ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

The reward prediction error hypothesis of dopamine neurons

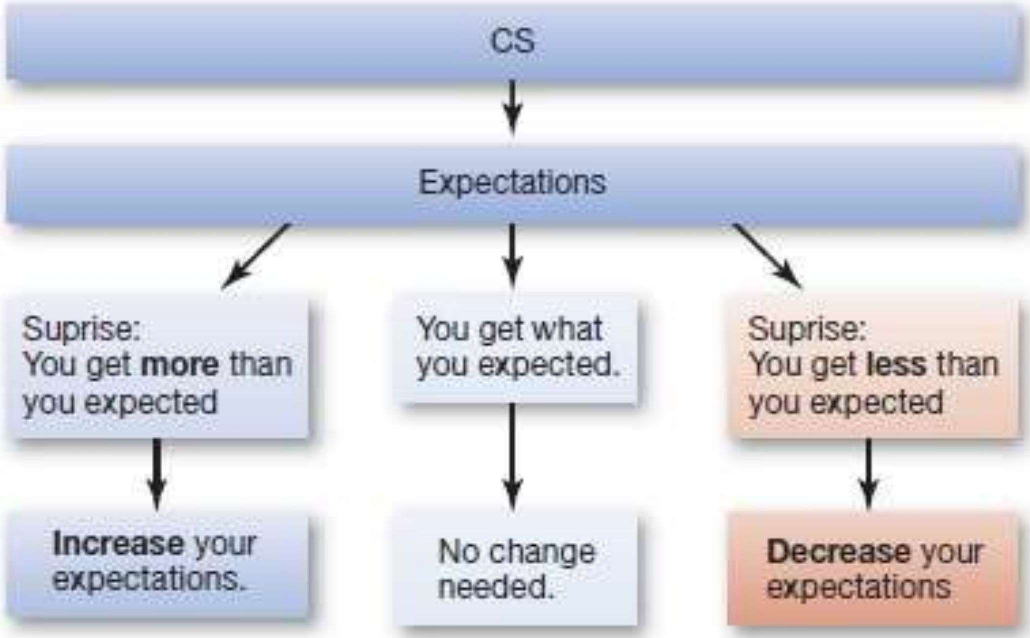
Cognition and Neuroscience
Academic year 2023/2024

Francesca Starita

francesca.starita2@unibo.it

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?

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

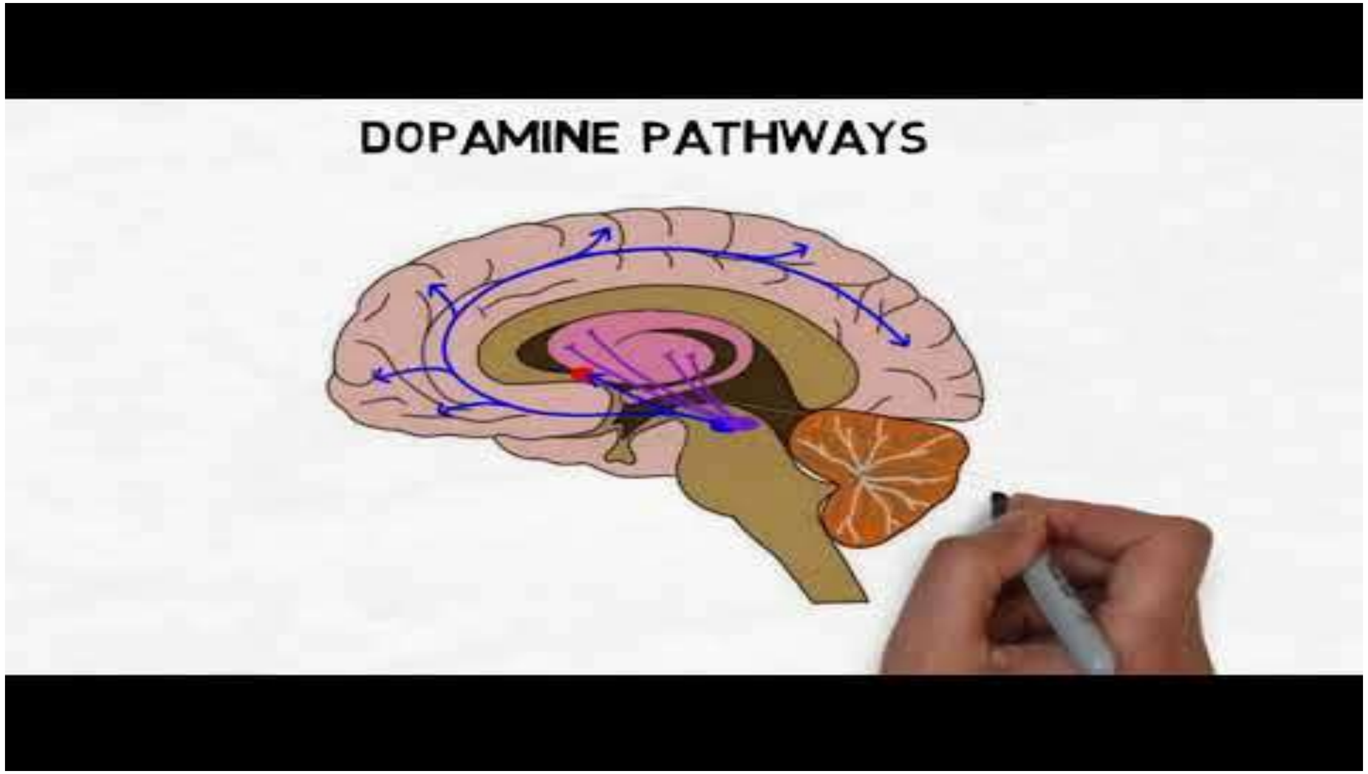
The parameters, or weights, adjusted by learning algorithms correspond to synaptic efficacies

The reward prediction error hypothesis of dopamine neuron activity

REWARD PREDICTION ERROR PROCESSING

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**

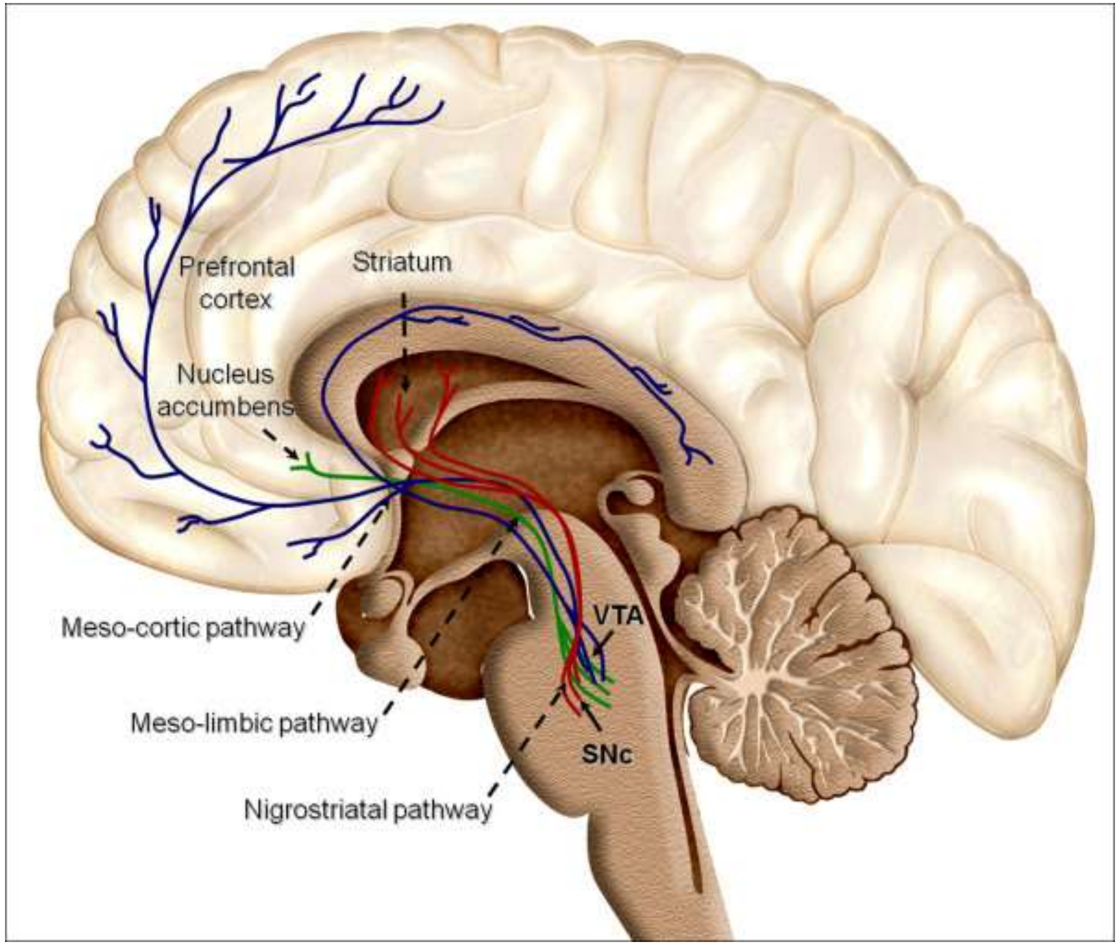
- originates in the substantia nigra pars compacta (SNc)
- projects primarily to the **caudate–putamen** (dorsal striatum in rodents)
- It is critical in the production of **movement** as part of the basal ganglia motor loop

2. **Mesolimbic pathway**

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

3. **Mesocortical pathway**

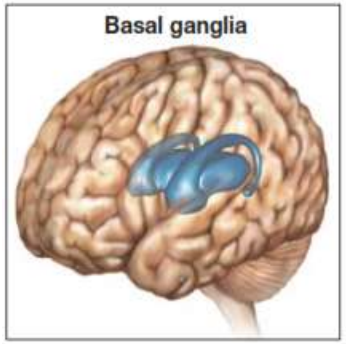
- Originates in the VTA
- projects to the medial prefrontal, cingulate, **orbitofrontal** and perirhinal cortex

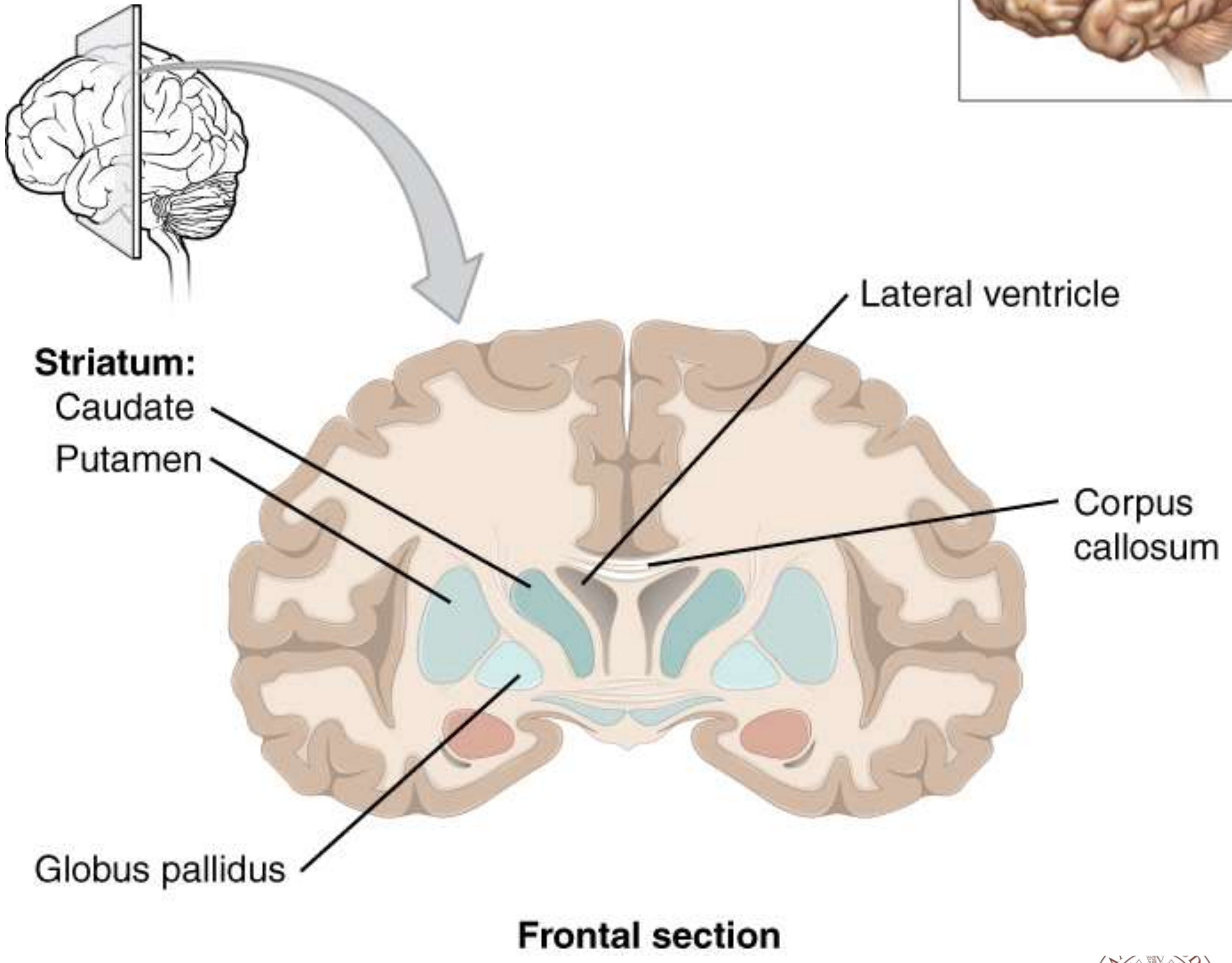


https://www.brainfacts.org/3d-brain#intro=false&focus=Brain

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



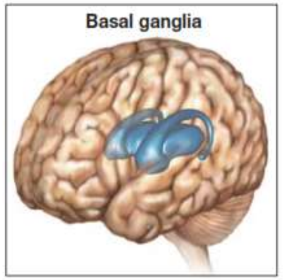


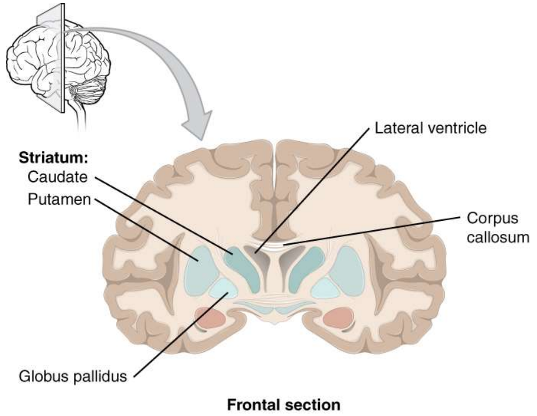
ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

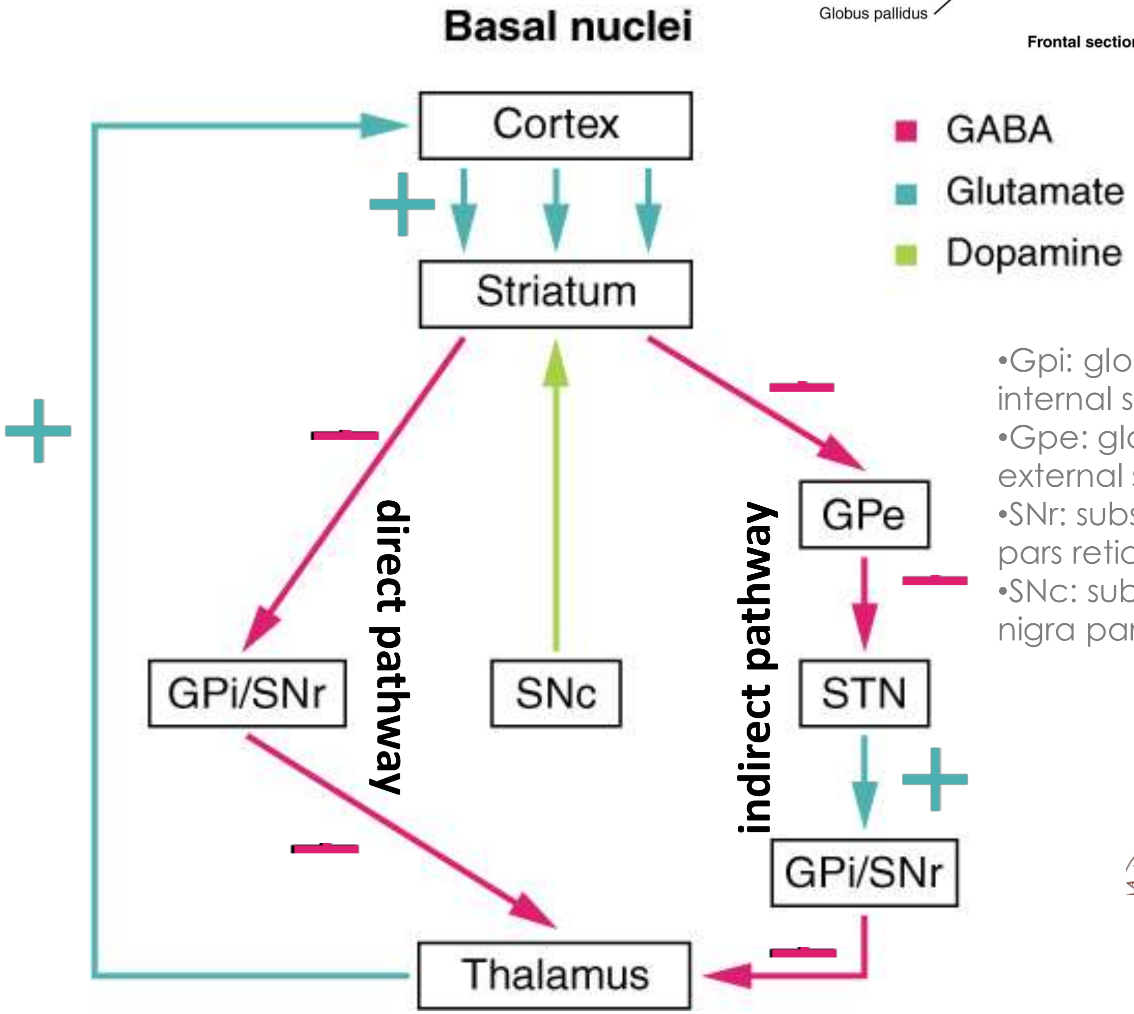
Basal Ganglia have crucial role in motor control & reinforcement learning

- 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**





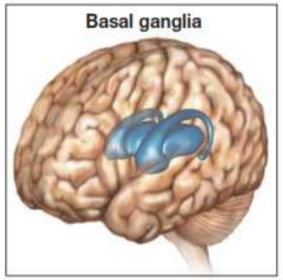


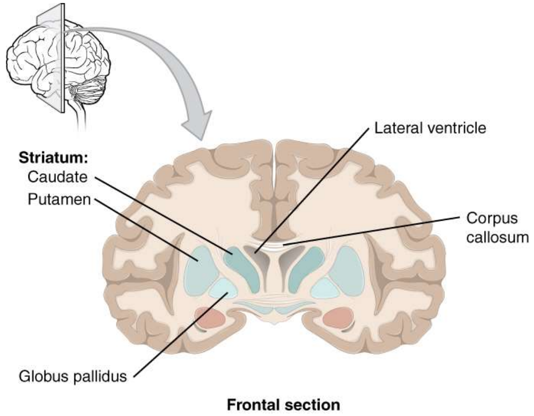
•Gpi: globus pallidus internal segment
•Gpe: globus pallidus external segment
•SNr: substantia nigra pars reticulata
•SNc: substantia nigra pars compacta

