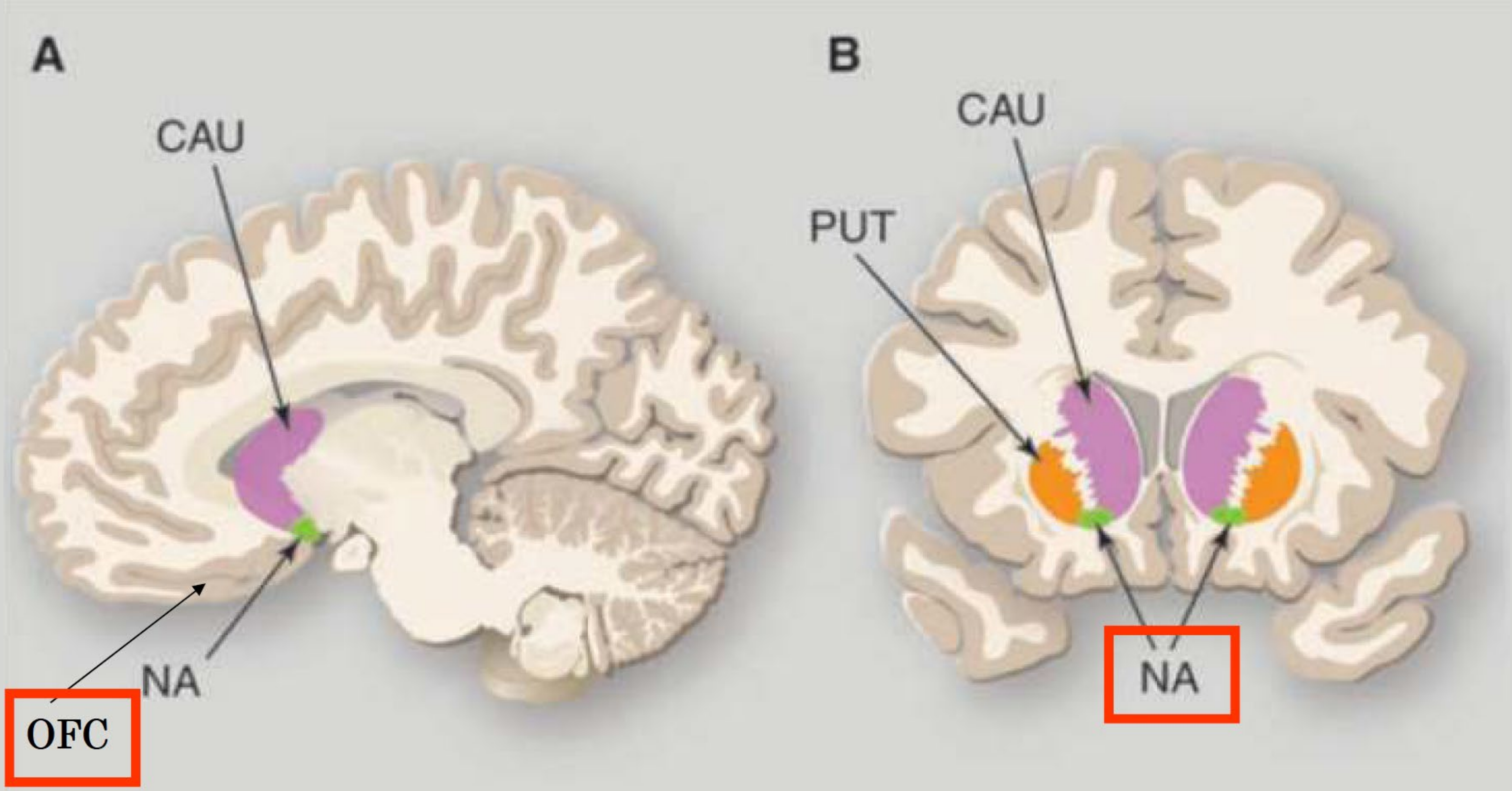
10

reward prediction error

The OFC and NA connection:



OFC: Orbital frontal cortex   NA: Nucleus Accumbens   PUT: Putamen   CAU: Caudate

Stereotactic and Functional Neurosurgery

Review

Stereotact Funct Neurosurg 2015;93:75–93
DOI: 10.1159/000368279

Received: March 8, 2014
Accepted after revision: September 10, 2014
Published online: February 18, 2015

The Nucleus Accumbens: A Comprehensive Review

Sanjay Salgado Michael G. Kaplitt

Laboratory of Molecular Neurosurgery, Department of Neurological Surgery, Weill Cornell Medical College, New York, N.Y., USA

Afferent and efferent connections to NAc.

