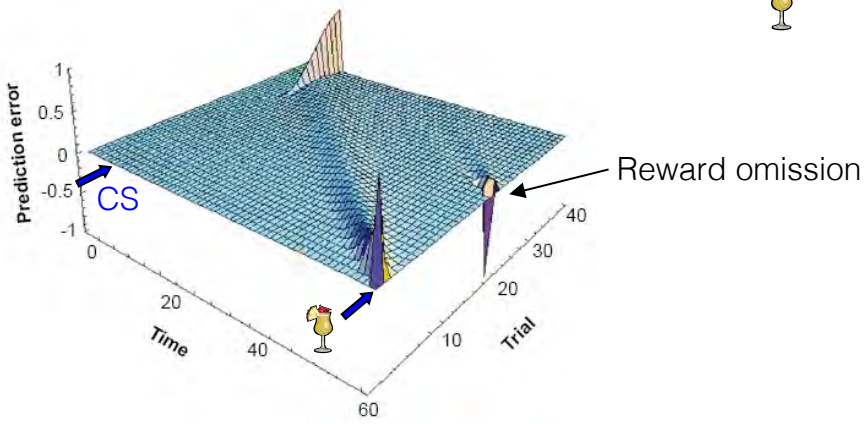
- Learning to predict future rewards

(Schultz, Dayan, Montague, 1997)

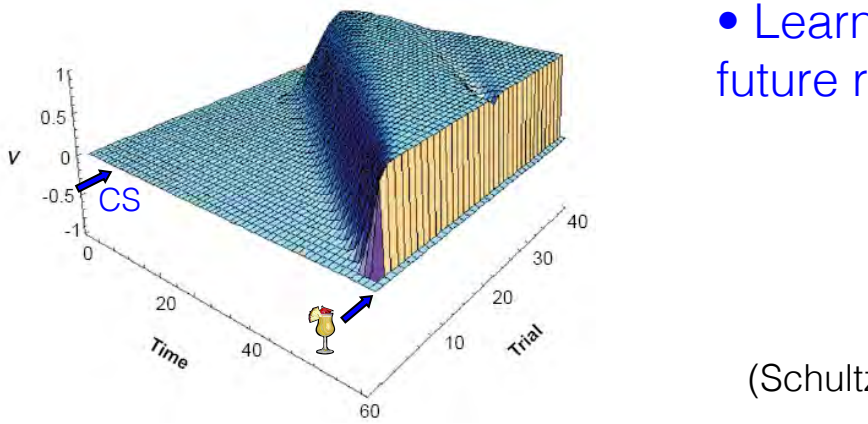
Dopamine as temporal difference (TD) error



TD-error signal



Value

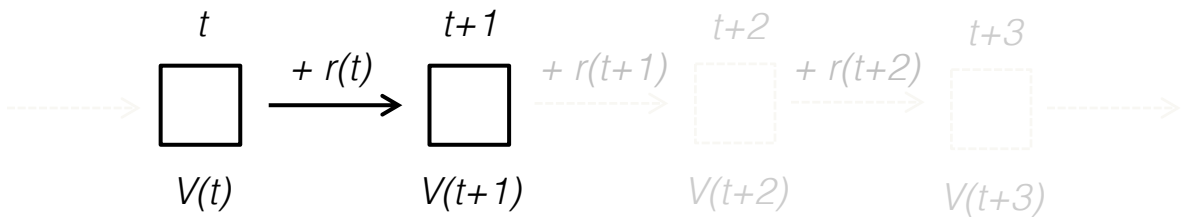


- Learning to predict future rewards

(Schultz, Dayan, Montague, 1997)

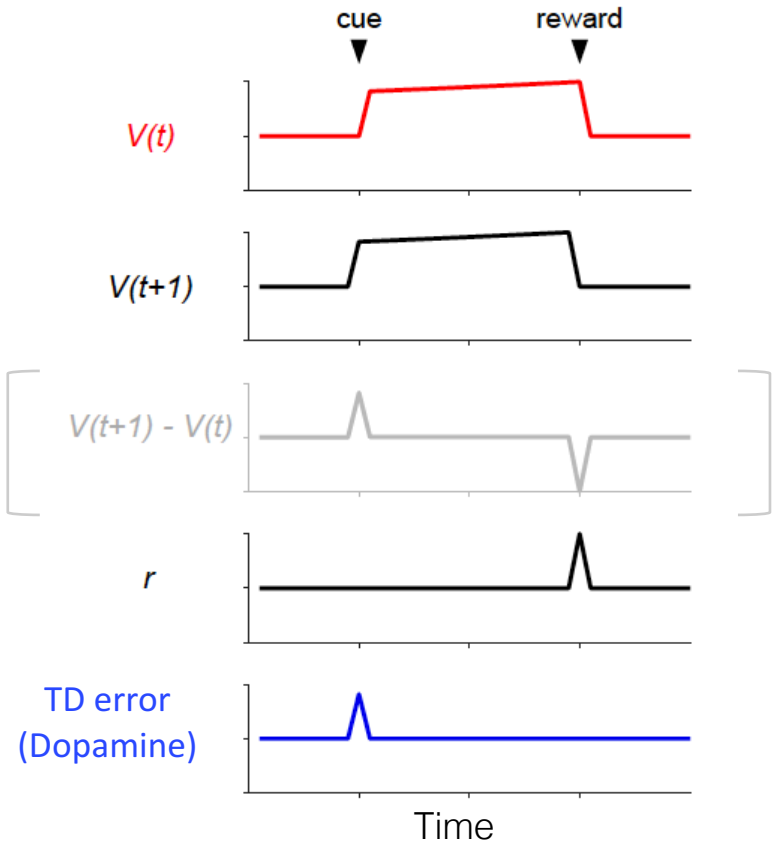
Dopamine as temporal difference (TD) error

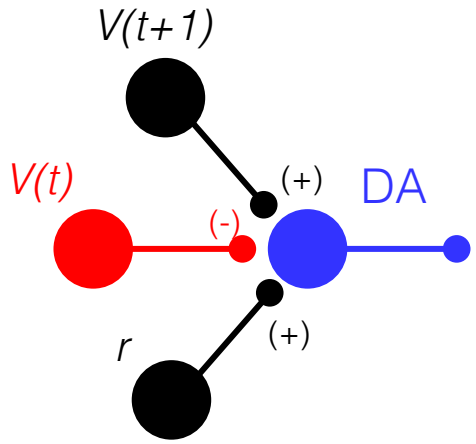
State transitions:



t: time
r: reward
V: value

derivative of *V(t)*

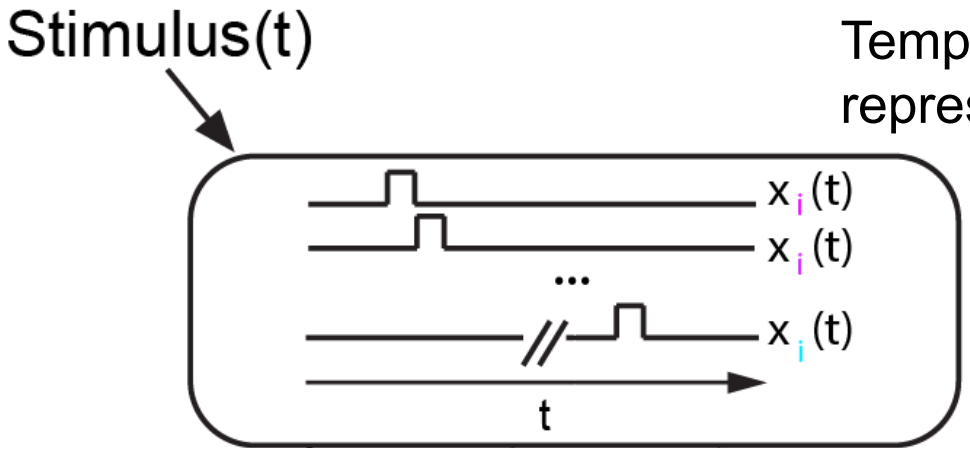




$$\delta = r(t) + \hat{V}(t+1) - \hat{V}(t)$$

(Sutton, 1988; Sutton & Barto, 1998; Montague et al., 1996; Schultz et al., 1997, Watabe-Uchida et al., 2017)

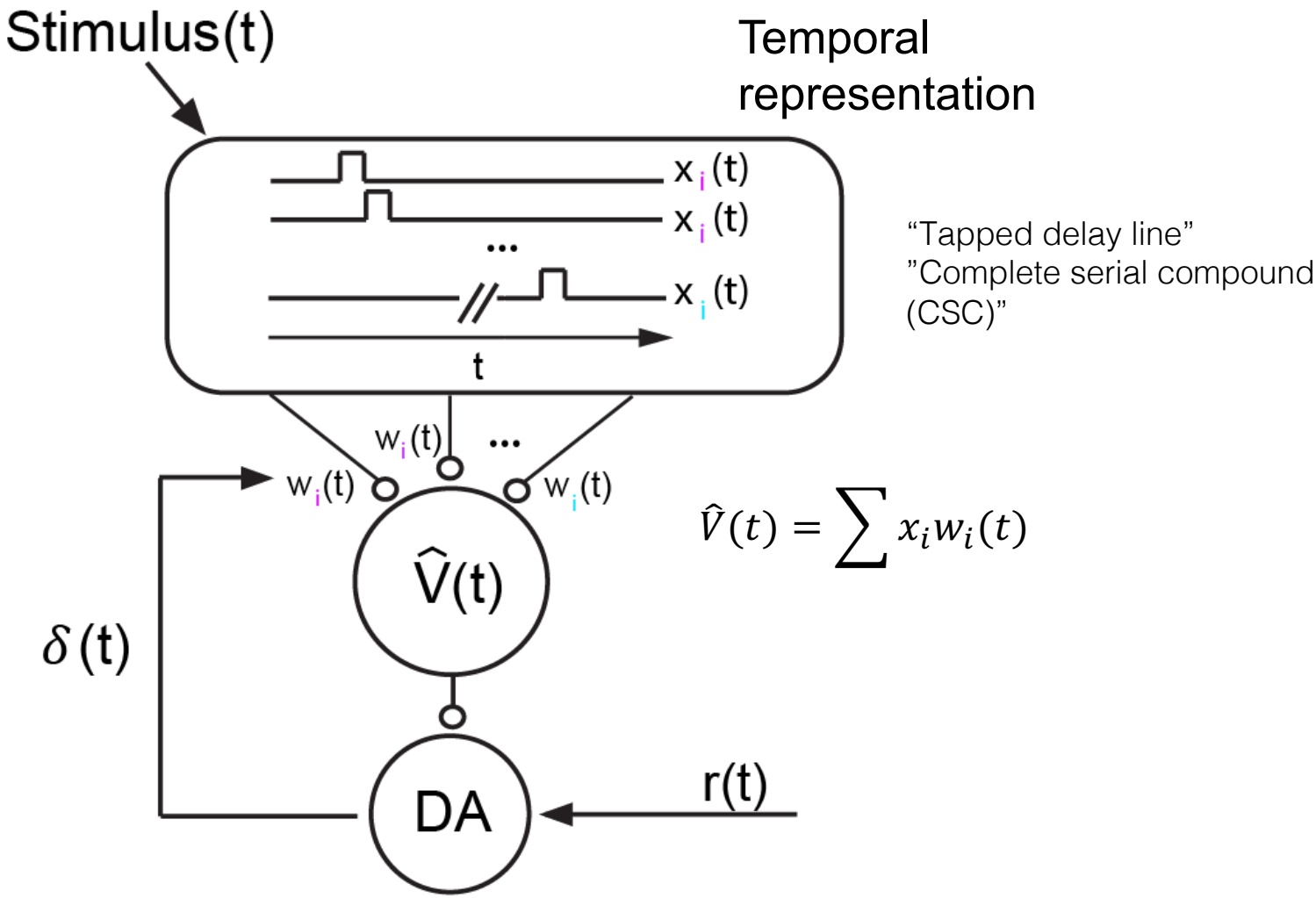
Dopamine as temporal difference (TD) error



"Tapped delay line"
"Complete serial compound (CSC)"

(Sutton, 1988; Sutton & Barto, 1998; Montague et al., 1996; Schultz et al., 1997)

Dopamine as temporal difference (TD) error



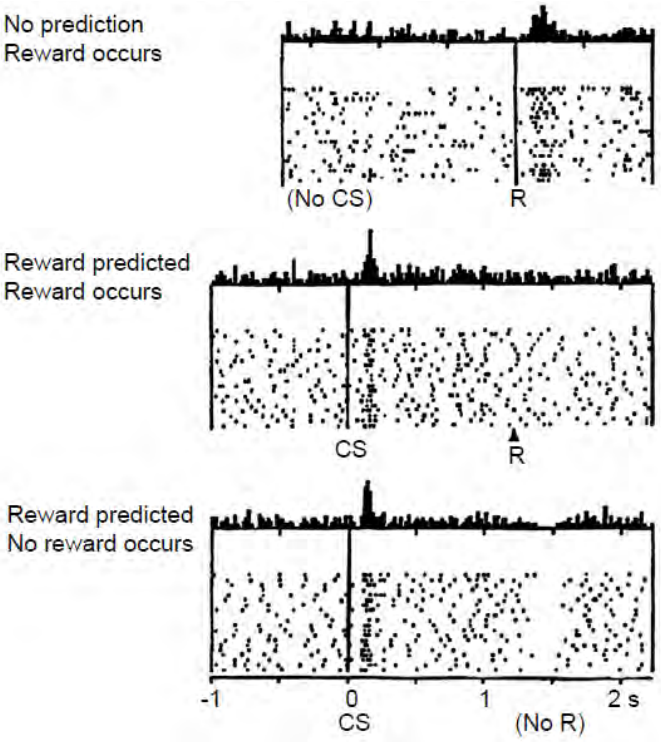
(Sutton, 1988; Sutton & Barto, 1998; Montague et al., 1996; Schultz et al., 1997)

SCIENCE • VOL. 275 • 14 MARCH 1997

A Neural Substrate of Prediction and Reward

Wolfram Schultz, Peter Dayan, P. Read Montague*

- Neurobiology
  Phasic dopamine

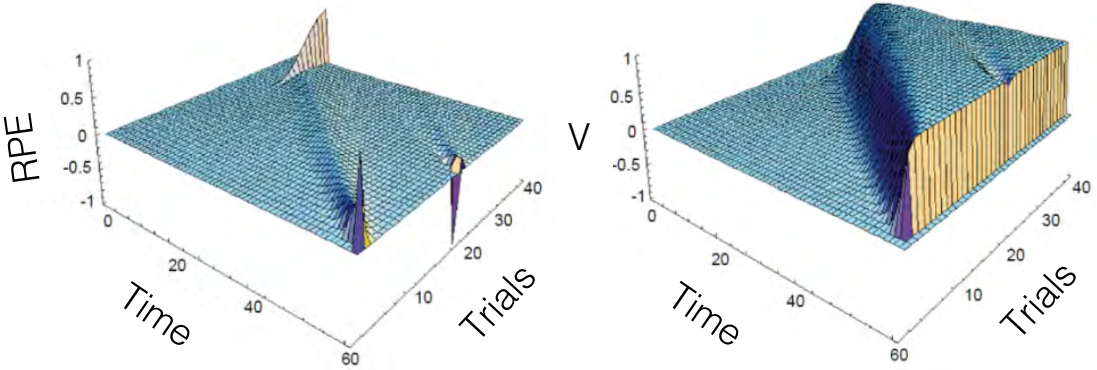


- Reinforcement learning
  Temporal difference (TD) learning theory

$$V(t) = E[\gamma^0 r(t) + \gamma^1 r(t+1) + \gamma^2 r(t+2) + \cdots] \quad (1)$$

$$V(t) = E[r(t) + \gamma V(t+1)] \quad (2)$$

$$\delta(t) = r(t) + \gamma \hat{V}(t+1) - \hat{V}(t) \quad (3)$$



- Animal learning theory
  Learning is driven by prediction errors

Kamin, Rescorla, Wagner

Dopamine as a TD error

- Magnitude
- Probability
- Temporal discounting
- Temporal uncertainty
- Long term value
- Cost / efforts (?)
- Safety
- Axiomatic proof

(Schultz, Glimcher, Fiorillo, Hikosaka, Phillips, Roitman, Cheer, Uchida, …)

Dopamine as a TD error

- Supporting evidence
- Minor problems
- Serious problems

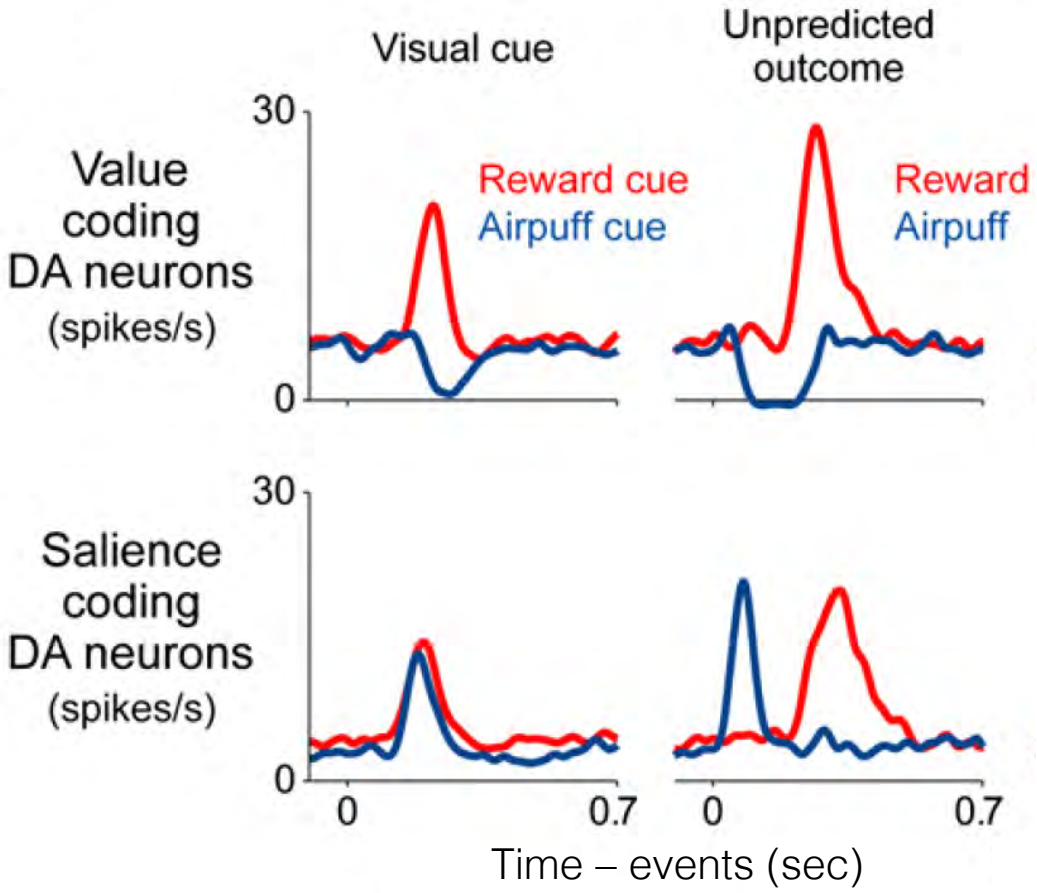
Topics

- A mouse model for studying dopamine RPE
- Do all dopamine neurons signal RPEs?
- What is the “state” in reinforcement learning?
- How are RPEs computed?
- Diversity of dopamine neurons

Topics

- A mouse model for studying dopamine RPE
- Do all dopamine neurons signal RPEs?
- What is the “state” in reinforcement learning?
- How are RPEs computed?
- Diversity of dopamine neurons

The problem of punishment

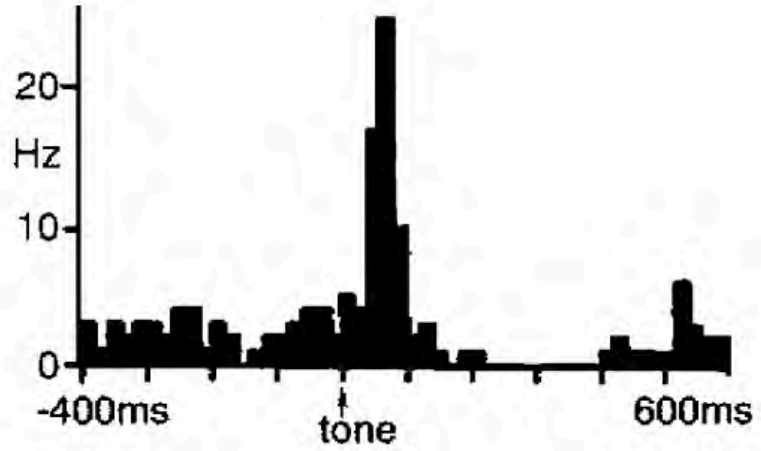


- Medial VTA

- Lateral VTA
- SNc

(Matsumoto & Hikosaka, 2008; Bromberg-Martin, Matsumoto, Hikosaka, 2010)

Novel stimuli activate some dopamine neurons



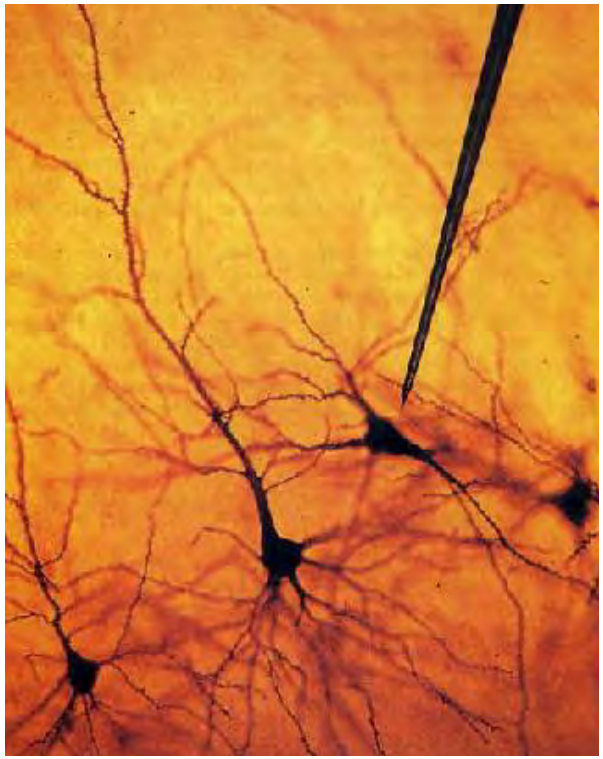
(adapted from
Kakade and Dayan, 2002)

Fig. 6. Novelty response with phasic activation and depression. This shows a histogram of the activity of a single dopamine cell in cat VTA in response to repetitions of an initially novel tone. This neuron shows a clear pattern of activation and depression in response to the stimulus. Adapted from Horvitz, Steward, and Jacobs (1997).

Attempts to incorporate novelty response into the value framework

- Potential reward or novelty itself is rewarding
- Positive value of exploration

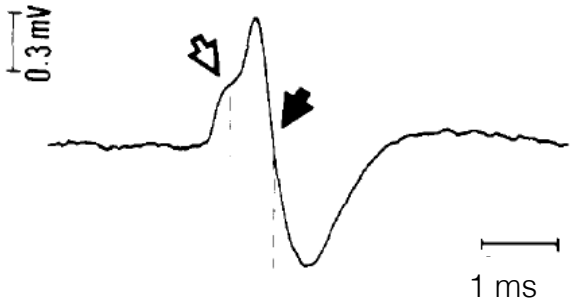
“The cell identification problem”



Extracellular recording

“The cell identification problem”

- Waveforms, spontaneous firing rates, pharmacology

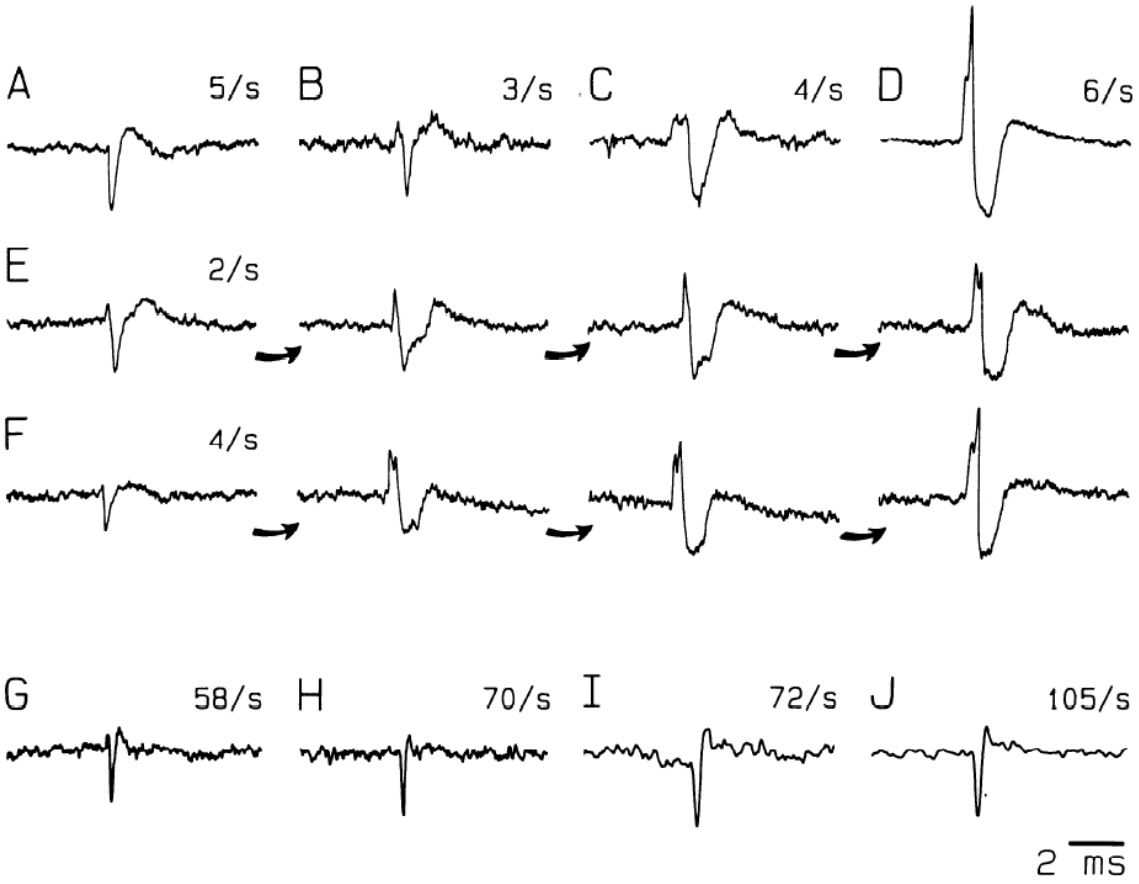


(Grace and Bunney, 1983; also see Ungless and Grace, 2012)

- Recent studies indicate diverse waveforms in dopamine neurons (e.g. Margolis et al., 2006; Lammel et al. 2008)

“The cell identification problem”

- Dopamine

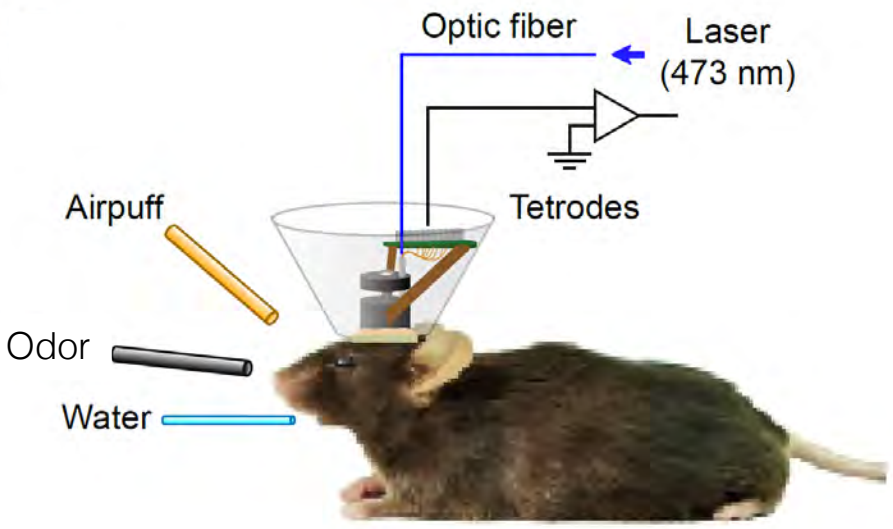


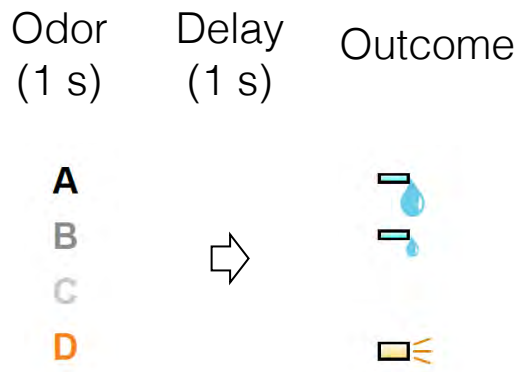
- Non-dopamine

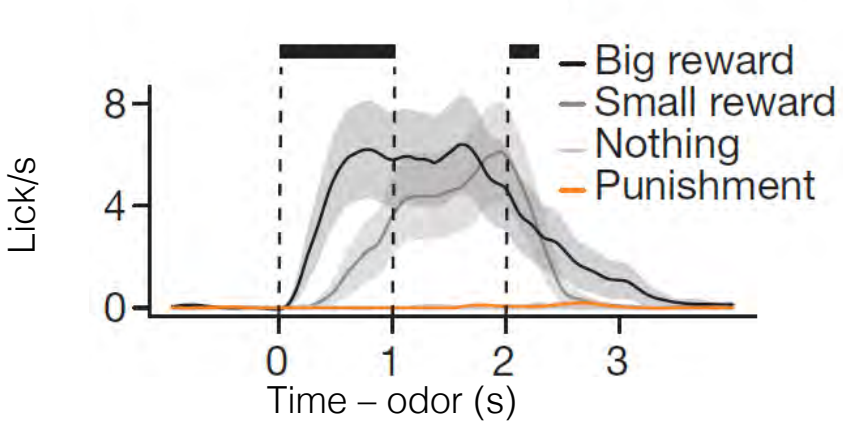
- Spike waveforms differ not only by neuron types but also relative locations to electrode

(Schultz, 1986)

Odor-value association task in mice





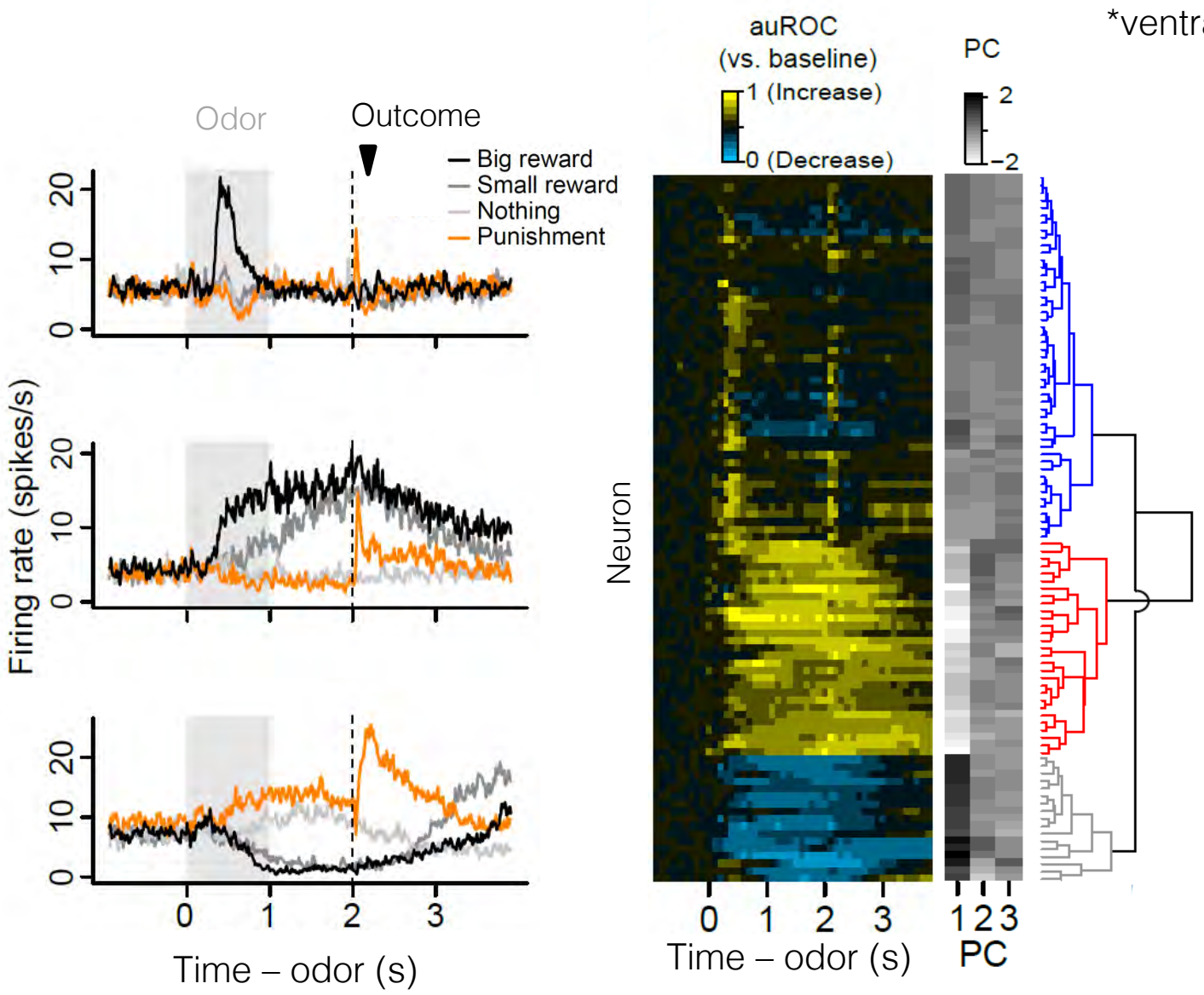


- Head-restrained mice
- Efficient training (1-2 days)
- Electrophysiology in behaving mice

(Cohen, Haesler et al., *Nature*, 2012)

Three response types in VTA*

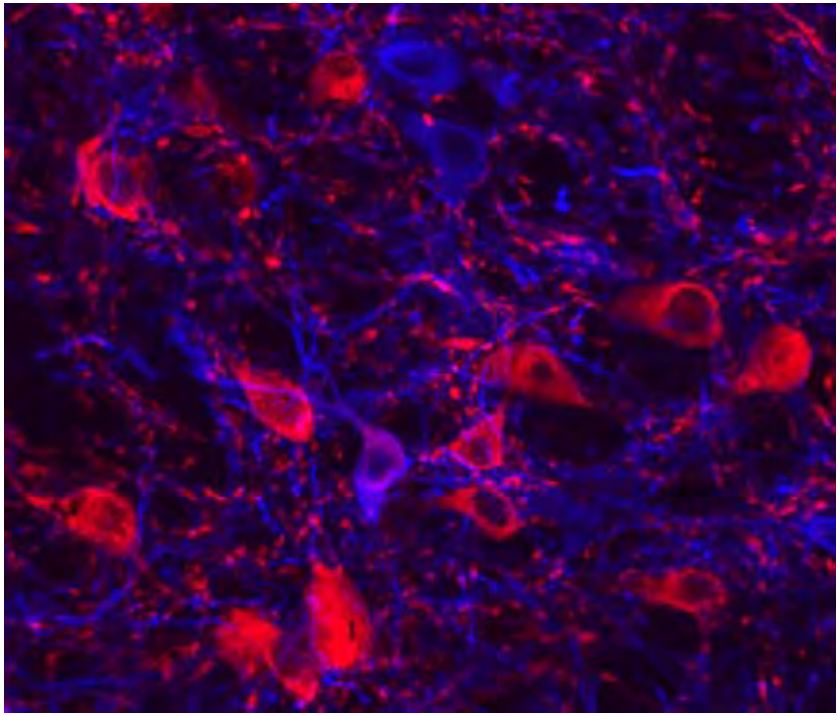
*ventral tegmental area



(big reward trials)

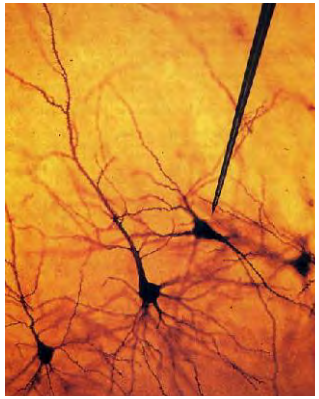
(Cohen, Haesler et al., *Nature*, 2012)

VTA contains multiple neuron types

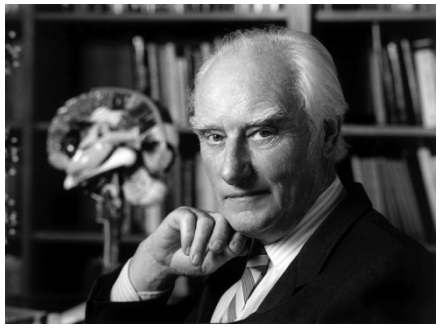


Dopamine neurons (55-65 %)

GABA neurons (30-40 %)



“The recording problem”



It is one of the peculiar features of most modern neurophysiology that the experimentalist…seldom knows which type of neuron he or she is listening to… It is common for experimentalist to record that, say, 25% of the neurons studied behave in a particular way, 37% in a different way and a further 15% in a third way… There is no indication where these different sets of neurons are sending their information, let alone exactly what type of neuron they are. This is not science but rather natural history. Rutherford would probably have called it stamp collecting.

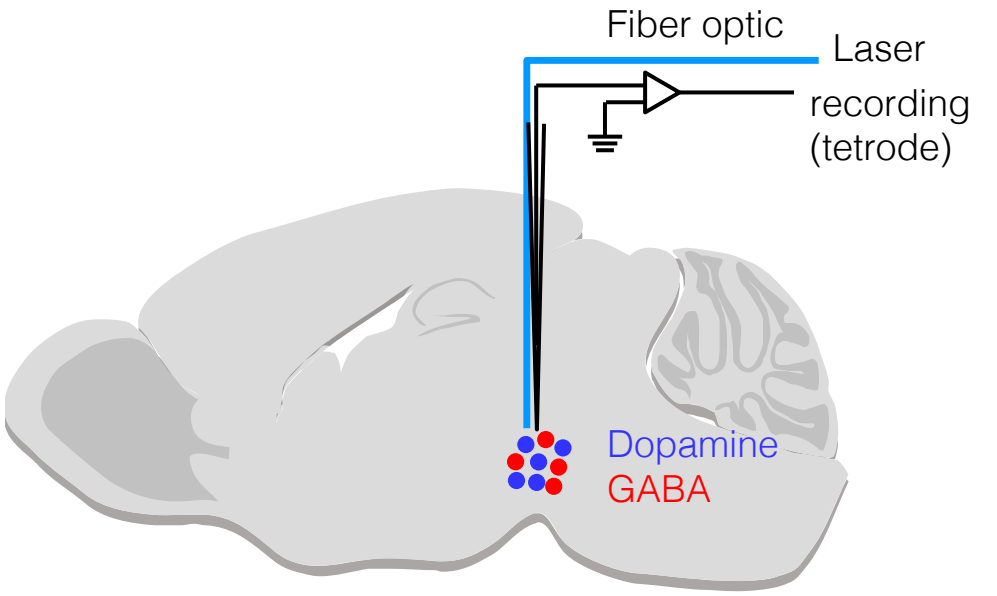
(Francis Crick, 1999, The impact of molecular biology on neuroscience)

Ernst Rutherford

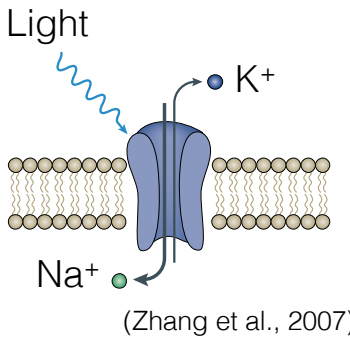
All science is either physics or stamp collecting

(As quoted in Rutherford at Manchester (1962) by J. B. Birks)

Optogenetic identification of neuron types



- DAT-Cre (dopamine)
- Vgat-Cre (GABA)
- AAV-FLEX-ChR2



(Zhang et al., 2007)

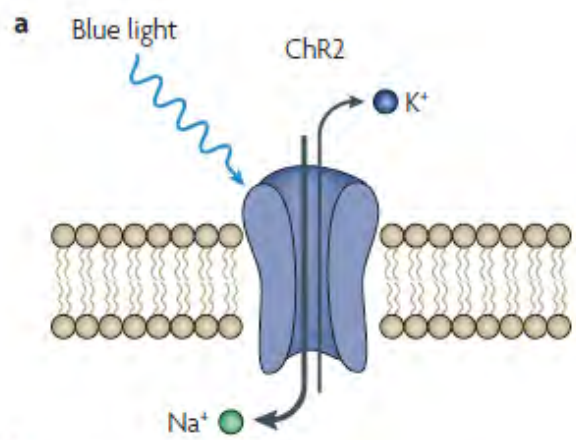
- Channelrhodopsin-2 (ChR2) in dopamine neurons

(Cohen, Haesler et al., *Nature*, 2012)

(Also see; Lima et al., PLoS One, 2009; Cardin et al., Nature, 2010)

Optogenetics

Channelrhodopsin-2

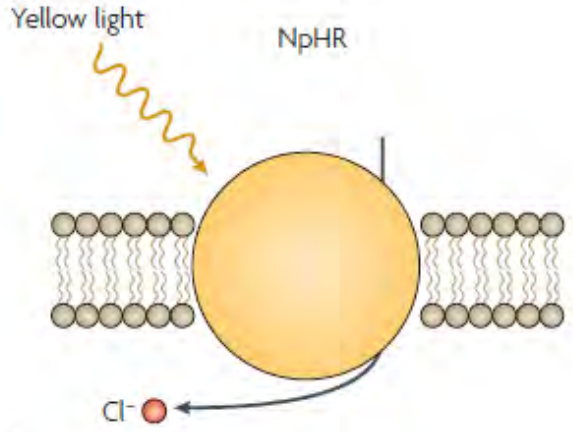


Light-gated ion-channel

Cation channel (H^+, Na^+, K^+, and Ca^{2+})

from green algae

Halorhodopsin



Light-activated Cl^- pump

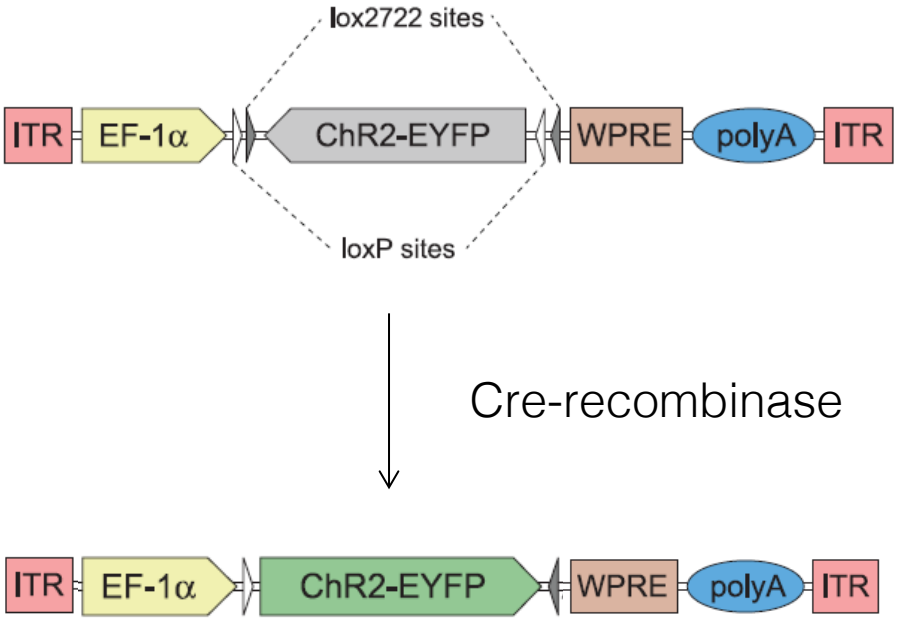
from halobacteria

Activation

Inactivation

Cre/loxP recombination system

- Transgenic mice: Cre-recombinase is expressed in a specific neural population.
- Inject adeno-associated virus carrying floxed-channerhodopsin-2 (ChR2)



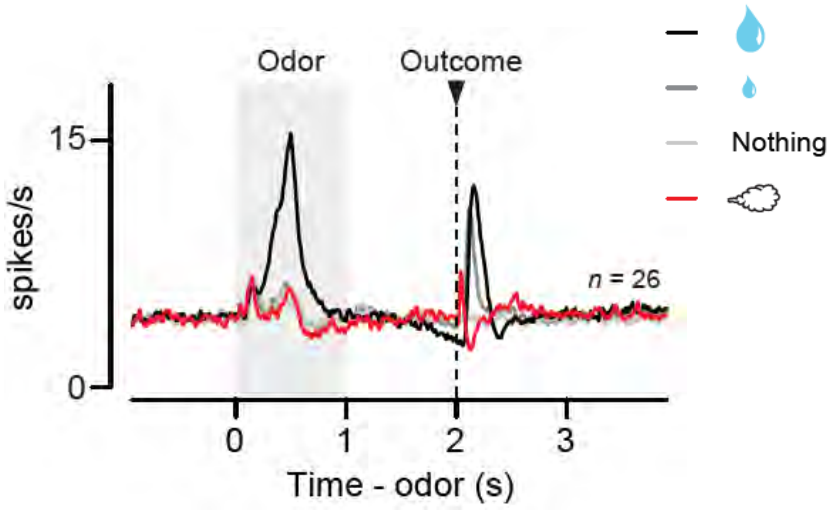
Optogenetic identification of neuron types

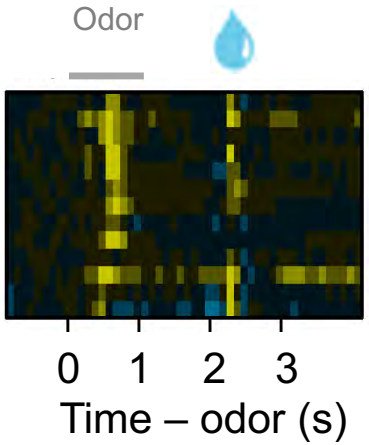
Criteria

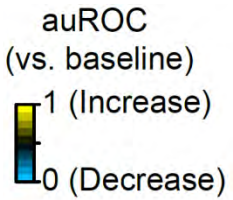
1) Latency < 2-4 ms
2) Follow high frequency stimulation
3) Spike shapes are almost identical between light-evoked and spontaneous spikes

Neuron-type-specific signals in VTA

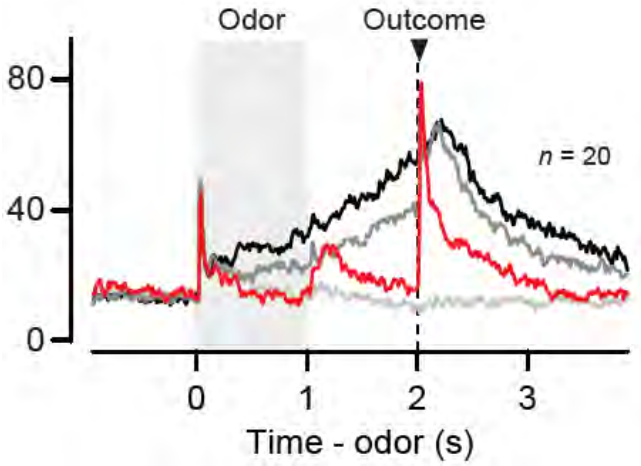
- Dopamine neurons

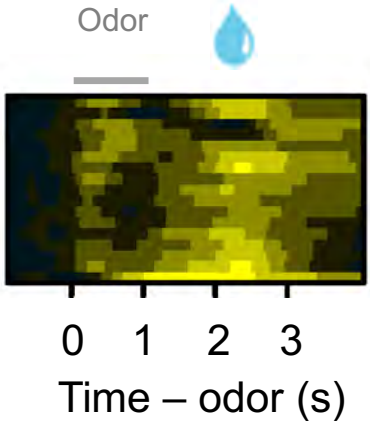






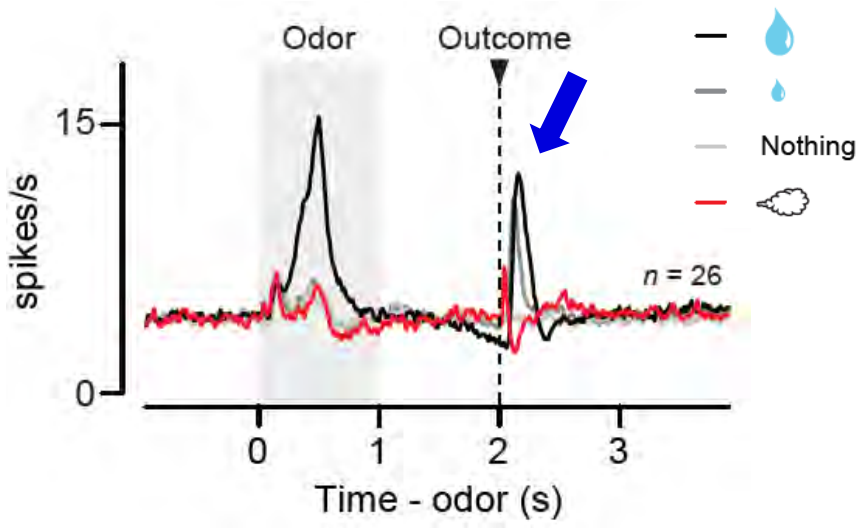
- GABA neurons

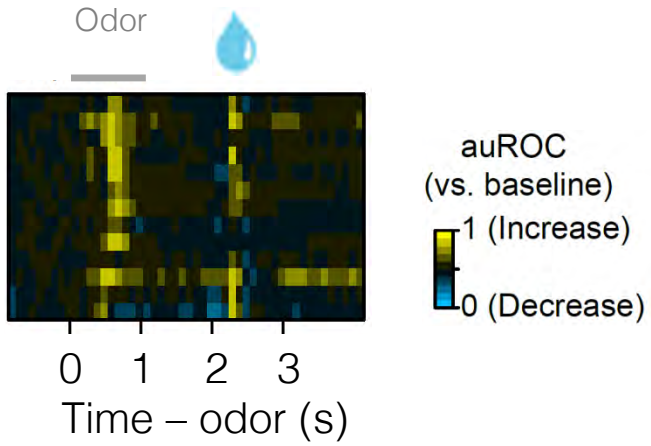




(Cohen, Haesler et al., *Nature*, 2012)

RPE?

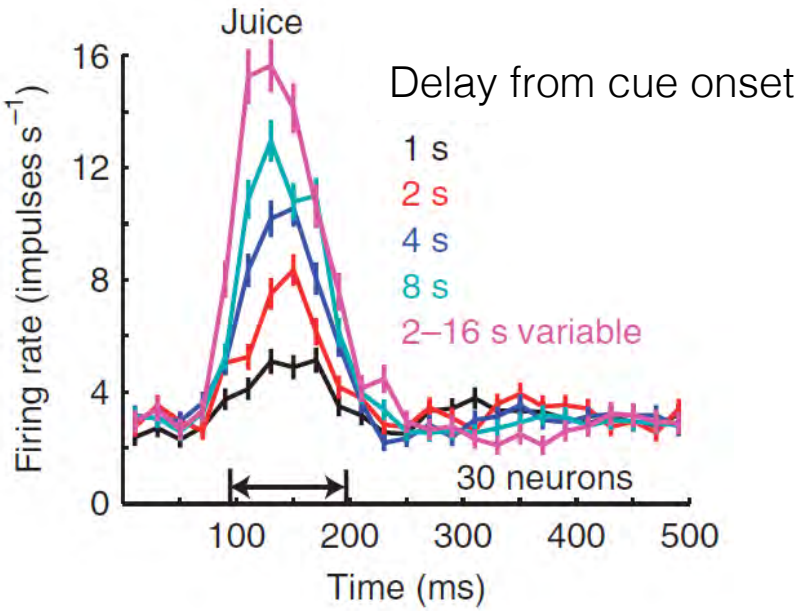




(Cohen, Haesler et al., *Nature*, 2012)

RPE?

- Effect of delays

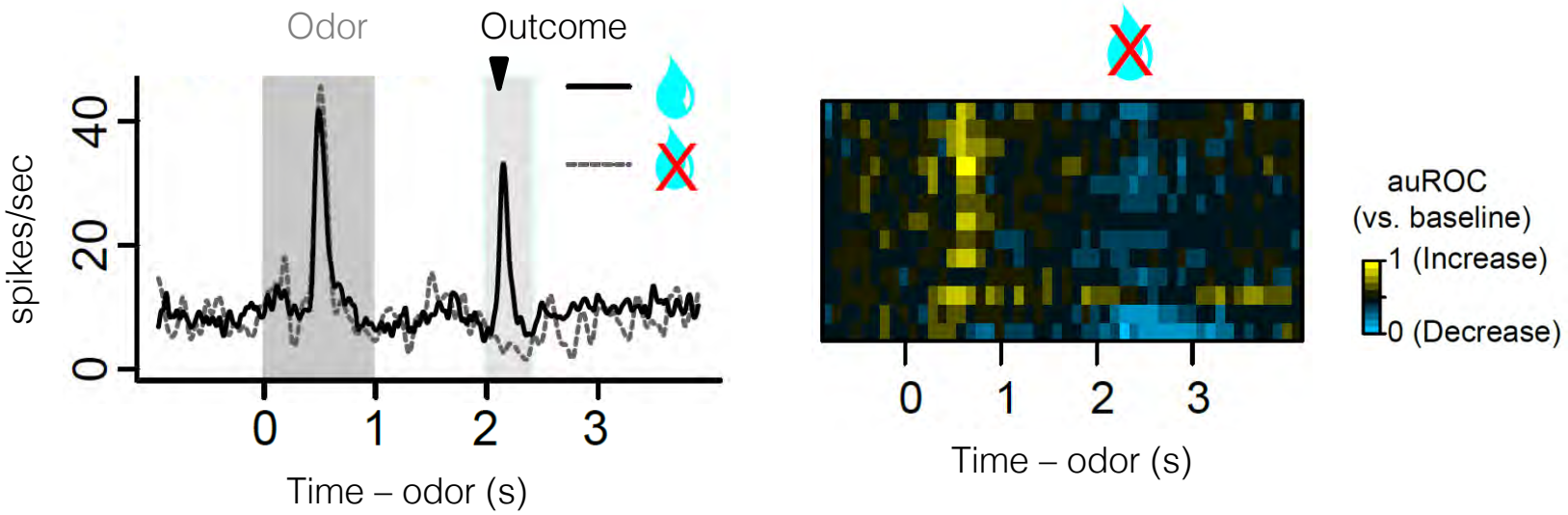


Delay → reduced reward predictability → increased dopamine RPEs

(Fiorillo, Newsome and Schultz, 2008; also see, Kobayashi et al., 2008)

RPE?

Reward omission (negative prediction error)

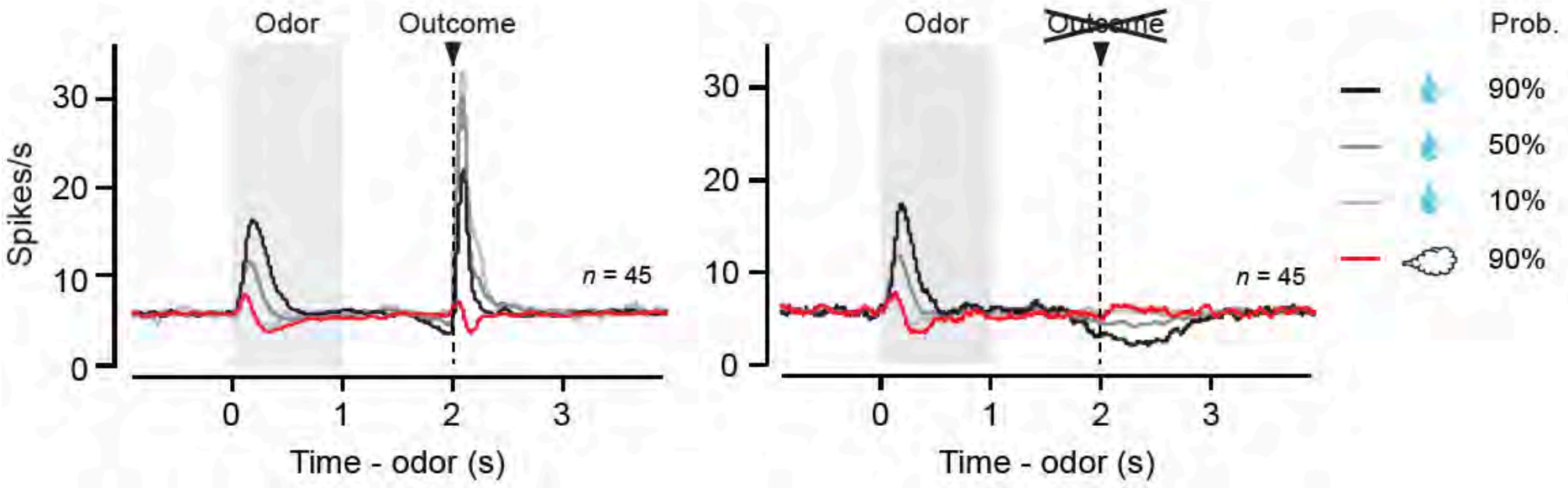


- Omission of reward causes suppression of firing.

