Reinforcement learning in neuroscience - II

Rita Almeida

October 2018

Dopamine pathways

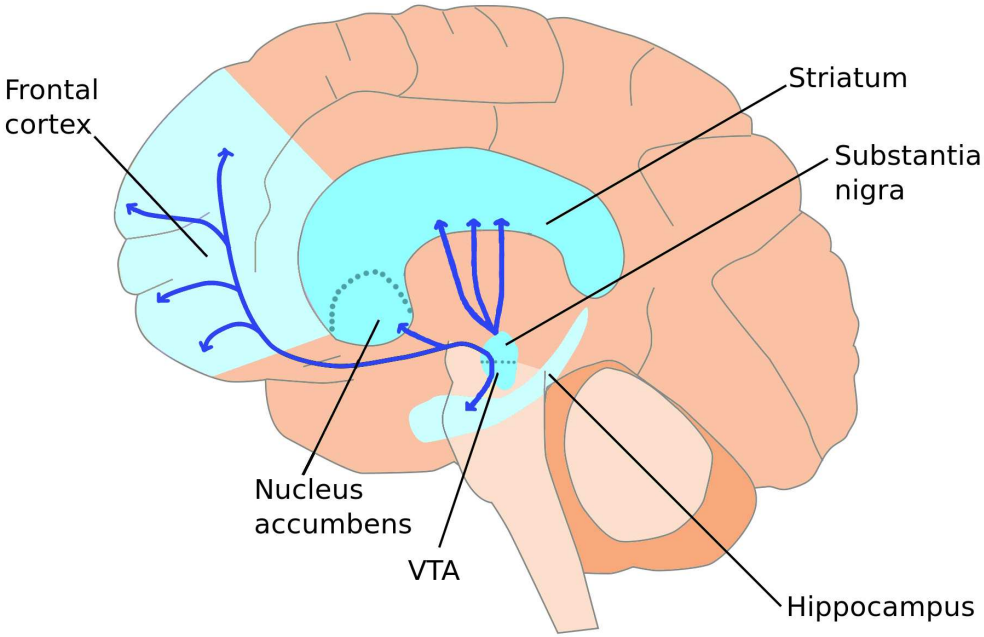


Figure from OIST (www.oist.jp)

Dopamine is thought to have a role in:

- movement control
- reward and motivation
  - substance abuse
- other:
  - working-memory
  - schizophrenia ...

- Dopamine was first hypothesized to be the reward system,
- now is associated with prediction error of reward learning.

Dopamine response

Dopamine release after stimulation of the axon:

- rise in extracelullar dopamine concentration of several nM in few ms,
- concentration becomes homogeneous on a sphere, max diffusion after 75 ms, over a diameter of 7-12 μm, 80 nM,
- reuptake takes concentration to baseline values after a few 100 ms.

Dopamine activity after a reward predicting event:

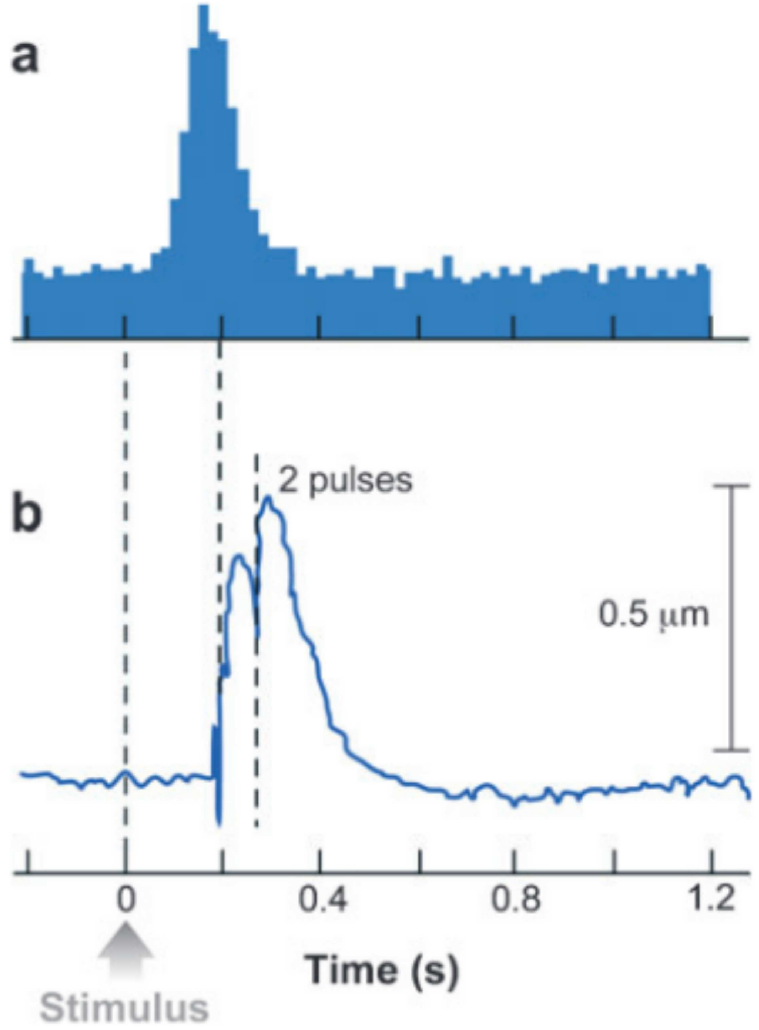


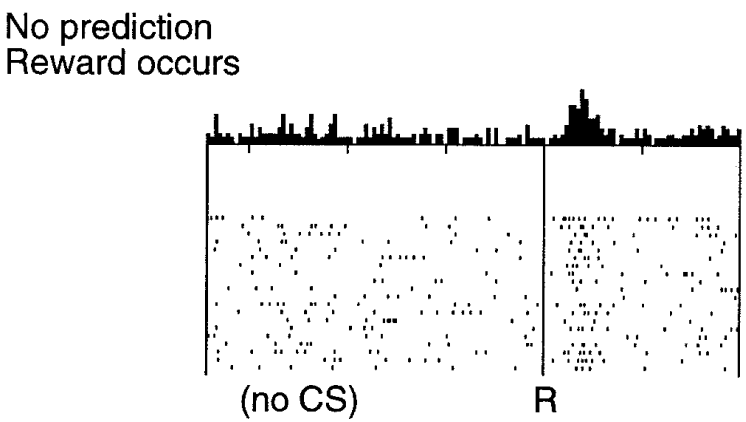
Figure from Schultz, 2007.