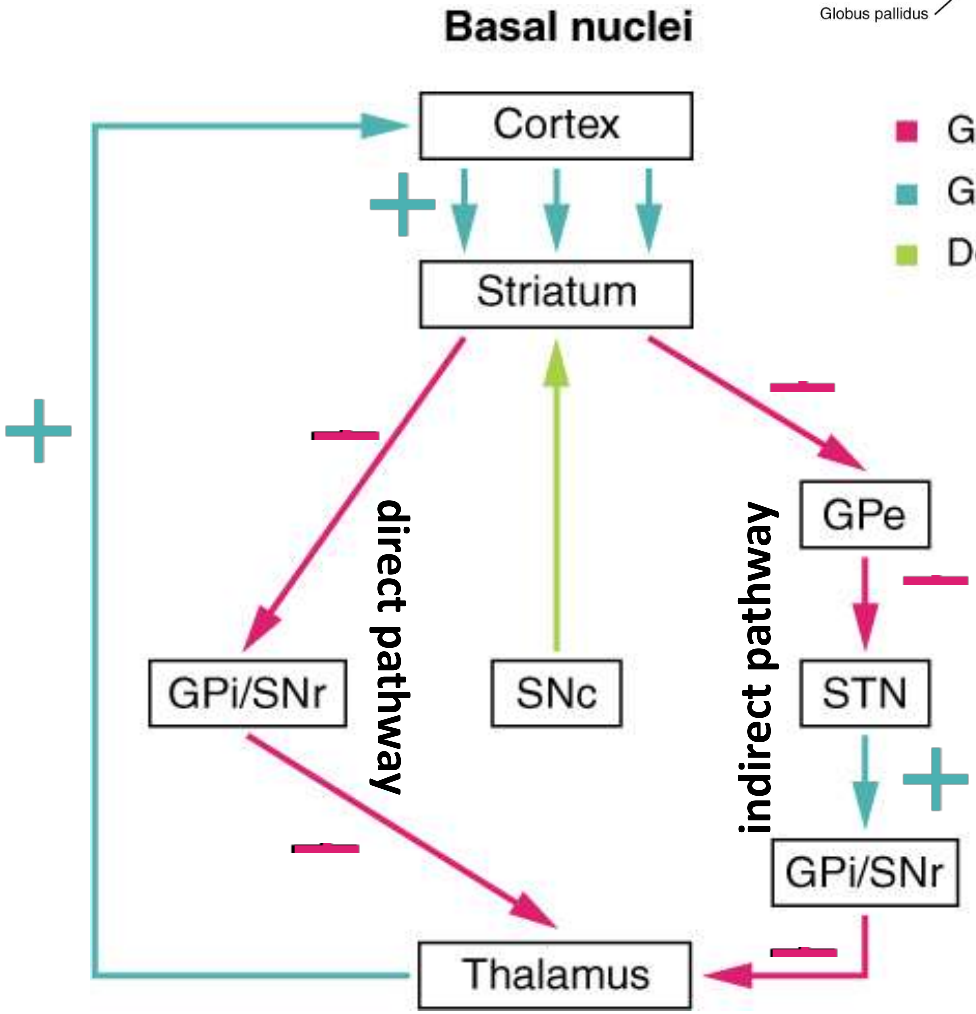
Basal Ganglia have crucial role in motor control & reinforcement learning

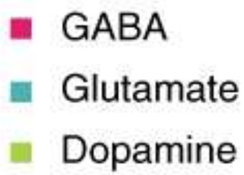
The switch between the two pathways is the **substantia nigra pars compacta:**

- 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**







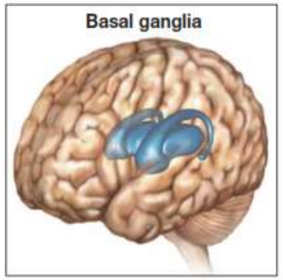


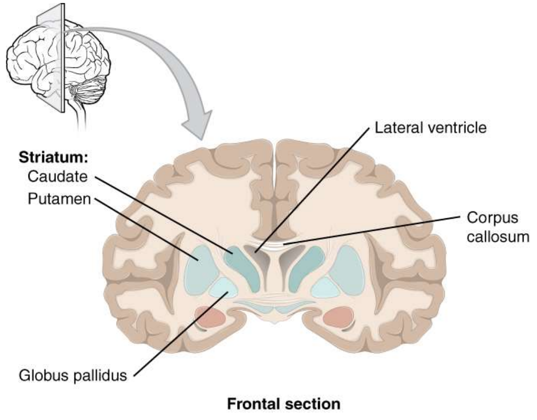
- Gpi: globus pallidus internal segment
- Gpe: globus pallidus external segment
- SNr: substantia nigra pars reticulata
- SNc: substantia nigra pars compacta

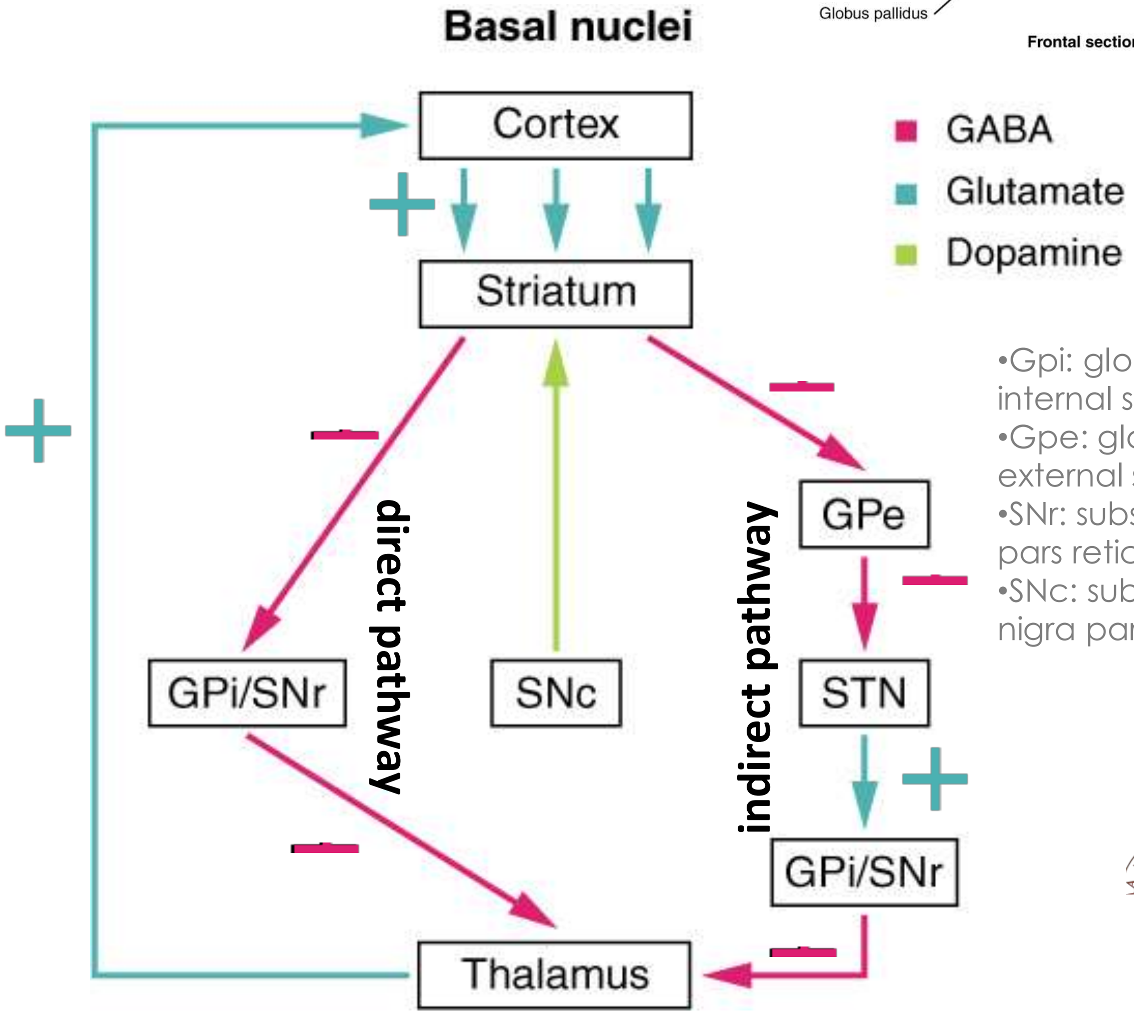
Basal Ganglia have crucial role in motor control & reinforcement learning

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
  - 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

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

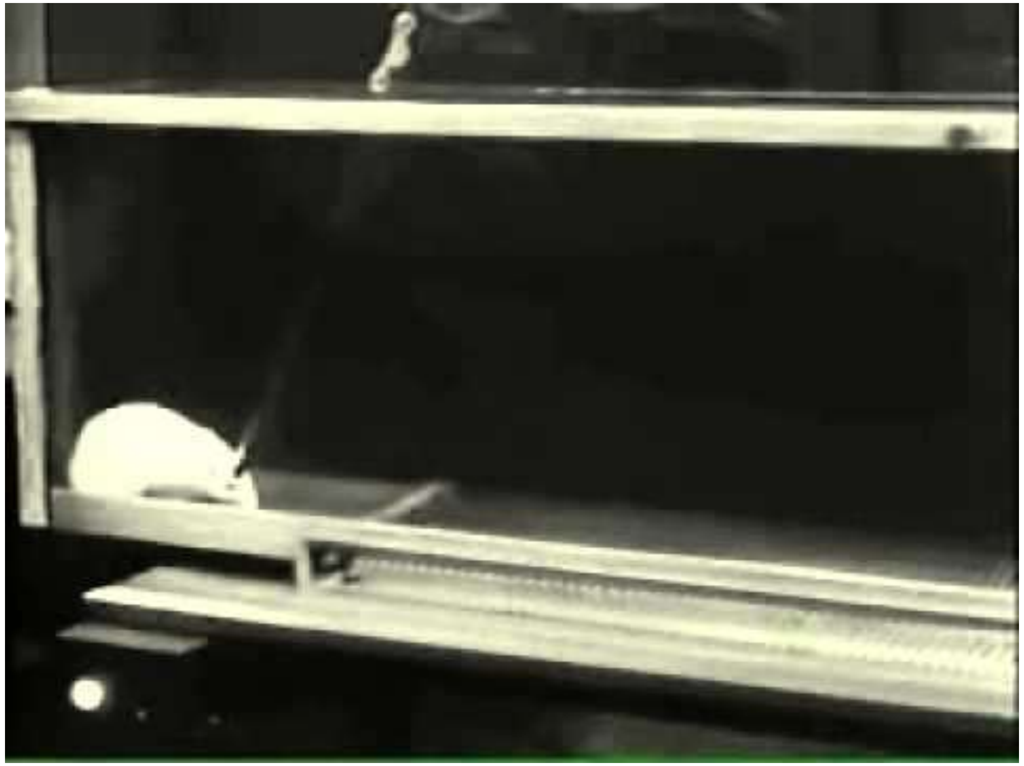
ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

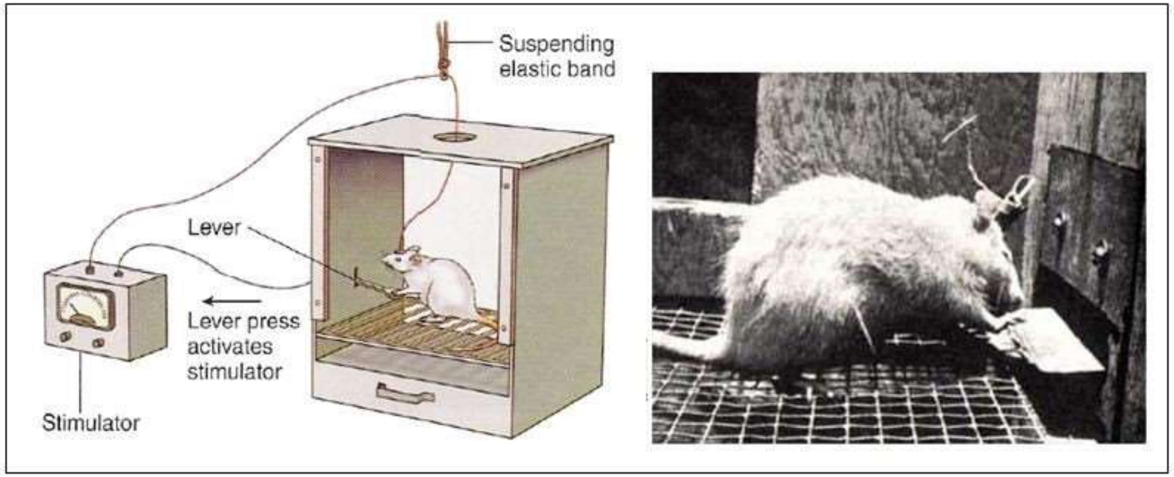
The early view: Dopamine neurons broadcast a reward signal

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.

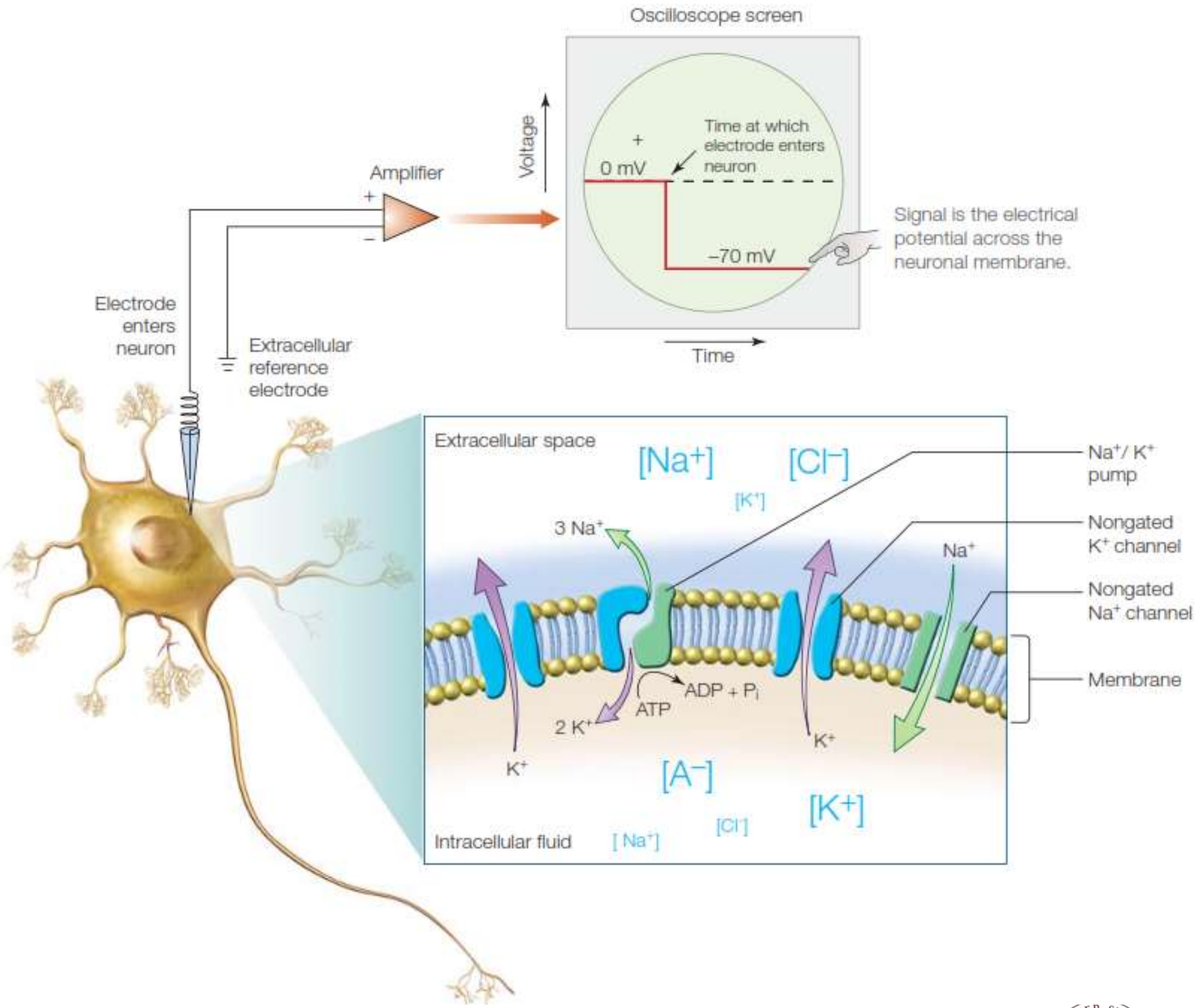




https://youtu.be/uofQPLuLV9A

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.

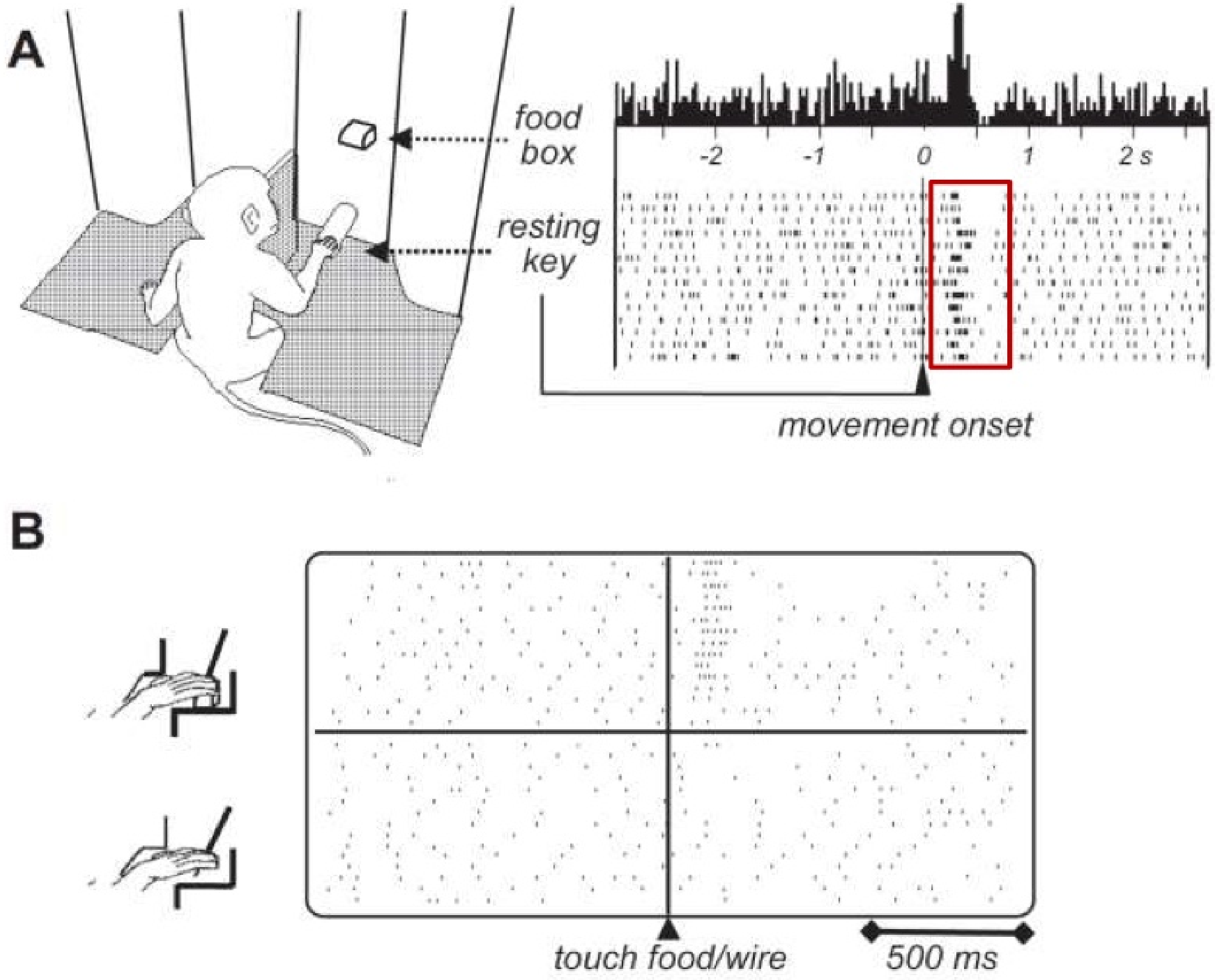


Dopaminergic neurons exhibit a strong phasic response to an unexpected reward...