(Cohen, Haesler et al., *Nature*, 2012)

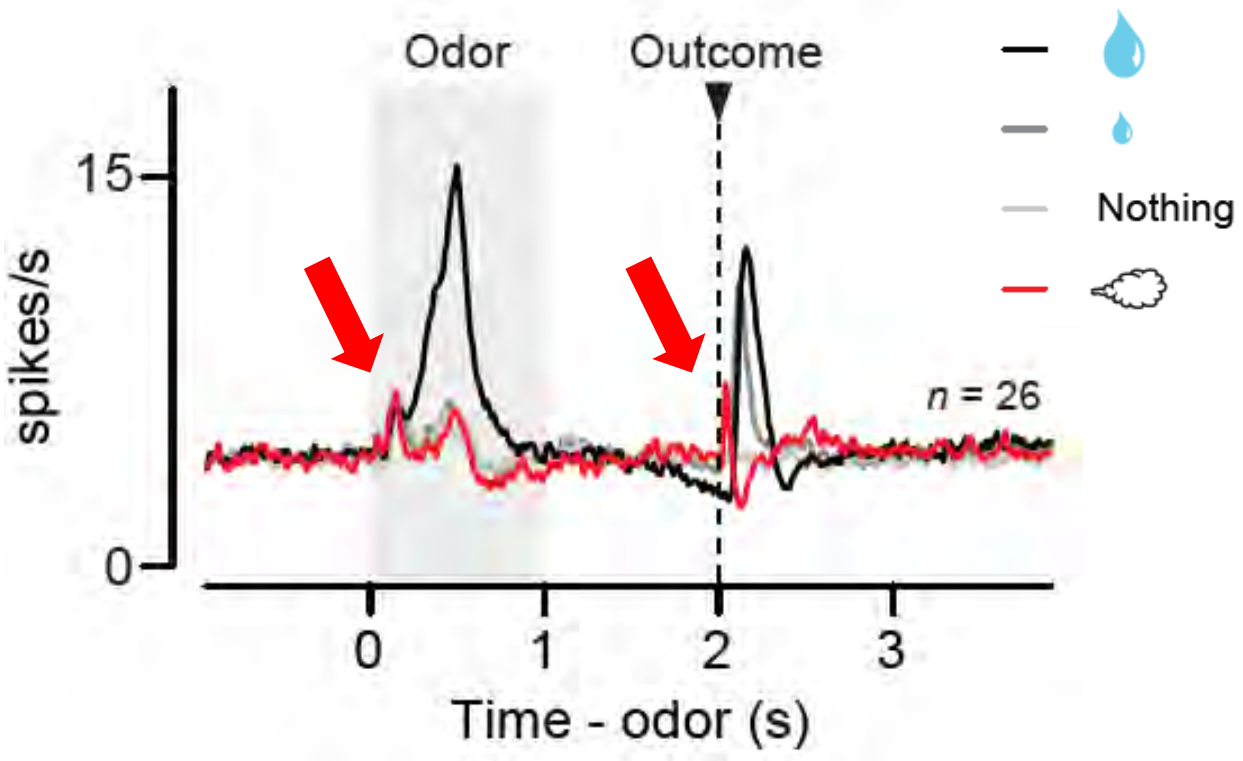
Recording from optogenetically-identified dopamine neurons

RPE coding by VTA dopamine neurons



(Tian and Uchida, *Neuron*, 2015)

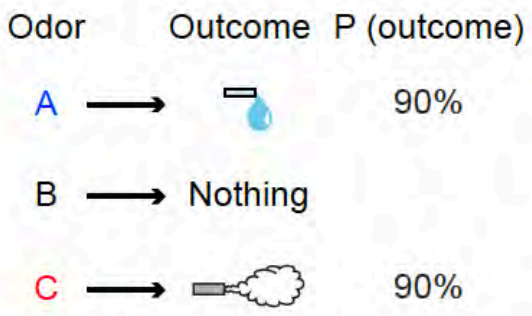
What about aversive stimuli?

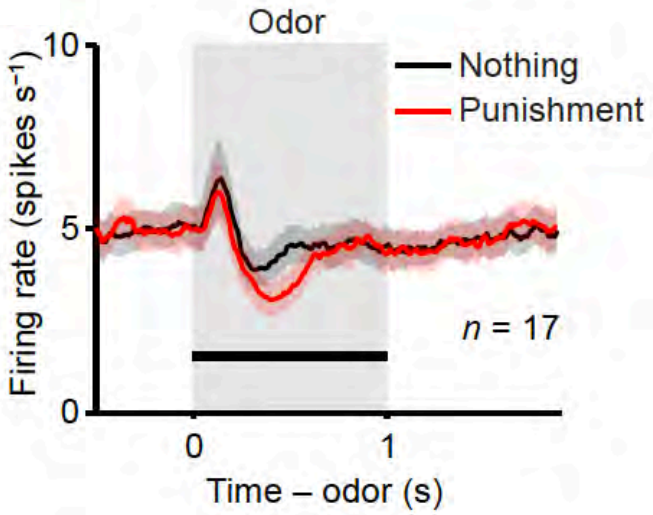


Biphasic dopamine response to air puff *or* air puff-predictive cues

(Cohen, Haesler et al., *Nature*, 2012)

Context-dependency of dopamine aversion response



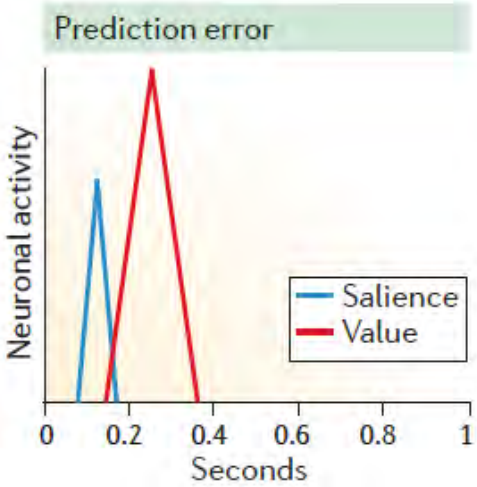


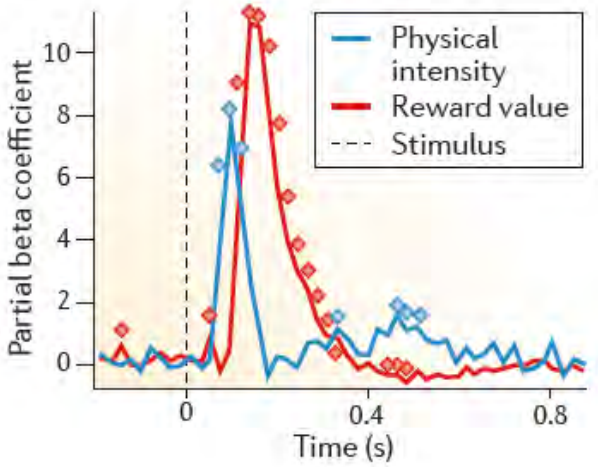
Nature Reviews Neuroscience, 2016

OPINION

Dopamine reward prediction-error signalling: a two-component response

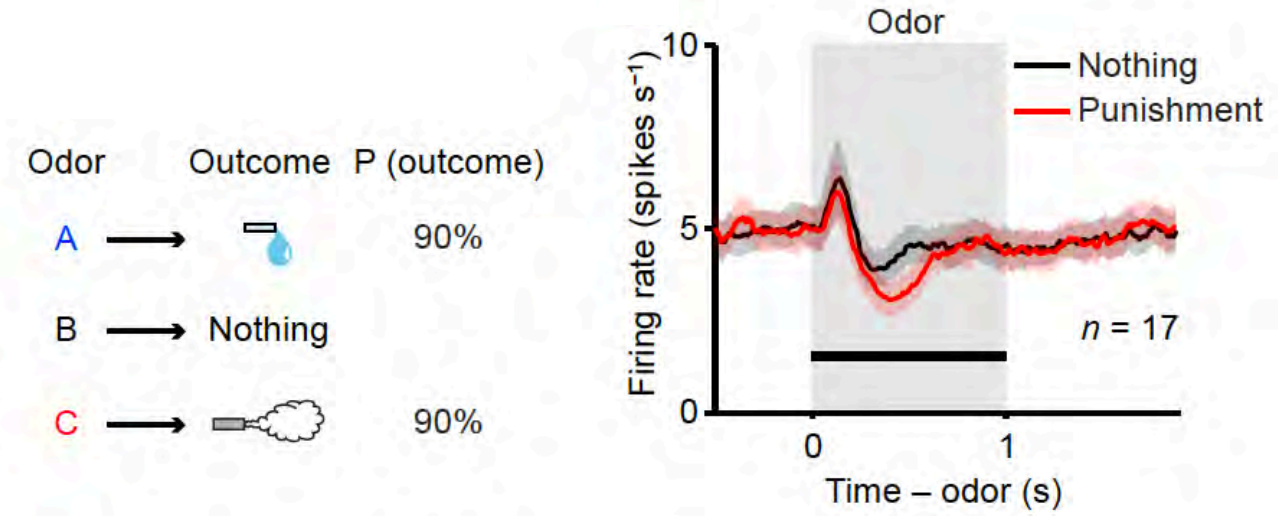
Wolfram Schultz





(Matsumoto, Tian, Uchida and Watabe-Uchida, *eLife*, 2016)

Context-dependency of dopamine aversion response



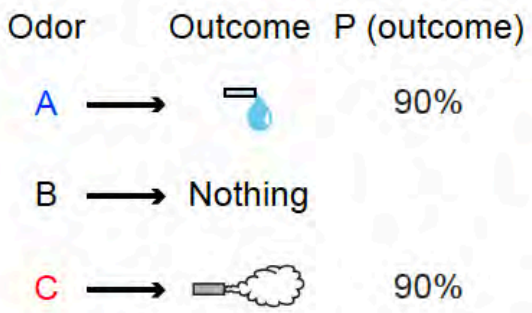
Kobayashi and Schultz, 2014

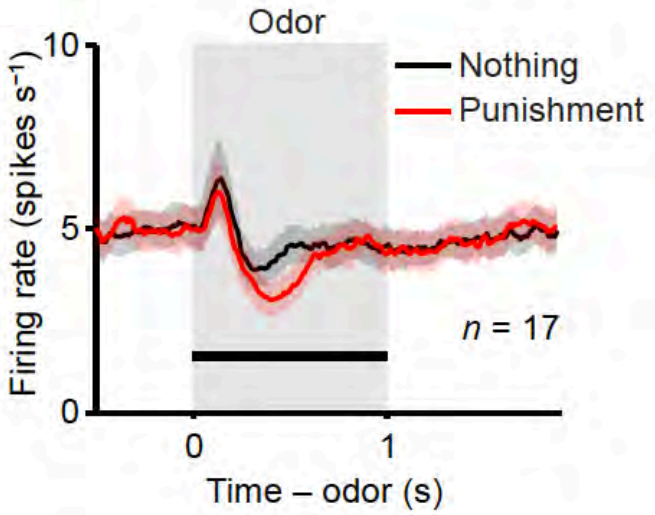
Reward contexts extends dopamine signals to unrewarded stimuli

(Matsumoto, Tian, Uchida and Watabe-Uchida, *eLife*, 2016)

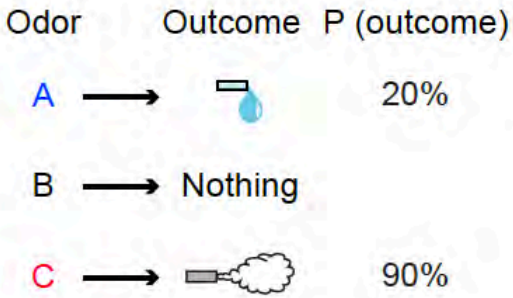
Context-dependency of dopamine aversion response

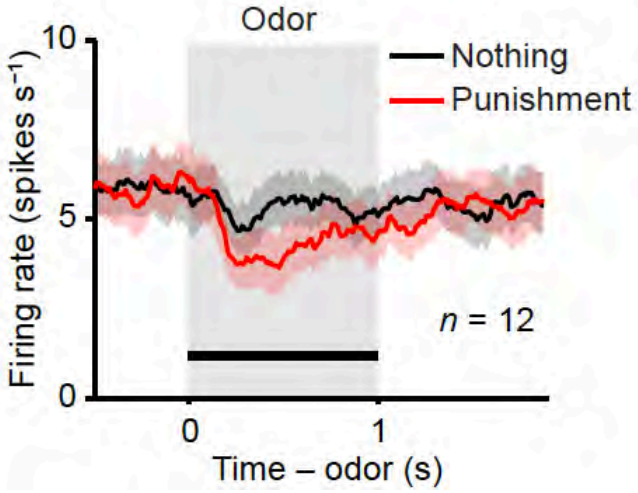
- High reward context





- Low reward context



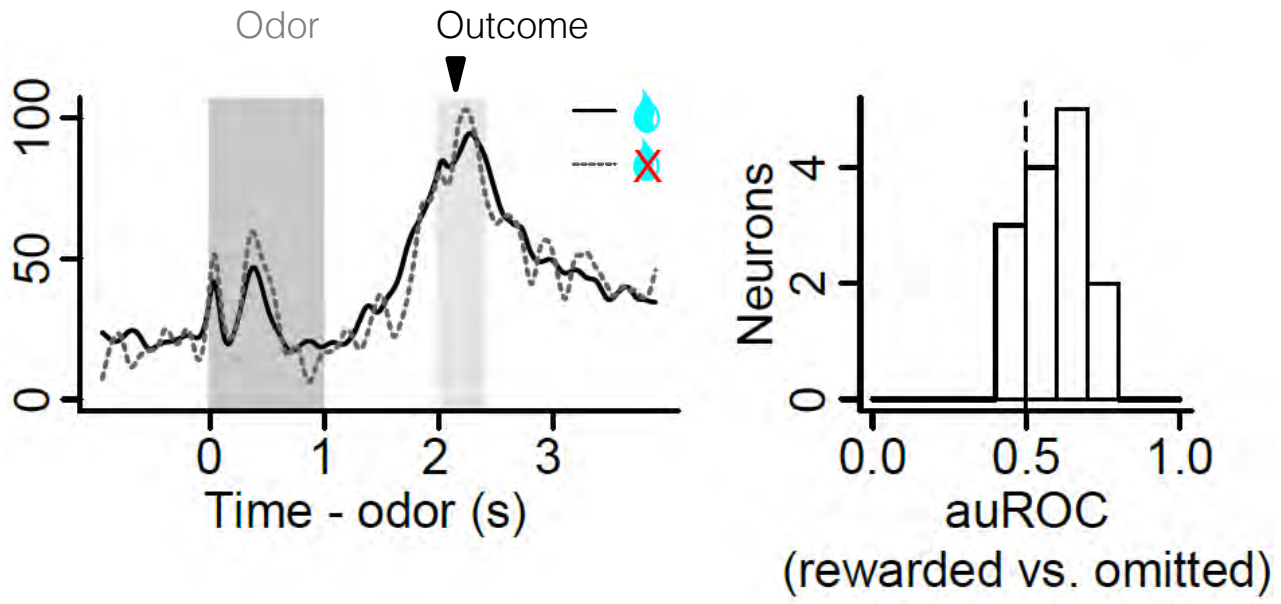


(Matsumoto, Tian, Uchida and Watabe-Uchida, *eLife*, 2016)

What do GABA neurons signal?

VTA GABA neurons signal reward expectation

- GABA neuron

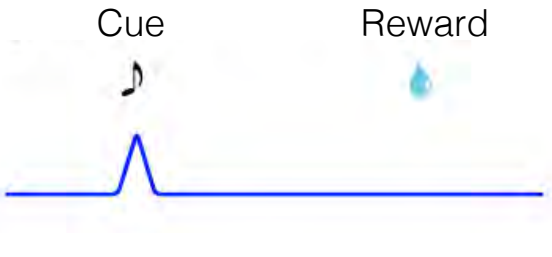


- Delivery of reward has little impact on firing.
- The activity reflects the expected time of reward.

(Cohen, Haesler et al., *Nature*, 2012)

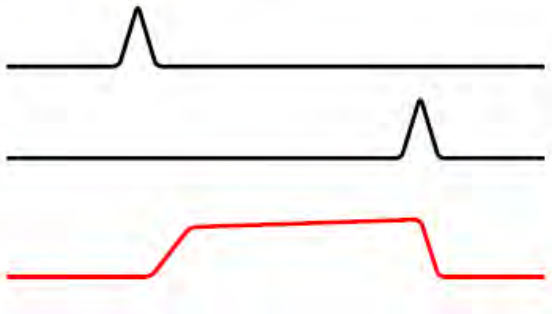
VTA GABA neurons may provide the reward expectation signal to dopamine neurons

- Output (dopamine neurons)



Reward prediction error ⬅ VTA Dopamine neurons

- Model (inputs to dopamine neurons)



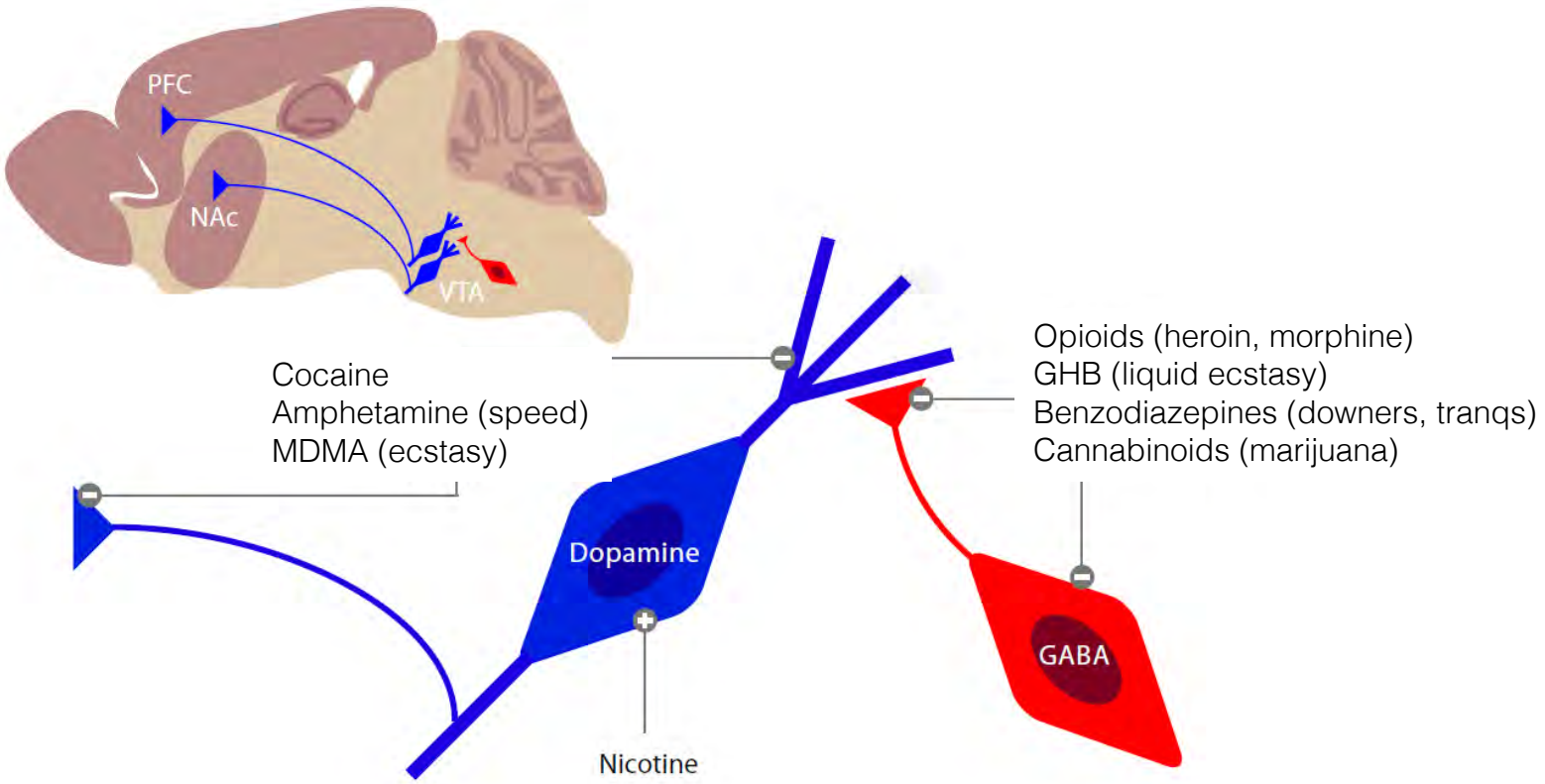
Cue

Reward

Reward expectation ⬅ VTA GABA neurons

(after Houk, Adams, Barto, 1995; Hazy, Frank, O'Reilly, 2010)

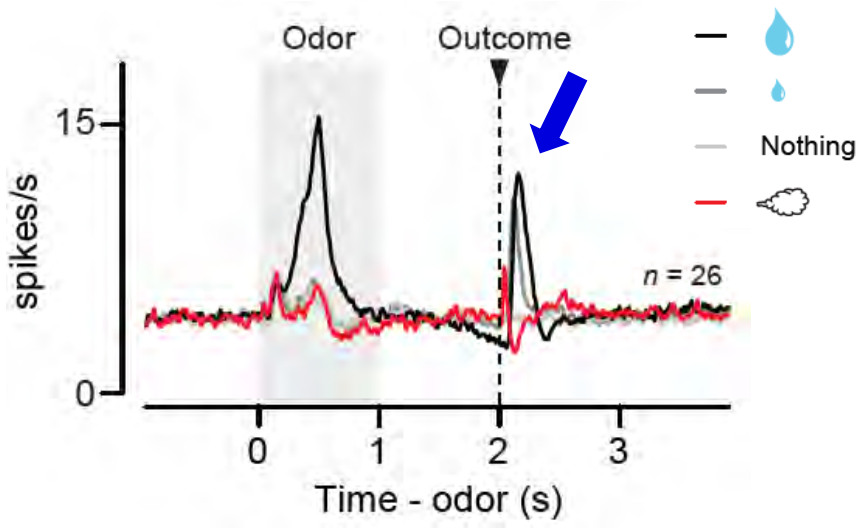
Addiction as impaired prediction error signaling

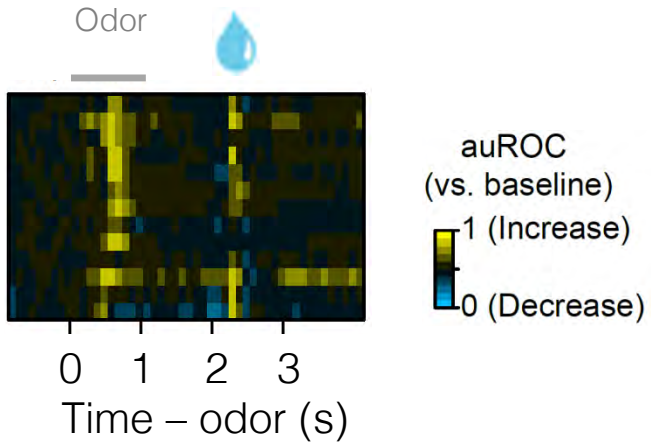


(Luscher & Malenka, 2011; Hyman, Malenka & Nestler, 2006)

- Addictive drugs inhibit VTA GABA neurons
- Persistent reinforcement signals by addictive drugs (Redish, 2004)

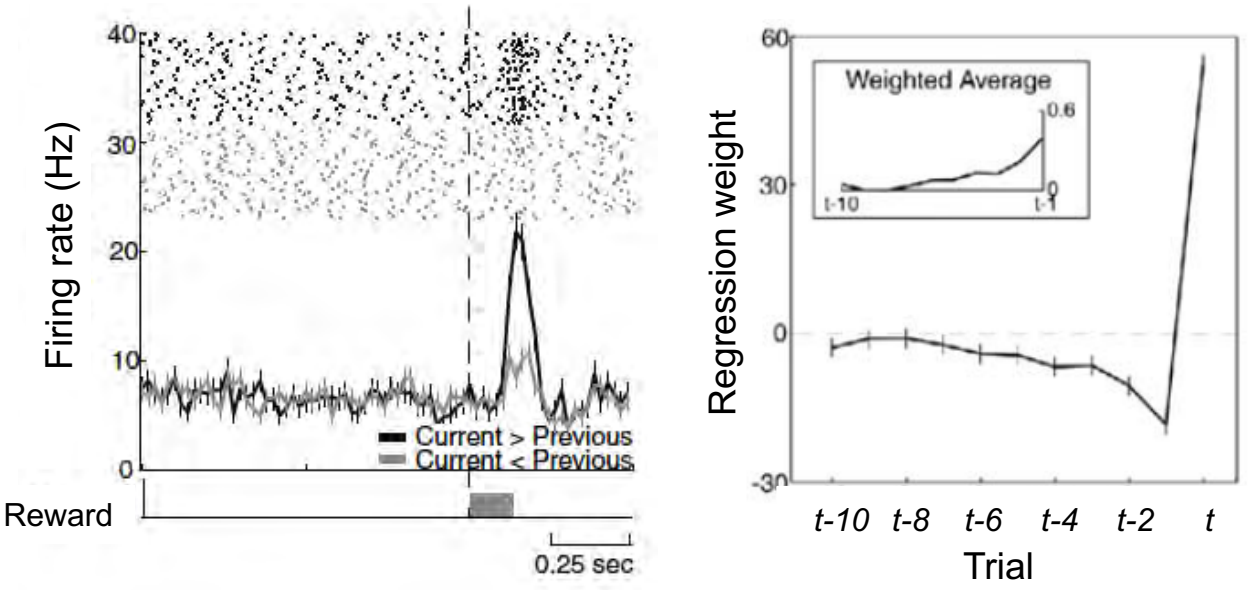
RPE?





(Cohen, Haesler et al., *Nature*, 2012)

Dopamine = Present reward – weighted sum of previous rewards



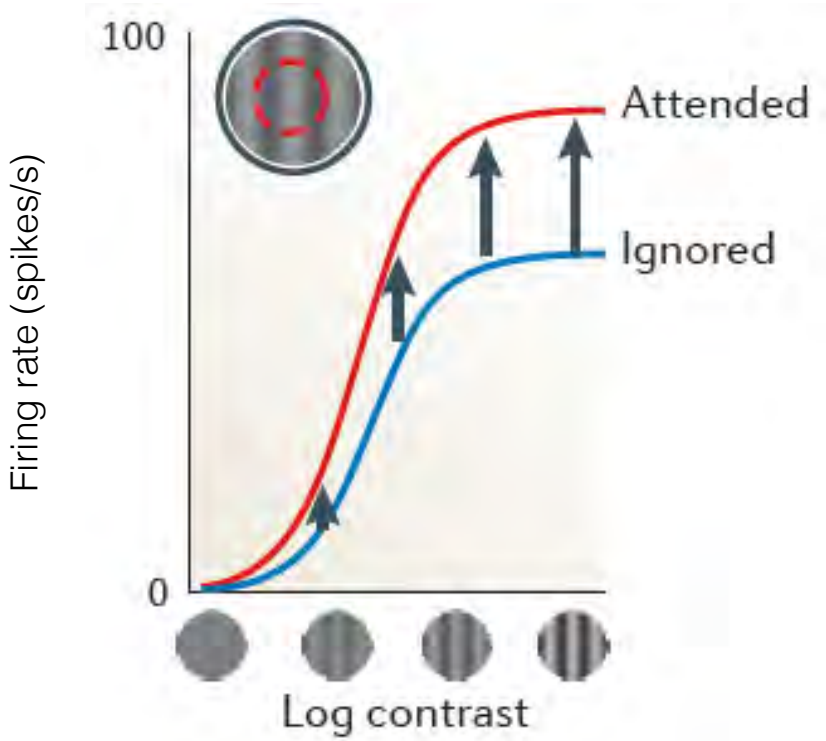
- Regression analysis

$$FR = b_0 + b_1 r_t + b_2 r_{t-1} + b_3 r_{t-2} + \cdots$$

FR: firing rate of dopamine neurons
r_t: reward size at trial t
b: regression weight

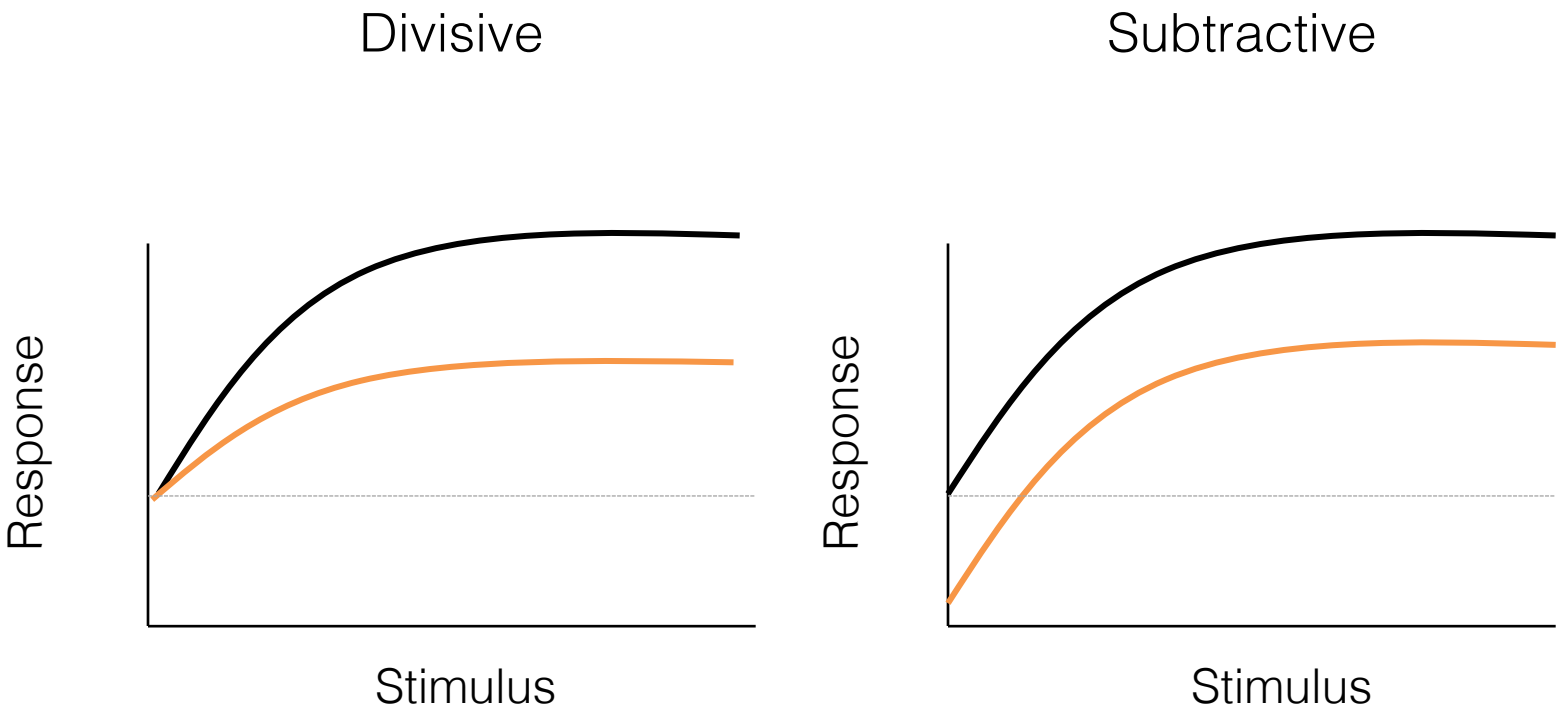
(Bayer and Glimcher, 2005)

Divisive (multiplicative) “gain” change



Carandini & Heeger, 2011; Bonin et al., 2005 Williford & Maunsell, 2006

Testing computations

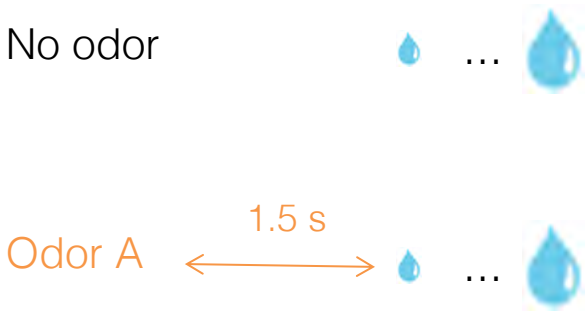


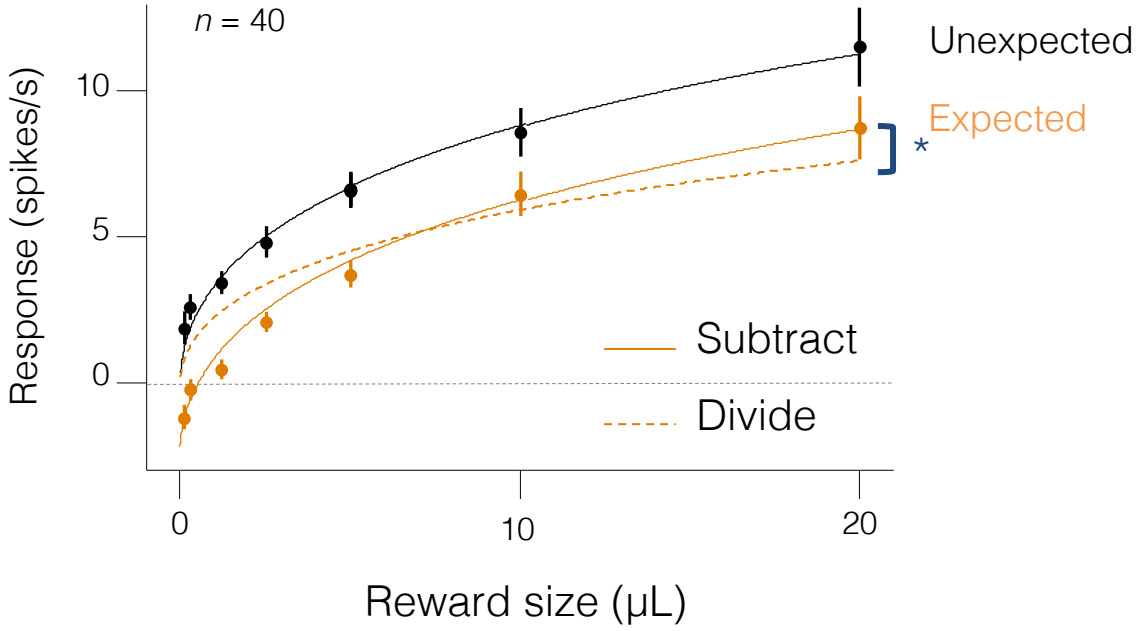
(Eshel, Bukwich, Rao, Hemmelder, Tian and Uchida, *Nature*, 2015)

Reward expectation triggers subtraction

(0-600ms after reward onset)

- Task



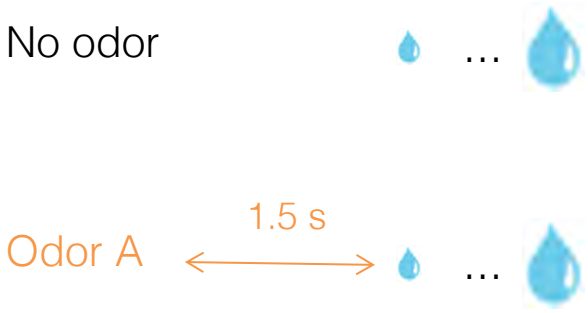


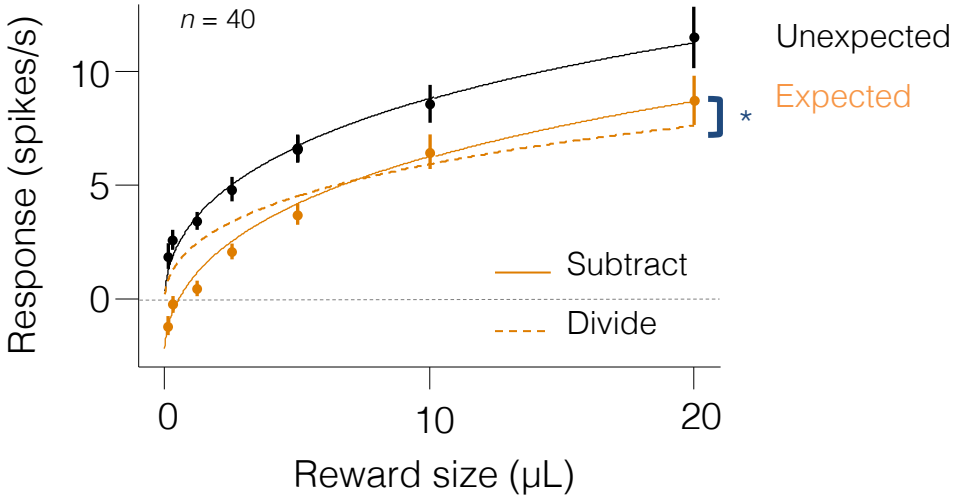
*P < 0.001, bootstrap to compare model fits

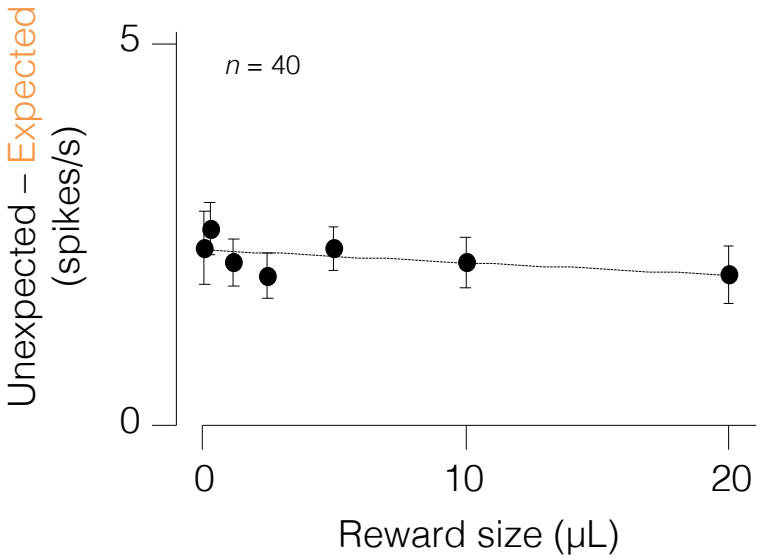
(Eshel, Bukwich, Rao, Hemmelder, Tian and Uchida, *Nature*, 2015)

Reward expectation triggers subtraction

- Task



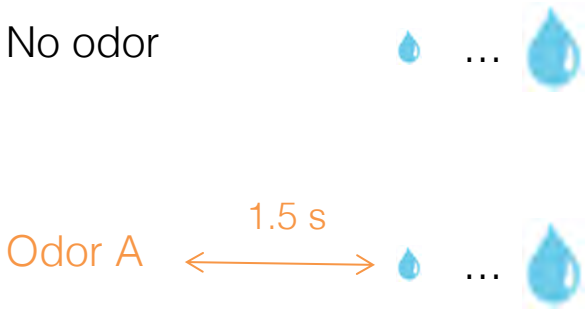


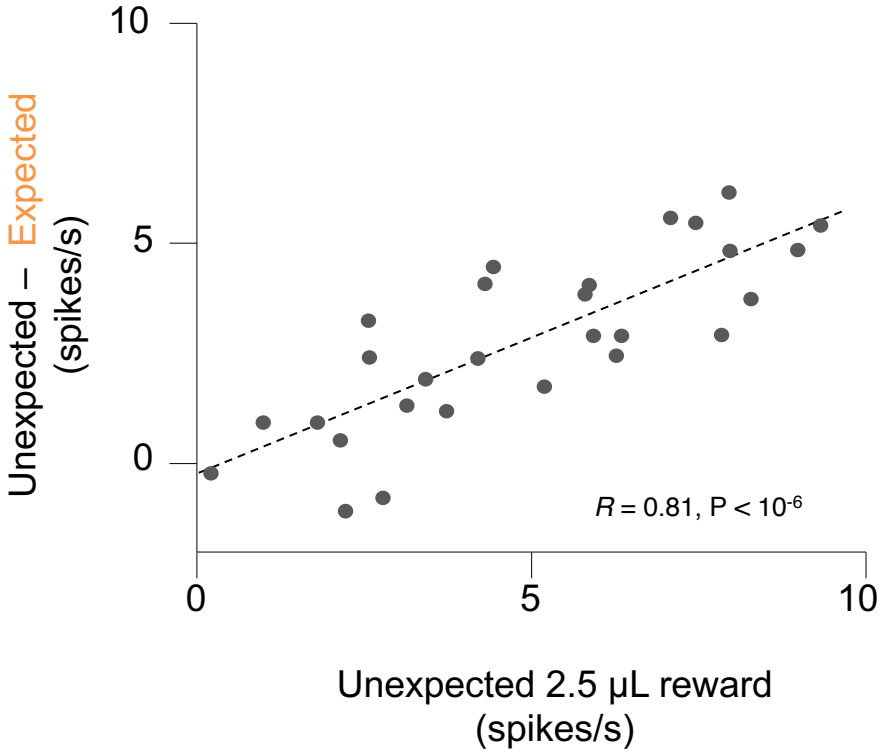


(Eshel, Bukwich, Rao, Hemmelder, Tian and Uchida, *Nature*, 2015)

Subtraction is **scaled** by reward response

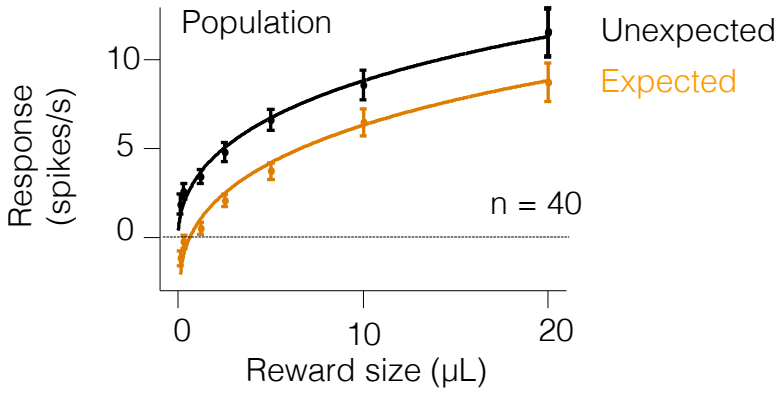
- Task





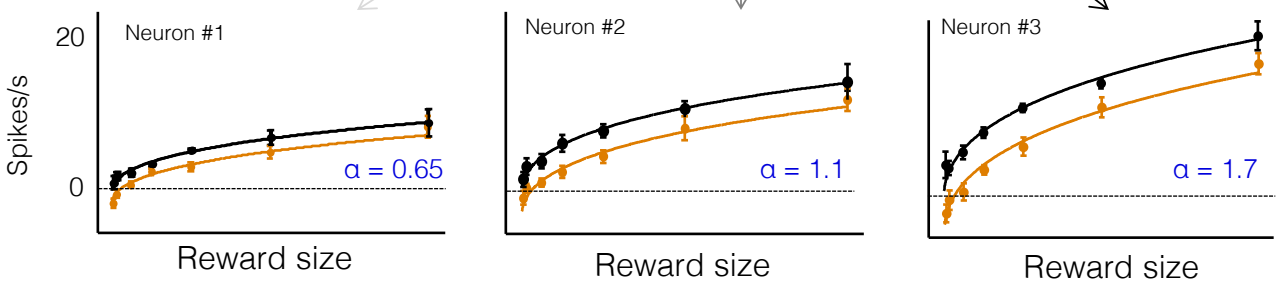
(Eshel, Bukwich, Rao, Hemmelder, Tian and Uchida, *Nature*, 2015)

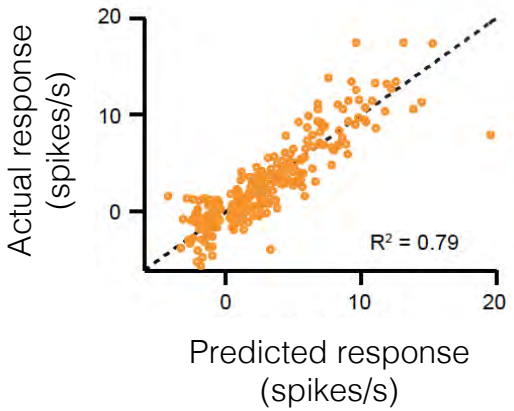
Dopamine neurons follow a universal template



"Universal function"

x α



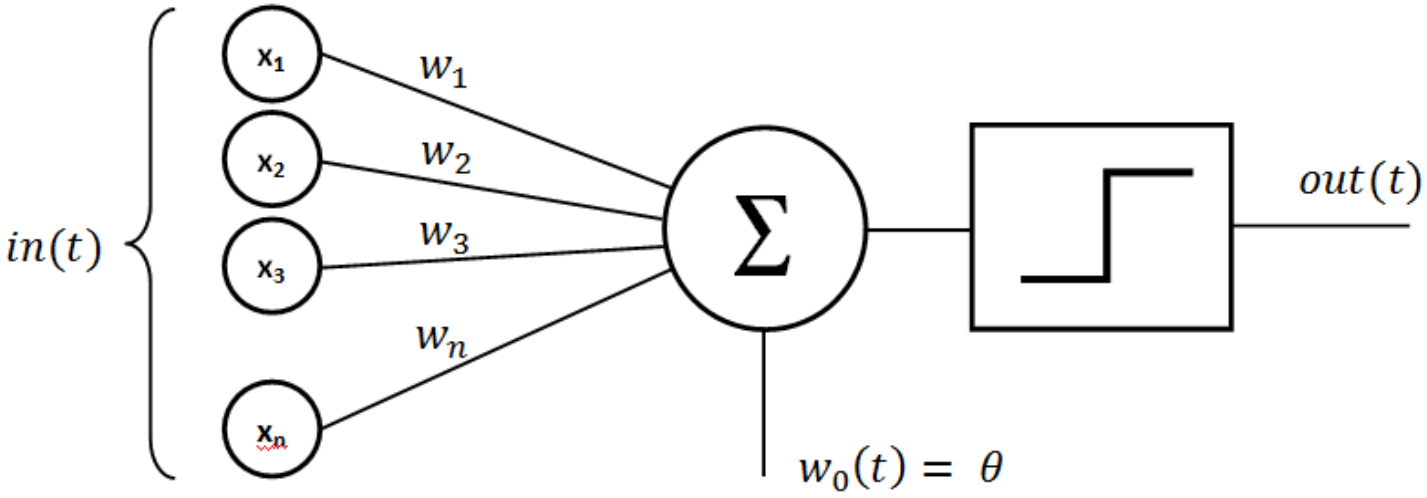


(Eshel, Bukwich, Tian and Uchida, *Nature Neurosci*. 2016)

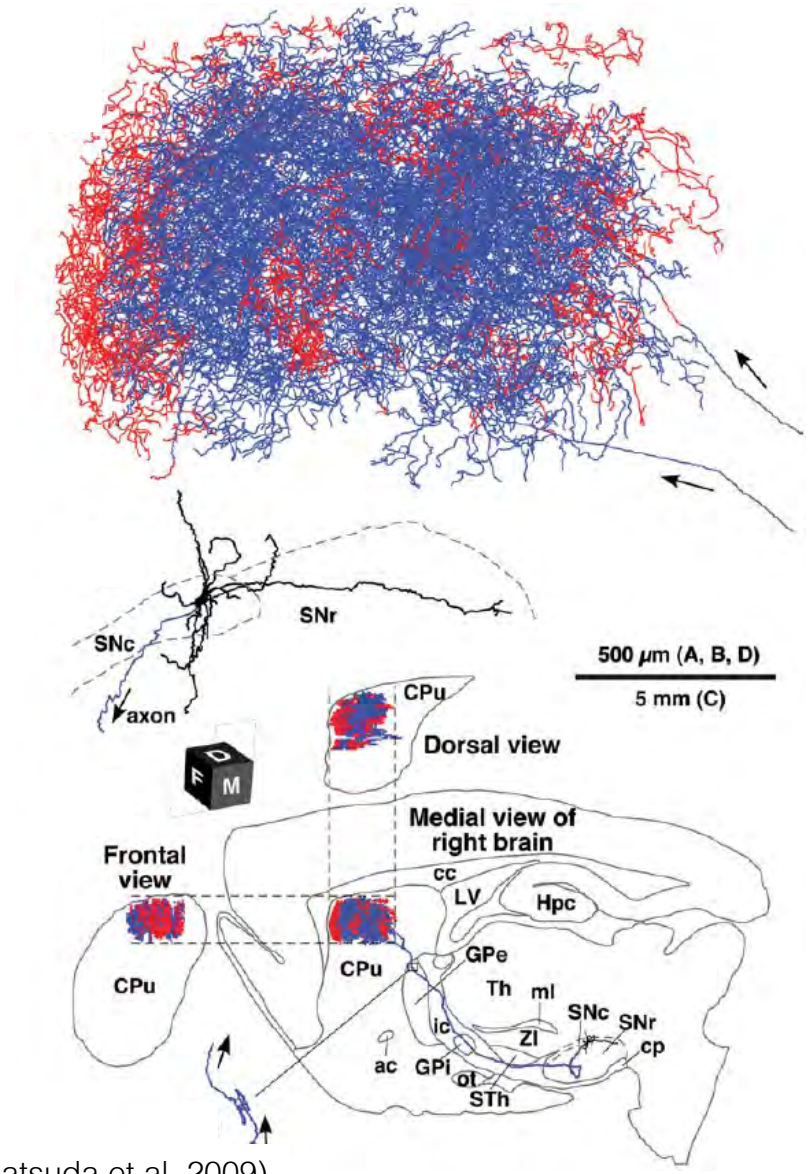
Summary

- Nearly all dopamine neurons exhibited phasic excitations to reward-predictive cues and reward.
- Some dopamine neurons show biphasic responses to aversive events but these responses are diminished in low reward contexts.
- VTA GABA neurons signal reward expectation, which can suppress dopamine reward responses when reward is expected.
- Reward expectation reduces dopamine reward responses in a subtractive fashion.
- Subtraction was scaled by a neuron's responsiveness to reward. This relationship may naturally arise from balanced excitation and inhibition.

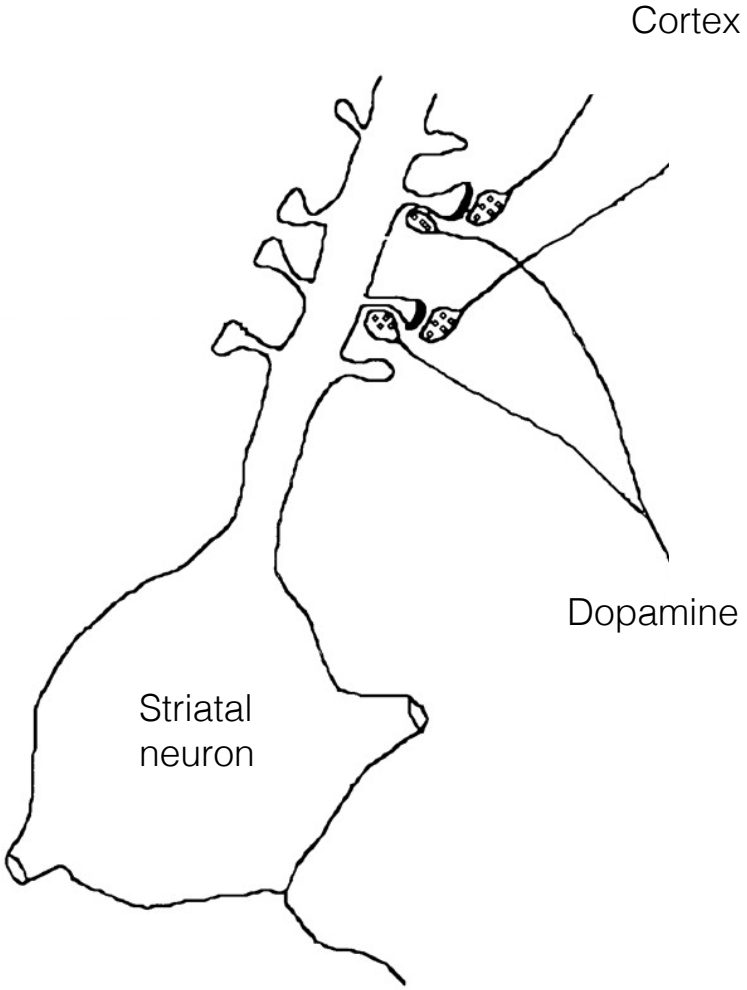
Population coding



What is a good teaching signal?
Many synapses, each synapse

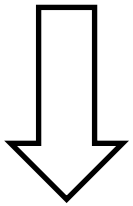


(Matsuda et al. 2009)

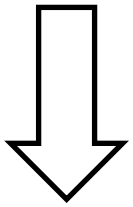


(Schultz, 1999)

Common inputs
balanced excitation/inhibition

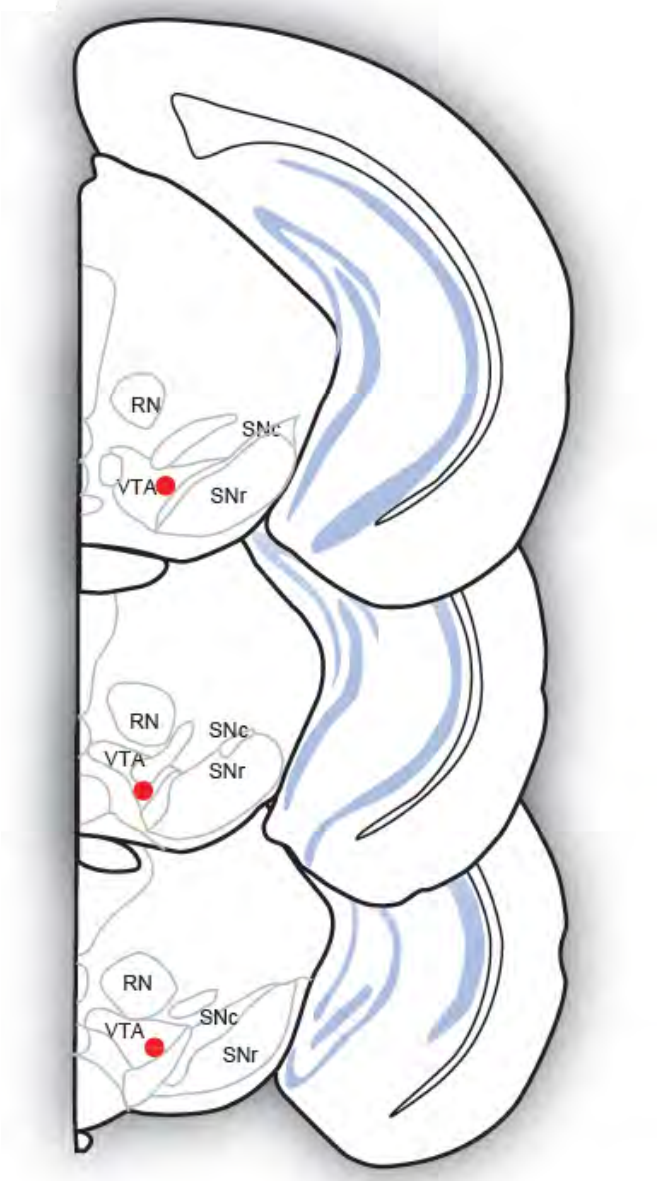


Homogeneity of dopamine signals



Consistency

Recording sites



Topics

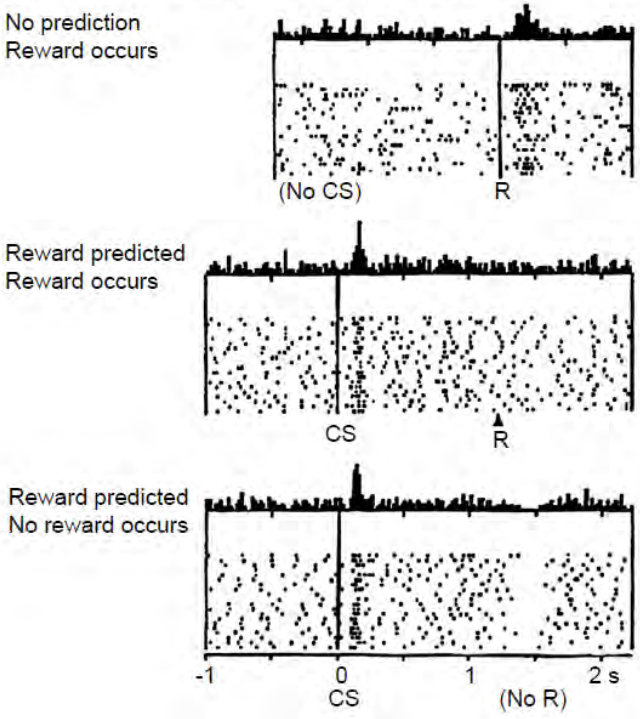
- A mouse model for studying dopamine RPE
- Do all dopamine neurons signal RPEs?
- What is the “state” in reinforcement learning?
- How are RPEs computed?
- Diversity of dopamine neurons

SCIENCE • VOL. 275 • 14 MARCH 1997

A Neural Substrate of Prediction and Reward

Wolfram Schultz, Peter Dayan, P. Read Montague*

- Phasic dopamine

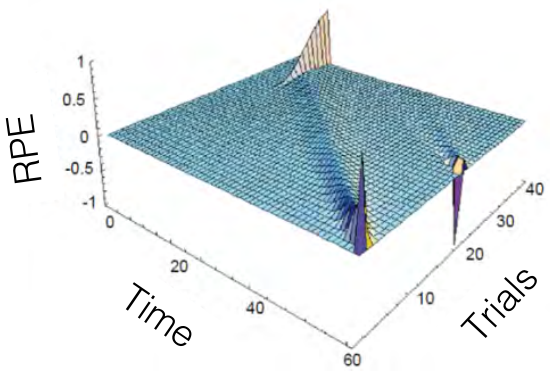


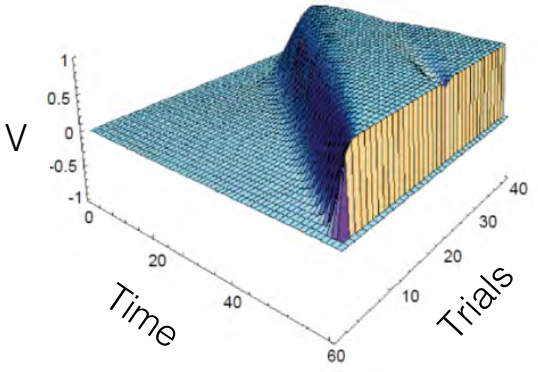
- Temporal difference (TD) learning theory

$$V(t) = E[\gamma^0 r(t) + \gamma^1 r(t+1) + \gamma^2 r(t+2) + \cdots] \quad (1)$$

$$V(t) = E[r(t) + \gamma V(t+1)] \quad (2)$$

$$\delta(t) = r(t) + \gamma \hat{V}(t+1) - \hat{V}(t) \quad (3)$$



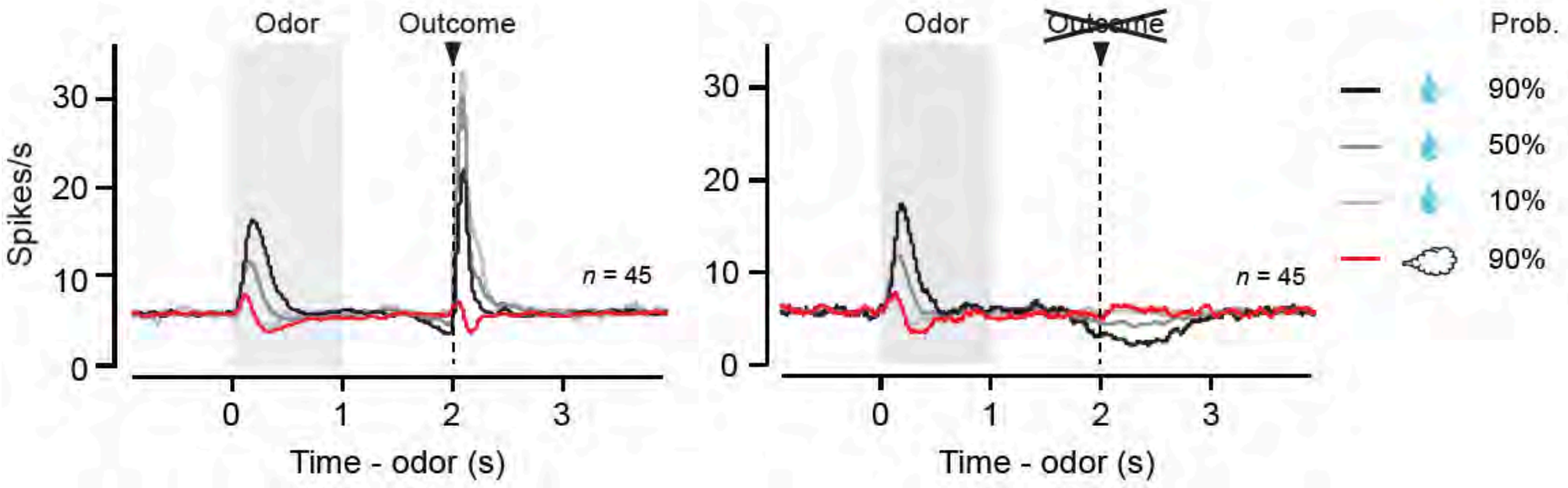


- Animal learning theory

Kamin, Rescorla, Wagner

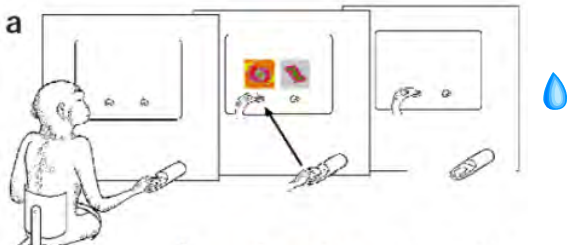
Recording from optogenetically-identified dopamine neurons