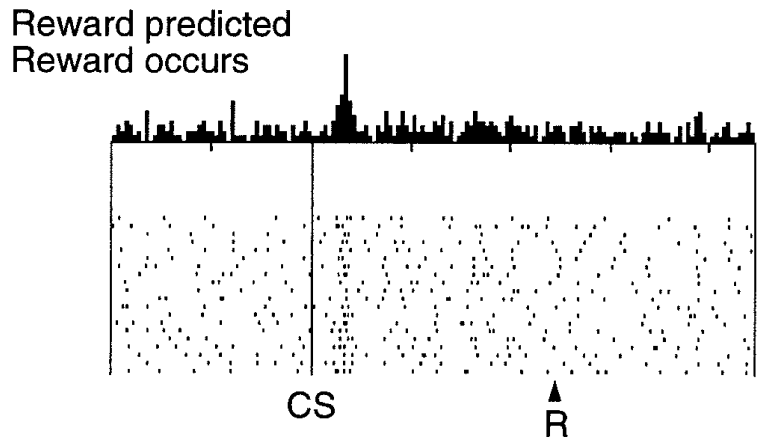
Dopamine and prediction error

Phasic activity of dopamine neurons proposed as reflecting prediction error:

- series of experiments by Schultz and coworkers,
- theoretical work by Montague, Dayan and coworkers.

- Before learning - reward unexpected
- After learning - reward expected
- After learning - no reward is unexpected





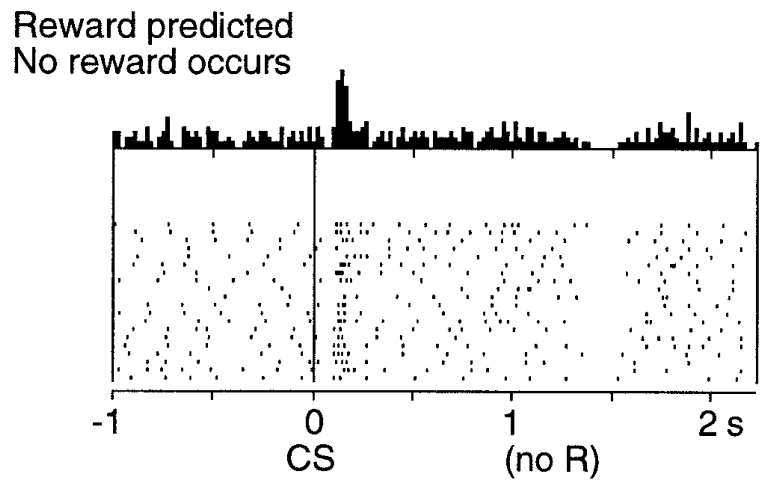


Figure adapted from Schultz 1998.

Dopamine and prediction error - second order conditioning

- Dopamine neurons response transfers to earliest predictive stimulus.

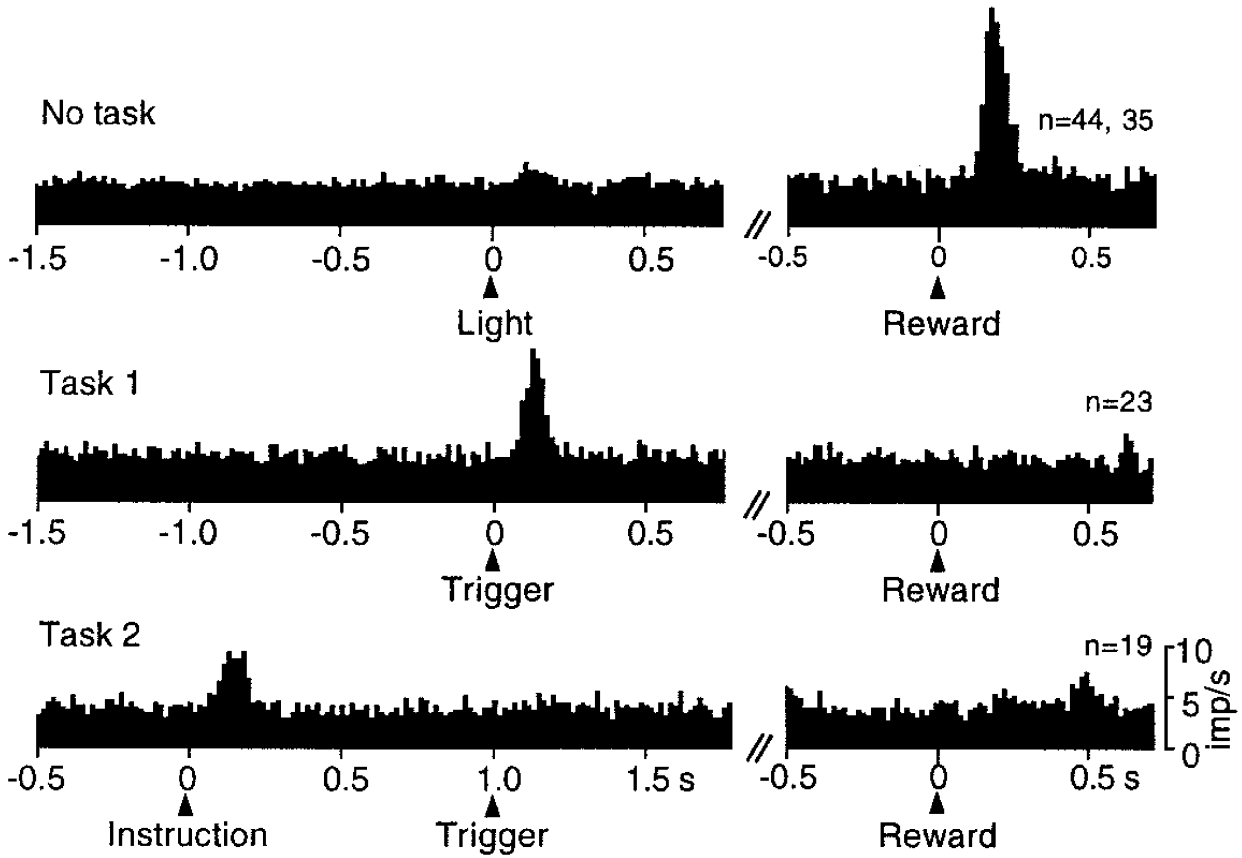


Figure adapted from Schultz 1998.

- Dopamine neurons activity reflects reward probability.

