The reward prediction error hypothesis of dopamine neurons

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

Neuroscience and Cognition
Academic year 2024/2025

Francesca Starita

francesca.starita2@unibo.it

Learning objectives

- Characterize dopamine as a neuromodulator
- Name and describe the main dopaminergic pathways, including their origins and projection areas, and discuss the primary function of the nigrostriatal pathway.
- Explain the crucial functions of the basal ganglia in motor control, reinforcement learning, and goal-oriented behavior.
- Describe how dopamine, released by the substantia nigra pars compacta (SNc), modulates the direct and indirect pathways within the basal ganglia to influence the initiation and inhibition of movement.
- Articulate how the release of dopamine is dependent on the discrepancy between predicted future reward and the actual reward received.
- Summarize the reward prediction error hypothesis of dopamine neuron activity as a synthesis of computational reinforcement learning principles and findings from neuroscience experiments.
- Explain how dopamine's modulation of synaptic plasticity serves as a plausible neural mechanism for the implementation of learning algorithms.

Learning objectives

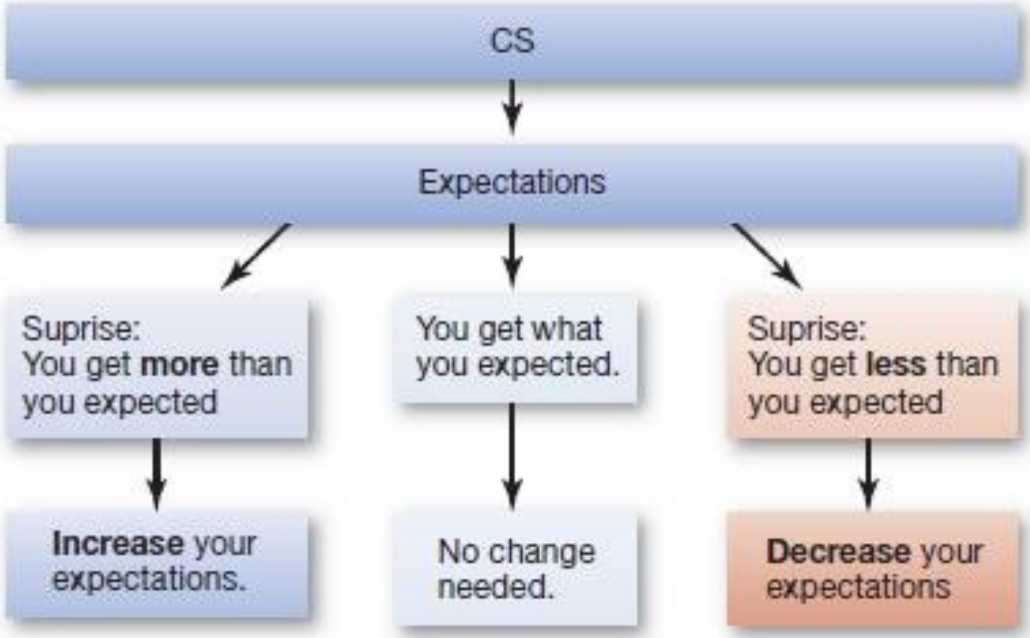
- Identify the dopaminergic system as the primary neural substrate for reward and reinforcement processing, encompassing both natural rewards and addictive drugs.
- Distinguish between the earlier view that dopamine neurons broadcast a reward signal and the more recent reward prediction error hypothesis.
- Describe the strong phasic response of dopamine neurons to unexpected rewards.
- Explain how dopamine neurons differentiate between the presence and absence of reward.
- Describe how dopamine neurons discriminate between different magnitudes of reward in a relative rather than absolute manner.
- Explain how dopamine neurons encode the probability or uncertainty associated with reward.
- Describe the signed (bidirectional) response of dopamine neurons to prediction errors, including increased firing to unexpected reward delivery (positive PE) and paused firing to unexpected omission (negative PE).
- Explain how the dopamine response shifts from the reward itself to the conditioned stimulus (CS) that predicts reward occurrence as learning progresses.

Learning objectives

- Describe the involvement of dopamine in both Pavlovian (prediction) learning and instrumental (control) learning.
- Discuss the implications of the reward prediction error hypothesis for understanding addiction, focusing on the differences between dopamine signals elicited by natural rewards and addictive drugs.
- Explain that dopamine's role extends beyond simply inducing pleasure, and how it functions to highlight unexpected valuable events and predictive cues, driving learning about where to find desired outcomes.

Prediction errors in computational models

- **Signed difference between the expected and delivered outcome**
- Functions as **teaching signal** to update expectations and reduce following prediction errors
- **Used to update predicted value**
  - The value signal produced by the reward itself transfers back to events that reliably precede reward delivery (e.g. CSs)
  - Thus, **the rewarding value transfers from the reward to the CS that predicts reward**



7.13 The role of expectations and surprise This figure shows the (automatic, unconscious) process through which expectations can be adjusted, trial by trial, in a classical conditioning experiment. The one complication not shown here is that bigger surprises (greater departures from expectations) will trigger larger adjustments; smaller surprises will lead to smaller adjustments.

https://www.brainkart.com/article/Classical-Conditioning--Role-of-Surprise_29314/

How could signed prediction errors be implemented in neurons?

How could signed prediction errors be implemented in neurons?

Neurons could **change their firing rate** when predictions do not meet reality

- increase their firing rates when the error is positive
- decrease their firing rates when the error is negative

Synaptic plasticity

- **Changes in synaptic efficacy** through changes in the amount
    - of neurotransmitter that is released, directly affecting excitation or inhibition of postsynaptic neuron
    - of **neuromodulator**, which is a neurotransmitter having effects other than, or in addition to, direct neural excitation or inhibition

The parameters, or weights, adjusted by learning algorithms correspond to synaptic efficacies

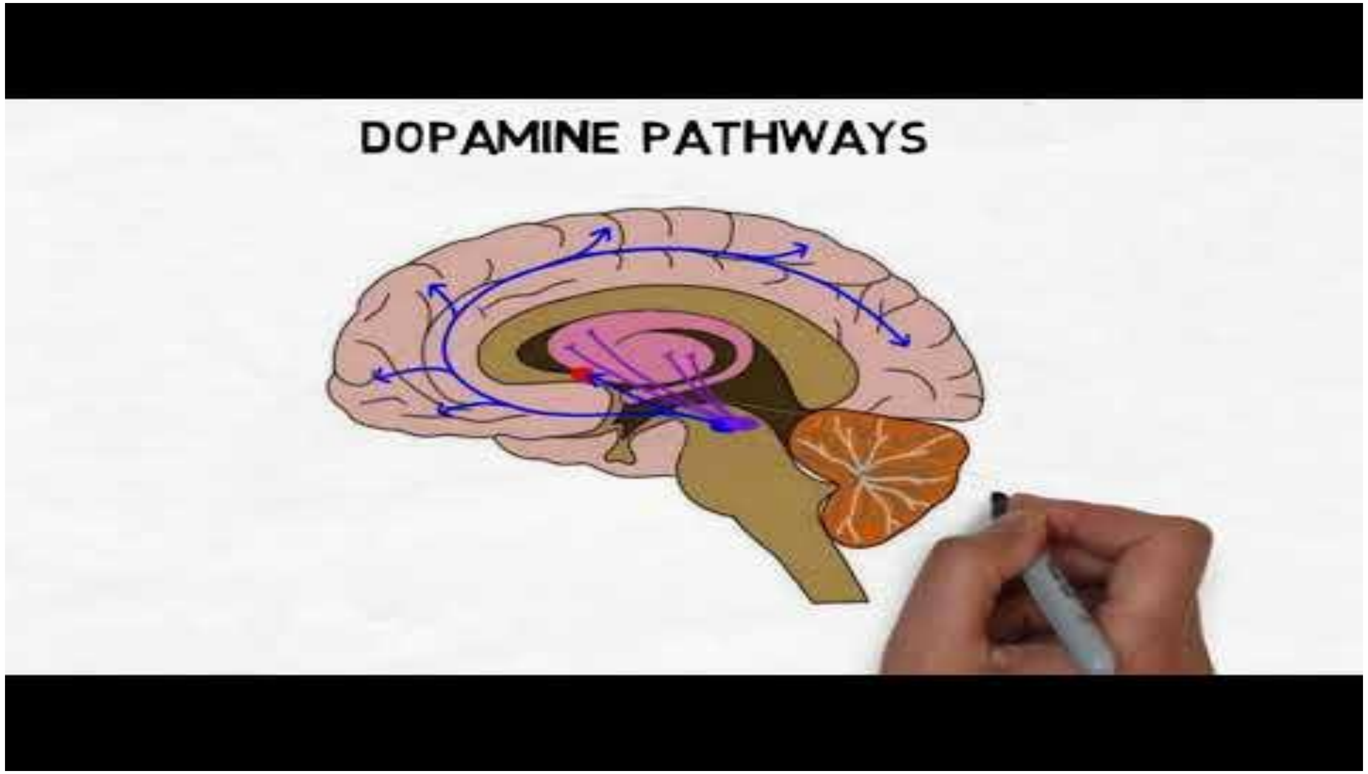
The reward prediction error hypothesis of dopamine neuron activity

Dopamine

It's a **neuromodulator**, i.e.
a neurotransmitter having
effects other than, or in addition
to, direct neural excitation or inhibition

Plays essential roles in many processes

- Motivation
- Learning
- action-selection, decision-making
- most forms of addiction
- Parkinson’s disease
- Huntington's disease



https://youtu.be/Wa8_nLwQIpg

The dopaminergic pathways

1. Nigrostriatal pathway

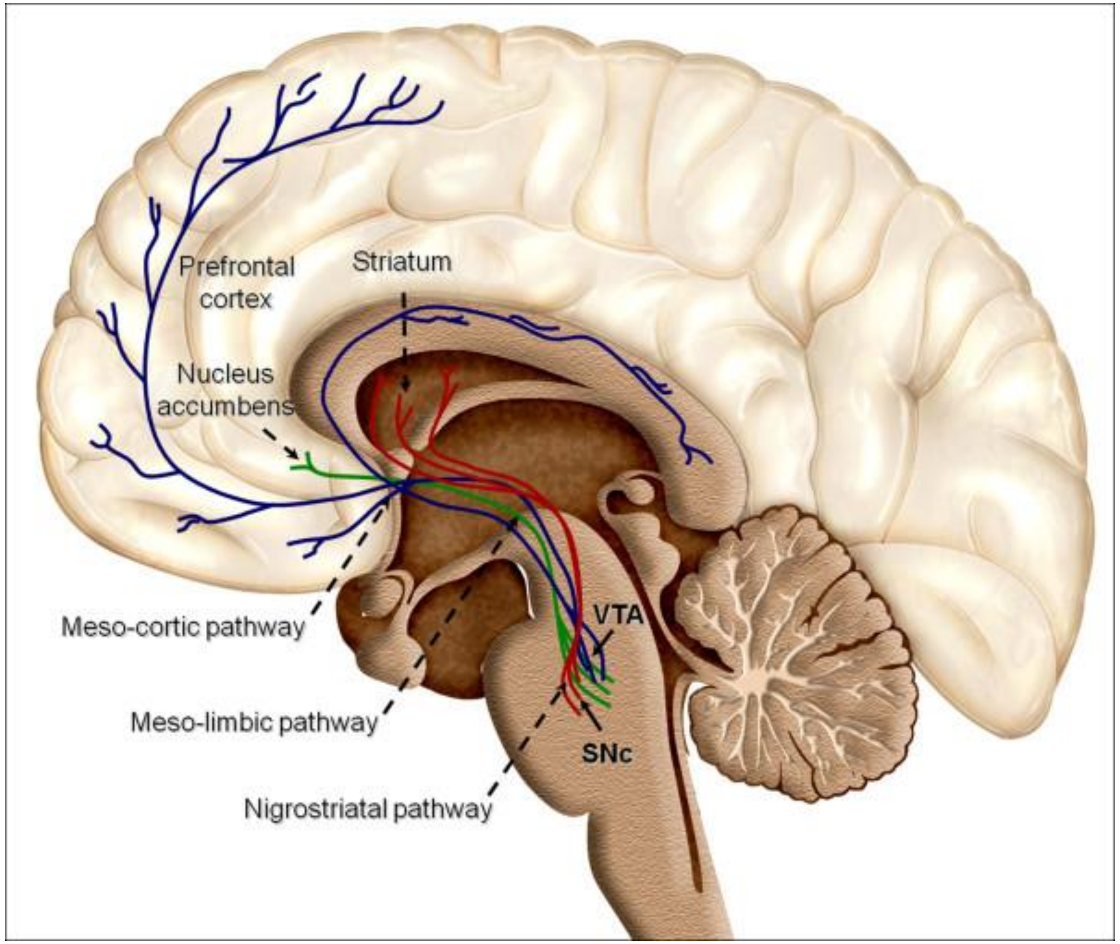
- originates in the substantia nigra pars compacta (SNc)
- projects primarily to the **caudate–putamen** (dorsal striatum in rodents)
- It is critical in the production of **movement** as part of the basal ganglia motor loop

2. Mesolimbic pathway

- originates in the VTA
- projects to the **nucleus accumbens**, septum, amygdala and hippocampus

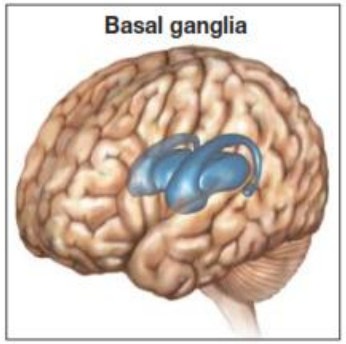
3. Mesocortical pathway

- Originates in the VTA
- projects to the medial prefrontal, cingulate, **orbitofrontal** and perirhinal cortex

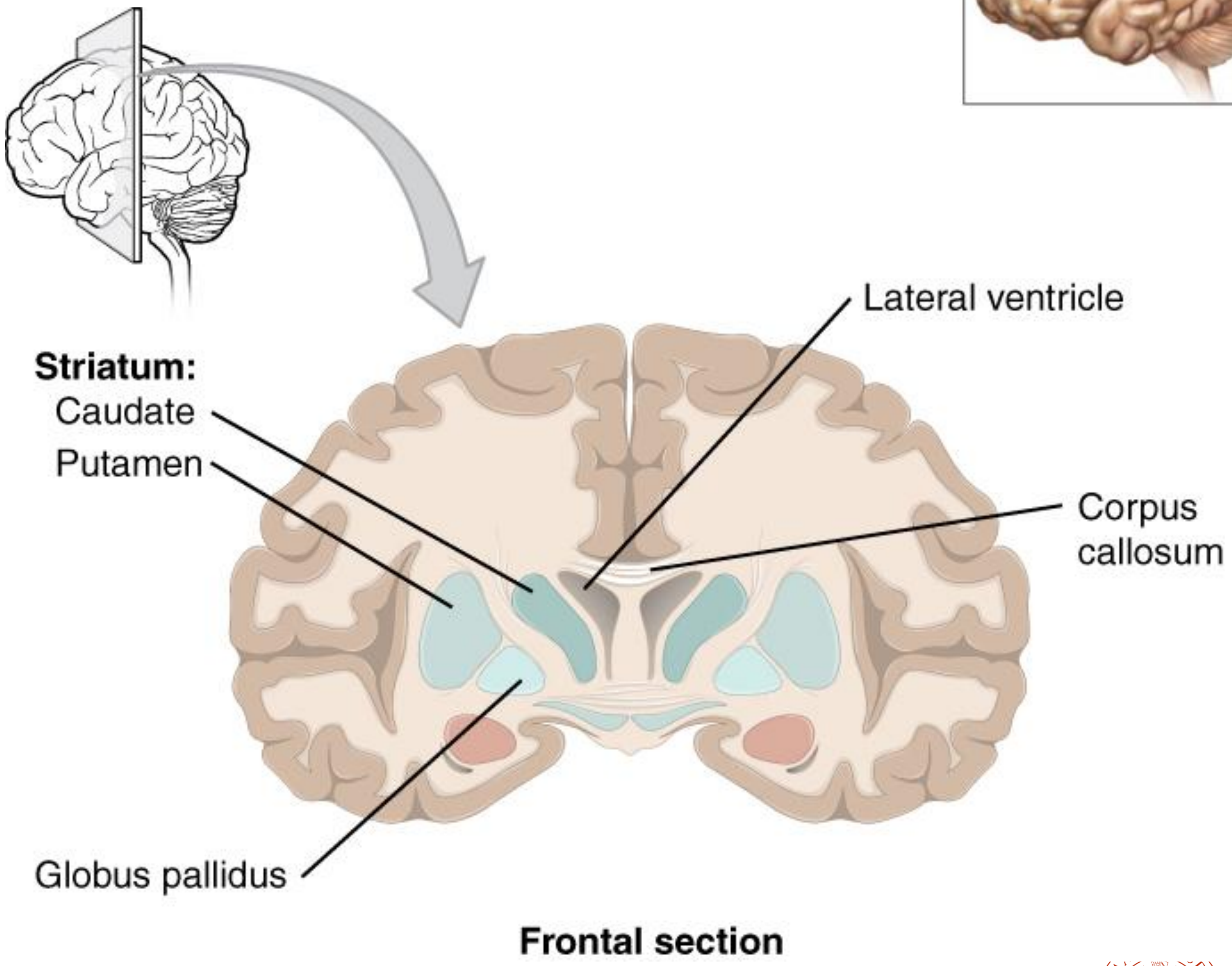


https://www.brainfacts.org/3d-brain#intro=false&focus=Brain

Basal Ganglia motor loop



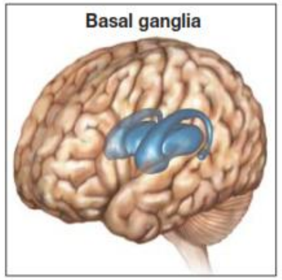
- Collection of subcortical nuclei
- Receive inputs from sensory and motor areas
- Send output largely through the thalamus to the frontal lobe
- Extensively interconnected
- Have crucial role in **motor control**
- Have crucial role in **reinforcement learning** and goal-oriented behavior