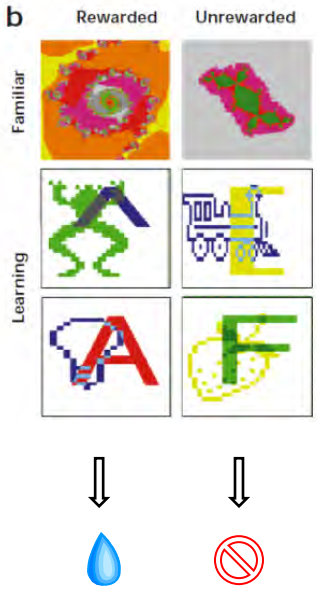
RPE coding by VTA dopamine neurons

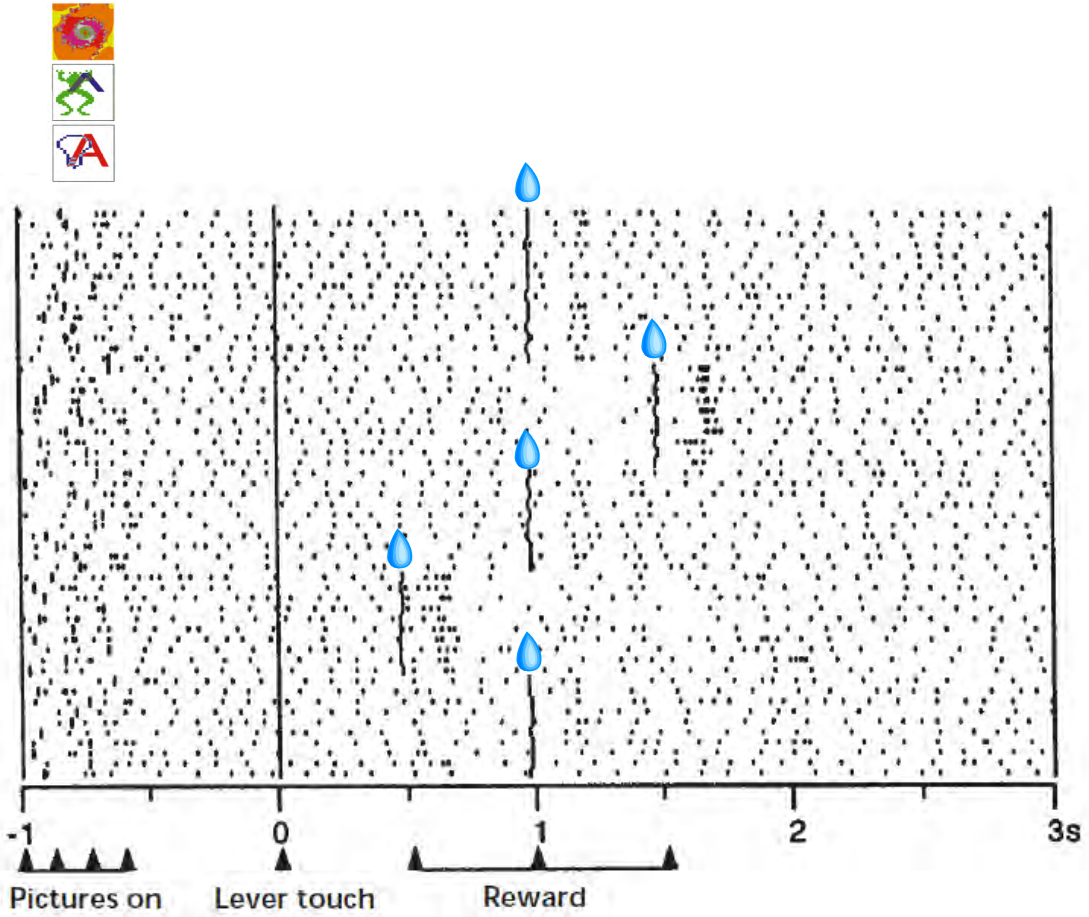


(Tian and Uchida, *Neuron*, 2015)

Dopamine neurons exhibit exquisite sensitivity to reward expectation over time

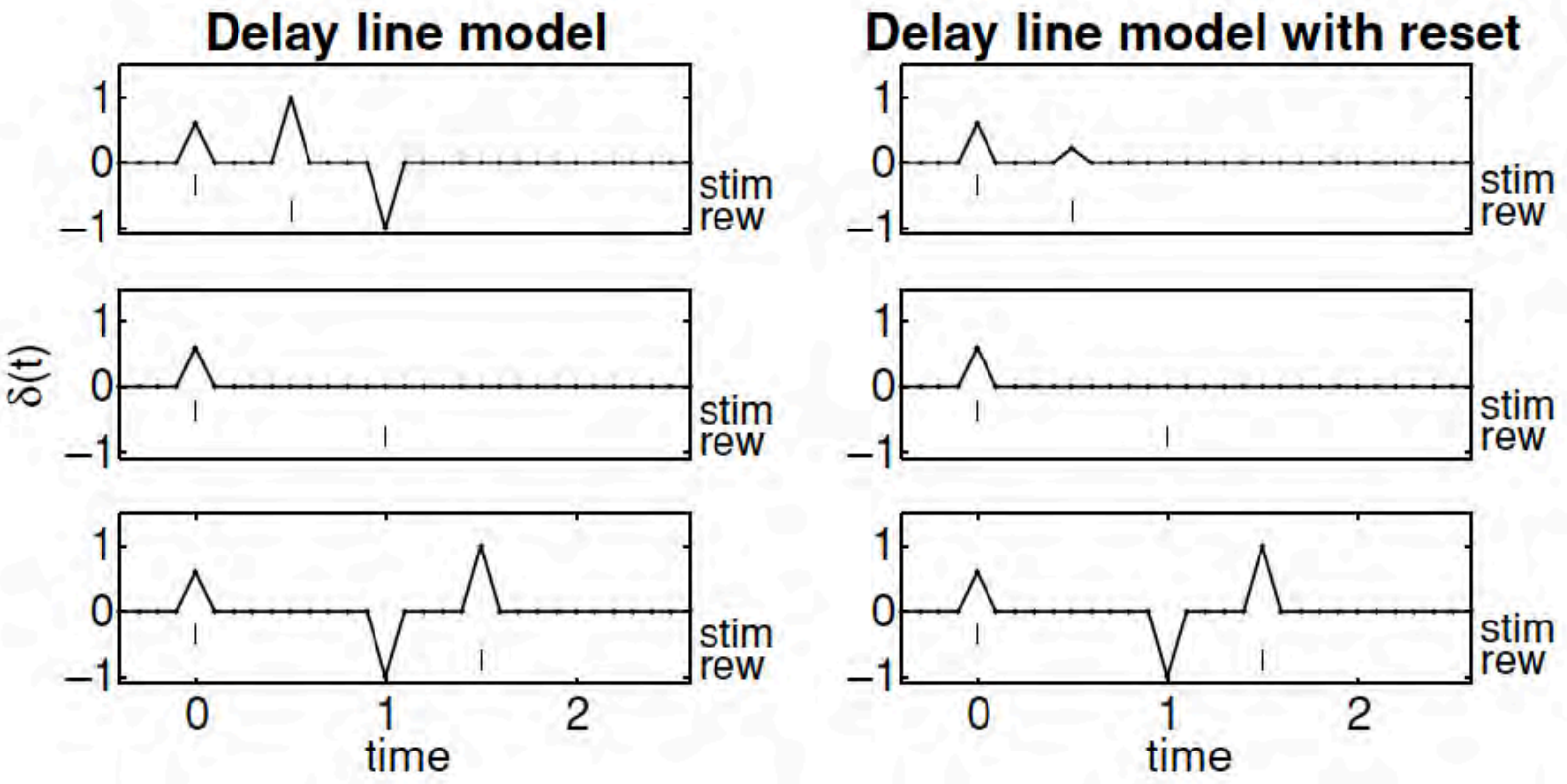






(Hollerman & Schultz, 1998, Dopamine neurons report an error in the temporal prediction of reward during learning)

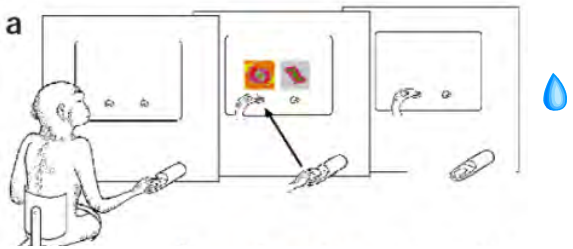
The TD model does not explain Hollerman & Schultz (2008)

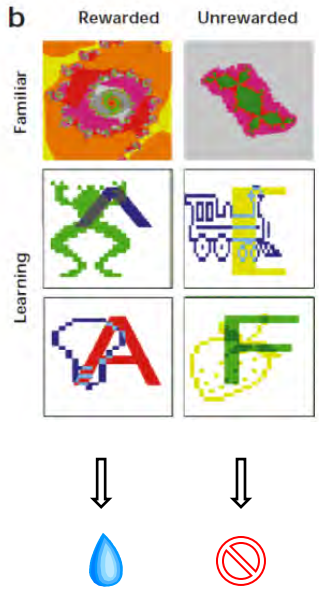


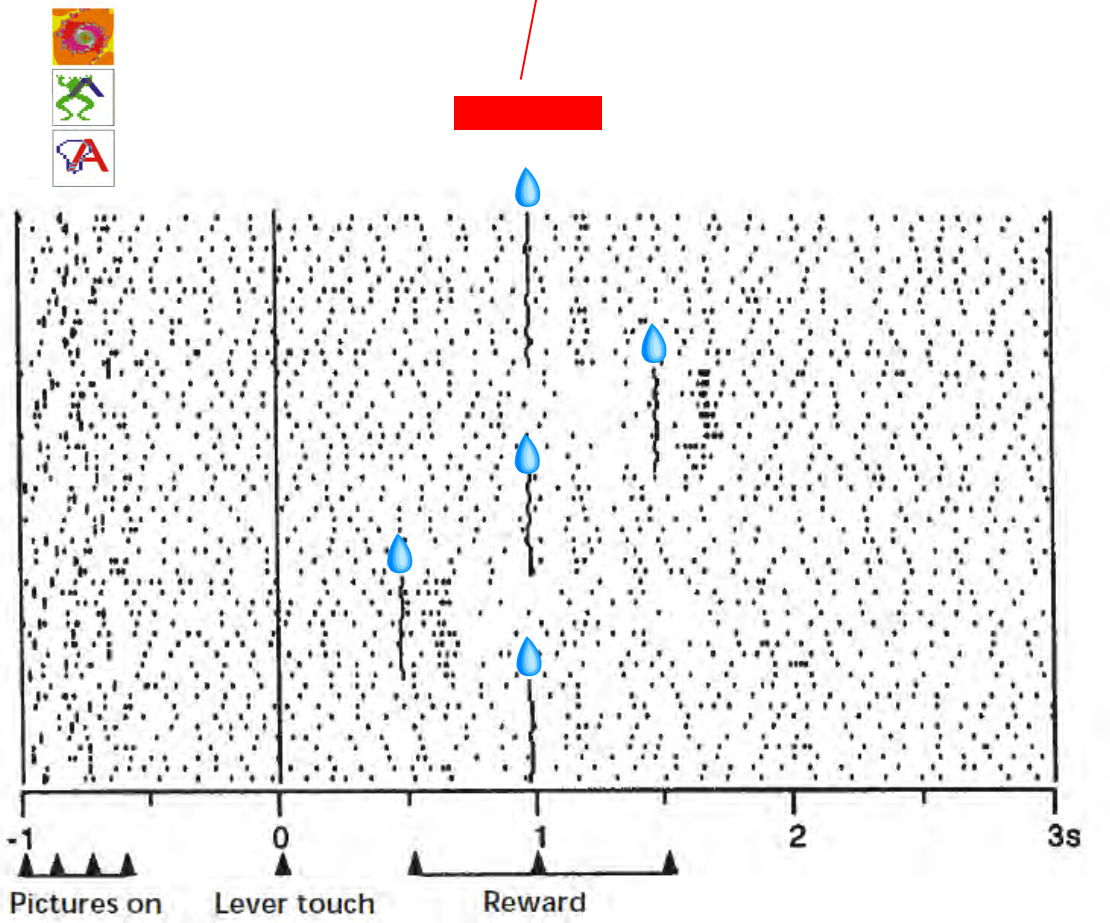
(Daw et al., 2006)

Dopamine neurons exhibit exquisite sensitivity to reward expectation over time

- Narrow time window for suppression





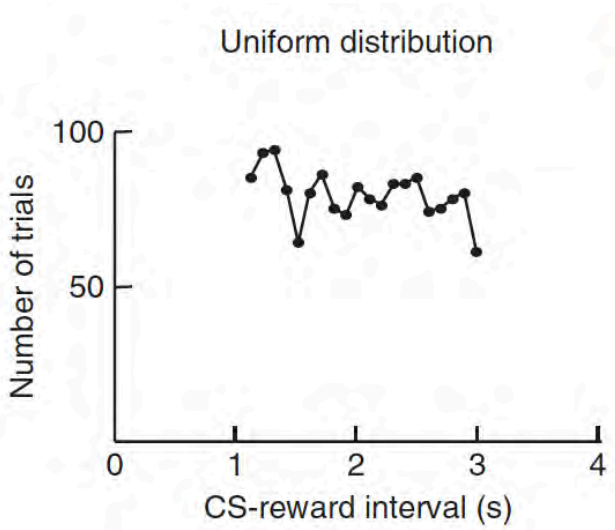


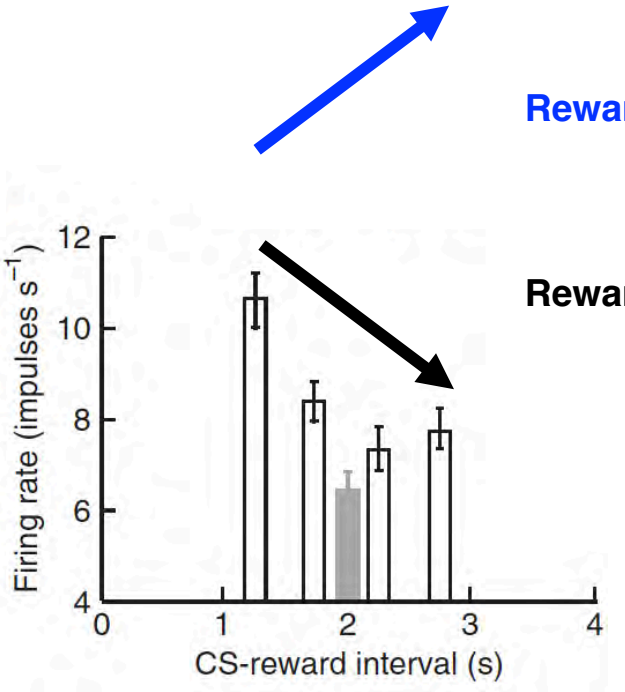
(Hollerman & Schultz, 1998, Dopamine neurons report an error in the temporal prediction of reward during learning)

Dopamine response in a variable delay condition

Cue → Reward

Variable delay





(Fiorillo, Newsome, Schultz, 2008)

Expectation over time follows hazard rate

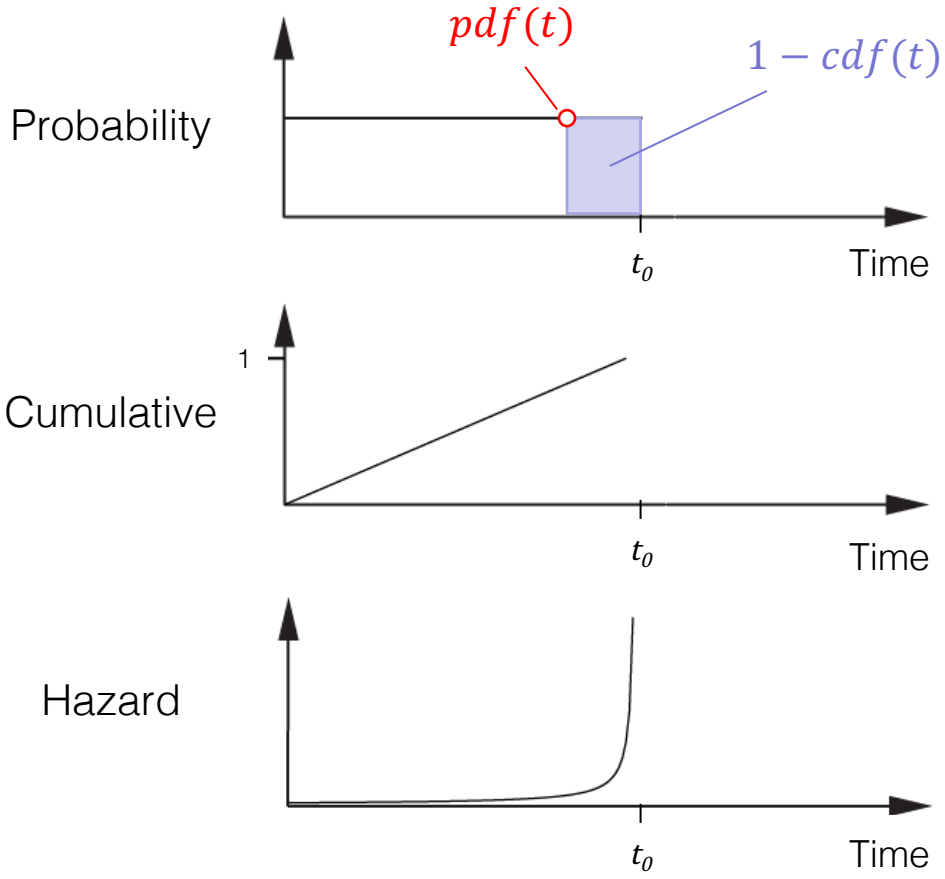
- **Hazard rate**: Likelihood that an event will occur given that it has not yet happened.

$$h(t) = \frac{pdf(t)}{1 - cdf(t)}$$

$h(t)$: Hazard rate

$pdf(t)$: Probability density

$cdf(t)$: Cumulative distribution



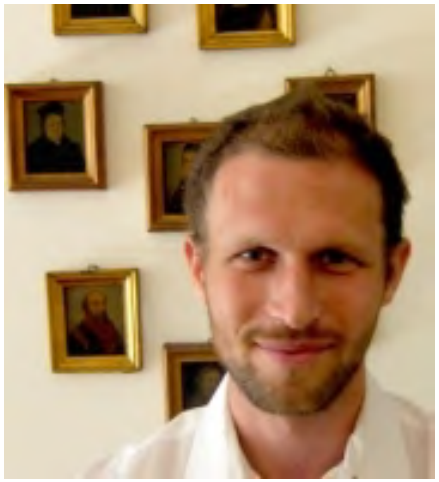
Timing sensitivity of dopamine neurons

- Dopamine “dip” at the time of expected reward
  - Schultz, Dayan, Montague (1997)
  - Hollerman and Schultz (1998)
  - Lateral habenula: Matsumoto and Hikosaka (2008), Tian and Uchida (2015)
- Hazard-like modulation of dopamine RPE
  - Nakahara et al. 2004
  - Fiorillo, Newsome, Schultz, 2008
  - Nomoto et al, 2010
  - Pasquereau and Turner, 2015

Question

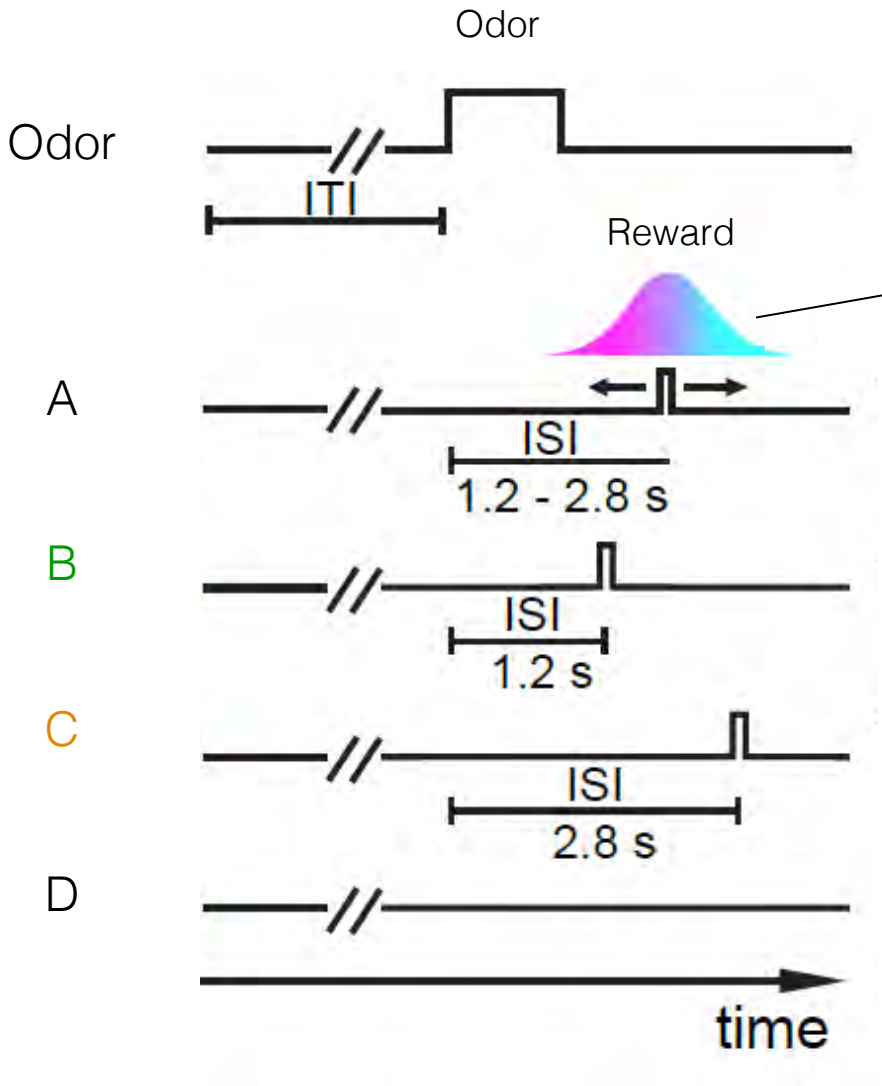
- What is the shape of the expectation function across time?
- Is hazard rate a good explanation for it?
- Which brain areas are involved in computing reward expectation across time?

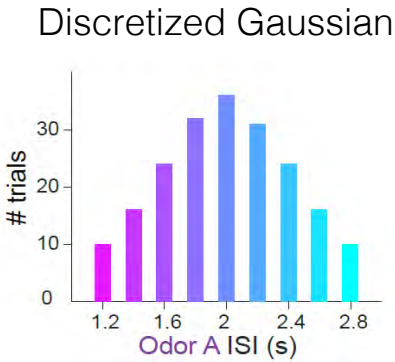
Clara Starkweather



Samuel Gershman

Task





- Task 1: P(reward) = 1
- Task 2: P(reward) = 0.9

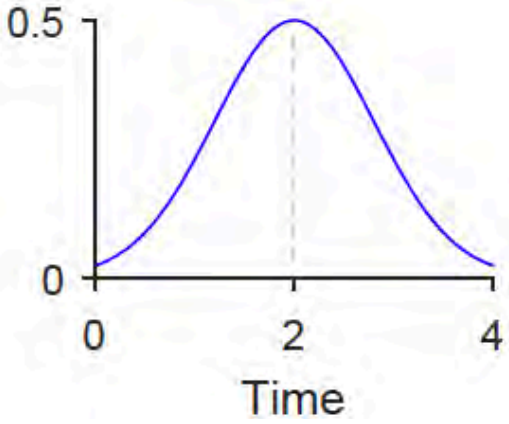
(Starkweather, Babayan, Uchida, Gershman, *Nature. Neurosci.* 2017)

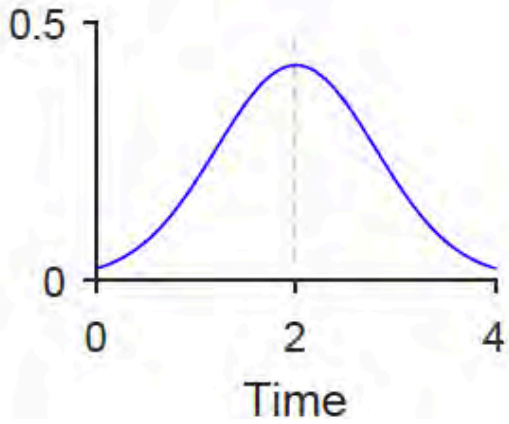
Prediction

Task 1
100% Rewarded

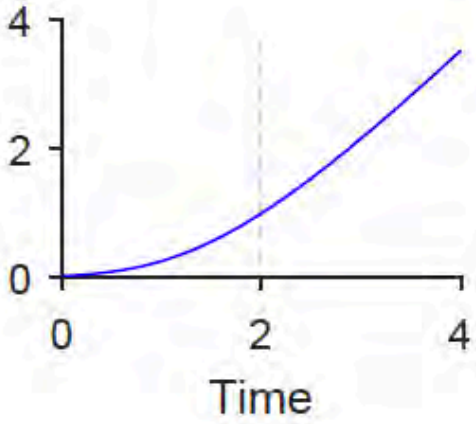
Task 2
90% Rewarded

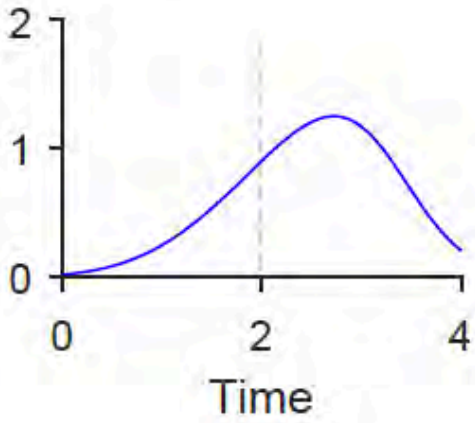
Probability





Hazard





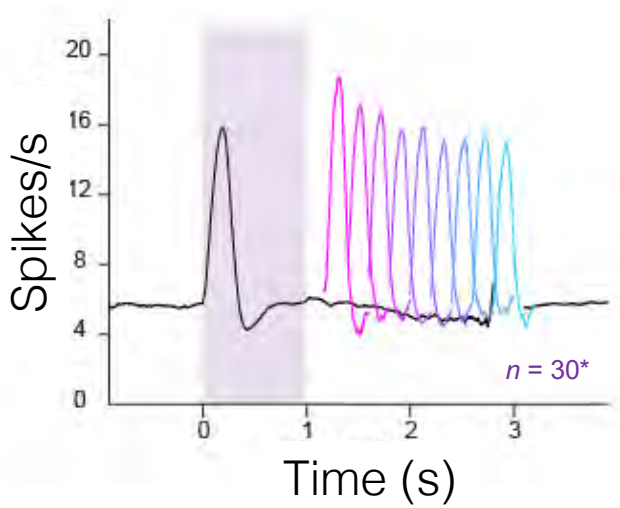
(Starkweather, Babayan, Uchida, Gershman, *Nature. Neurosci*. 2017)

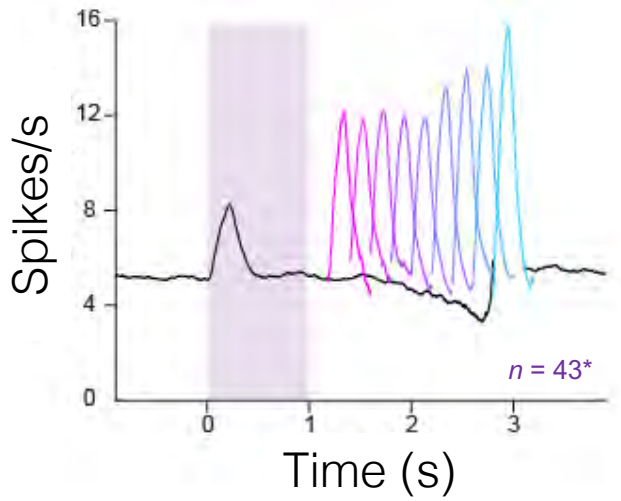
Data cannot be explained by hazard rate

Task 1
100% Rewarded

Task 2
90% Rewarded

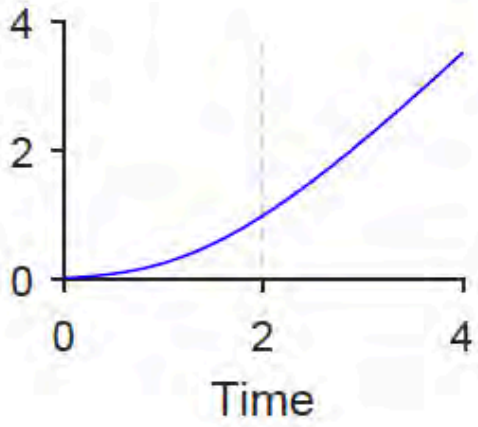
Dopamine (Data)

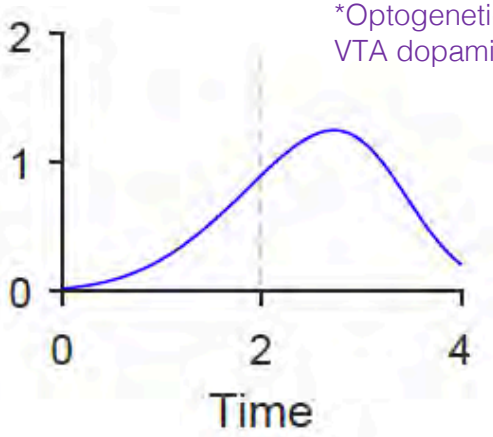




*Optogenetically-identified VTA dopamine neurons

Hazard





(Starkweather, Babayan, Uchida, Gershman, *Nature. Neurosci.* 2017)

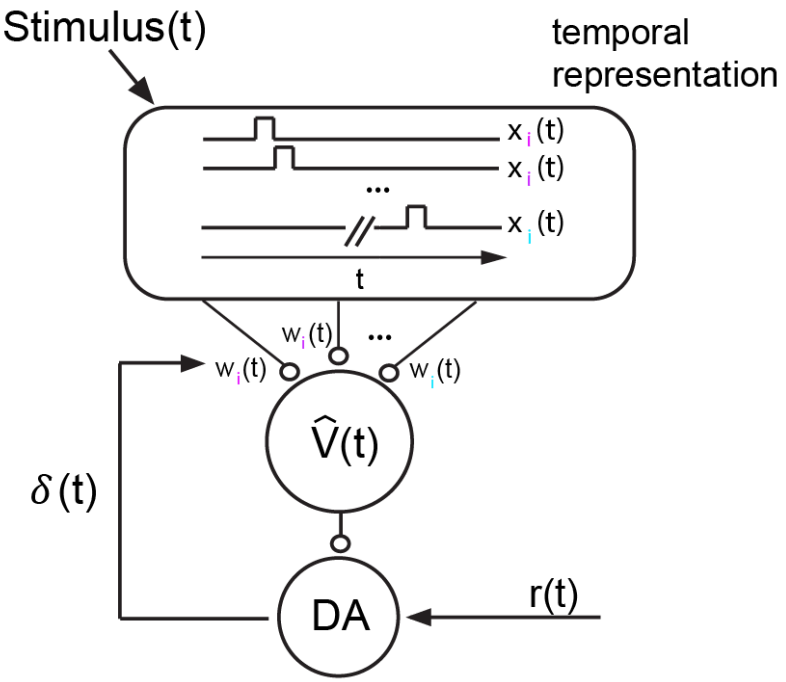
Can a simple learning algorithm explain our data?

- Temporal Difference (TD) learning
  - Moore et al. (1998), Sutton and Barto (1990), Montague et al. (1996), Schultz, Dayan, Montague (1997)
  - Keeps track of time after stimulus onset as a series of time steps
    - "Complete serial compound (CSC)"

$$V(t) = r(t) + \gamma^1 r(t+1) + \gamma^2 r(t+2) + \cdots$$

$$\hat{V}(t) = r(t) + \gamma \hat{V}(t+1)$$

$$\delta(t) = r(t) + \gamma \hat{V}(t+1) - \hat{V}(t)$$

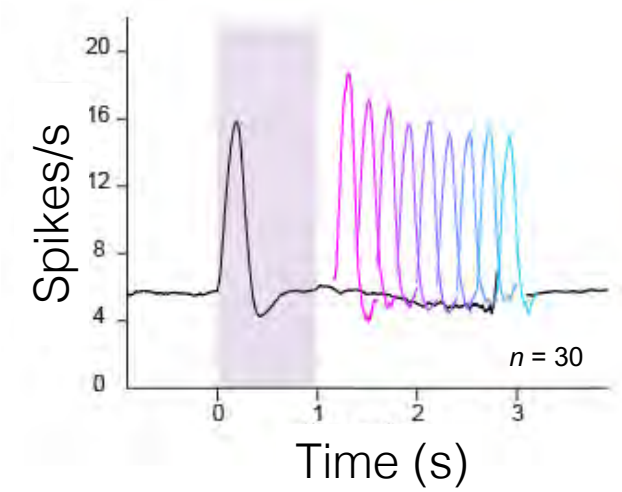


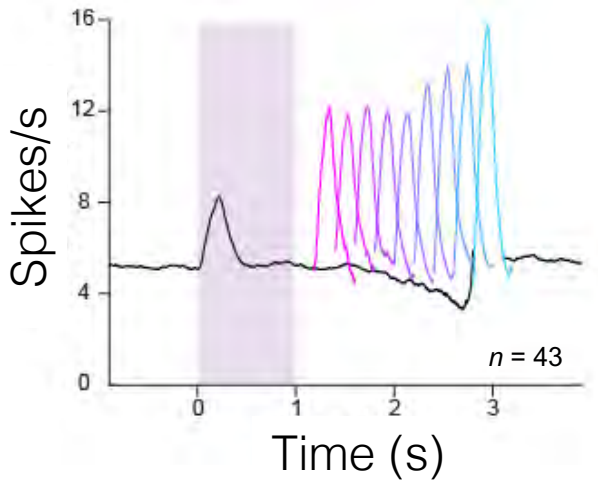
Classic TD learning cannot explain our data

Task 1
100% Rewarded

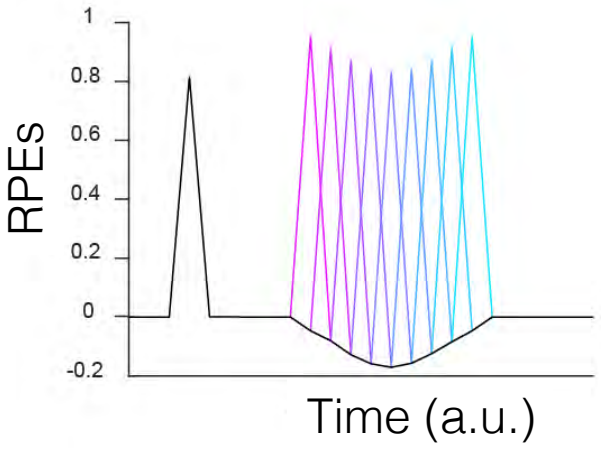
Task 2
90% Rewarded

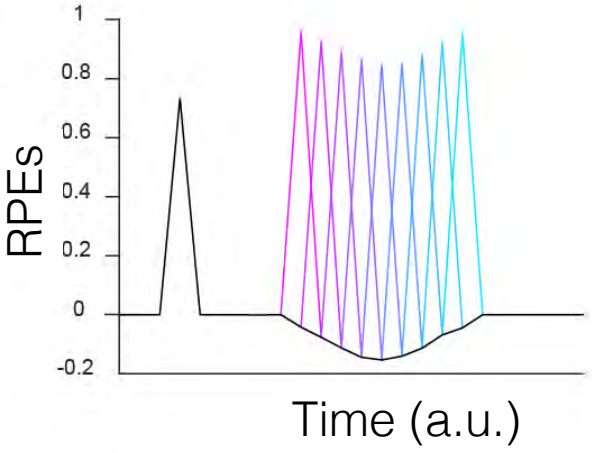
Dopamine (Data)





TD Model





(Starkweather, Babayan, Uchida, Gershman, *Nature. Neurosci*. 2017)

TD learning with CSC *with reset* generates *hazard-like* dynamics

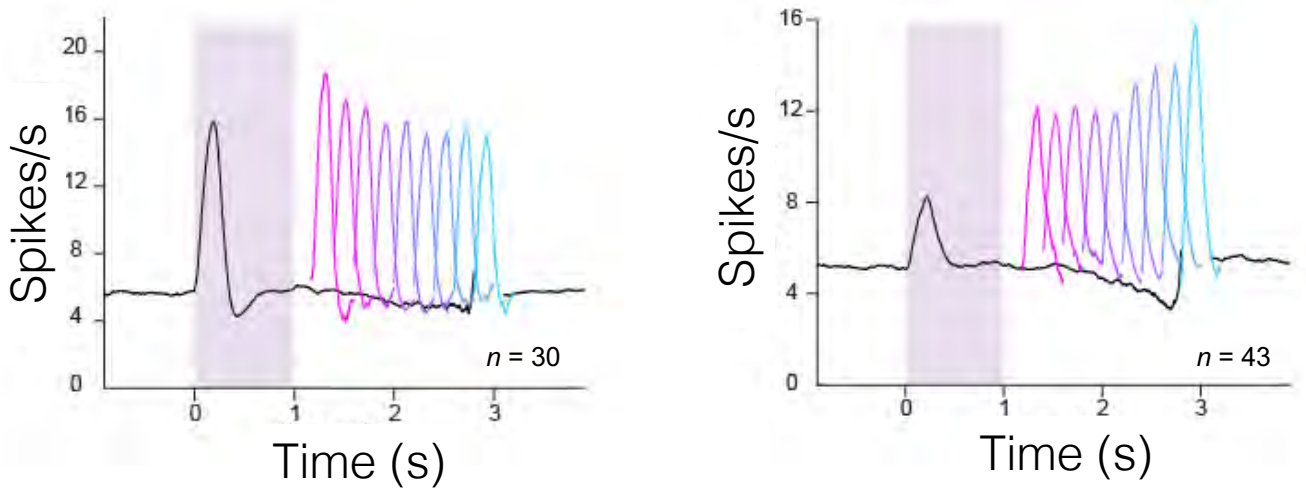
Task 1

100% Rewarded

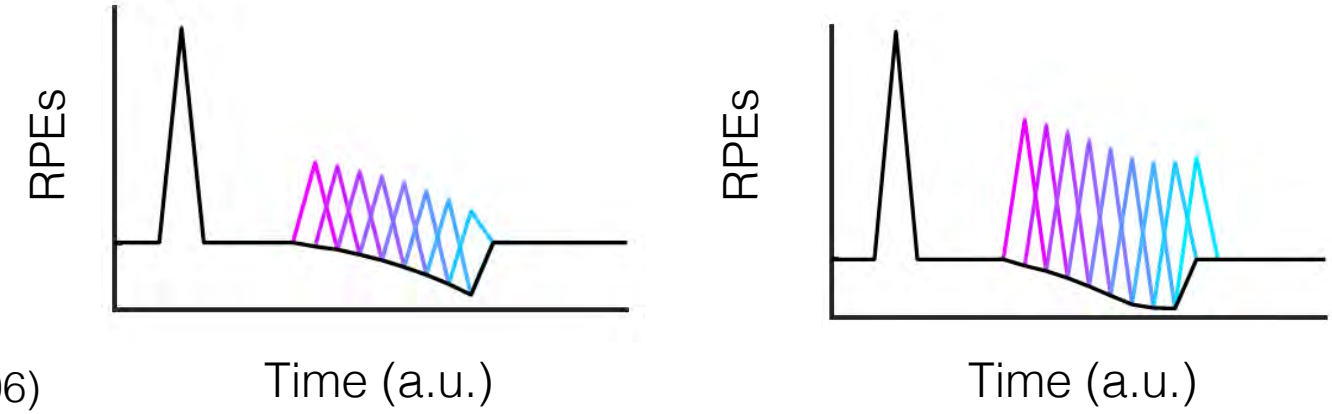
Task 2

90% Rewarded

Dopamine (Data)



TD Model + "reset"



(cf. Daw et al. 2006)

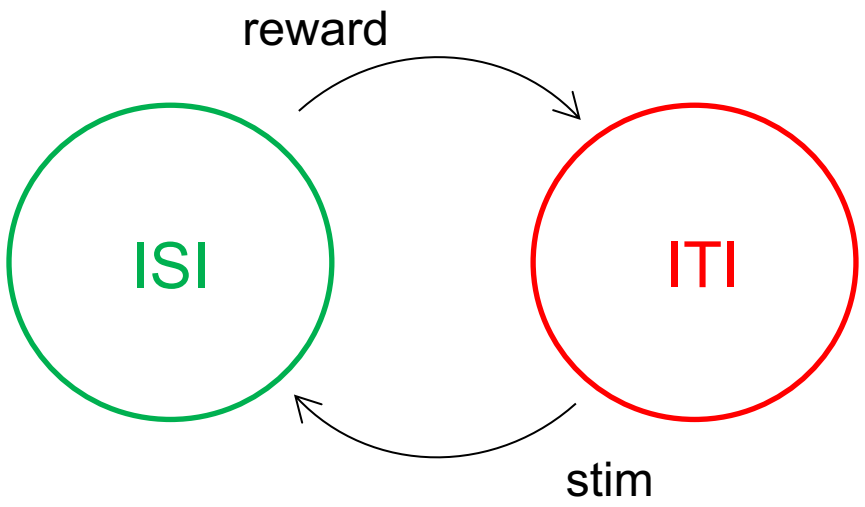
(Starkweather, Babayan, Uchida, Gershman, *Nature. Neurosci.* 2017)

Can a proposed modification to TD learning explain our data?

- Animals have to infer which “state” they are in.
- Proposed amendment
  - TD + belief state (rather than CSC)
  - Belief state = p(state | observations)
  - Daw et al, 2006; Rao et al, 2010
- RL models must operate on a belief state

(Starkweather, Babayan, Uchida, Gershman, *Nature. Neurosci*. 2017)

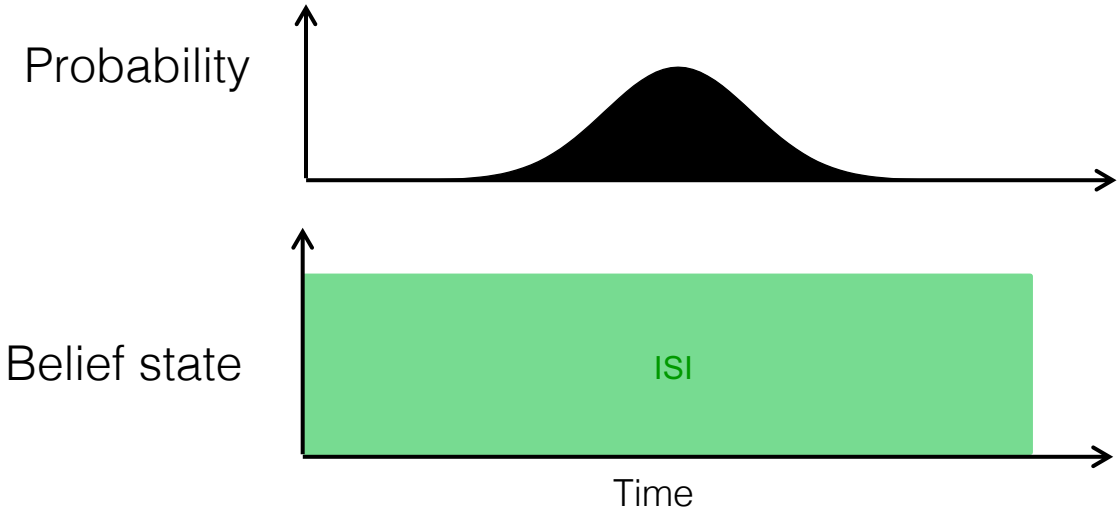
Belief state model



States:

- Reward will come (ISI)
- Reward won't come (ITI)

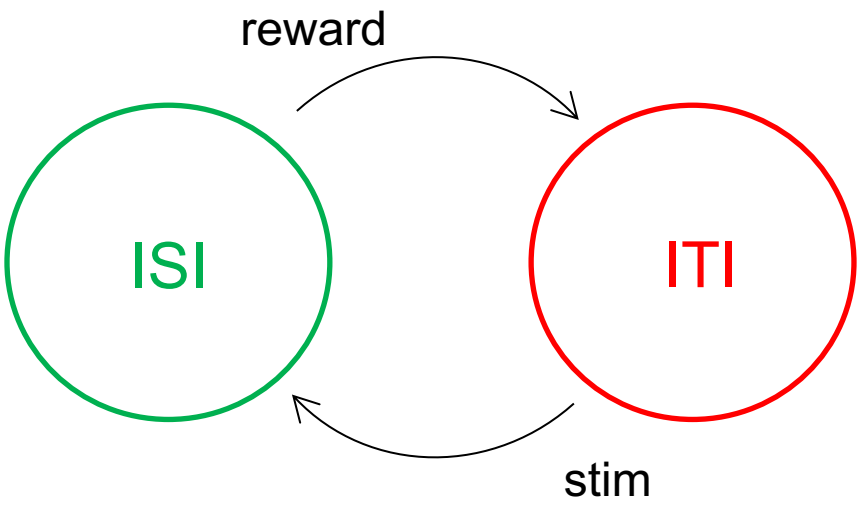
(Daw et al., 2006)



100% reward
(deterministic)

(Starkweather, Babayan, Uchida, Gershman, *Nature. Neurosci*. 2017)

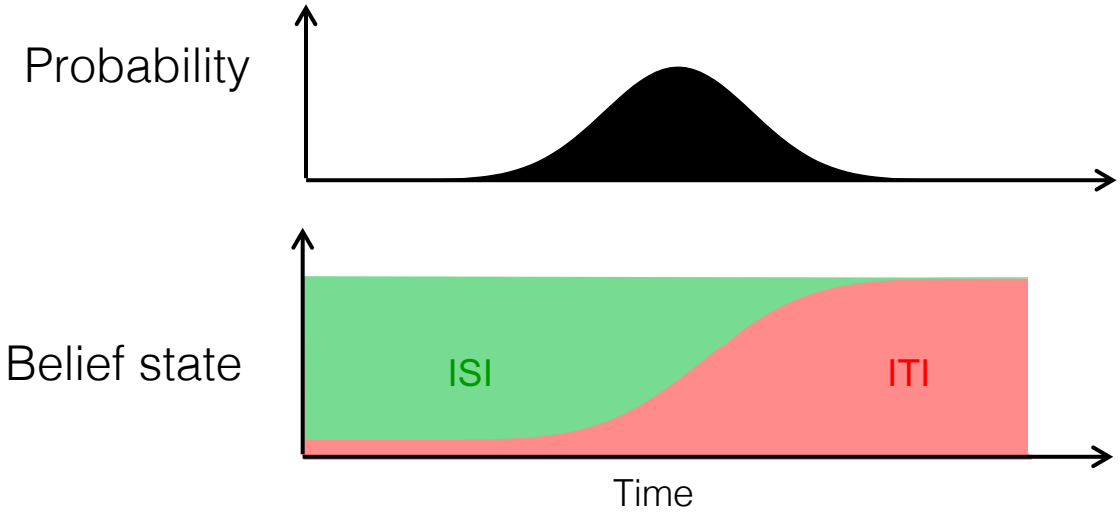
Belief state model



States:

- Reward will come (ISI)
- Reward won't come (ITI)

(Daw et al., 2006)



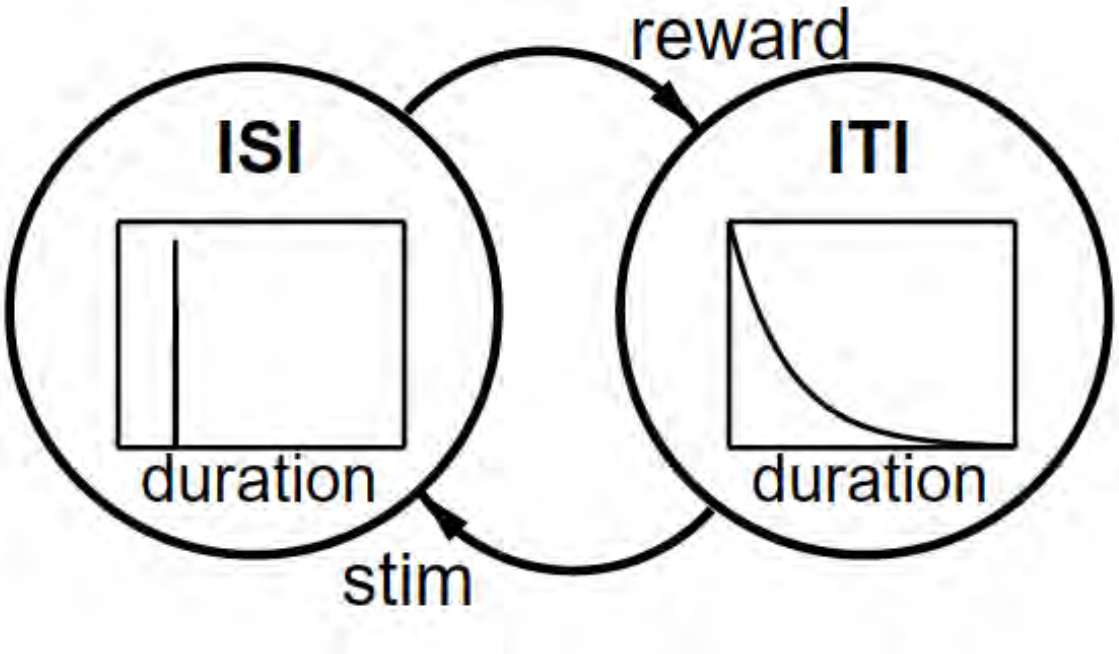
90% reward
(probabilistic)

(Starkweather, Babayan, Uchida, Gershman, *Nature. Neurosci*. 2017)

Expectation across time

- Dynamic Bayesian Inference
- Prior knowledge

A semi-Markov model of trace conditioning



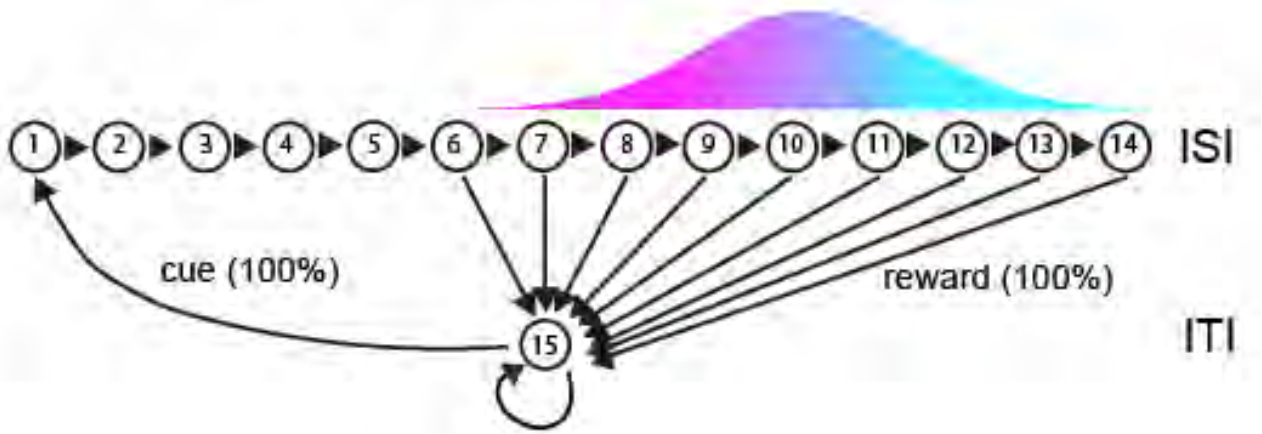
- State transitions
- Dwell time distribution

(Daw et al., 2006)

Belief state TD model

- Partially observable Markov decision process

Task 1 (100% reward)

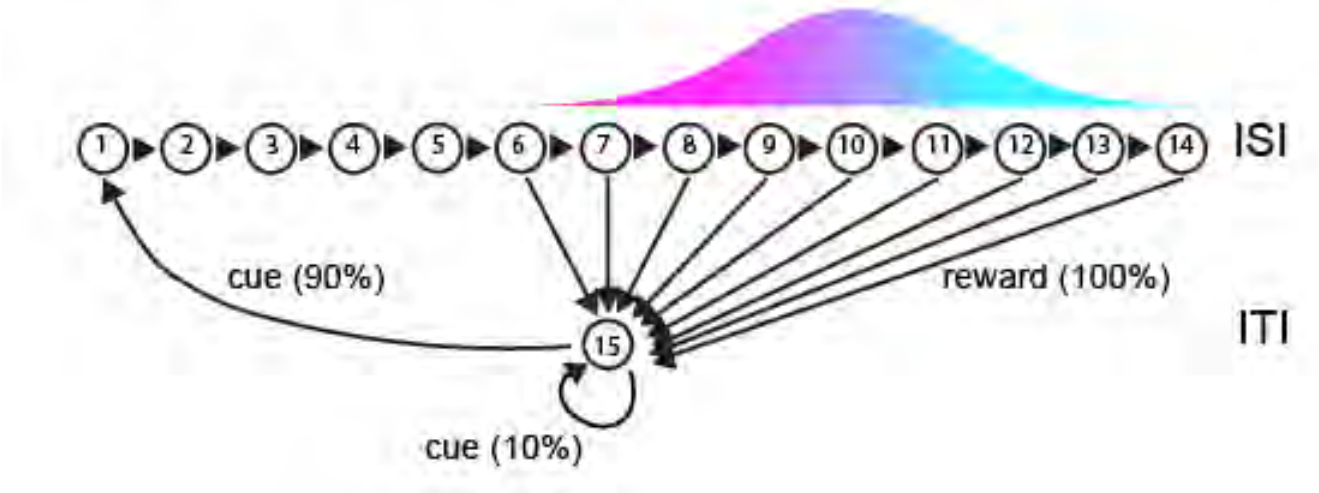


(Starkweather, Babayan, Uchida, Gershman, *Nature. Neurosci*. 2017)

Belief state TD model

- Partially observable Markov decision process

Task 2 (90% reward)

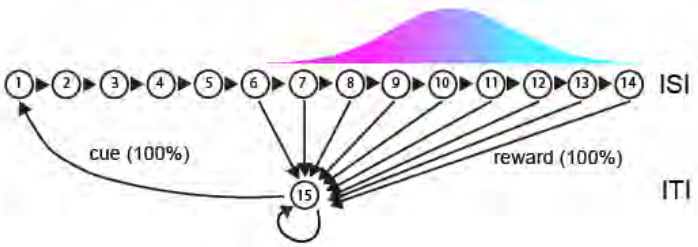


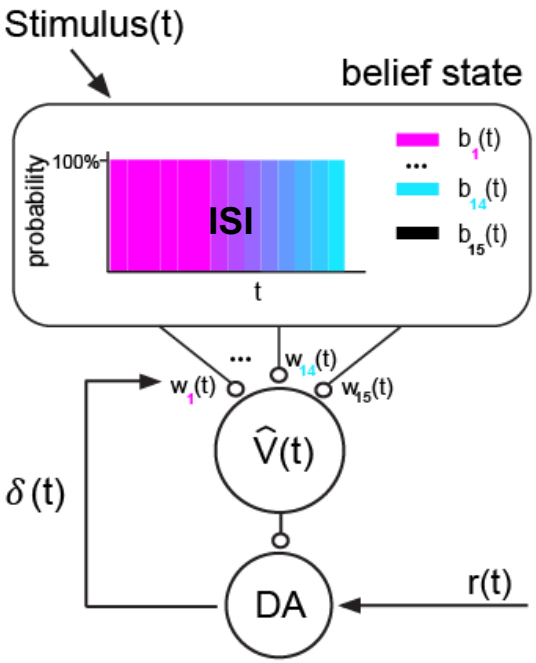
(Starkweather, Babayan, Uchida, Gershman, *Nature. Neurosci*. 2017)