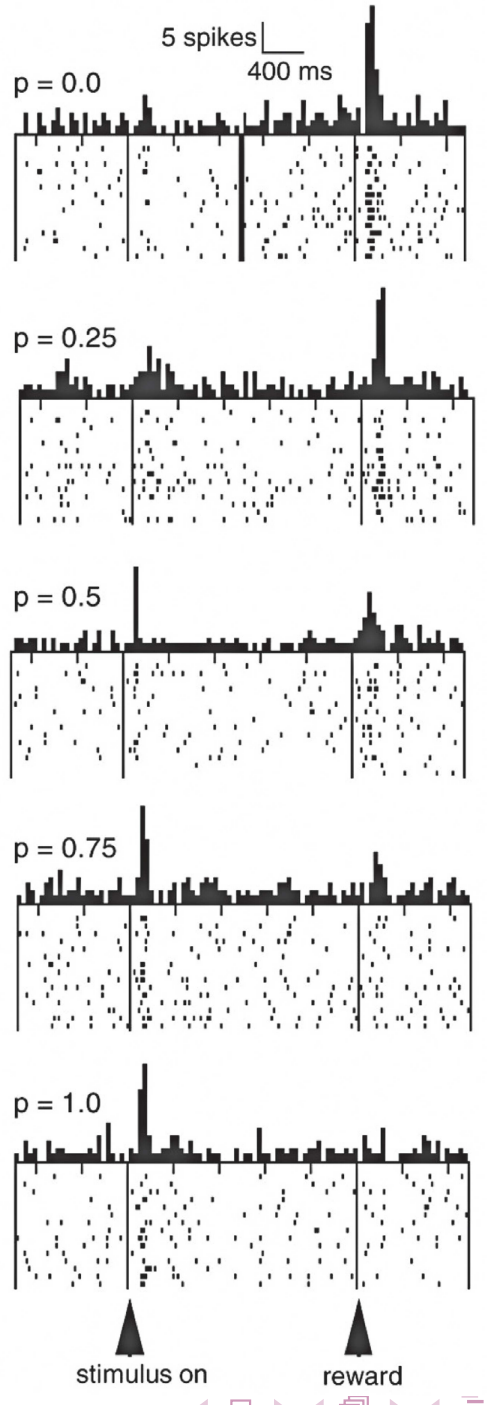
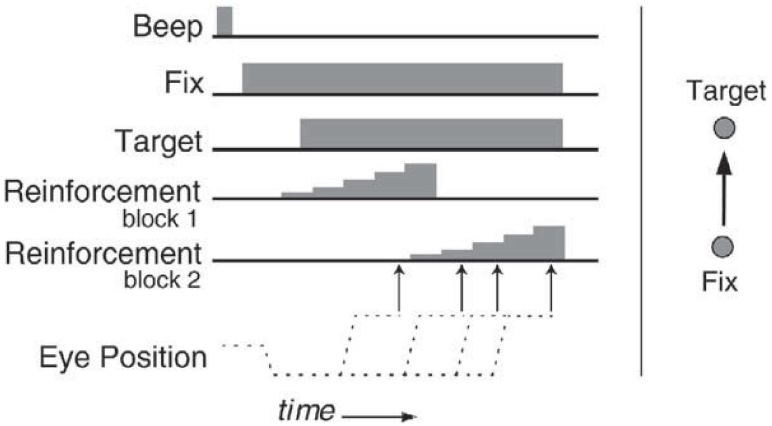


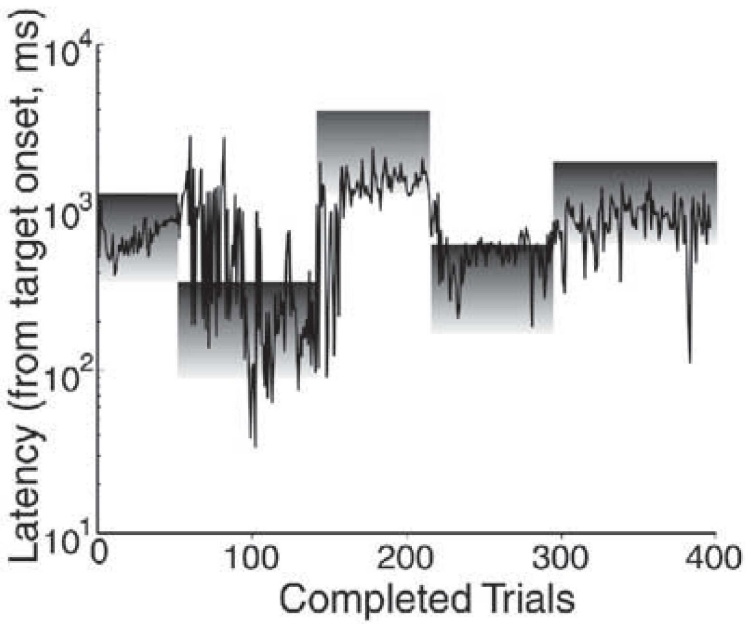
Figure adapted from Fiorillo et al 2003.

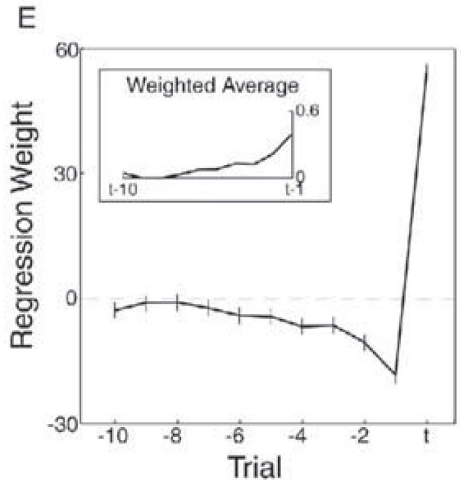
Dopamine and prediction error - history of reward

- Dopamine neurons activity reflects history of previous rewards (for reward higher than expected).



$$y = \beta_0 r_t + \beta_1 r_{t-1} + \cdots + \beta_{10} r_{t-10} + k$$





Figures adapted from Bayer and Glimcher 2005.

Causal role of dopamine neurons activity in RL

A causal link between prediction errors, dopamine neurons and learning

Elizabeth E Steinberg[1,2,11], Ronald Keiflin[1,11], Josiah R Boivin[1,2], Ilana B Witten[3,4], Karl Deisseroth[5–8] & Patricia H Janak[1,2,9,10]