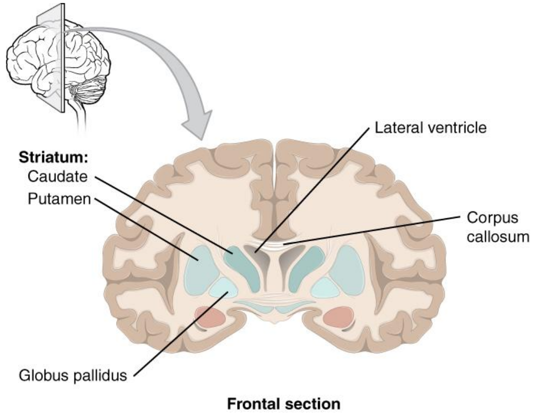


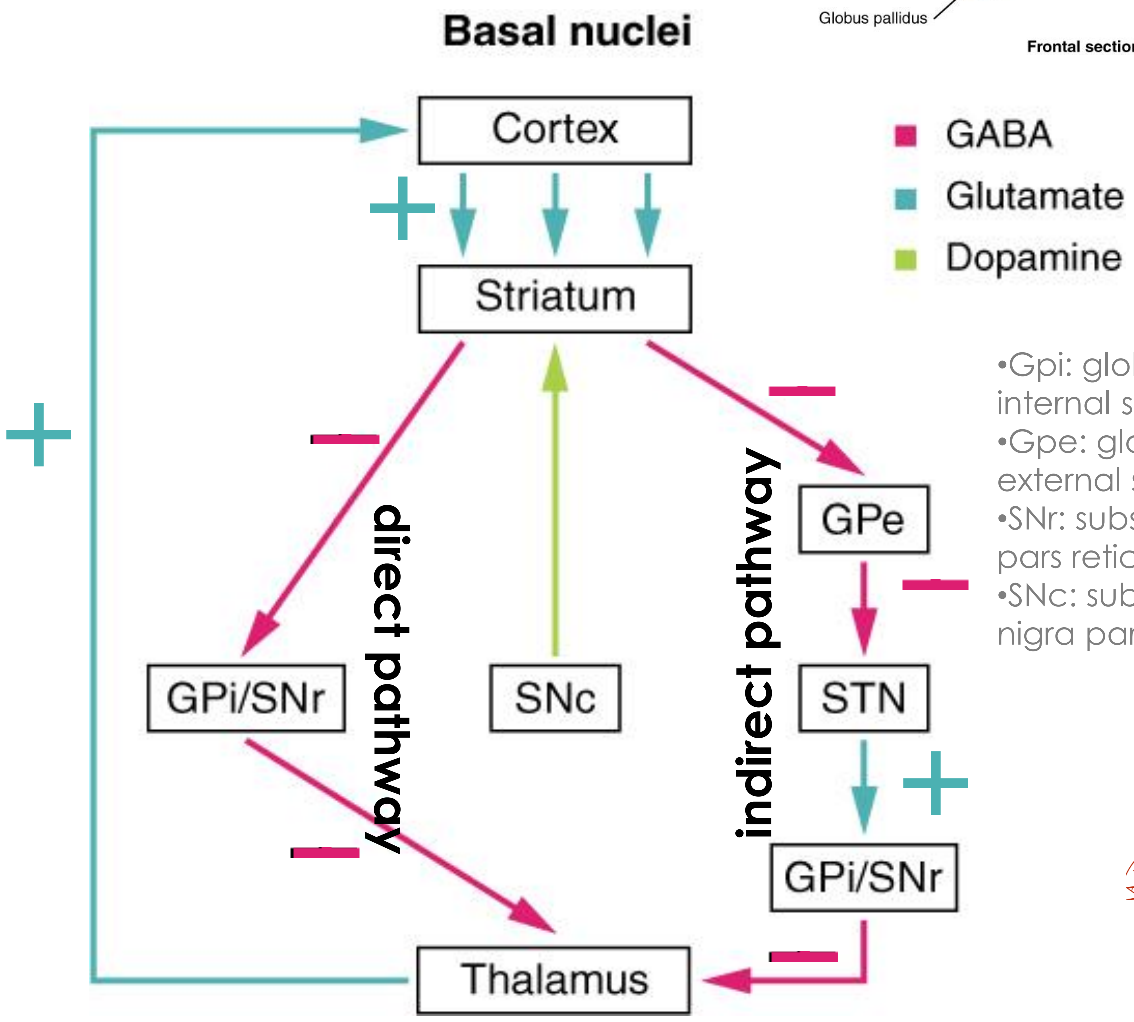
Basal Ganglia have crucial role in motor control & reinforcement learning

- the **direct pathway** causes the disinhibition of the thalamus → **initiates movement**

- the **indirect pathway** causes, or reinforces, the normal inhibition of the thalamus → **inhibits movement**





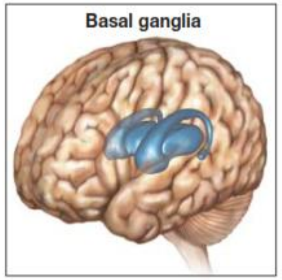


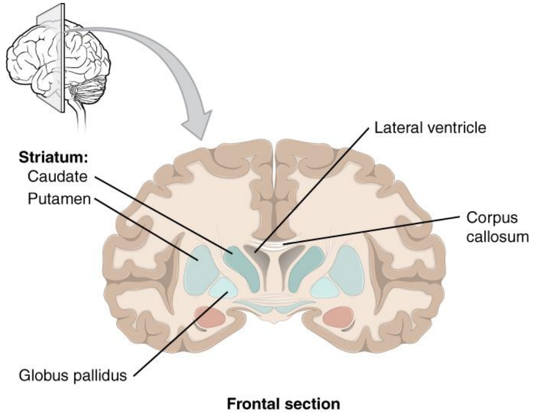
•Gpi: globus pallidus internal segment
•Gpe: globus pallidus external segment
•SNr: substantia nigra pars reticulata
•SNc: substantia nigra pars compacta

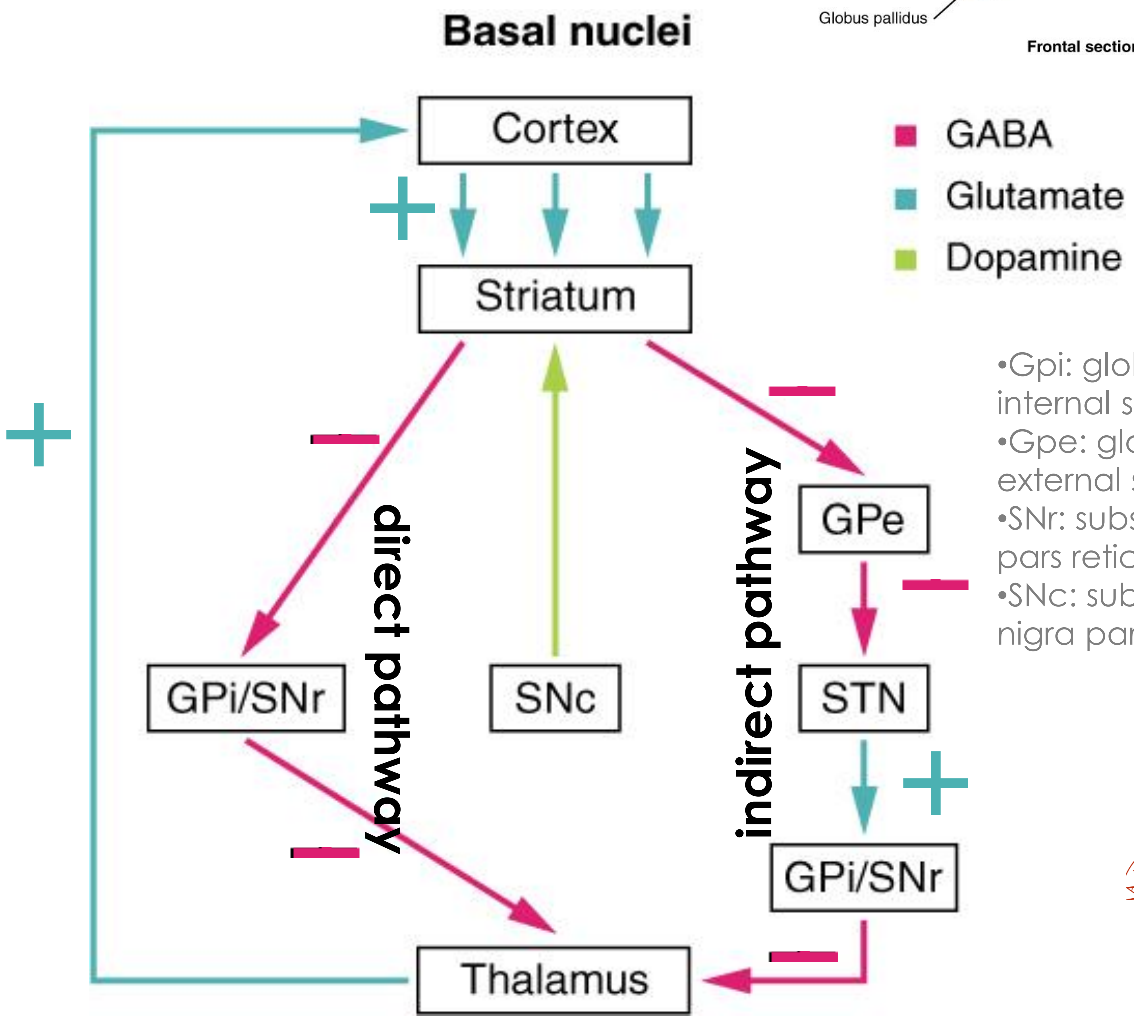
Basal Ganglia have crucial role in motor control & reinforcement learning

The switch between the two pathways is the **substantia nigra pars compacta:**

- which **projects to the striatum**
- and releases the neurotransmitter **dopamine**:
    - Activates the direct pathway
    - Inhibits the indirect pathway
    - **Dopamine release depends on the error between predicted future reward and actual reward**





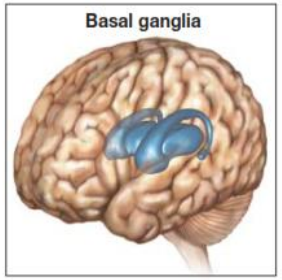


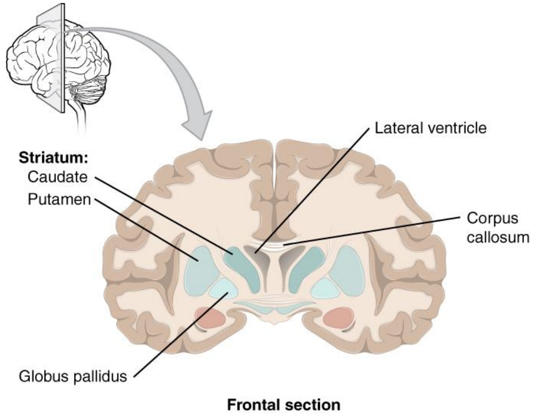
•Gpi: globus pallidus internal segment
•Gpe: globus pallidus external segment
•SNr: substantia nigra pars reticulata
•SNc: substantia nigra pars compacta

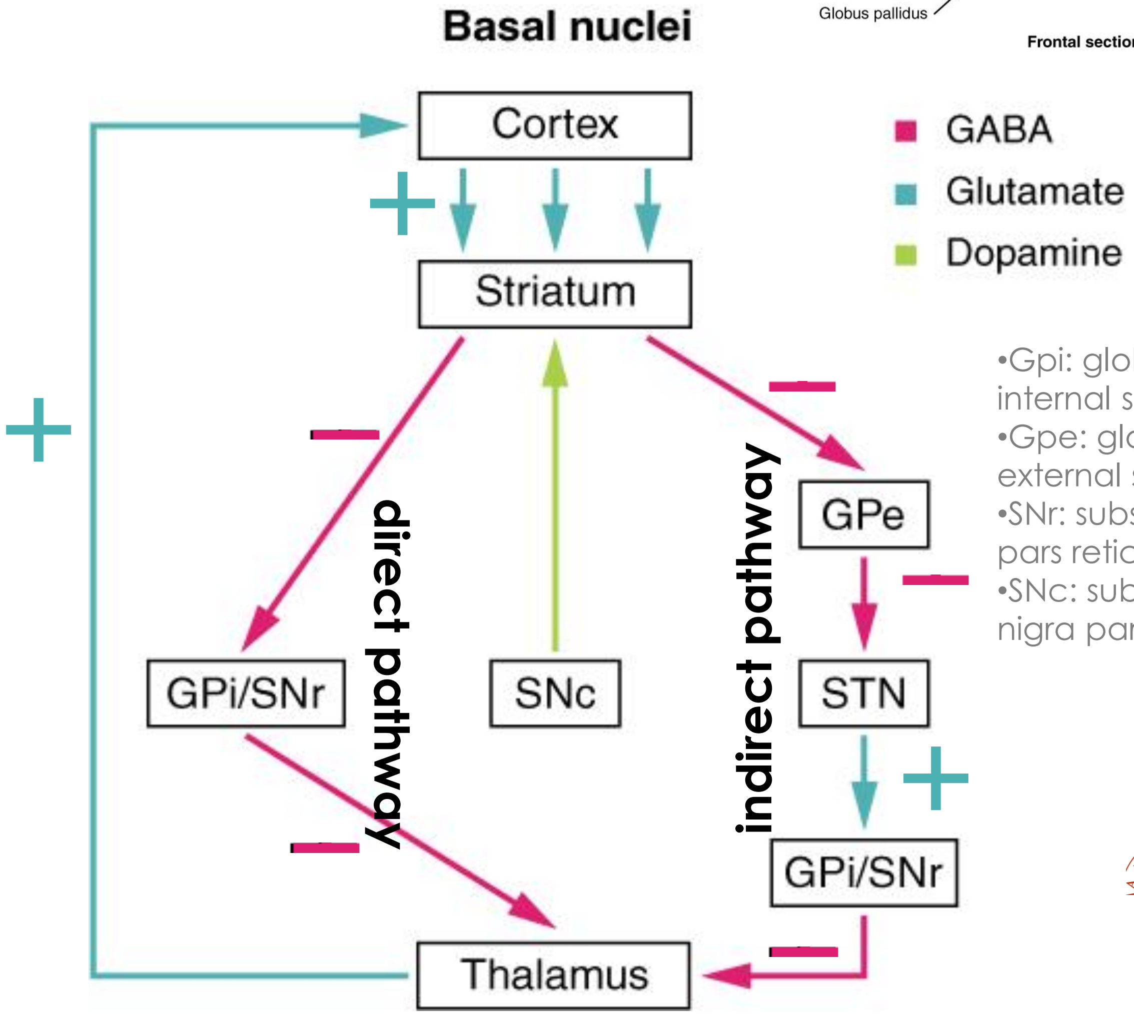
Basal Ganglia have crucial role in motor control & reinforcement learning

The striatum may be the interface where reward influences action

- The basal ganglia are involved in the selection of actions
- Rewards may influence which actions are selected
    - o by affecting plasticity in the striatum, so as to reinforce rewarded actions and make them more likely to recur





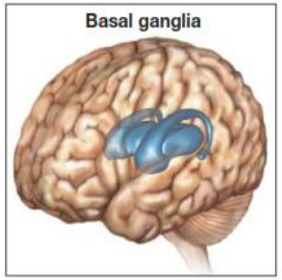


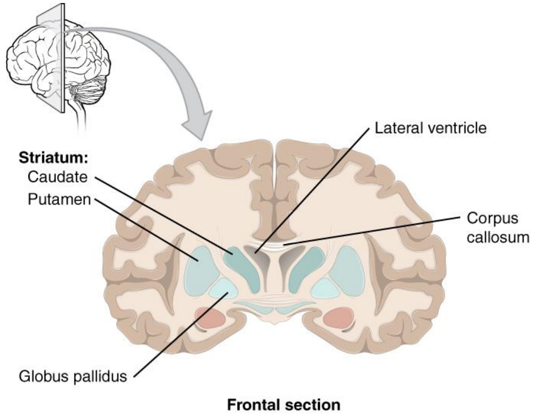
•Gpi: globus pallidus internal segment
•Gpe: globus pallidus external segment
•SNr: substantia nigra pars reticulata
•SNc: substantia nigra pars compacta

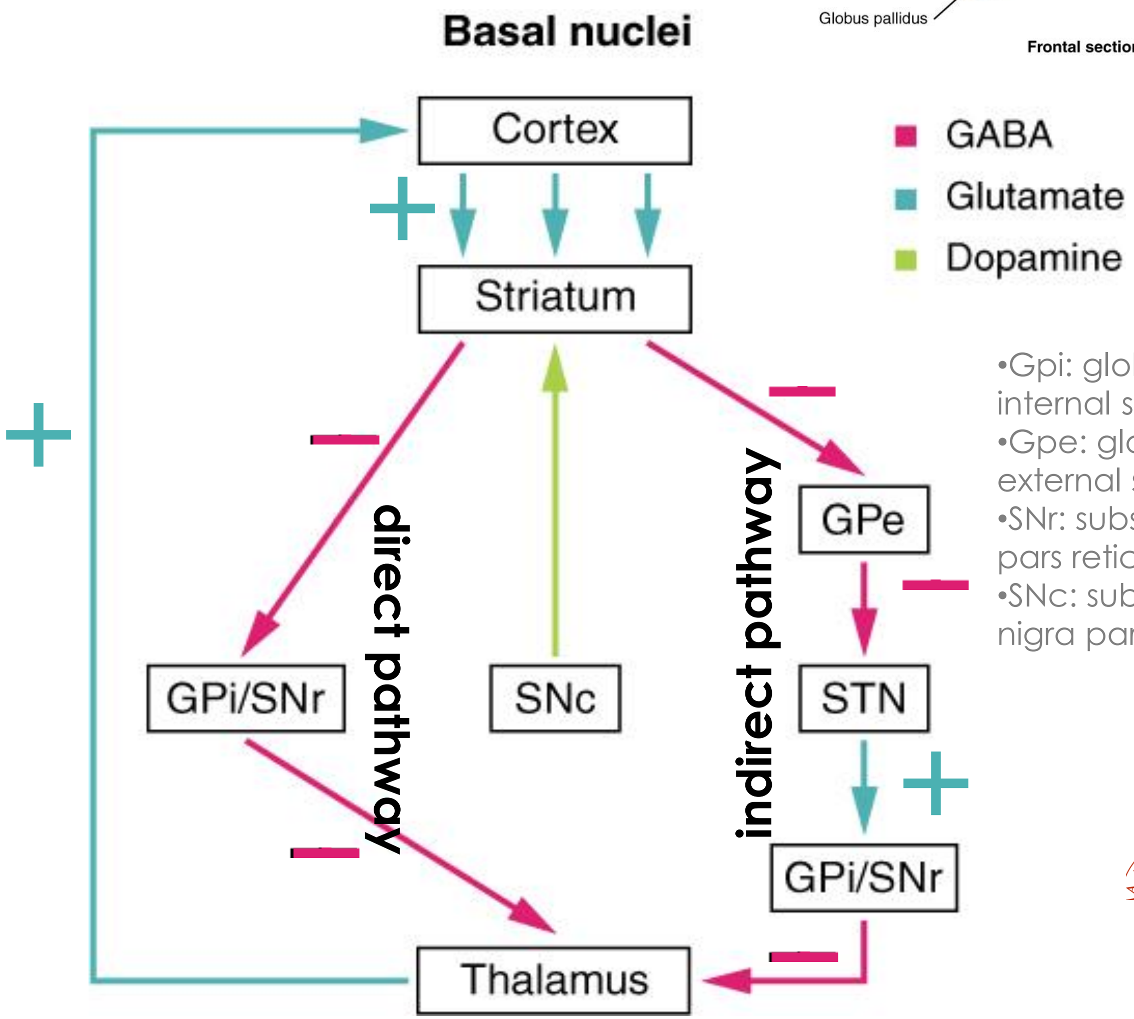
Basal Ganglia have crucial role in motor control & reinforcement learning

Parkinson's disease

- loss of dopaminergic neurons of the SNc → overactivation of the indirect pathway → decrease of movement (hypokinesia)
- But also pathological gambling







- Gpi: globus pallidus internal segment
- Gpe: globus pallidus external segment
- SNr: substantia nigra pars reticulata
- SNc: substantia nigra pars compacta

The reward prediction error hypothesis of dopamine neuron activity

- Resulted from the convergence of computational reinforcement learning and results of neuroscience experiments

- Modulation of synaptic plasticity via the neuromodulator dopamine is a plausible mechanism for how the brain might implement learning algorithms

- There is strong evidence that the dopaminergic system is the major neural substrate of reward and reinforcement for both natural rewards and addictive drugs

Wolfram Schultz Raymond Dolan Peter Dayan

2017

for 'their multidisciplinary analysis of brain mechanisms that link learning to reward, which has far-reaching implications for the understanding of human behaviour, including disorders of decision-making in conditions such as gambling, drug addiction, compulsive behaviour and schizophrenia.'