Belief state TD model

Task 1

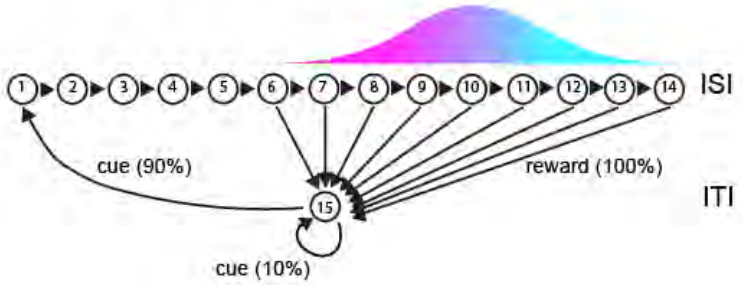
100% Rewarded

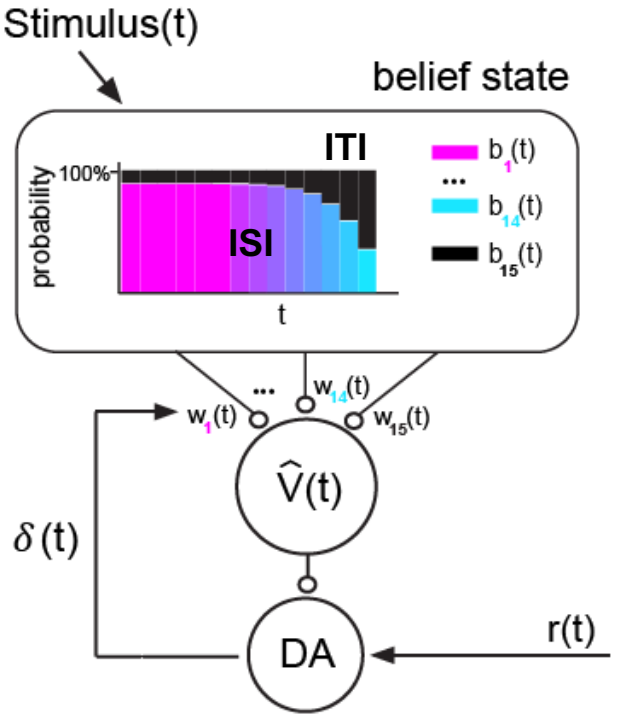




Task 2

90% Rewarded



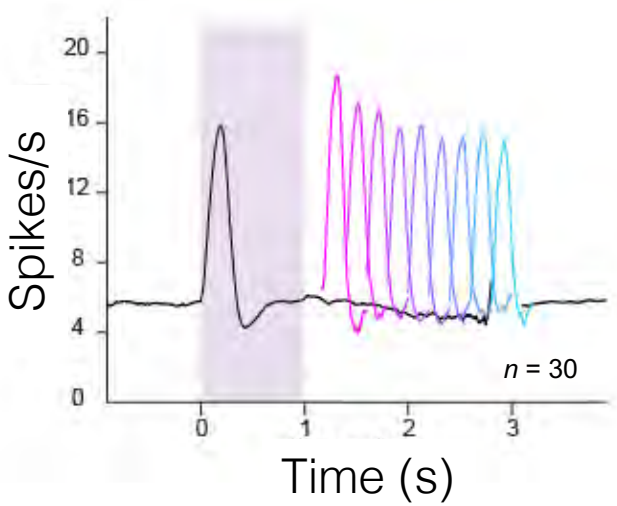


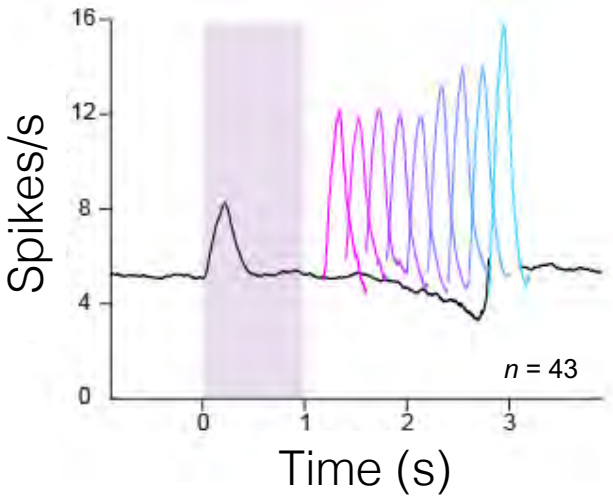
Belief state TD model captures data

Task 1
100% Rewarded

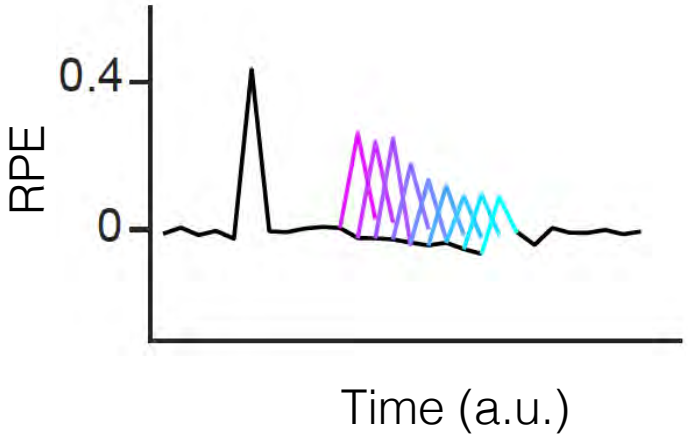
Task 2
90% Rewarded

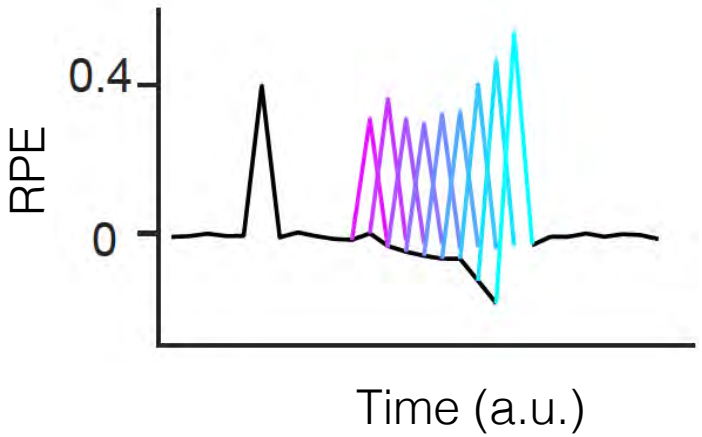
Dopamine (Data)





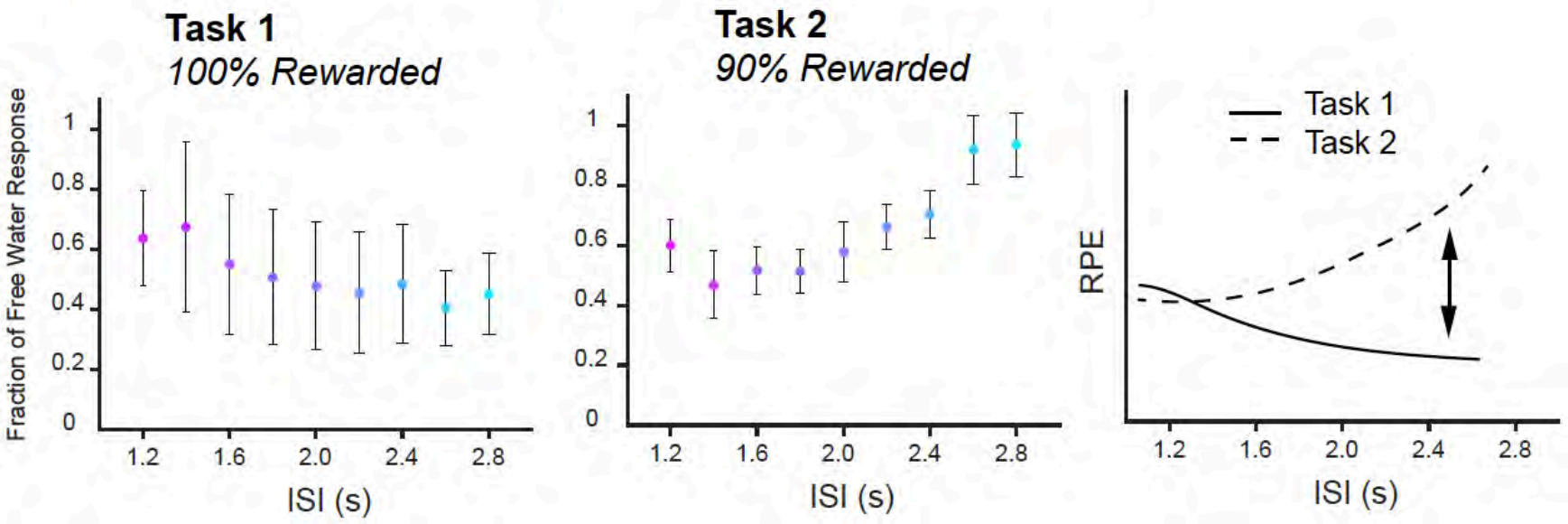
Belief state TD





(Starkweather, Babayan, Uchida, Gershman, *Nature. Neurosci.* 2017)

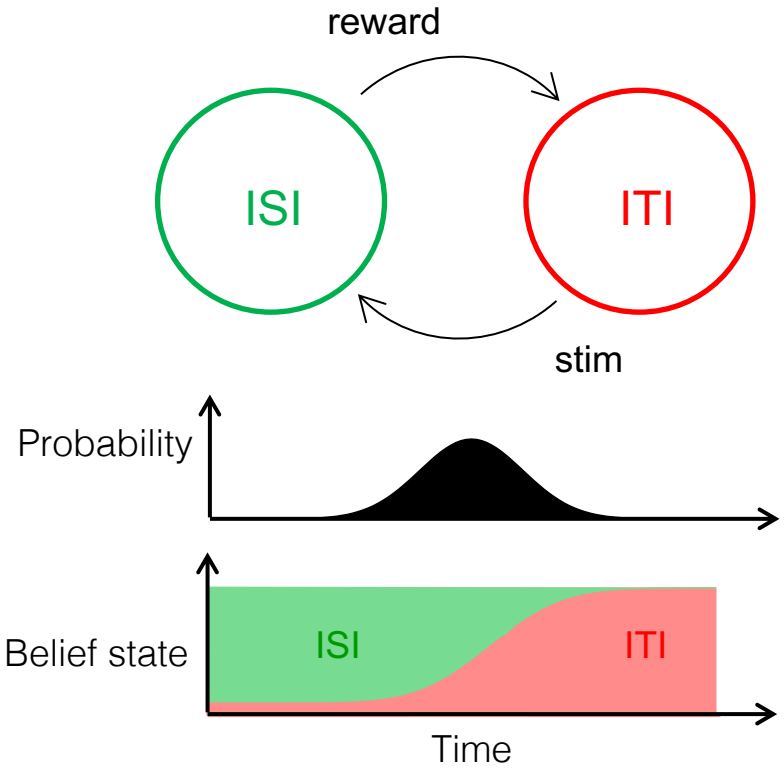
- Normalized by the responses to unexpected reward



(Starkweather, Babayan, Uchida, Gershman, *Nature. Neurosci*. 2017)

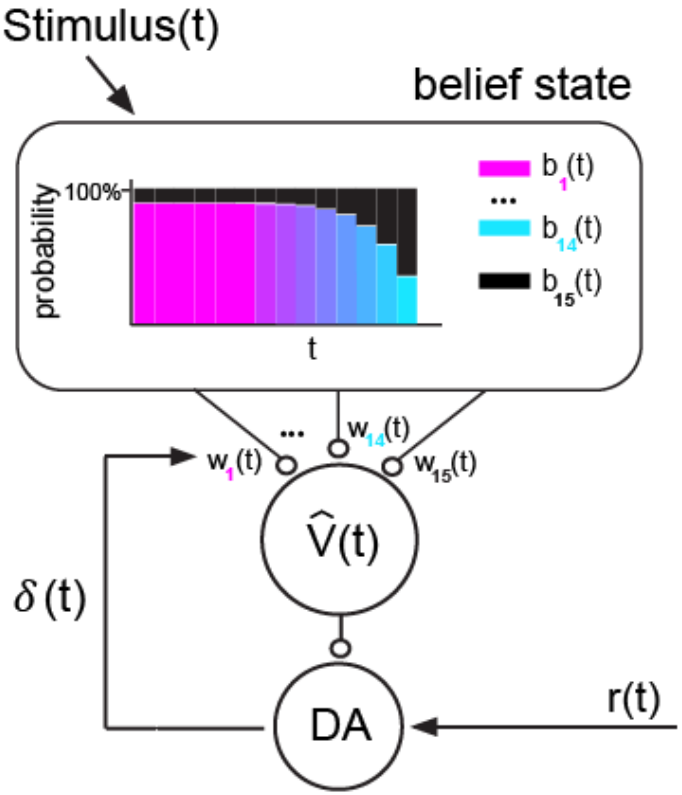
Summary

- Dopamine RPEs are shaped by
  - Interval timing (hazard rate)
  - Hidden state inference (belief state)

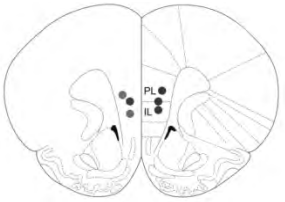


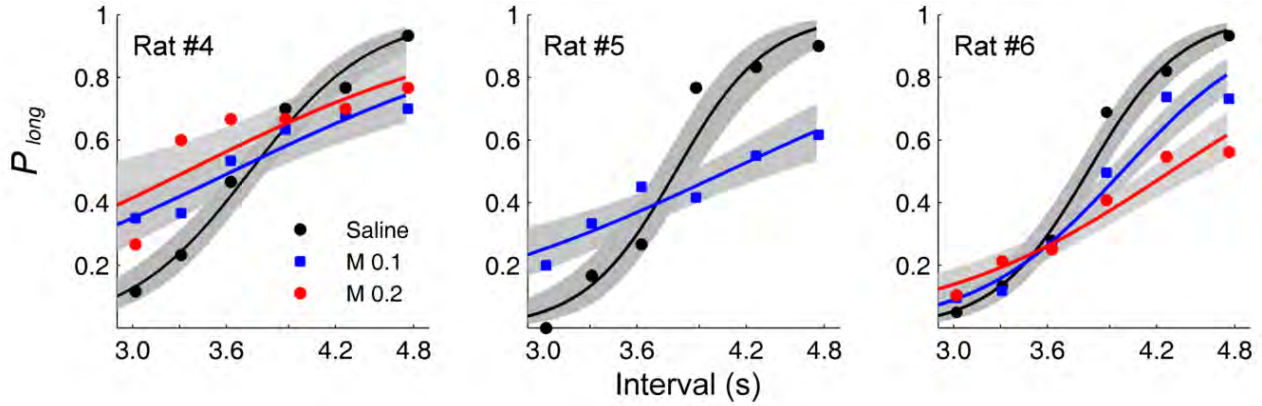
What brain area conveys belief state to dopamine system?

- Hidden state inference tied to timing
  - Candidate area: mPFC
    - Kim et al, 2009; 2013
    - Xu et al, 2014

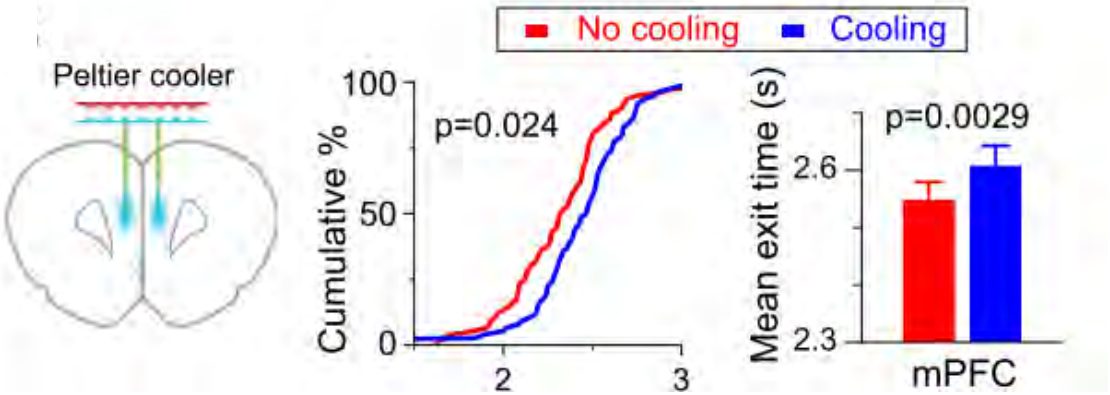


mPFC inactivation→impairment of interval timing behavior



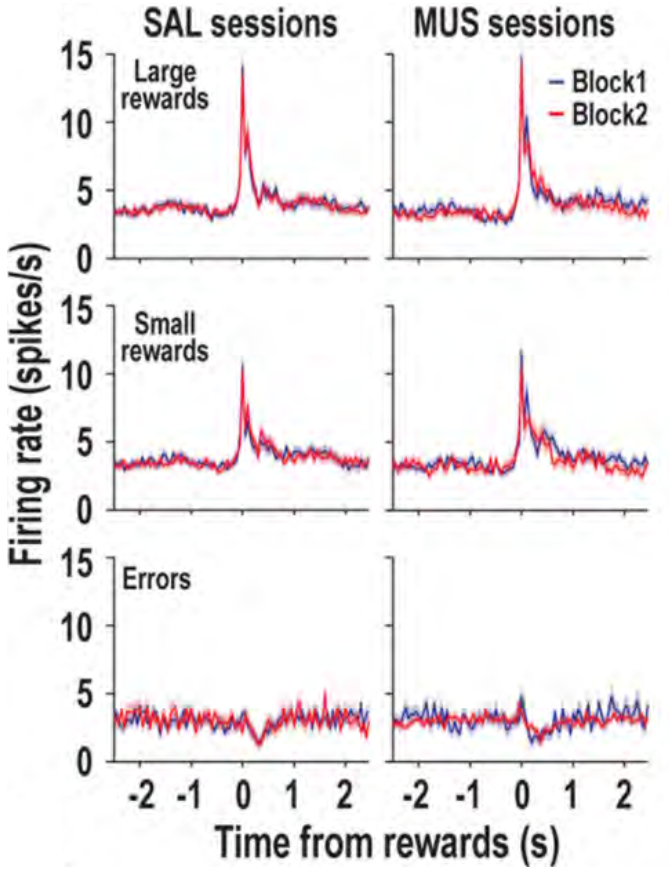


Kim et al, 2009

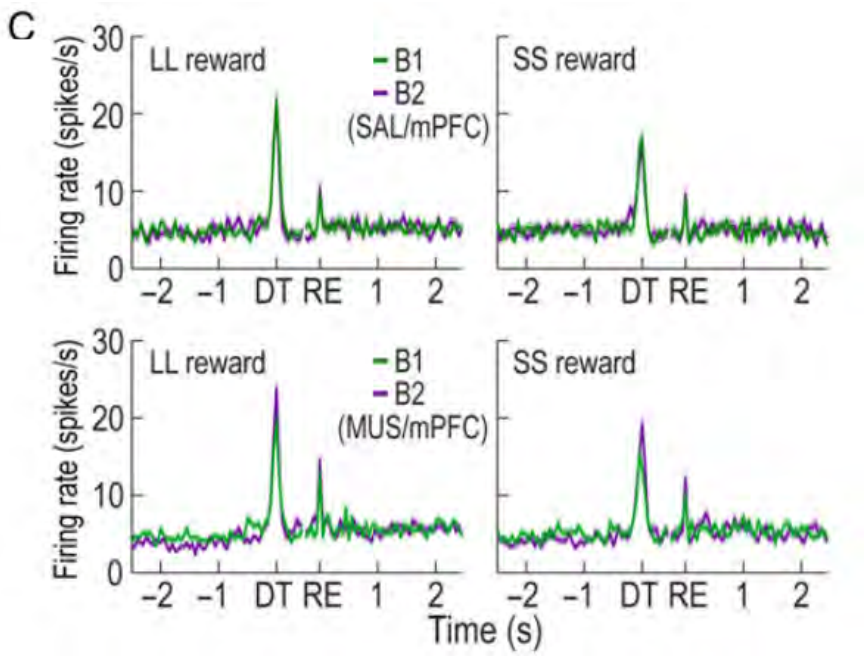


Xu et al, 2014

mPFC inactivation → mild effect on dopamine RPEs



Jo and Mizumori, 2013

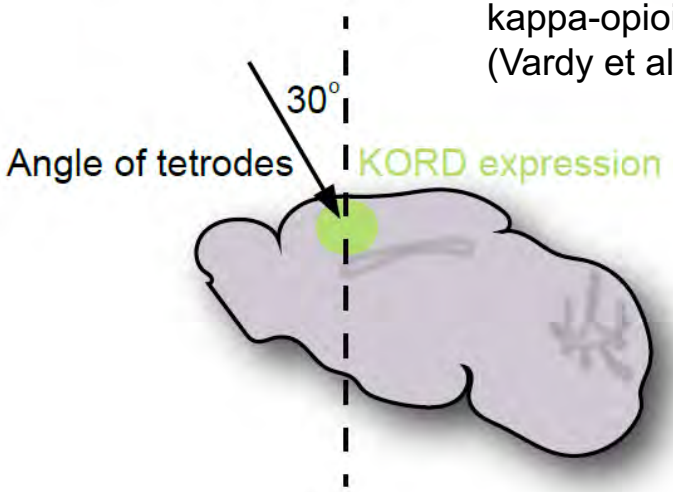


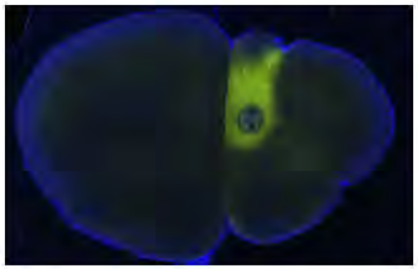
Jo and Mizumori, 2015

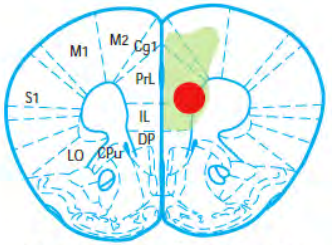
Chemogenetic inactivation of mPFC

Tool 3

KORD:
kappa-opioid receptor-based DREADD
(Vardy et al., 2016)

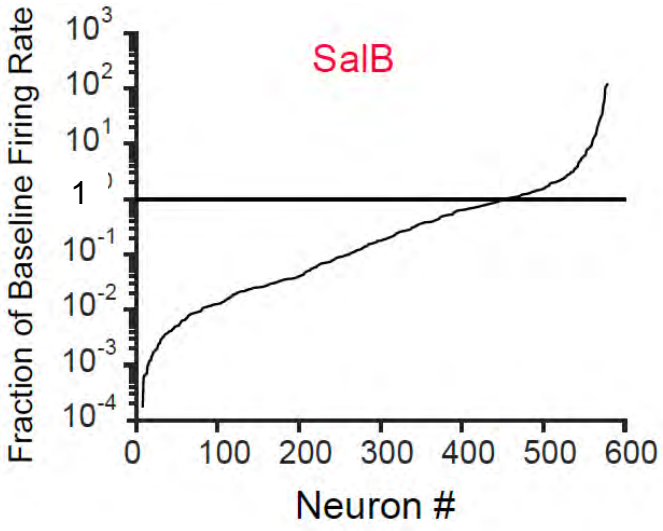


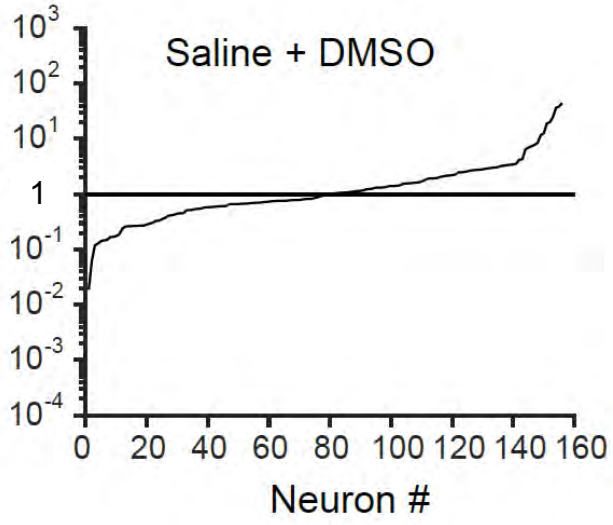




Bregma 1.94 mm

1 mm

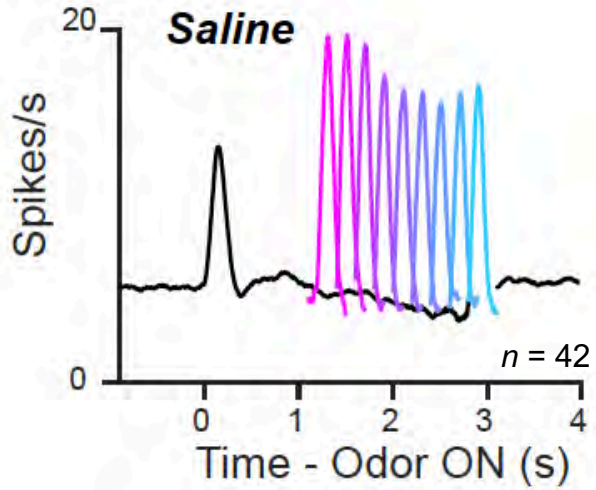


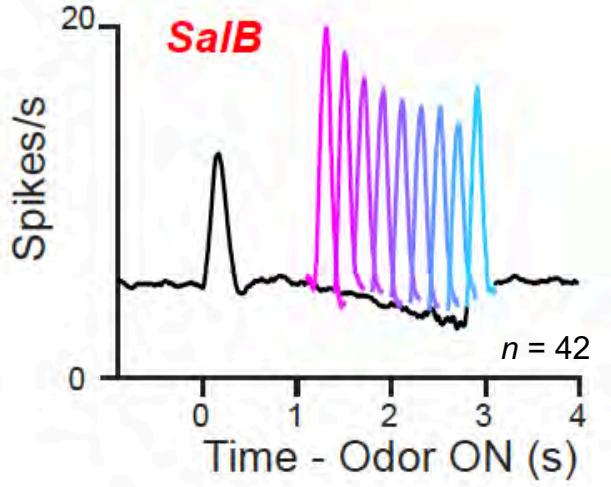


(Starkweather, Gerhsman, Uchida, unpublished)

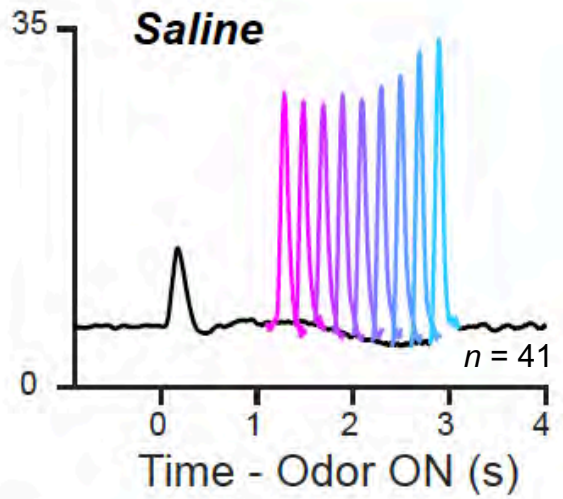
mPFC inactivation

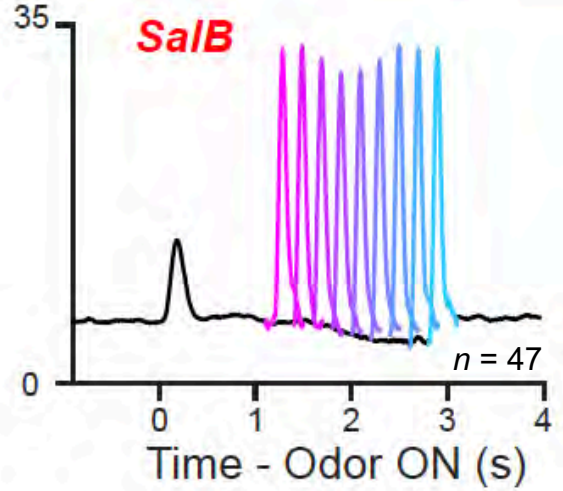
Task 1 (100% reward)





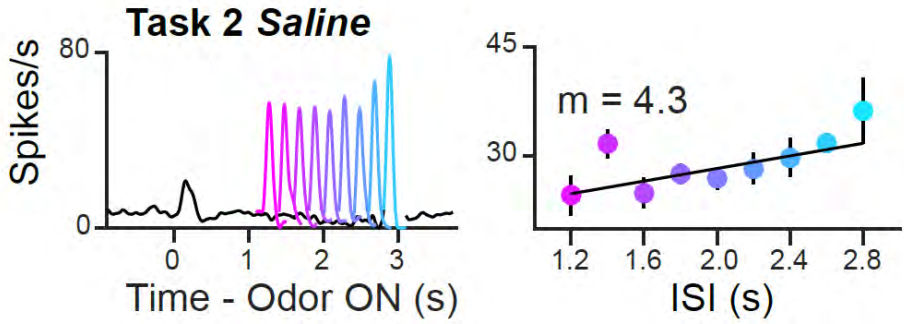
Task 2 (90% reward)



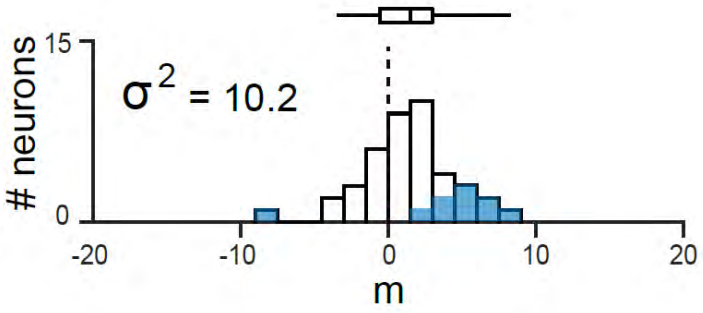


Changes of individual neurons in Task 2

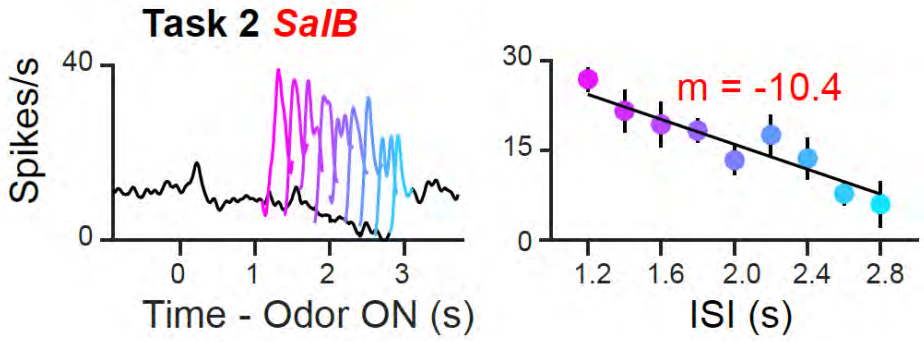
1 example neuron:



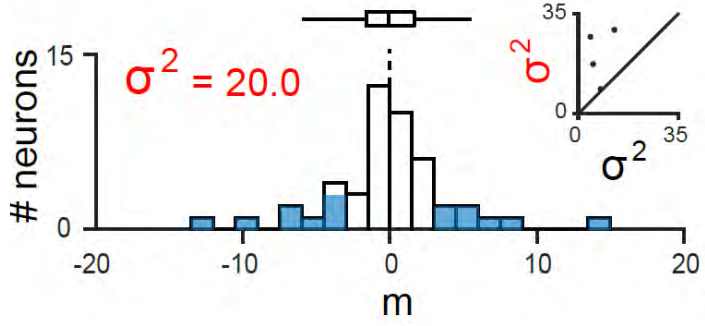
Saline



1 example neuron:



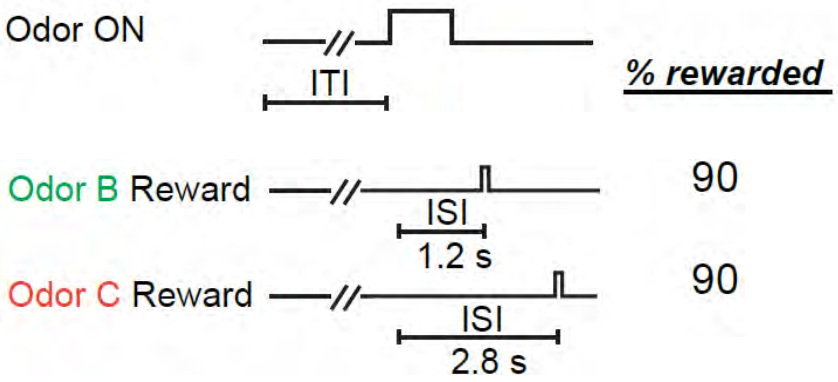
SalB



(Starkweather, Gerhsman, Uchida, unpublished)

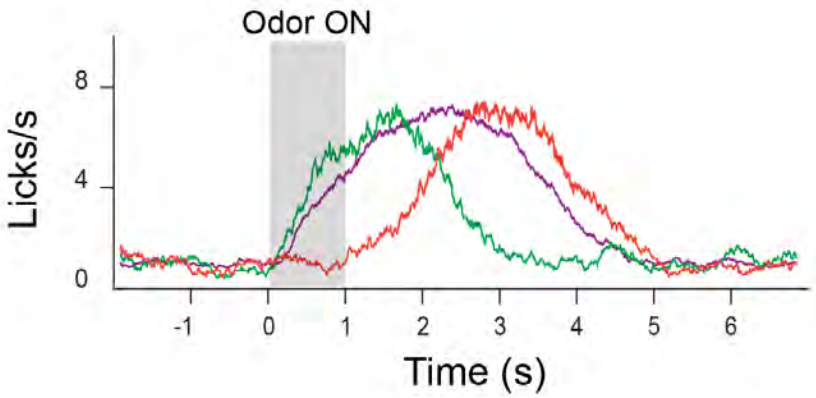
mPFC inactivation does not affect the time course of anticipatory licking

Task 2: 90% Rewarded

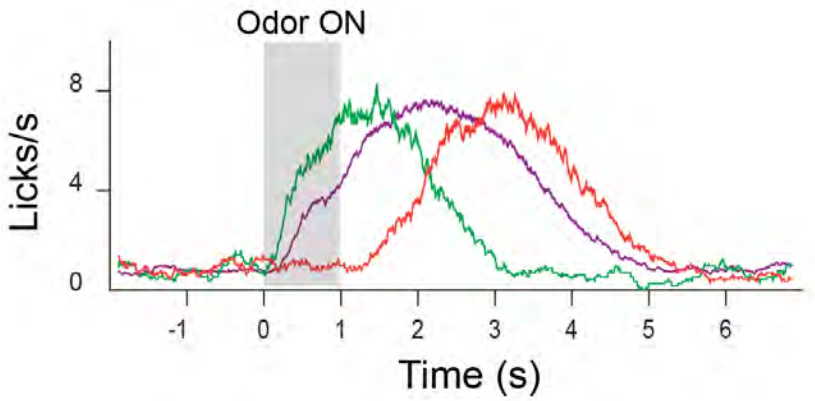


*Licking behavior on 10% **reward omission** trials

Saline

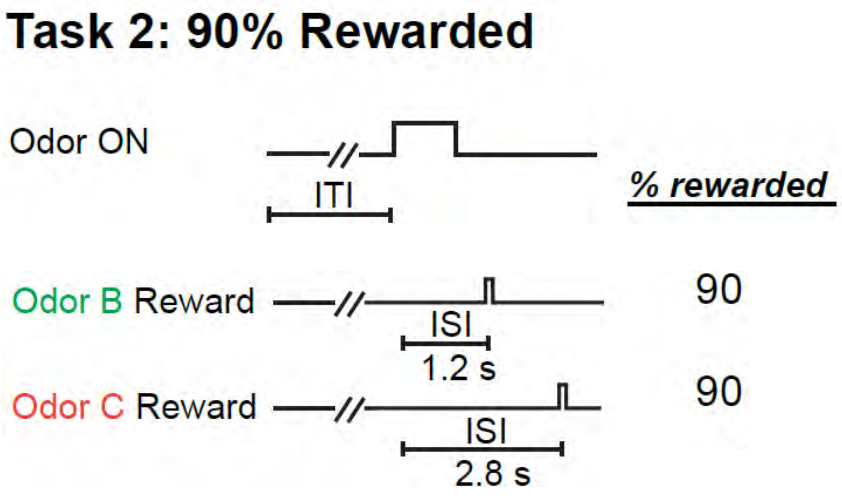


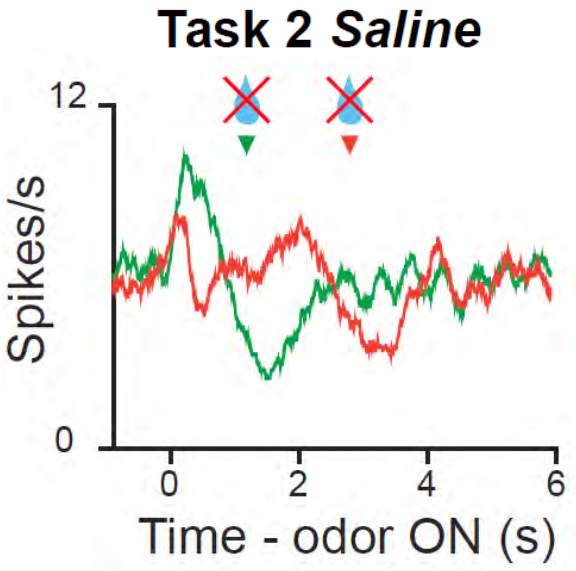
SalB

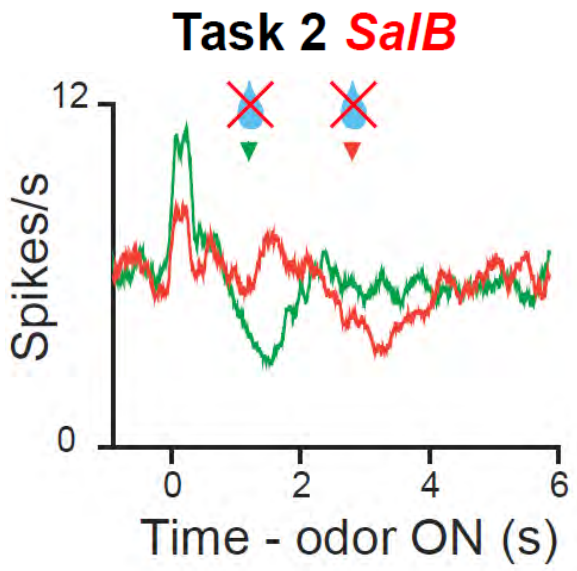


(Starkweather, Gerhsman, Uchida, unpublished)

mPFC inactivation does not impair timing-related aspects of RPE







(Starkweather, Gerhsman, Uchida, unpublished)

Model-based predictions

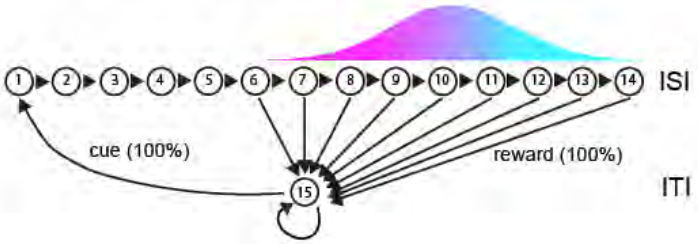
What changes in the model recapitulate the observations?

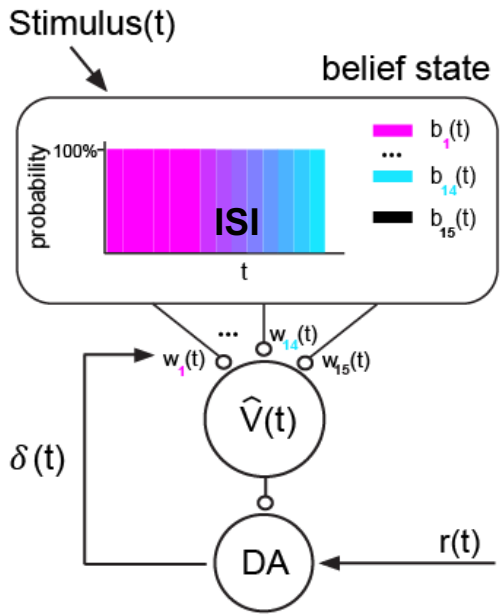
(Starkweather, Gerhsman, Uchida, unpublished)

Simulating specific impairments in the model

Task 1

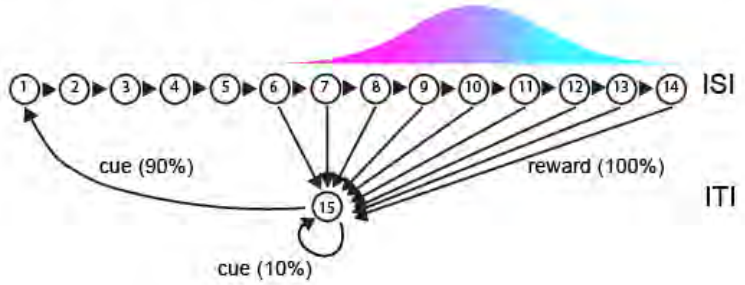
100% Rewarded

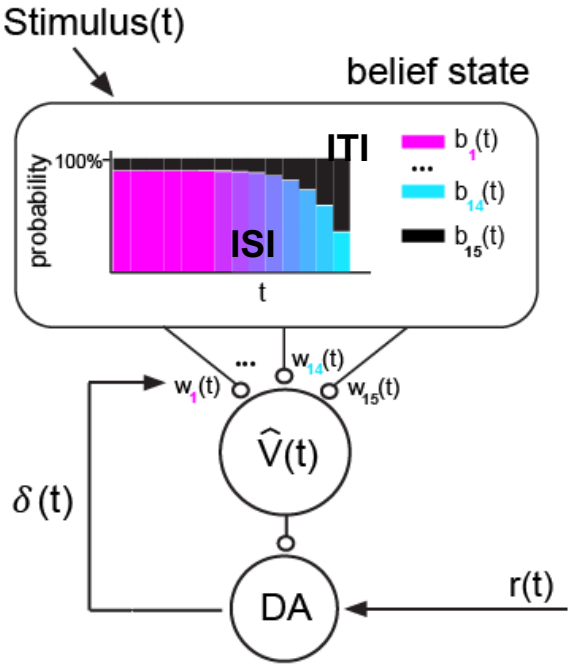




Task 2

90% Rewarded

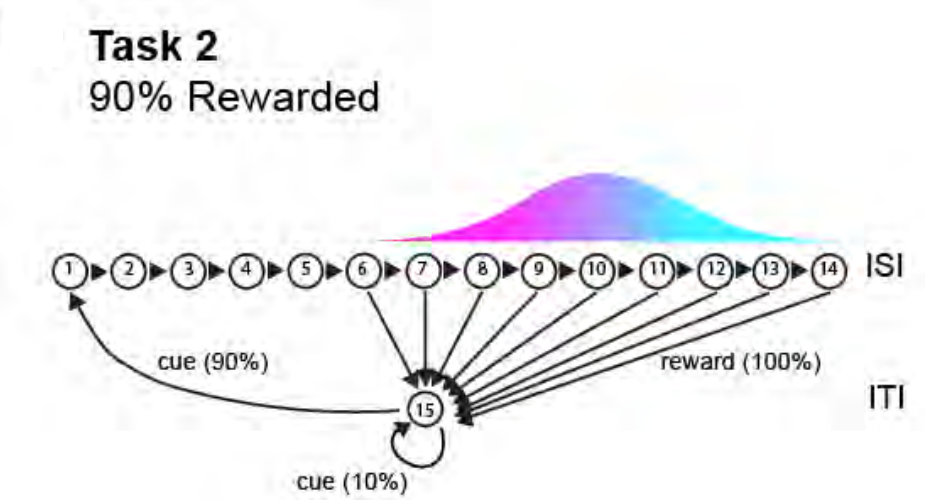


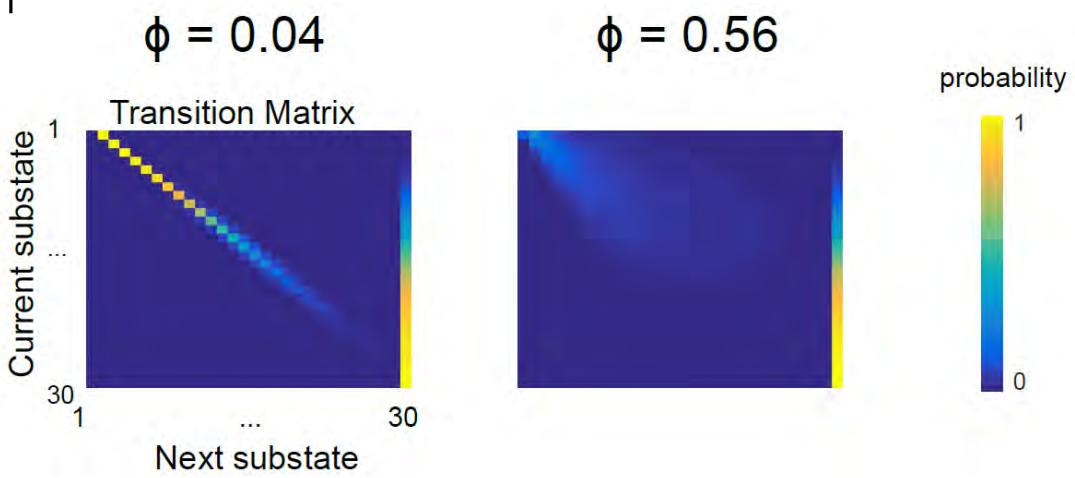


(Starkweather, Gerhsman, Uchida, unpublished)

Modeling mPFC inactivation as a timing deficit

- Transition matrix stores knowledge of dwell time
  - Timing 'lesion' simulated by blurring transition matrix (Takahashi et al, 2016)
  - Ways to simulate timing deficit:
    - Increase Weber fraction

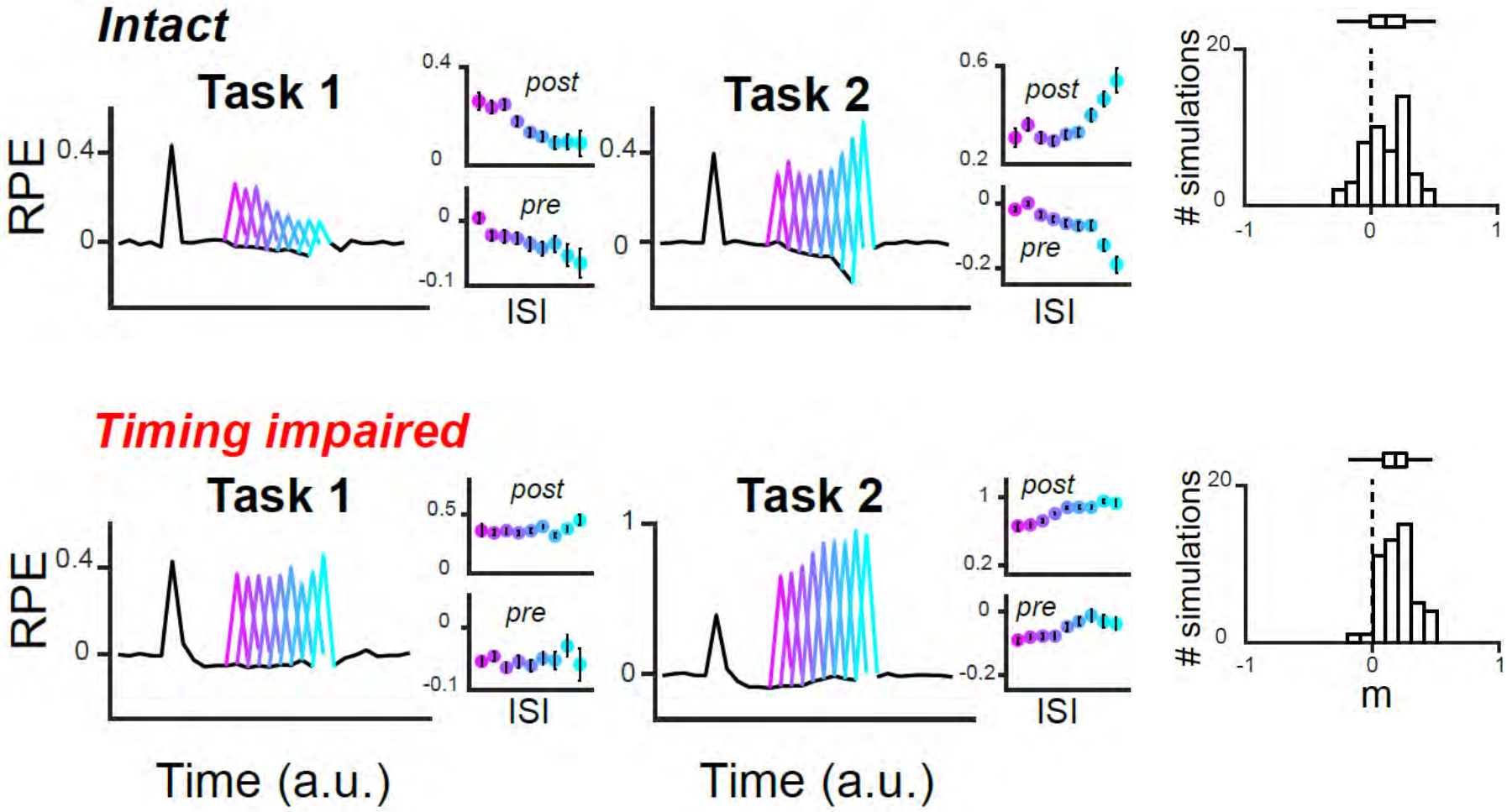




(Starkweather, Gerhsman, Uchida, unpublished)

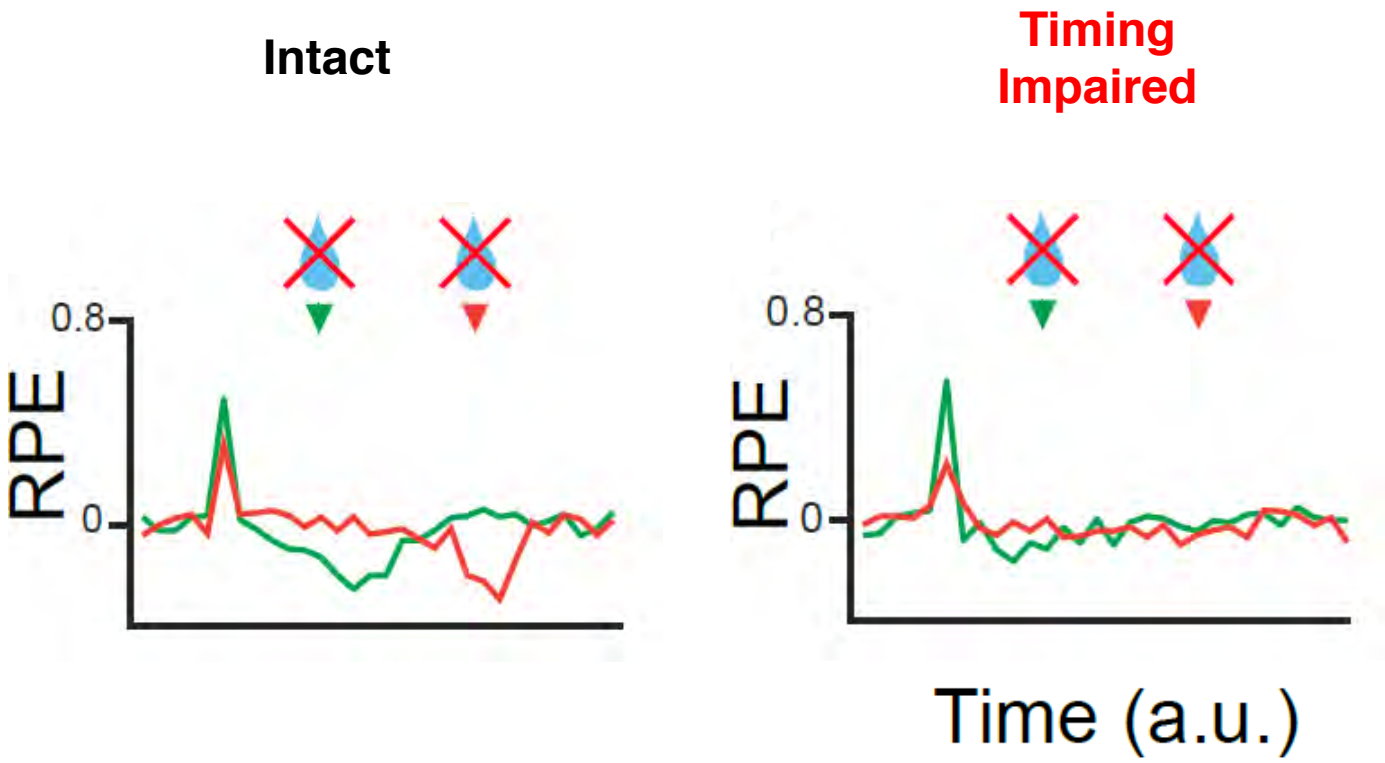
Modeling mPFC inactivation as a timing deficit

Model:



(Starkweather, Gerhsman, Uchida, unpublished)

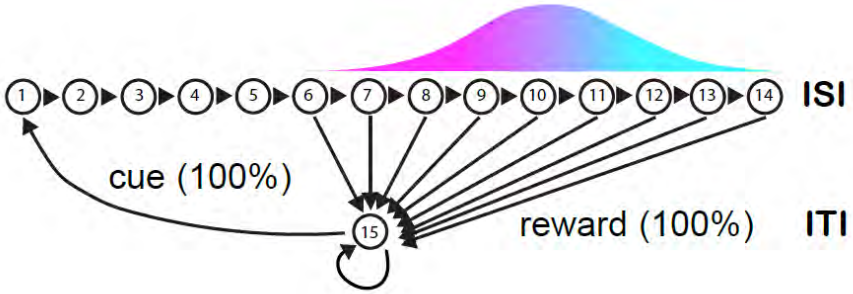
Modeling mPFC inactivation as a timing deficit



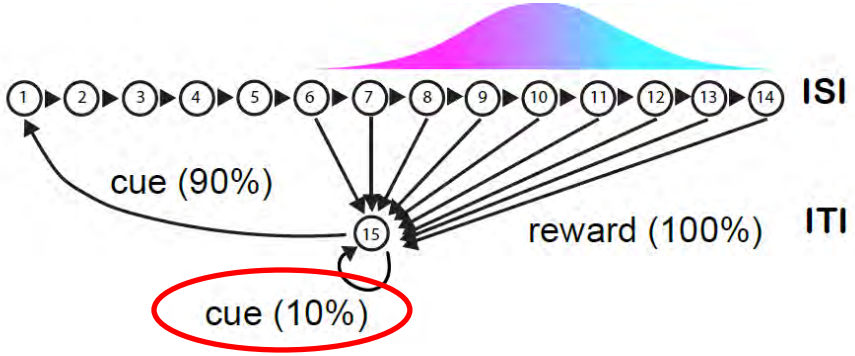
(Starkweather, Gerhsman, Uchida, unpublished)

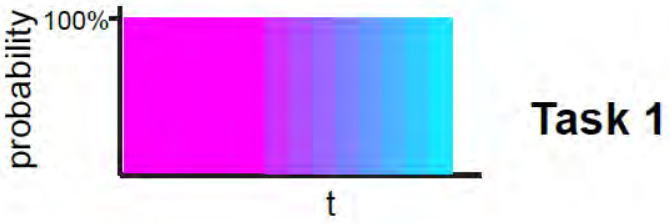
Modeling mPFC inactivation: Hidden state inference deficit

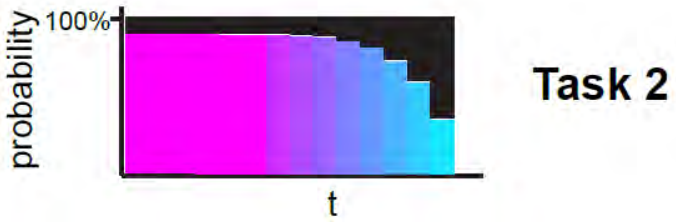
100% Rewarded



90% Rewarded



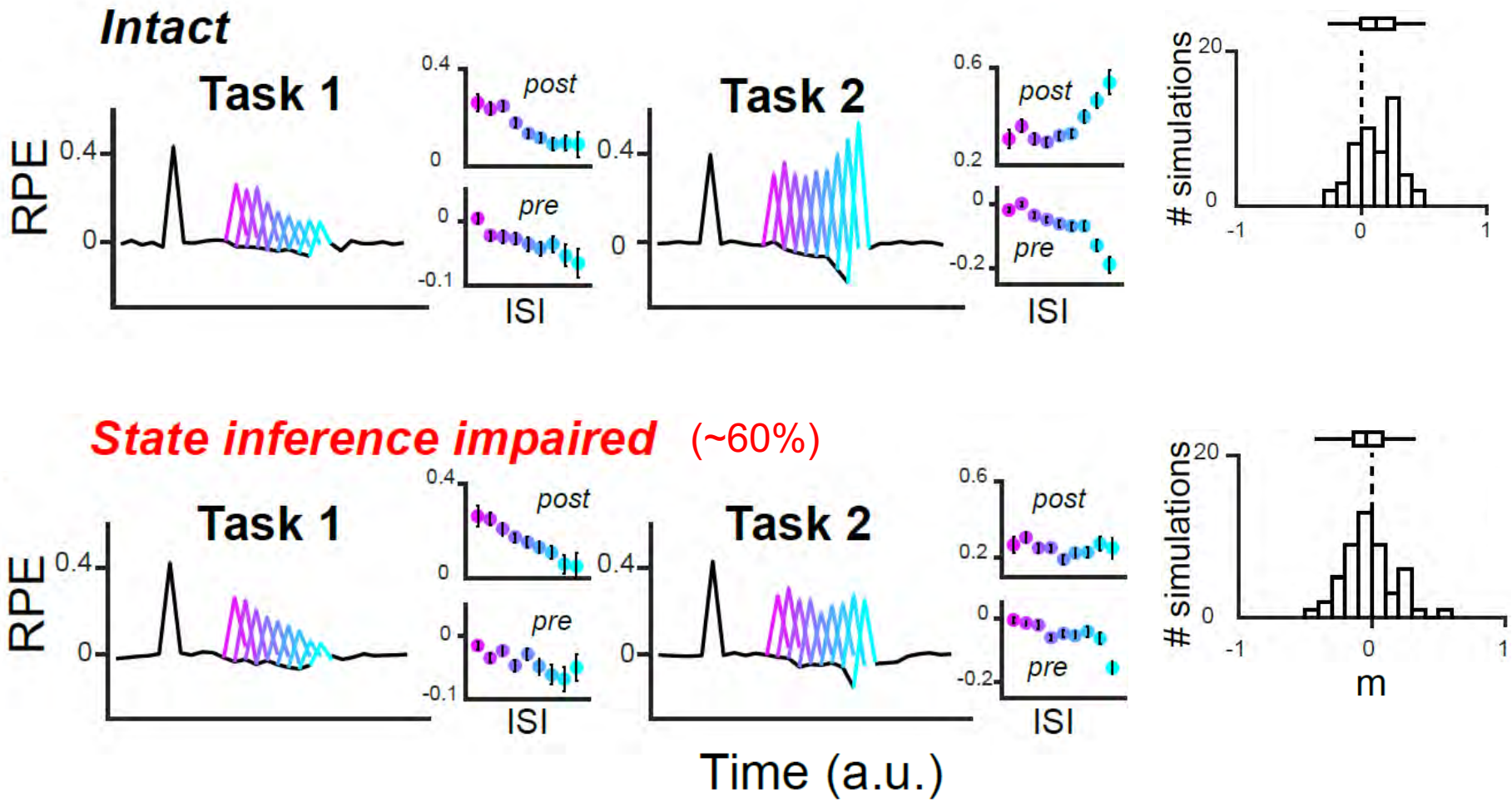




(Starkweather, Gerhsman, Uchida, unpublished)

Modeling mPFC inactivation as a hidden state inference deficit

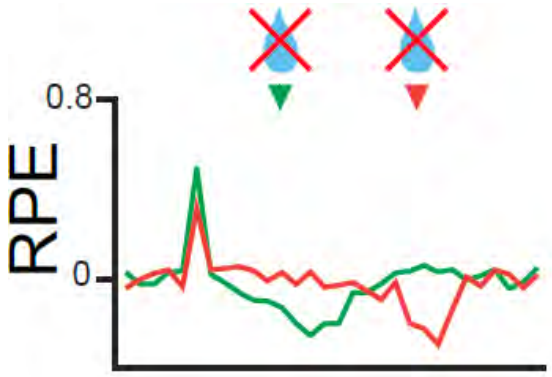
Model:



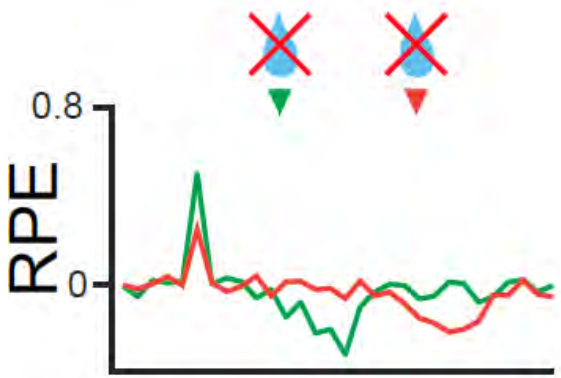
(Starkweather, Gerhsman, Uchida, unpublished)

Modeling mPFC inactivation as a hidden state inference deficit

Intact



State inference impaired



Time (a.u.)