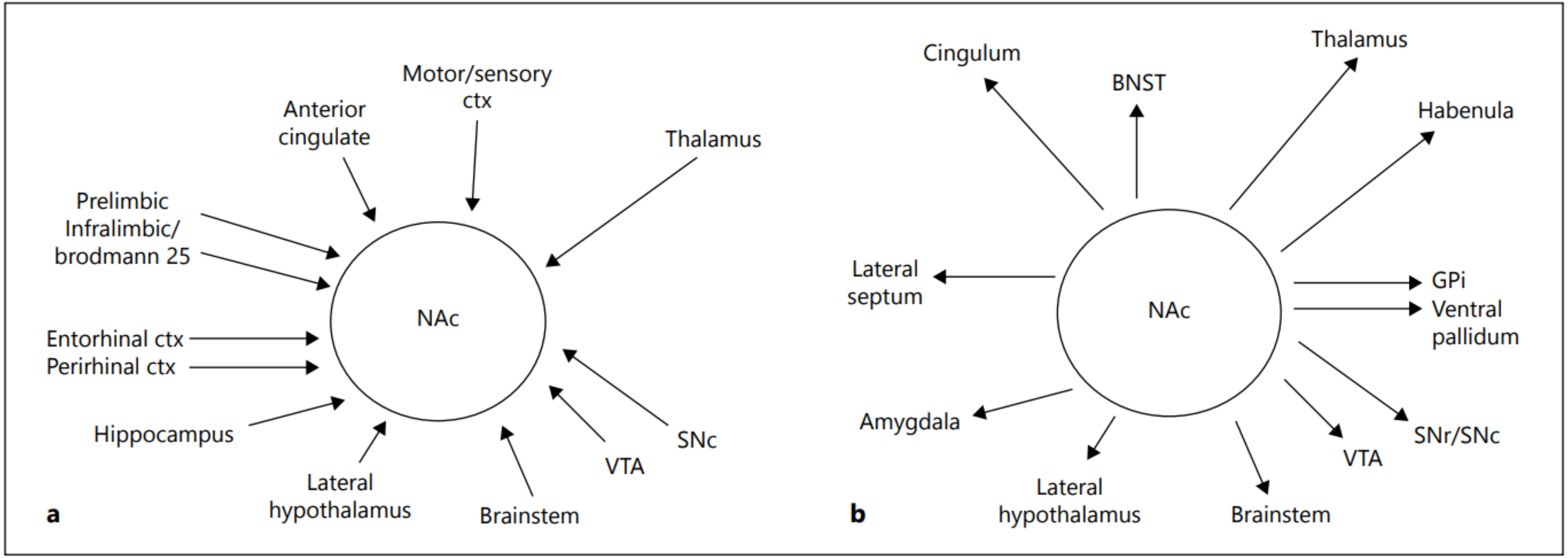


Fig. 1. Schematic representation of major afferent (**a**) and efferent (**b**) connections of the NAc. SNc = Substantia nigra pars compacta; BNST = bed nucleus of the stria terminalis; GPi = globus pallidus internus; SNr = substantia nigra pars reticulata; ctx = cortex.