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



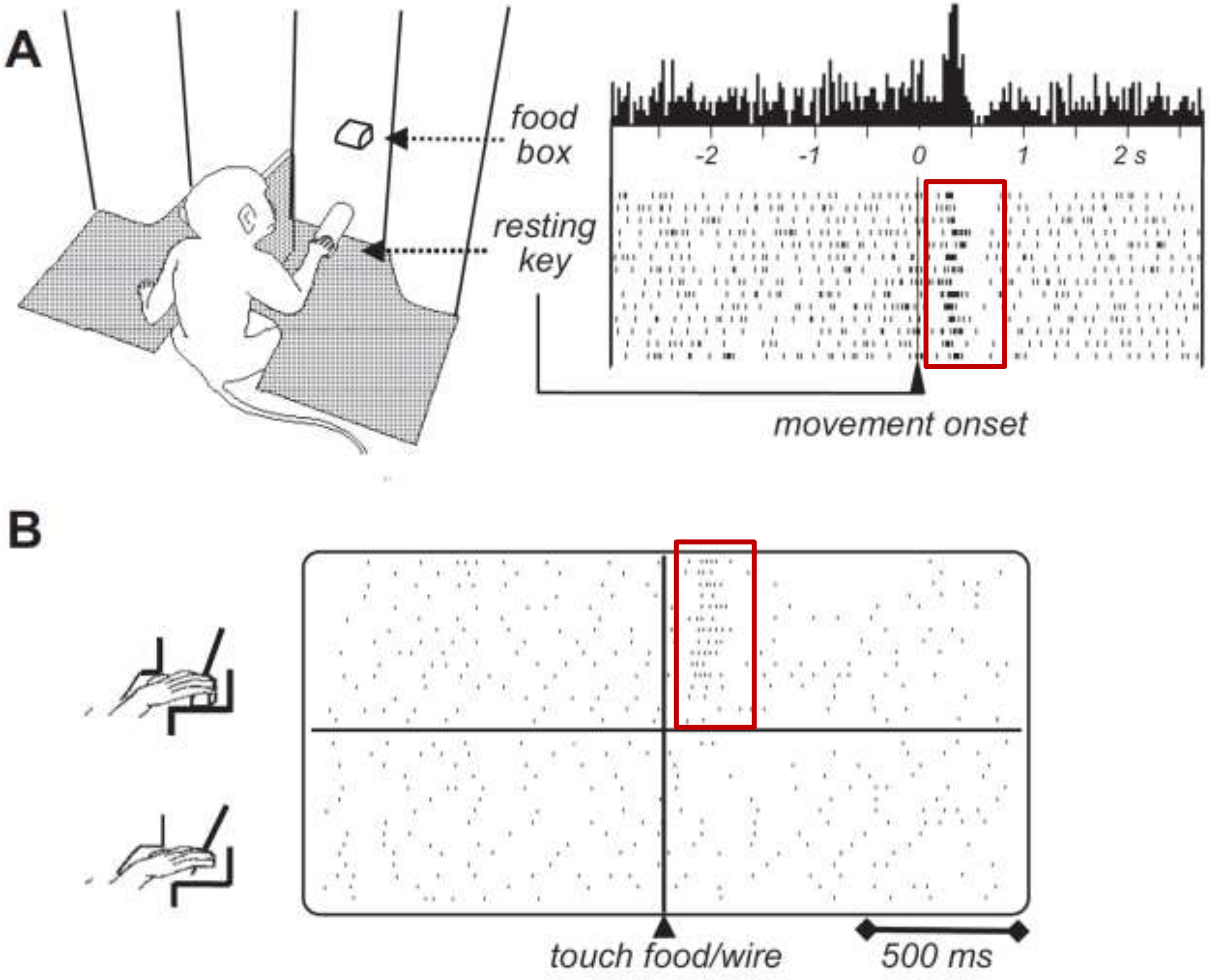
ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

...discriminate between reward & non-reward

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

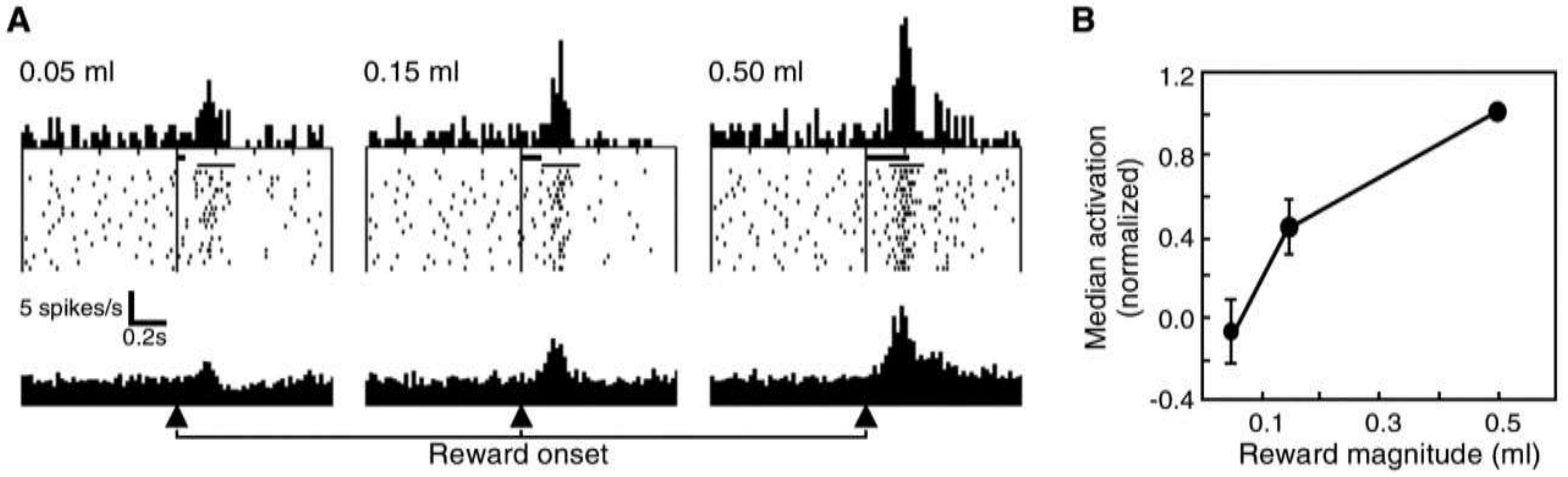
B) Differential response of dopamine neuron to touch of a wire holding a piece of apple (top), or touch of a wire holding an inedible objects (bottom).

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)

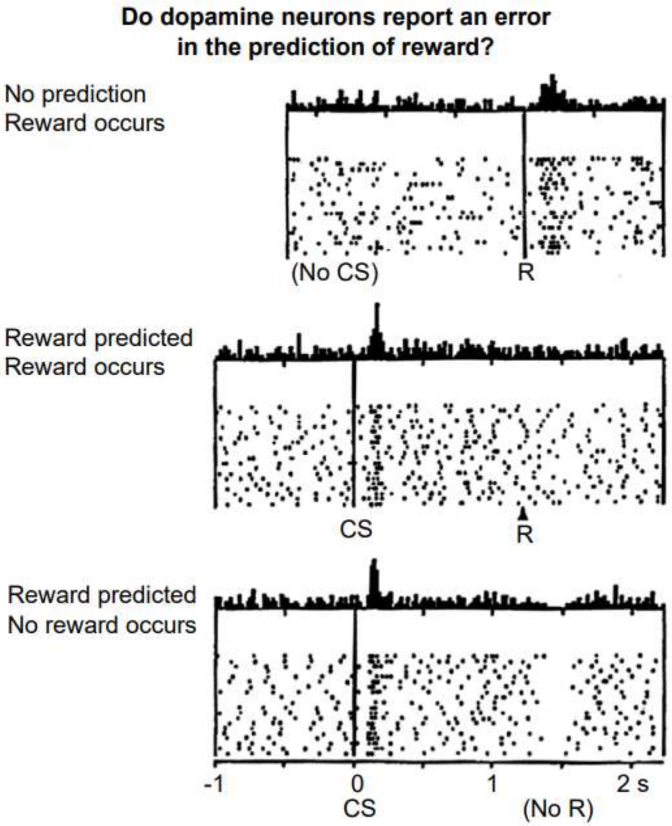


Fig. 1. Changes in dopamine neurons' output code for an error in the prediction of appetitive events. (**Top**) Before learning, a drop of appetitive fruit juice occurs in the absence of prediction—hence a positive error in the prediction of reward. The dopamine neuron is activated by this unpredicted 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*

1594 SCIENCE • VOL. 275 • 14 MARCH 1997 • http://www.scie

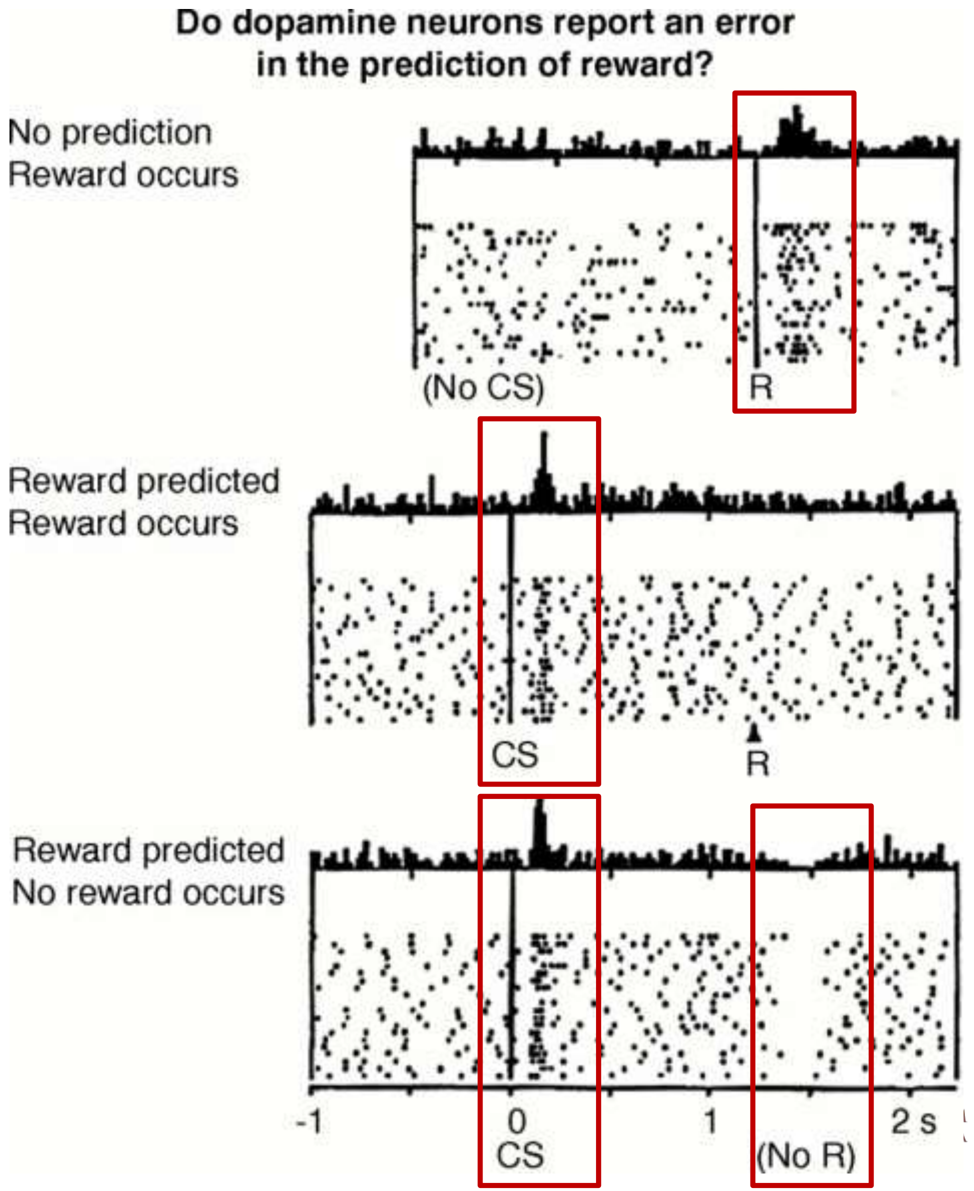
...transfers back to a cue which predicts reward occurrence (i.e. CS)

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

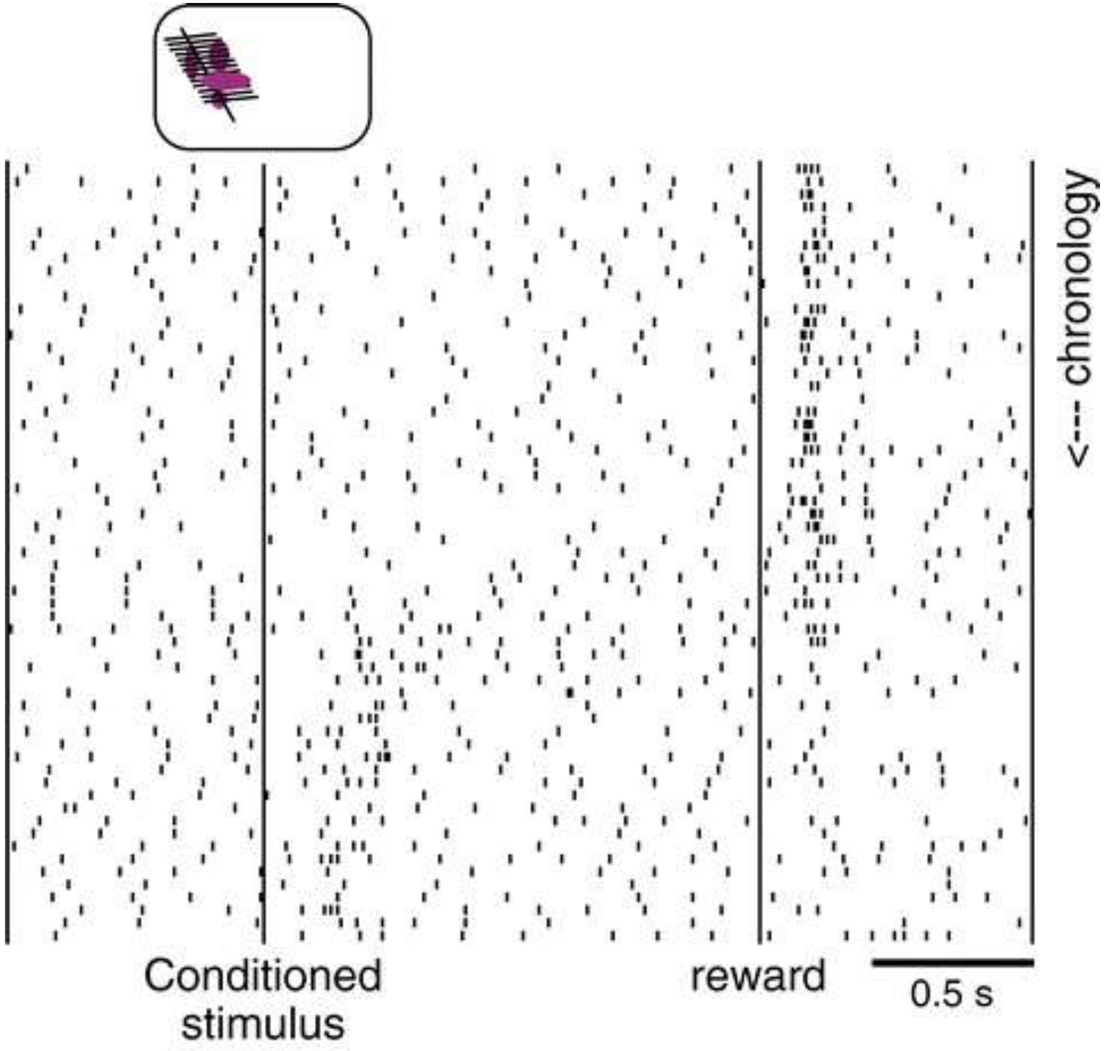


...transfers back to a cue which predicts reward occurrence (i.e. CS)

A dopamine neuron that responds initially to a liquid or food reward acquires a response to the CS after some tens of paired CS-reward trials.

Each line of dots represents a trial, each dot represents the time of the discharge of the dopamine neuron, the vertical lines indicate the time of the stimulus and juice reward, and the picture above the raster shows the visual conditioned stimulus presented to the monkey on a computer screen.

Chronology of trials is from top to bottom. **The top trial shows the activity of the neuron while the animal saw the stimulus for the first time in its life, whereas it had previous experience with the liquid reward.**



Schultz W. 2006.
Annu. Rev. Psychol. 57:87–115

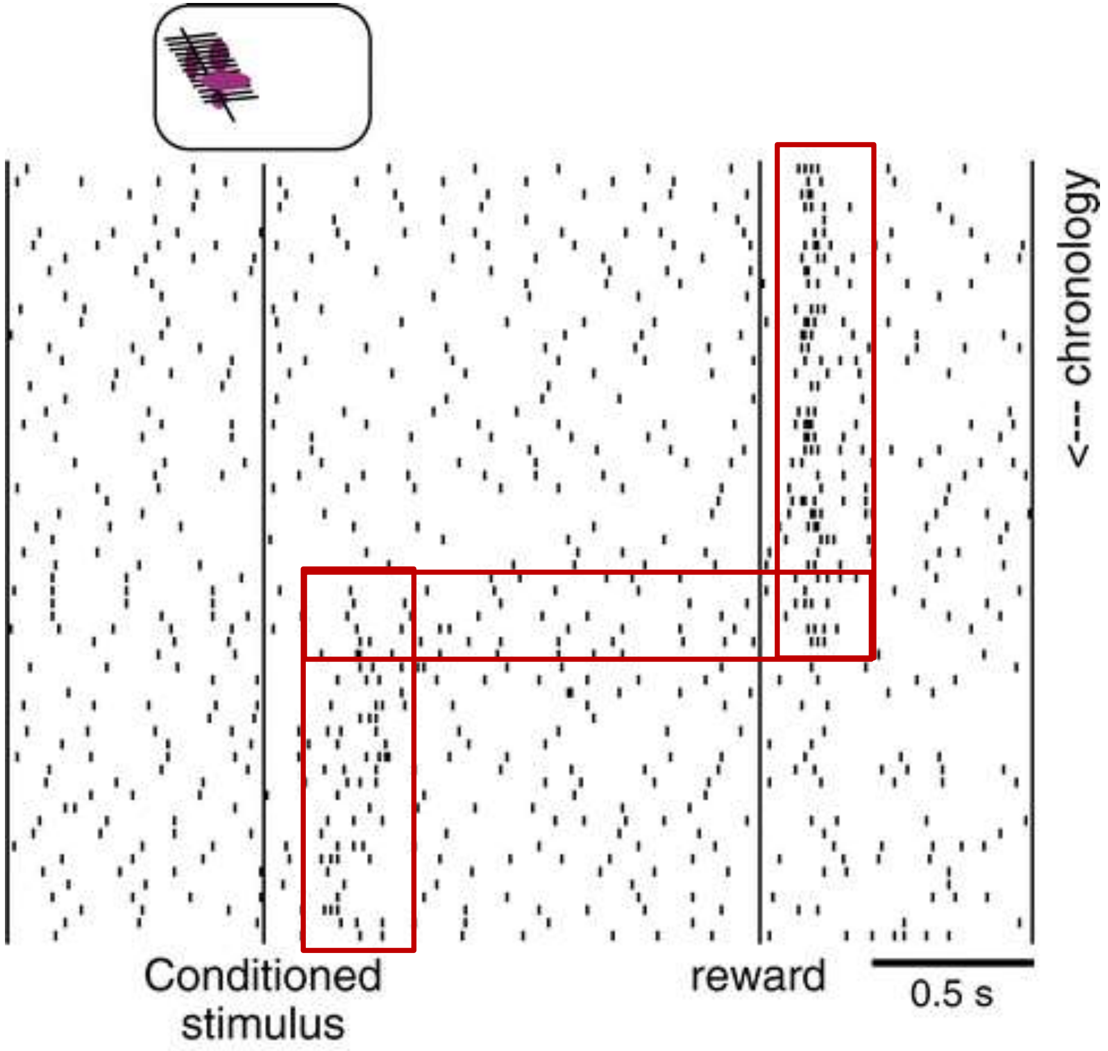
Data from Waelti (2000).

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

A dopamine neuron that responds initially to a liquid or food reward acquires a response to the CS after some tens of paired CS-reward trials.

Each line of dots represents a trial, each dot represents the time of the discharge of the dopamine neuron, the vertical lines indicate the time of the stimulus and juice reward, and the picture above the raster shows the visual conditioned stimulus presented to the monkey on a computer screen.

Chronology of trials is from top to bottom. **The top trial shows the activity of the neuron while the animal saw the stimulus for the first time in its life, whereas it had previous experience with the liquid reward.**

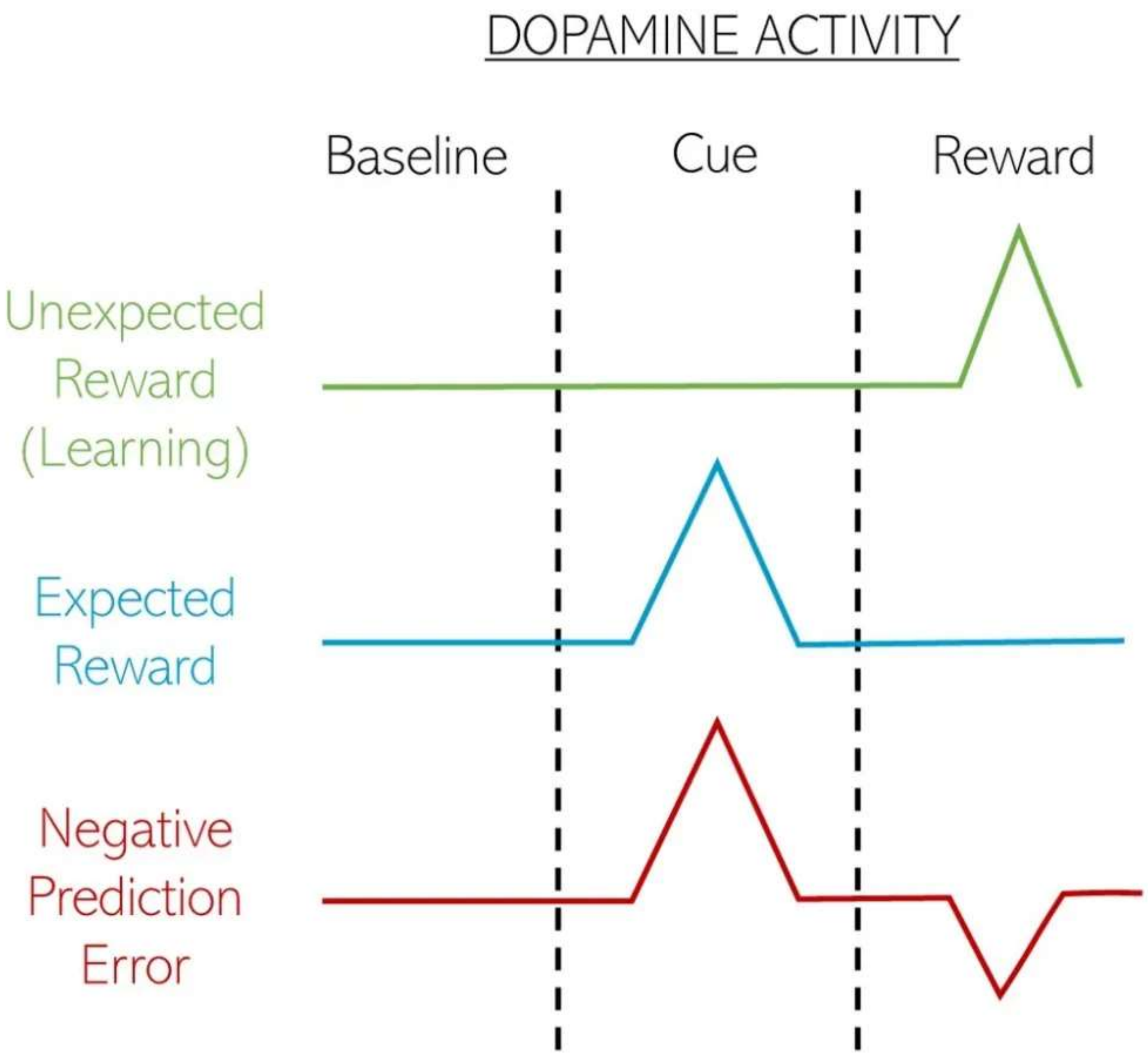


Schultz W. 2006.
Annu. Rev. Psychol. 57:87–115

Data from Waelti (2000).