(Starkweather, Gerhsman, Uchida, unpublished)

Summary

- Dopamine RPEs are shaped by
  - Interval timing (hazard rate)
  - Hidden state inference (belief state)

- A TD model in which reward expectation is computed over belief states uniquely explains dopamine responses.

- Processes of interval timing and hidden state inference can be separated experimentally.

Dopamine RPE and inference

J Neurophysiol 104: 1068–1076, 2010.
First published June 10, 2010; doi:10.1152/jn.00158.2010.

A Pallidus-Habenula-Dopamine Pathway Signals Inferred Stimulus Values

Ethan S. Bromberg-Martin,[1] Masayuki Matsumoto,[1,2] Simon Hong,[1] and Okihide Hikosaka[1]
[1]*Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, Maryland; and* [2]*Primate Research Institute, Kyoto University, Inuyama, Aichi, Japan*

CellPress

Neuron
Article

Temporal Specificity of Reward Prediction Errors Signaled by Putative Dopamine Neurons in Rat VTA Depends on Ventral Striatum

Yuji K. Takahashi,[1,5,*] Angela J. Langdon,[2,5] Yael Niv,[2] and Geoffrey Schoenbaum[1,3,4,*]

Current Biology
Article

CellPress

Midbrain Dopamine Neurons Signal Belief in Choice Accuracy during a Perceptual Decision

Armin Lak,[1,2] Kensaku Nomoto,[3,4] Mehdi Keramati,[5] Masamichi Sakagami,[3] and Adam Kepecs[1,6,*]

Topics

- A mouse model to study dopamine RPE
- Do all dopamine neurons signal RPEs?
- What is the “state” in reinforcement learning?
- How are RPEs computed?
- Diversity of dopamine neurons

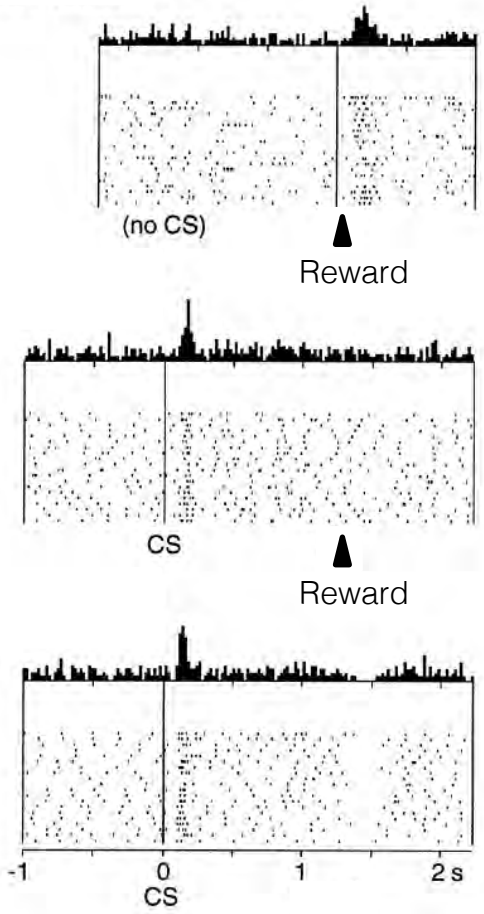
Firing of putative dopamine neurons

No prediction
Reward occurs

CS predicts reward
Reward occurs

(CS: conditioned stimulus)

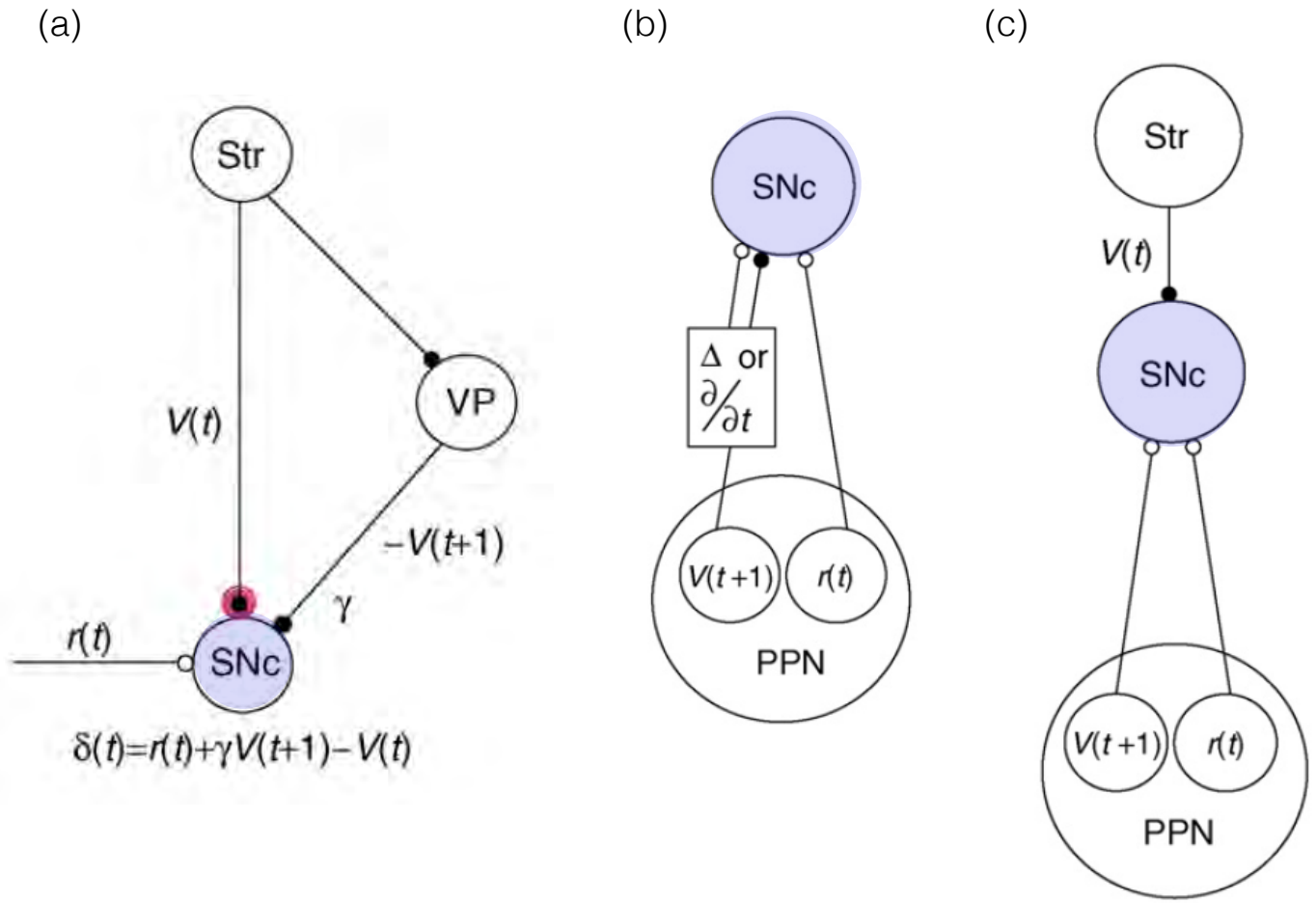
CS predicts reward
No reward occurs



$$\delta(t) = r(t) + \gamma \hat{V}(t+1) - \hat{V}(t)$$

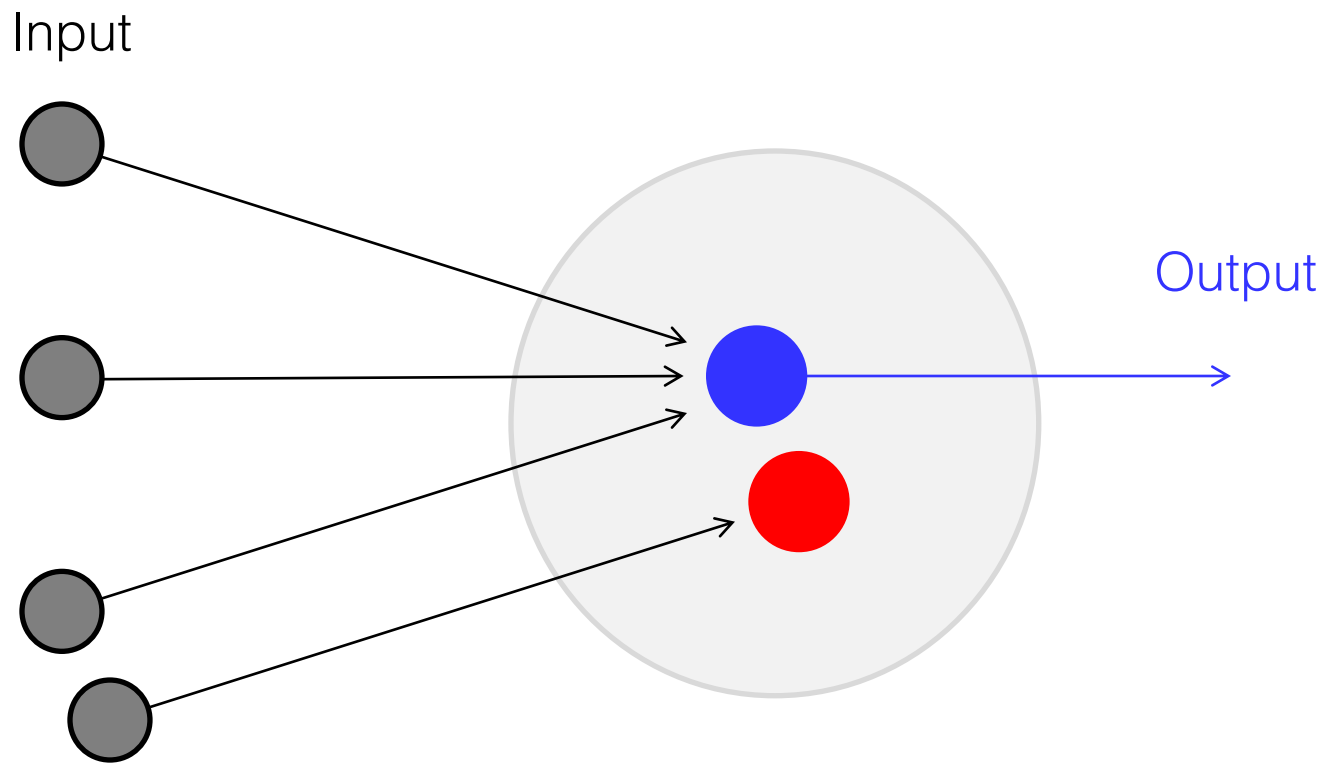
(Schultz, Dayan, Montague, 1997)

Models of RPE computation



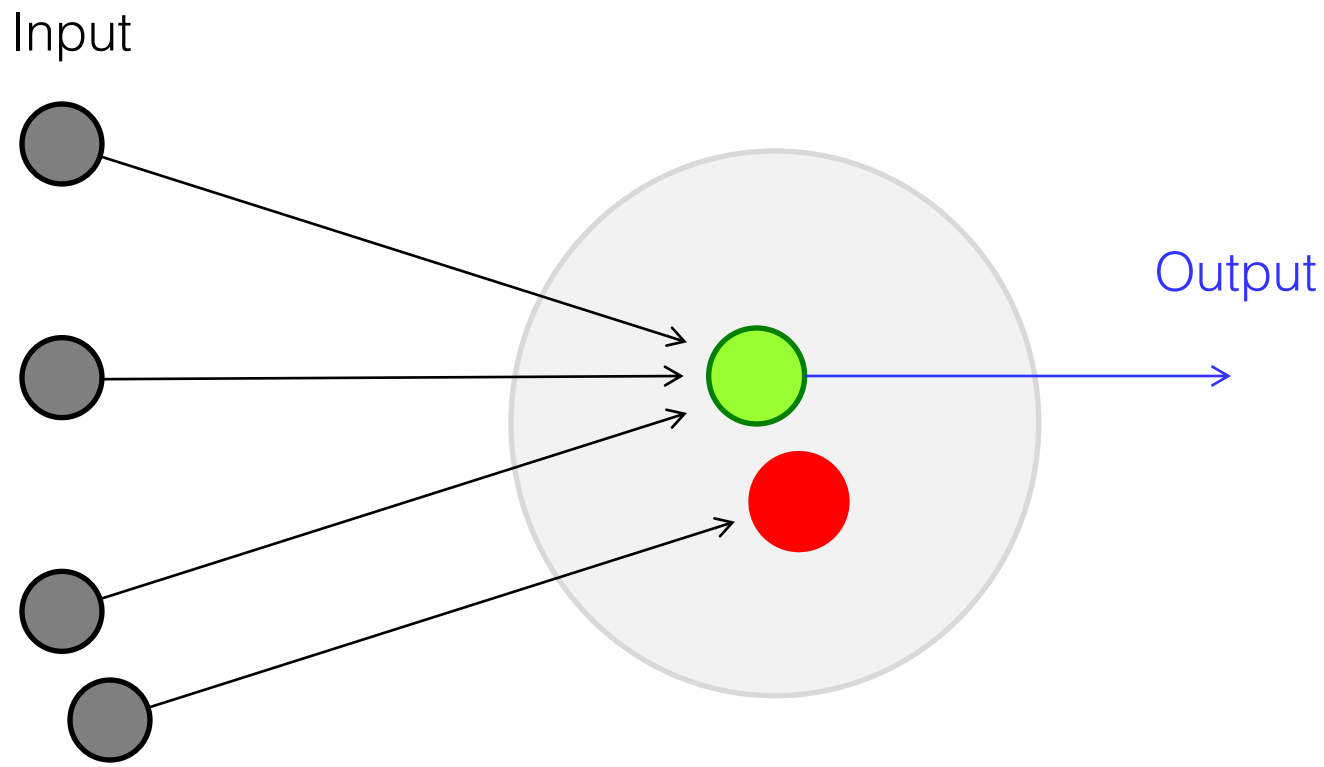
(Kawato and Samejima, 2007)

Where do other inputs come from?

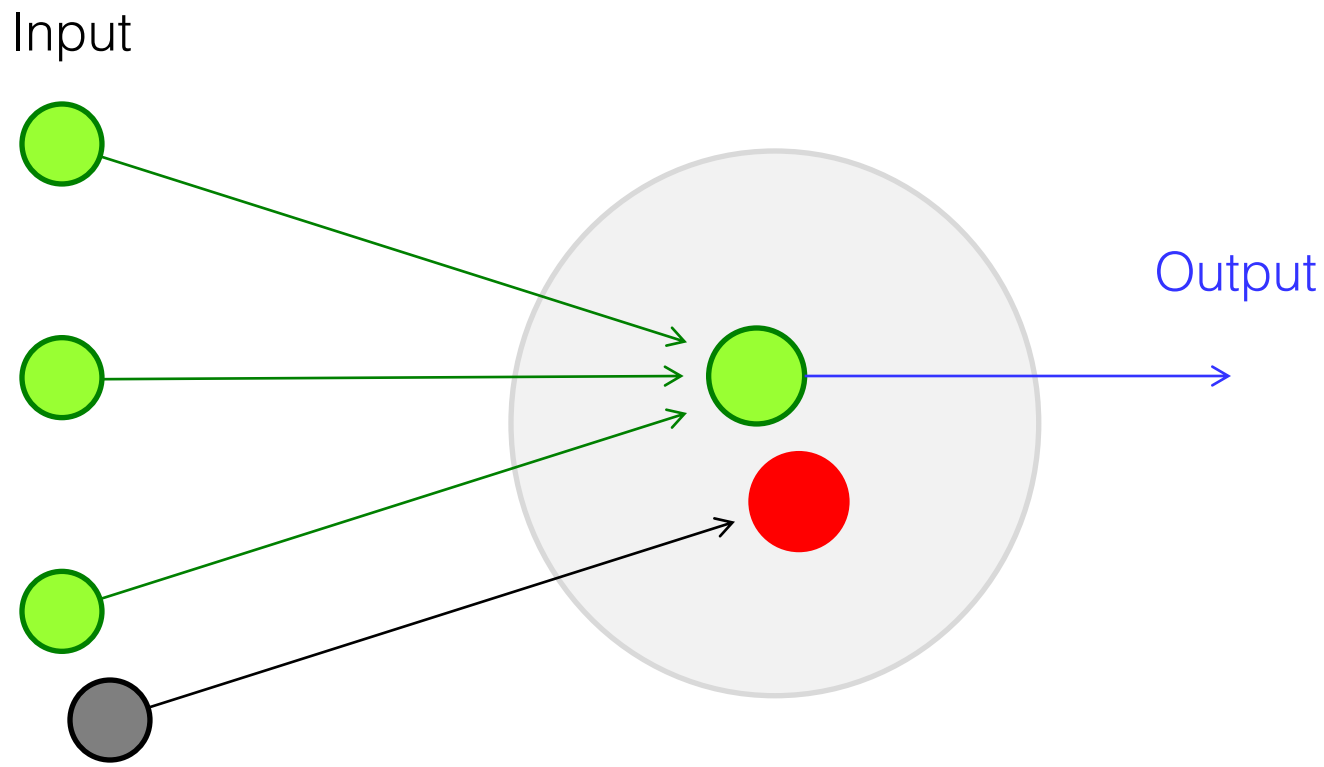


- Conventional tracers are taken up non-specifically…

Where do other inputs come from?



Where do other inputs come from?



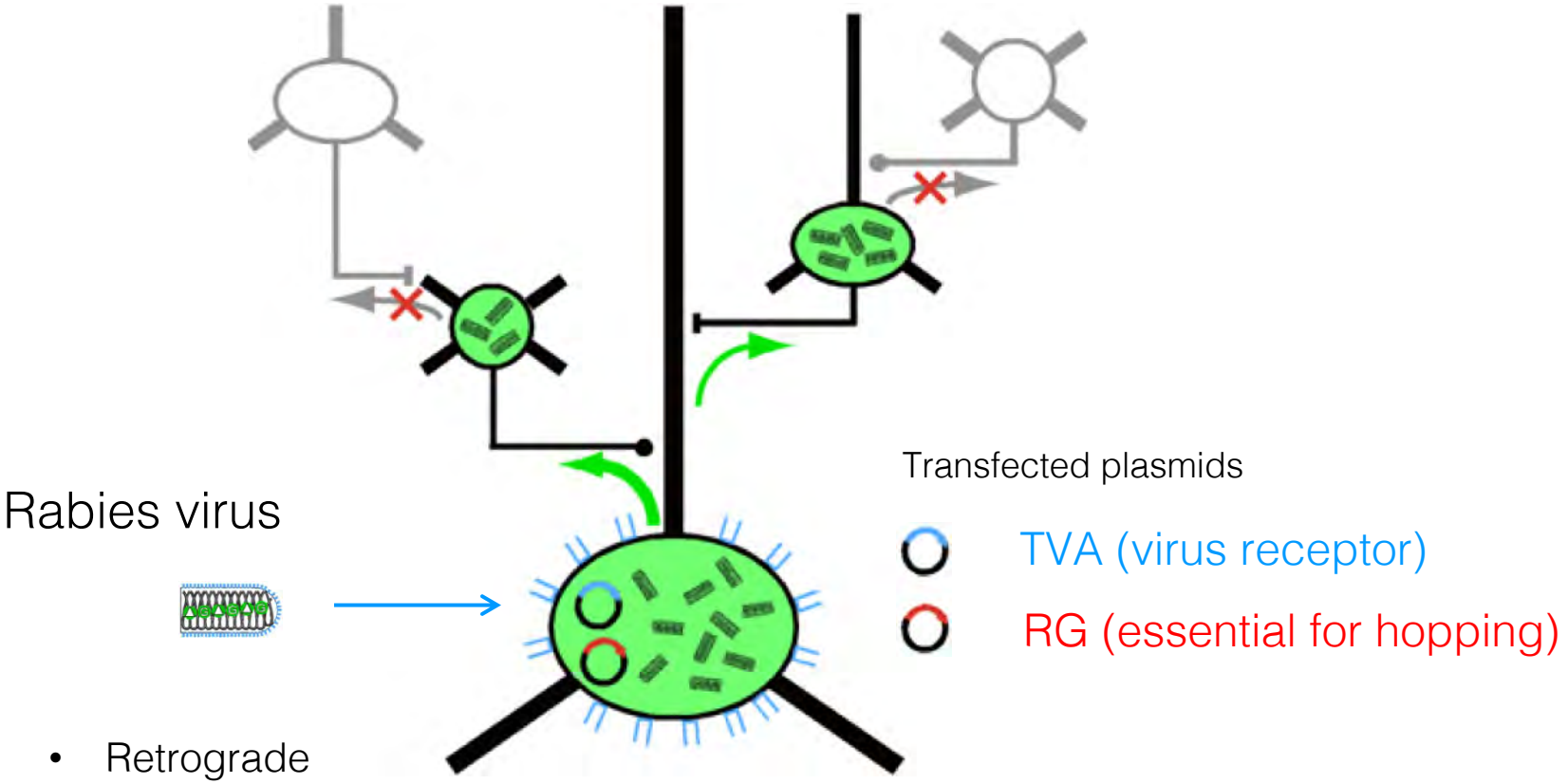
Neuron 53, 639–647, March 1, 2007

Tool 4

Monosynaptic Restriction of Transsynaptic Tracing from Single, Genetically Targeted Neurons

Ian R. Wickersham,[1,*] David C. Lyon,[1,4] Richard J.O. Barnard,[2,5] Takuma Mori,[1] Stefan Finke,[3,6] Karl-Klaus Conzelmann,[3] John A.T. Young,[2] and Edward M. Callaway[1]

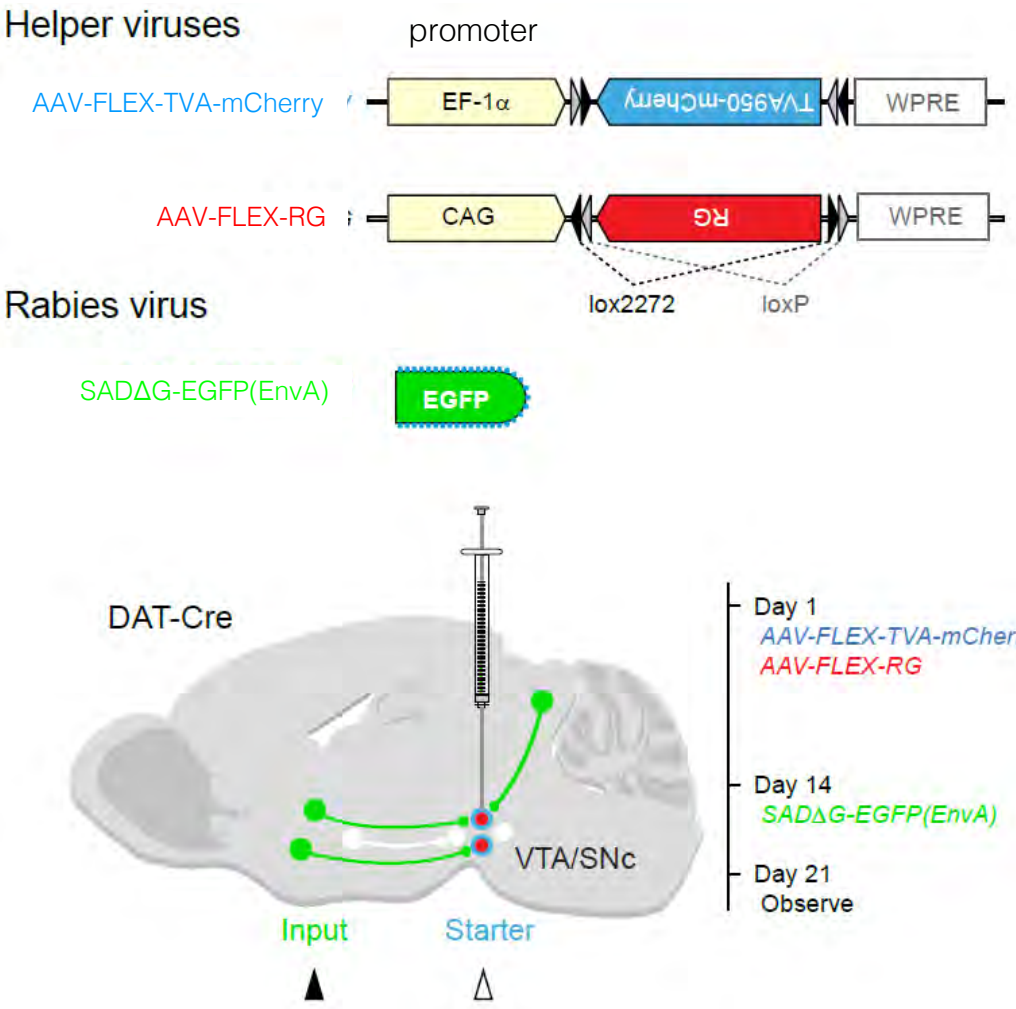
Rabies virus-mediated monosynaptic input tracing



- Retrograde
- Monosynaptic

Rabies virus-mediated input tracing

AAV: adeno-associated virus



TVA: Virus receptor

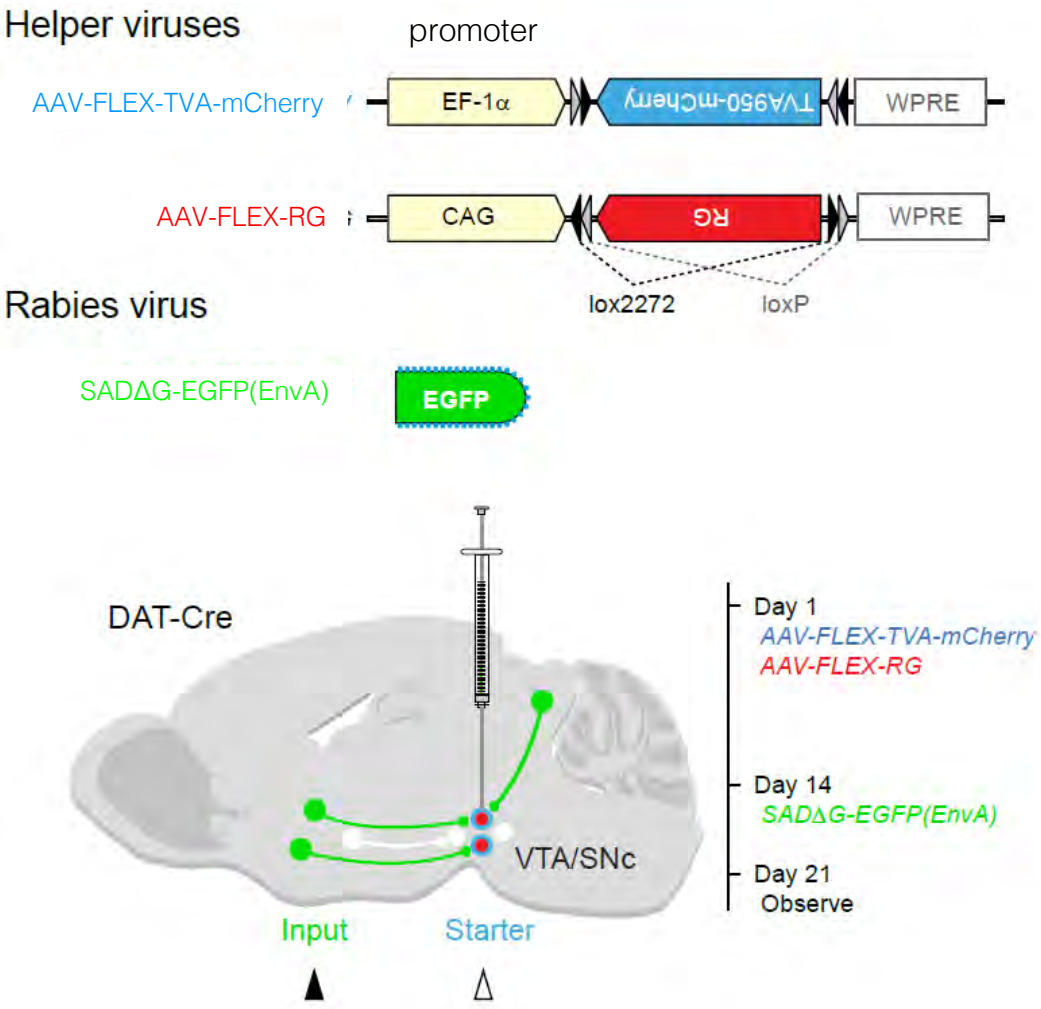
RG (Rabies glycoprotein):
essential for transsynaptic hopping

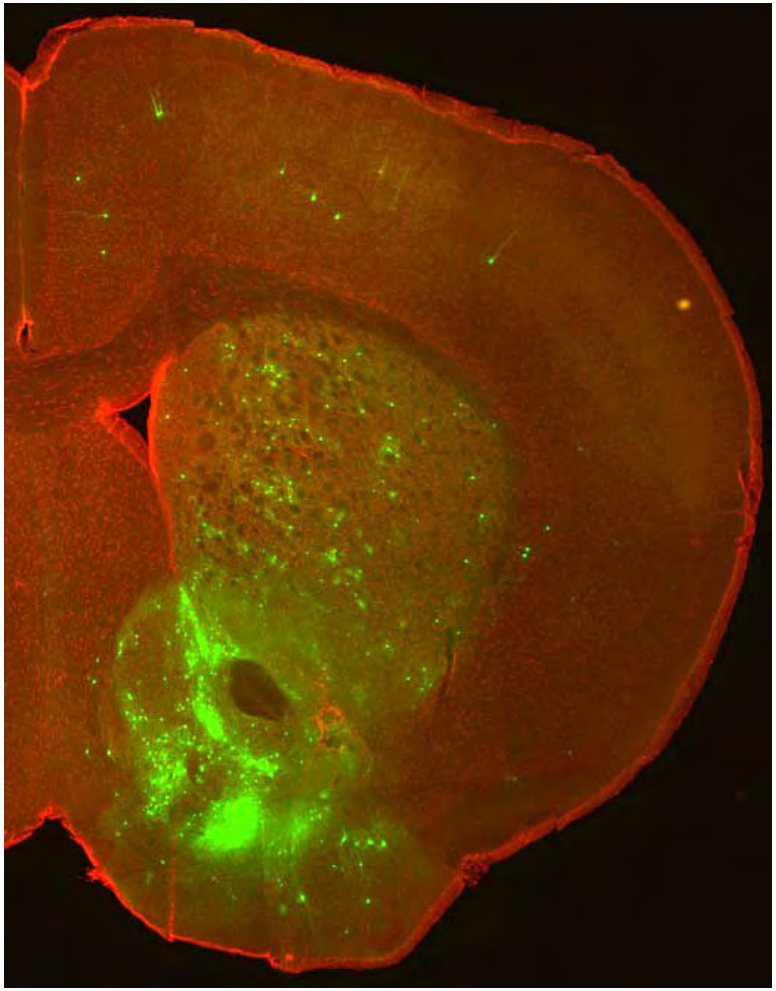
Miyamichi et al. 2010
Haubensak et al. 2010
Wall et al. 2010

(Watabe-Uchida et al., *Neuron*, 2012)

Rabies virus-mediated input tracing

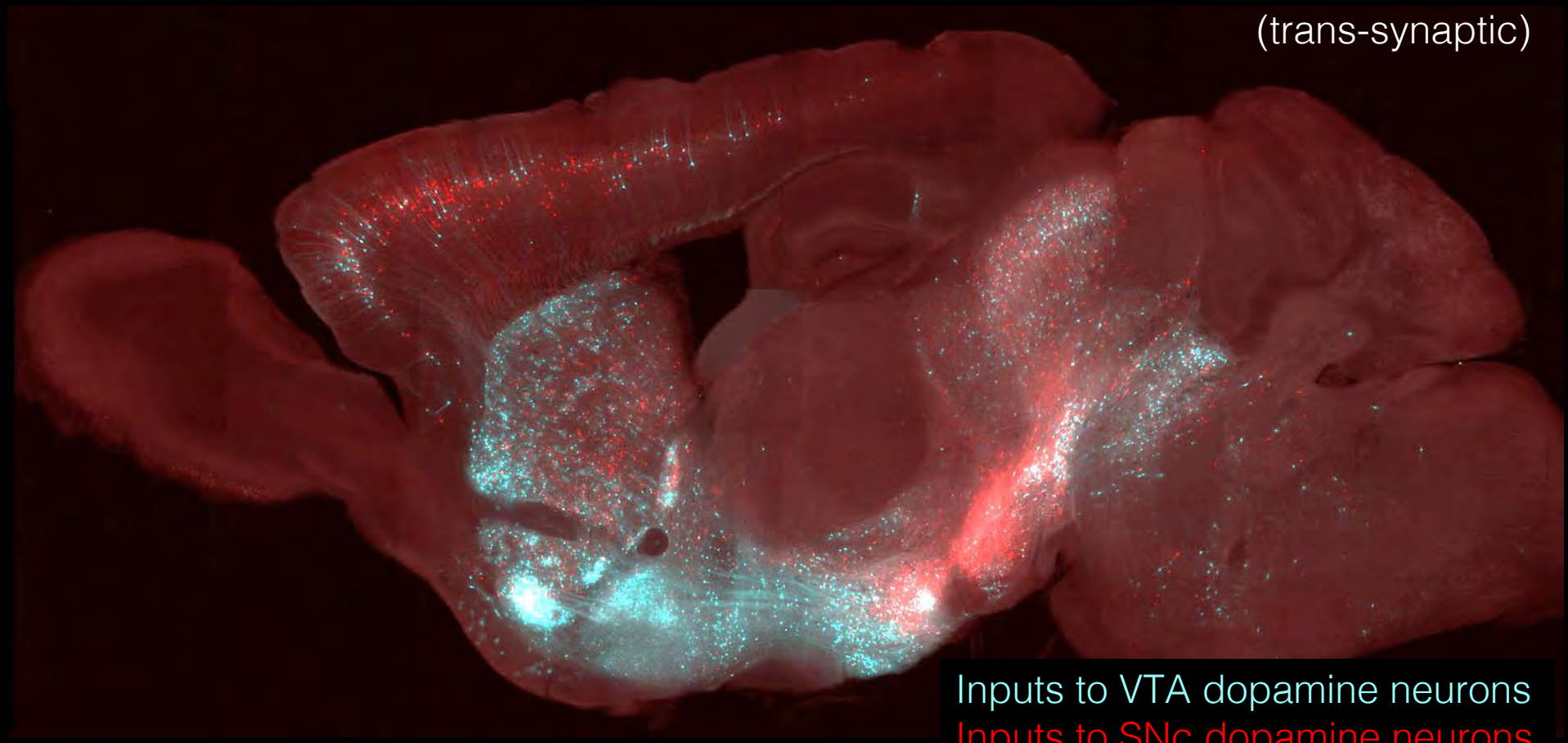
AAV: adeno-associated virus





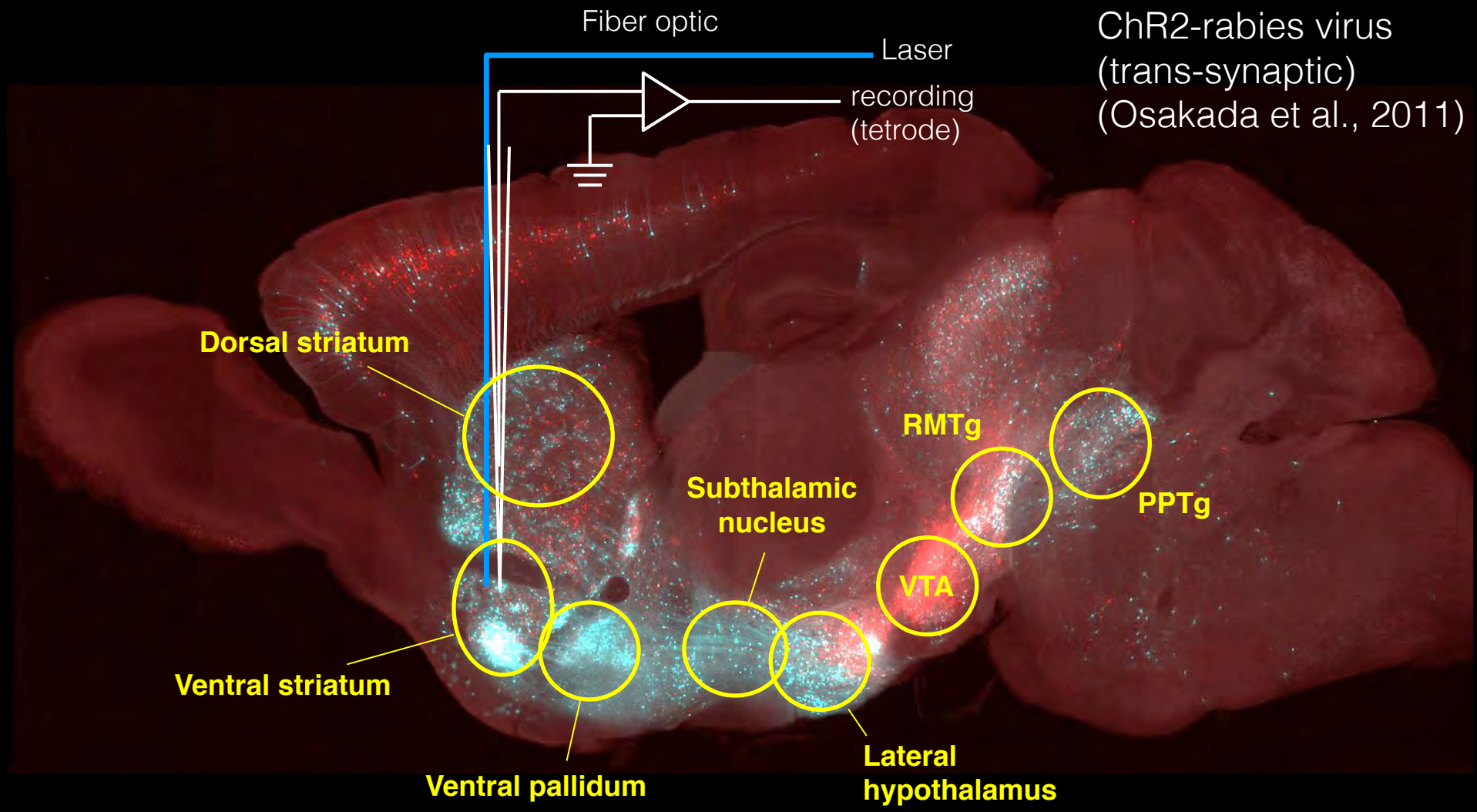
(Watabe-Uchida et al., *Neuron*, 2012)

Direct inputs to dopamine neurons



(Watabe-Uchida et al., *Neuron*, 2012)

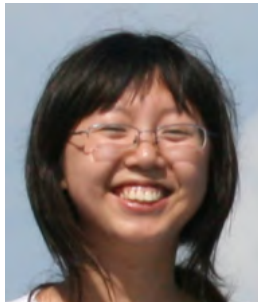
Recording from direct inputs of dopamine neurons



(Tian et al., *Neuron*, 2016)

PPTg: pedunculopontine nucleus
RMTg: rostromedial tegmental area

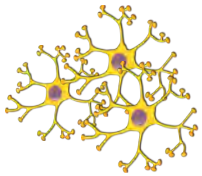
Recording from direct inputs to dopamine neurons



Ju Tian

Mitsuko Watabe-Uchida

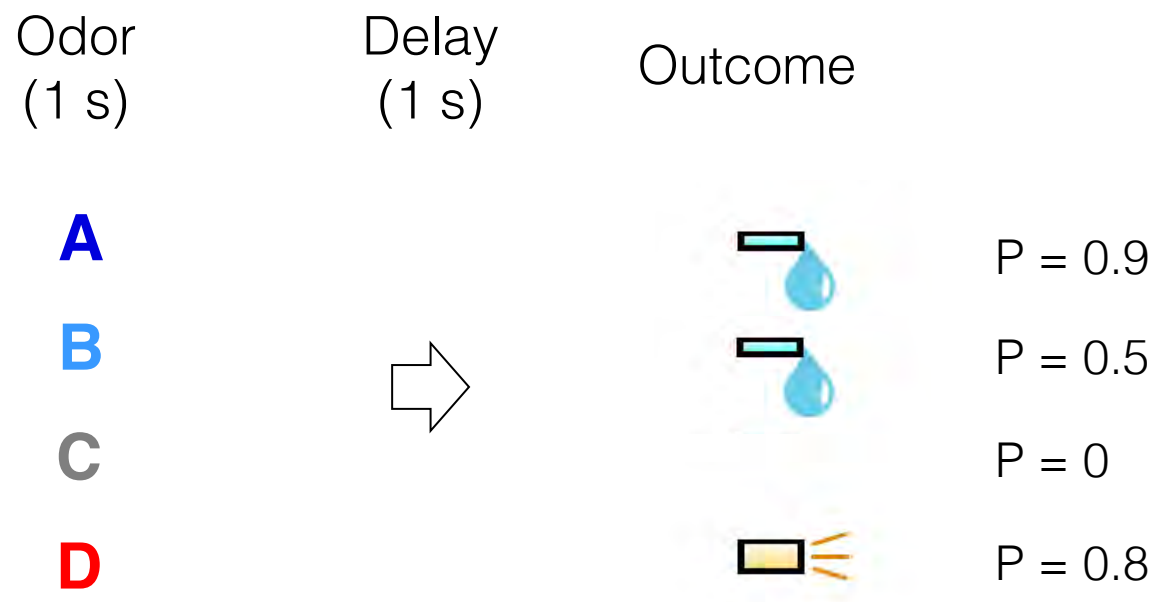
99 mice



1,931 neurons from 7 input areas

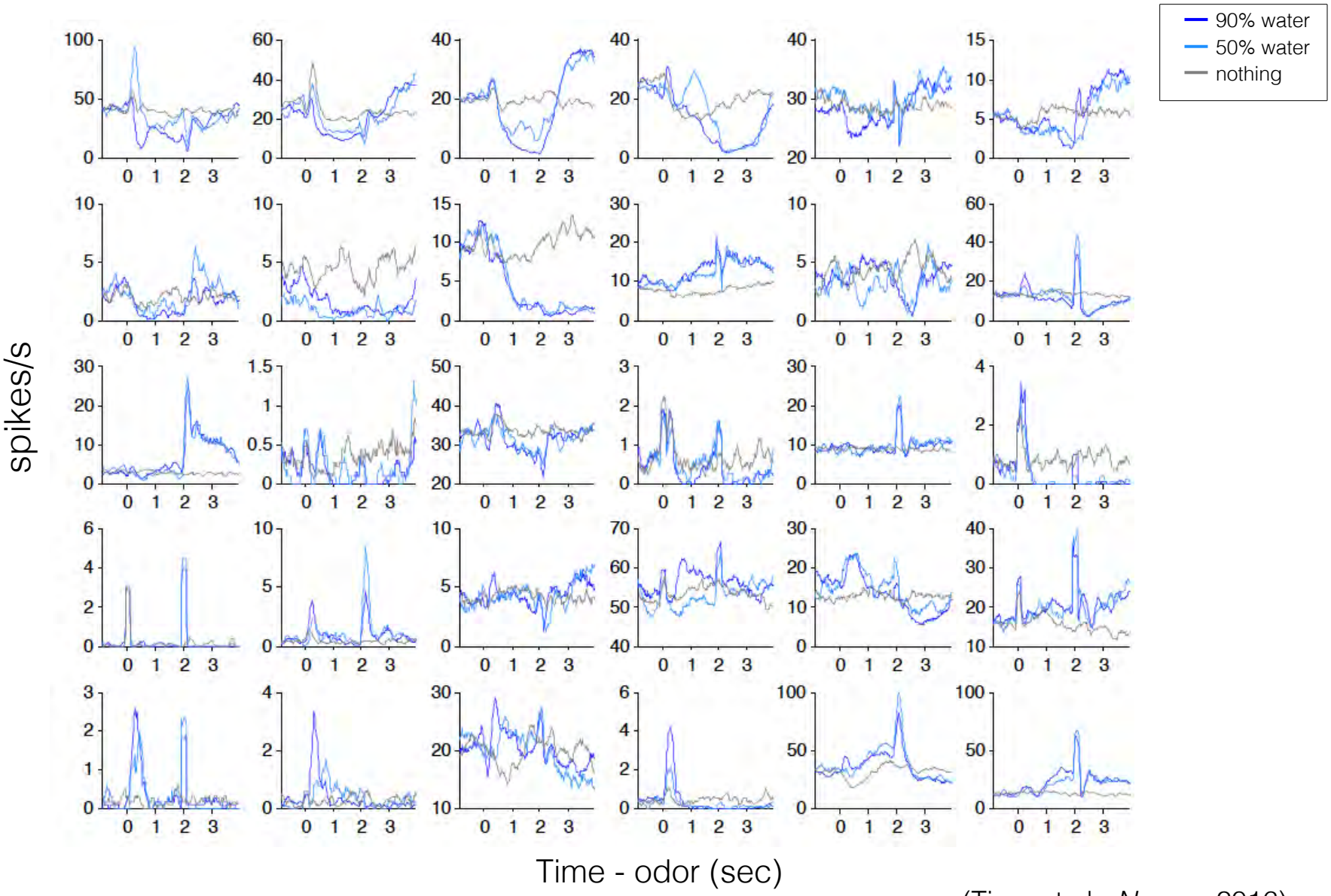
205 identified input neurons

(Tian, et al., *Neuron* 2016)



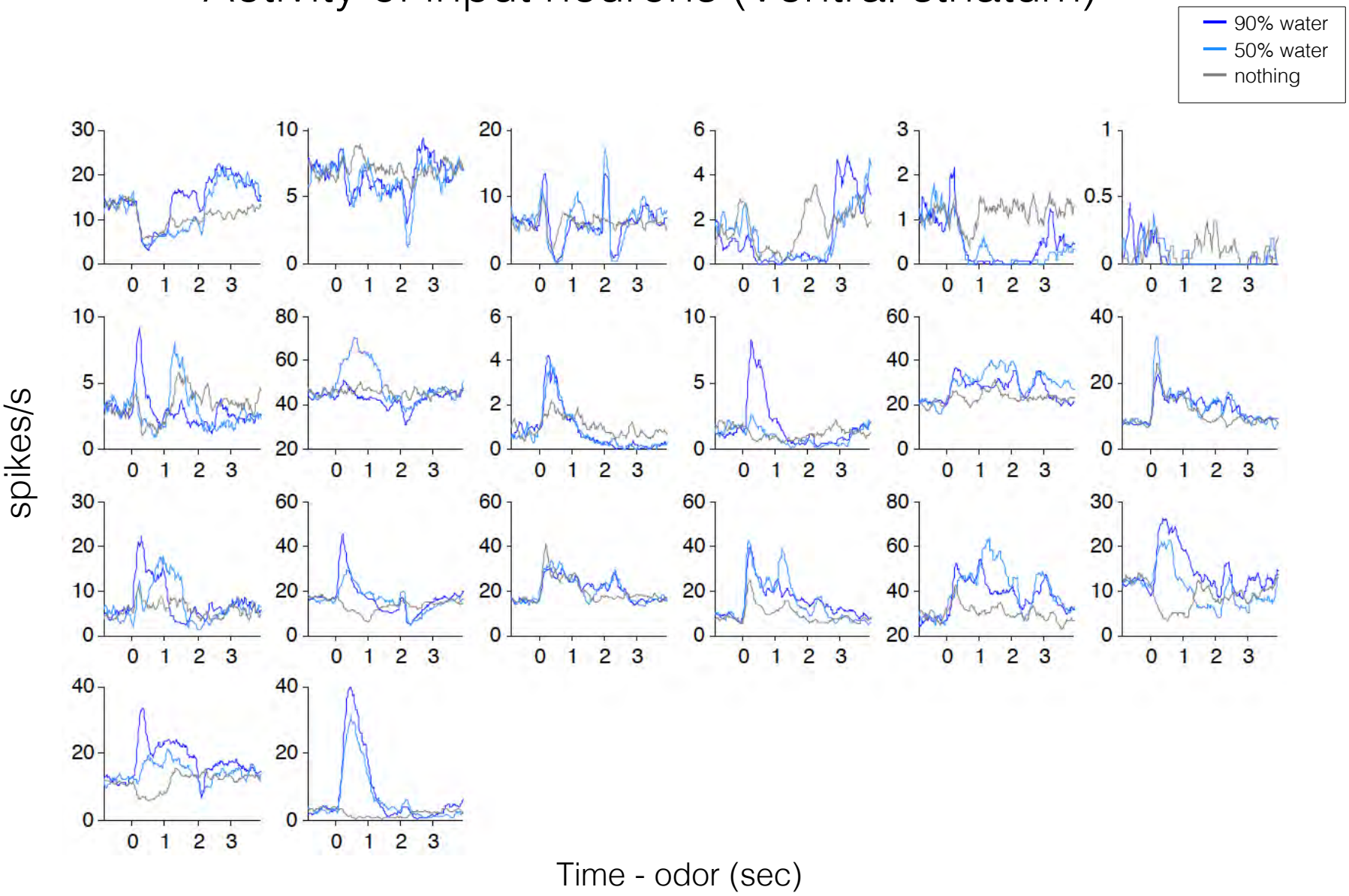
(Tian, et al., *Neuron* 2016)

Activity of input neurons (Pedunculopontine tegmental area)



(Tian, et al., *Neuron* 2016)

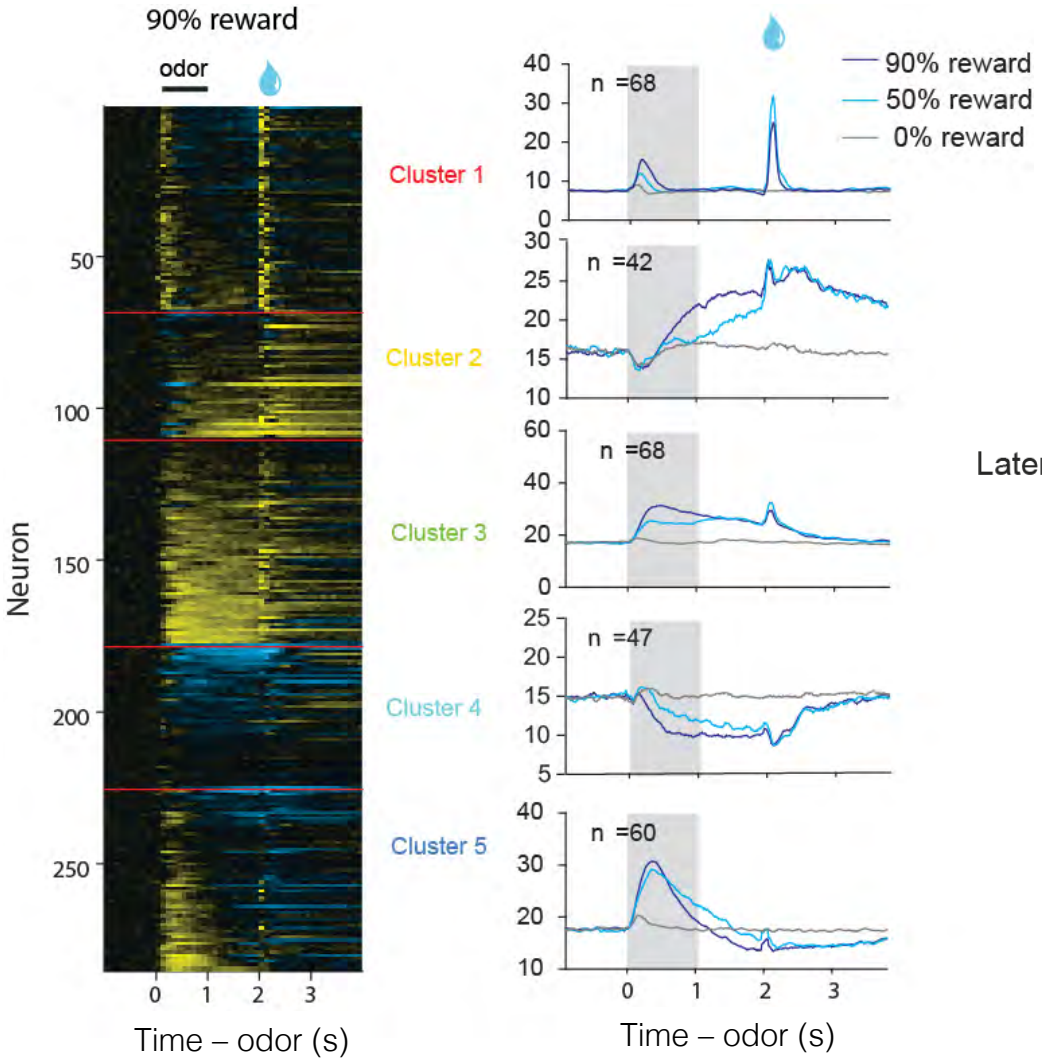
Activity of input neurons (Ventral striatum)

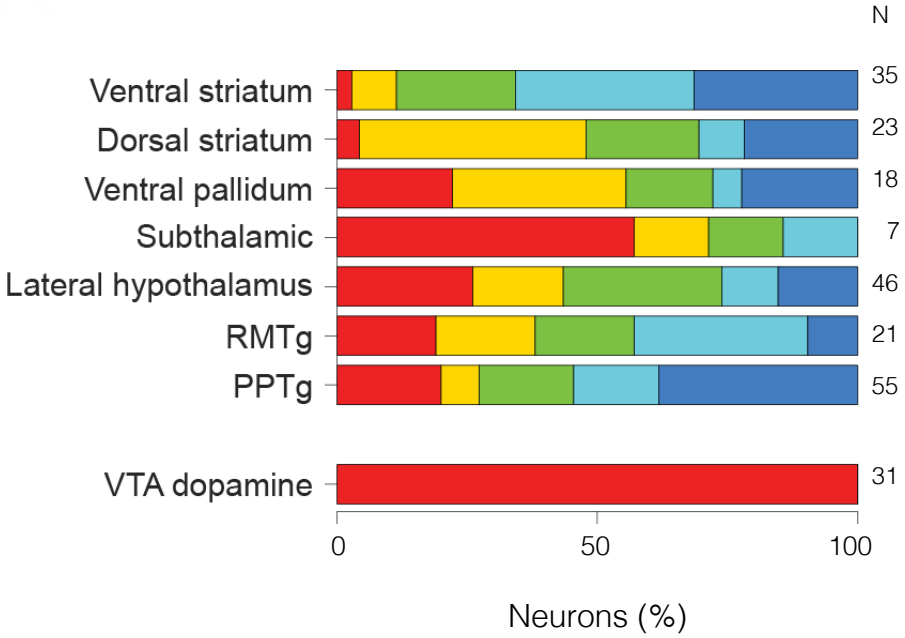


(Tian, et al., *Neuron* 2016)

Diverse firing patterns of identified input neurons

All identified inputs (n = 205) and VTA neurons

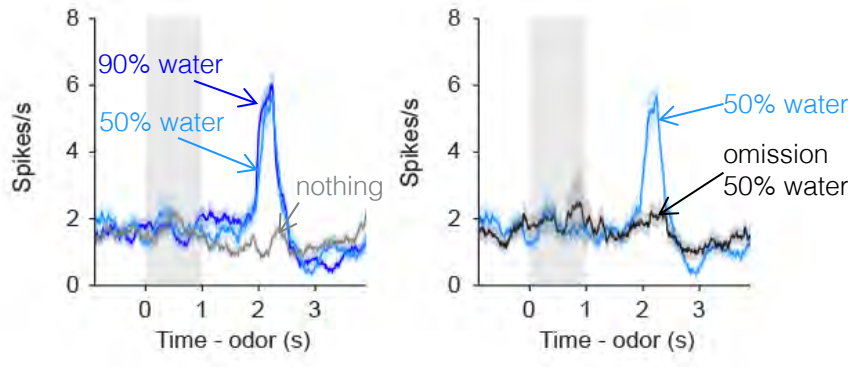


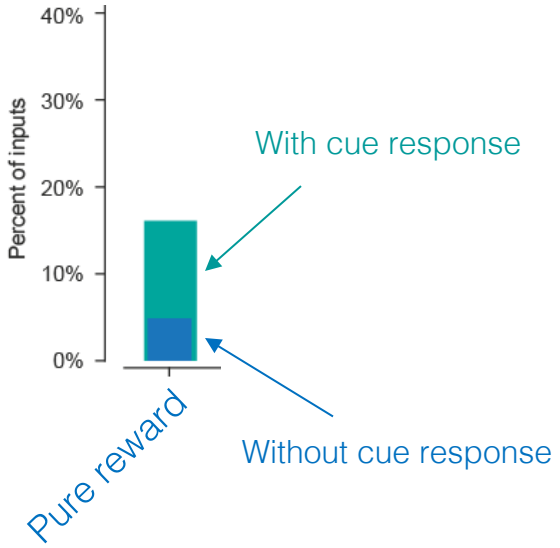


(Tian, et al., *Neuron* 2016)

Pure reward coding neurons?

Pure reward



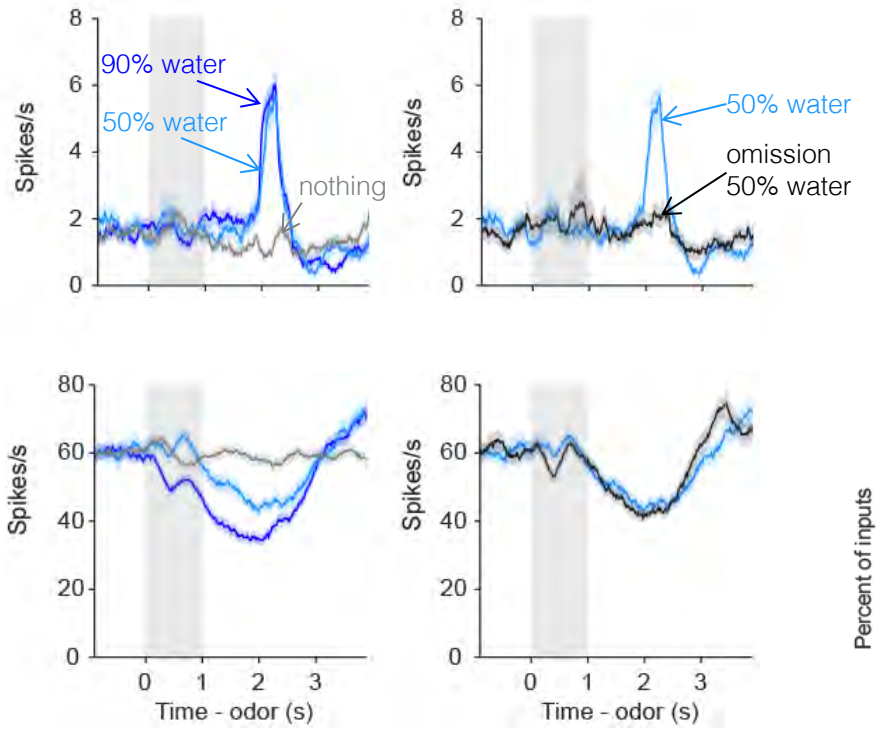


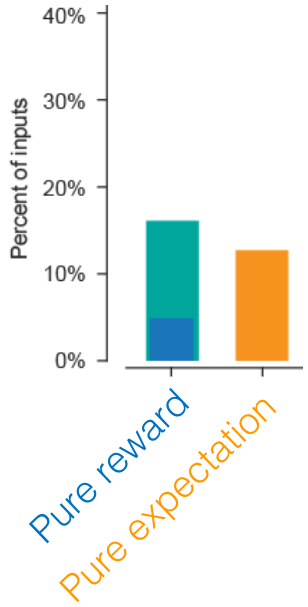
(Tian, et al., *Neuron* 2016)

Pure expectation coding neurons?