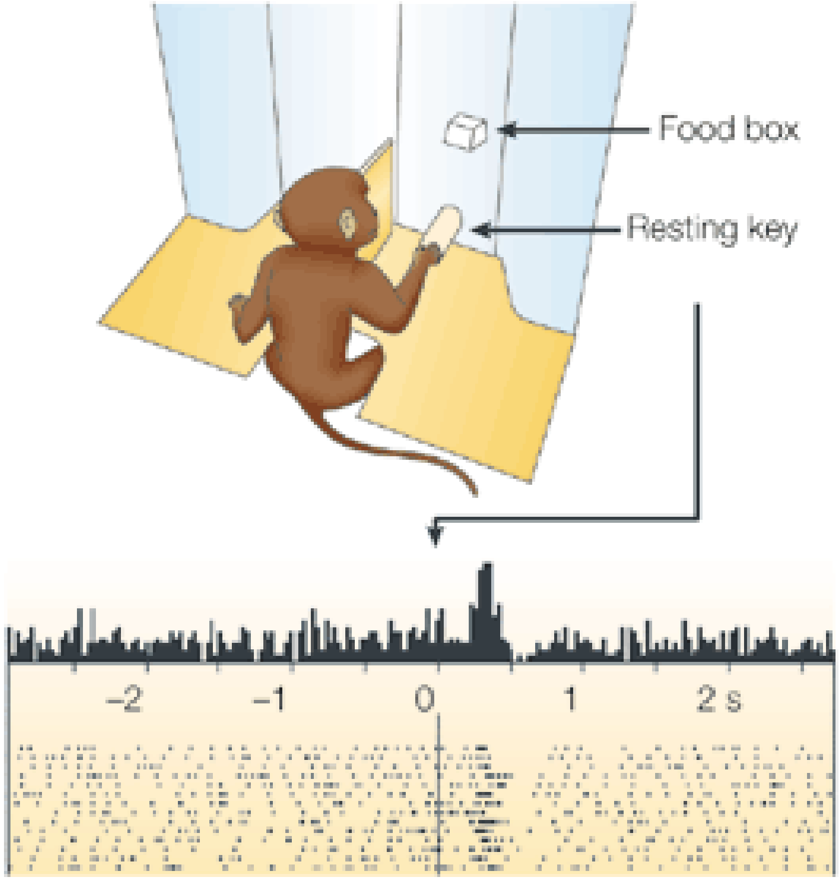
Salgado and Kaplitt (2015) Stereotact Funct Neurosurg 93:75-93

MULTIPLE REWARD SIGNALS IN THE BRAIN

Wolfram Schultz

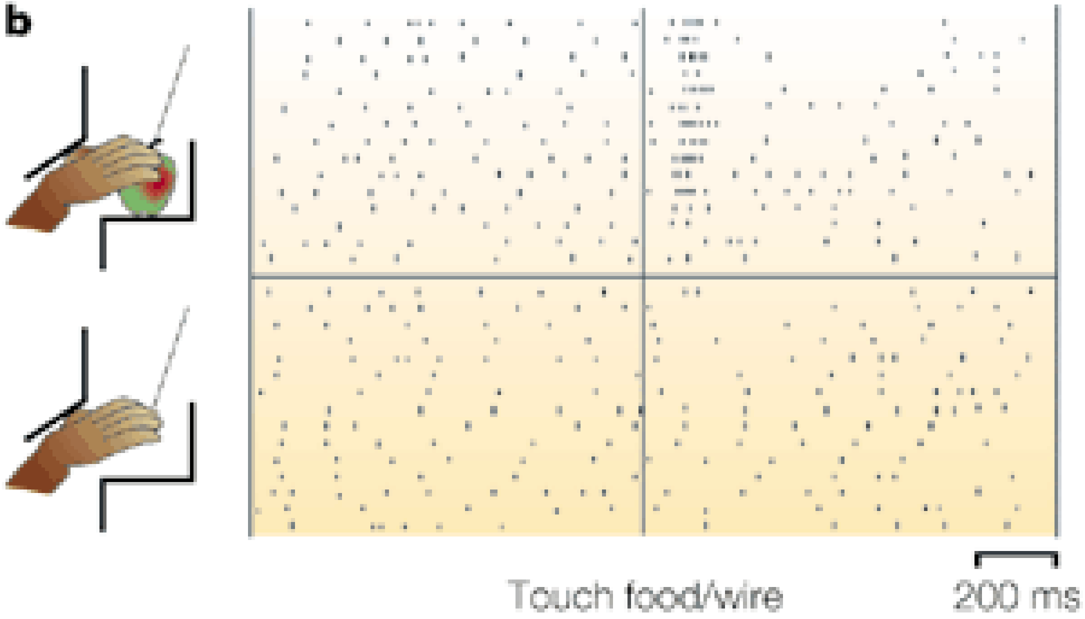
The fundamental biological importance of rewards has created an increasing interest in the neuronal processing of reward information. The suggestion that the mechanisms underlying drug addiction might involve natural reward systems has also stimulated interest. This article focuses on recent neurophysiological studies in primates that have revealed that neurons in a limited number of brain structures carry specific signals about past and future rewards. This research provides the first step towards an understanding of how rewards influence behaviour before they are received and how the brain might use reward information to control learning and goal-directed behaviour.