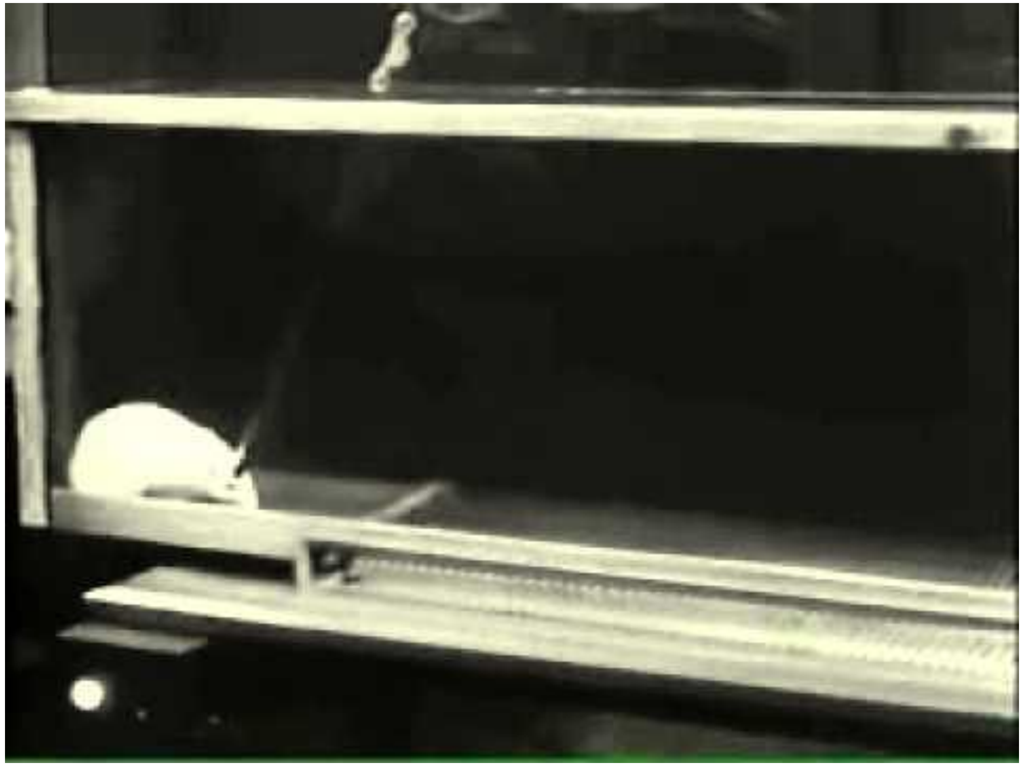
https://youtu.be/8n1-lv1_wCY

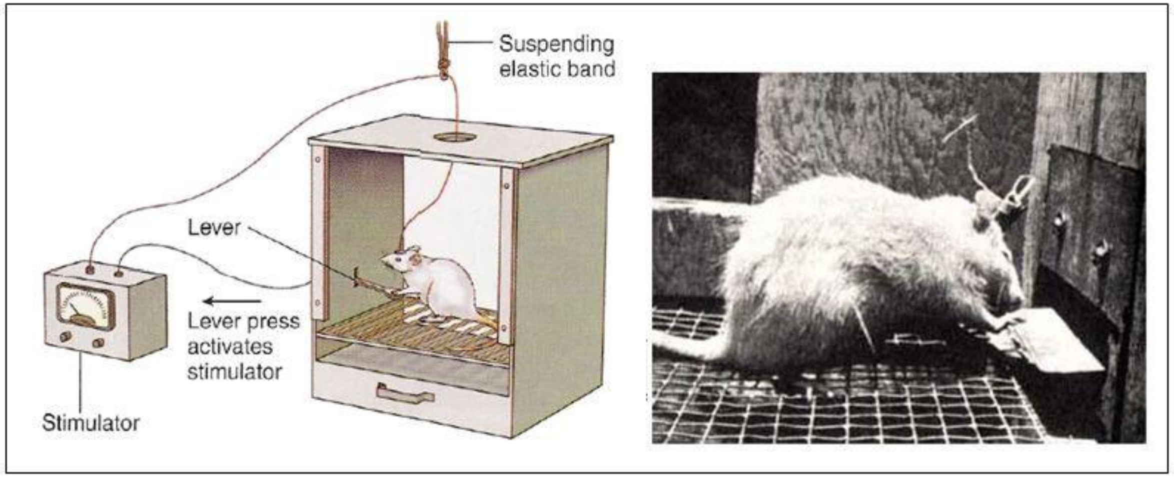
The early view: Dopamine neurons broadcast a reward signal

1954: James Olds and Peter Milner publish a paper that describes the effects of electrical stimulation on certain areas of a rat's brain

"... the control exercised over the animal's behavior by means of this reward is extreme, possibly exceeding that exercised by any other reward previously used in animal experimentation"

the sites at which stimulation was most effective in producing this rewarding effect excited dopamine pathways, either directly or indirectly, that ordinarily are excited by natural rewarding stimuli.

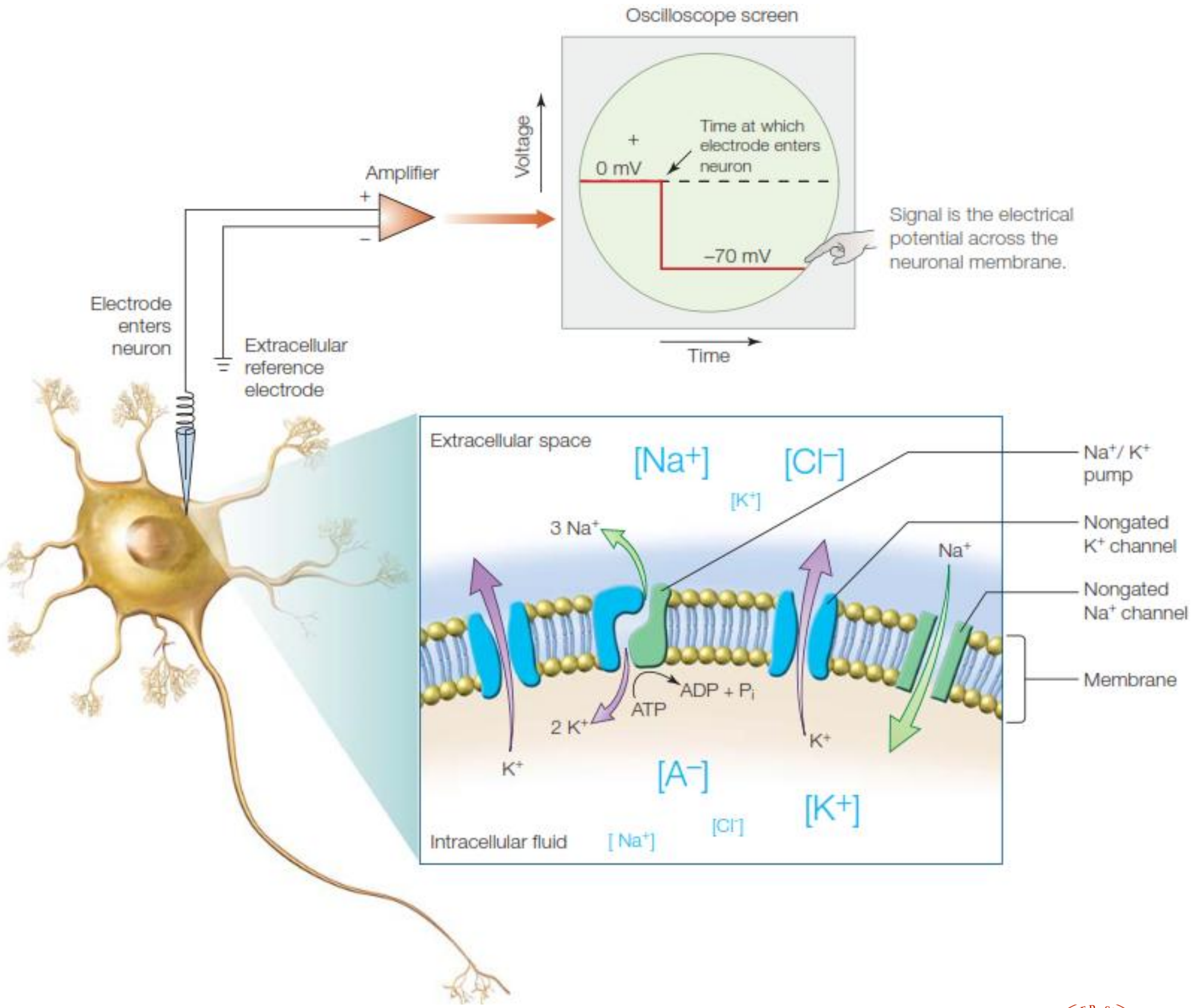




https://youtu.be/uofQPLuLV9A

Measuring neuronal signalling

Idealized neuron shown with intracellular recording electrode penetrating the neuron. The electrode measures the difference between the voltage inside versus outside the neuron and this difference is amplified and displayed on an oscilloscope screen (top). The oscilloscope screen shows voltage over time, and shows that prior to the electrode entering the neuron, voltage between the electrode and the extracellular reference electrode is zero, but when the electrode is pushed into the neuron, the difference becomes –70 mV, which is the resting membrane potential.

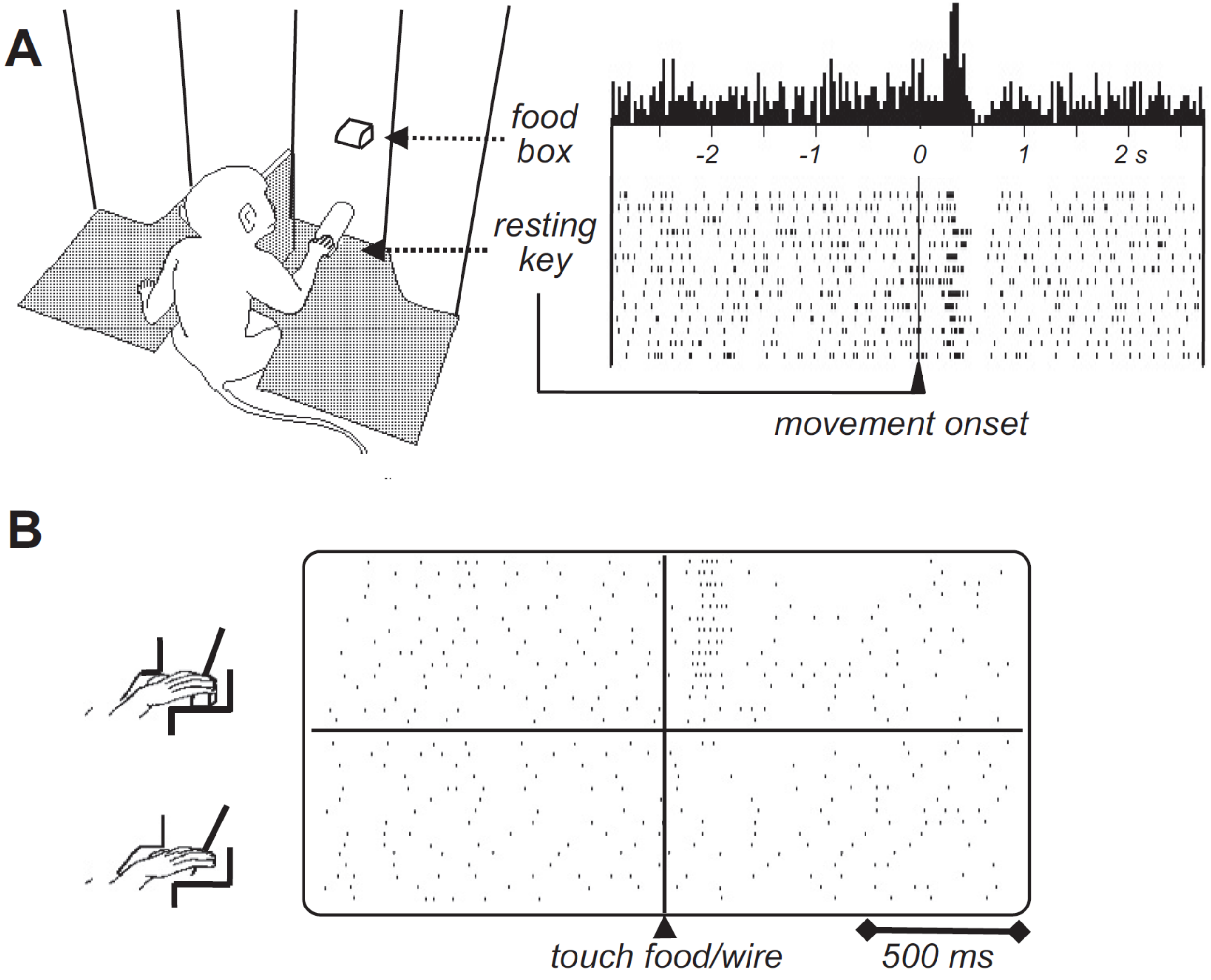


Dopaminergic neurons exhibit a strong phasic response to an unexpected reward...

Activity of single dopamine neurons is recorded in alert monkeys while they perform behavioral acts and receive rewards.

A) Dopamine responses to touch of food in absence of any stimuli predicting the reward. The food inside the box is invisible but is touched by the hand underneath the cover.

Romo & Schultz , J. Neurophysiology, 1990

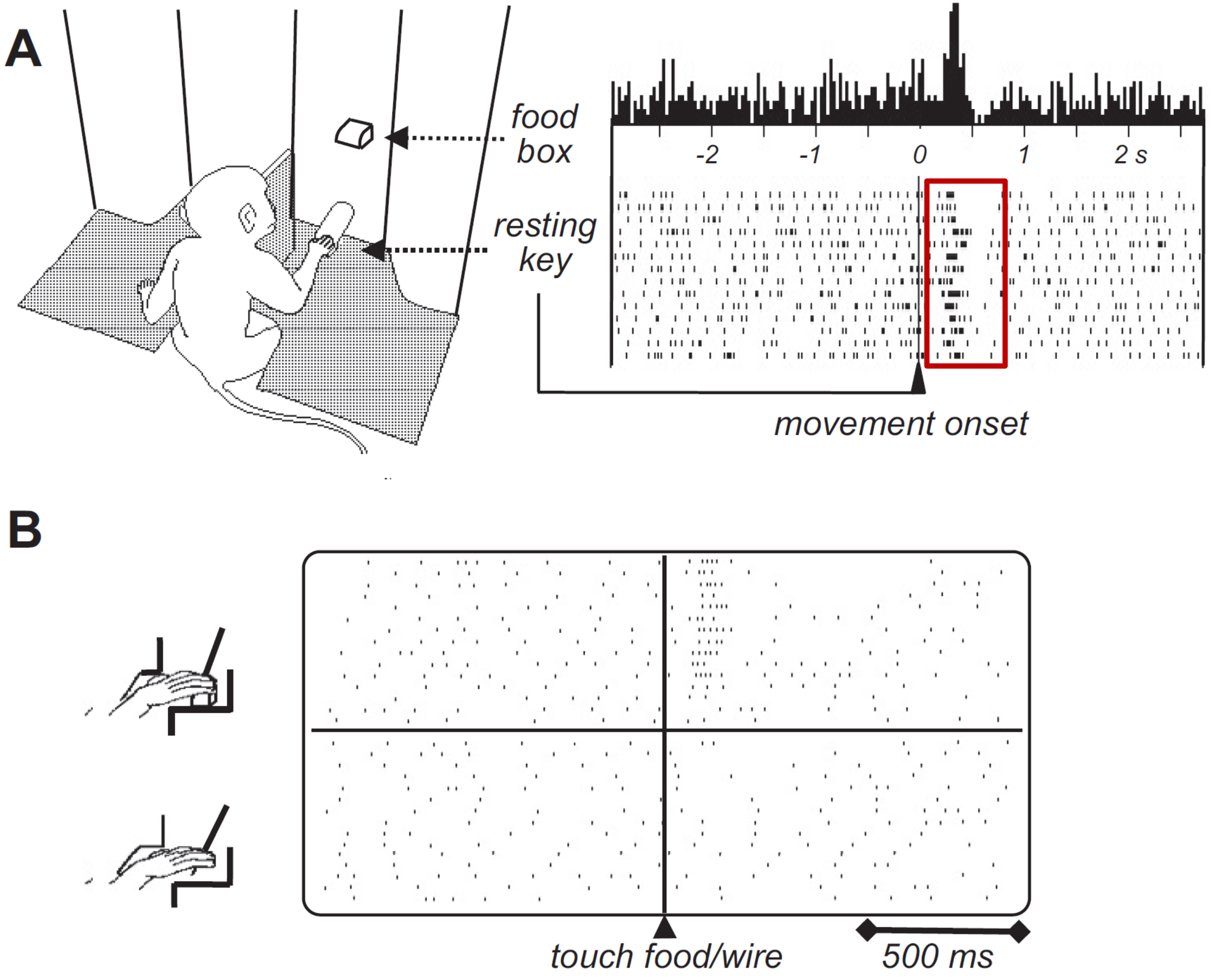


Dopaminergic neurons exhibit a strong phasic response to an unexpected reward...

Activity of single dopamine neurons is recorded in alert monkeys while they perform behavioral acts and receive rewards.

A) Dopamine responses to touch of food in absence of any stimuli predicting the reward. The food inside the box is invisible but is touched by the hand underneath the cover.

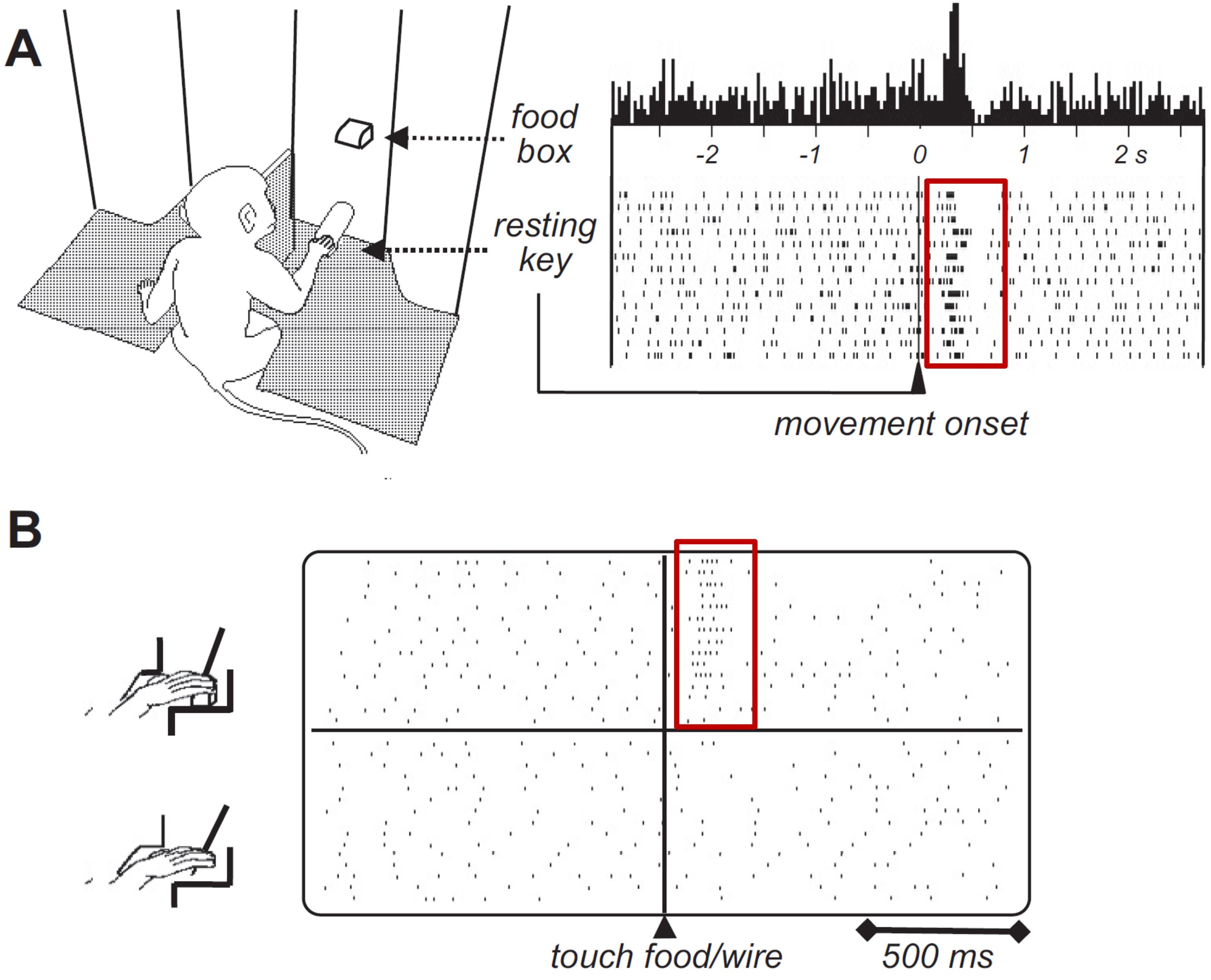
Romo & Schultz , J. Neurophysiology, 1990



...discriminate between reward & non-reward

Activity of single dopamine neurons is recorded in alert monkeys while they perform behavioral acts and receive rewards.