Pure reward

Pure expectation





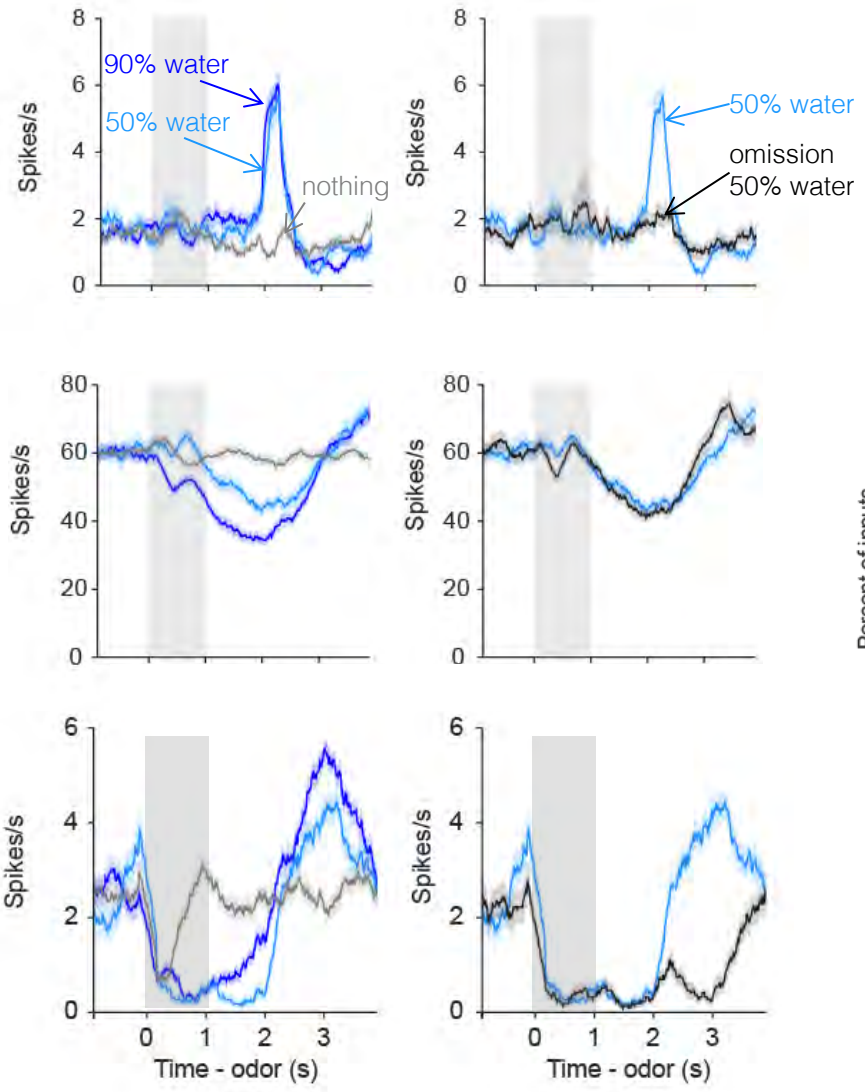
(Tian, et al., *Neuron* 2016)

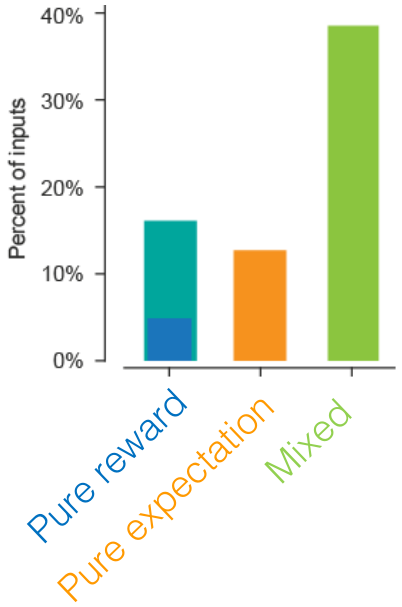
Many input neurons encode both reward and expectation

Pure reward

Pure expectation

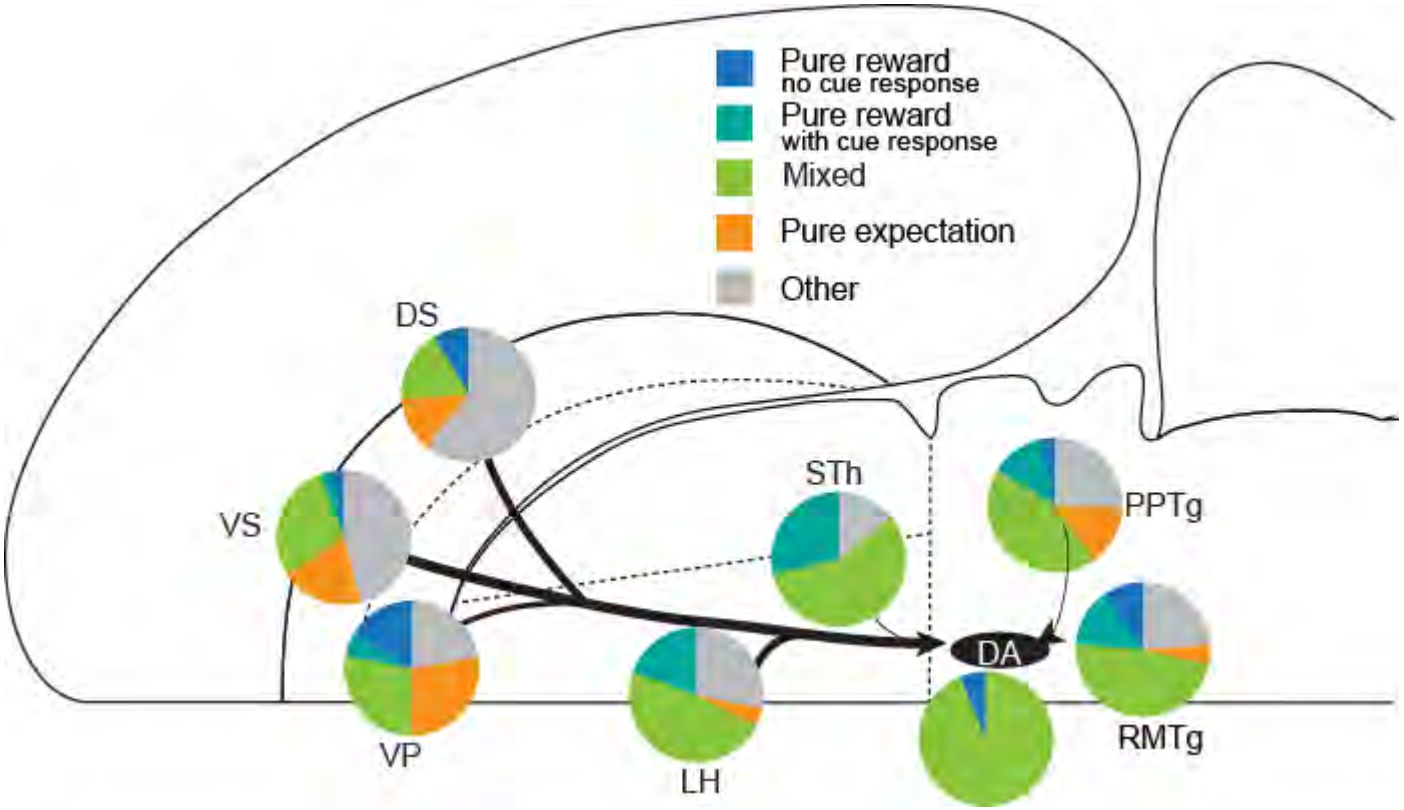
Mixed
(reward, expectation)





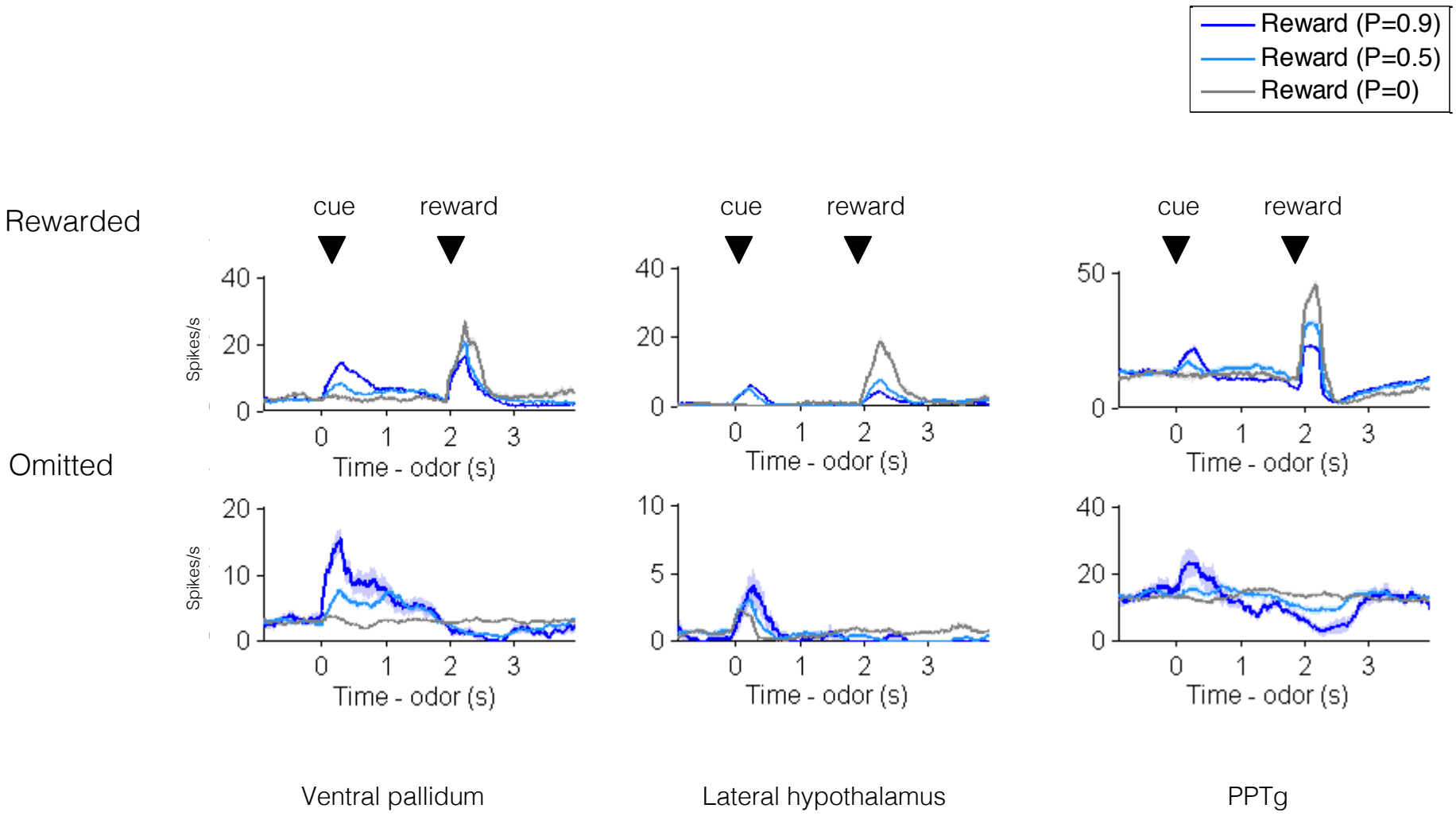
(Tian, et al., *Neuron* 2016)

Mixed coding of reward and expectation signals in identified input neurons



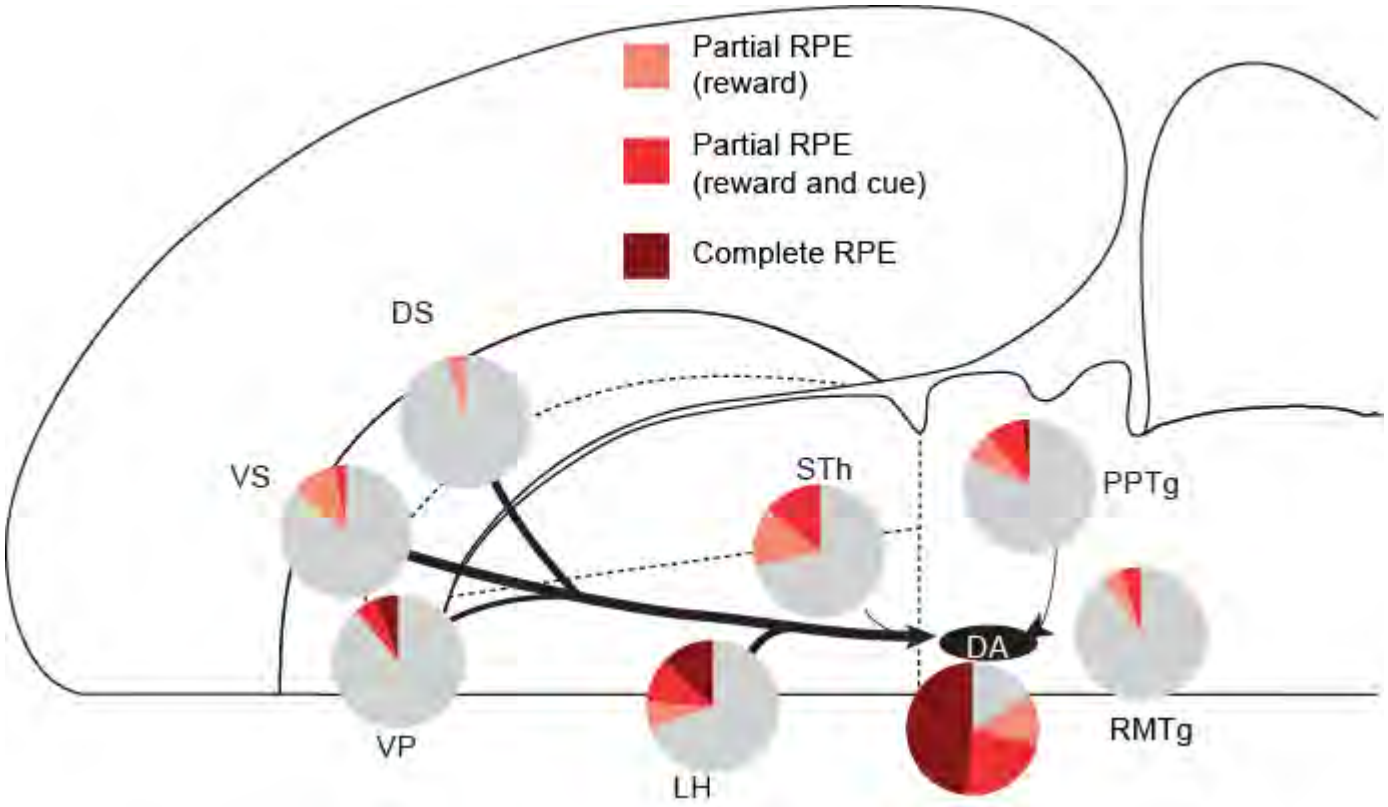
(Tian, et al., *Neuron* 2016)

Some input neurons already have RPE signals



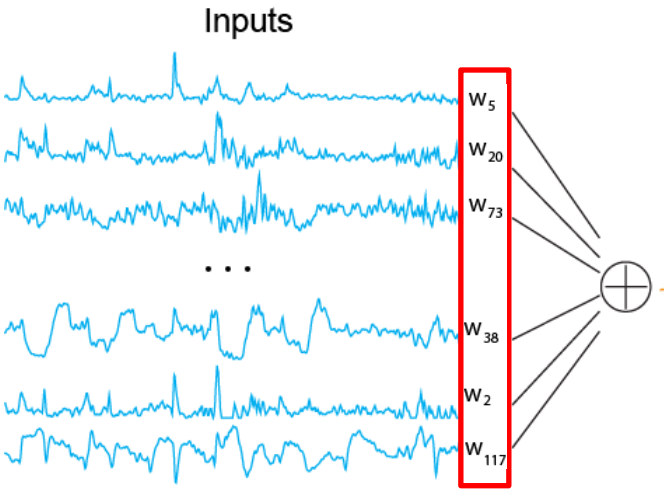
(Tian, et al., *Neuron* 2016)

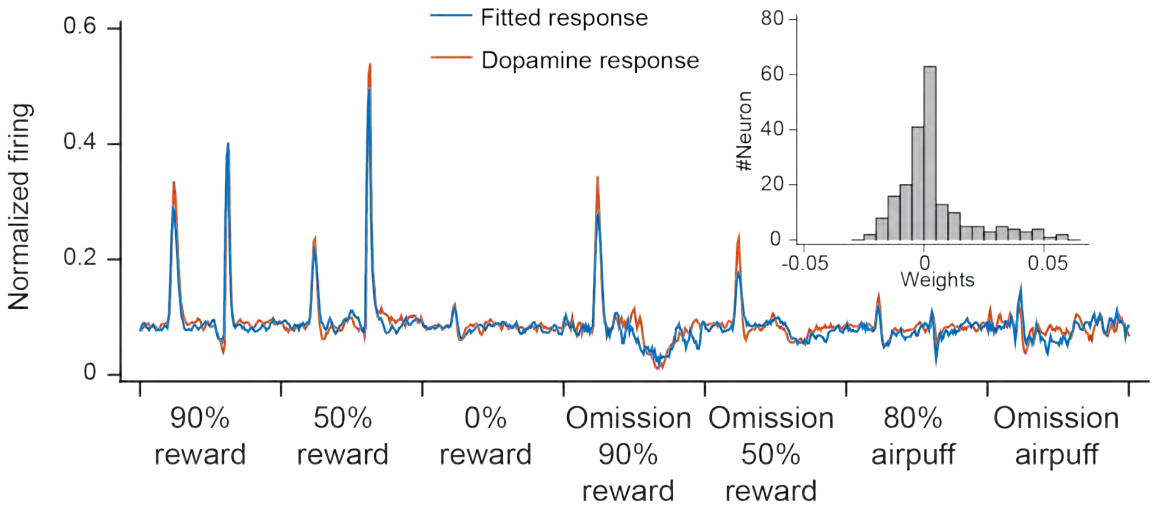
Identified input neurons encode RPE signals



(Tian, et al., *Neuron* 2016)

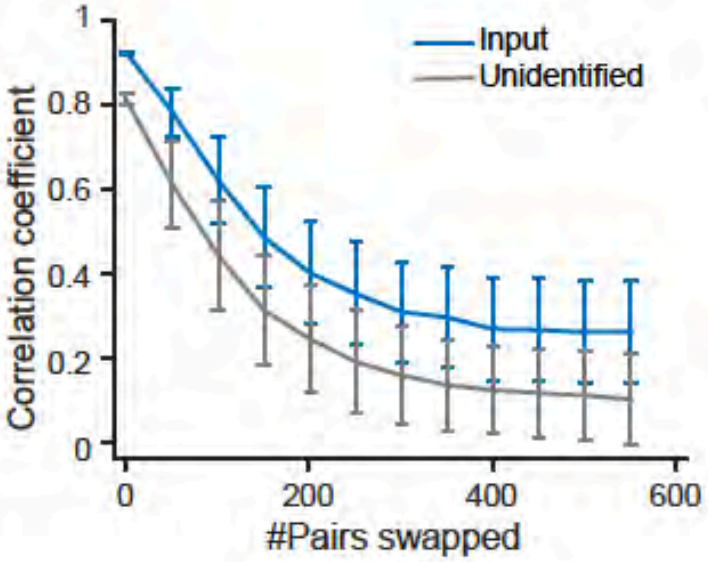
Precise weights of inputs are not required for reading out of RPE signals





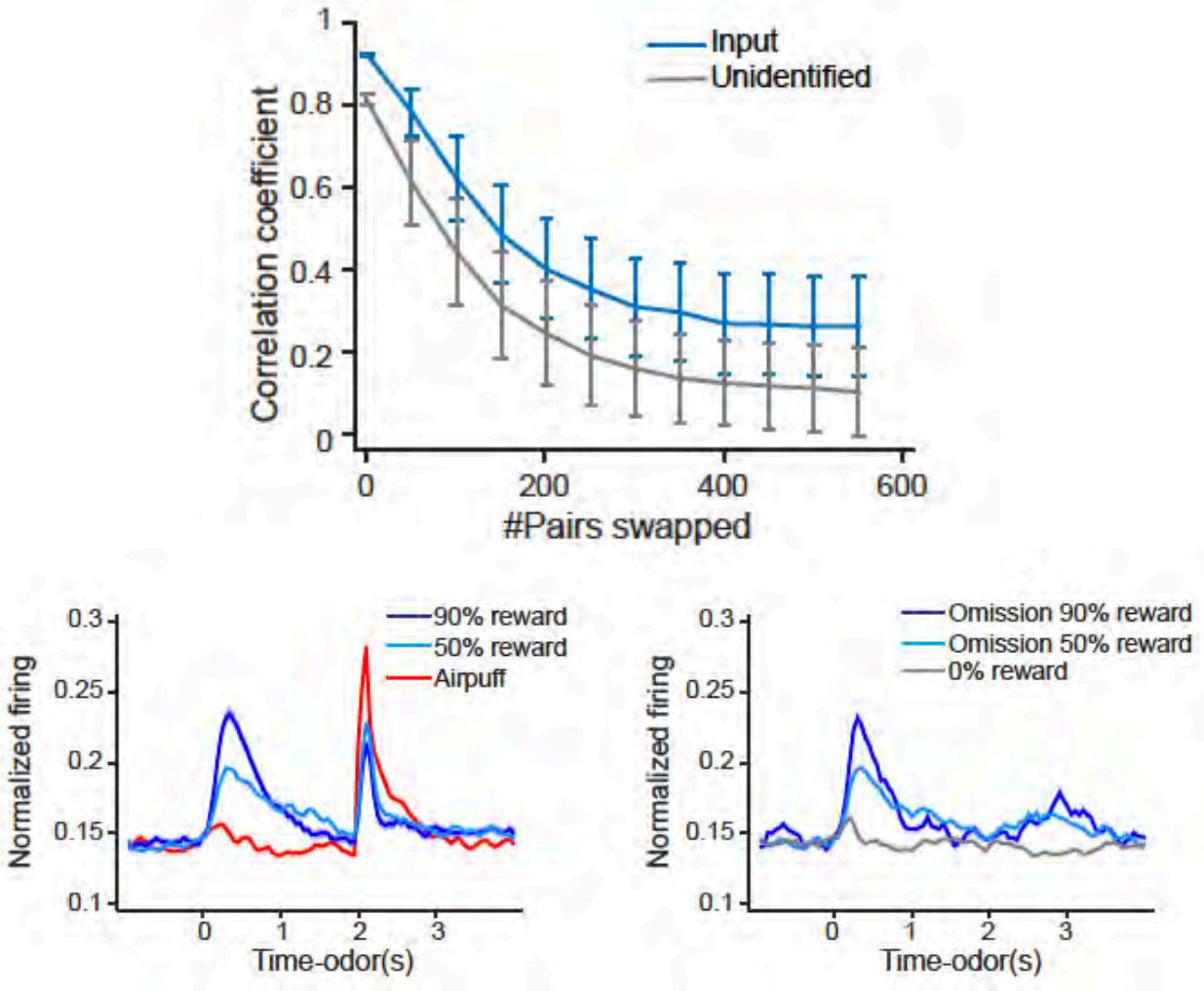
(Tian, et al., *Neuron* 2016)

RPE-like responses without fine tuning



(Tian, et al., *Neuron* 2016)

RPE-like responses without fine tuning

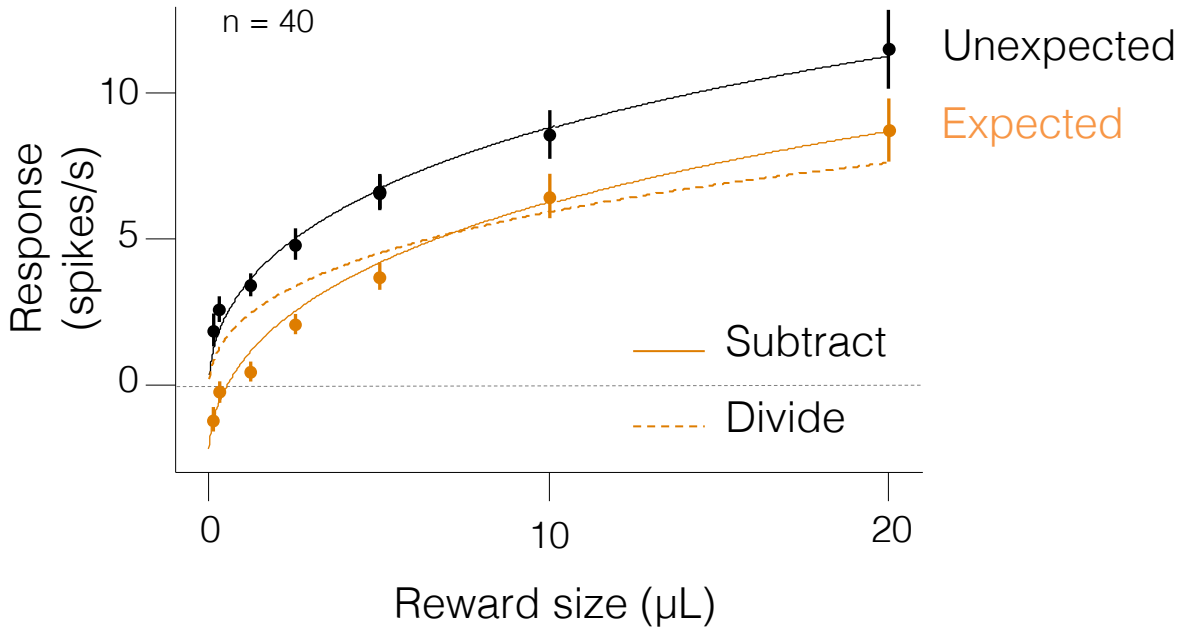


(Tian, et al., *Neuron* 2016)

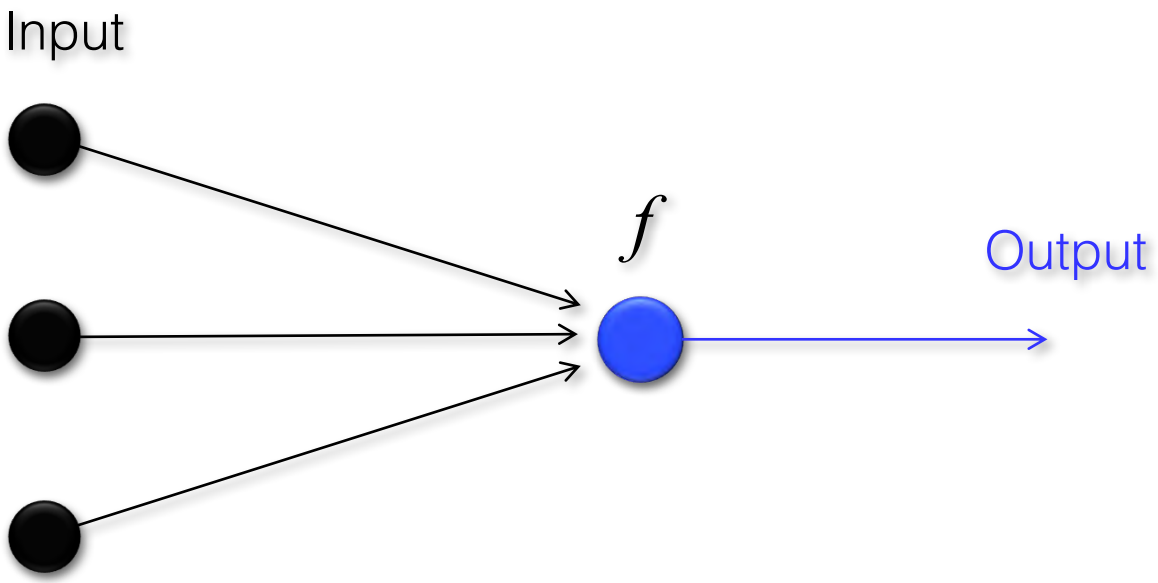
Activity of input neurons (summary)

- All tested areas contain neurons with diverse firing patterns.
- Most neurons do not encode a single variable. Information about reward and expectation are already combined in a complex manner.
- RPEs are already partially computed in some input neurons.
- Despite these complexities at the level of inputs, once they are combined, dopamine neurons signal RPEs in an extremely homogeneous fashion.
- Random combinations of inputs can recapitulate aspects of RPEs except for some specific components such as their response during air puff or reward omission.

Reward expectation triggers subtraction



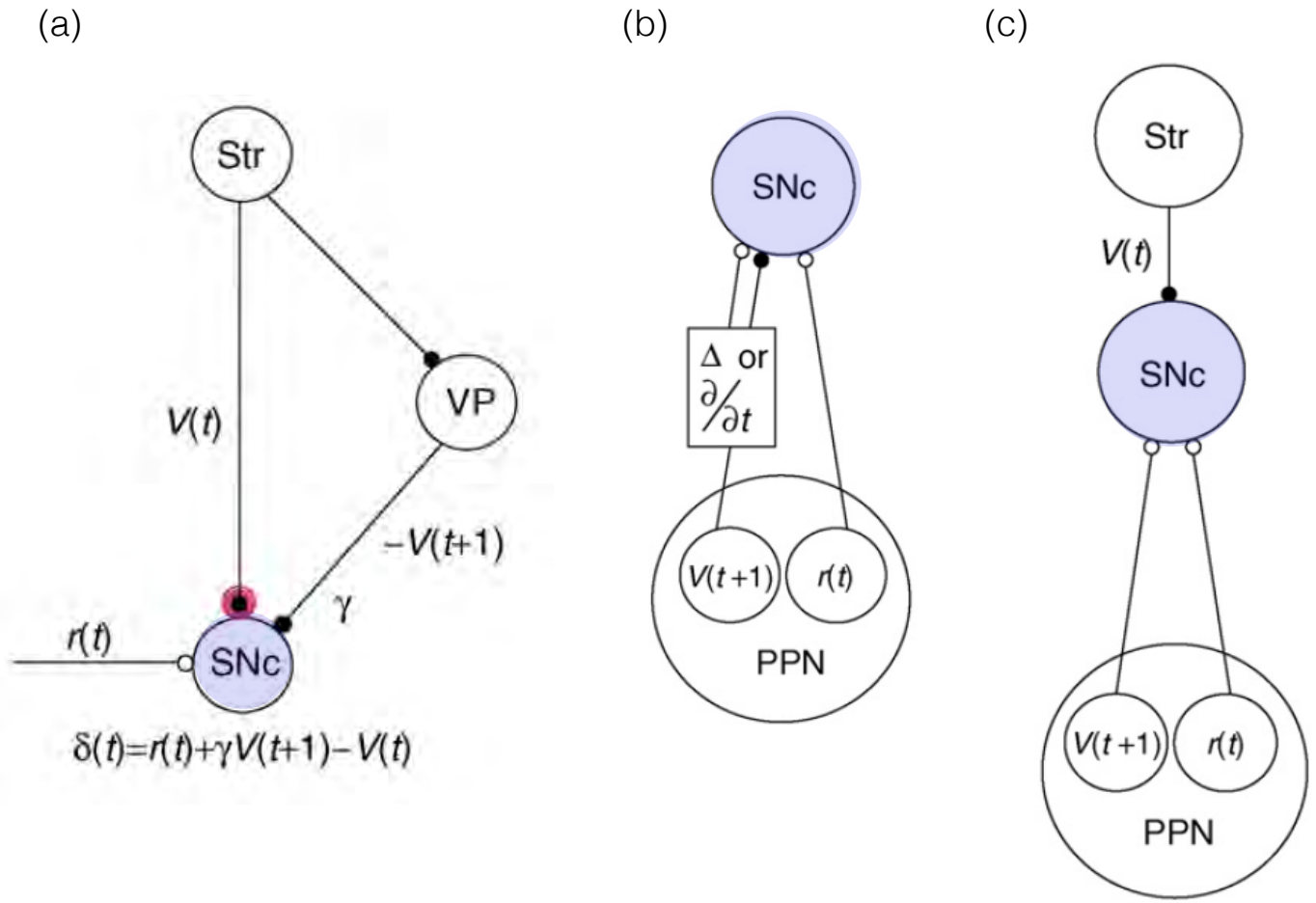
Understanding computation (arithmetic) in the brain



$$\delta = V_{actual} - V_{expected}$$

$$\delta = r + V(t) - V(t\text{-}1)$$

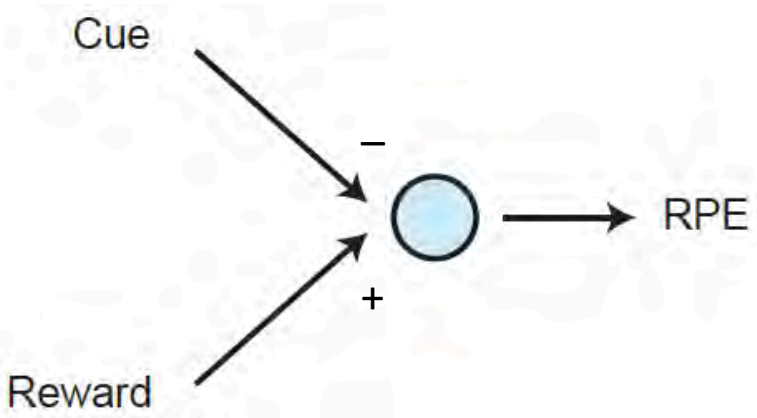
Models of RPE computation

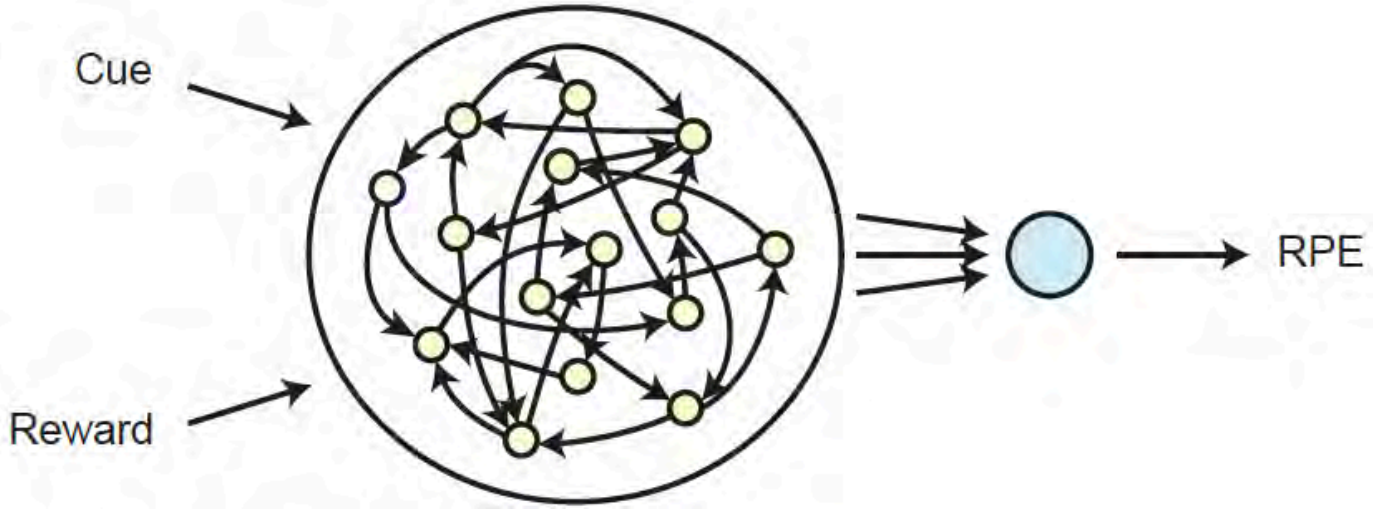


(Kawato and Samejima, 2007)

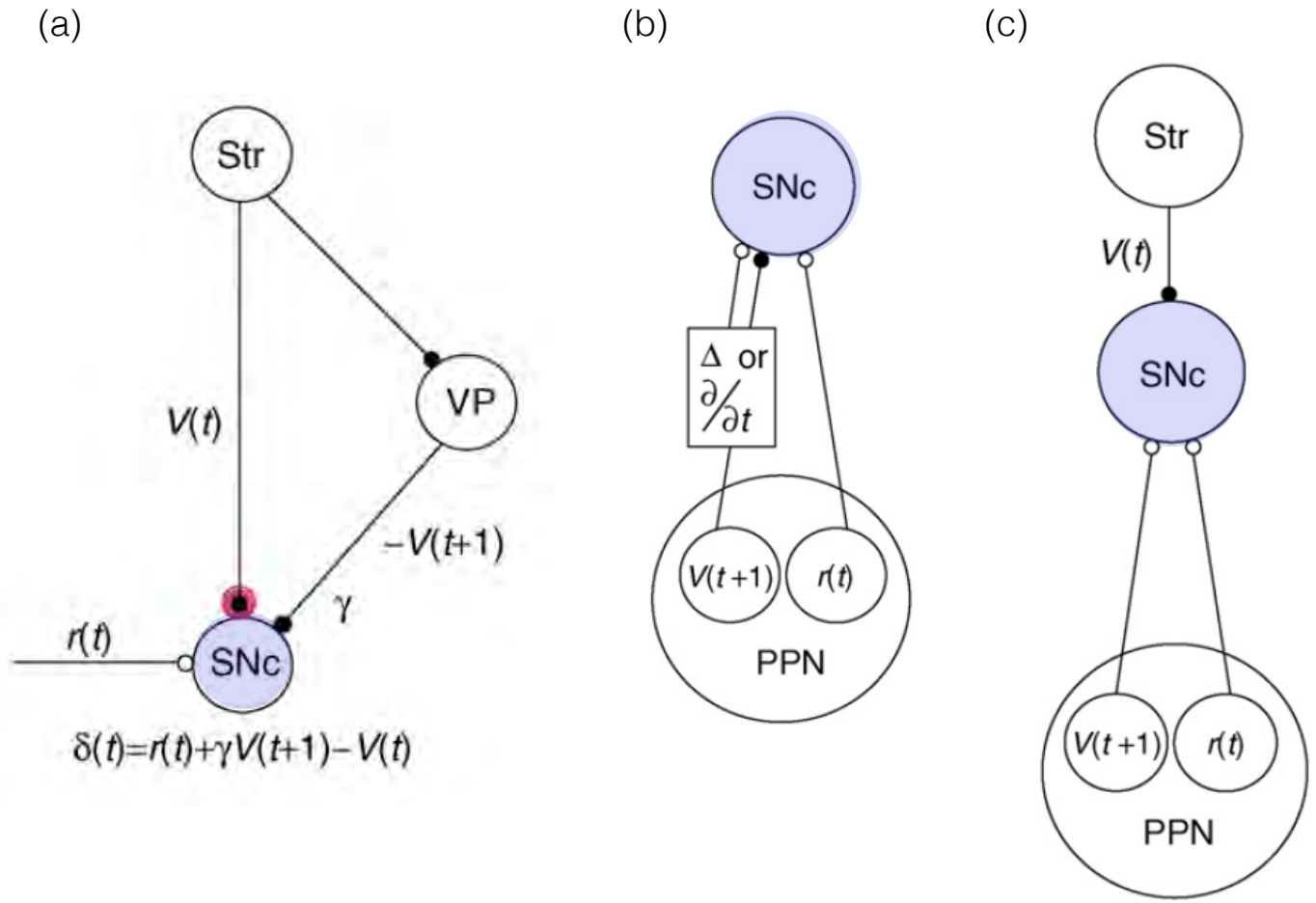
Arithmetic in the brain

- Representation might not be simple (pure)
- Computation might be distributed and/or gradual
  - Mixing information (“mixed selectivity”) in non-linear way may expand the space of information coding and/or facilitate the readout of downstream neurons (Fusi et al., 2016).
  - Representation and computation may be embedded in patterns of population activities (Rigotti et al. 2013; Mante et al., 2013; etc.)
- Computing with recurrent network and loop. Theory??

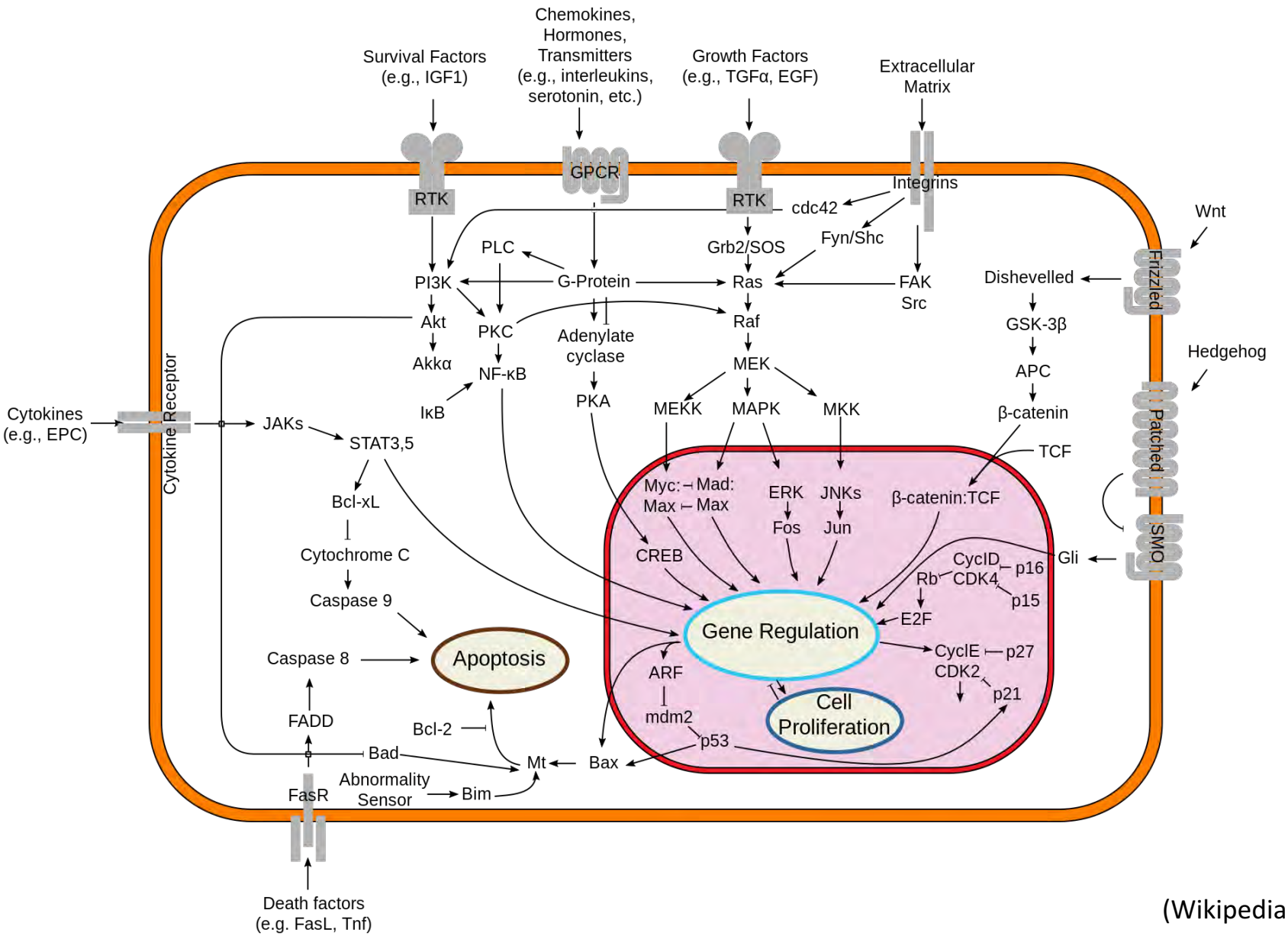

Cue
−
RPE
+
Reward


Cue
RPE
Reward

Models of RPE computation

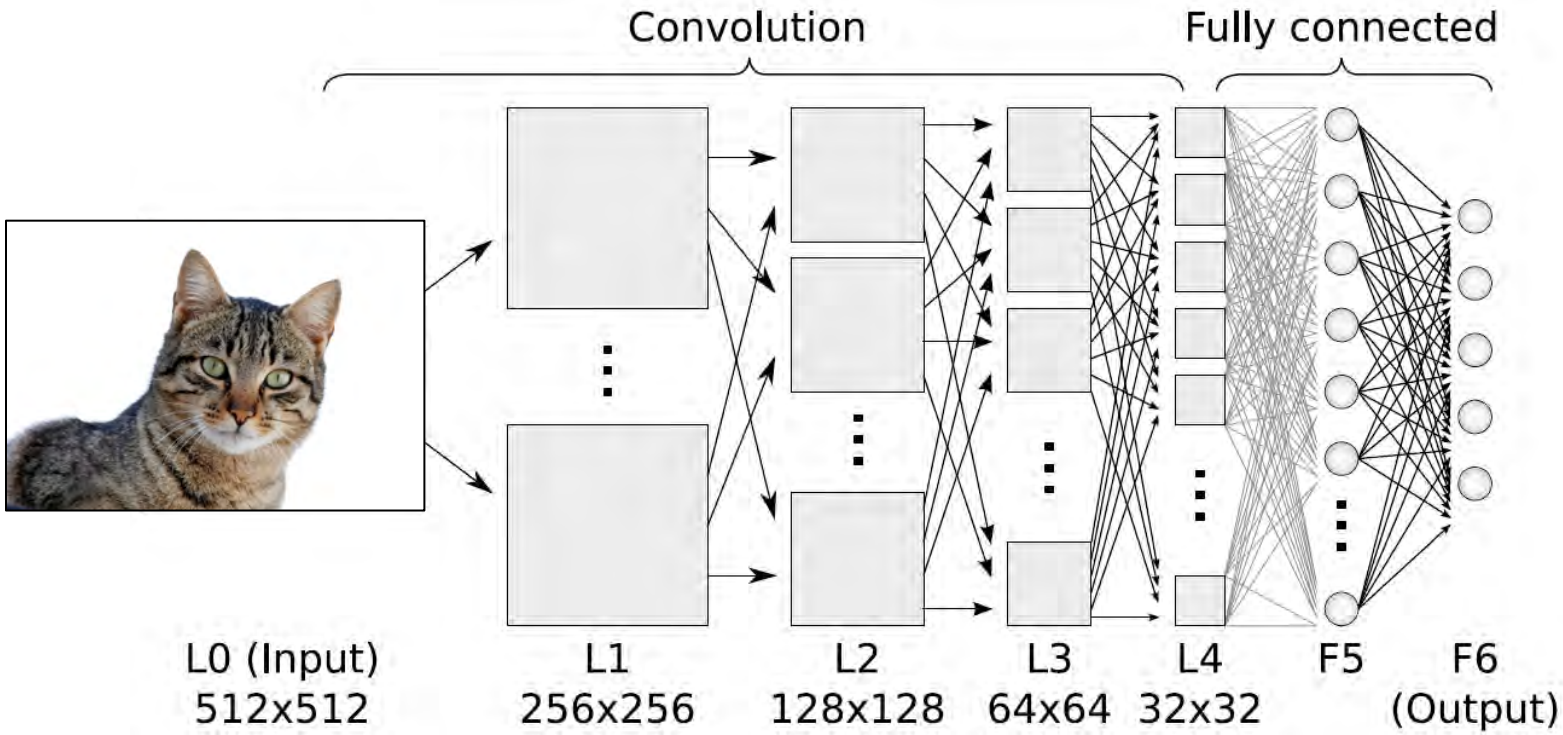


(Kawato and Samejima, 2007)


Survival Factors
(e.g., IGF1)
Chemokines,
Hormones,
Transmitters
(e.g., interleukins,
serotonin, etc.)
Growth Factors
(e.g., TGFα, EGF)
Extracellular
Matrix
RTK
GPCR
RTK
Integrins
cdc42
Wnt
Grb2/SOS
Fyn/Shc
PLC
PI3K
G-Protein
Ras
FAK
Src
Frizzled
Dishevelled
Akt
PKC
Adenylate
cyclase
Raf
GSK-3β
Hedgehog
Akkα
NF-κB
MEK
APC
Cytokines
(e.g., EPC)
Cytokine Receptor
JAKs
IκB
PKA
MEKK
MAPK
MKK
β-catenin
Patched
STAT3,5
TCF
Myc:⊣Mad:
Max⊢Max
ERK
JNKs
β-catenin:TCF
SMO
Bcl-xL
Fos
Jun
Gli
CREB
CycID
Rb
CDK4
p16
p15
Cytochrome C
E2F
Caspase 9
Gene Regulation
Caspase 8
Apoptosis
CycIE
p27
ARF
CDK2
p21
mdm2
Cell
Proliferation
FADD
Bcl-2
p53
Bad
Mt
Bax
FasR
Abnormality
Sensor
Bim
Death factors
(e.g. FasL, Tnf)

(Wikipedia)

Deep learning



CAT!

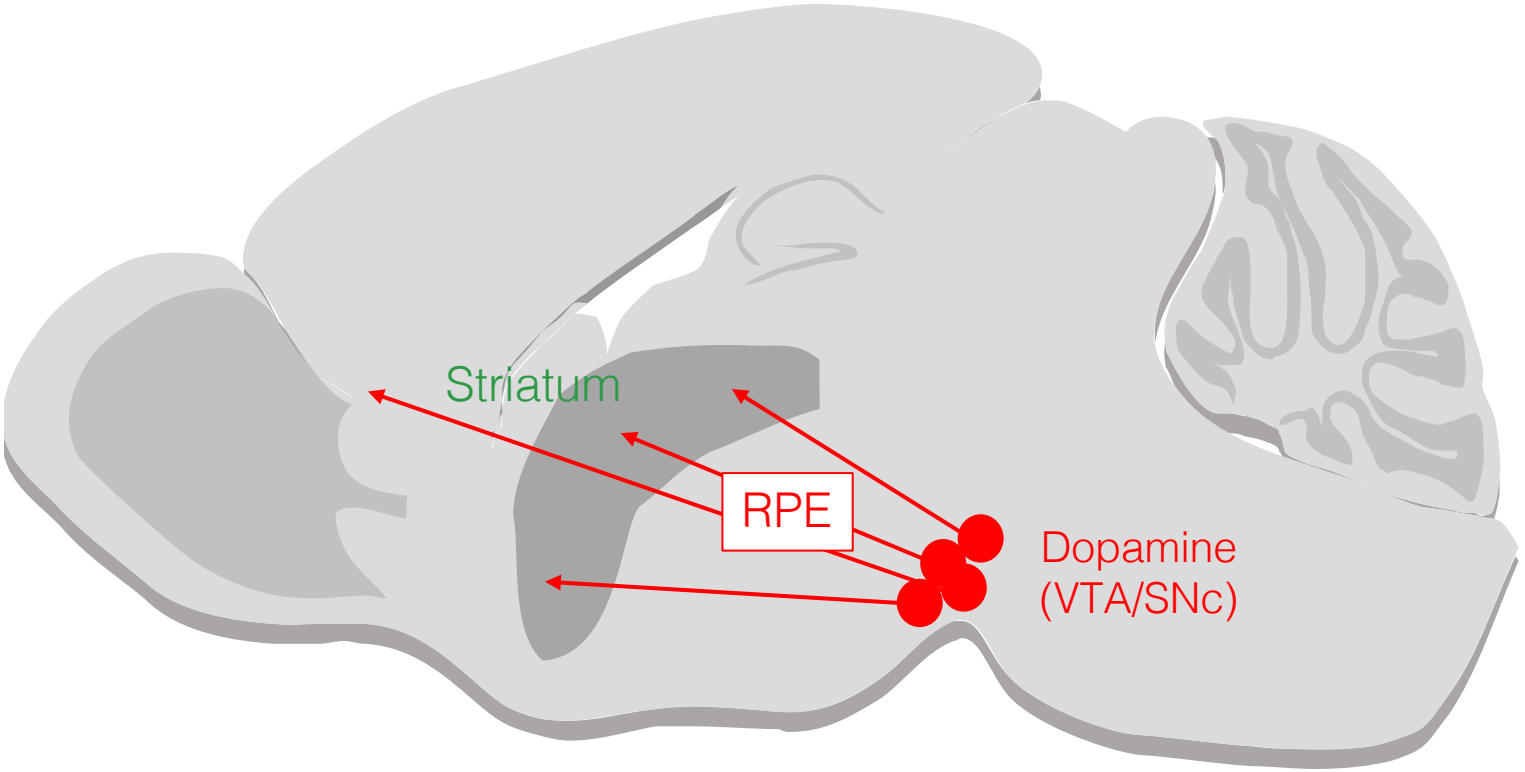
(recognition)

https://www.quora.com/What-is-a-convolutional-neural-network

Topics

- A mouse model to study dopamine RPE
- Do all dopamine neurons signal RPEs?
- What is the “state” in reinforcement learning?
- How are RPEs computed?
- Diversity of dopamine neurons

Dopamine neurons broadcast reward prediction error (RPE)

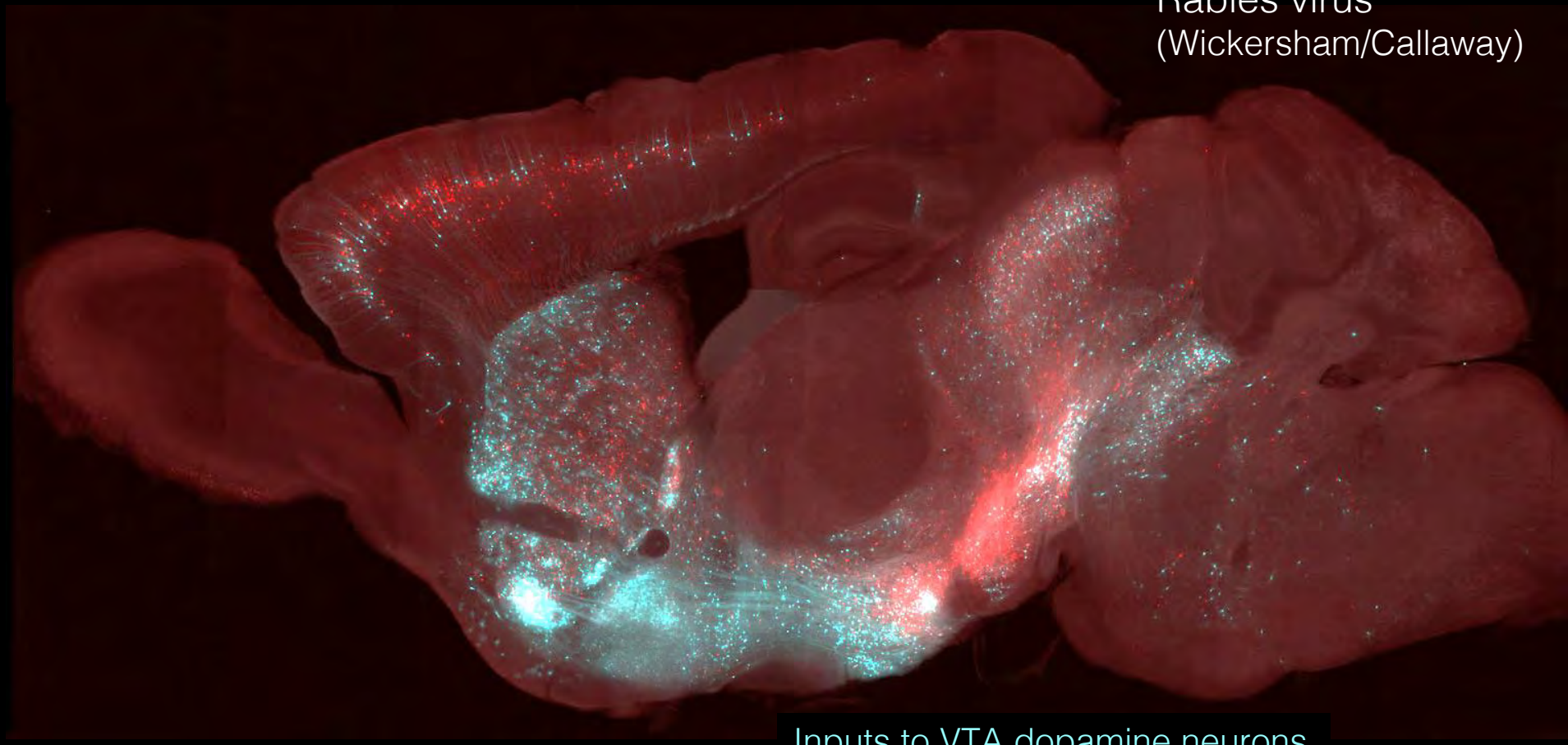


Do all dopamine neurons signal RPEs?