VOLUME 16 | NUMBER 7 | JULY 2013 NATURE NEUROSCIENCE

Causal role of dopamine neurons activity in RL

Blocking

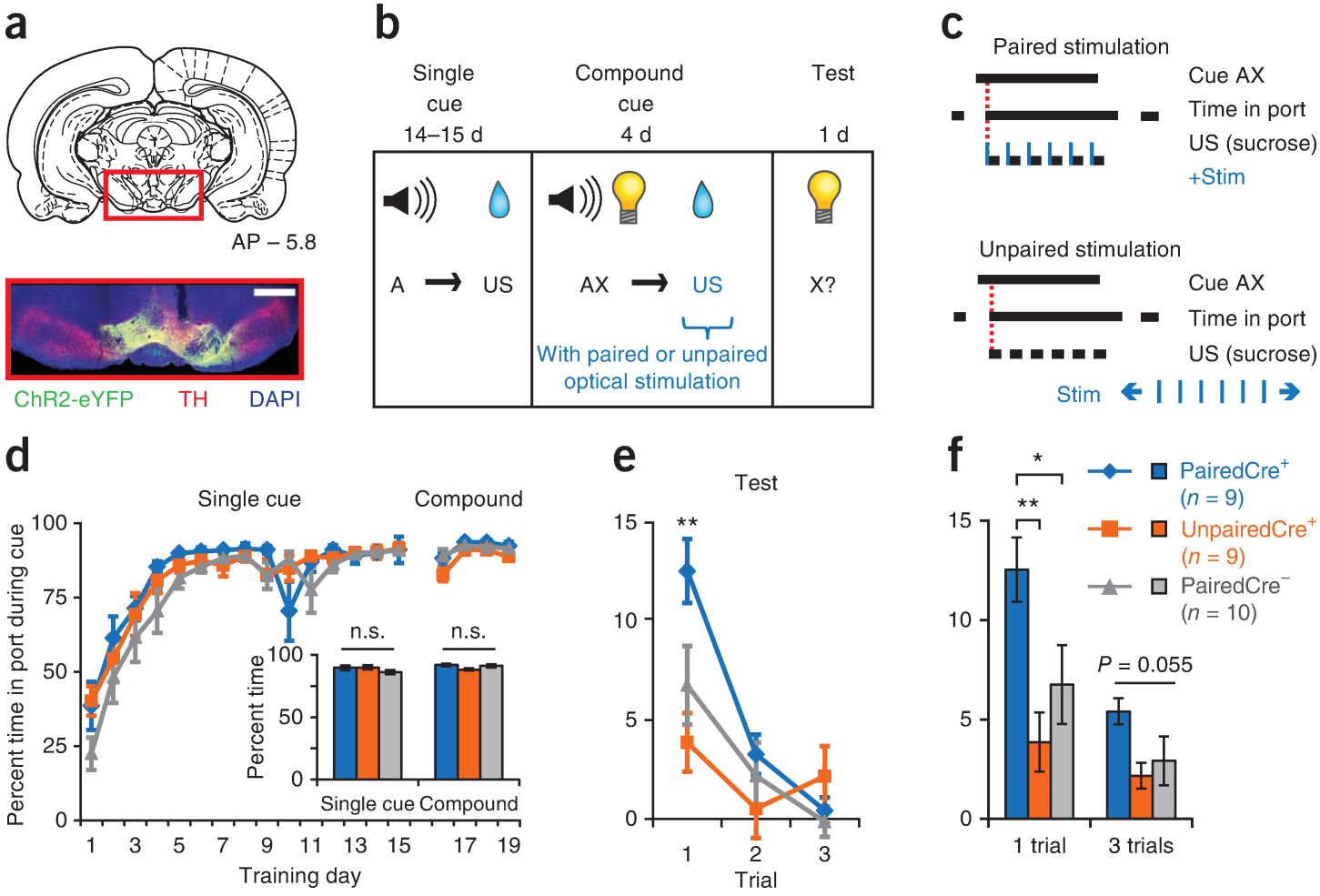


Figure from Steinberg et al. 2013.

Causal role of dopamine neurons activity in RL

Extinction

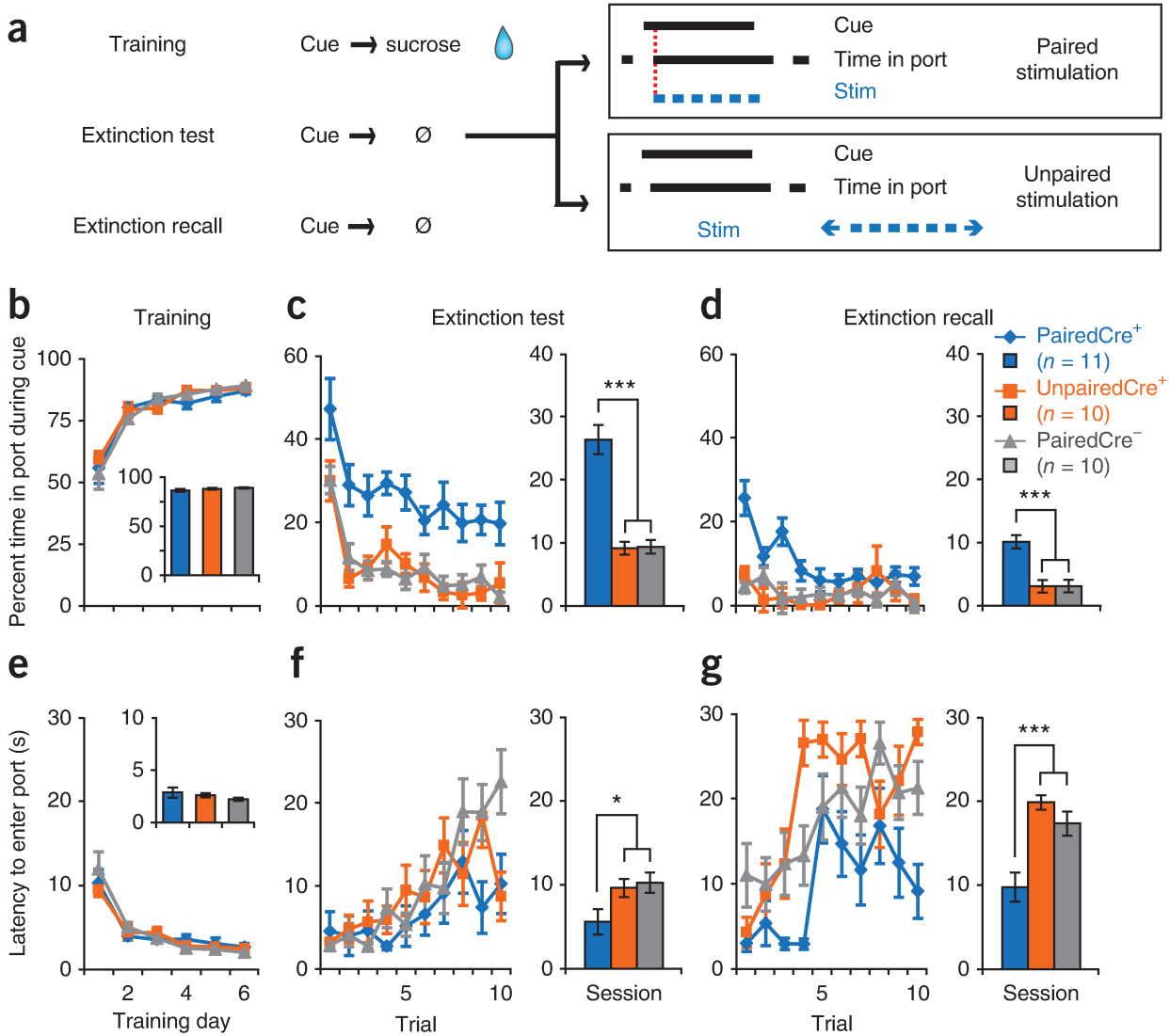


Figure from Steinberg et al. 2013.