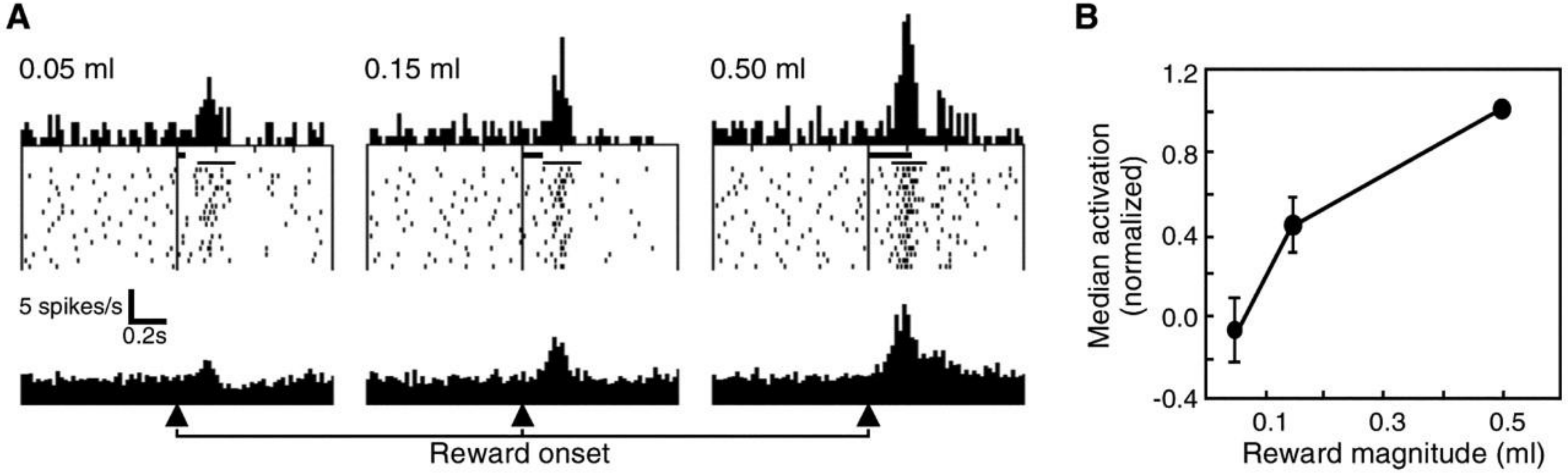
B) Differential response of dopamine neuron to touch of a wire holding a piece of apple (top), or touch of a wire holding an inedible objects (bottom).



Romo & Schultz , J. Neurophysiology, 1990

...discriminate between reward magnitude

(Tobler, Fiorillo and Schulz, Science, 2005)



Neural discrimination of liquid volume.

(**A**) (Top) Rasters and histograms of activity from a single dopamine neuron. (Bottom) Population histograms of activity from all neurons tested (*n* = 55 neurons). Three volumes of liquid were delivered in the absence of any explicit predictive stimuli.

(**B**) Neural response as a function of liquid volume. Median (±95% confidence intervals) percentage change in activity for the population of neurons (*n* = 55 neurons) was calculated for responses to each volume after normalization in each neuron to the response after delivery of 0.5 ml, which itself elicited a median activation of 159% above baseline activity.

How could signed prediction errors be implemented by dopamine neurons?

Neurons could change their firing rate when **predictions do not meet reality**

1. increase their firing rates when the error is positive (more than expected)
2. decrease their firing rates when the error is negative (less than expected)

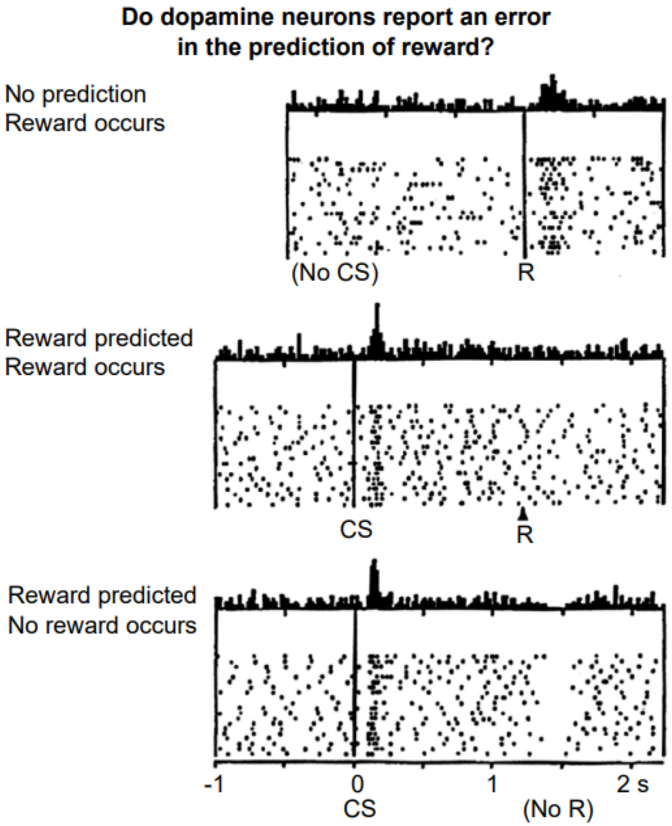
Dopamine neurons show phasic changes in firing rate when predictions do not meet reality

1. increase firing rate when
   - Reward is unexpectedly delivered
   - Reward is better than expected
2. Suppress firing rate when
   - Reward is unexpectedly omitted
   - Reward is worse than expected

Phasic responses of dopamine neurons signal reward prediction errors, not reward itself

...transfers back to a cue which predicts reward occurrence (i.e. CS)

Fig. 1. Changes in dopamine neurons' output code for an error in the prediction of appetitive events. (**Top**) Before learning, a drop of appetitive fruit juice occurs in the absence of prediction—hence a positive error in the prediction of reward. The dopamine neuron is activated by this unpredicted occurrence of juice. (**Middle**) After learning, the conditioned stimulus predicts reward, and the reward occurs according to the prediction—hence no error in the prediction of reward. The dopamine neuron is activated by the reward-predicting stimulus but fails to be activated by the predicted reward (right). (**Bottom**) After learning, the conditioned stimulus predicts a reward, but the reward fails to occur because of a mistake in the behavioral response of the monkey. The activity of the dopamine neuron is depressed exactly at the time when the reward would have occurred. The depression occurs more than 1 s after the conditioned stimulus without any intervening stimuli, revealing an internal representation of the time of the predicted reward. Neuronal activity is aligned on the electronic pulse that drives the solenoid valve delivering the reward liquid (top) or the onset of the conditioned visual stimulus (middle and bottom). Each panel shows the peri-event time histogram and raster of impulses from the same neuron. Horizontal distances of dots correspond to real-time intervals. Each line of dots shows one trial. Original sequence of trials is plotted from top to bottom. CS, conditioned, reward-predicting stimulus; R, primary reward.

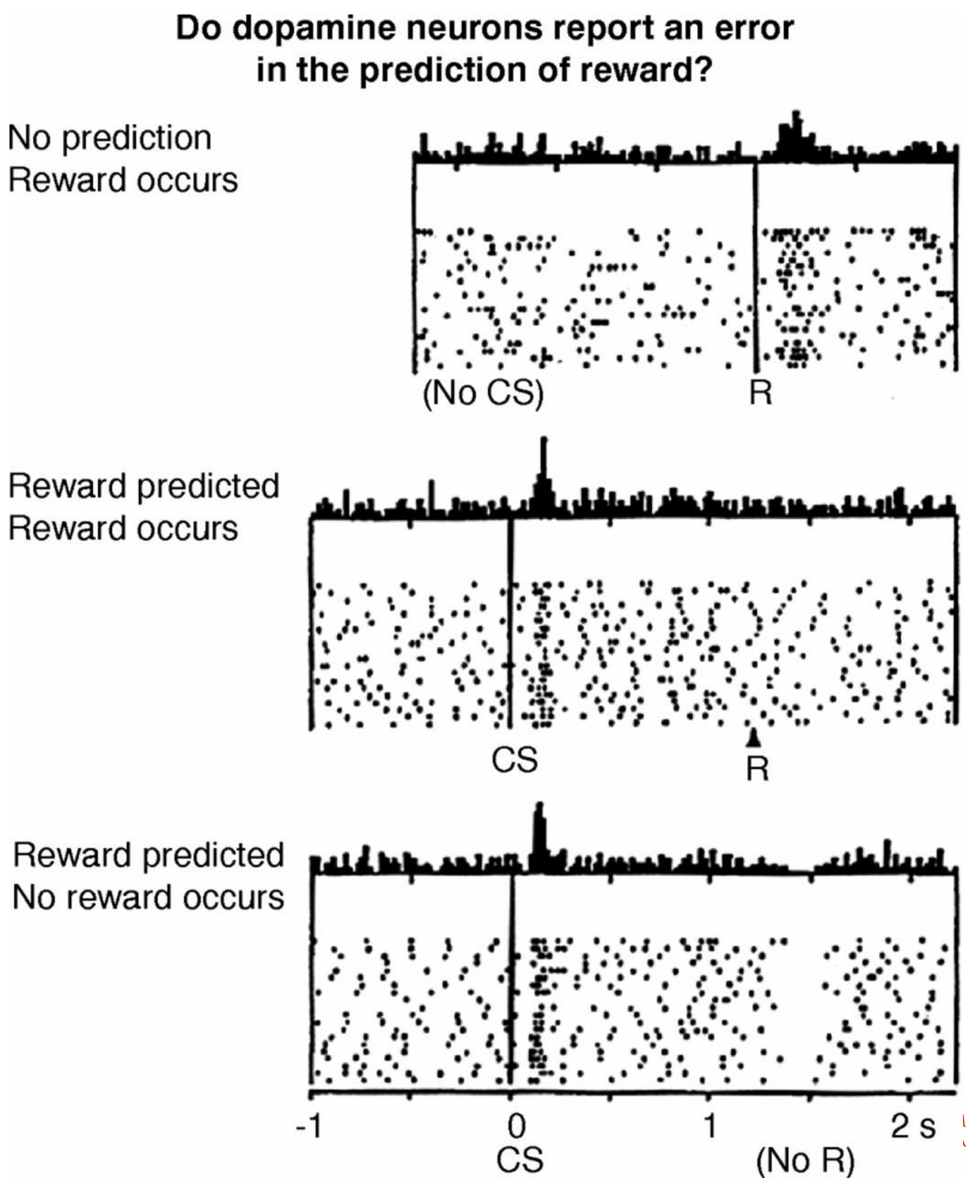


A Neural Substrate of Prediction and Reward

Wolfram Schultz, Peter Dayan, P. Read Montague*

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

...transfers back to a cue which predicts reward occurrence (i.e. CS)



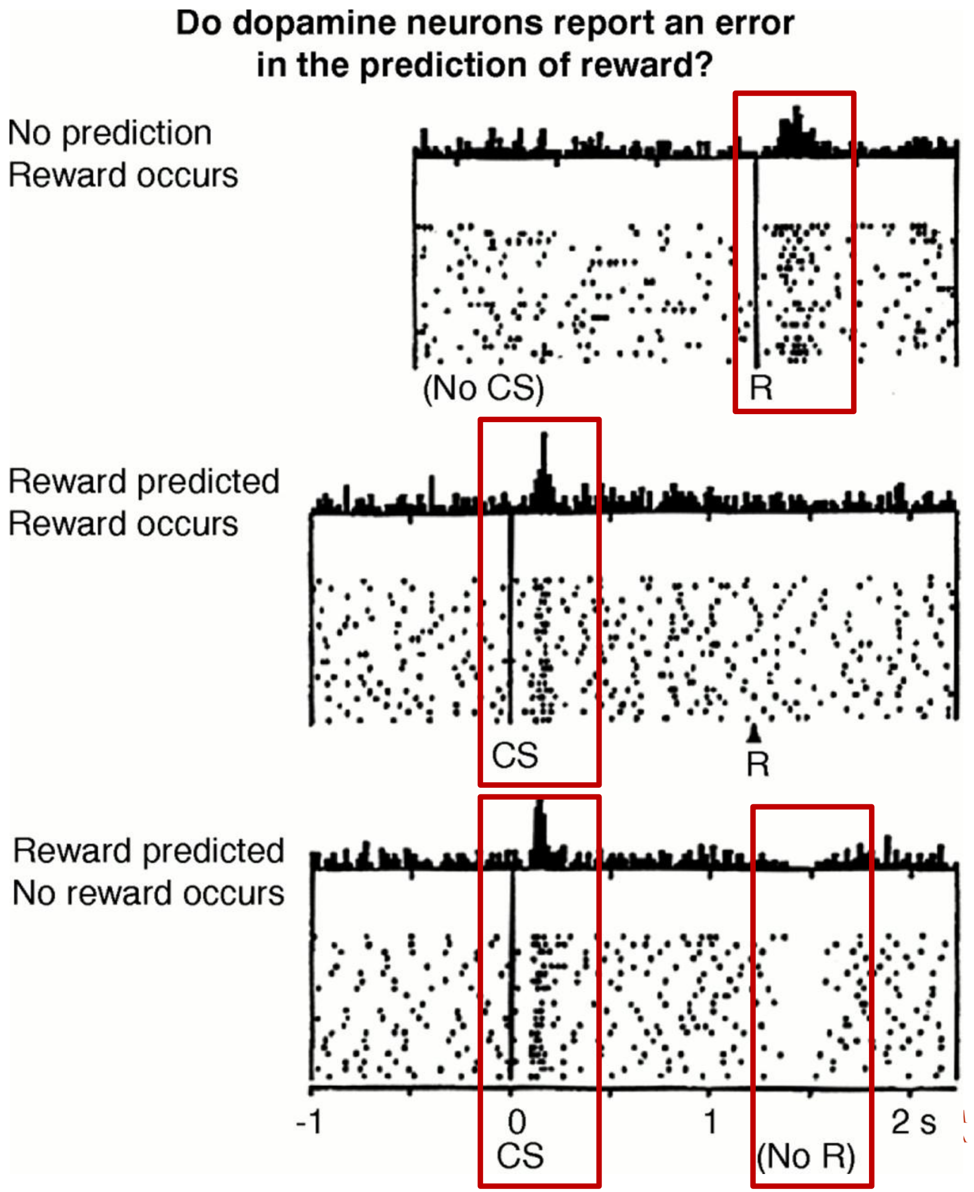
...transfers back to a cue which predicts reward occurrence (i.e. CS)

Before learning

- unexpected rewards occurs --> + PE
    - Dopamine neuron firing is increased following reward

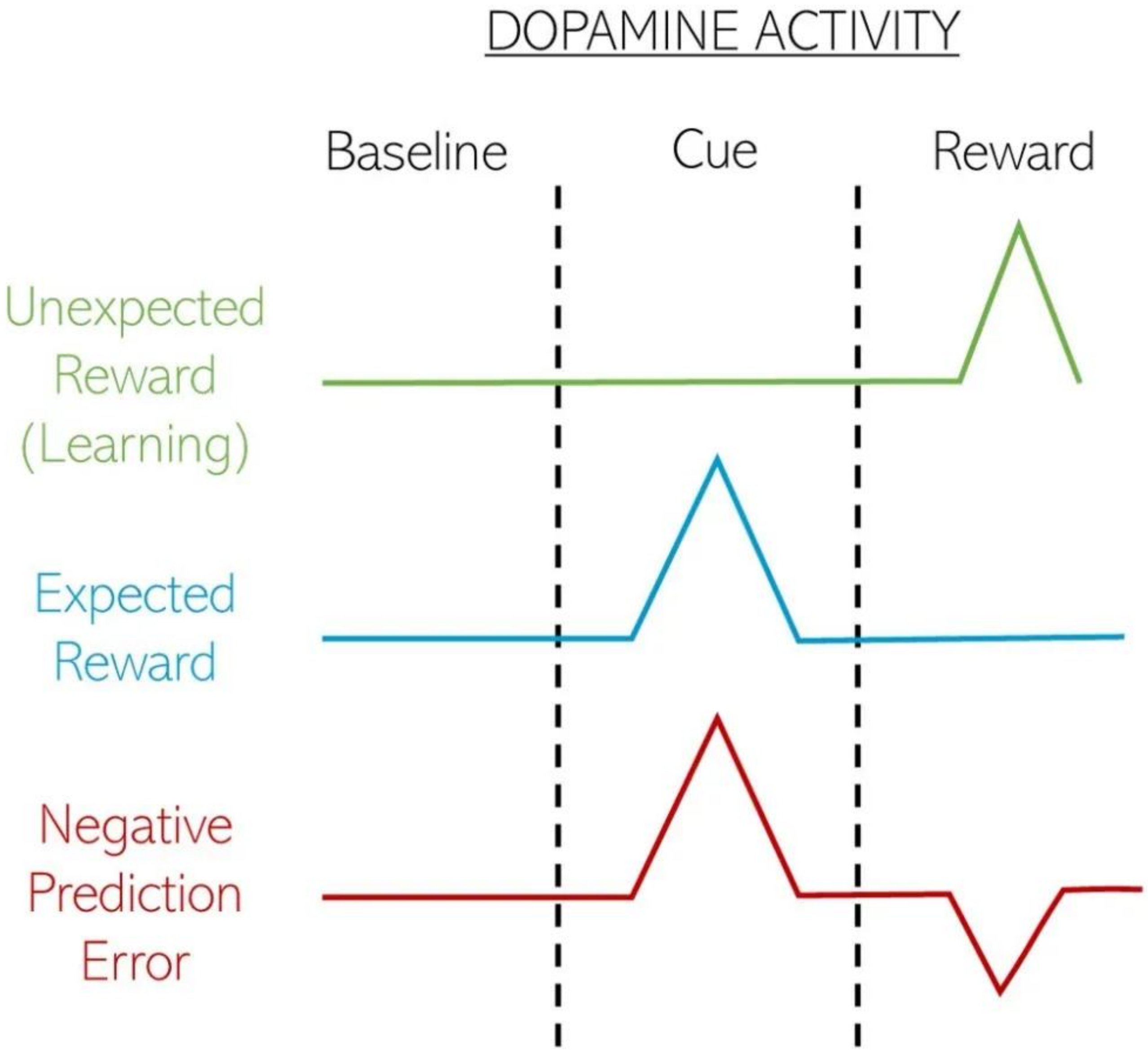
After learning

- the CS predicts reward, and the reward occurs --> no PE
    - Dopamine neuron firing is increased following CS but **not following reward.** So, **<u>dopamine ≠ reward signal</u>**
- the CS predicts reward, and the reward does not occur --> - PE
    - Dopamine neuron firing is increased following CS but decreased following omitted reward
    - Exactly at the time when reward was expected

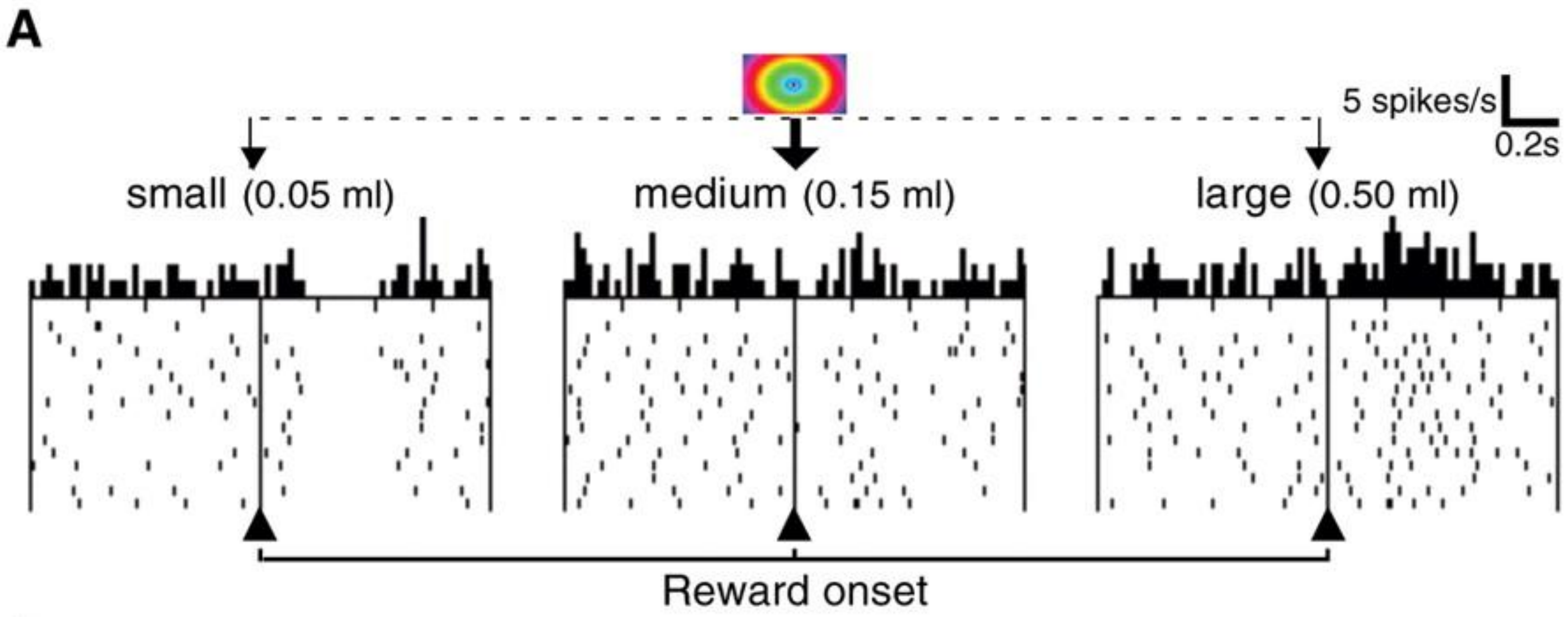


Reward prediction and subsequent dopamine activity

- Unexpected rewards increase the activity of dopamine neurons, acting as positive feedback signals for the brain regions associated with the preceding stimulus/behavior.
- As learning takes place, the timing of activity will shift until it occurs upon the cue alone, with the expected reward having no additional effect.
- Should the expected reward not be received, dopamine activity drops, sending a negative feedback signal to the relevant parts of the brain, weakening the positive association.