Schultz, W. (2000) NATURE REVIEWS | NEUROSCIENCE VOLUME 1

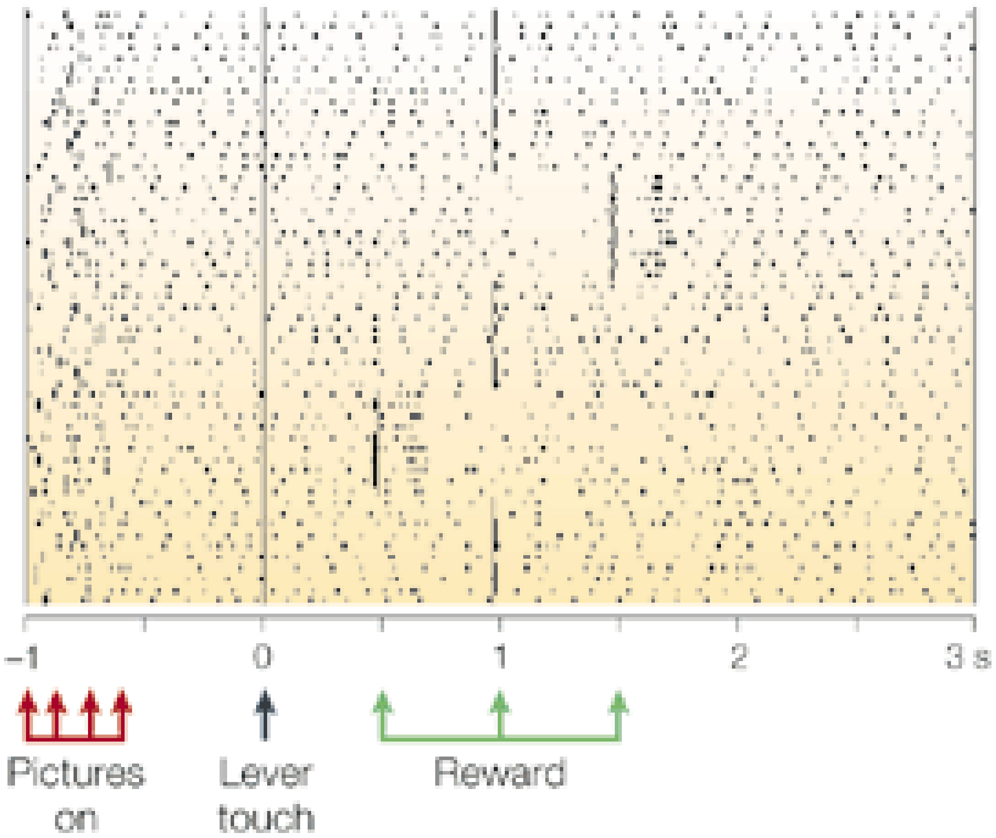


“The food is invisible to the monkey but the monkey can touch the food by placing its hand underneath the protective cover. The peri-event time histogram of the neuronal impulses is shown above the raster display, in which each dot denotes the time of a neuronal impulse in reference to movement onset (release of resting key). Each horizontal line represents the activity of the same neuron on successive trials, with the first trials presented at the top and the last trials at the bottom of the raster display. **a | Touching food reward in the absence of stimuli that predict reward** produces a brief increase in firing rate within 0.5 s of movement initiation.”

Schultz, W. (2000) NATURE REVIEWS | NEUROSCIENCE VOLUME 1



"Touching a **piece of apple (top) enhances the firing rate** but touching the bare wire or an inedible object that the monkey had previously encountered does not. The traces are aligned to a temporal reference point provided by touching the hidden object (vertical line)."



Dopamine neurons encode an error in the temporal prediction of reward.
The firing rate is depressed when the reward is delayed beyond the expected time-point (1 s after lever touch). The **firing rate is enhanced** at the new time of reward delivery whether it is **delayed** (1.5 s) or **precocious** (0.5 s). The three arrows indicate, from left to right, the time of precocious, **habitual and delayed reward delivery**. The original trial sequence is from top to bottom. Data are from a two-picture discrimination task.