Evidence for action selection signals in the brain

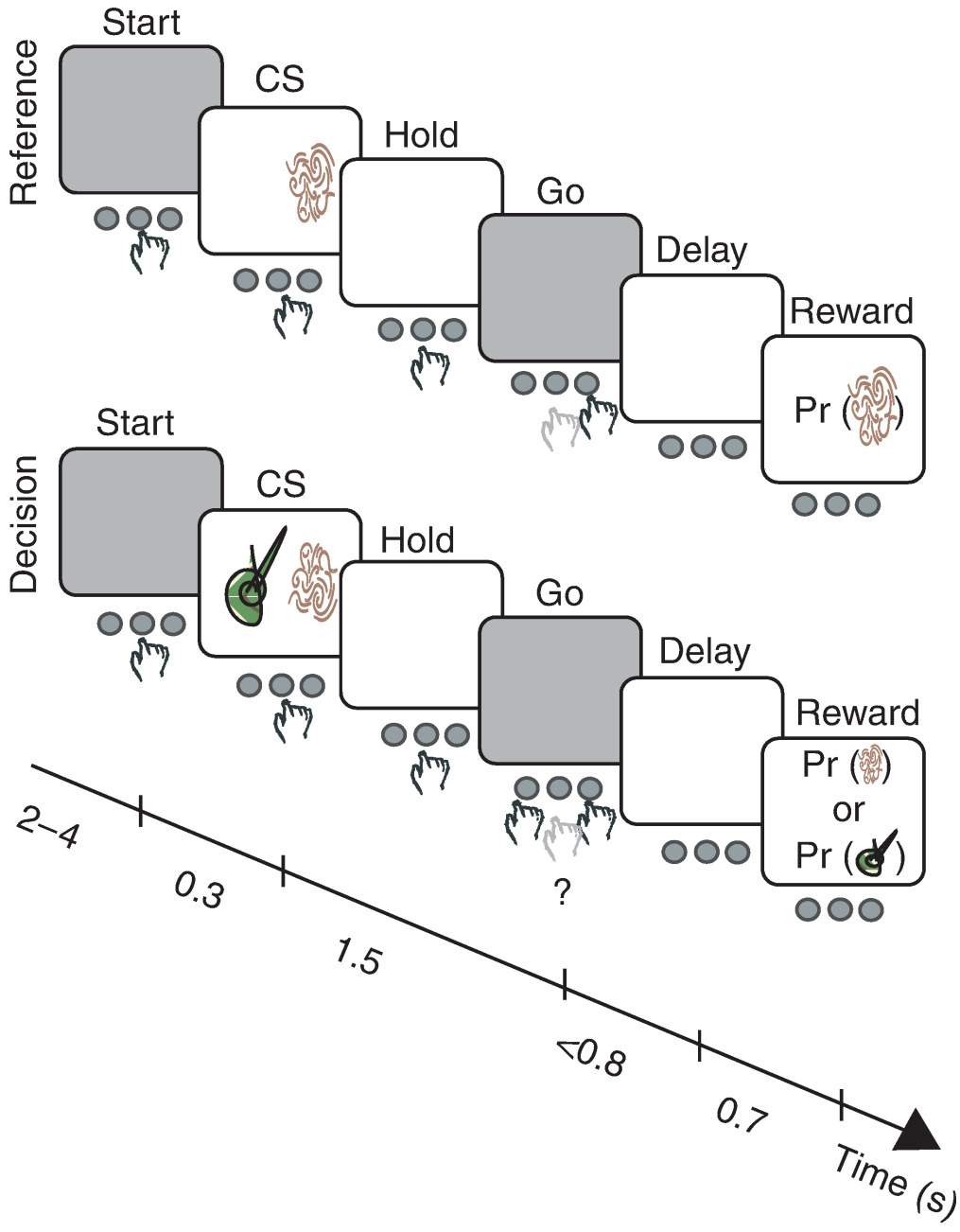
Midbrain dopamine neurons encode decisions for future action

Genela Morris[1,2], Alon Nevet[2], David Arkadir[2], Eilon Vaadia[1,2] & Hagai Bergman[1,2]

NATURE NEUROSCIENCE VOLUME 9 | NUMBER 8 | AUGUST 2006

Evidence for action selection

Figure from Morris et al. 2006.



Evidence for action selection

Probability of choosing one alternative in the decision trials as a function of quantities derived from the instructed trials:

- relative frequency of reward for that alternative,
- relative dopamine response.

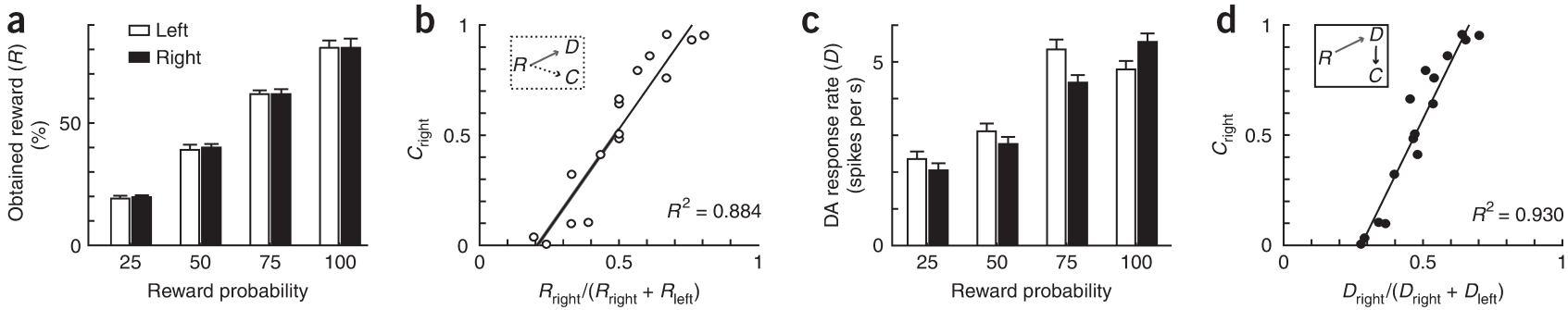


Figure from Morris et al. 2006.

Relative dopamine response predicts better the choice probabilities.

Evidence for action selection

Dopamine responses in decision trials separated according to chosen action.

Dopamine neurons encode values of future actions.

Filled circles: Decision action values.

Open circles: Reference action values.