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.

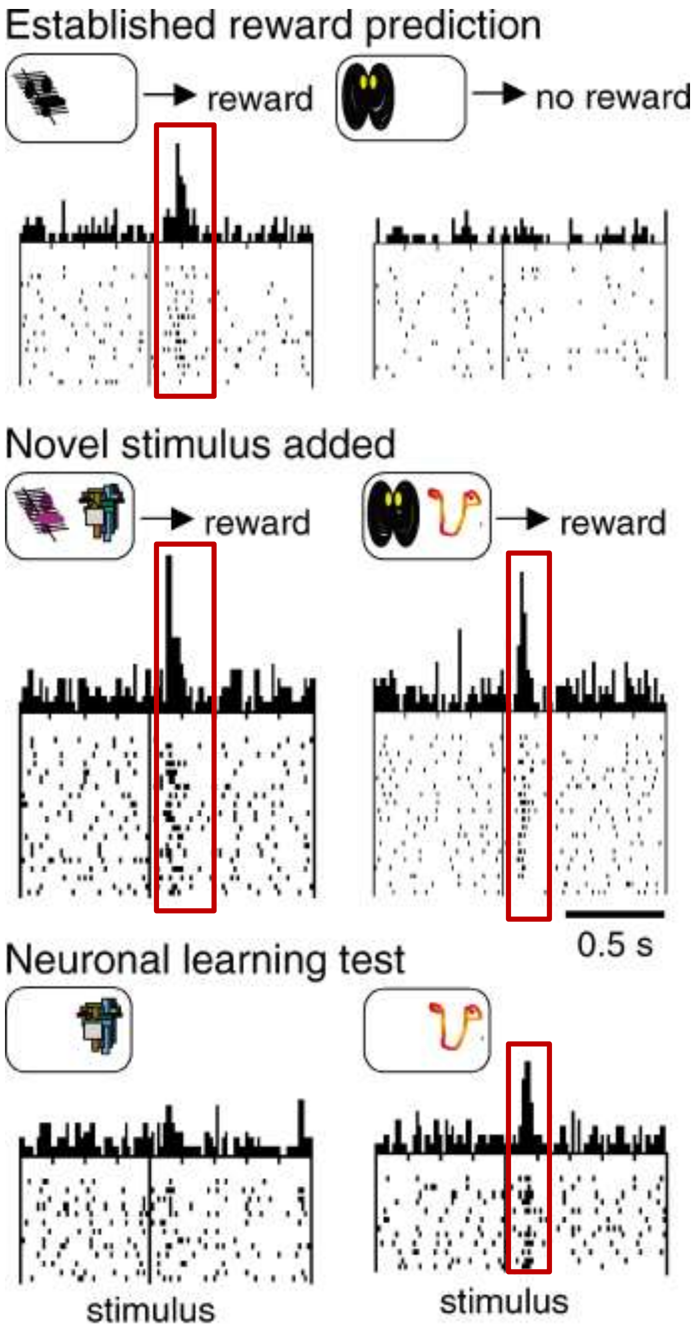


Dopaminergic neurons response conforms to blocking

Neural learning is blocked when the reward is predicted by another stimulus (left) but is intact in the same neuron when reward is unpredicted in control trials with different stimuli (right).

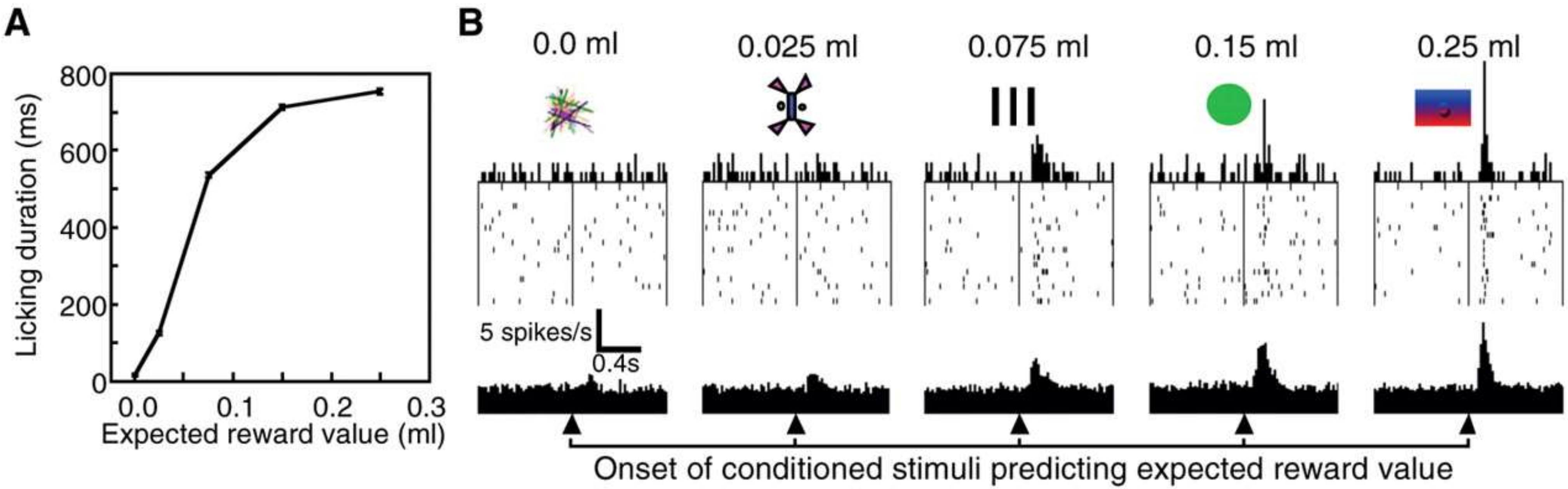
- The neuron has the capacity to respond to reward-predicting stimuli (top left) and discriminates against unrewarded stimuli (top right).
- The addition of a second stimulus results in maintenance and acquisition of response, respectively (middle).
- Testing the added stimulus reveals absence of learning when the reward is already predicted by a previously conditioned stimulus (bottom left).

Data from Waelti et al. (2001).



Schultz W. 2006.
Annu. Rev. Psychol. 57:87–115

Dopaminergic neurons firing to the CS increases with expected reward value (Tobler, Fiorillo and Schultz, Science, 2005)

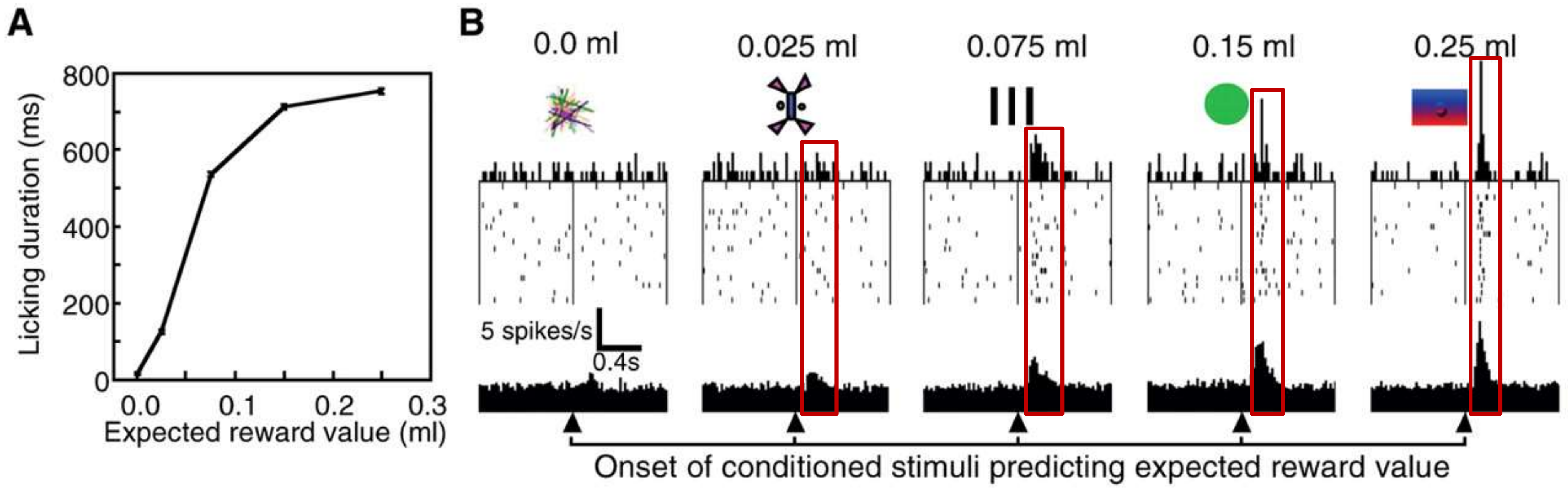


(**A**) Anticipatory licking responses during the 2-s delay between the conditioned stimuli and liquid delivery.

(**B**) Single-neuron (top) and population responses (bottom) (*n* = 57 neurons) from the experiment in (A). Visual conditioned stimuli with their expected magnitude of reward are shown above the rasters.

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

Dopaminergic neurons firing to the CS increases with expected reward value (Tobler, Fiorillo and Schultz, Science, 2005)

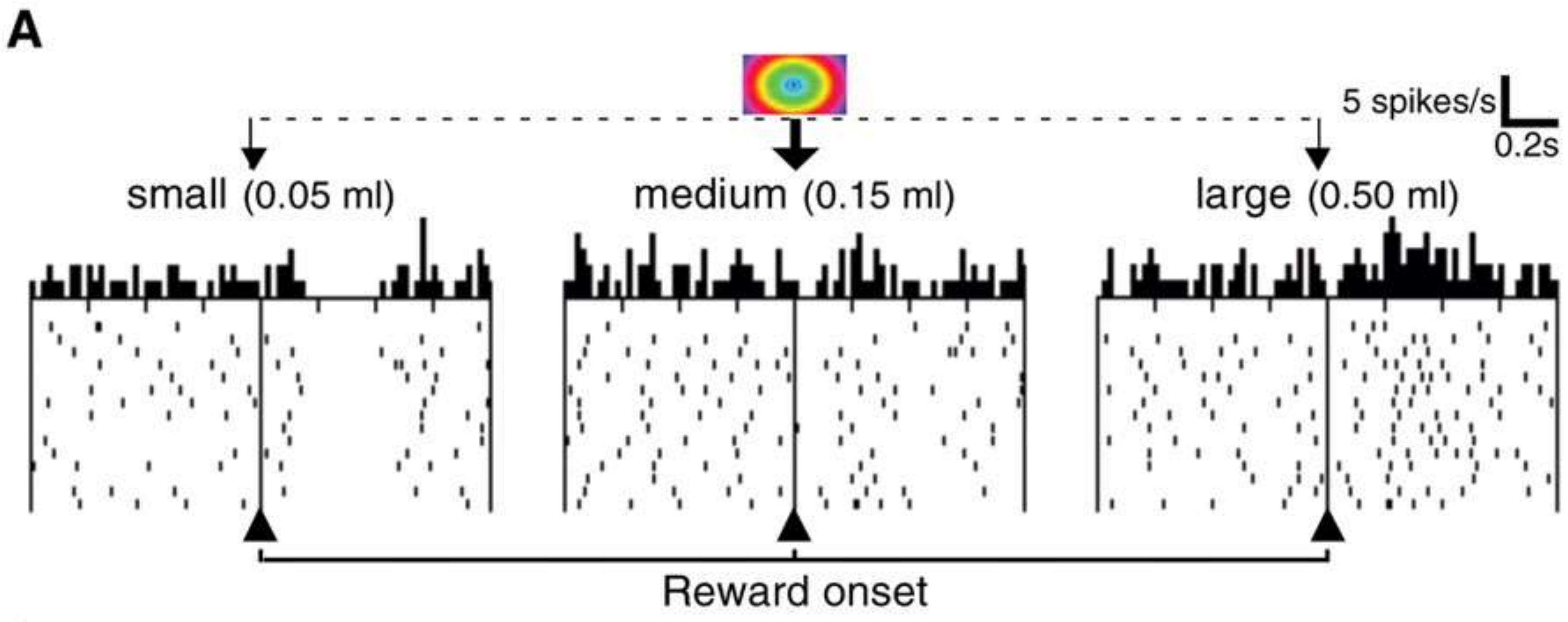


(**A**) Anticipatory licking responses during the 2-s delay between the conditioned stimuli and liquid delivery.

(**B**) Single-neuron (top) and population responses (bottom) (*n* = 57 neurons) from the experiment in (A). Visual conditioned stimuli with their expected magnitude of reward are shown above the rasters.

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

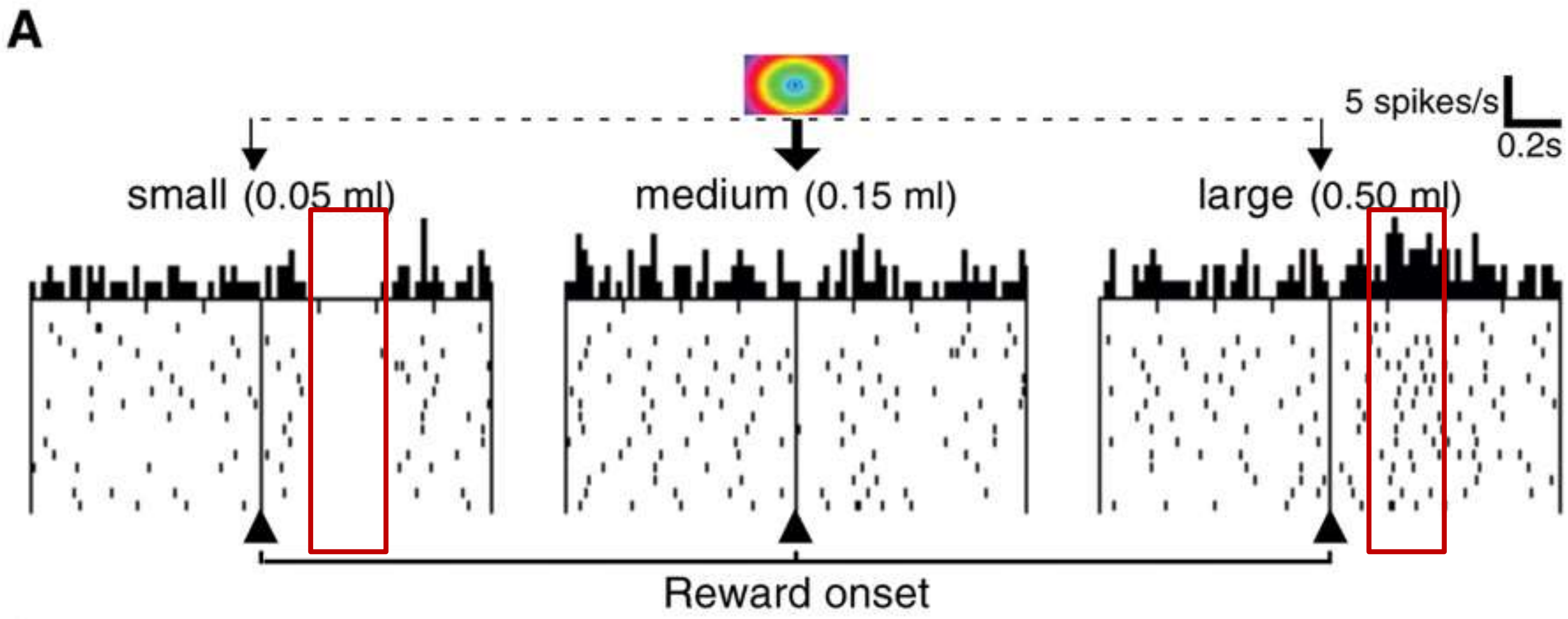
Dopaminergic neurons exhibit a signed (bidirectional) response to unexpected quantity of reward... (Tobler, Fiorillo and Schulz, Science, 2005)



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**

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

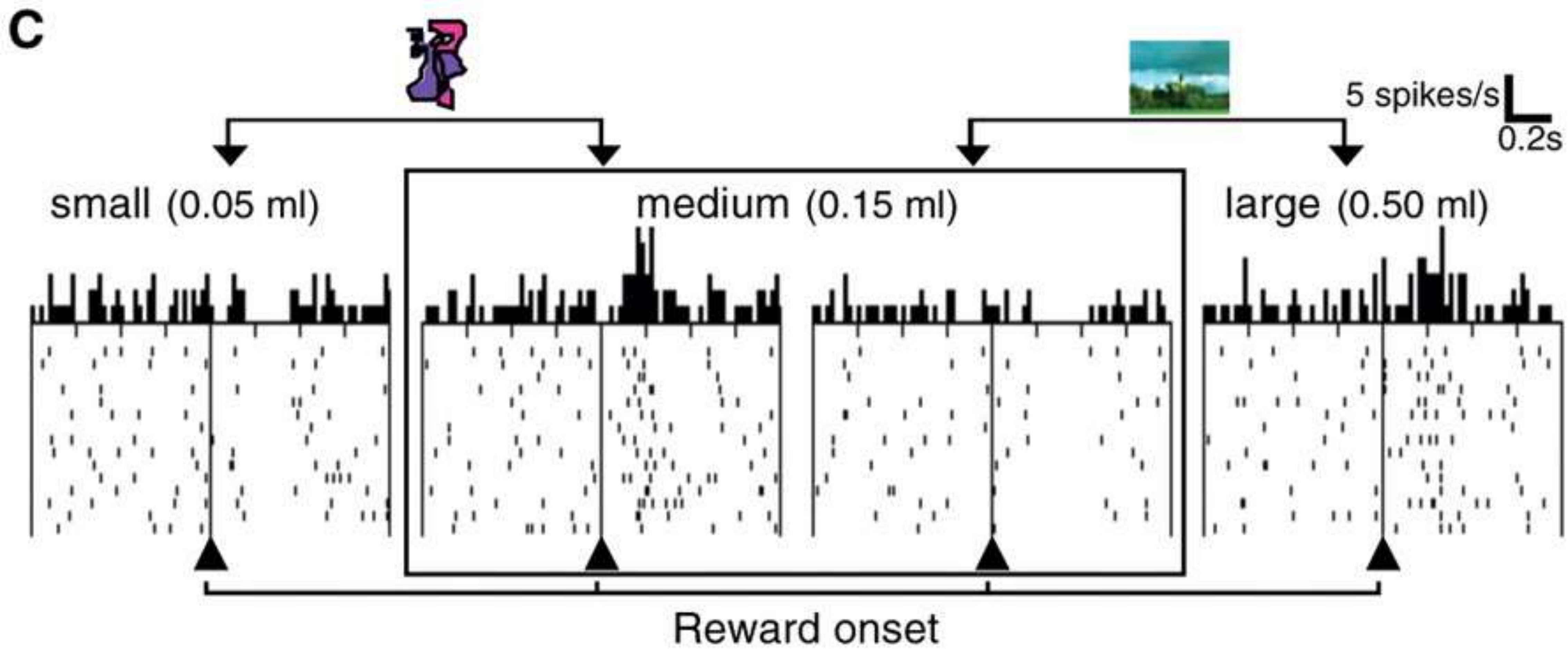
Dopaminergic neurons exhibit a signed (bidirectional) response to unexpected quantity of reward... (Tobler, Fiorillo and Schulz, Science, 2005)