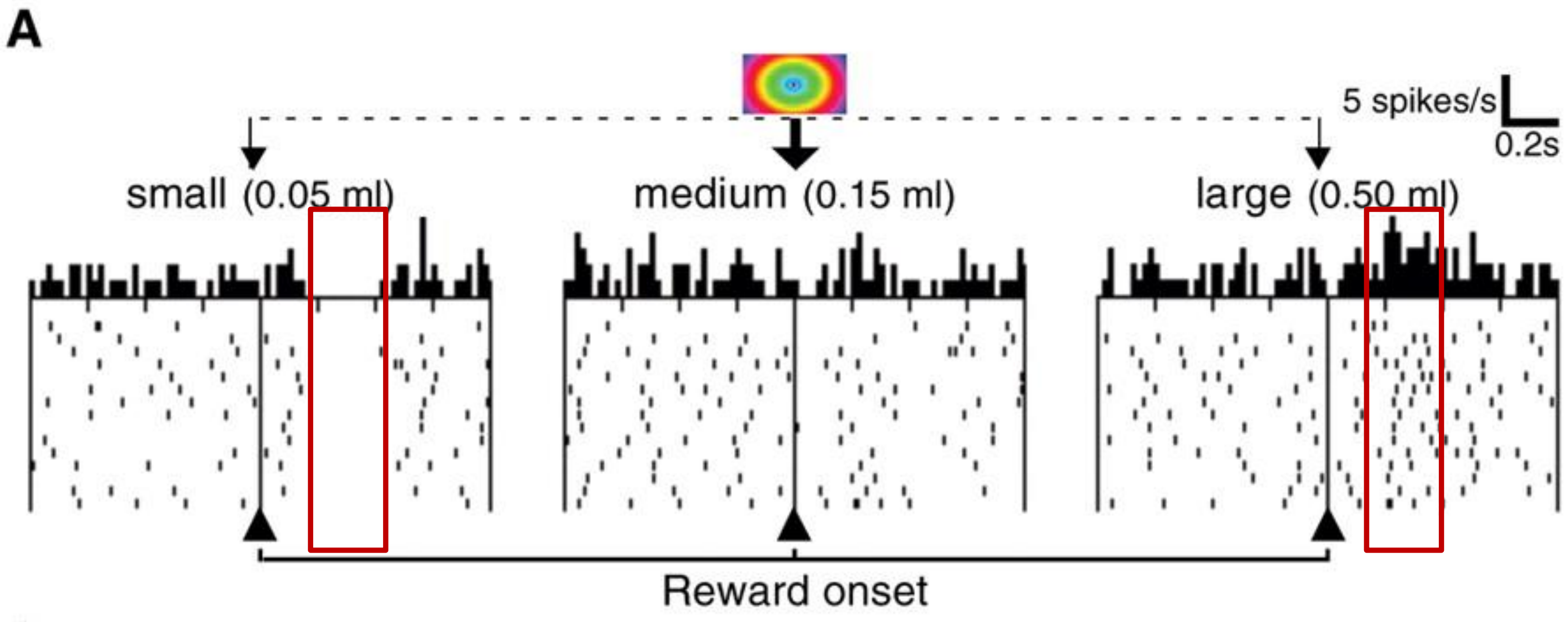
Dopaminergic neurons exhibit a signed (bidirectional) response to unexpected quantity of reward... (Tobler, Fiorillo and Schulz, Science, 2005)



A single CS was usually followed by an intermediate volume of liquid (0.15 ml) that elicited no change in the neuron's activity (center). However, on a small minority of trials, **smaller** (0.05 ml) or **larger** (0.50 ml) volumes were **unpredictably substituted, and neural activity decreased (left) or increased (right), respectively**

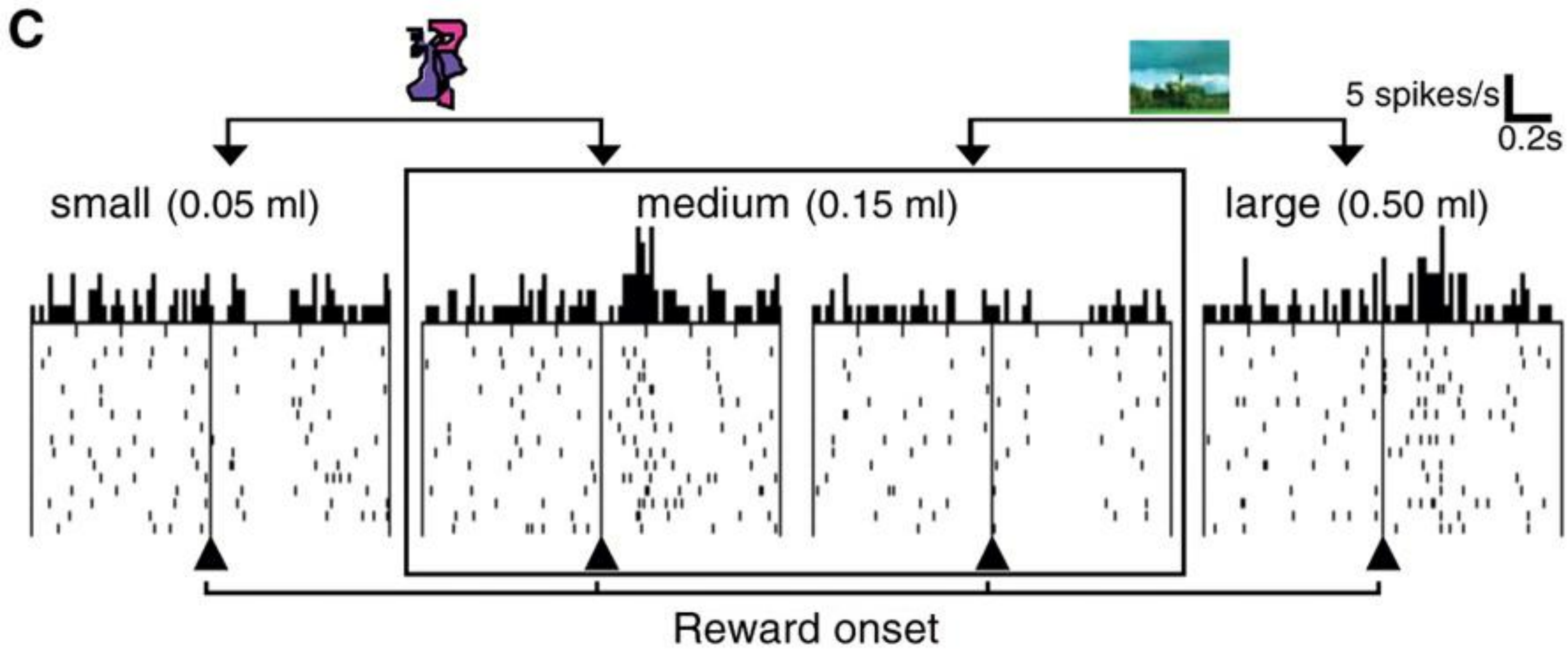
ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

Dopaminergic neurons exhibit a signed (bidirectional) response to unexpected quantity of reward... (Tobler, Fiorillo and Schulz, Science, 2005)

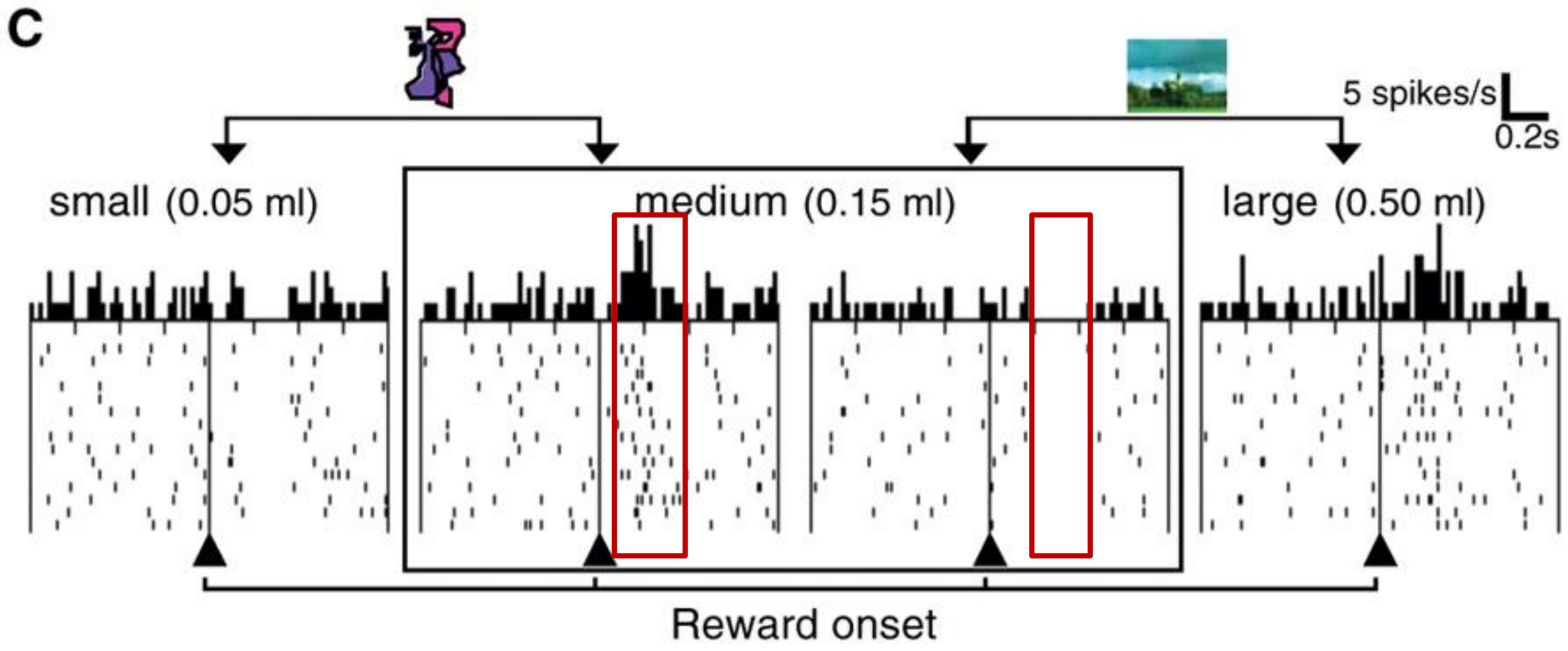


A single CS was usually followed by an intermediate volume of liquid (0.15 ml) that elicited no change in the neuron's activity (center). However, on a small minority of trials, **smaller** (0.05 ml) or **larger** (0.50 ml) volumes were **unpredictably substituted, and neural activity decreased (left) or increased (right), respectively**

...which is relative to a predicted magnitude (Tobler, Fiorillo and Schulz, Science, 2005)



...which is relative to a predicted magnitude (Tobler, Fiorillo and Schulz, Science, 2005)



Responses of a single neuron to three liquid volumes, delivered in the **context of two different predictions**. One stimulus predicted small or medium volume with equal probability, whereas another stimulus predicted medium or large volume. **The medium volume activated the neuron in one context, but suppressed activity in the other.**

- Dopamine neurons process reward magnitude relative to a predicted magnitude
- **A reward outcome that is positive on an absolute scale can nonetheless suppress the activity of dopamine neurons**

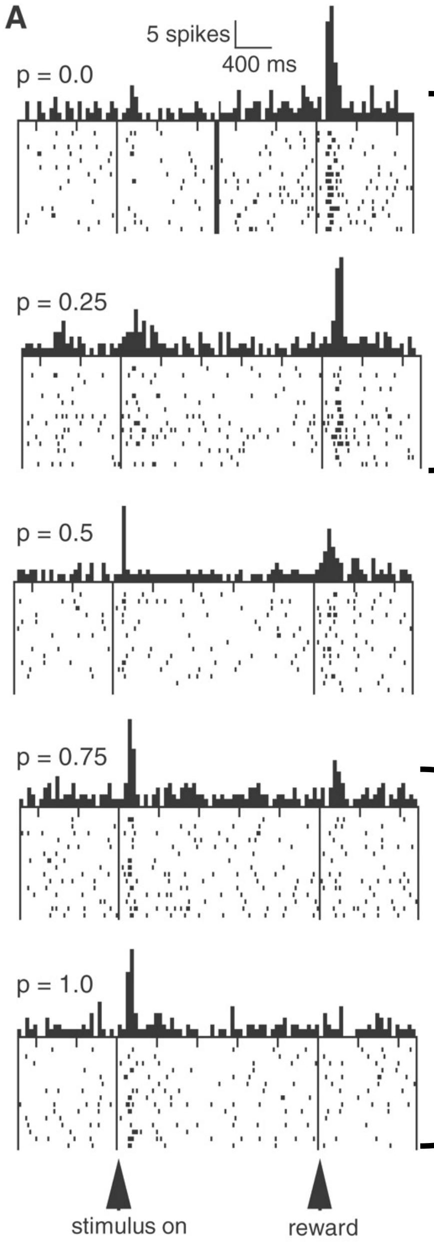
Dopaminergic neurons encode reward probability (or uncertainty)

(Fiorillo, Tobler and Schultz, Science, 2003)

<u>Phasic</u> activation of dopamine neurons vary monotonically with reward probability

Rasters and histograms of activity in a single cell, illustrating responses to the conditioned stimuli and reward at various reward probabilities, increasing from top to bottom.

Reward at *P* = 0.0 was given in the absence of any explicit stimulus at a rate constant of 0.02 per 100 ms and thus presumably occurred with a low subjective probability. Only rewarded trials are shown at intermediate probabilities. Bin width = 20 ms



No reward after CS more likely

Max uncertainty

Reward after CS more likely

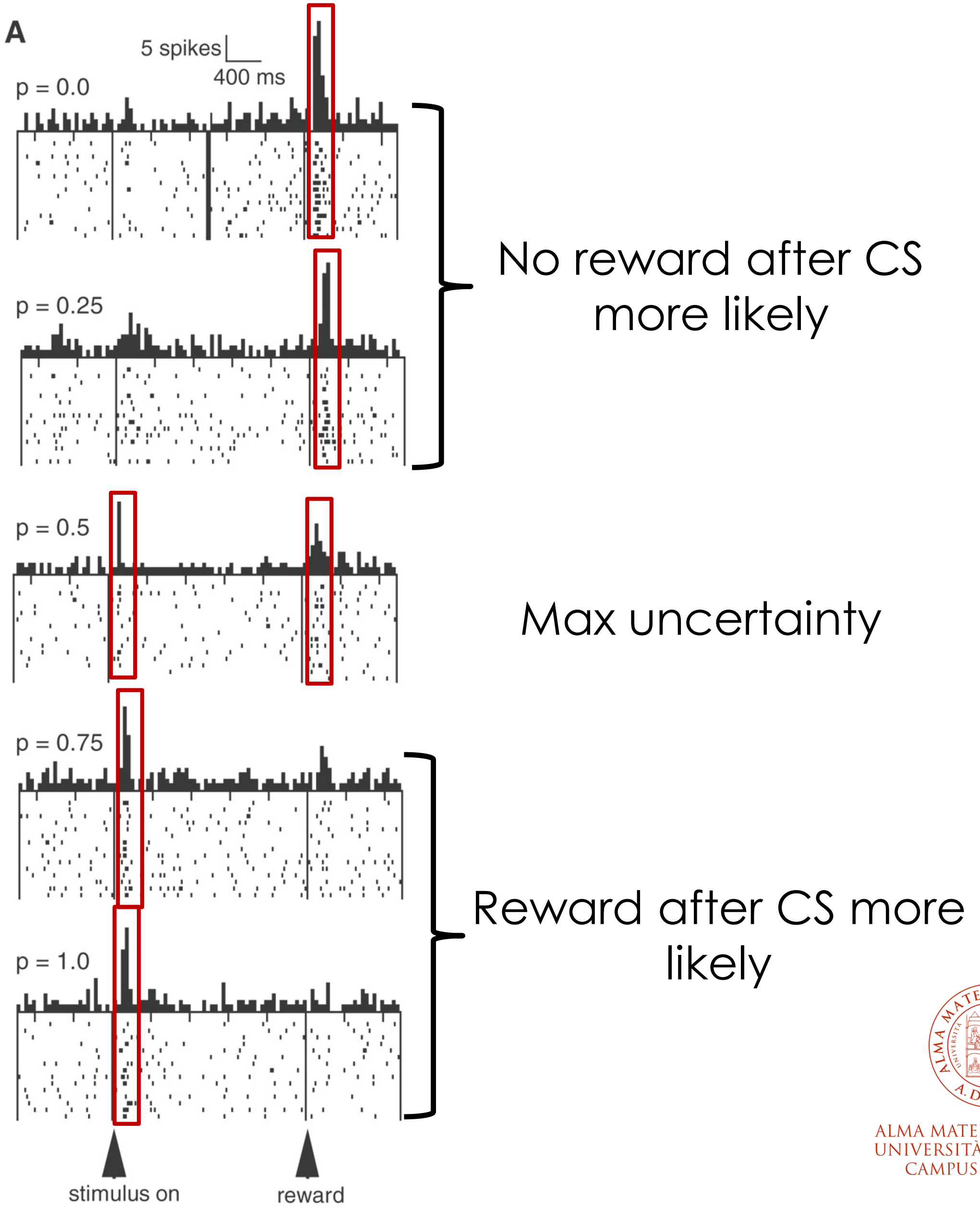
Dopaminergic neurons encode reward probability (or uncertainty)

(Fiorillo, Tobler and Schultz, Science, 2003)

<u>Phasic</u> activation of dopamine neurons vary monotonically with reward probability

Rasters and histograms of activity in a single cell, illustrating responses to the conditioned stimuli and reward at various reward probabilities, increasing from top to bottom.