Diversity of dopamine neurons

- Anatomy (inputs)
- Dopamine signals (fiber fluorometry)

Direct inputs to dopamine neurons



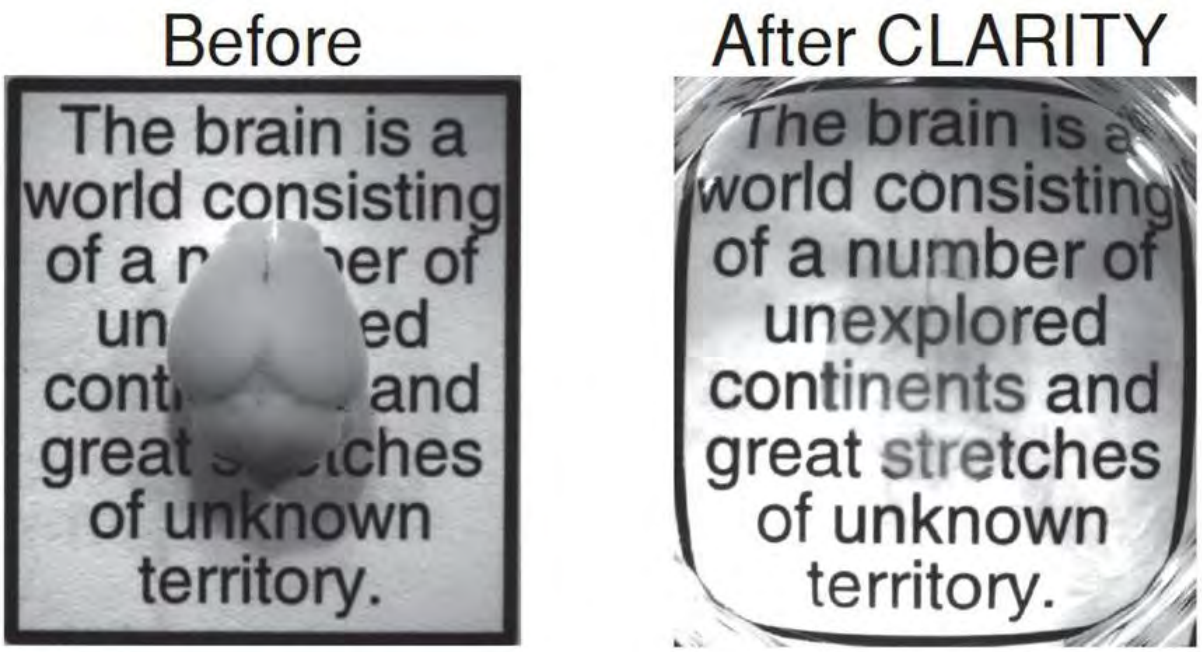
(Watabe-Uchida et al., *Neuron*, 2012)

Toward large-scale mapping

- CLARITY (brain-clearing method)
- Light-sheet imaging
- Automated analysis
  - Registering each brain to a standard brain (atlas)
  - Cell counting

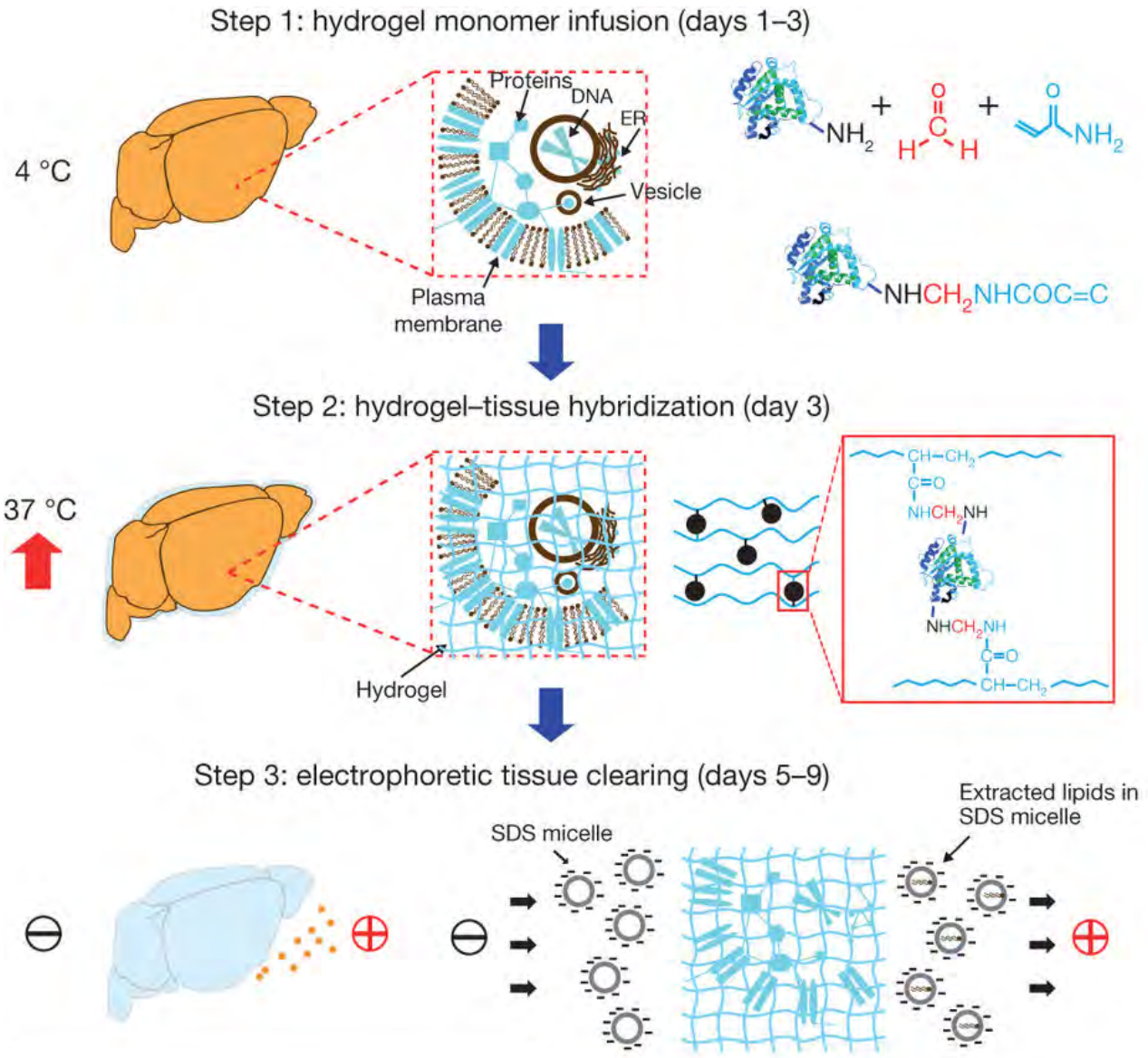
(Menegas et al., *eLife*, 2015)

CLARITY: a brain-clearing method

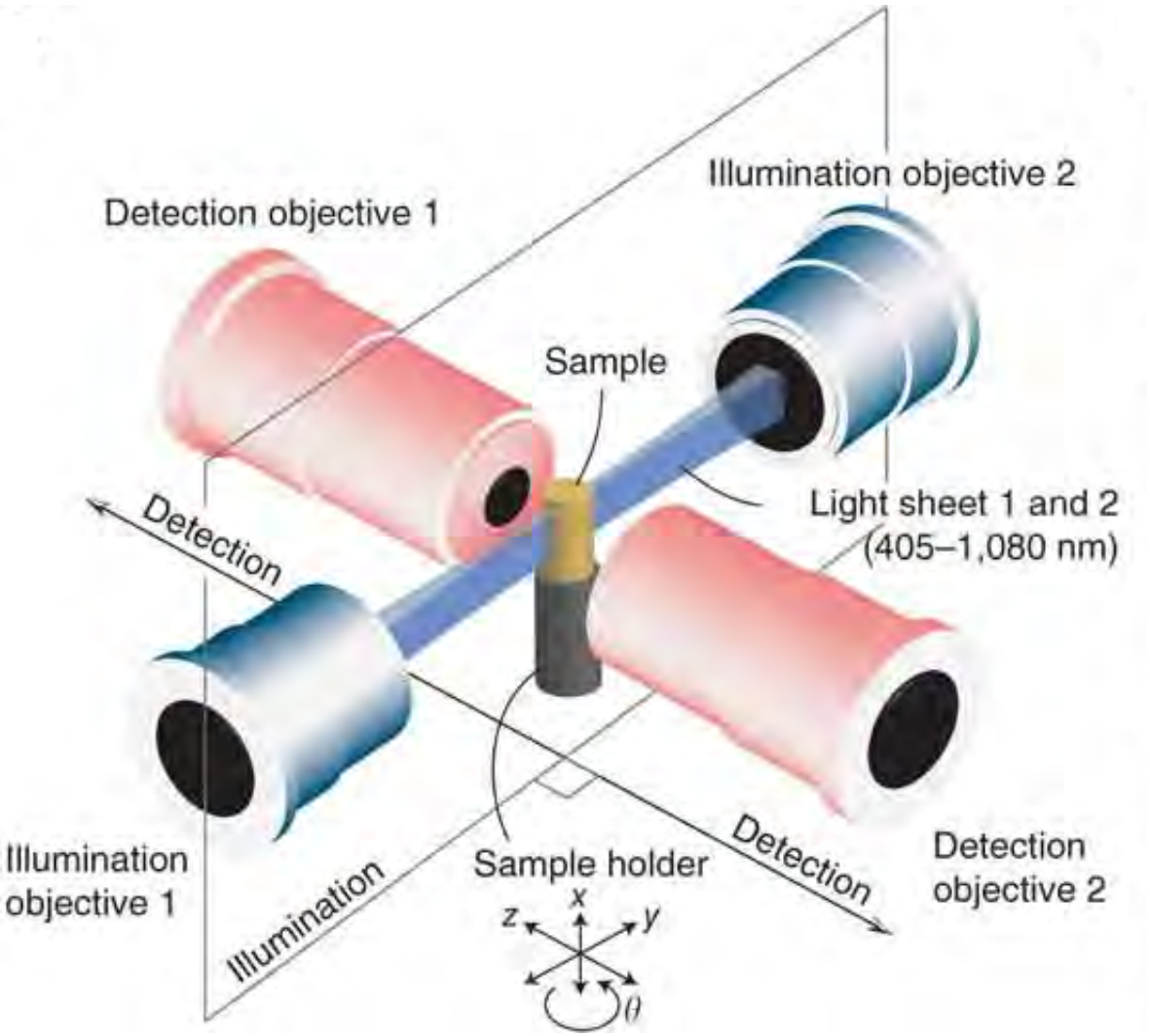


(Chung et al., *Nature*, 2015)

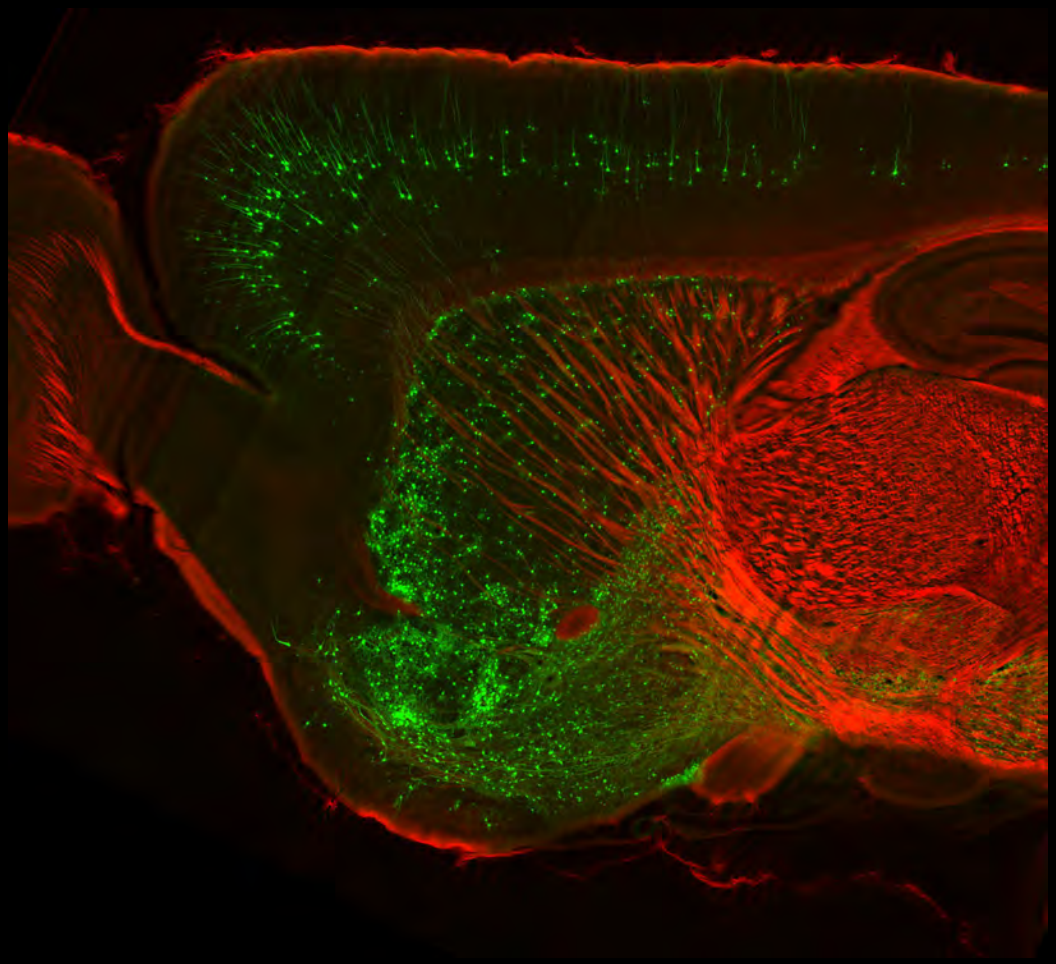
CLARITY: a brain-clearing method

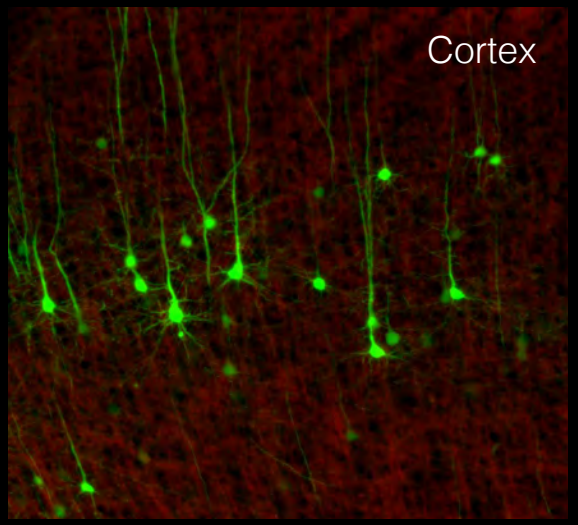


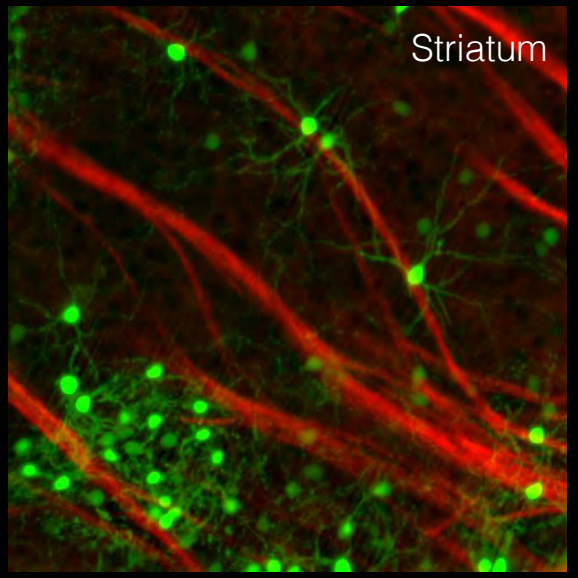
Light-sheet microscopy



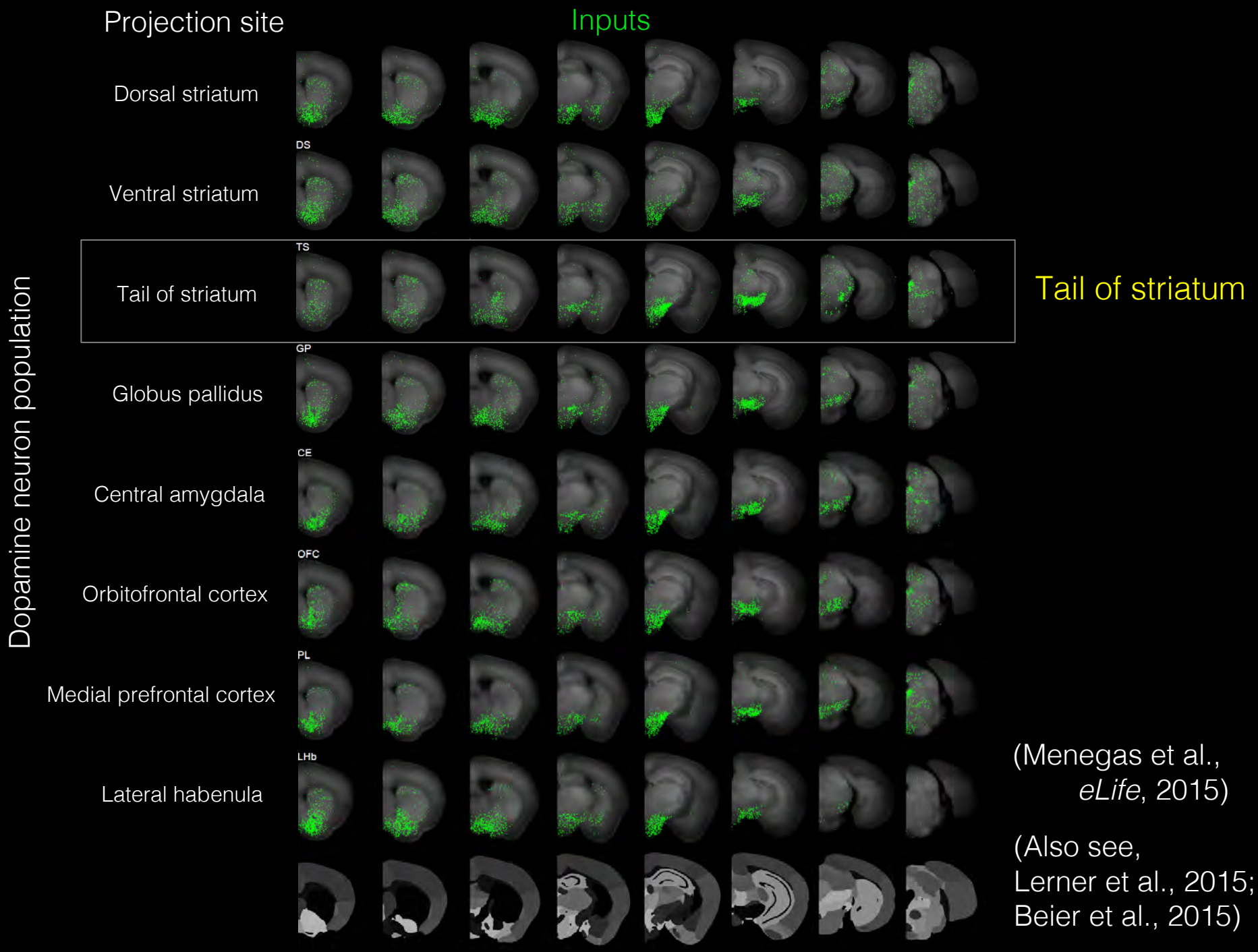
CLARITY + Light-sheet imaging





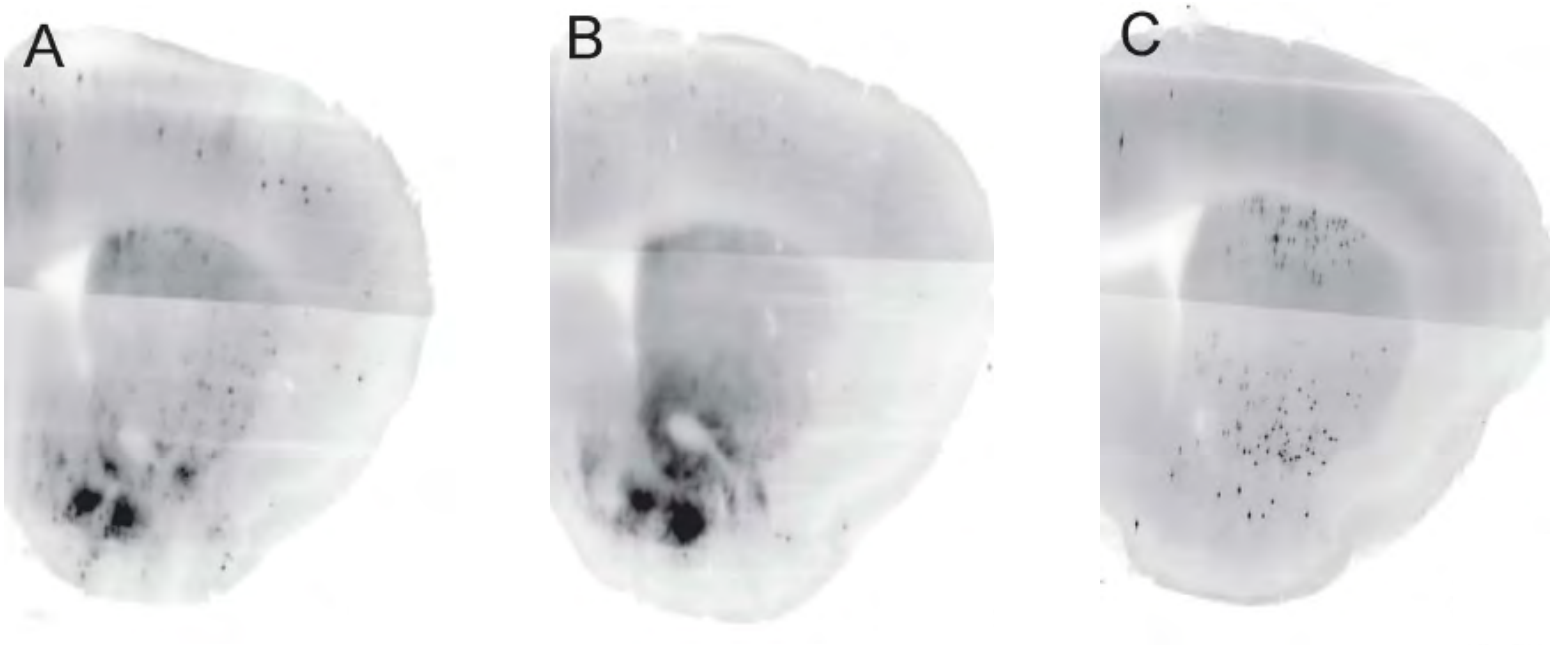


(Menegas et al., *eLife*, 2015)


Projection site
Inputs
Dopamine neuron population
Dorsal striatum
Ventral striatum
Tail of striatum
Globus pallidus
Central amygdala
Orbitofrontal cortex
Medial prefrontal cortex
Lateral habenula
DS
TS
GP
CE
OFC
PL
LHb
Tail of striatum
(Menegas et al., *eLife*, 2015)
(Also see, Lerner et al., 2015; Beier et al., 2015)

Unique input pattern for TS dopamine

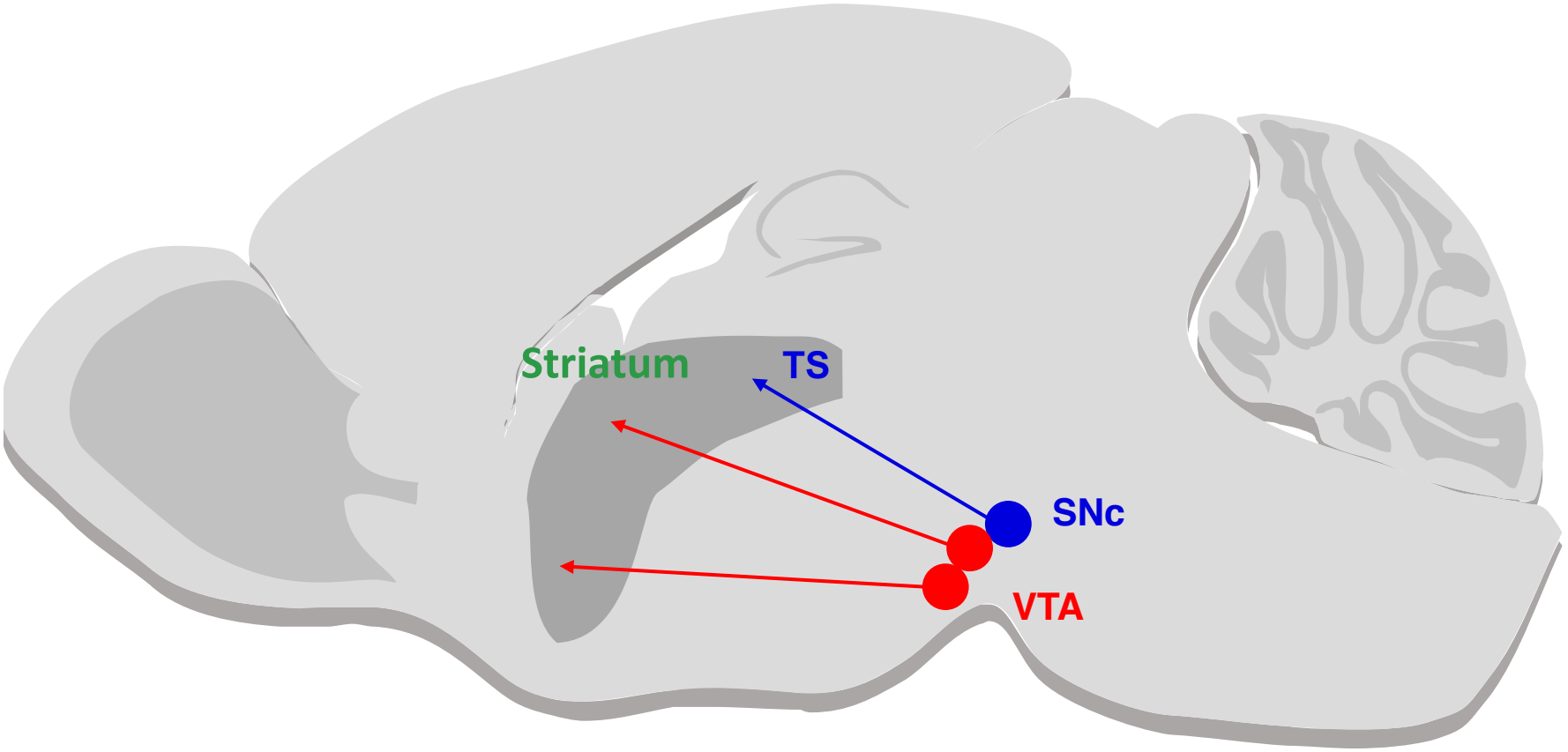
- Inputs to dopamine neurons



Projection target	Ventral striatum (VS)	Dorsal striatum (DS)	Tail of the striatum (TS)

(Menegas et al., *eLife*, 2015)

TS dopamine: anatomically unique population



Controversy

Do all dopamine neurons signal reward prediction errors?

- Aversive stimuli

- Some dopamine neurons are activated by aversive stimuli

(Matsumoto & Hikosaka, 2009; Lerner et al., 2015)

- Novel stimuli

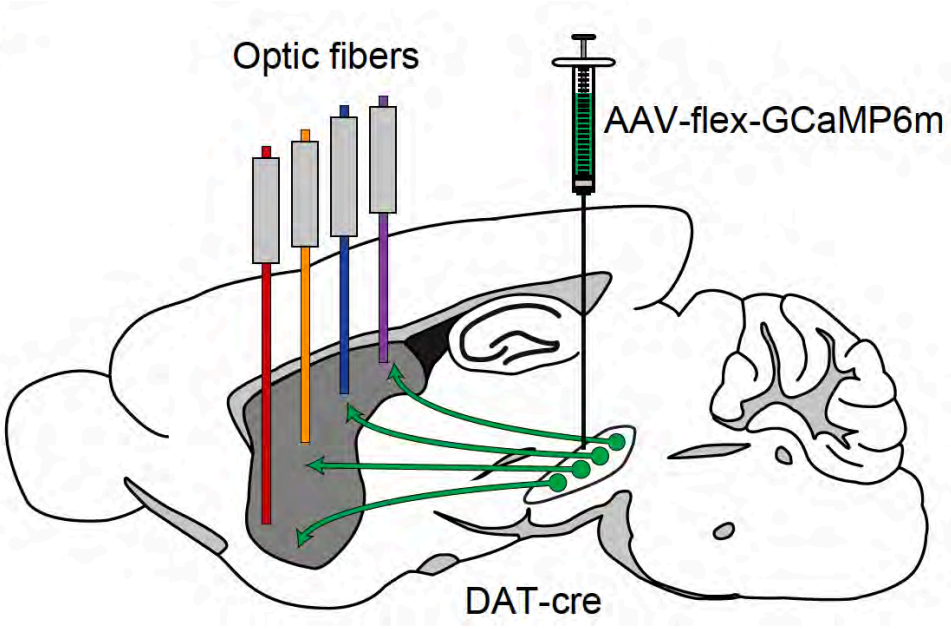
- Potential reward or novelty itself is rewarding

- Positive value of exploration

(Kakade & Dayan, 2002; Horvitz et al., 1997)

Calcium imaging at dopamine projection targets

- **Fiber fluorometry** (photometry)



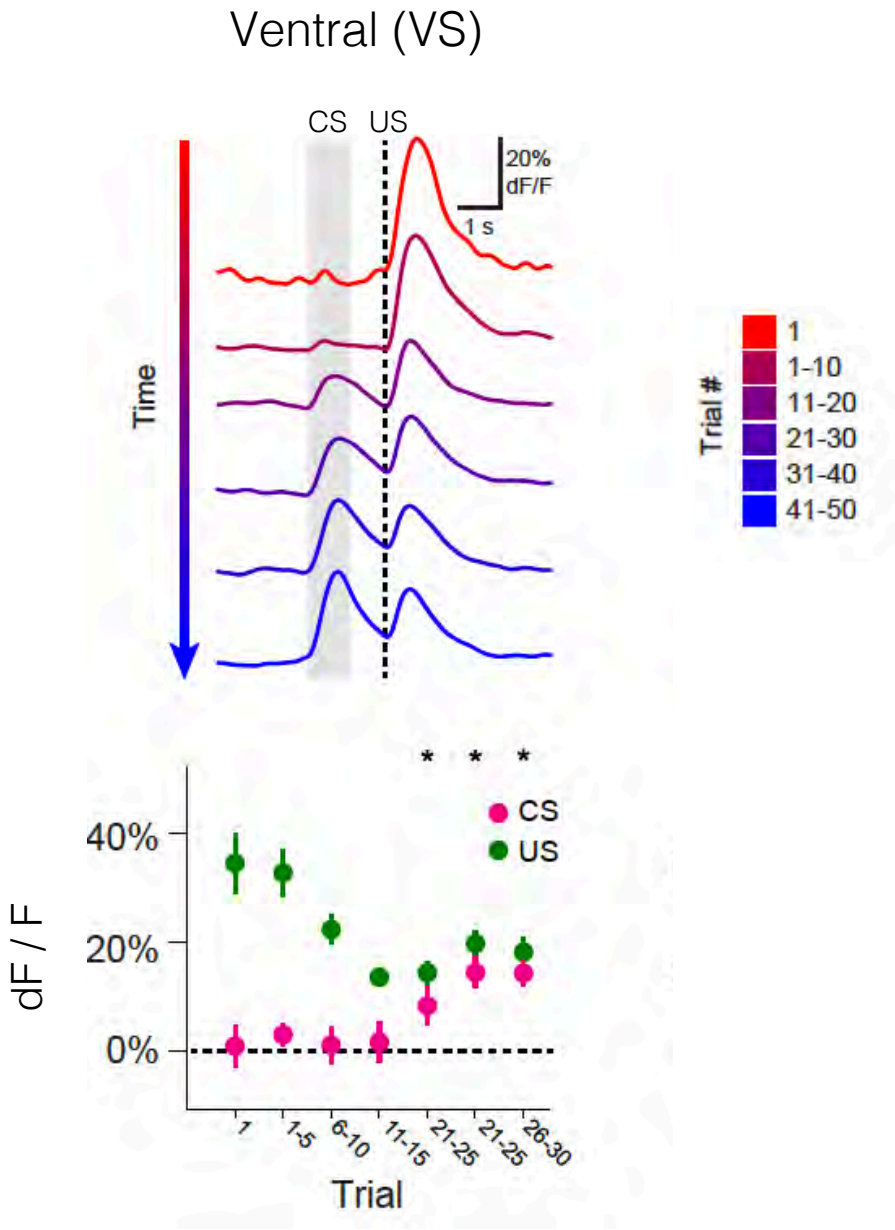
Kudo et al. (1992)
Davis and Schmidt (2000)
Adelsberger et al. (2005)
Murayama et al. (2007)

Cui et al. (2013)
Gunaydin et al. (2014)
Lerner et al. (2015)
Parker et al. (2016)
Howe and Dombeck (2016)

- Calcium indicator (GCaMP6m) in dopamine neurons
- Head-fixed mice
- Classical conditioning (reward, air puff, neutral stimuli)

(Menegas, Babayan, Uchida, Watabe-Uchida, *eLife*, 2017)

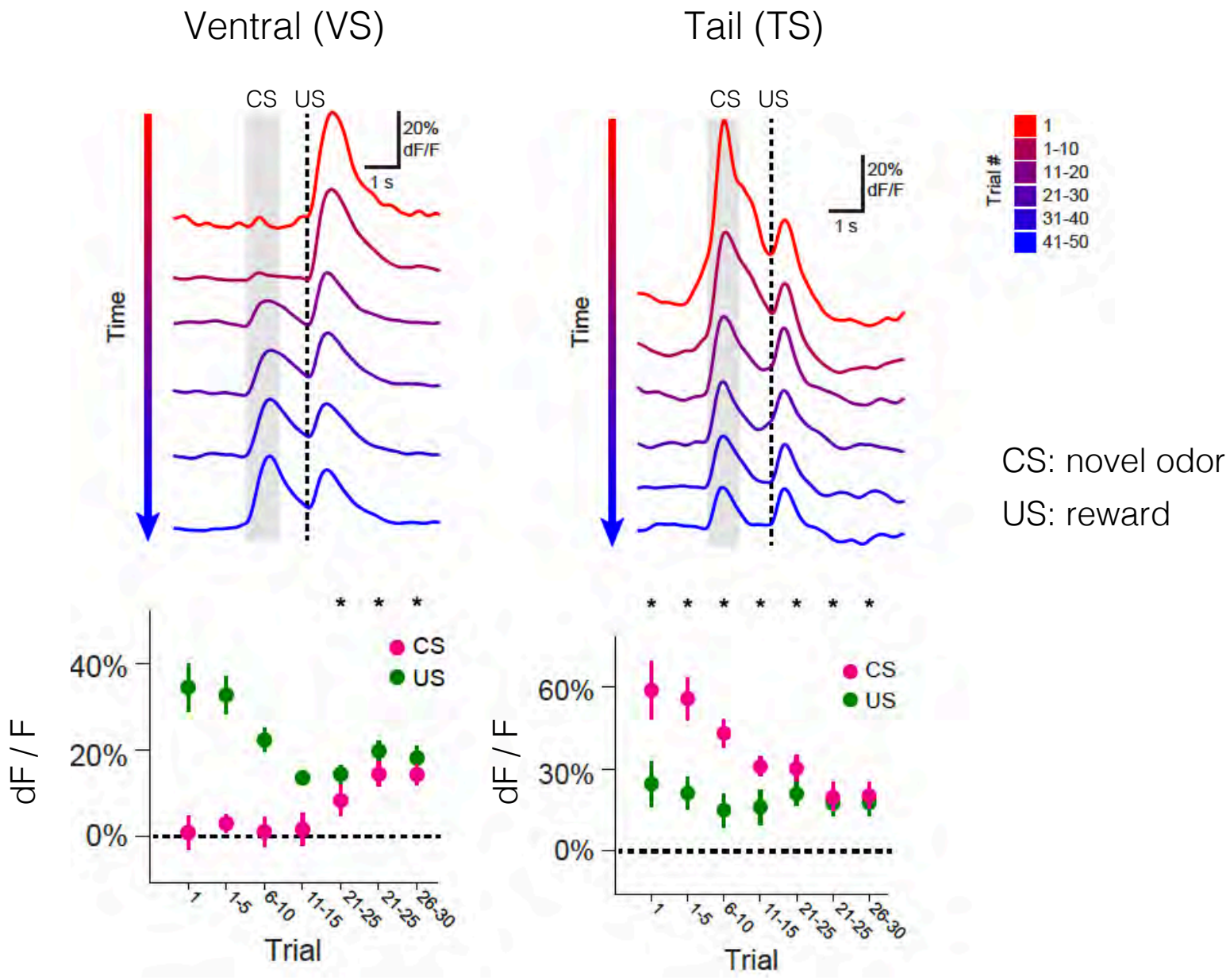
Ventral striatum (VS) dopamine signals RPE



CS: novel odor

US: reward

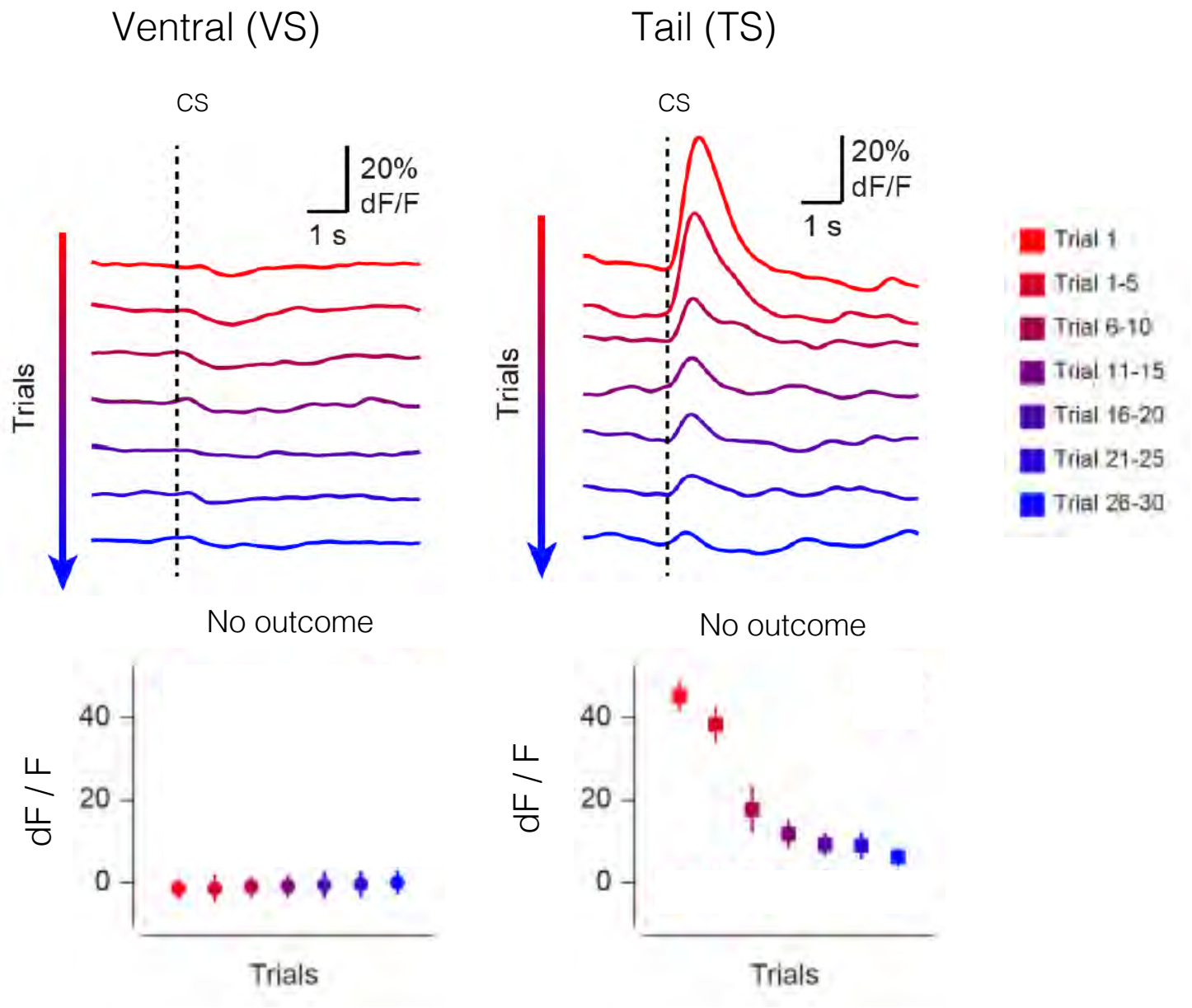
Novel stimuli activate the tail of the striatum (TS)



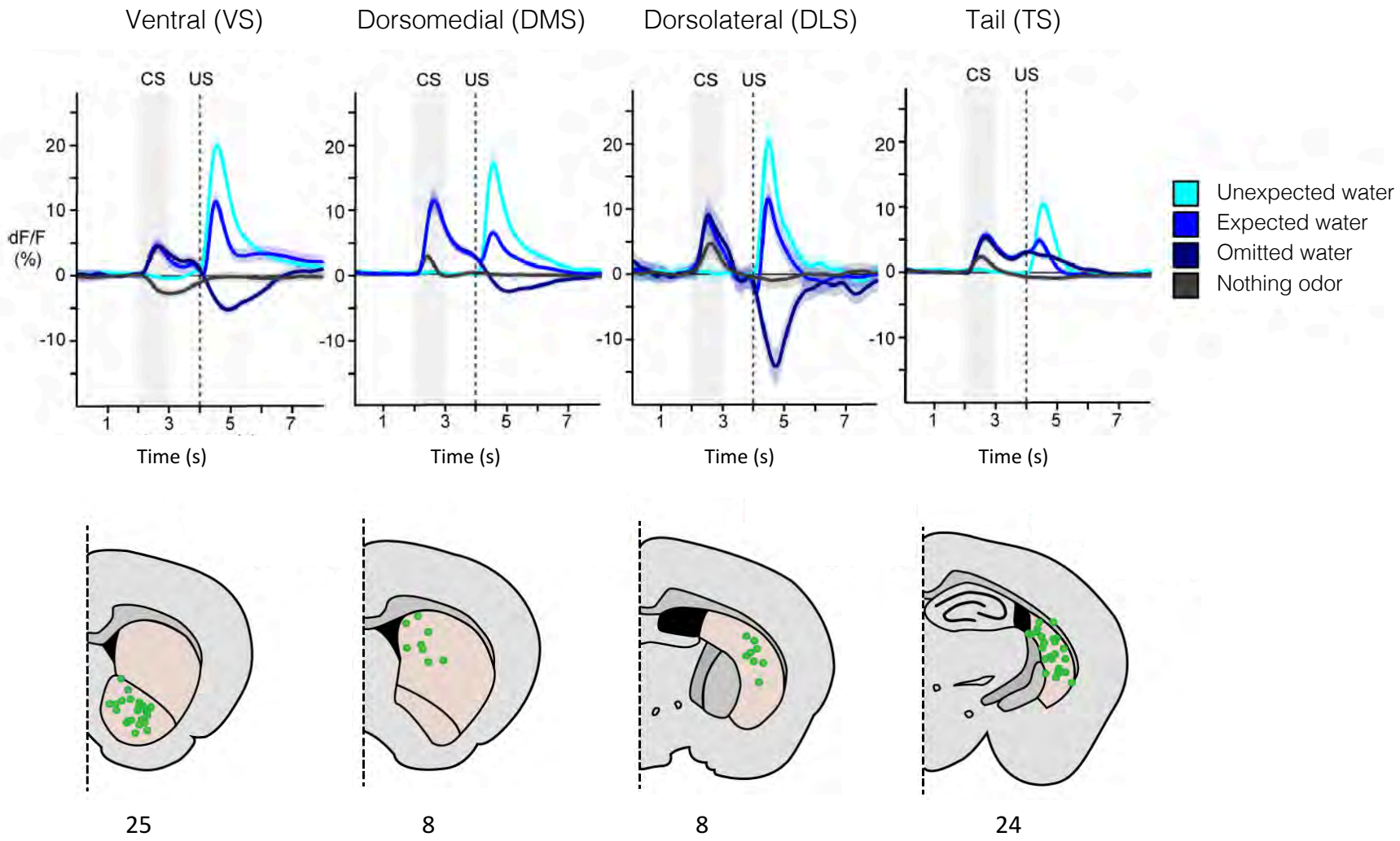
CS: novel odor

US: reward

Novel stimuli activate TS but not VS dopamine

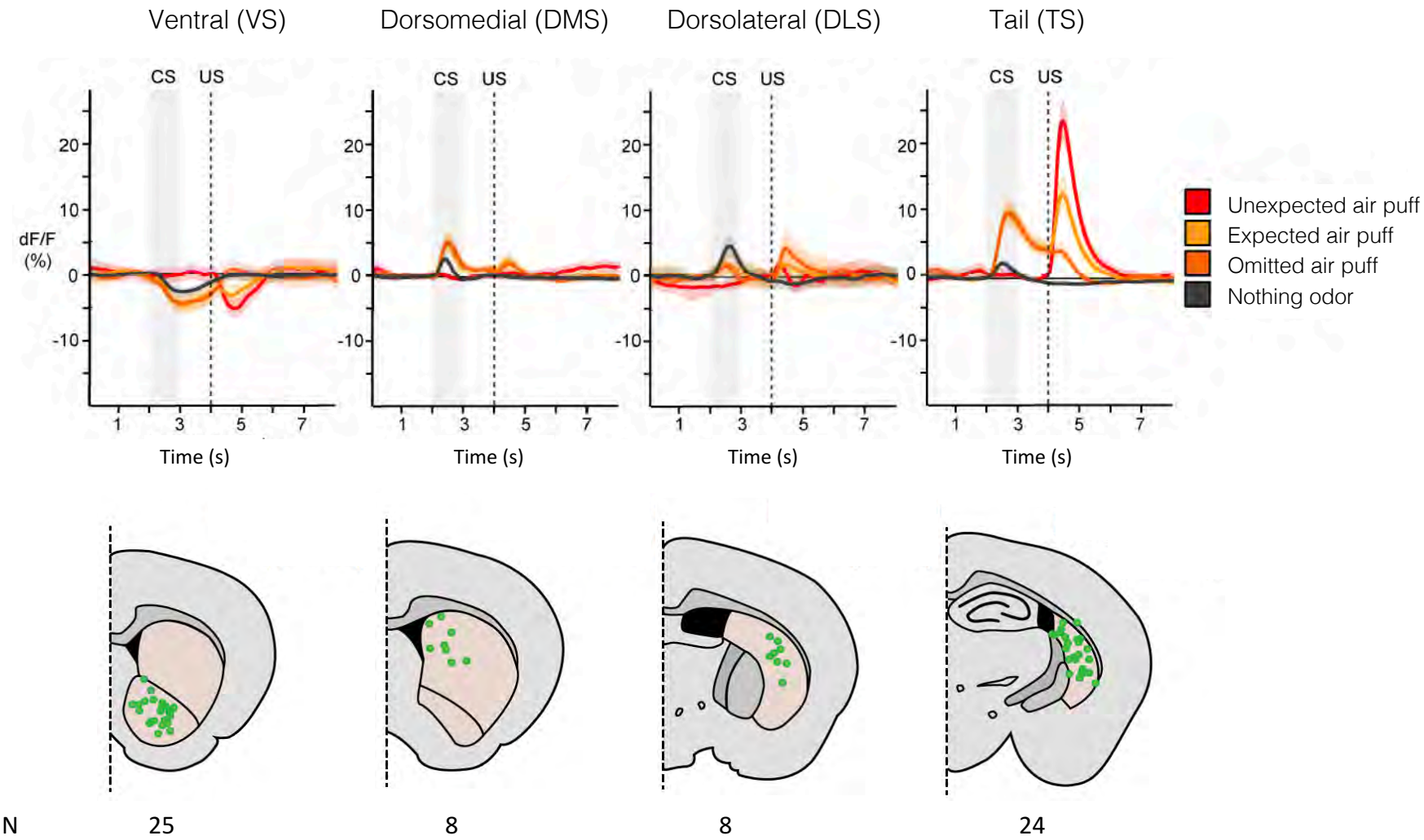


Responses to **reward** are relatively similar across areas



(Menegas, Babayan, Uchida, Watabe-Uchida, *eLife*, 2017)

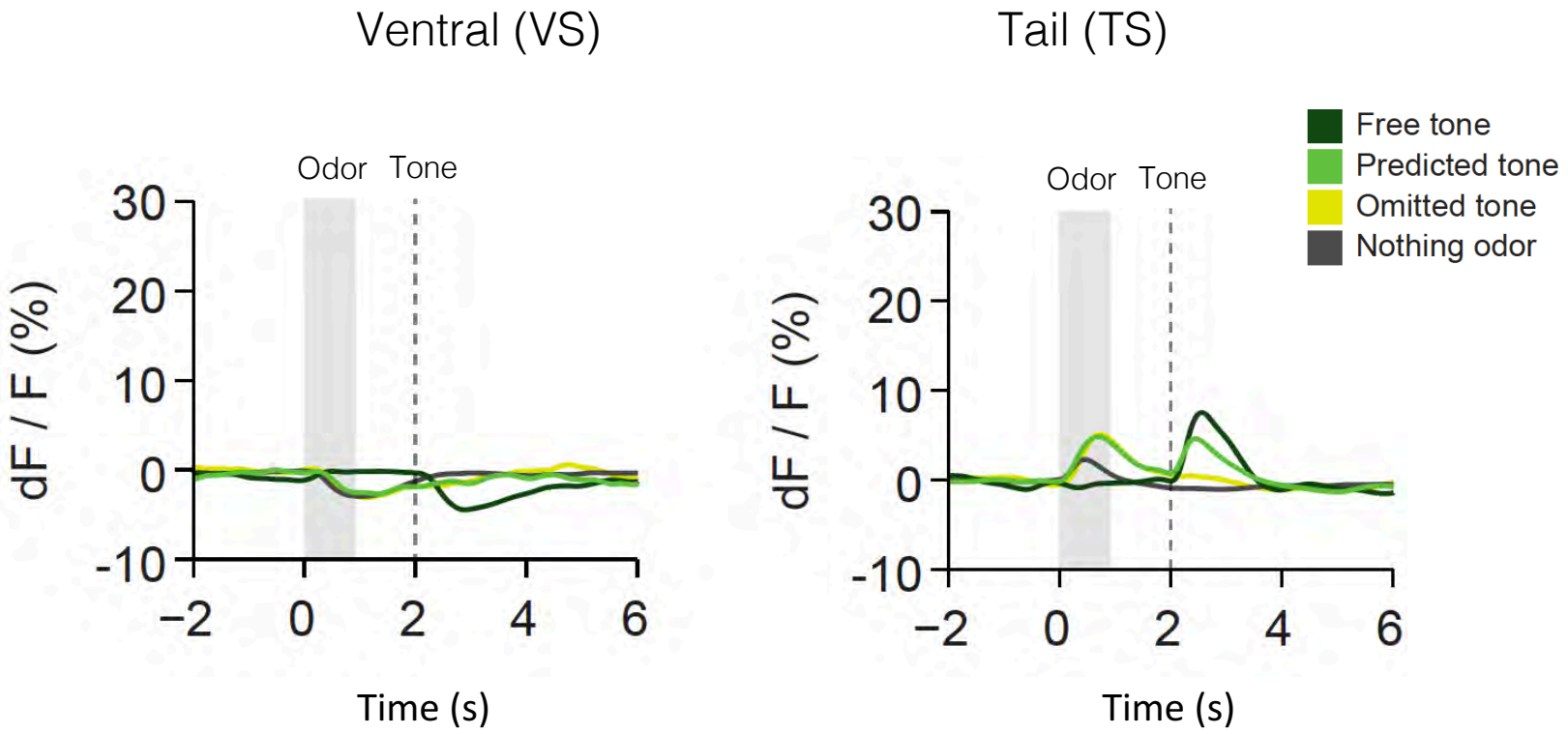
Responses to **air puff** are distinct among striatal areas



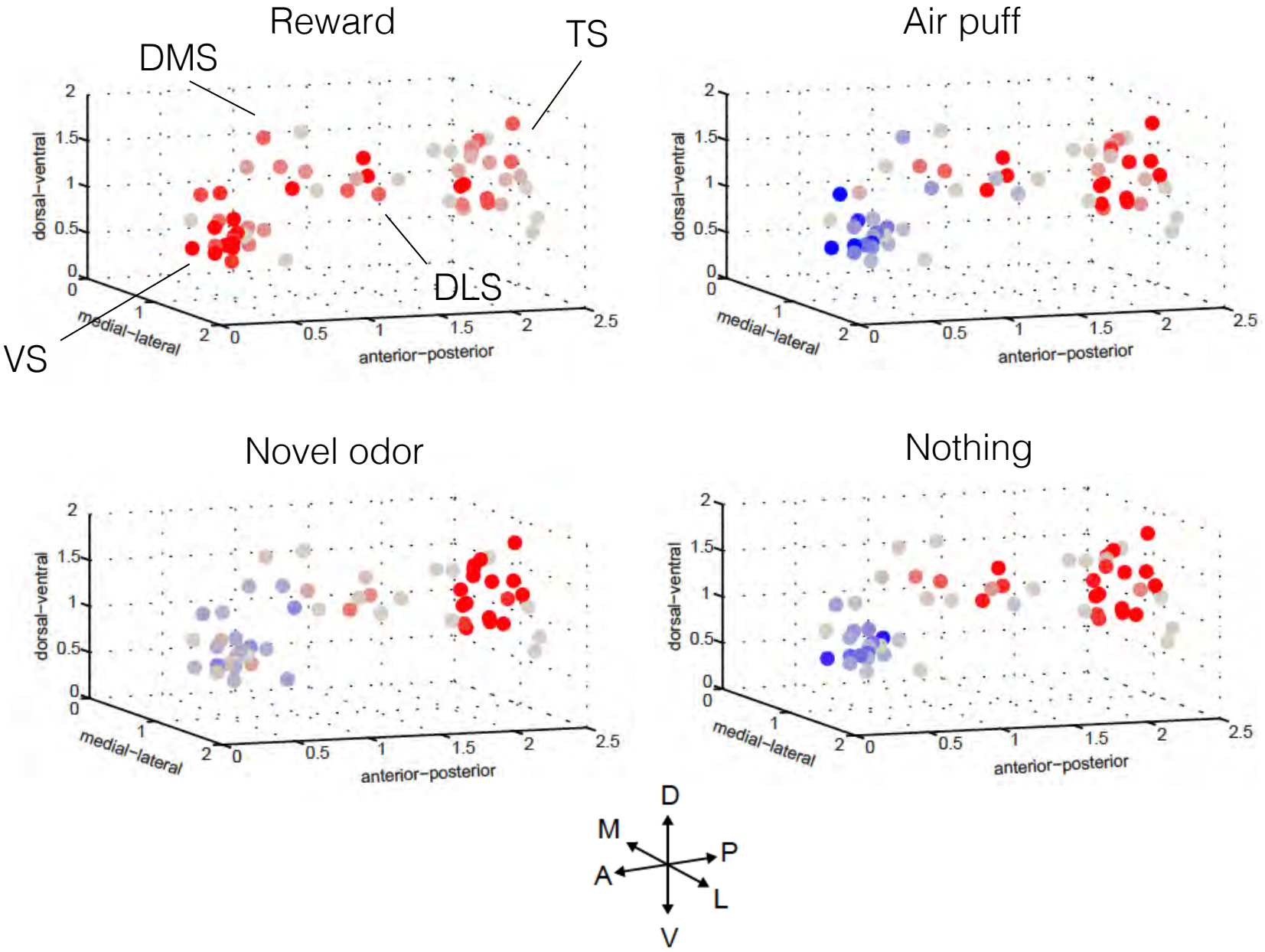
(Menegas, Babayan, Uchida, Watabe-Uchida, *eLife*, 2017)

Mild tones activate TS but not VS dopamine

~ 55 dB



(Menegas, Babayan, Uchida, Watabe-Uchida, *eLife*, 2017)


Reward
DMS
TS
DLS
VS
Air puff
Novel odor
Nothing
dorsal–ventral
medial–lateral
anterior–posterior
Response
(dF / F)
20%
0%
-20%
D
M
A
P
L
V

Conclusion

- Dopamine neurons projecting to the ventral striatum (VS) signal canonical value prediction errors.

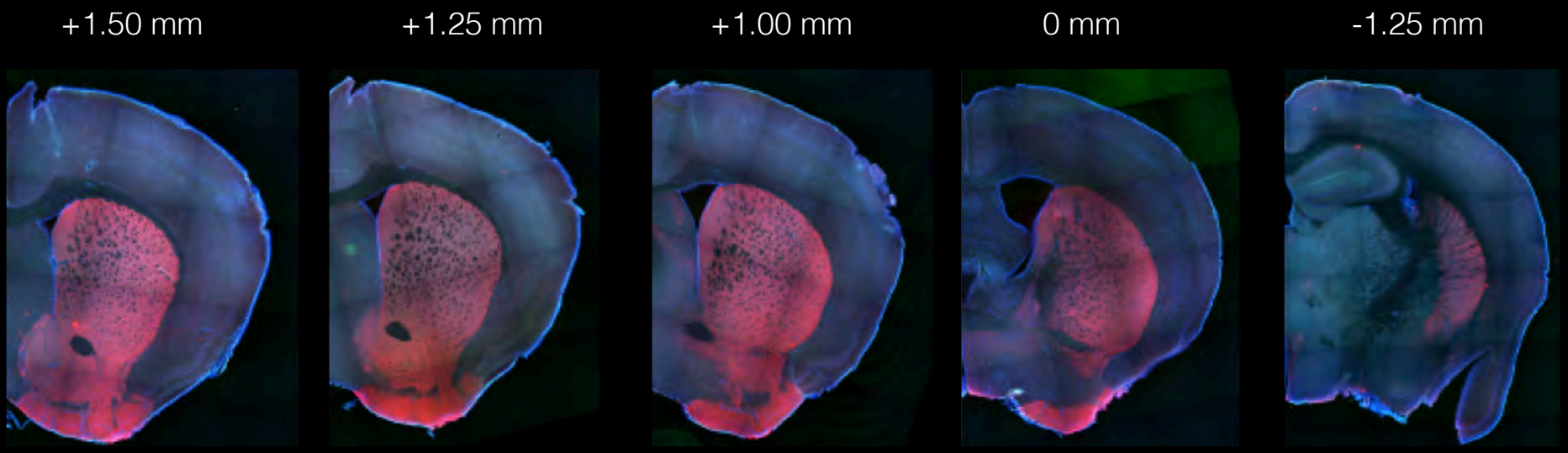
  (Schultz, Glimcher, Phillips, Roitman, Cheer etc.)

- Dopamine neurons projecting to the tail of the striatum (TS) form a distinct population both anatomically and functionally.
- TS dopamine is activated by salient stimuli

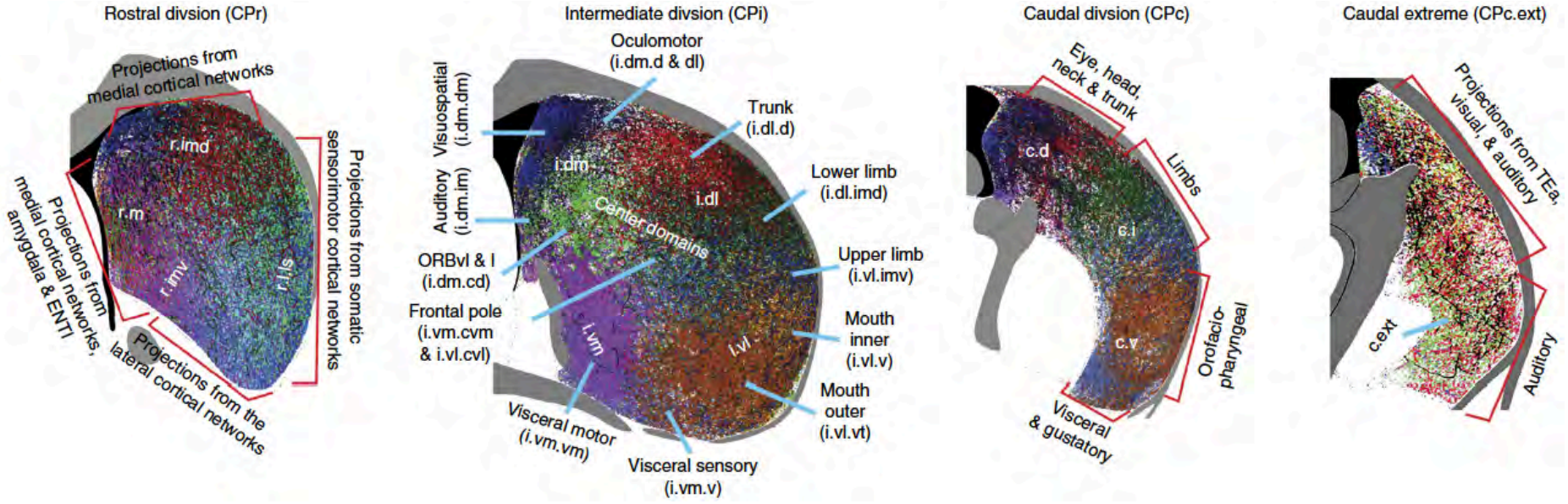
  (cf. lateral SNc: Matsumoto & Hikosaka, 2009)

- Novel stimuli activate TS but not VS dopamine
- Dopamine neurons' responses to novel stimuli can be understood as a part of general salience signal in TS but unlikely to be a positive value of exploration ("novelty bonus").