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**

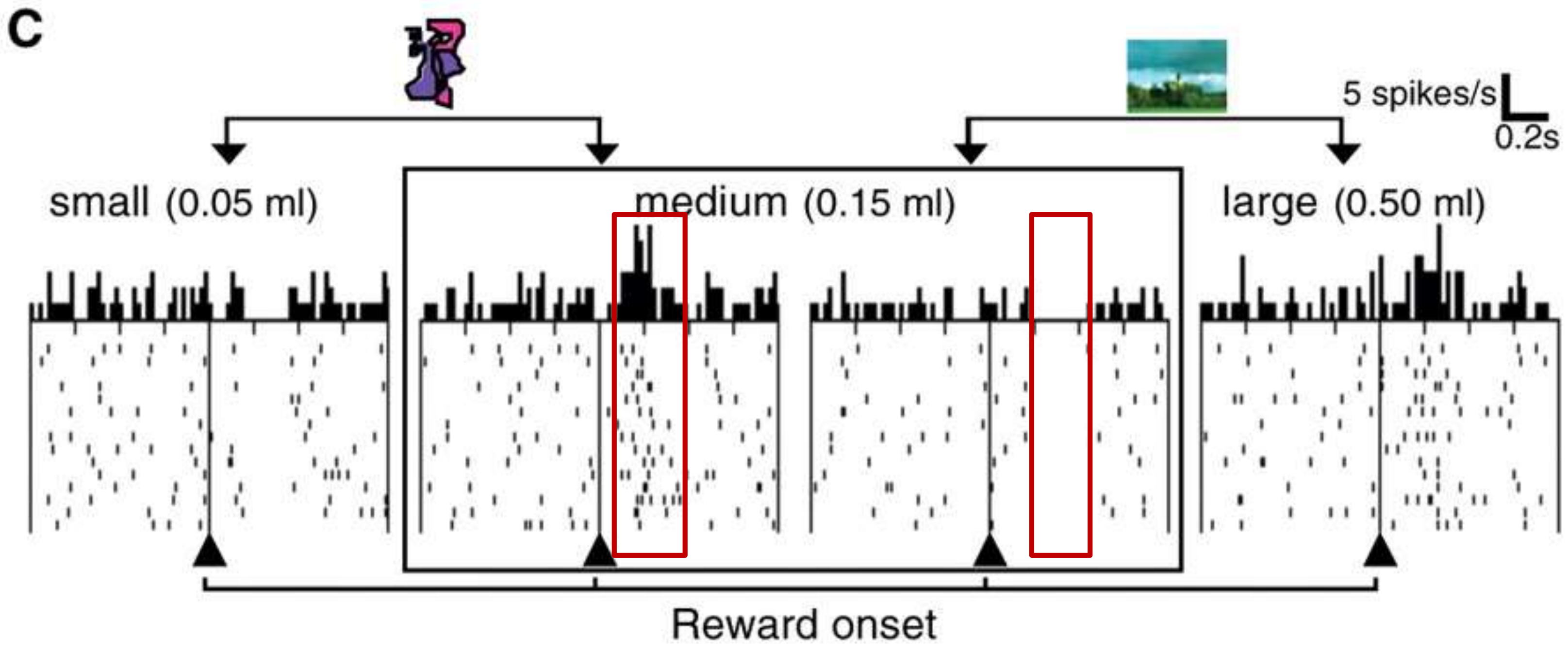
ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

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



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

...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**

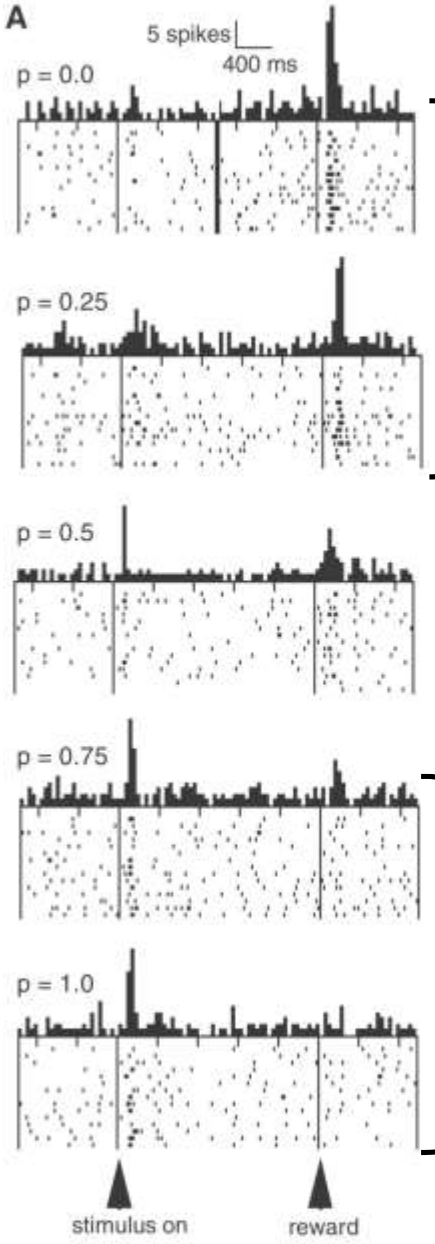
Dopaminergic neurons encode reward probability (or uncertainty)

(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



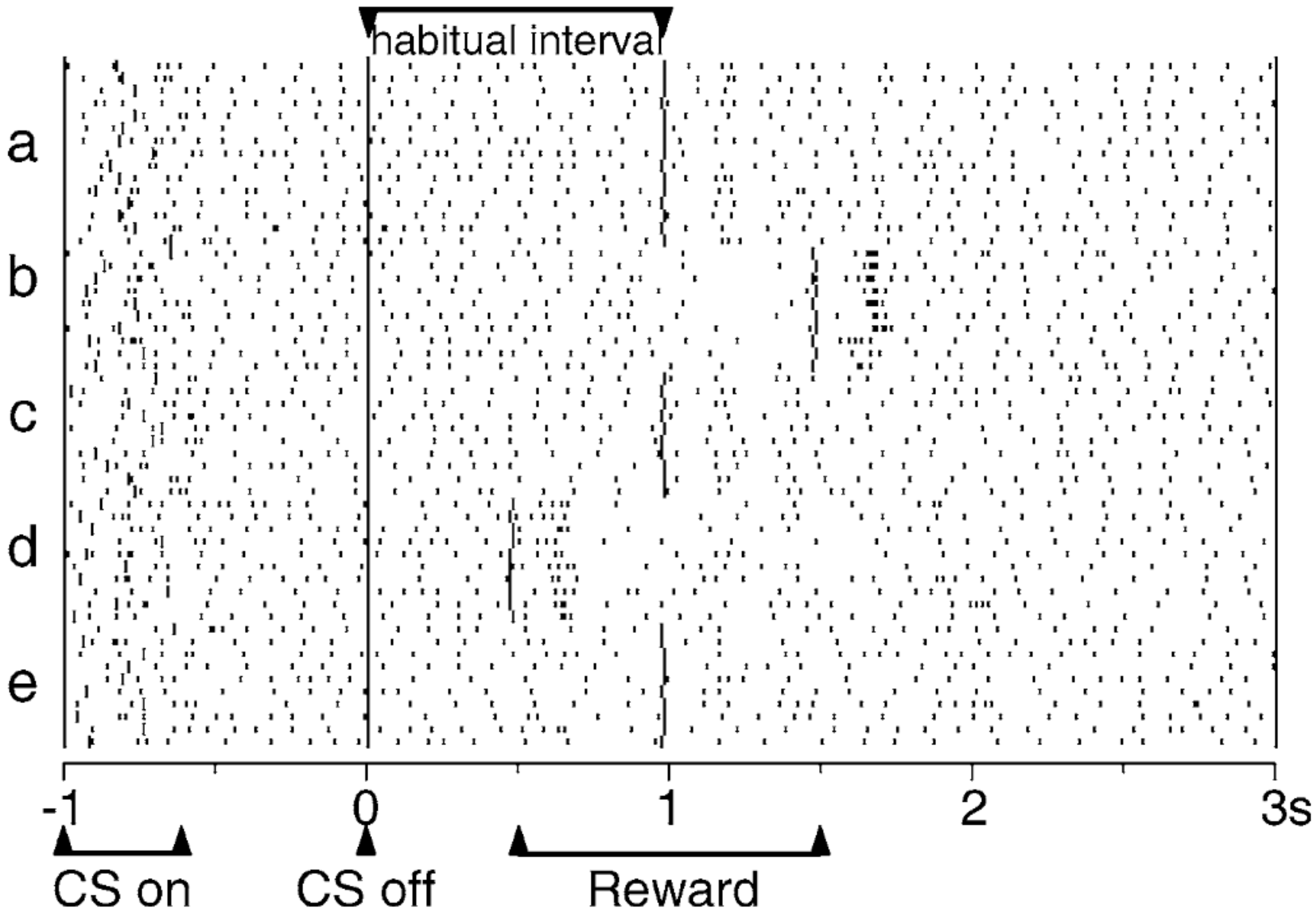
No reward after CS more likely

Max uncertainty

Reward after CS more likely

Dopaminergic neurons encode expectations of reward timing

(Hollerman and Schultz, Nature Neuroscience, 1998)



Dopaminergic neurons encode expectations of reward timing

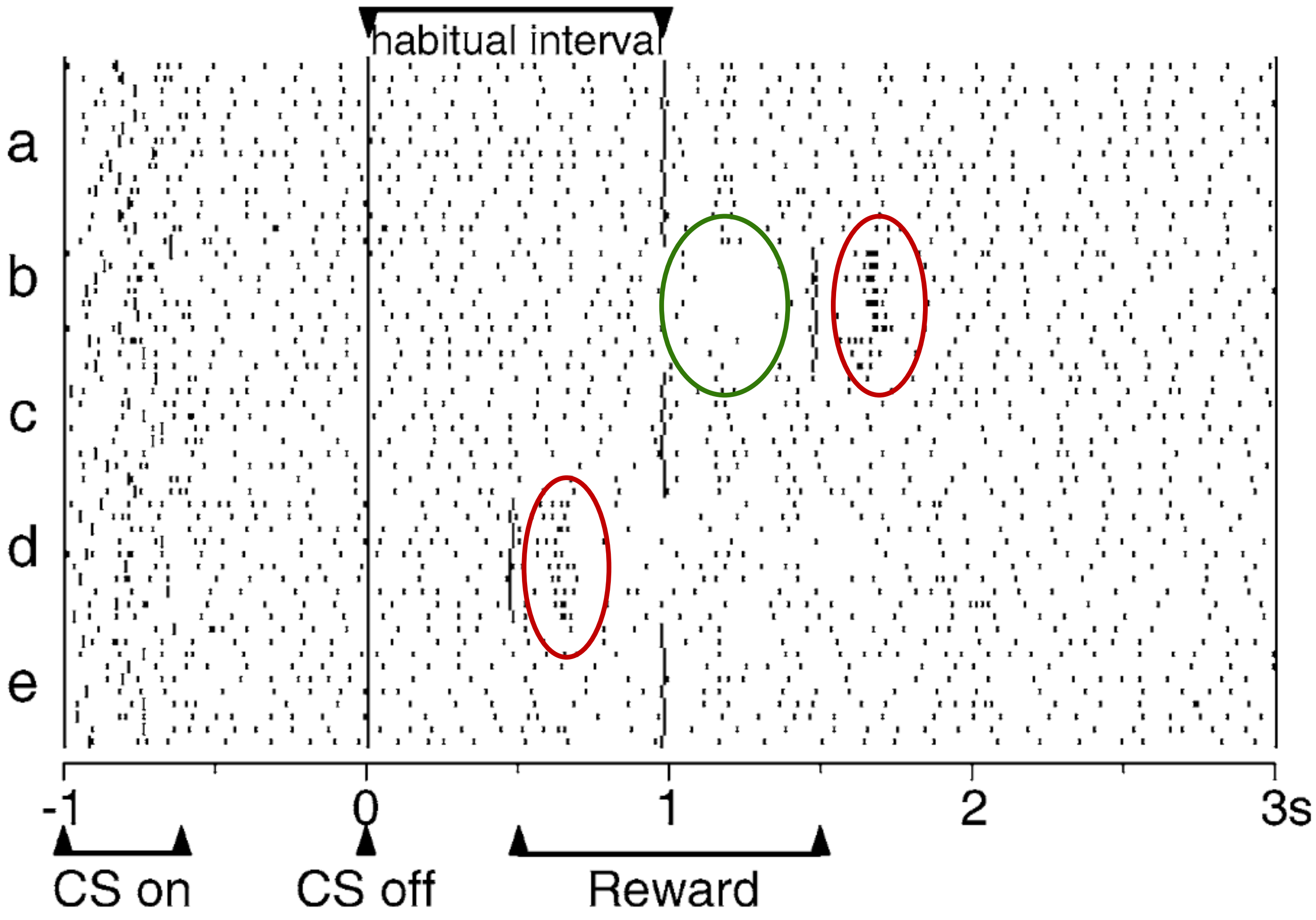
(Hollerman and Schultz, Nature Neuroscience, 1998)

Following a correct response:

- Expected time: the reward was delivered after 1.0 s
- Unexpected delay: 1.5 s
- Unexpected anticipation: 0.5 s

Firing of a dopamine neuron was

- depressed (green) when reward failed to occur at the expected time
- increased (red) when reward unexpectedly occurred at a new time, either earlier or later.

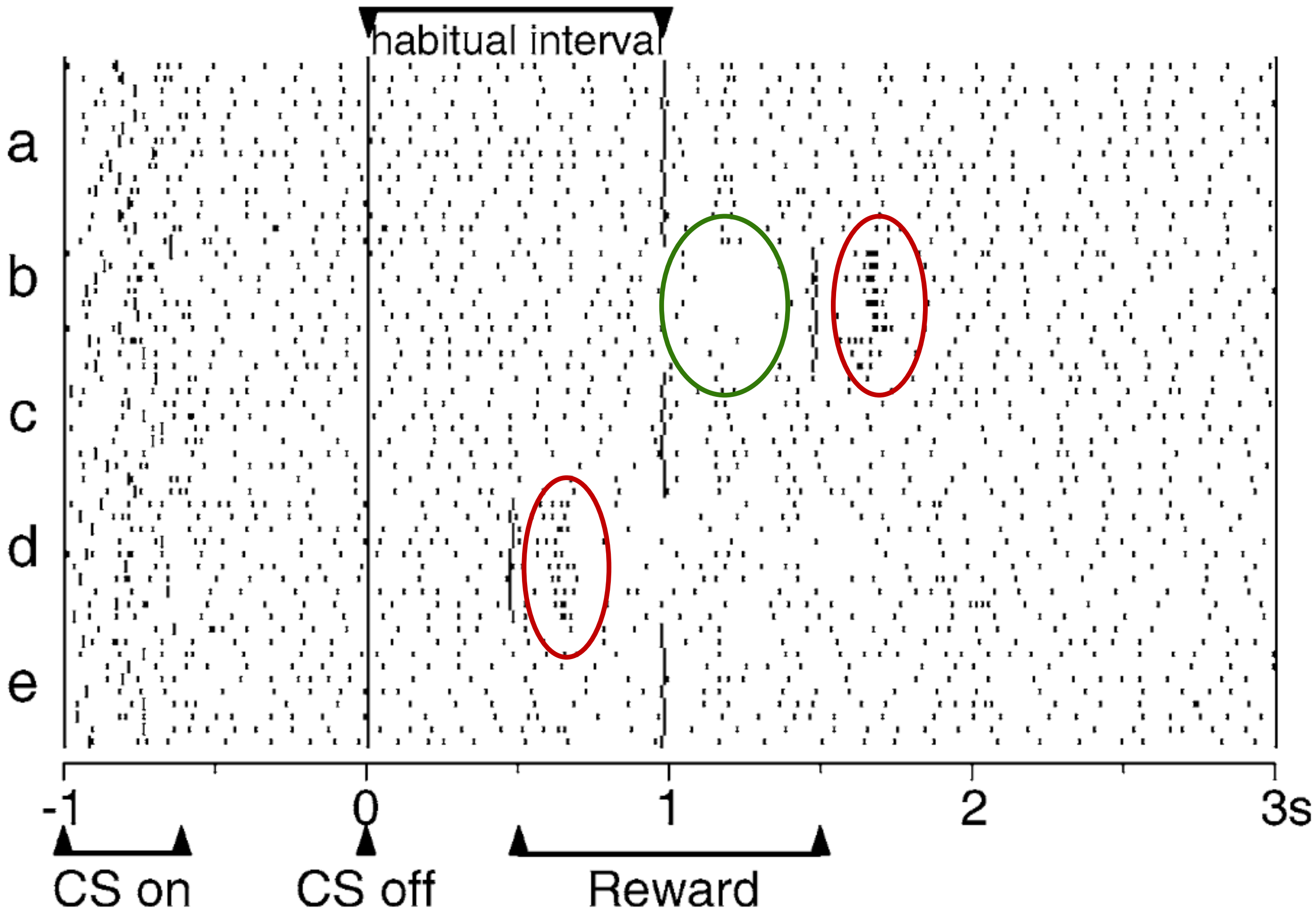


Dopaminergic neurons encode expectations of reward timing

(Hollerman and Schultz, Nature Neuroscience, 1998)

Temporal prediction error

- Unexpectedness, is not limited to event occurrence (e.g., a reward is delivered or omitted unexpectedly), but also includes the **time of reward**
- Rewards elicit transient activations when they are delivered **earlier or later than predicted**, **even though it is certain that the reward will occur**.

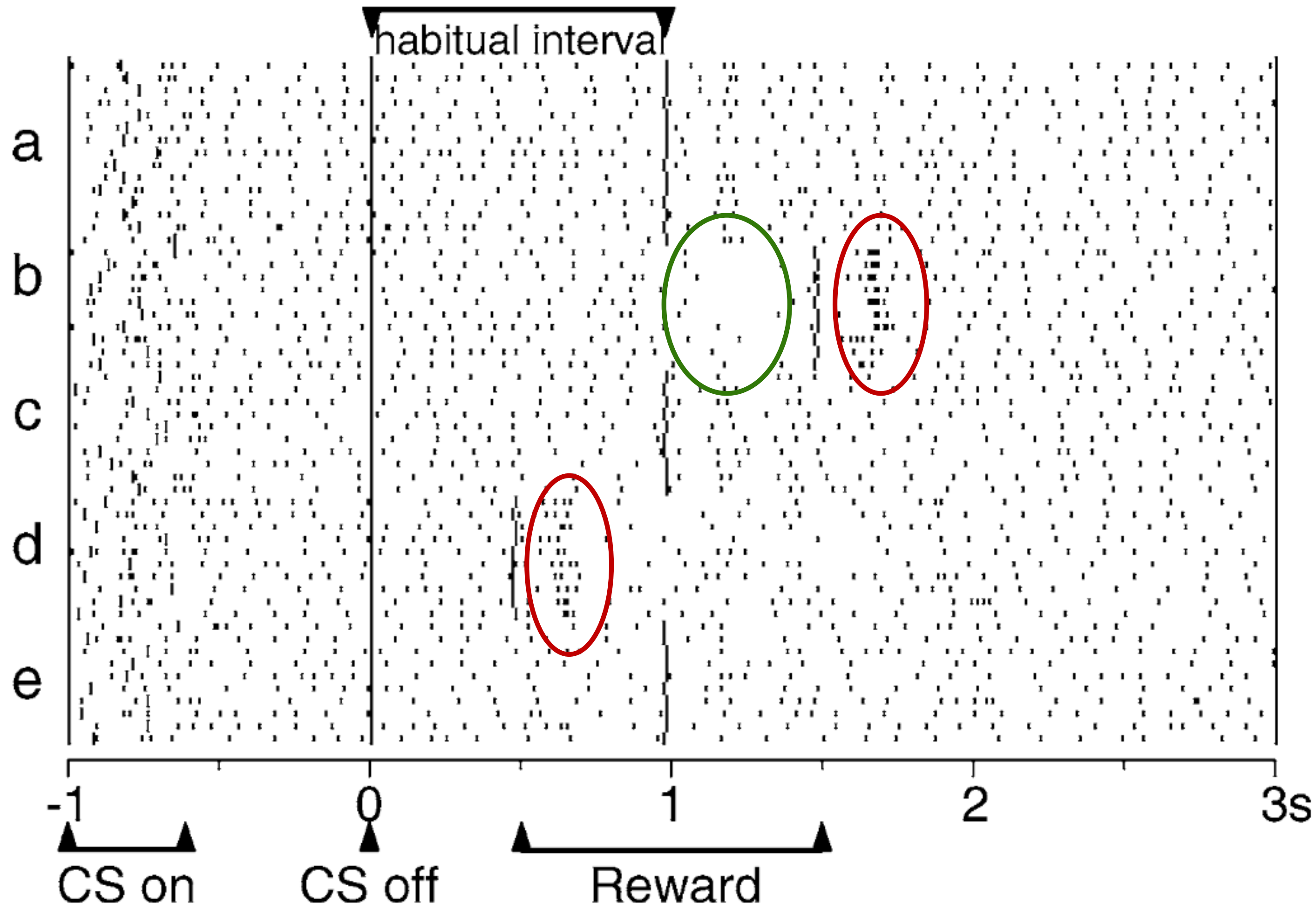


Dopaminergic neurons encode expectations of reward timing

(Hollerman and Schultz, Nature Neuroscience, 1998)

Temporal prediction error

- Firing is depressed exactly at the time of the usual occurrence of reward when a predicted reward is omitted.
- The **depression occurs even in the absence of any stimuli** at the time of the omitted reward
- The depression does not constitute a neuronal response to a stimulus but reflects an **expectation process based on an internal clock tracking the precise time of predicted reward**.