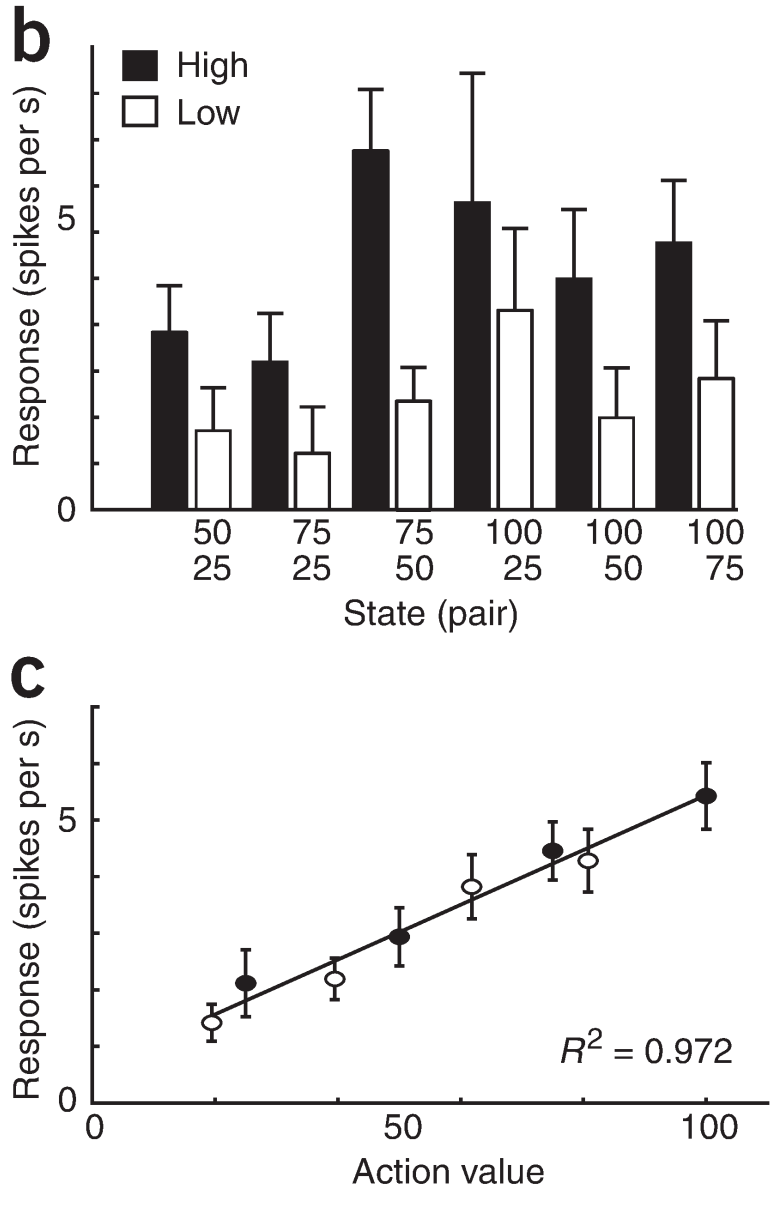


Figure from Morris et al. 2006.

Evidence for Actor/Critic

Dissociable Roles of Ventral and Dorsal Striatum in Instrumental Conditioning

John O'Doherty,[1]* Peter Dayan,[2] Johannes Schultz,[1]
Ralf Deichmann,[1] Karl Friston,[1] Raymond J. Dolan[1]

16 APRIL 2004 VOL 304 SCIENCE www.sciencemag.org

Actor/critic method

The critic evaluates using TD error:

$\delta_t = r_{t+1} + \eta V(s_{t+1}) - V(s_t)$

Given a_t, s_t:

- $\delta_t > 0 \rightarrow$ increase probability of selecting a
- $\delta_t < 0 \rightarrow$ decrease probability of selecting a

For example:

$p(s_t, a_t) \leftarrow p(s_t, a_t) + \beta\delta_t$

for a policy $\pi(s, a) = p(a|s)$

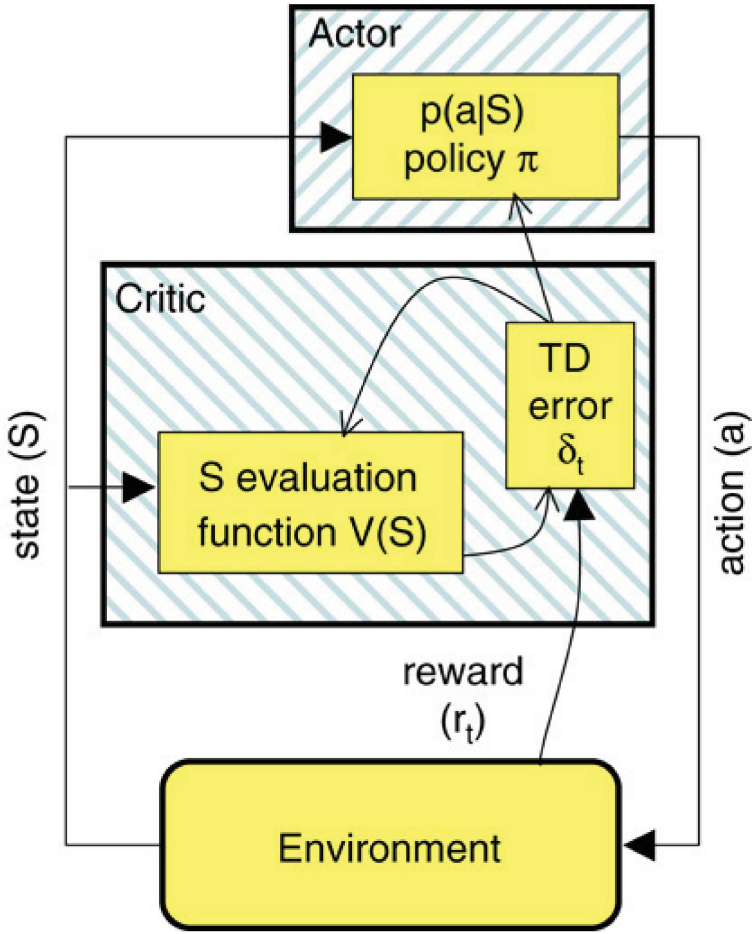


Figure from Niv, 2009.

Evidence for Actor/Critic

fMRI experiment in Humans.