(Menegas, Babayan, Uchida, Watabe-Uchida, *eLife*, 2017)


+1.50 mm
+1.25 mm
+1.00 mm
0 mm
-1.25 mm
DAT/mCherry

The tail of the striatum (TS) receives inputs from sensory cortices



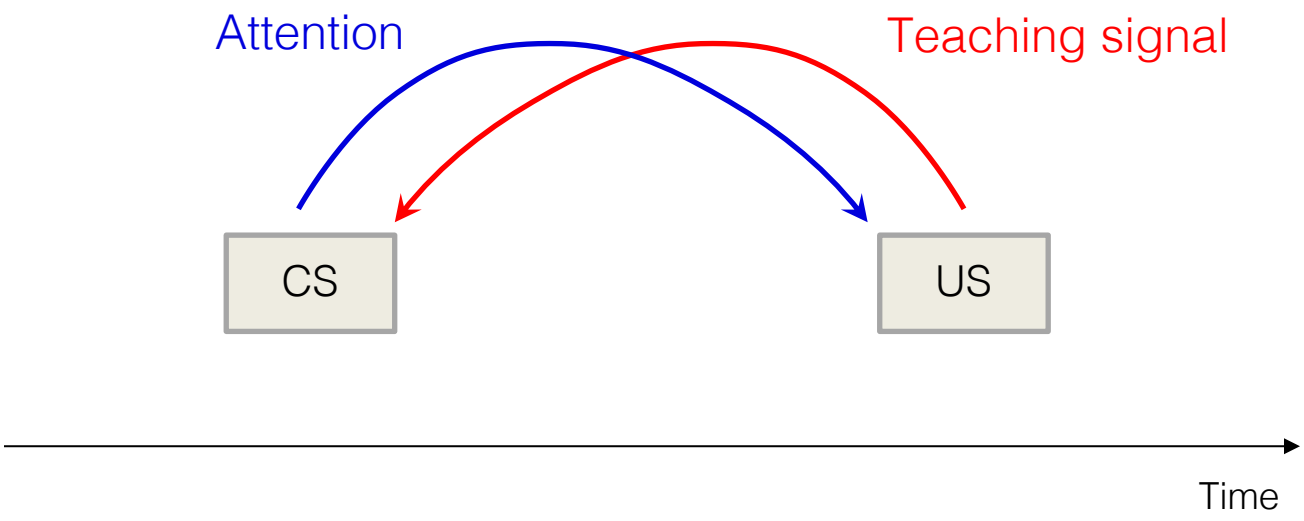
Cortex		Striatum
Visual Auditory Temporal association (TEa) Entorhinal, lateral Perirhinal	➡	TS

(Hintiryan et al., 2016; also see Hunnicutt et al. 2016)

- **Associability** (the effectiveness of associative learning)
  - **US associability** $\propto US\ unpredictability$

    (Widrow and Hoff, 1960; Rescorla & Wagner, 1972)
  - **CS associability** $\propto CS\ unpredictability\ or\ US\ unpredictability$

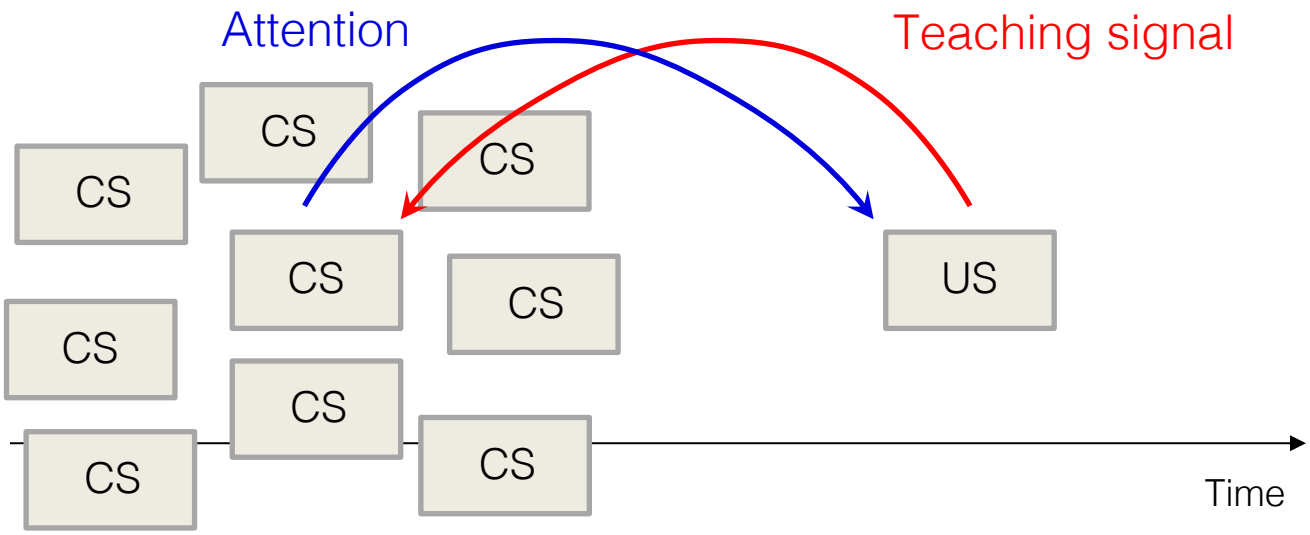
    (Macintosh, 1975; Wagner, 1978; Pearce & Hall, 1980)



- **Associability** (the effectiveness of associative learning)
  - **US associability** $\propto US\ unpredictability$

    (Widrow and Hoff, 1960; Rescorla & Wagner, 1972)
  - **CS associability** $\propto CS\ unpredictability\ or\ US\ unpredictability$

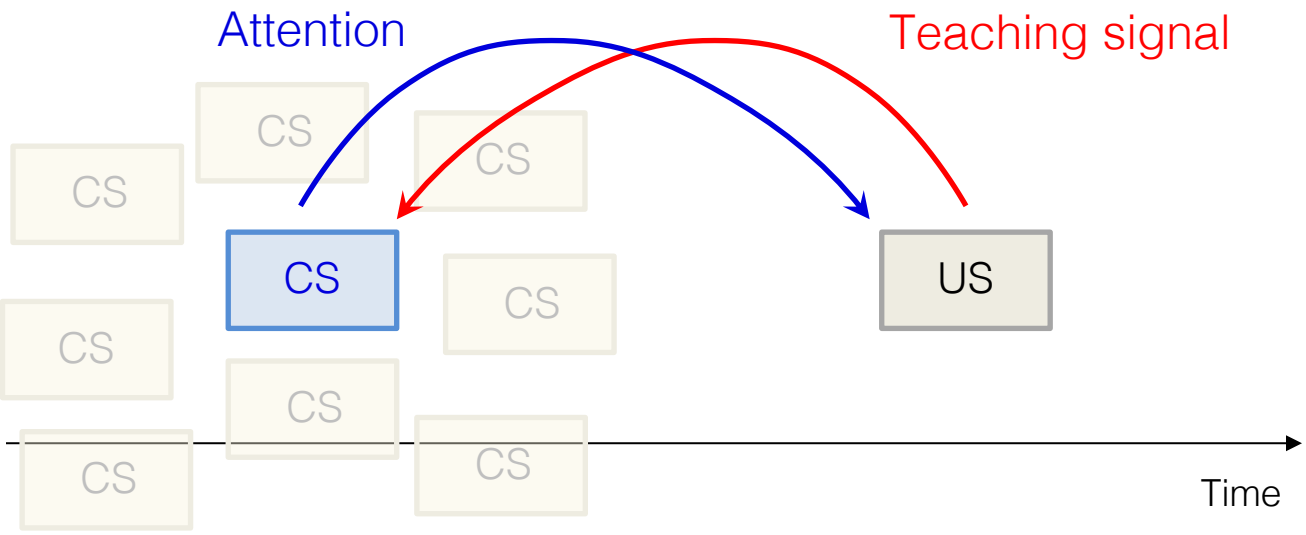
    (Macintosh, 1975; Wagner, 1978; Pearce & Hall, 1980)



- **Associability** (the effectiveness of associative learning)
  - **US associability** $\propto US\ unpredictability$

    (Widrow and Hoff, 1960; Rescorla & Wagner, 1972)
  - **CS associability** $\propto CS\ unpredictability\ or\ US\ unpredictability$

    (Macintosh, 1975; Wagner, 1978; Pearce & Hall, 1980)



- **Associability** (the effectiveness of associative learning)
  - **US associability** $\propto US\ unpredictability$

    (Widrow and Hoff, 1960; Rescorla & Wagner, 1972)

  - **CS associability** $\propto CS\ unpredictability\ or\ US\ unpredictability$

    (Macintosh, 1975; Wagner, 1978; Pearce & Hall, 1980)

- **VS dopamine**: US associability (reward prediction error)
- **TS dopamine**: CS associability (attention)
- TS attention system may help select behaviorally-relevant stimuli to increase the efficiency of reinforcement learning.

(Menegas, Babayan, Uchida, Watabe-Uchida, *eLife*, 2017)

Distinct dopamine signals in striatal regions

- Ventral striatum
  - Value prediction error (VPE)
- Dorsal striatum (dorsomedial and dorsolateral)
  - Mixture of VPE, salience and motion*(?)
- Tail of the striatum
  - General salience (attention)
    - Attention for associative learning (CS associability?)

*cf. Jin and Costa (2010)
Howe and Dombeck (2016)
Parker et al. (2016)
Barter et al. (2015)
Also see, Lerner et al. (2015)

(Menegas, Babayan, Uchida, Watabe-Uchida, *eLife*, 2017)

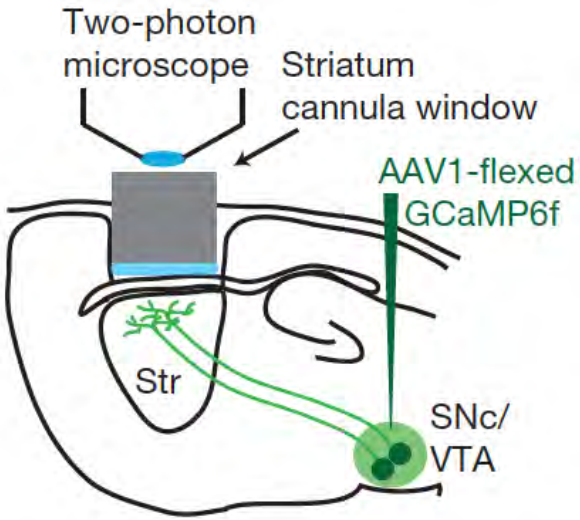
ARTICLE

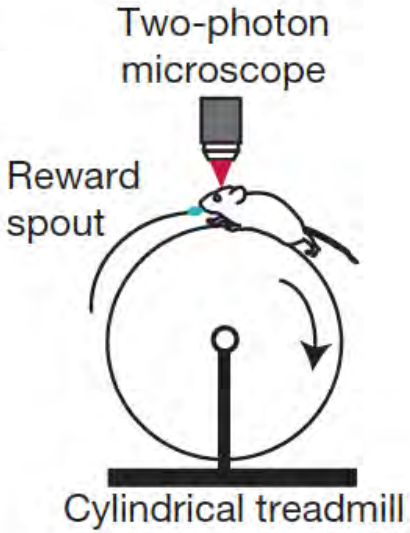
NATURE 2016

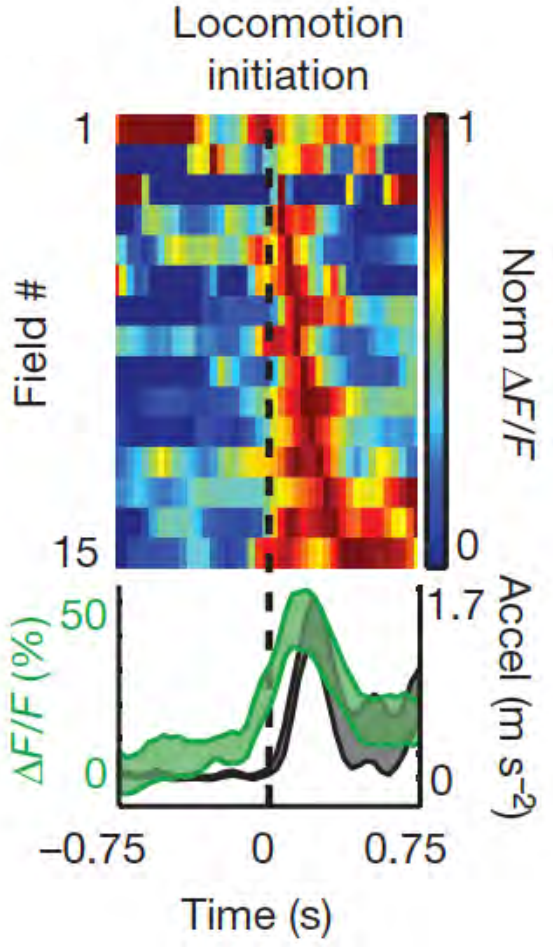
doi:10.1038/nature18942

Rapid signalling in distinct dopaminergic axons during locomotion and reward

M. W. Howe[1] & D. A. Dombeck[1]







Dopamine as a TD error

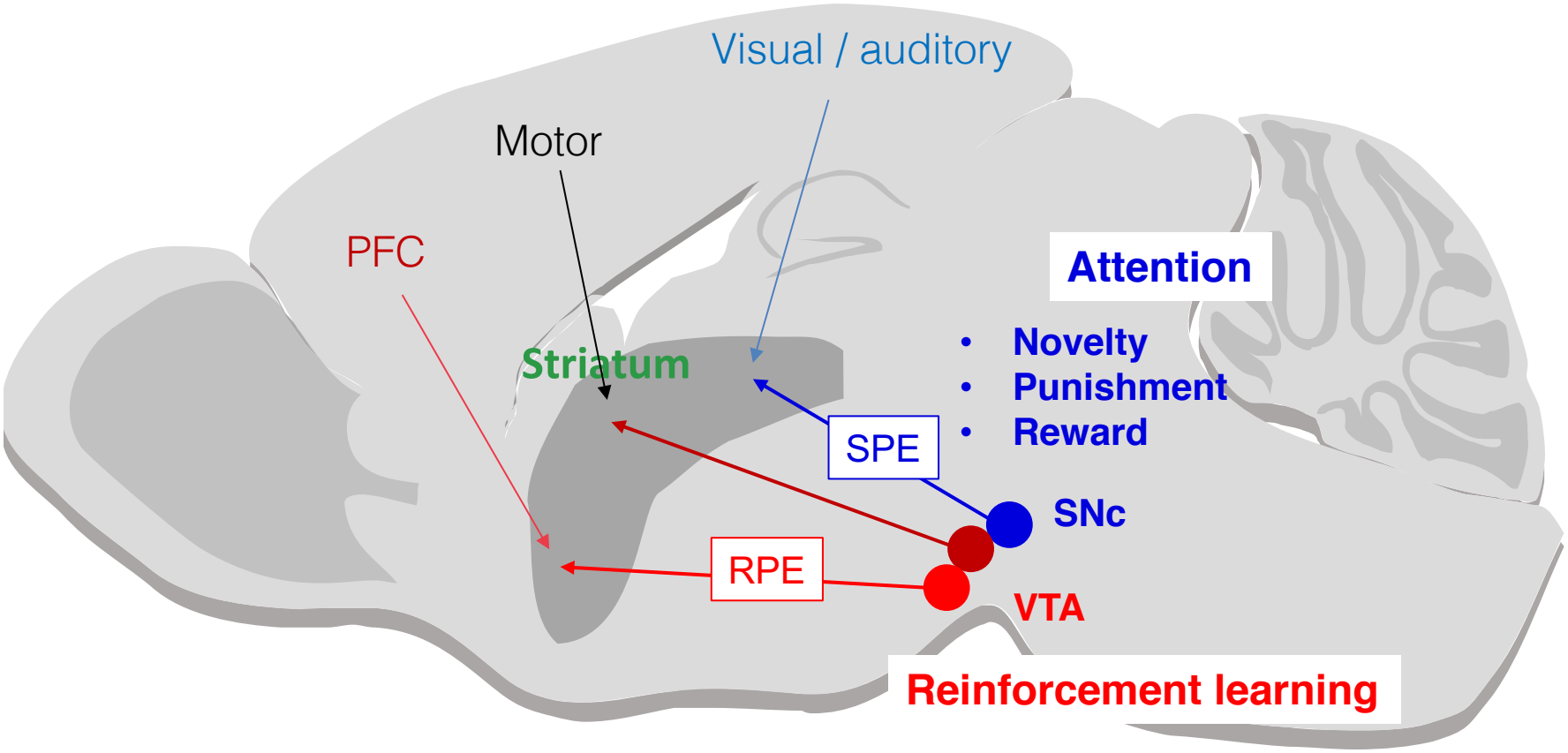
- Supporting evidence
- Minor problems
- Serious problems

Future questions

- Defining dopamine signals at different targets.
- What does dopamine do in each target?

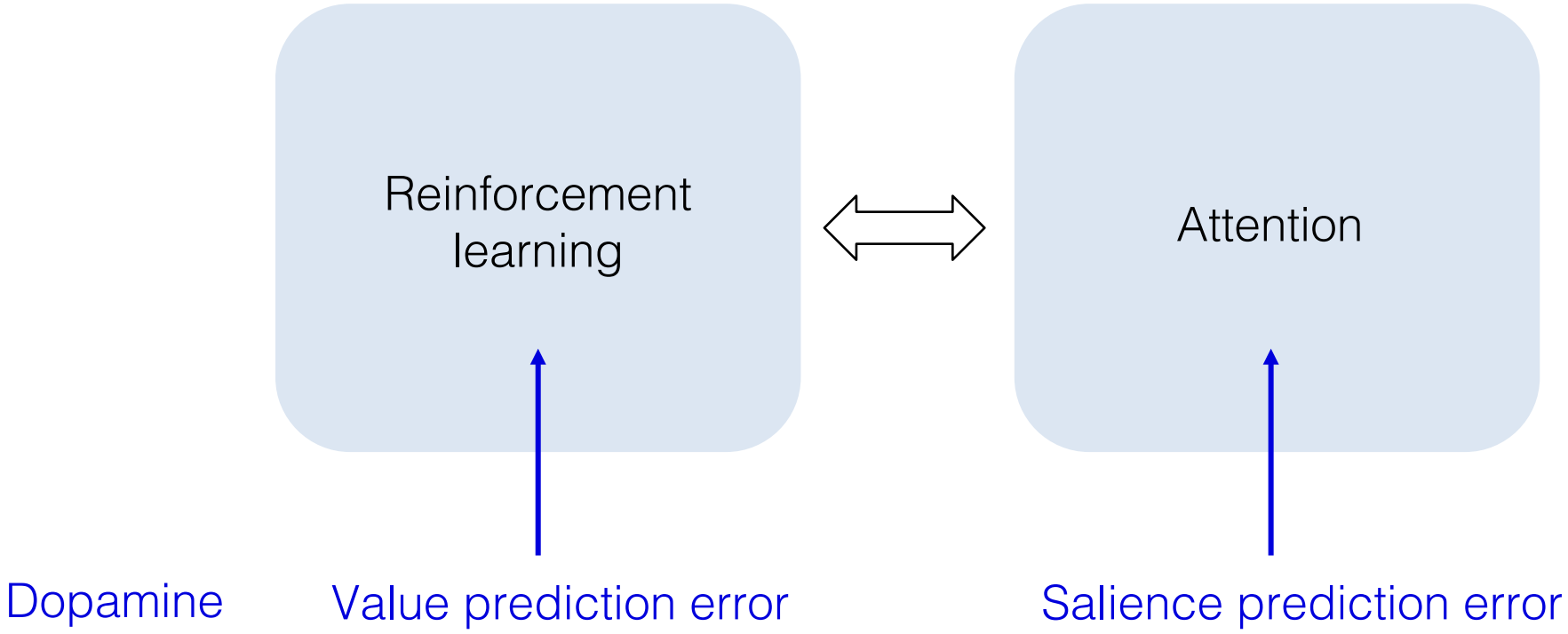
  Regulate plasticity, ongoing activity
- Do different regions of the striatum use the same plasticity rule?
- What are the functions of dopamine in each target?

Distinct cortico-basal ganglia systems



SPE: salience prediction error

Dopamine diversity



(Menegas, Babayan, Uchida, Watabe-Uchida, *eLife*, 2017)

Acknowledgements

Uchida lab:

Jeremiah Cohen (Johns Hopkins)

Sebastian Haesler (Leuven University)

Hideyuki Matsumoto (Osaka City University)

Neir Eshel (Stanford University)

Michael Bukwich

Vinod Rao

Vivian Hemmelder

Ju Tian

Sachie Ogawa

William Menegas

Benedicte Babayan

Mitsuko Watabe-Uchida

Samuel Gershman

Christian Machens

Dmitry Kobak

Reagents

Edward Callaway (rabies virus)

Fumitaka Osakada (rabies virus)

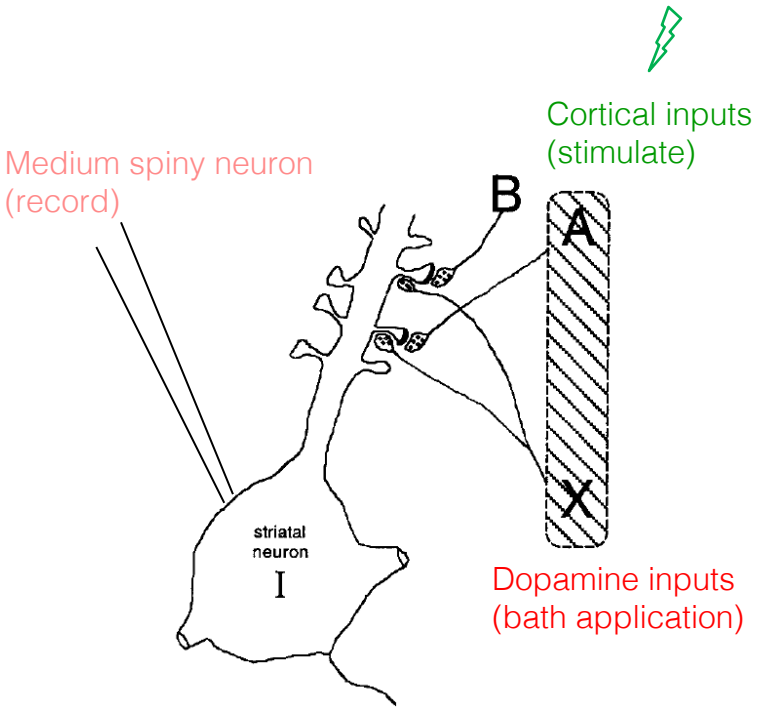
Bradford Lowell (vGAT-Cre)

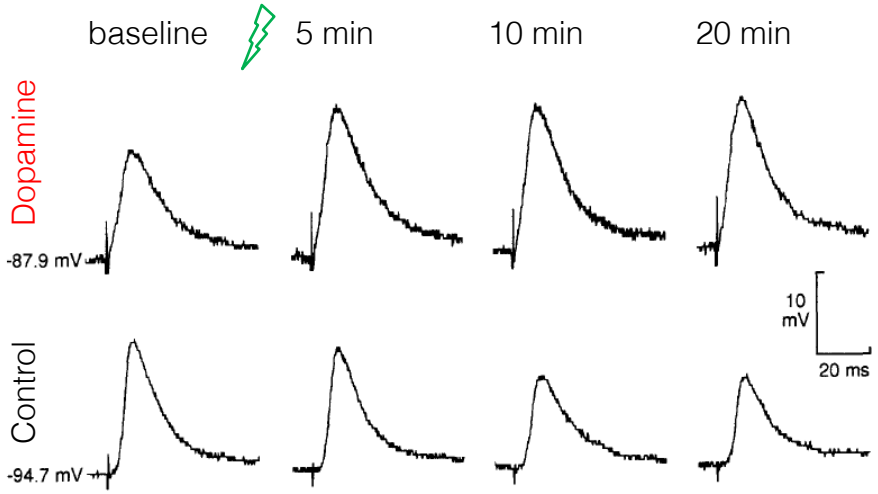
Linh Vong (vGAT-Cre)

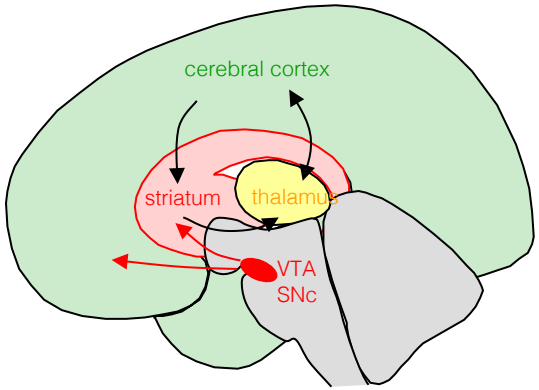
Funding

- National Institute of Mental Health

Dopamine regulates neural plasticity in the striatum







- membrane potential depolarization

  (EPSP: excitatory postsynaptic synaptic potentials)

- With dopamine, depressing synapse gets facillitatory

(Wickens et al., 1996)

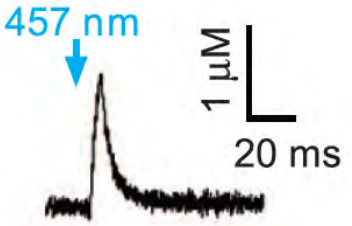
SCIENCE, 2014, VOL. 345 ISSUE 6204

NEUROSCIENCE

A critical time window for dopamine actions on the structural plasticity of dendritic spines

Sho Yagishita,[1,2] Akiko Hayashi-Takagi,[1,2,3] Graham C.R. Ellis-Davies,[4] Hidetoshi Urakubo,[5] Shin Ishii,[5] Haruo Kasai[1,2]*

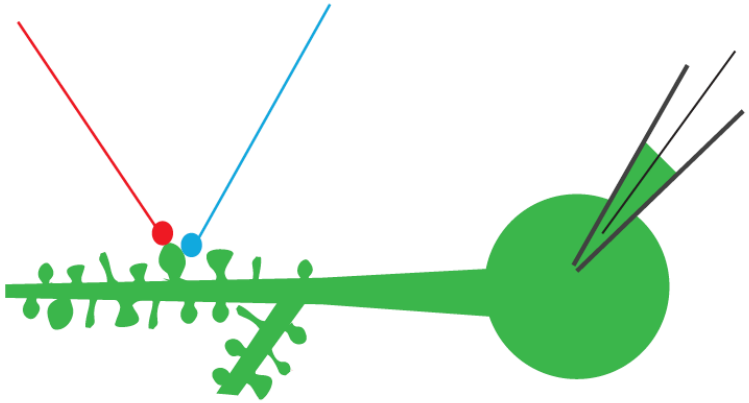
Optogenetics



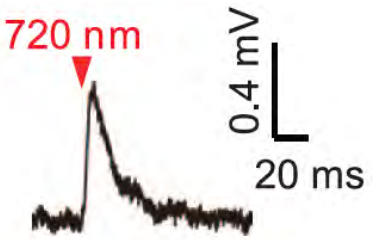
Amperometry

Glutamatergic input (Glu-uncaging)

Dopamine input (ChR2)

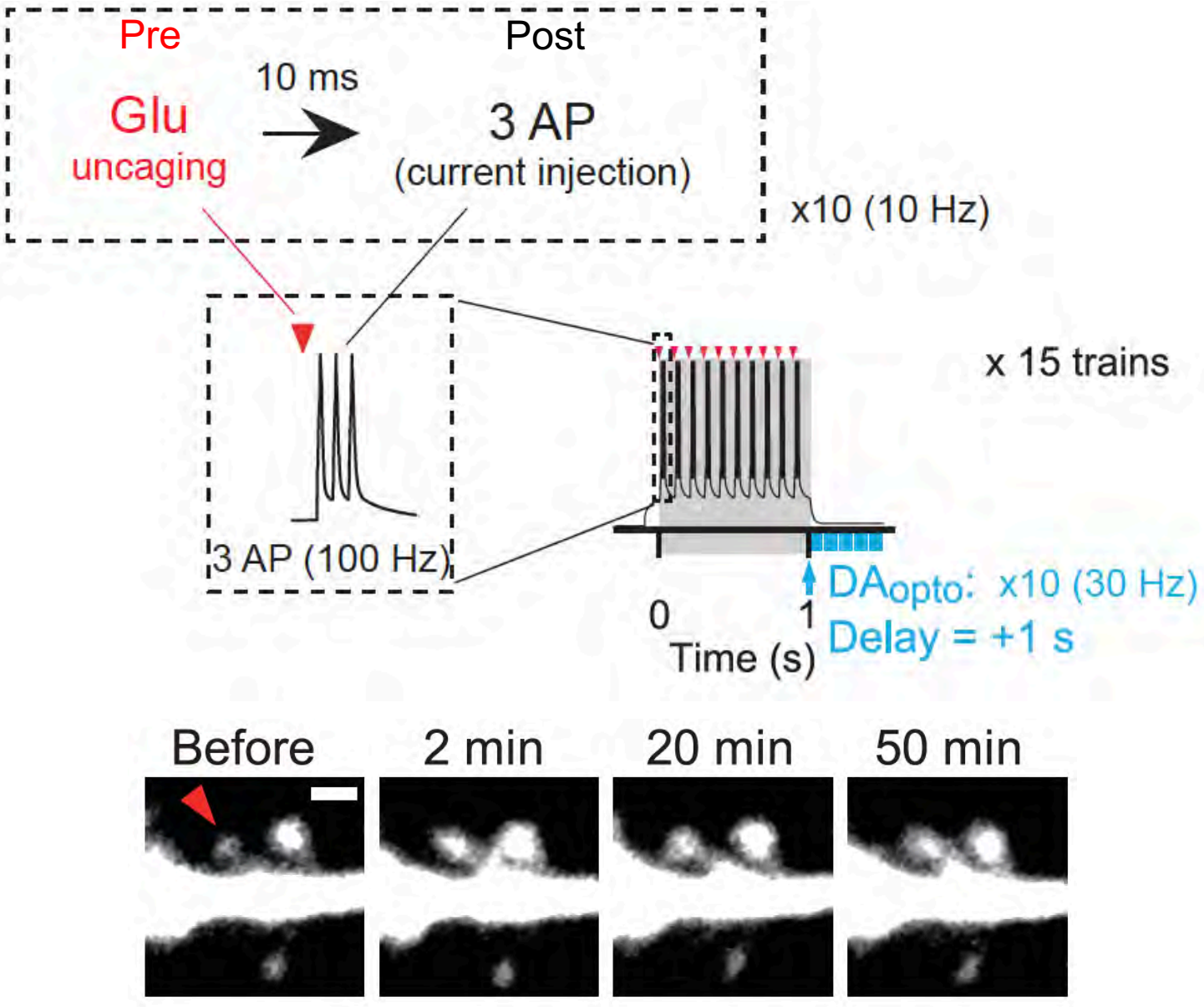


Uncaging

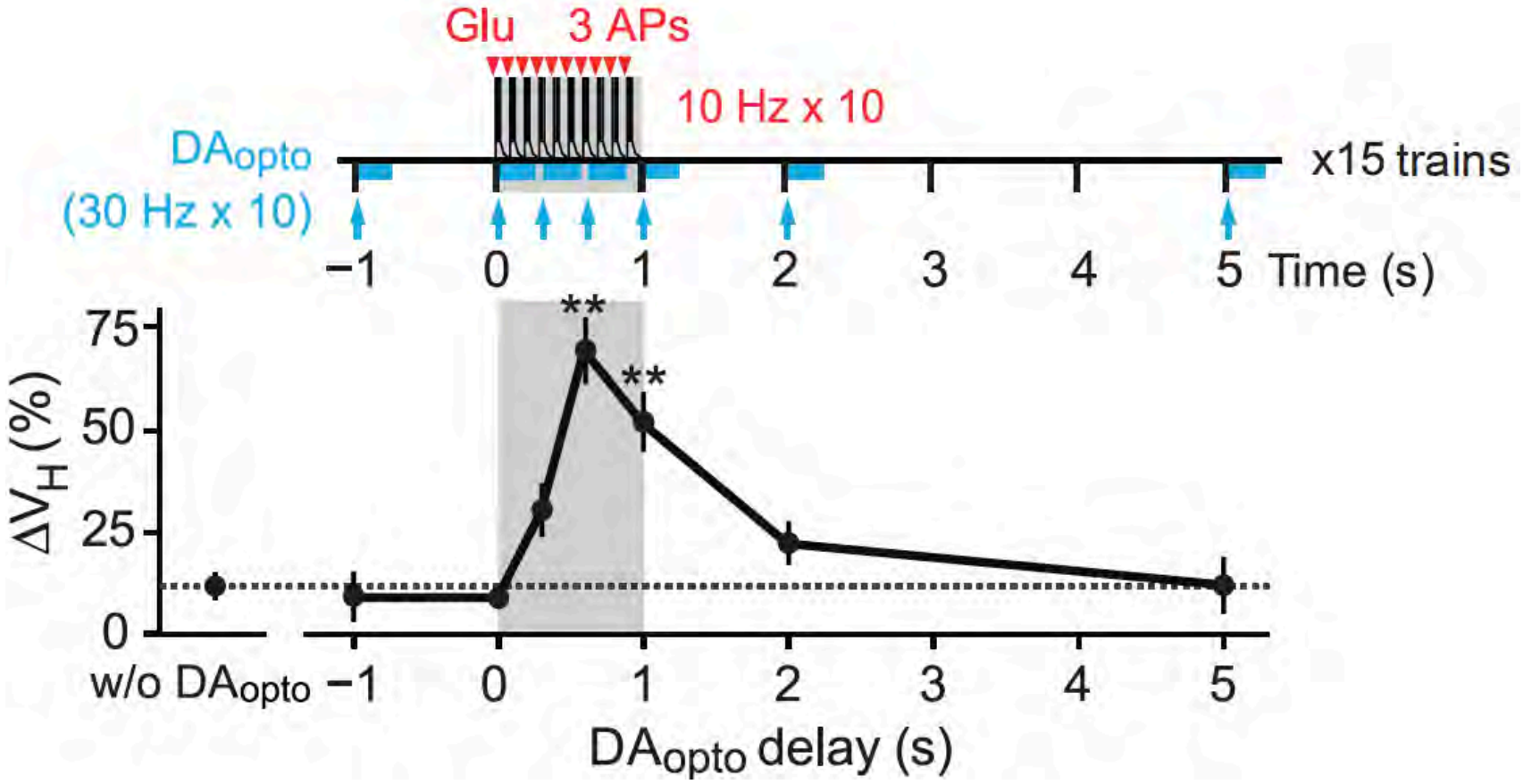


EPSP

- STDP (spike-timing dependent plasticity)



A window for dopamine reinforcement



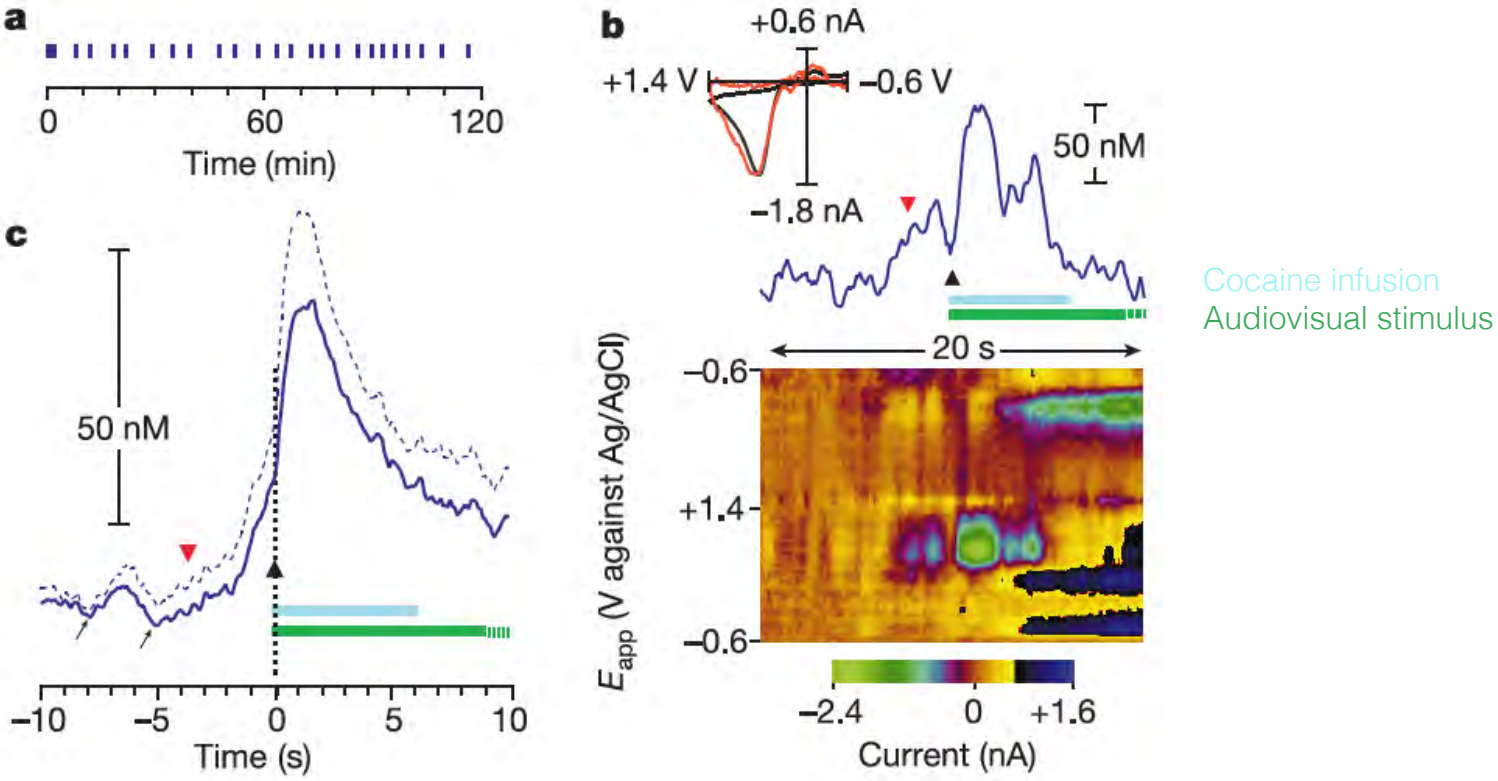
Non-canonical firing patterns

Subsecond dopamine release promotes cocaine seeking

Paul E. M. Phillips*†‡, **Garret D. Stuber**‡§, **Michael L. A. V. Heien**†, **R. Mark Wightman**†‡§ & **Regina M. Carelli***§

* *Department of Psychology,* † *Department of Chemistry,* ‡ *Neuroscience Center, and* § *Curriculum in Neurobiology, University of North Carolina, Chapel Hill, North Carolina 27599, USA*

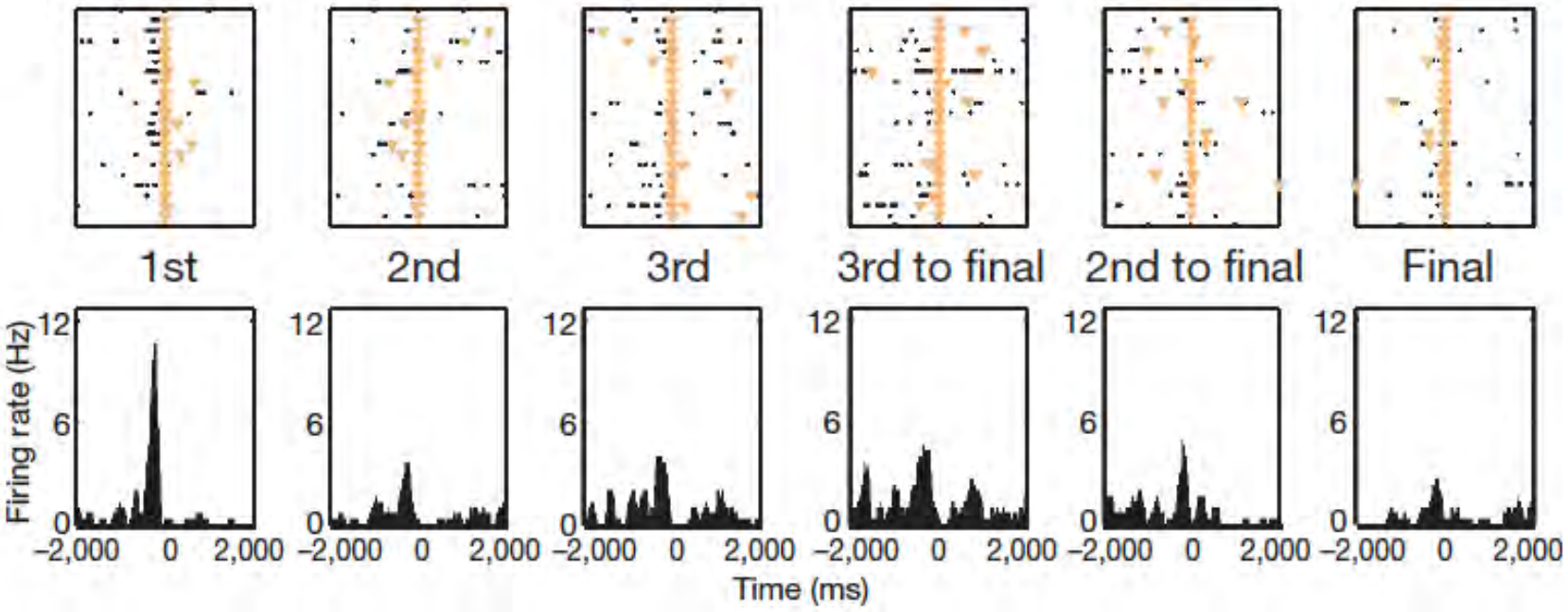
- Dopamine increase before initiation of movement



- Cyclic voltametry

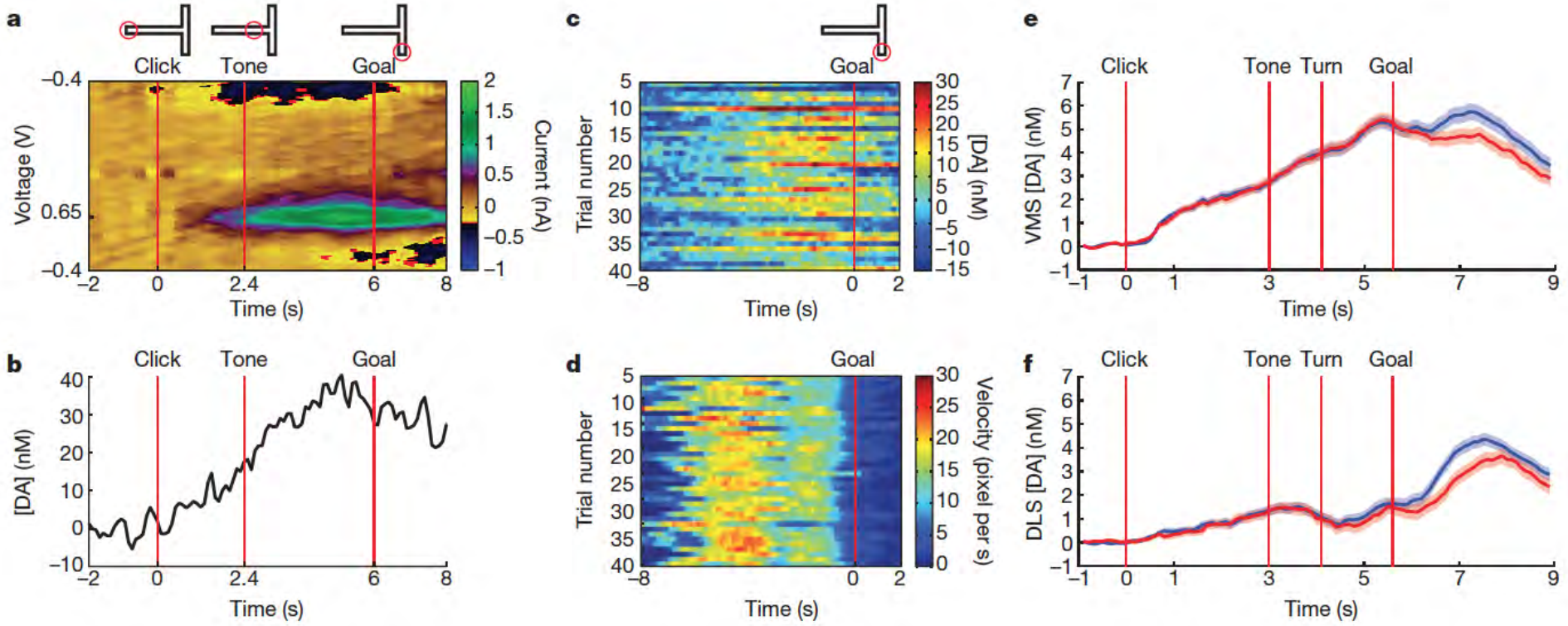
Start/stop signal

- Mouse performing lever pressing with a fix-ratio (8) schedule



(Xin and Costa, 2010)

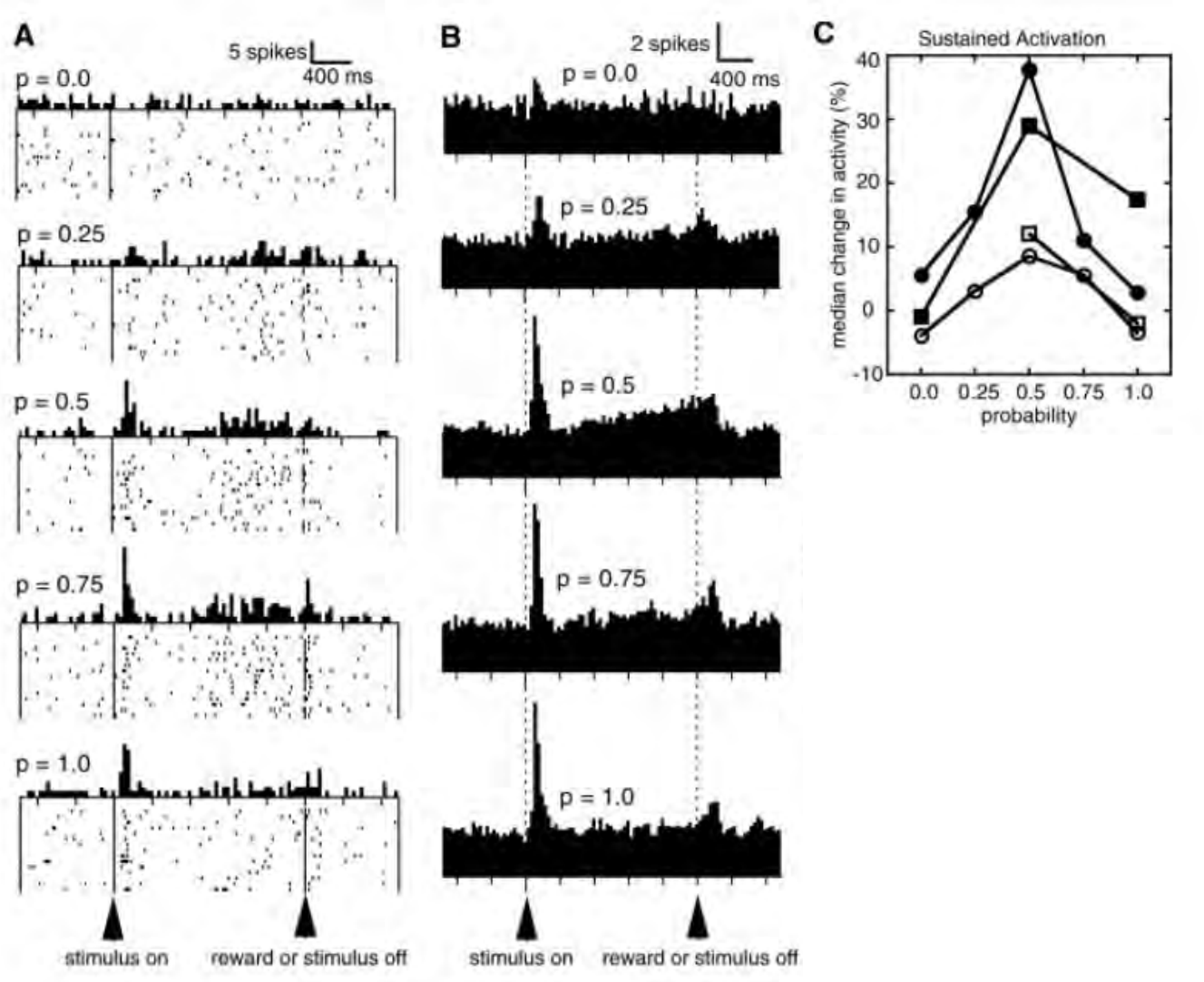
Ramping dopamine



- Freely-moving rats
- T-maze
- Cyclic voltammetry

(Howe et al., 2013)

Ramping dopamine



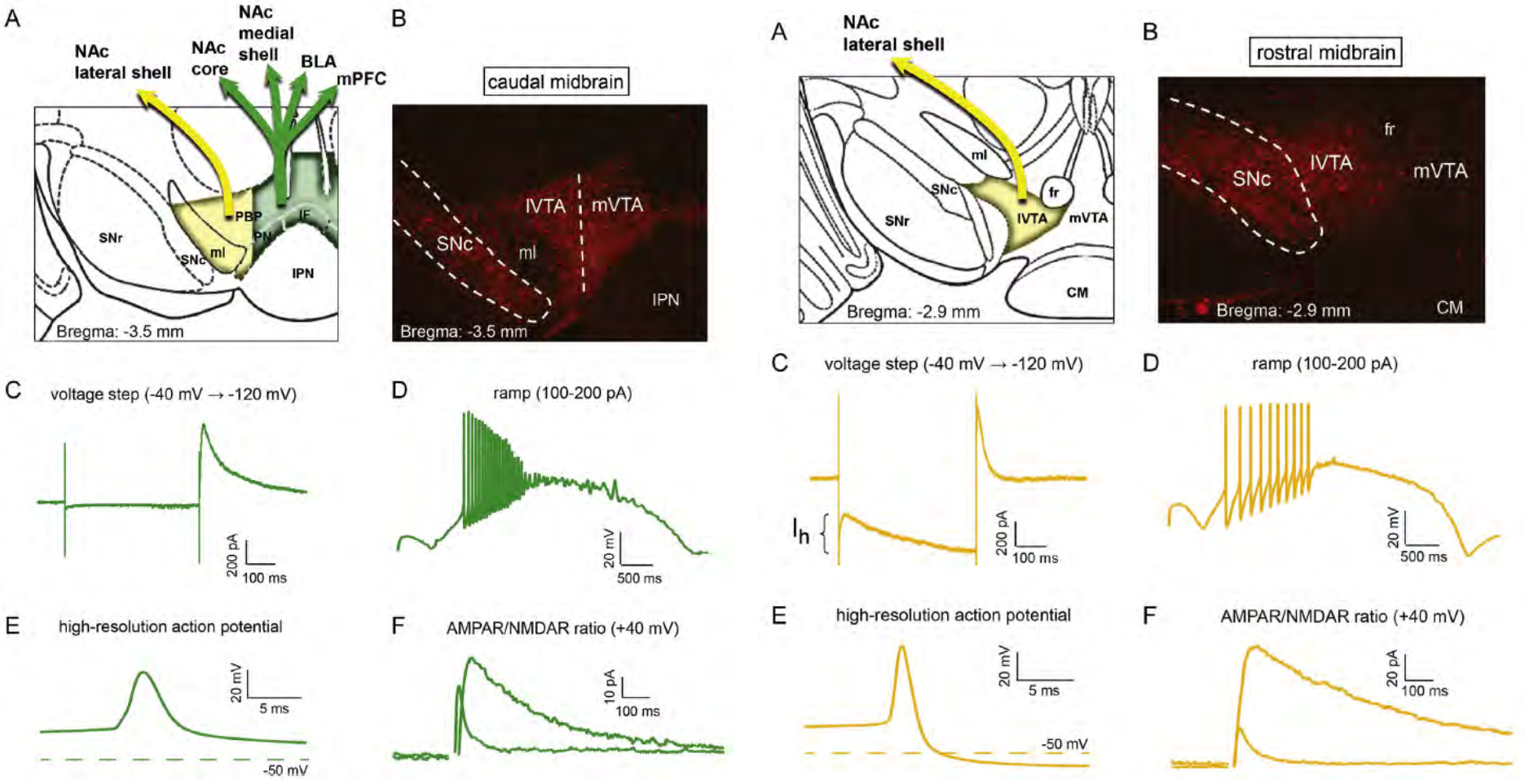
(Fiorillo et al., 2003)

Neuron 57, 760–773, March 13, 2008

Unique Properties of Mesoprefrontal Neurons within a Dual Mesocorticolimbic Dopamine System

Stephan Lammel,[1,3] Andrea Hetzel,[3,4] Olga Häckel,[3,4] Ian Jones,[3,5] Birgit Liss,[2,3,*] and Jochen Roeper[1,3,*]

Dopamine neurons projecting to different areas have different properties



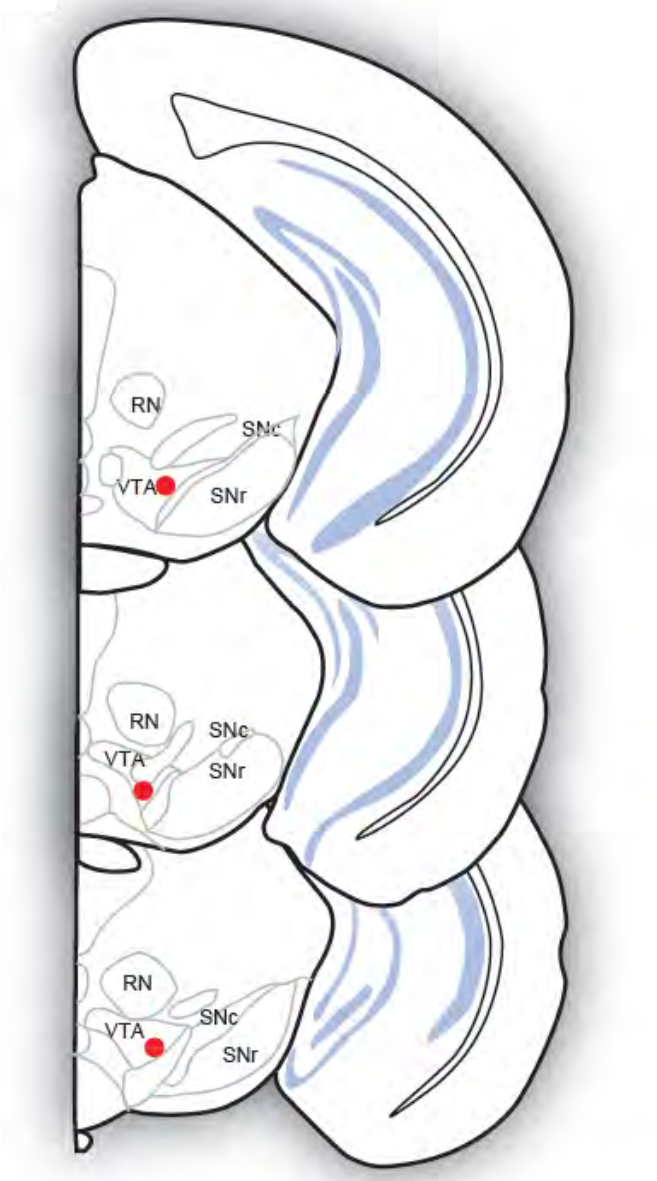
(Lammel et al. 2014)

Multiple dopamine hypothesis

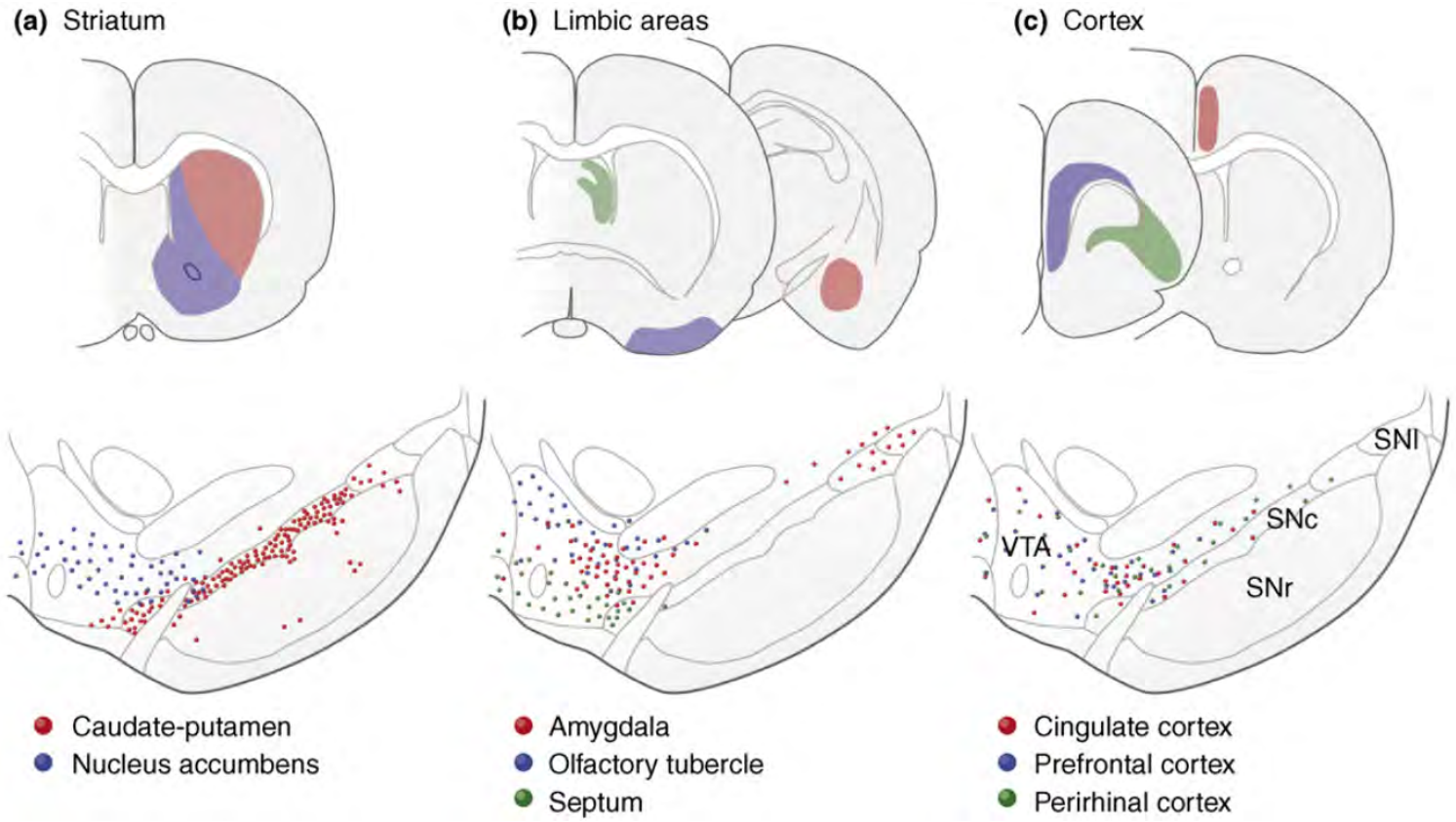
- Movement initiation/termination
- Reward
- Wanting, seeking
- Pleasure, hedonic
- Prediction error (learning, action selection)
- Salience/attention
- Incentive salience
- Motivation/energizing behavior
- Uncertainty
- Cost/benefit computation

(Wise, Berridge, Dayan, Niv, Schultz etc., ….)

Recording sites



Origins of dopamine projections



(Björklund & Dunnett, 2007)