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 signed manner
        - 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


Something called the expected value
of the gamble.

wooclap

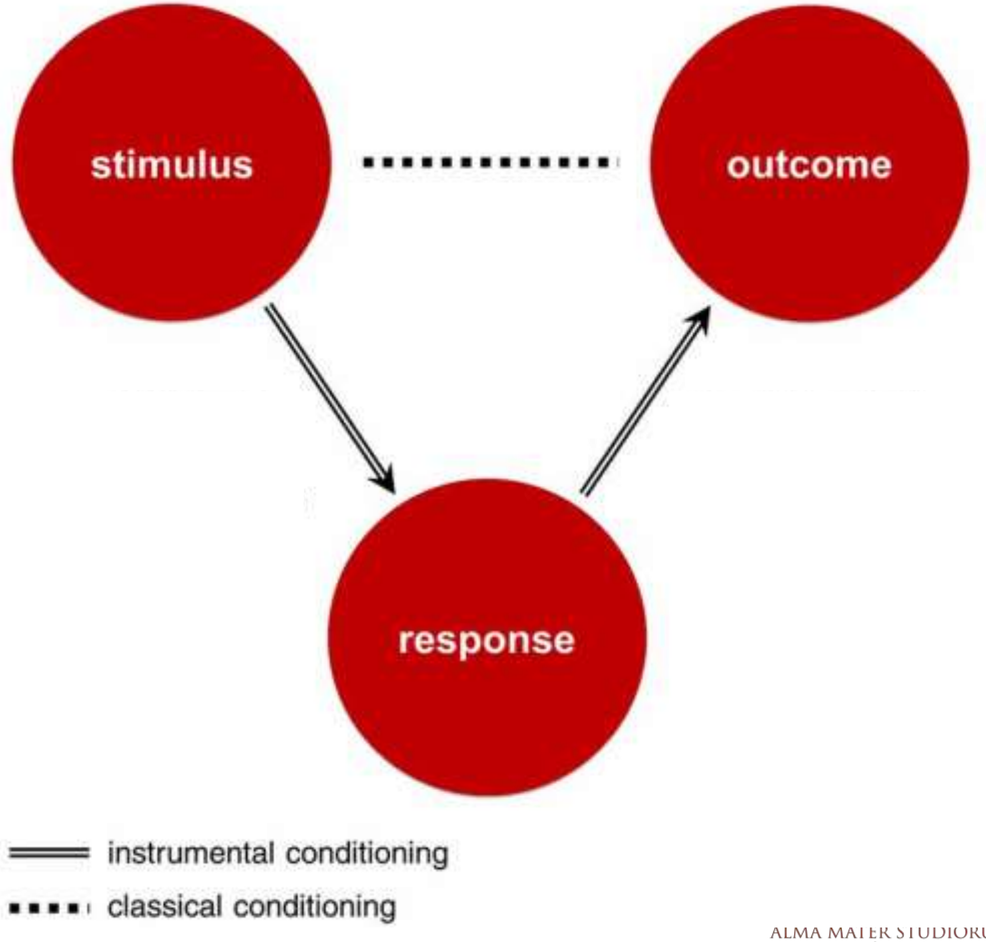
ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

From predictions to control and choice

Predictions are for control

So far we have talked about the role of dopamine only when learning to predict rewards (Pavlovian learning)

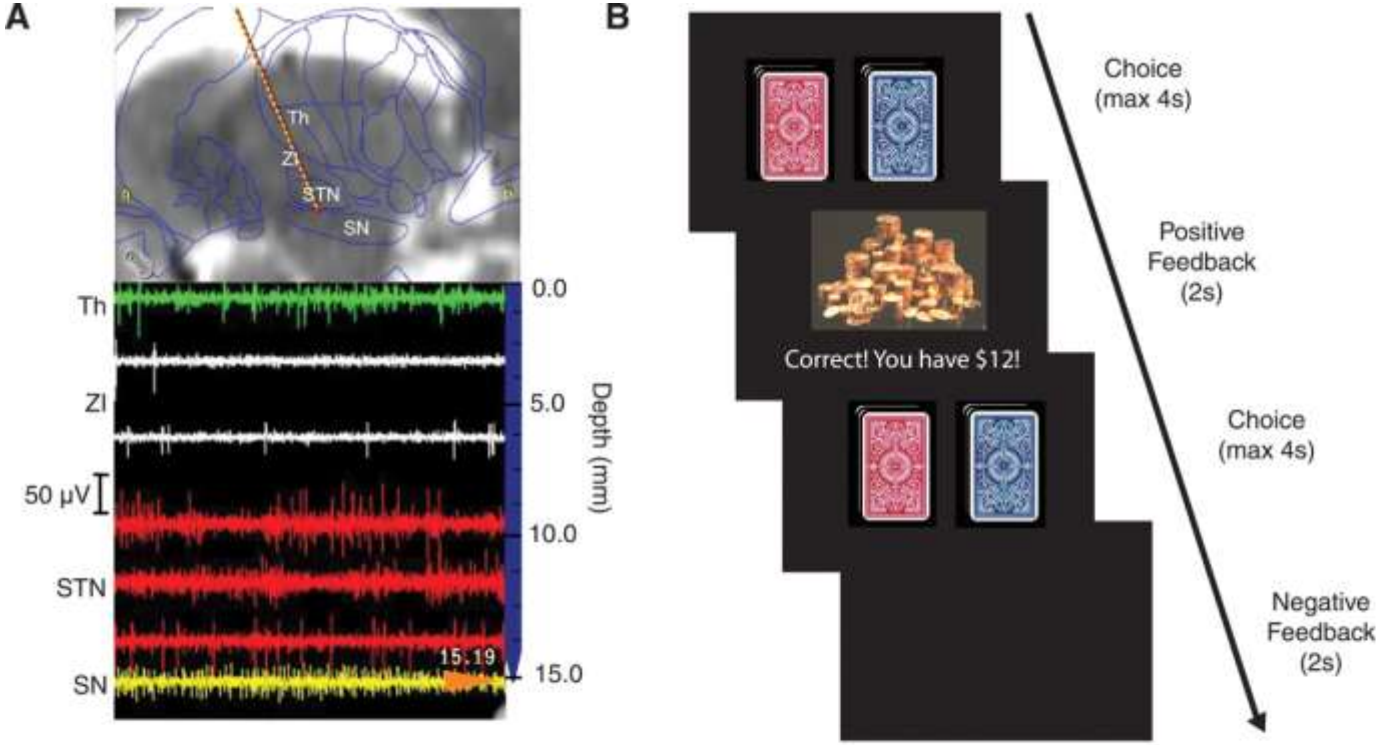
- But is there evidence that learned predictions actually drive choices (instrumental learning) in the way we have described?
- And if dopamine carries prediction errors that drive learning about reward predictions, is it causally involved in choice?



Electrophysiological correlational evidence

Microelectrode recordings during deep brain stimulation surgery to study neuronal activity in the **human substantia nigra (SN)** while patients with Parkinson's disease engaged in a **probabilistic instrumental learning task** motivated by virtual financial rewards.

Participants are presented with two decks of cards on a computer screen. They are instructed to repeatedly draw cards from either deck to determine which deck yields the higher reward probability.



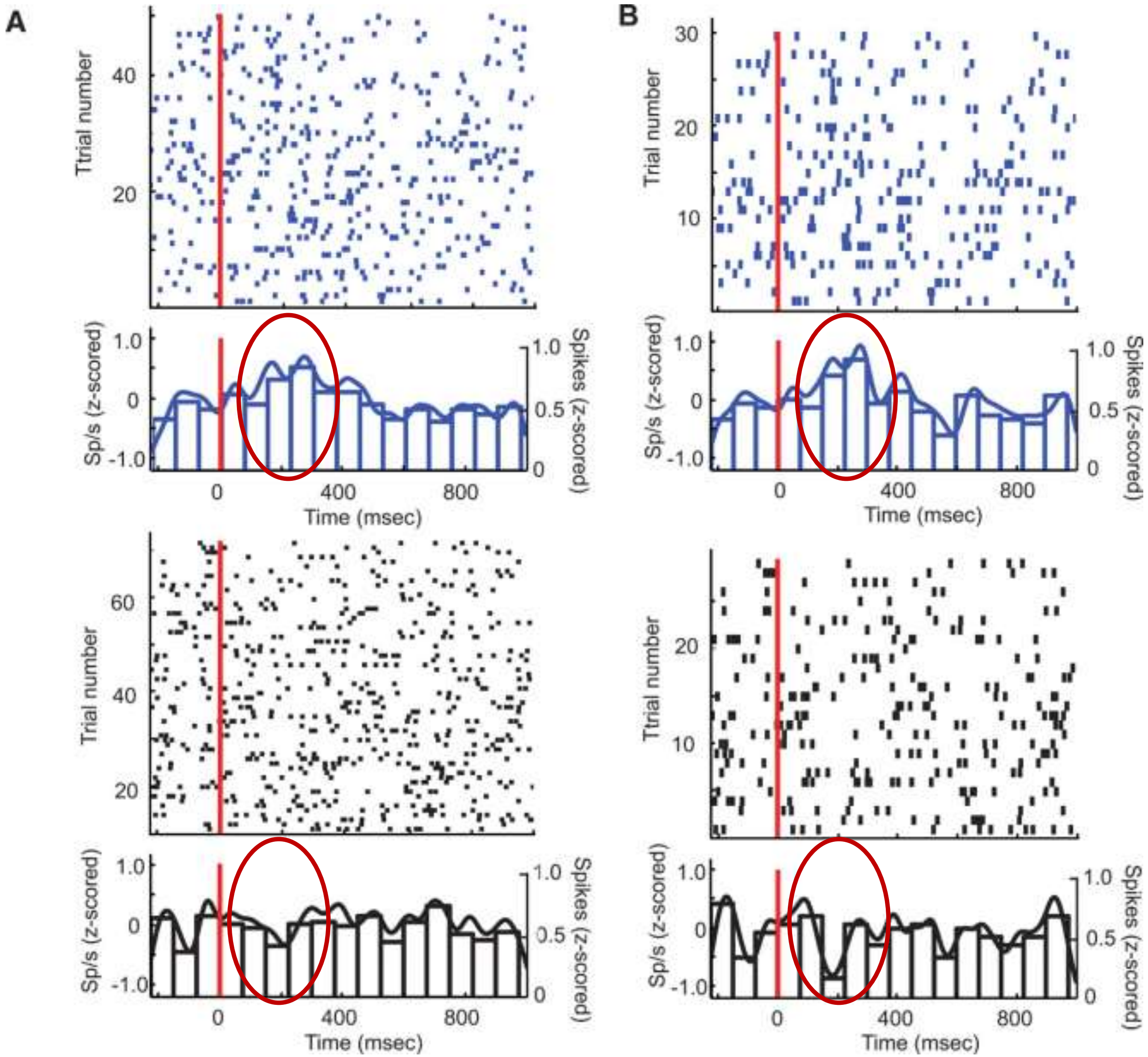
Zaghloul, Blanco, Weidemann, et al. Science. 2009

Electrophysiological correlational evidence

(**A**) Spike raster for a single experiment from one participant. Individual spike activity recorded from SN for trials during **positive (blue) and negative (black) feedback** is shown for each trial as a function of time. Below each spike raster is the average z-scored continuous-time firing rate (continuous trace) and histogram (bars, 75-msec intervals). The red vertical line indicates feedback onset.

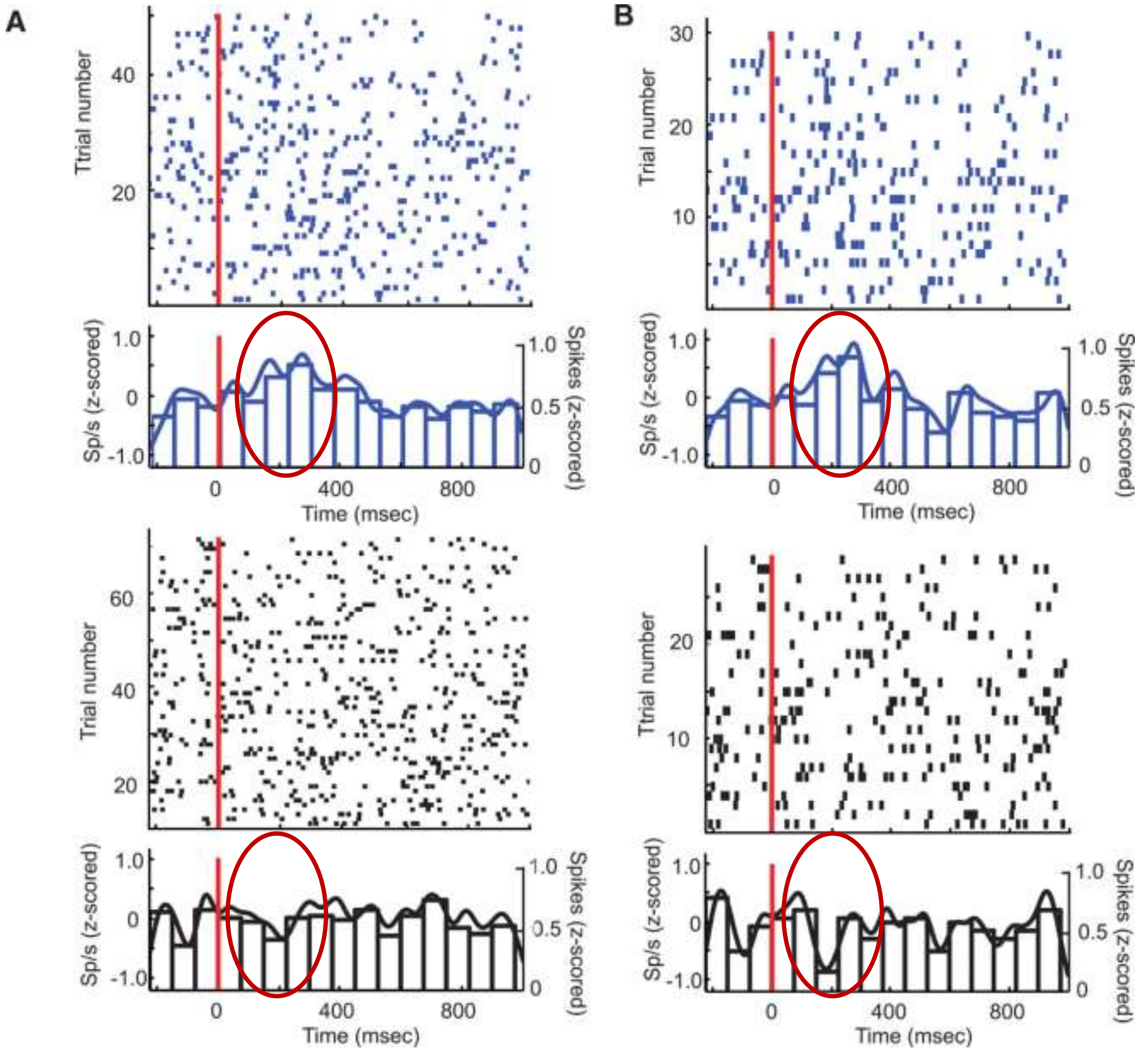
(**B**) Individual spike activity, recorded from the same cell as shown in Fig. 3A, for trials in response to **unexpected gains (blue) and losses (black)** is shown for each trial as a function of time.

Zaghloul, Blanco, Weidemann, et al. Science. 2009



Electrophysiological correlational evidence

- Raw spike count increased in response to positive feedback and decreased in response to negative feedback during this interval
- The difference in activity between responses to unexpected gains and losses was clearer than the difference between positive and negative feedback.



Zaghloul, Blanco, Weidemann, et al. Science. 2009

Notice that so far

We showed evidence that striatal activity changes to unexpected events

but

no direct causal evidence of a link with dopamine and of an effect on behavior

Probabilistic selection task (PST)

Learning phase

- Participants complete an instrumental learning task, which includes three pairs of stimuli (AB, CD, and EF). By trial and error, participants are required to learn the stimulus in each pair more likely to lead to reward.
- Within each pair, choosing one stimulus is more likely to lead to reward (and less likely to lead to punishment) than choosing the other.
- Importantly, the probability of reward and punishment differs for each stimulus (Figure 1), so that each stimulus and the choice associated to it acquire a more or less positive or negative value compared to the remaining ones.
- Figure 2: On each trial, participants choose one stimulus of the pair and reward (positive feedback) or punishment (negative feedback) following the choice is provided.

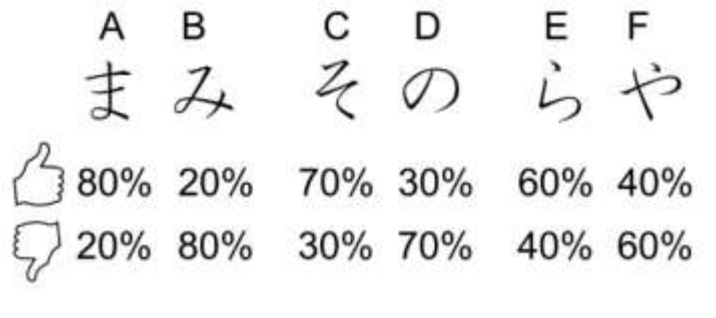


FIGURE 1 | Illustration of examples of stimuli and the feedback probability associated with each stimulus for each type of feedback.

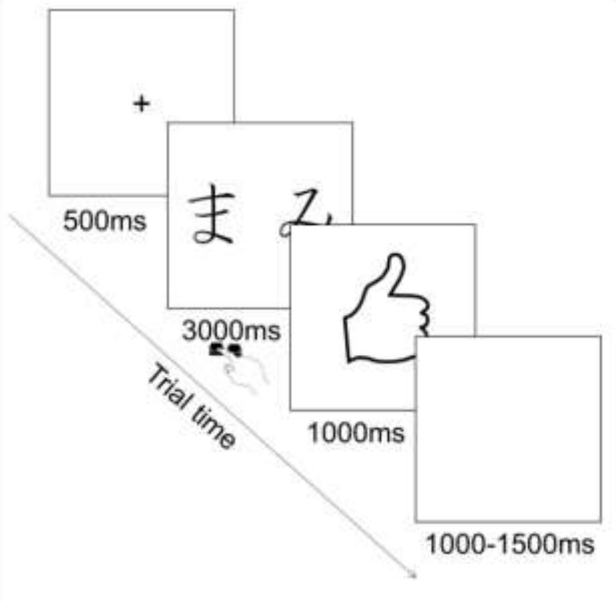


FIGURE 2 | Illustration of experimental trial for the learning phase.

Probabilistic selection task (PST)

Note that learning to choose the more rewarding stimulus in each pair could be accomplished either by learning that:

- choosing the more rewarding stimulus (e.g. A) leads to positive feedback
- choosing the less rewarding stimulus (e.g. B) leads to negative feedback
- Both

Thus, a testing phase is completed to assess whether participants learned more from positive (reward) or negative (punishment) feedback

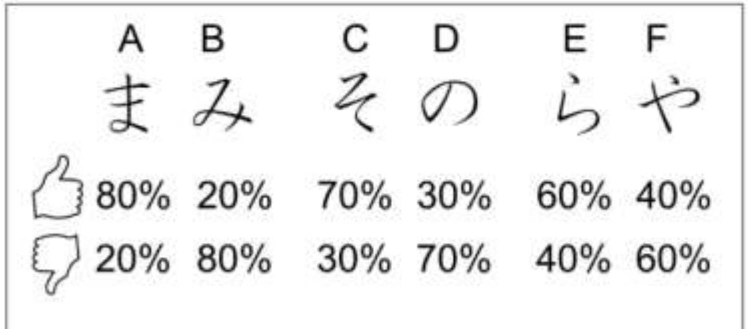


FIGURE 1 | Illustration of examples of stimuli and the feedback probability associated with each stimulus for each type of feedback.

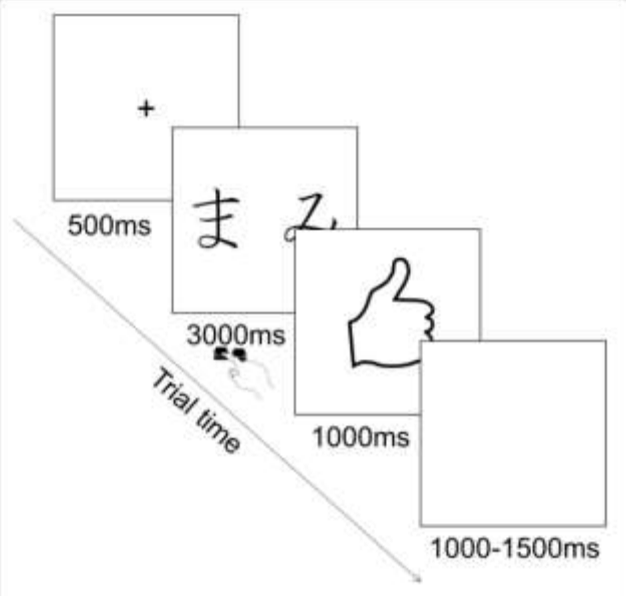


FIGURE 2 | Illustration of experimental trial for the learning phase.

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

Probabilistic selection task (PST)

Testing phase

- Participants are again faced with pairs of stimuli; however, all possible combinations of the stimuli encountered during learning are presented.
- Participants' task remains to choose the stimulus in each pair more likely to lead to reward.
- No feedback is provided about the choice.