Conditions:

- Instrumental task - probabilistic (30%, 60%)
  - reward
  - neutral
- Pavlovian task - probabilistic (30%, 60%)
  - reward
  - neutral

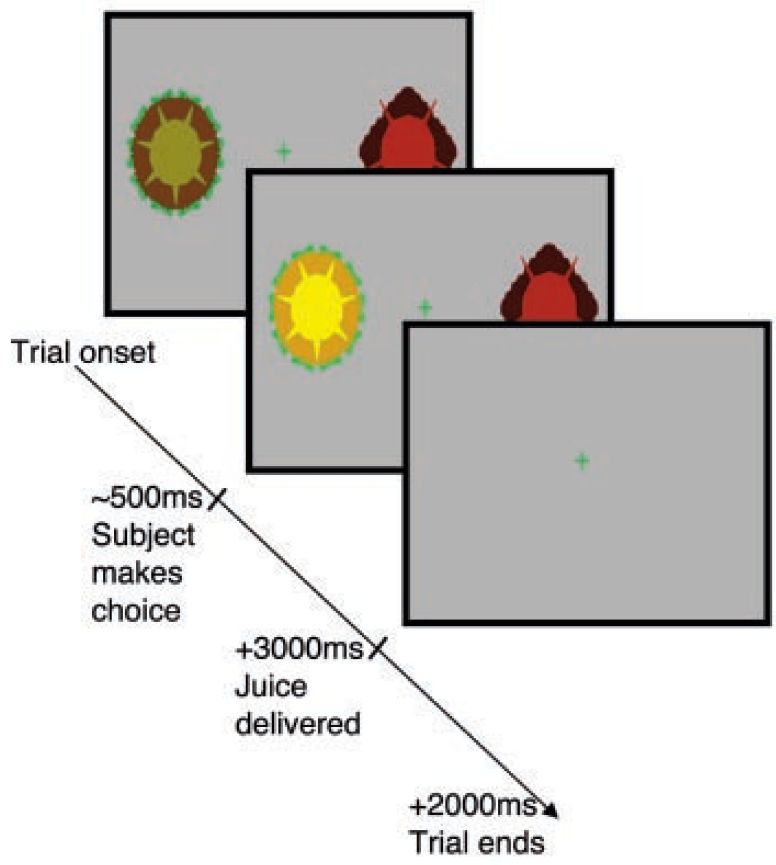


Figure from O'Doherty et al. 2004.

BOLD fMRI to study reinforcement learning

- Data analysis idea: look for signals that correlate with the predicted bold response.

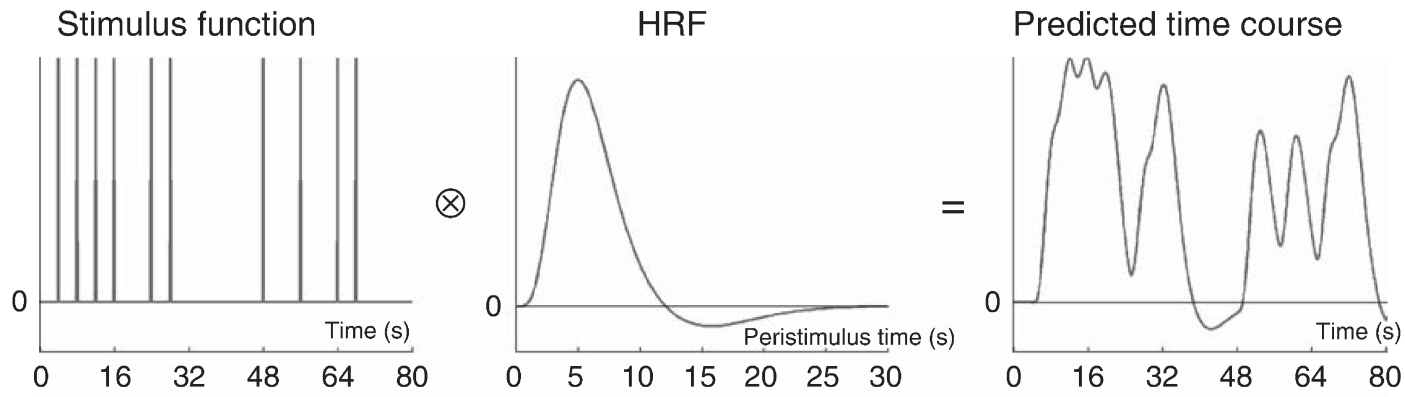
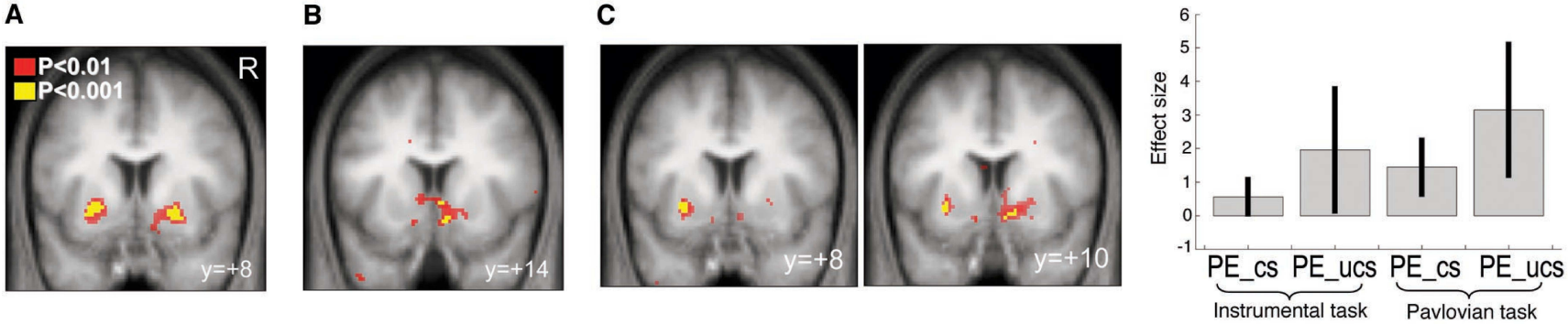


Figure adapted from *Statistical Parametric Mapping* Friston et al. 2007

- For reinforcement learning studies
  - Estimate model parameters to fit the behavior.
  - Derive time changing variables describing quantities as: value of actions, prediction error.
  - Find if/where are signals that correlate with such variables.