Reward at *P* = 0.0 was given in the absence of any explicit stimulus at a rate constant of 0.02 per 100 ms and thus presumably occurred with a low subjective probability. Only rewarded trials are shown at intermediate probabilities. Bin width = 20 ms



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

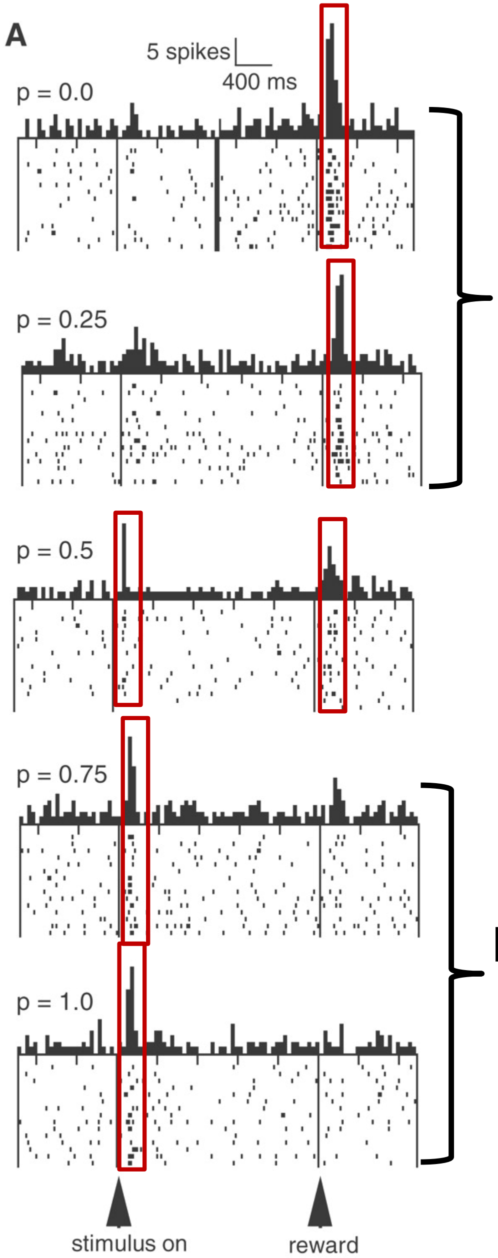
Dopaminergic neurons encode reward probability (or uncertainty)

(Fiorillo, Tobler and Schultz, Science, 2003)

<u>Phasic</u> activation of dopamine neurons vary monotonically with reward probability

Rasters and histograms of activity in a single cell, illustrating responses to the conditioned stimuli and reward at various reward probabilities, increasing from top to bottom.

Reward at *P* = 0.0 was given in the absence of any explicit stimulus at a rate constant of 0.02 per 100 ms and thus presumably occurred with a low subjective probability. Only rewarded trials are shown at intermediate probabilities. Bin width = 20 ms



No reward after CS more likely

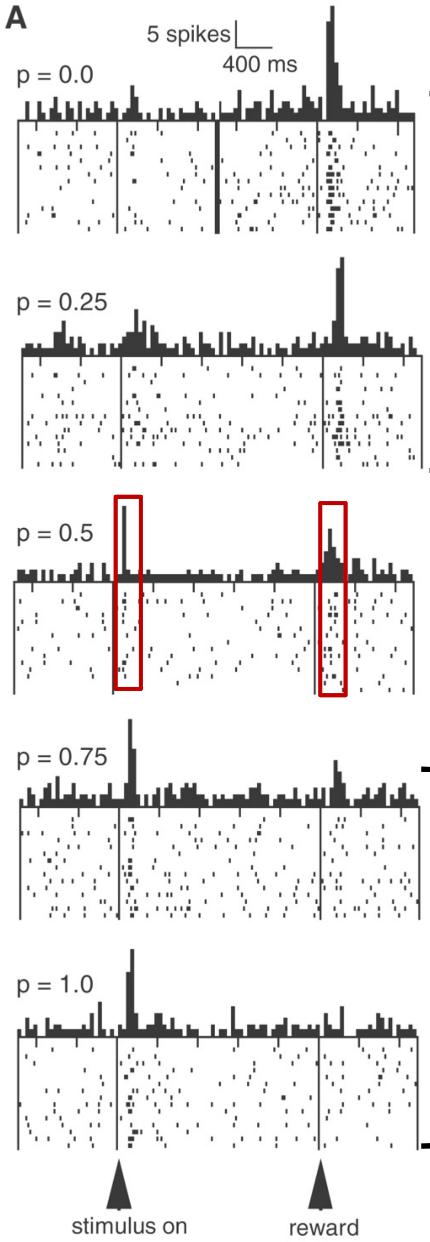
Max uncertainty

Reward after CS more likely

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

Dopaminergic neurons encode reward probability (or uncertainty)

(Fiorillo, Tobler and Schultz, Science, 2003)

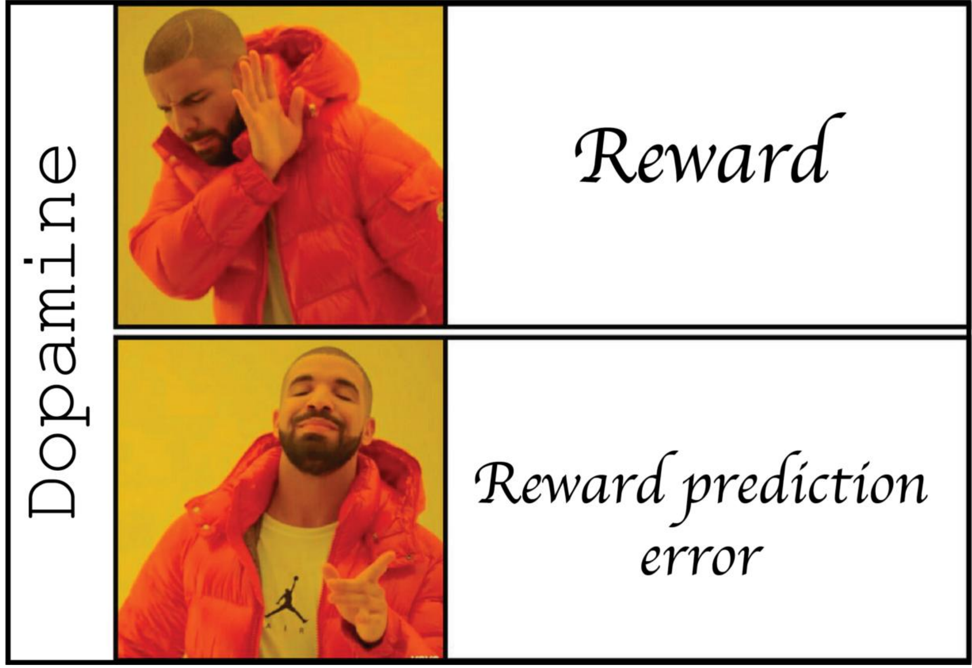


No reward after CS more likely

Max uncertainty

Reward after CS more likely

Summary



Dopamine neurons

- **DO NOT** broadcast a **reward signal**
- **DO** broadcast a **prediction error signal**

Dopaminergic neurons exhibit changes in phasic response to

- an unexpected reward
    - discriminating between **reward & no-reward**
    - discriminating between reward **magnitude** (more or less reward)
        - In a relative way, rather than absolute
    - discriminating between reward **probability** (more or less likely to get reward)
    - discriminating between reward **timing** (earlier or later reward)
    - in a <u>signed manner</u>
        - Increase firing to unexpected delivery (positive PE)
        - pause firing to unexpected omission (negative PE)
- **transfers back** to a cue which predicts reward occurrence (i.e. CS)
    - enabling associative (reinforcement) learning

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

From predictions to control and decision-making

So far we have talked about the role of dopamine only when learning to predict rewards (Pavlovian learning), showing evidence that dopaminergic activity changes to unexpected reward delivery and omission

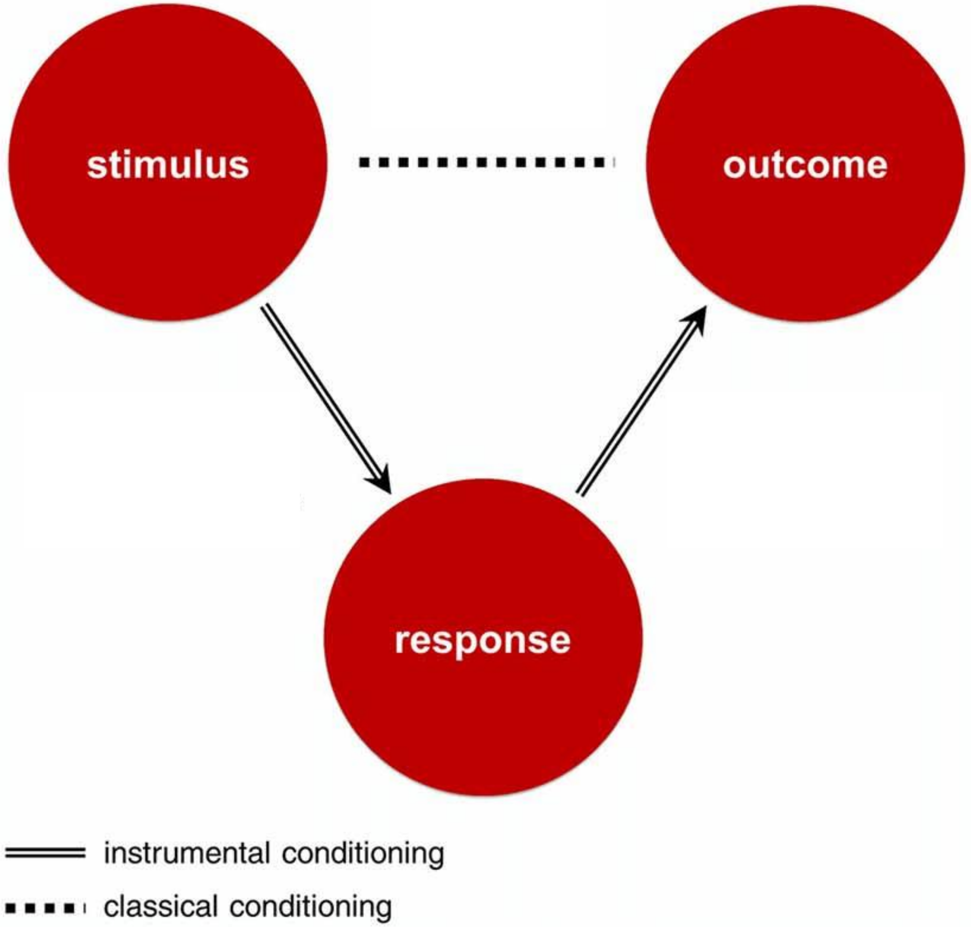
BUT

- Is there evidence that learned predictions actually drive decision-making (choice behavior / instrumental learning)?

AND

- If dopamine encodes prediction errors that drive learning about reward predictions, is dopamine causally involved in decision-making?

Predictions are for control



The causal role of dopamine in driving reinforcement learning and choice behavior

The behavioral and neural effects of **three different drug conditions** on instrumental learning were assessed in three groups of healthy participants:

1. Enhanced dopaminergic function: 3,4-dihydroxy-L-phenylalanine (L-DOPA)
2. Reduced dopaminergic function: Haloperidol
3. Control condition: Placebo

Pessiglione, Seymour, Flandin, Dolan & Frith. *Nature* (2006). https://doi-org.ezproxy.unibo.it/10.1038/nature05051

The causal role of dopamine in driving reinforcement learning and choice behavior

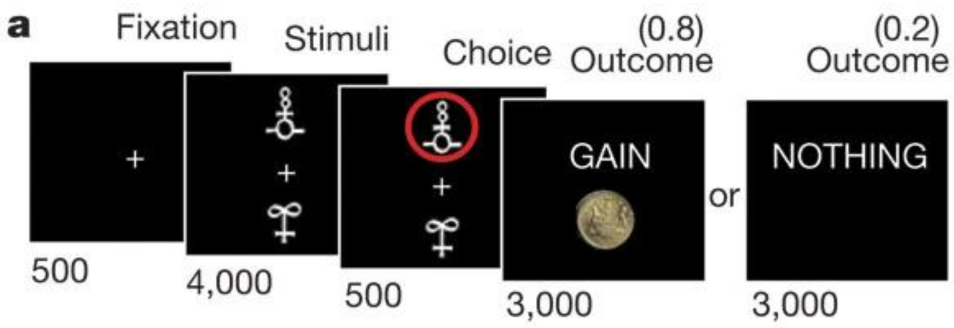


Fig. a - Experimental Task: Participants chose between two abstract visual stimuli displayed on a screen (upper or lower) and observed the outcome.

Each stimulus was linked to a specific probability of gain or loss to assess:

- *Reward-based learning – Gain condition*: One pair of stimuli was associated with gains (£1 or nothing) → Assessing drug effects on the ability to learn from rewards
- *Punishment-based learning – Loss condition*: A second pair was associated with losses (-£1 or nothing) → Assessing drug effects on the ability to learn from punishments
- *Neutral condition*: A third pair had no financial outcome → Serving as a neutral control

The causal role of dopamine in driving reinforcement learning and choice behavior

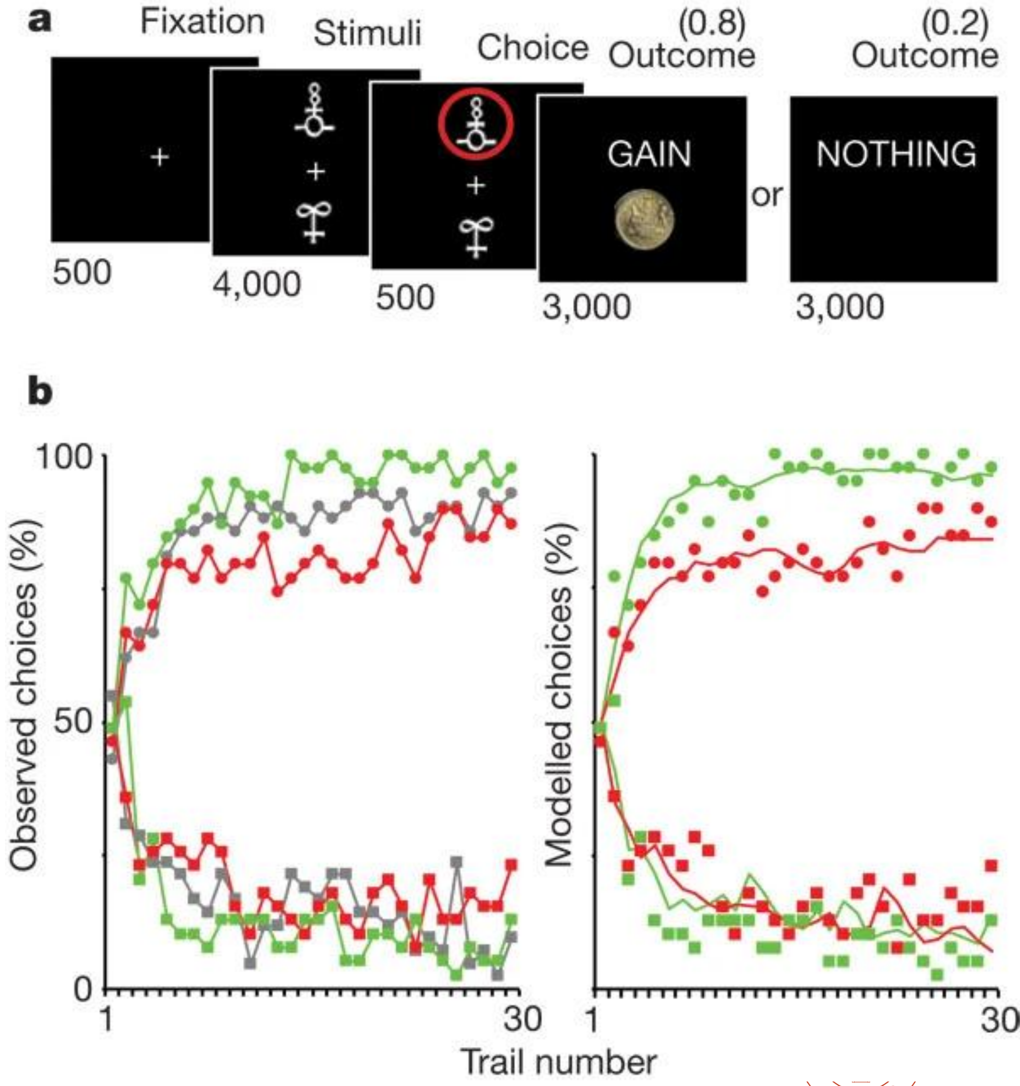


Fig. b - Behavioural Results

- **Left Panel: Observed behavioural choices under placebo (grey), L-DOPA (green), and haloperidol (red).**

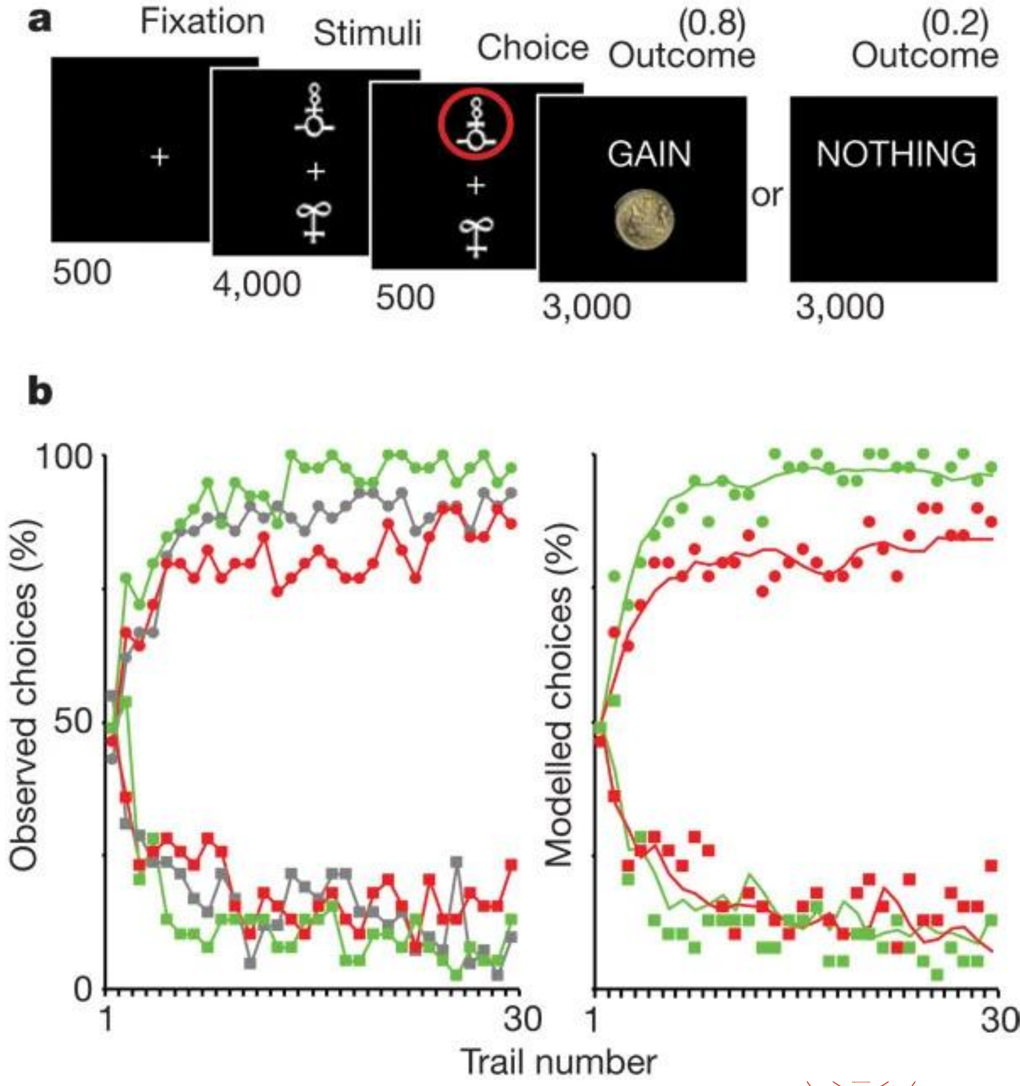
  The learning curves track, trial by trial, the proportion of subjects selecting:
  - The 'correct' stimulus in the *gain condition* (circles, upper graph).
  - The 'incorrect' stimulus in the *loss condition* (squares, lower graph).
- **Right Panel: Modelled behavioural choices for L-DOPA (green) and haloperidol (red) groups.**
  - The learning curves show probabilities predicted by a computational model.
  - Circles and squares representing observed choices are included for comparison.