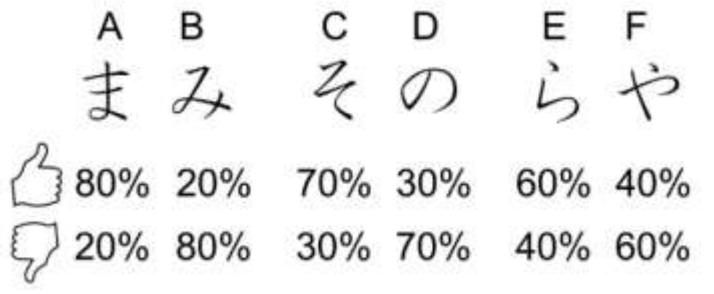


FIGURE 1 | Illustration of examples of stimuli and the feedback probability associated with each stimulus for each type of feedback.

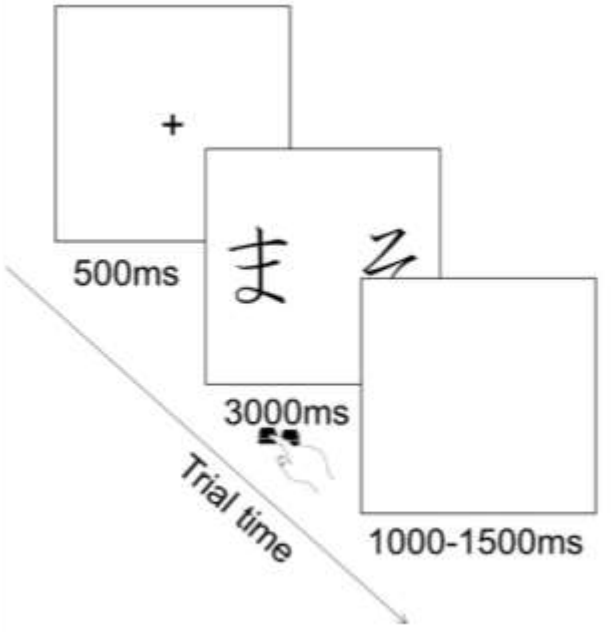


FIGURE 3 | Illustration of experimental trial for the testing phase.

Does dopamine affect performance on the PST?

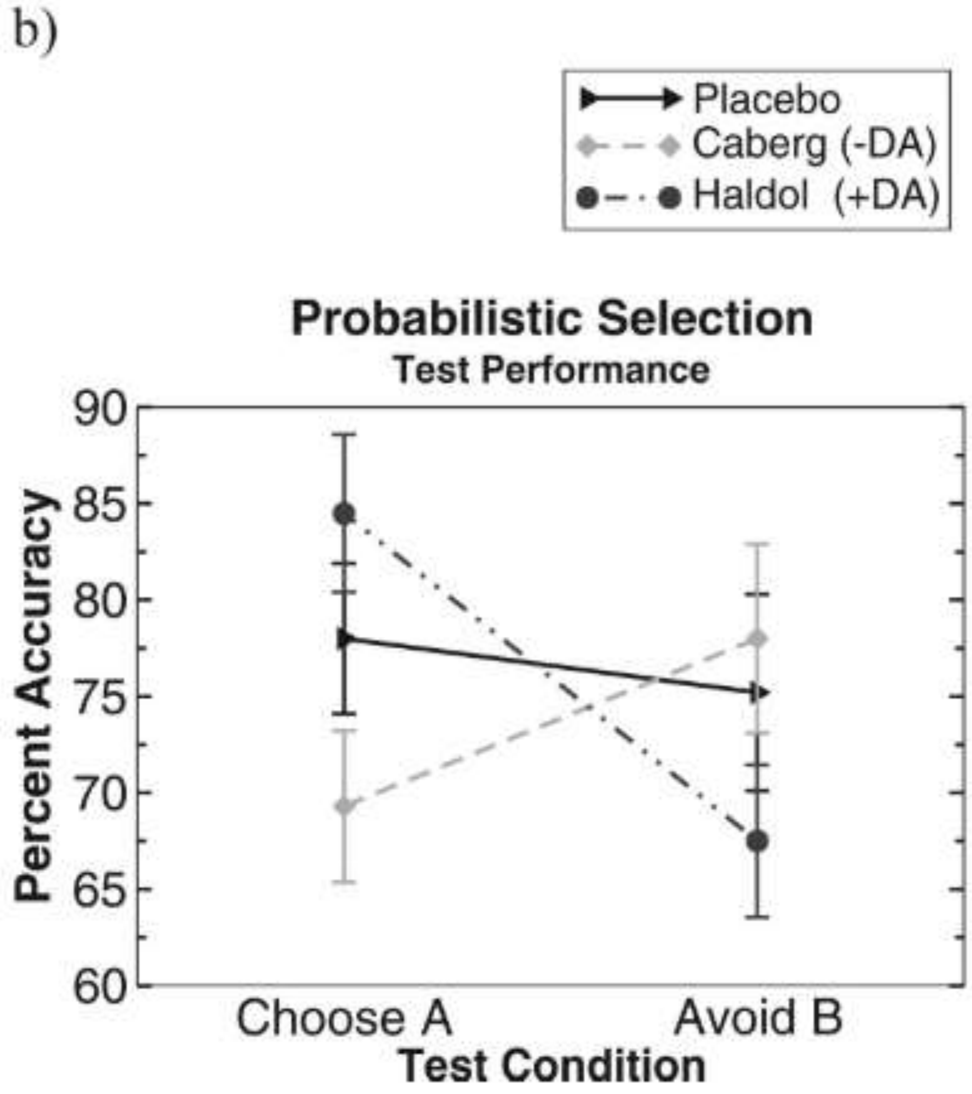
Frank, M. J., & O'Reilly, R. C. (2006). A mechanistic account of striatal dopamine function in human cognition: psychopharmacological studies with cabergoline and haloperidol. *Behavioral neuroscience*, *120*(3), 497.

Healthy participants taking:

- cabergoline (Caberg): dopamine antagonist
- haloperidol (Haldol; present study): dopamine agonist
- Placebo

There was an interaction between medication condition and tendency to learn from positive versus negative feedback.

- Participants on placebo learned equally as much about the positive consequences of choosing Stimulus A and the negative consequences of choosing Stimulus B.
- Cabergoline impaired positive feedback learning.
- Haloperidol enhanced positive feedback learning.

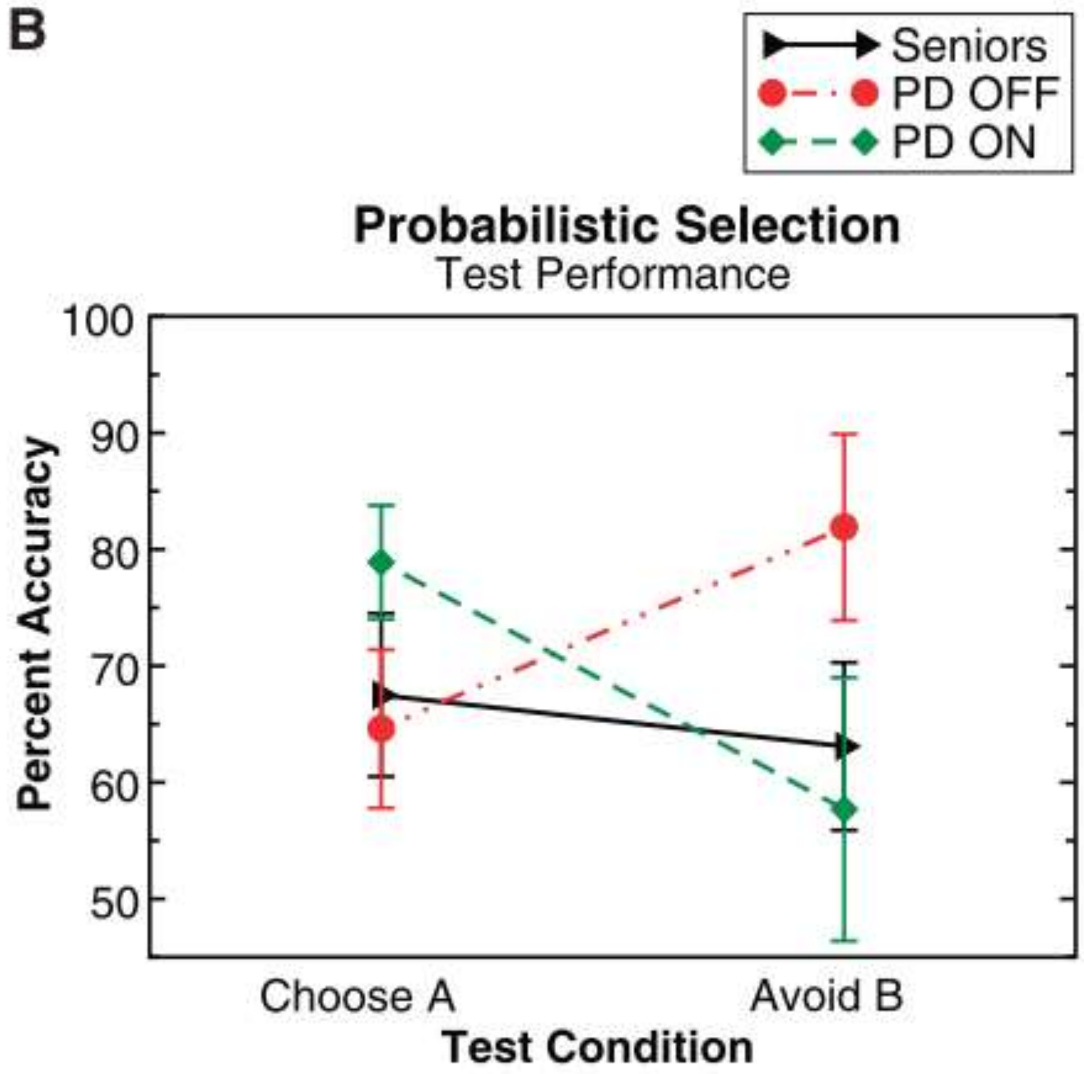


A.D. 1088
ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

Does dopamine affect performance on the PST?

These drug results are similar to those found previously in patients with Parkinson's disease (PD) when

- on dopaminergic (DA) medication patients learn more from positive feedback
- off dopaminergic (DA) medication learn more from negative feedback



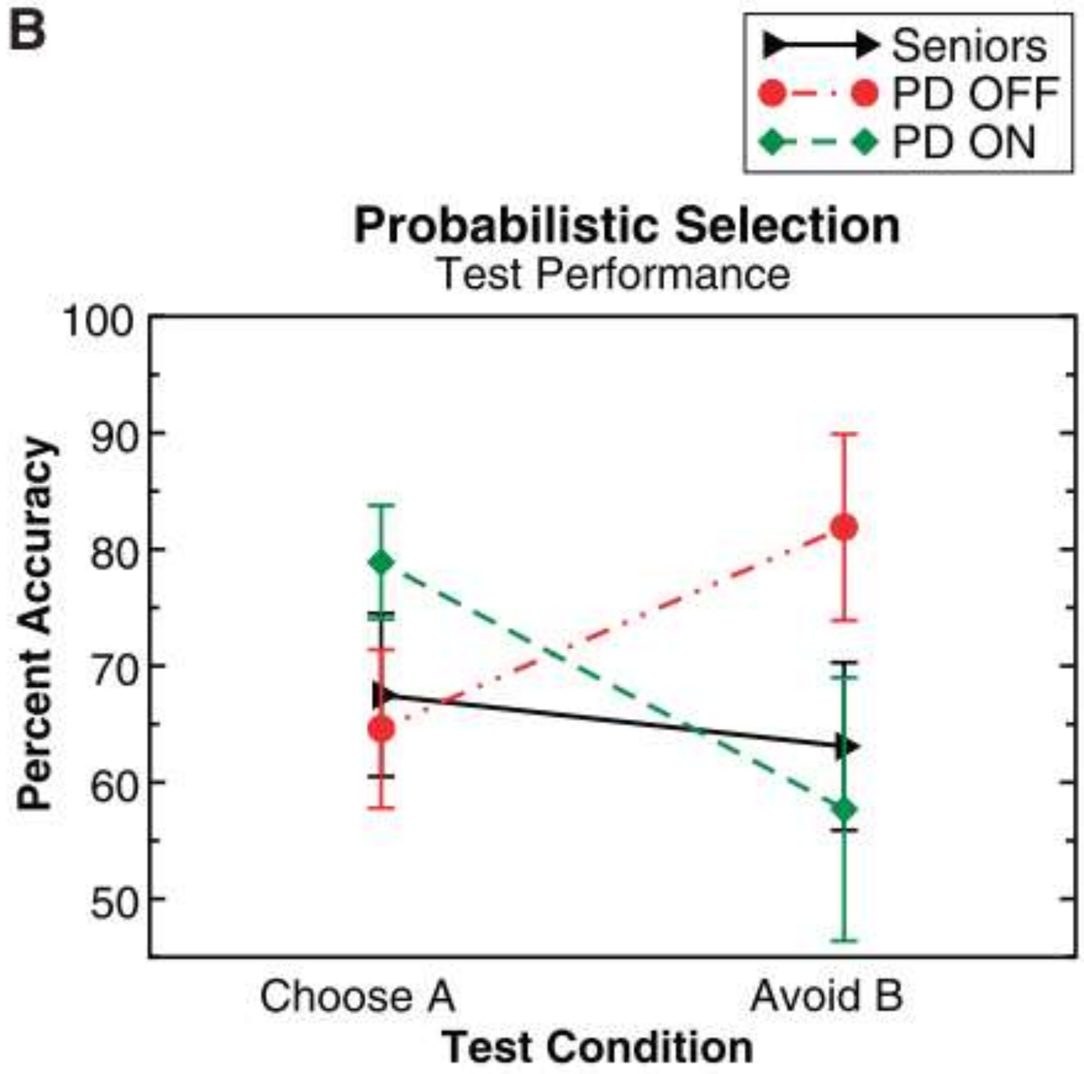
M. J. Frank, L. C. Seeberger, & R. C. O'Reilly, 2004, Science, 306, p. 1941

Does dopamine affect performance on the PST?

YES, BUT WHY?

- Dopamine antagonist/Non-medicated Parkinson's patients are impaired at learning from positive feedback: because of reduced levels of dopamine, +PE following positive feedback cannot occur, while –PE can still happen

- Dopamine agonist/Patients on medication have sufficient dopamine to learn from positive feedback (+PE), but are relatively impaired at learning from negative feedback because the medication blocks the effects of normal dopamine dips (-PE)



M. J. Frank, L. C. Seeberger, & R. C. O'Reilly, 2004, Science, 306, p. 1941

Behavioral & pharmacological causal evidence

During **instrumental learning**, the behavioral and neural effects of drugs

- enhancing (3,4-dihydroxy-L-phenylalanine;L-DOPA) dopaminergic function
- reducing (haloperidol) dopaminergic function
- placebo

was assessed in groups of healthy subjects

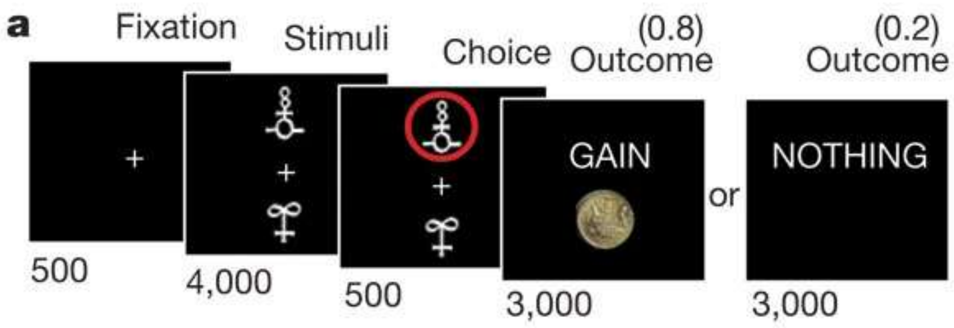
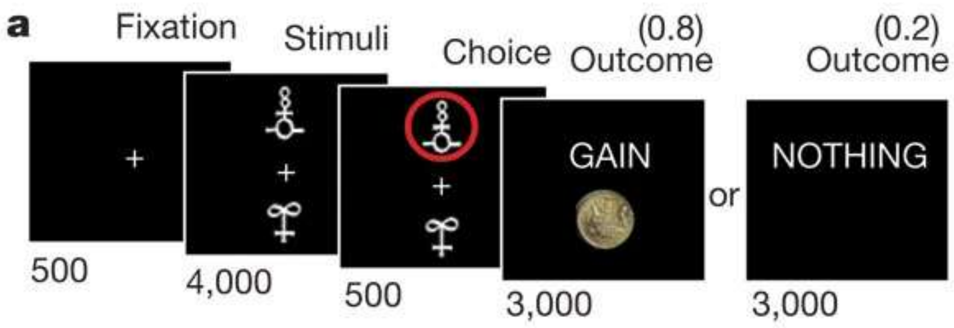


Fig. a, Experimental task. Subjects selected either the upper or lower of two abstract visual stimuli presented on a display screen, and subsequently observed the outcome.

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

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA
CAMPUS DI CESENA

Behavioral & pharmacological causal evidence



Each stimulus was associated with a certain probability of gain or loss:

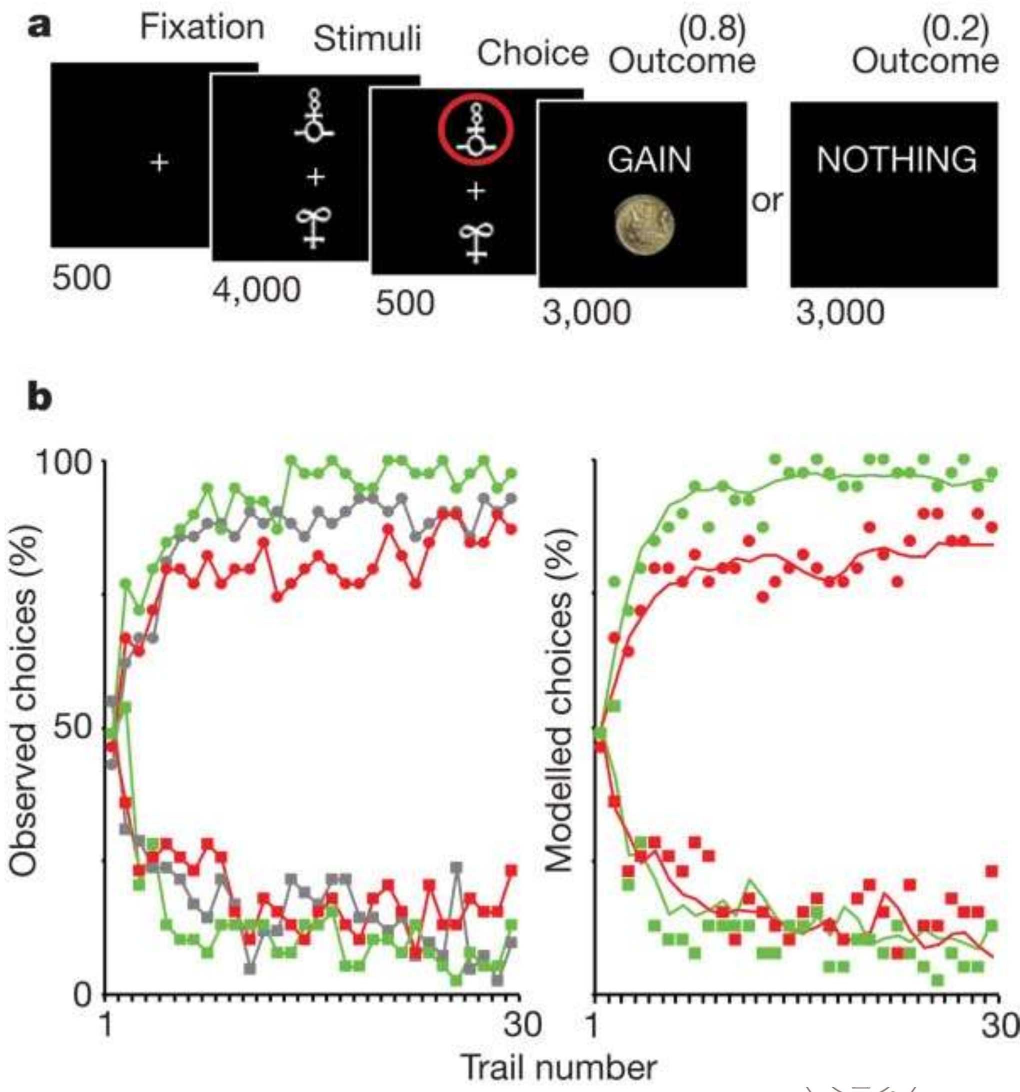
- one pair of stimuli was associated with **gains** (£1 or nothing) --> assess the effects of the drugs on the ability to **learn from rewards**
- a second pair was associated with **loss** (- £1 or nothing) --> assess the effects of the drugs on the ability to **learn from punishments**
- a third pair was associated with **no financial outcomes** --> **neutral control** condition

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

Behavioral & pharmacological causal evidence

Fig. b, Behavioural results.