Evidence for Actor/Critic

Critic - ventral striatum



Actor - dorsal striatum

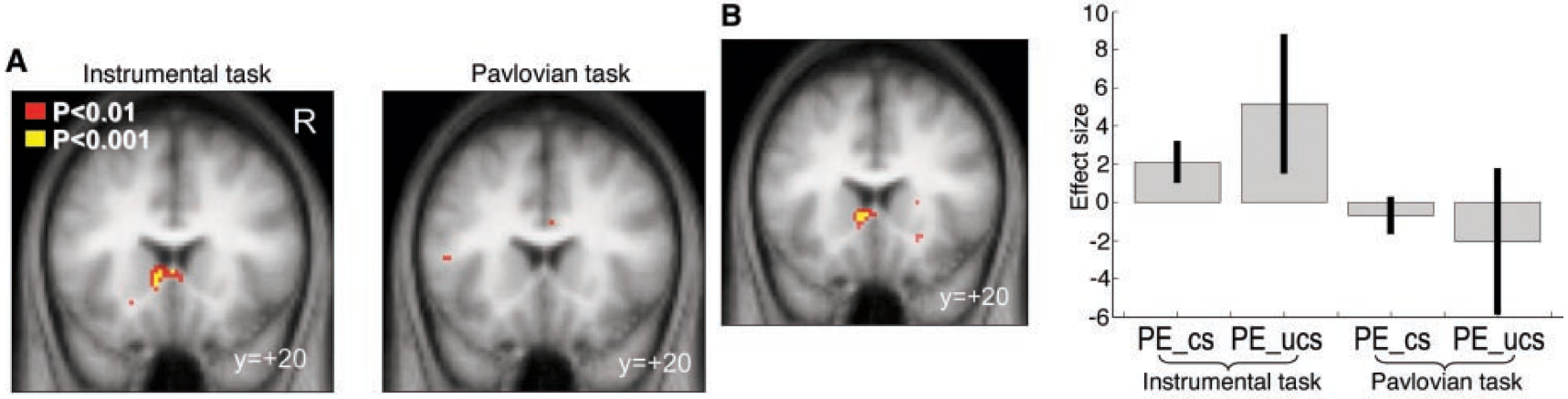


Figure from O'Doherty et al. 2004.

Other topics

- Predictive versus motivational role of dopamine
- Habitual versus goal-directed behaviour
- Model-free learning versus model-based learning
- Disorders: compulsive behaviors, ADHD, schizophrenia
- Substance abuse
- Changes over life-span: impulsivity during adolescence
- Tonic dopamine: response vigor, latency and rate
- ...

References

- H. Bayer and P. Glimcher, Midbrain dopamine neurons encode a quantitative reward prediction error signal, *Neuron*, 2005.
- P. Dayan and L. Abbott, Theoretical neuroscience, 2001.
- J. O'Doherty et al., Dissociable roles of ventral and dorsal Striatum in instrumental conditioning, *Science*, 2004.
- C. Fiorillo et al., Discrete coding of reward probability and uncertainty by dopamine neurons, *Science*, 2003.
- P. Montague et al., Computational roles for dopamine in behavioural control, *Nature*, 2004.
- G. Morris et al. Midbrain dopamine neurons encode decisions for future action, *Nature Neuroscience*, 2006.
- Y. Niv, Reinforcement learning in the brain, *Journal of Mathematical Psychology*, 2009.
- W. Schultz, Predictive reward signal of dopamine neurons, *Journal of Neurophysiology*, 1998.
- W. Schultz, Multiple dopamine functions at different time courses, *Annual Review of Neuroscience*, 2007.
- E. Steinberg et al., A causal link between prediction errors, dopamine neurons and learning, *Nature Neuroscience*, 2013.

Dopamine and synaptic plasticity

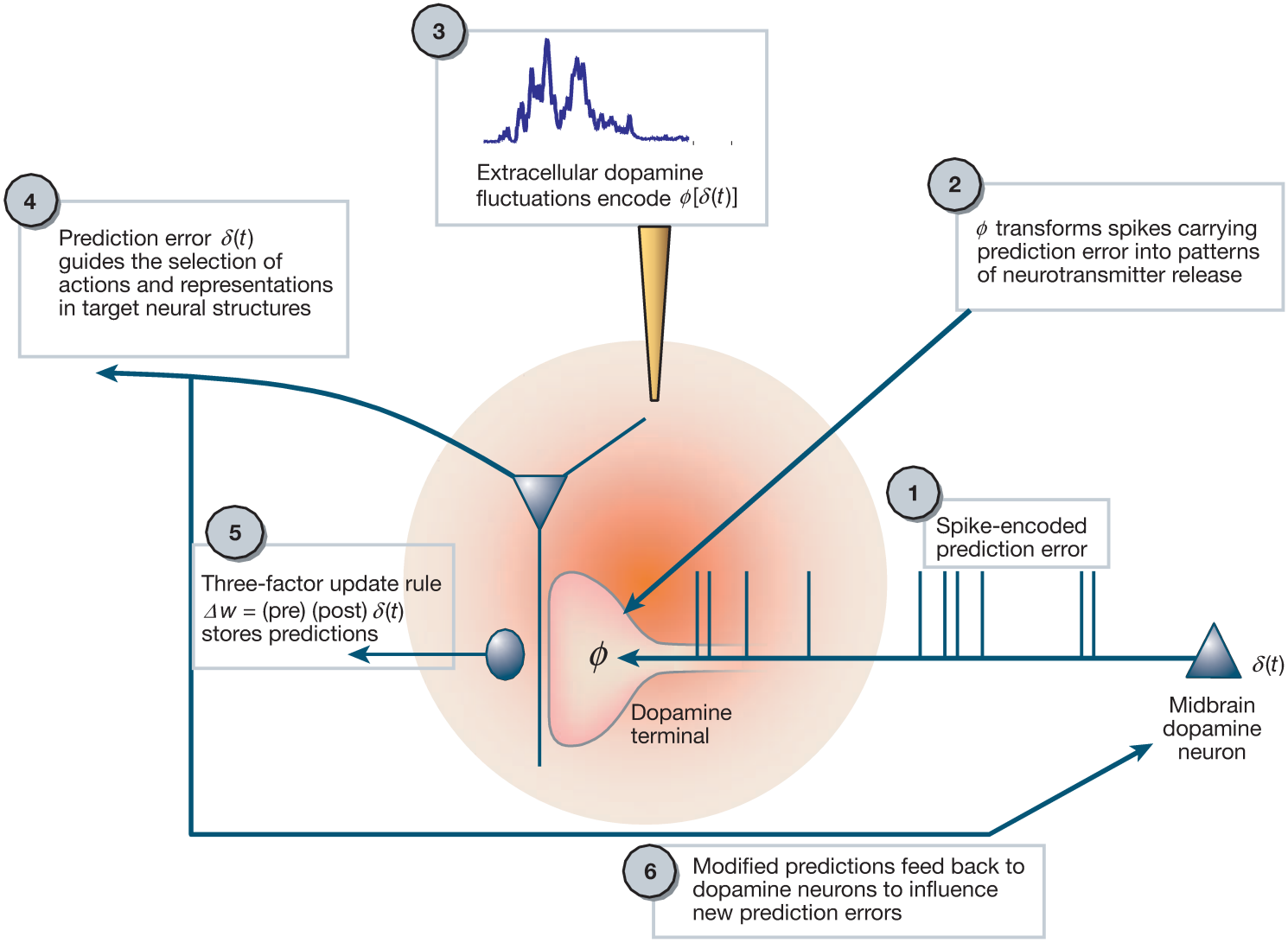


Figure from Montague et al. 2004.