Pessiglione, Seymour, Flandin, Dolan & Frith. *Nature* (2006). https://doi-org.ezproxy.unibo.it/10.1038/nature05051

The causal role of dopamine in driving reinforcement learning and choice behavior

Overall the study shows that **dopamine plays a causal role in reward-based learning and decision-making**:

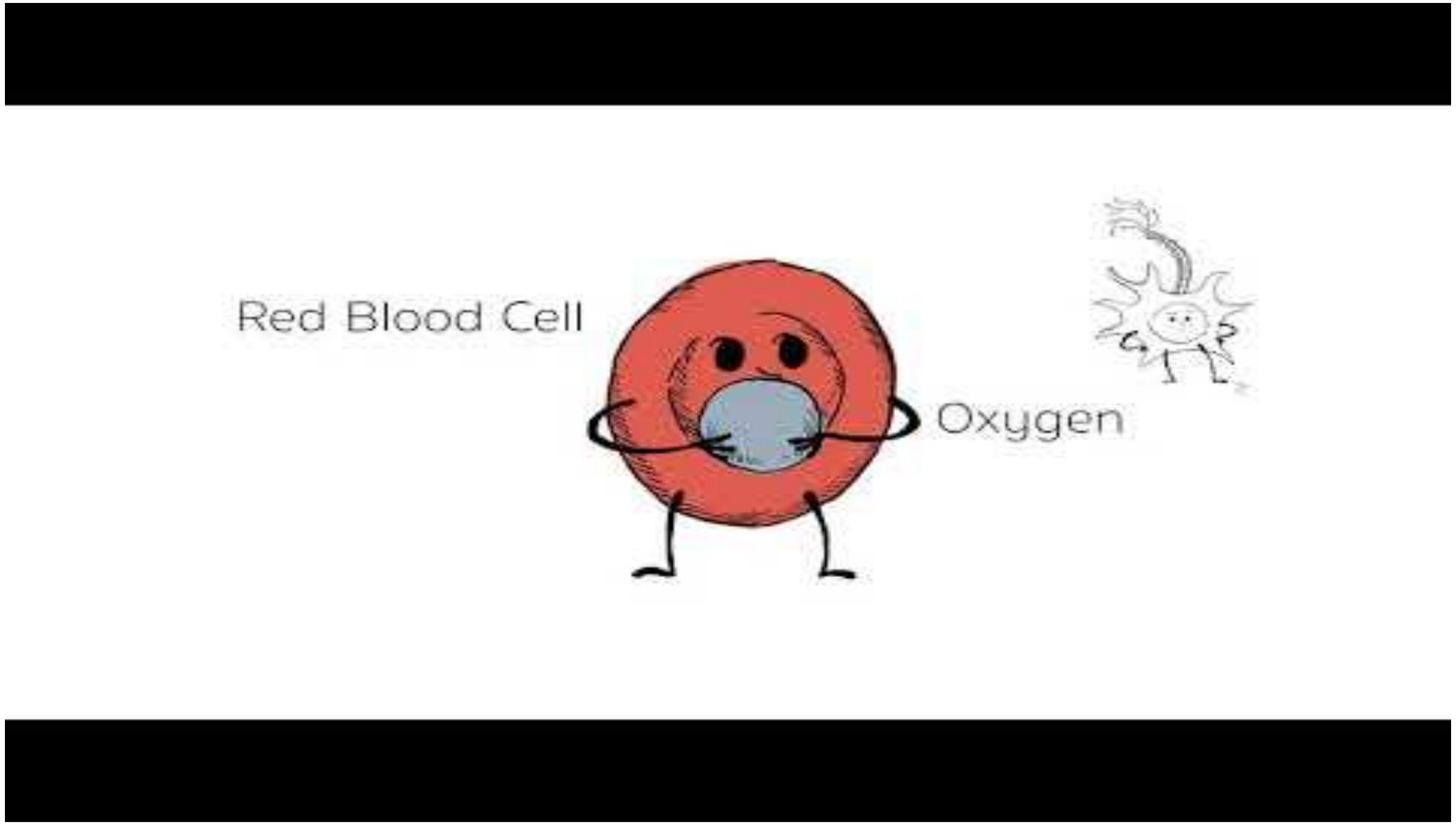
- Enhancing dopaminergic function with **L-DOPA** enhances learning from rewards, improving selection of high-reward stimuli.
- Reducing dopaminergic function with **haloperidol** impairs learning from rewards, reducing the ability to choose high-reward stimuli.
- Learning from punishment is less influenced by dopaminergic manipulation:
    - No significant effects of L-DOPA or haloperidol on avoidance of loss-related stimuli.
- Computational modelling confirms behavioural patterns: Changes in learning rates align with observed effects of L-DOPA and haloperidol.



Pessiglione, Seymour, Flandin, Dolan & Frith. *Nature* (2006). https://doi-org.ezproxy.unibo.it/10.1038/nature05051

Functional Magnetic Resonance Imaging (fMRI)

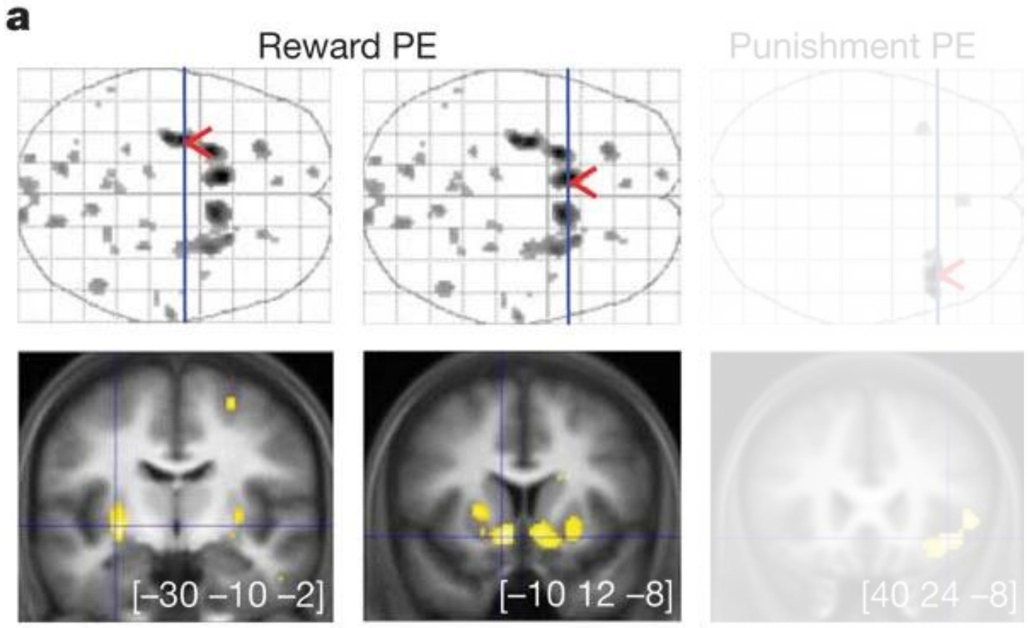
- Measures the ratio of oxygenated to deoxygenated hemoglobin
    - this value is referred to as the blood oxygen level–dependent (**BOLD**) signal
    - Indirect measure of neuronal activity
- Correlational evidence
- High spatial resolution but low temporal resolution
    - Appropriate to know where things happen but not when things happen



https://youtu.be/4UOeBM5BwdY

The causal role of dopamine in driving reinforcement learning and choice behavior

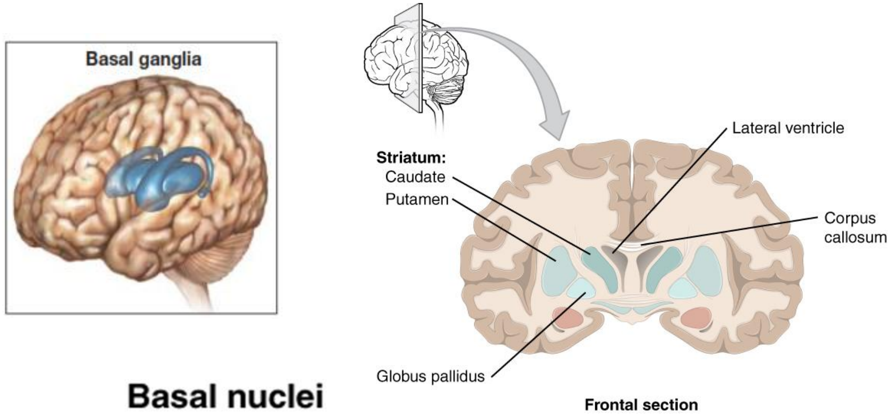
- Then, a standard action-value learning algorithm was fitted to the observed behaviour.
- Outcome prediction errors estimated from the model were used as statistical regressors in neuro-imaging data (fMRI) to find the **neural correlates of prediction errors.**
- **Reward prediction errors** (i.e. gain obtainment & loss avoidance) correlated with activity in:
  - Left posterior putamen
  - Left ventral striatum
- **Punishment prediction errors** correlated with activity in:
  - Right anterior insula
- These findings corroborate the role of the striatum in reinforcement learning and decision-making.

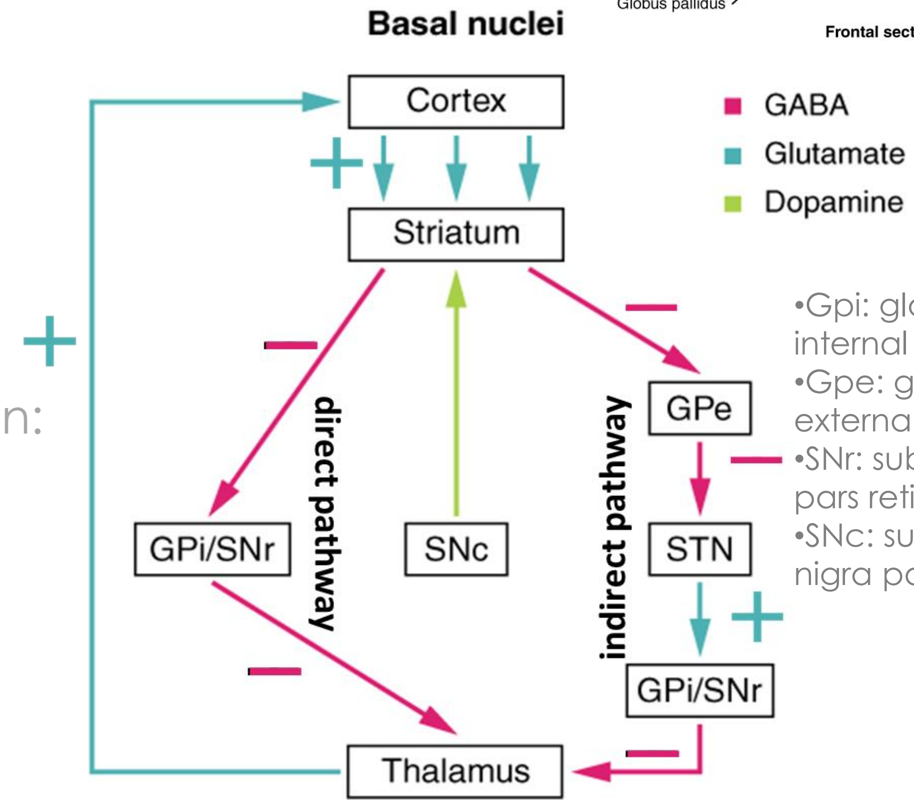


Pessiglione, Seymour, Flandin, Dolan & Frith. *Nature* (2006). https://doi-org.ezproxy.unibo.it/10.1038/nature05051

The causal role of dopamine in driving reinforcement learning and choice behavior

- Then, a standard action-value learning algorithm was fitted to the observed behaviour.
- Outcome prediction errors estimated from the model were used as statistical regressors in neuro-imaging data (fMRI) to find the **neural correlates of prediction errors.**
- **Reward prediction errors** (i.e. gain obtainment & loss avoidance) correlated with activity in:
    - Left posterior putamen
    - Left ventral striatum
- **Punishment prediction errors** correlated with activity in:
    - Right anterior insula
- These findings corroborate the role of the striatum in reinforcement learning and decision-making.





•Gpi: globus pallidus internal segment
•Gpe: globus pallidus external segment
•SNr: substantia nigra pars reticulata
•SNc: substantia nigra pars compacta

Pessiglione, Seymour, Flandin, Dolan & Frith. *Nature* (2006). https://doi-org.ezproxy.unibo.it/10.1038/nature05051

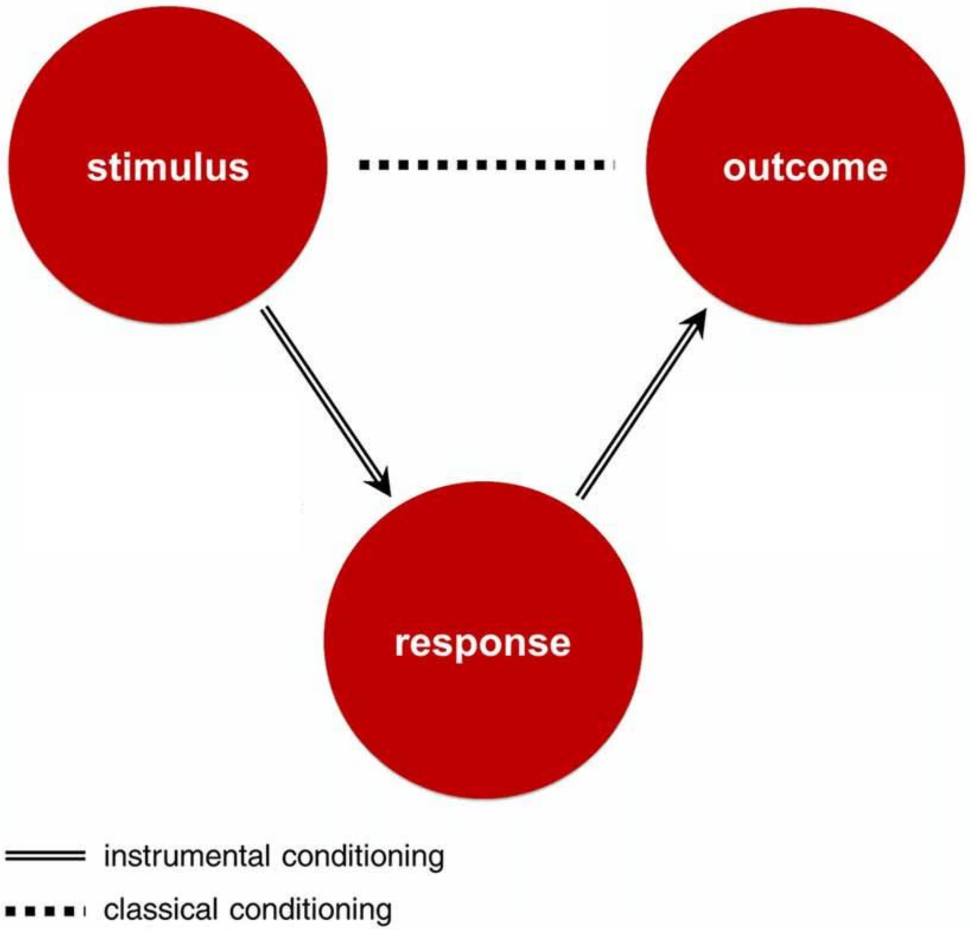
From predictions to control and decision-making

The experiments we have discussed show that dopamine has a crucial role in

- learning stimulus-outcome associations: Pavlovian/prediction learning
- learning response-outcome associations: instrumental/control learning

Together these learning system contribute to optimal decision-making

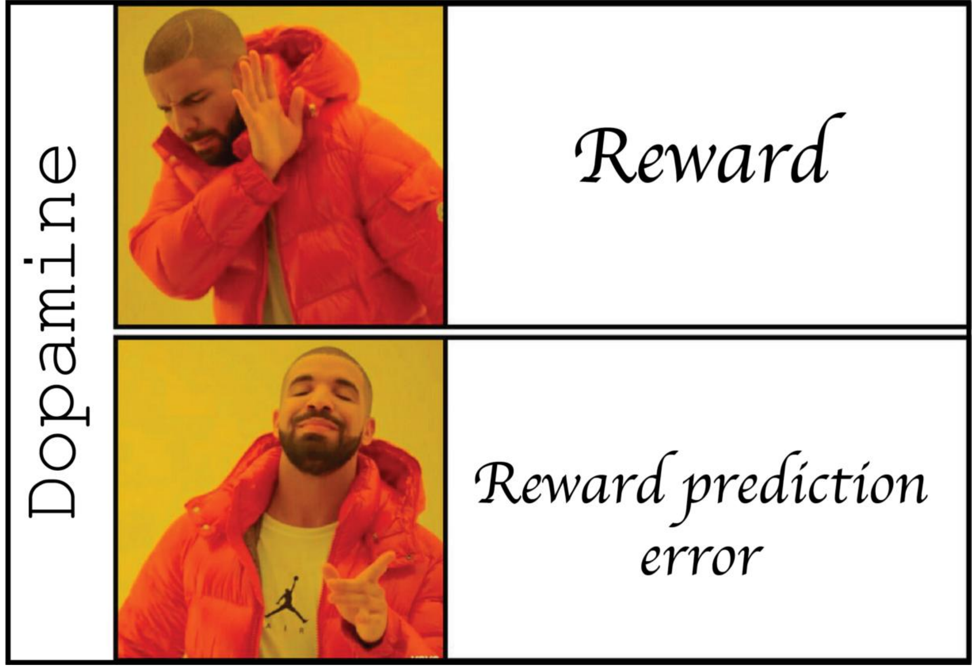
Predictions are for control



Dopaminergic encoding of reward prediction errors: insights into addiction mechanisms

https://youtu.be/aqXmOb_fuN4

Dopaminergic encoding of reward prediction errors: insights into addiction mechanisms



Dopamine
- is not the feel-good chemical
- does not simply make us feel good
- does not simply tell us how much we like something