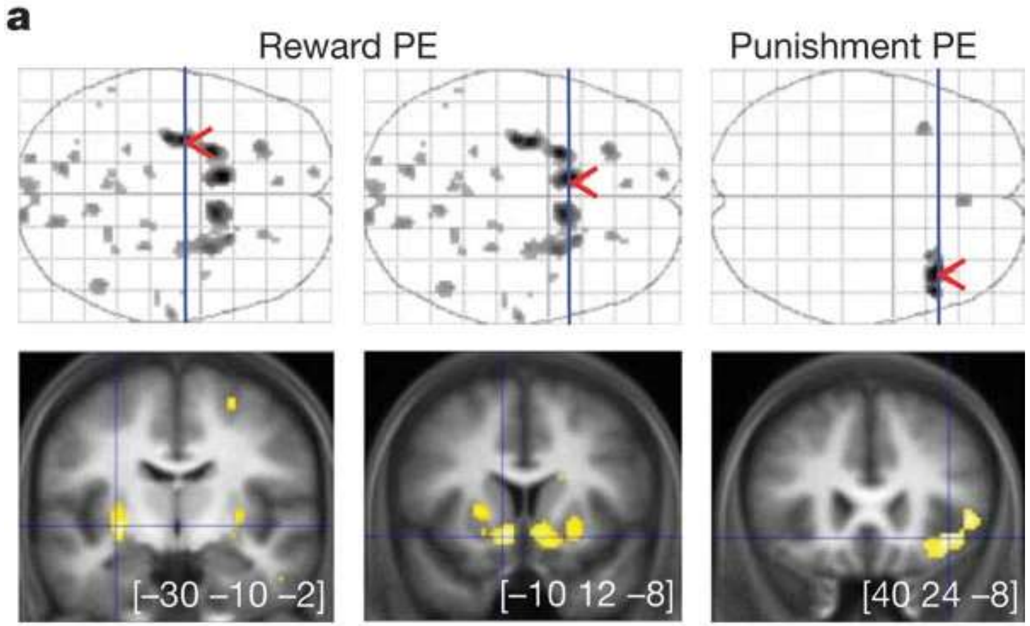
- Left: observed behavioural choices for **placebo (grey)**, **L-DOPA (green)** and **haloperidol (red)**. The learning curves depict, trial by trial, the **proportion of subjects that chose the 'correct' stimulus** (associated with a probability of 0.8 of winning £1) **in the gain condition** (circles, upper graph), **and the 'incorrect' stimulus** (associated with a probability of 0.8 of losing £1) **in the loss condition** (squares, lower graph).
- Right: modelled behavioural choices for L-DOPA (green) and haloperidol (red) groups. The learning curves represent the probabilities predicted by the computational model. Circles and squares representing observed choices have been left for the purpose of comparison.



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

Behavioral & pharmacological causal evidence

- A standard algorithm of action-value learning was then fitted to the observed behaviour. Outcome prediction errors estimated by the model were then used as a statistical regressor in the imaging data.
- Brain activity correlated with prediction errors derived from the computational model.
- **Reward prediction errors** (both from punishment and reward) correlated with activity in the **left posterior putamen, left ventral striatum**
- punishment prediction errors correlated with activity in the right anterior insula



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

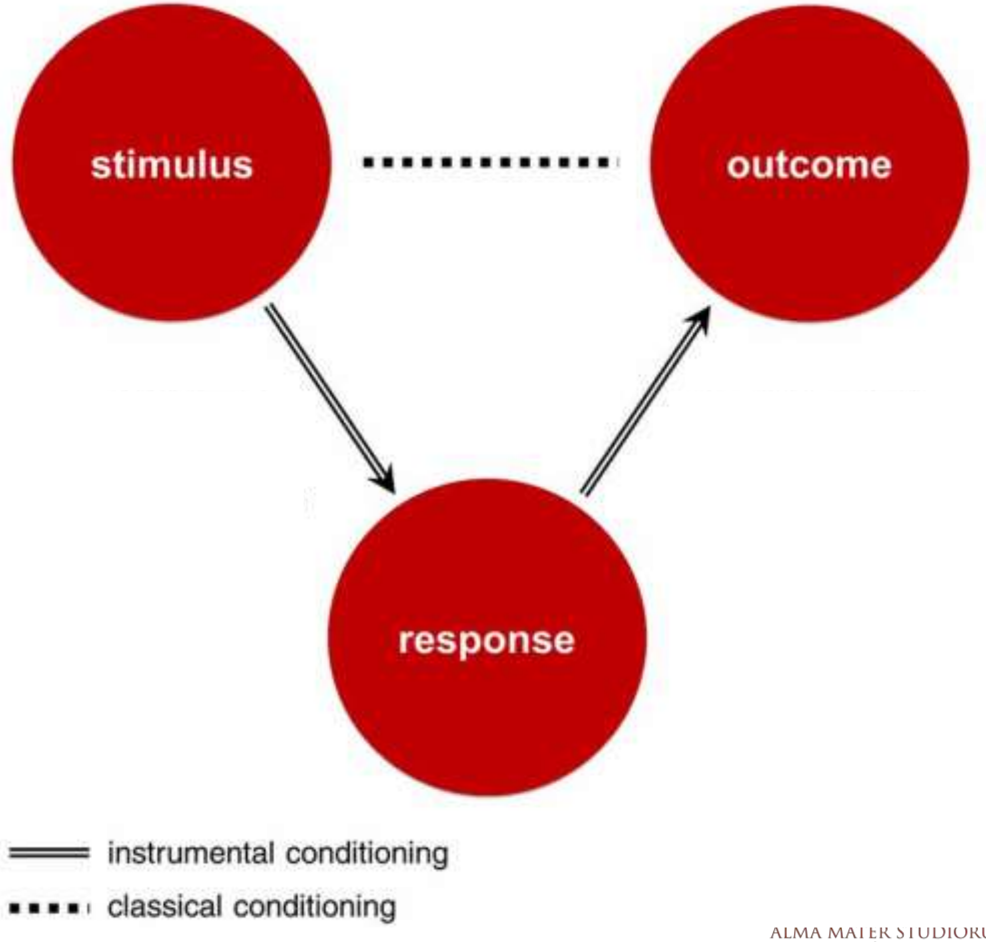
From predictions to control and choice

Predictions are for control

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

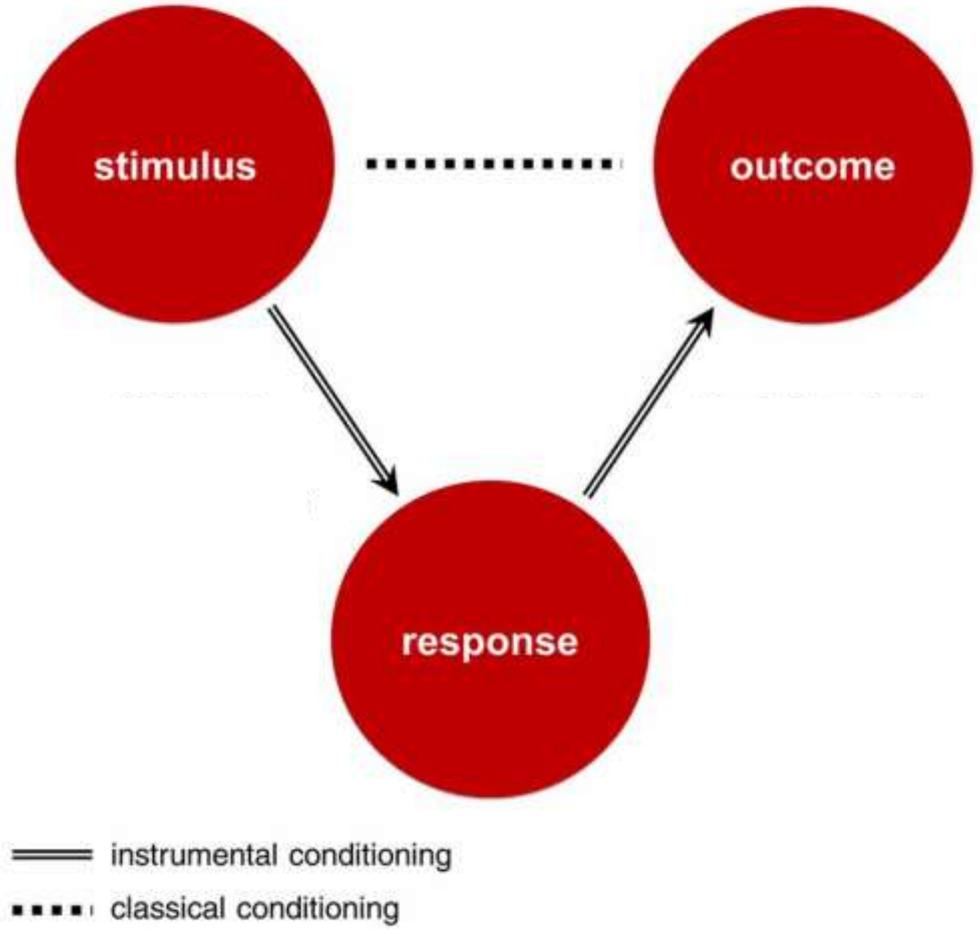


From predictions to control and choice: The actor-critic model

The cortex represents the current state, and the basal ganglia implement two computational modules:

1. **the critic**: learns stimulus-outcome associations (or state values)

2. **the actor**: learns stimulus-response associations

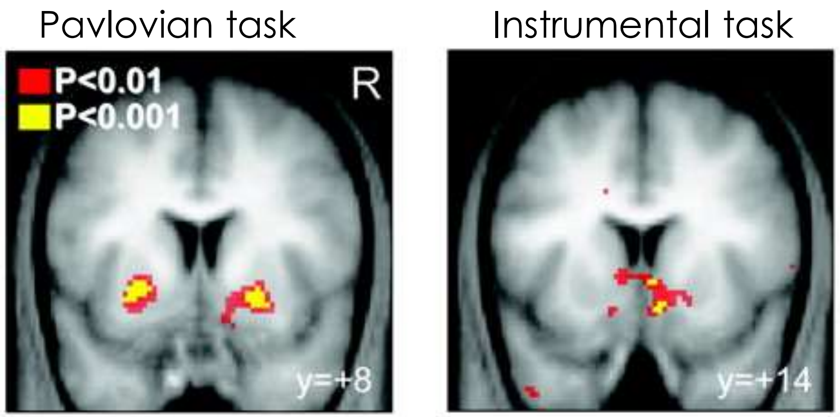


From predictions to control and choice: The actor-critic model

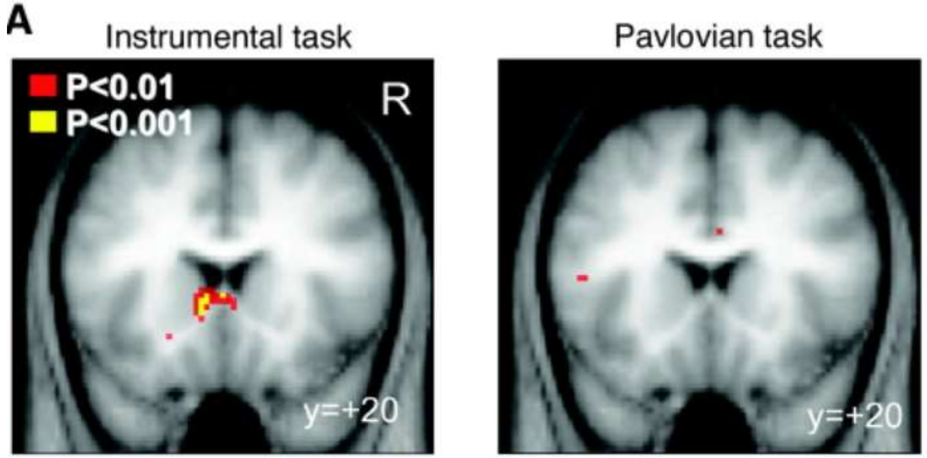
The cortex represents the current state and the basal ganglia implement two computational modules:

1. **the critic**: learns stimulus-outcome associations (or state values) and may be implemented in the **ventral striatum**, and possibly in the amygdala and orbitofrontal cortex (OFC)
2. **the actor**: learns stimulus-response associations and may be implemented in the **dorsal striatum**

The critic: ventral striatum active during both Pavlovian and instrumental task



The actor: dorsal striatum active during instrumental but not Pavlovian task



O'Doherty, J., Dayan, P., Schultz, J., Deichmann, R., Friston, K., & Dolan, R. J. (2004). Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *science*, *304*(5669), 452-454.

Implications of the reward prediction error hypothesis of dopamine neurons: a theory of addiction

Implications of the reward prediction error hypothesis of dopamine neurons

Dopamine

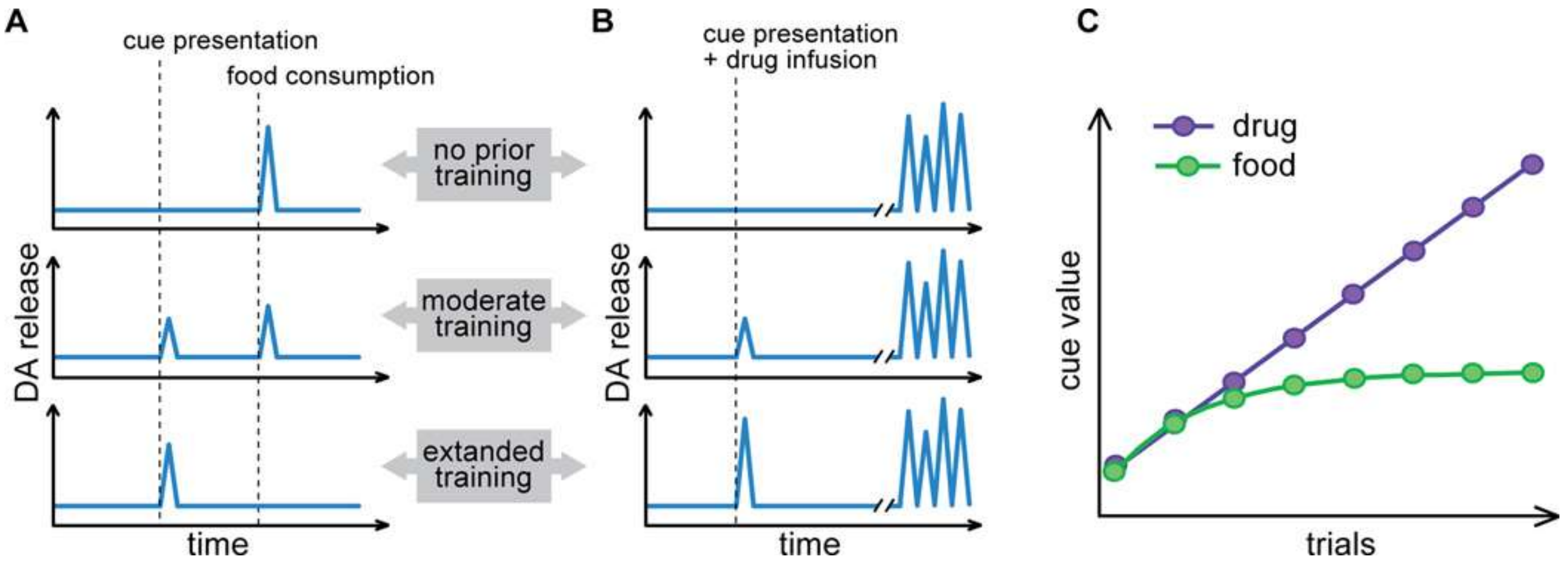
- is not the feel-good chemical
- does not simply make us feel good
- does not simply tell us how much we like something

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

Dopamine reward prediction errors and addiction

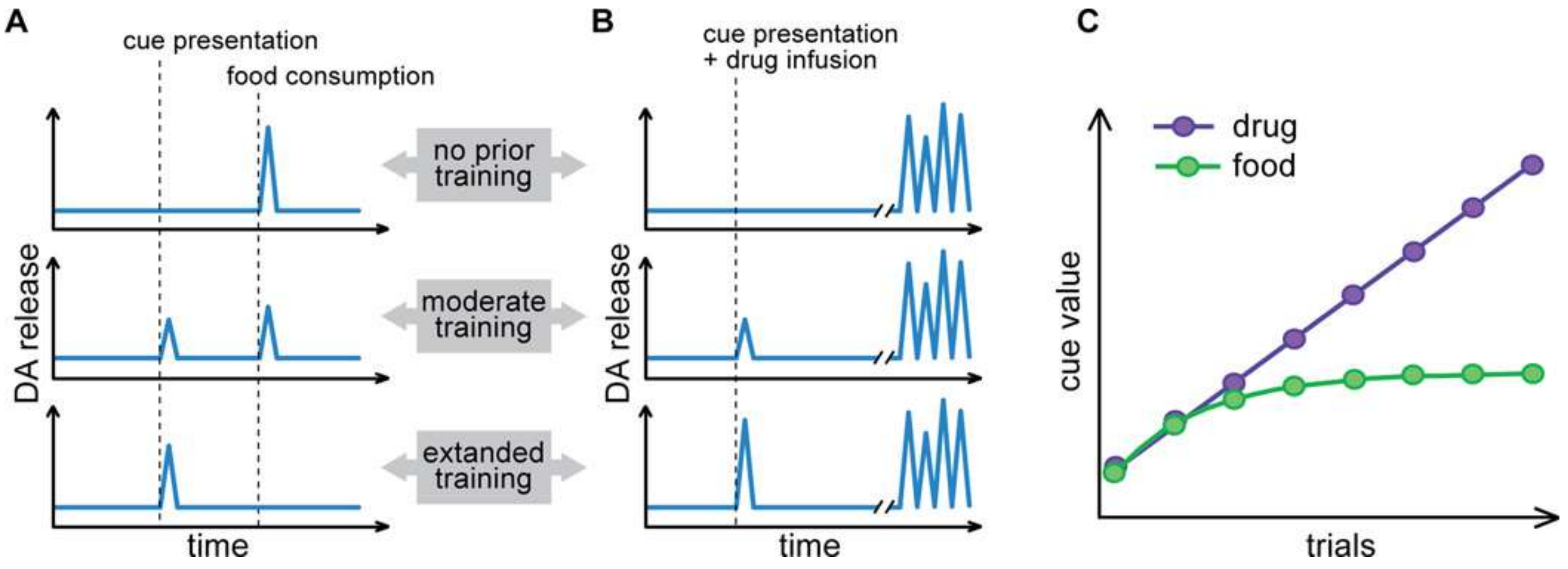


Differences between food-related and cocaine-related phasic DA signals

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

Dopamine reward prediction errors and addiction

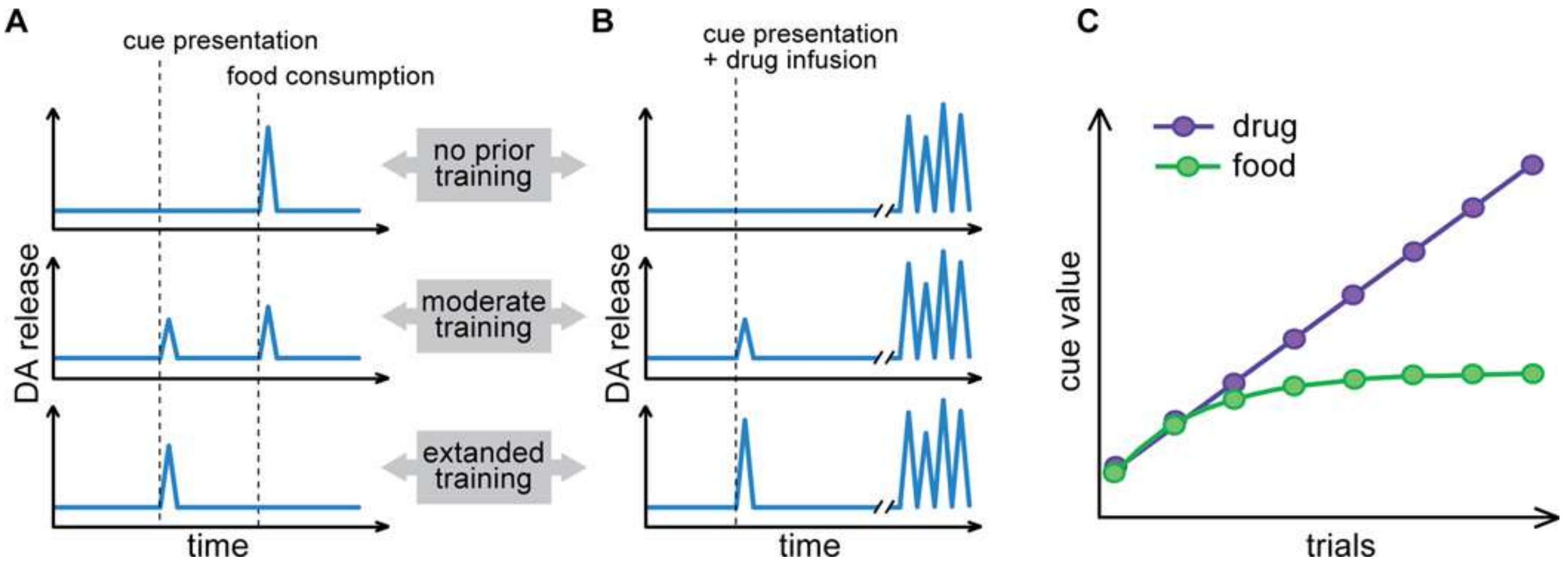


Differences between food-related and cocaine-related phasic DA signals

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

Dopamine reward prediction errors and addiction



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

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