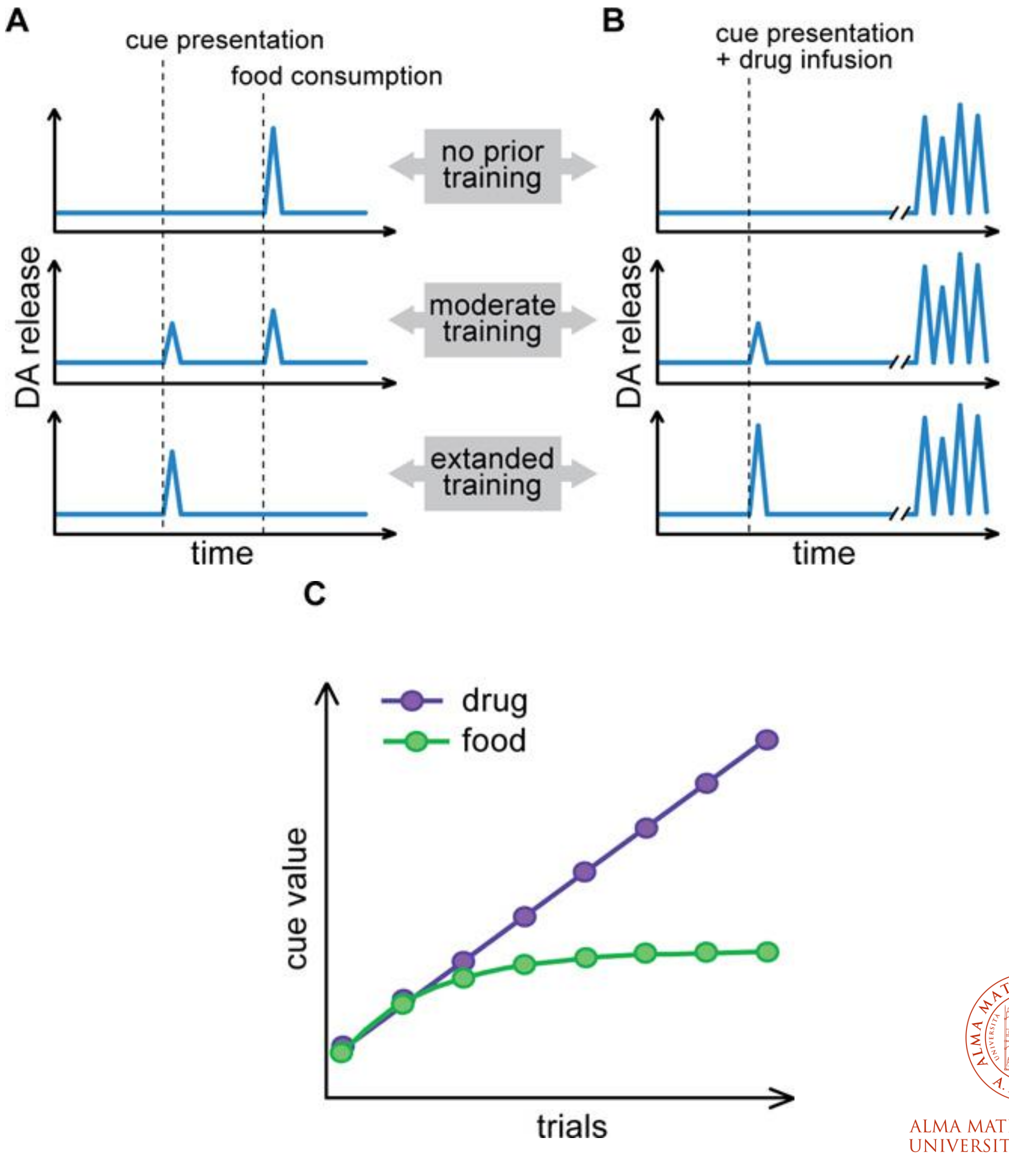
Dopamine
- make us feel good when something valuable unexpectedly happened
- make events that unexpectedly predicted rewards relevant (and disregard events that did not)
- teaches us where to find things we need or like
- by conveying Pes

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

Dopaminergic encoding of reward prediction errors: insights into addiction mechanisms

Differences between food-related and cocaine-related phasic DA signals

A. Prior to learning unexpected delivery of food reward results in phasic DA signals. As subjects learn that cue presentation signals food delivery, DA responses are transferred from the reward to the cue.

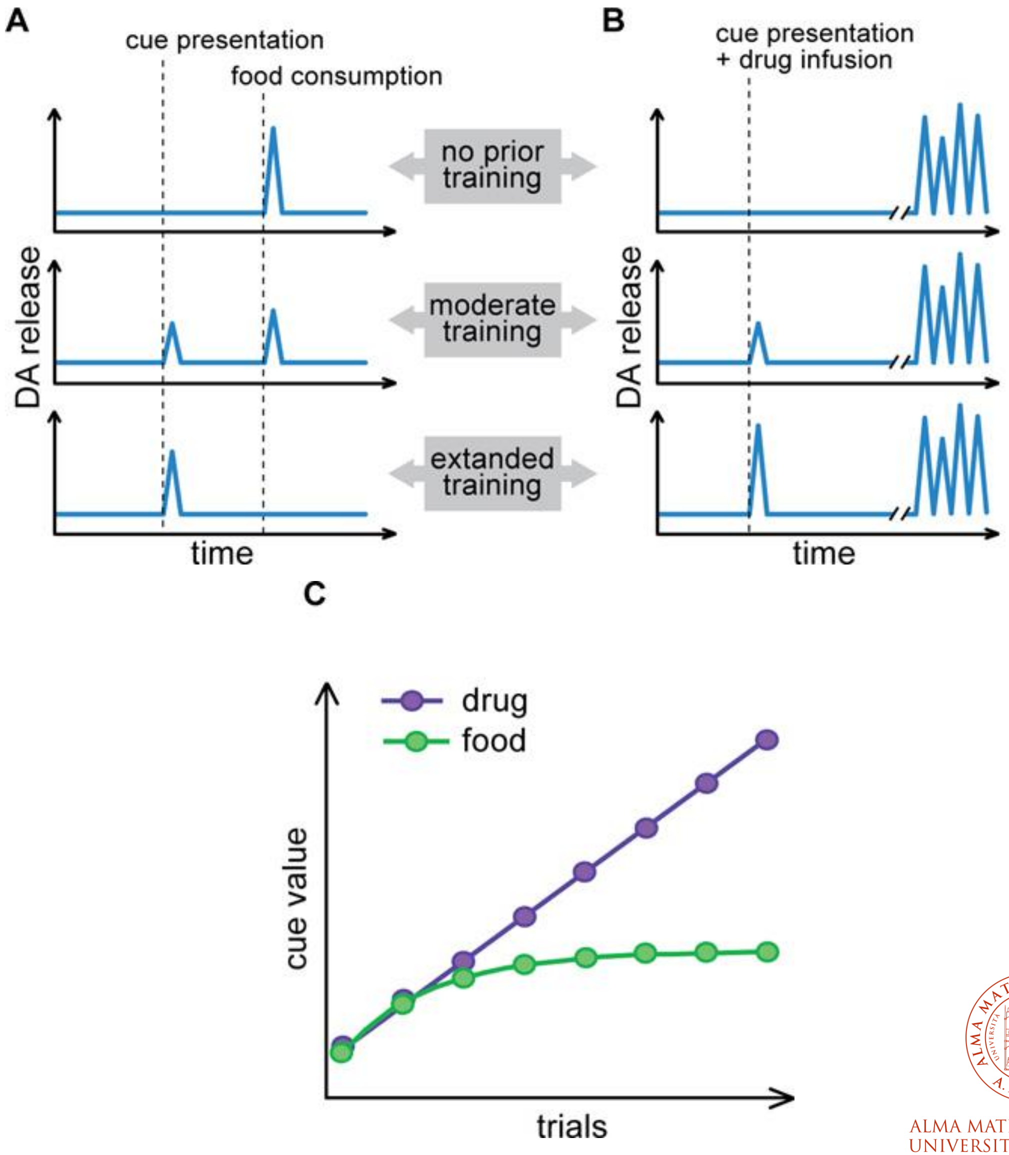


Keiflin R, Janak PH. Dopamine Prediction Errors in Reward Learning and Addiction: From Theory to Neural Circuitry. *Neuron*. 2015;88(2):247-263. doi:10.1016/j.neuron.2015.08.037

Dopaminergic encoding of reward prediction errors: insights into addiction mechanisms

Differences between food-related and cocaine-related phasic DA signals

B. When cocaine is the reward, each drug injection produces, with some delay, a burst of phasic DA events as a consequence of the pharmacological actions of the drug. As with natural reward, phasic DA responses progressively emerge to the cue. Unlike food-induced DA signals, drug-induced DA signals are not modulated by expectations and persist throughout learning.

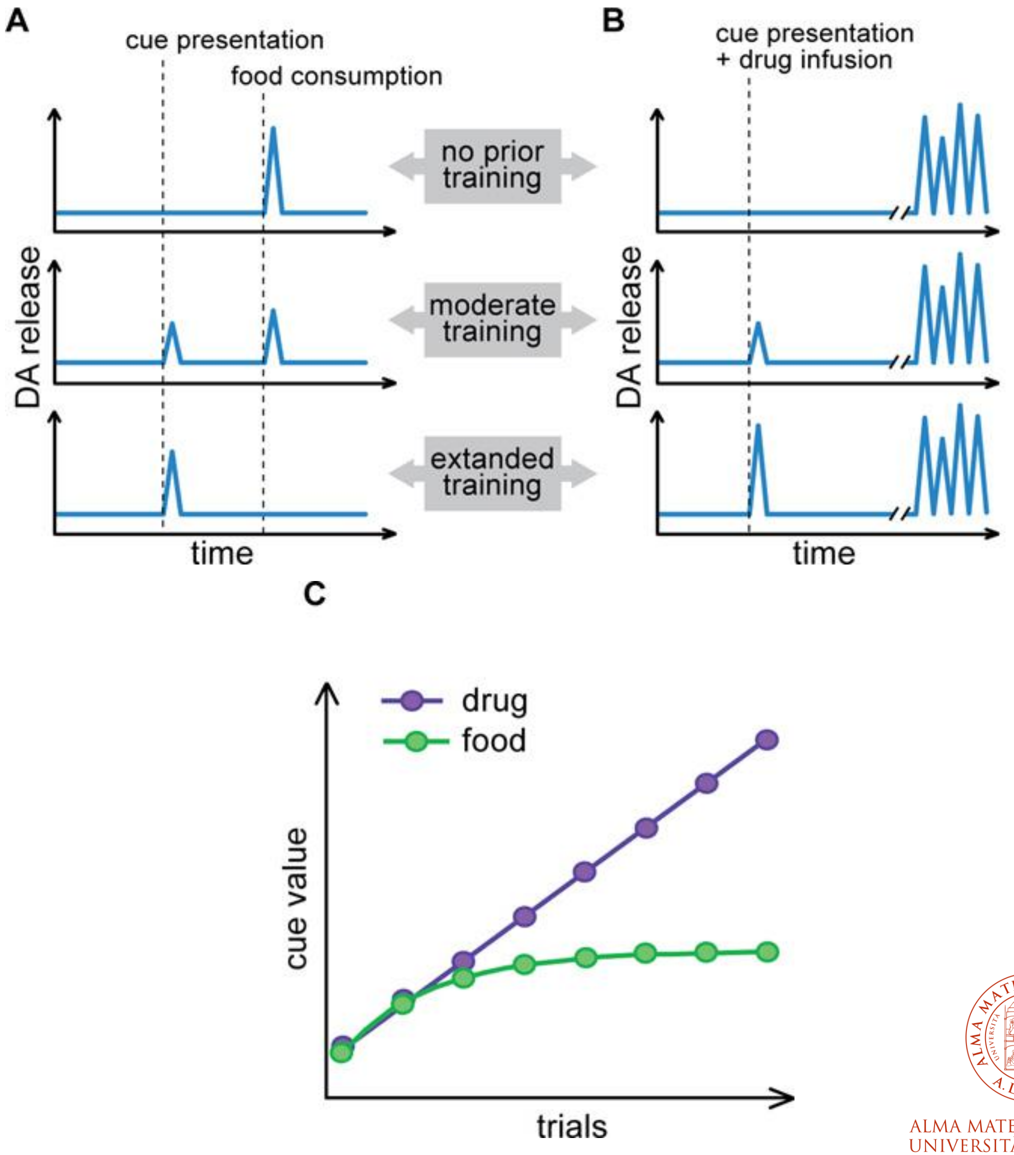


Keiflin R, Janak PH. Dopamine Prediction Errors in Reward Learning and Addiction: From Theory to Neural Circuitry. *Neuron*. 2015;88(2):247-263. doi:10.1016/j.neuron.2015.08.037

Dopaminergic encoding of reward prediction errors: insights into addiction mechanisms

Differences between food-related and cocaine-related phasic DA signals

C. Proposed consequences of these DA signals on learning. Food-evoked DA signals modulated by reward expectations promote learning until the prediction matches the actual outcome, resulting in stable cue value after a few trials. In contrast, persistent cocaine-evoked DA signals continue to increase the value of cocaine cues with every trial. Eventually, the value of cocaine cues surpasses the value of the food cues and can bias decision-making towards cocaine.

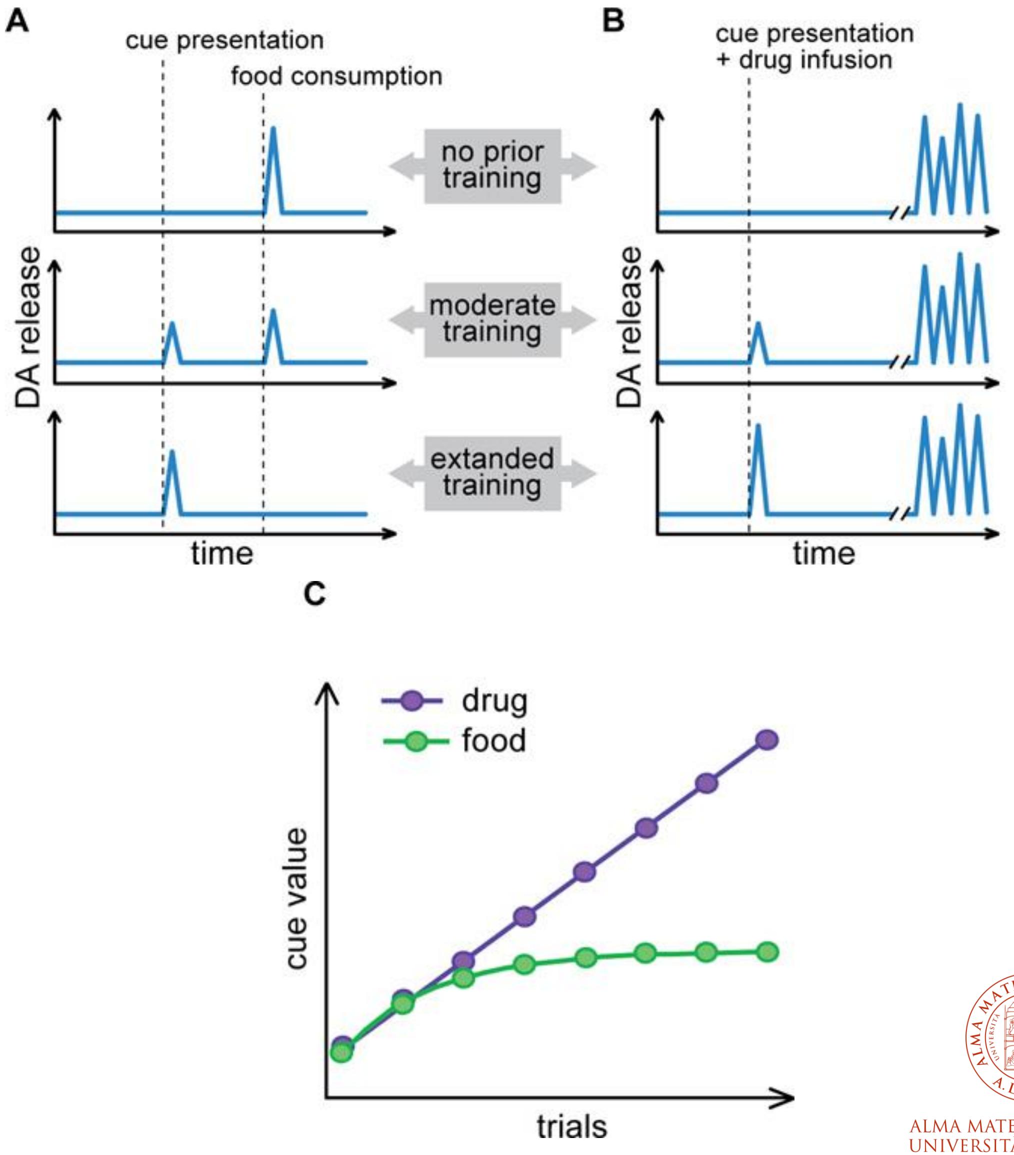


Keiflin R, Janak PH. Dopamine Prediction Errors in Reward Learning and Addiction: From Theory to Neural Circuitry. *Neuron*. 2015;88(2):247-263. doi:10.1016/j.neuron.2015.08.037

Dopaminergic encoding of reward prediction errors: insights into addiction mechanisms

Dopaminergic Dysregulation in Addiction

- Drugs of abuse hijack the RPE system, artificially triggering large dopamine surges.
- These exaggerated dopamine responses reinforce drug-seeking behaviours, even when the expected reward is fully predictable.
- Over time, natural rewards become less effective, as the brain's reward system adapts to drug-induced dopamine spikes.



Recommended readings

- Daw, N. D., & Tobler, P. N. (2014). Value learning through reinforcement: the basics of dopamine and reinforcement learning. In Neuroeconomics (pp. 283-298). Academic Press.

- Schultz, W. (2016). Dopamine reward prediction error coding. Dialogues in clinical neuroscience, 18(1), 23-32.

- Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction. MIT press.
  - Chapter 15
  - Sections 15.1, 15.2, 15.3, 15.4, 15.5, 15.6

Revision questions

- Discuss the experimental evidence that supports the reward prediction error hypothesis of dopamine neuron activity.
- Explain the role of dopamine and the basal ganglia in reinforcement learning and goal-oriented behavior. How do the direct and indirect pathways contribute to this process?
- Critically evaluate the statement: "Dopamine is the 'feel-good' neurotransmitter." How does the reward prediction error hypothesis challenge this view, and what are the implications for understanding reward processing?
- Describe how the reward prediction error framework can be applied to understand the neural mechanisms underlying drug addiction. What are the key differences in dopamine signaling between natural rewards and addictive drugs?
- Discuss the relationship between prediction errors, synaptic plasticity, and learning. How does dopamine contribute to these processes at a neural circuit level?

Glossary of key terms (not encountered in previous lectures)

- Caudate Nucleus: A component of the striatum, involved in motor control, learning, and reward.
- Disinhibition: The removal of inhibition, leading to an increase in activity in the target neurons.
- Globus Pallidus (GPe & Gpi): Parts of the basal involved in motor control, learning, and reward.
- Medial Prefrontal Cortex: The frontmost part of the prefrontal cortex, involved in executive functions and decision-making.
- Nucleus Accumbens: A key brain region in the reward pathway, receiving dopaminergic input from the VTA.
- Orbitofrontal Cortex: A part of the prefrontal cortex involved in decision-making and evaluating the value of rewards and punishments.
- Striatum: A major input structure of the basal ganglia, consisting of the caudate nucleus and putamen, involved in motor control, learning, and reward.
- Substantia Nigra Pars Compacta (SNc): A region in the midbrain containing dopamine-producing neurons that project to the striatum.

Glossary of key terms (not encountered in previous lectures)

- Substantia Nigra Pars Reticulata (SNr): A region in the midbrain that serves as a major output nucleus of the basal ganglia, involved in inhibiting unwanted movements.
- Thalamus: A brain structure that relays sensory and motor signals to the cerebral cortex and plays a role in regulating consciousness, sleep, and alertness.
- Ventral Striatum: Another term for the nucleus accumbens, emphasizing its location.
- Ventral Tegmental Area (VTA): A region in the midbrain containing dopamine-producing neurons that project to the nucleus accumbens, prefrontal cortex, and other areas involved in reward and motivation.