Schultz, W. (2000) NATURE REVIEWS | NEUROSCIENCE VOLUME 1

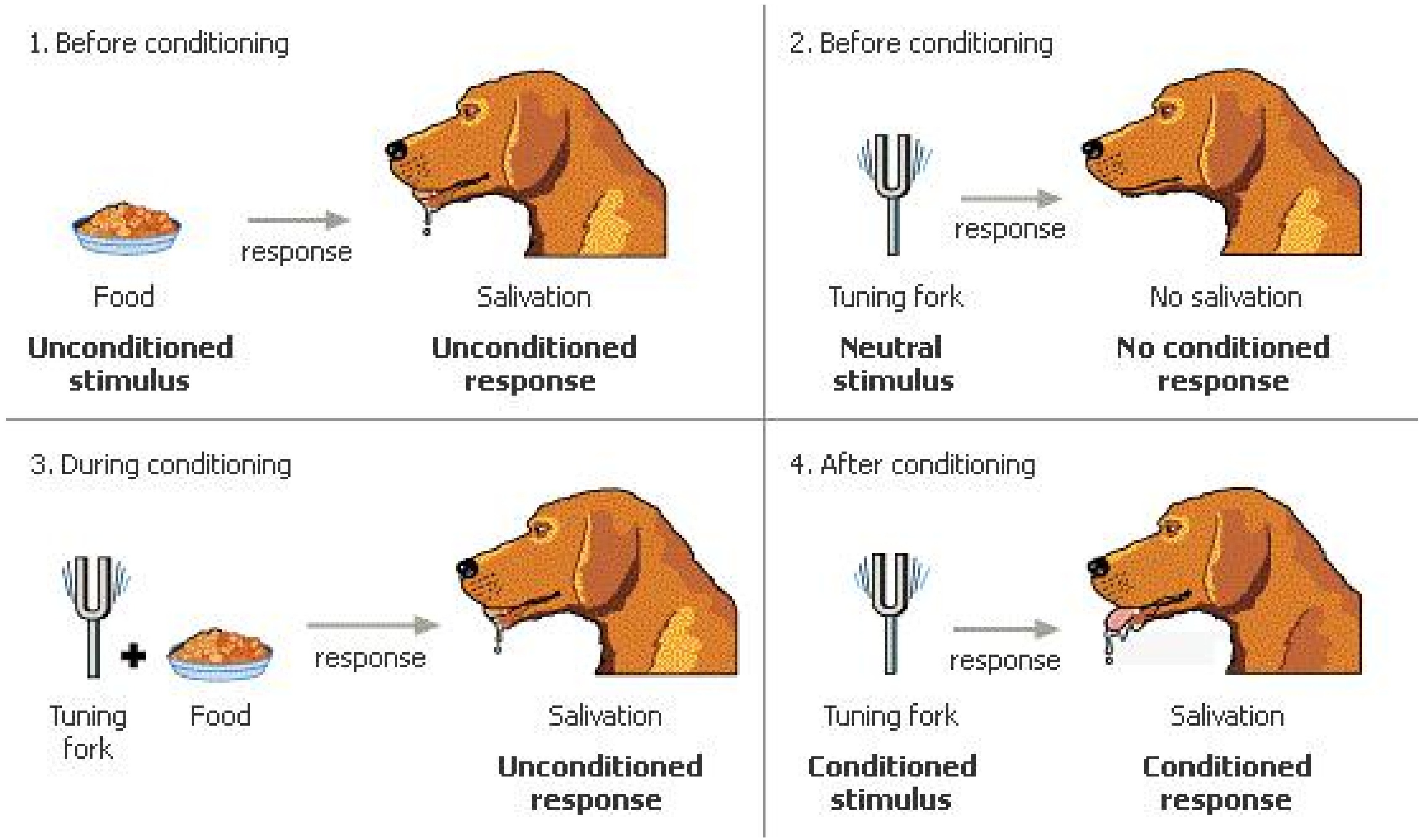
Annu. Rev. Psychol. 2006. 57:87–115
doi: 10.1146/annurev.psych.56.091103.070229

First published online as a Review in Advance on September 16, 2005

BEHAVIORAL THEORIES AND THE NEUROPHYSIOLOGY OF REWARD

Wolfram Schultz
Department of Anatomy, University of Cambridge, CB2 3DY United Kingdom;
email: ws234@cam.ac.uk

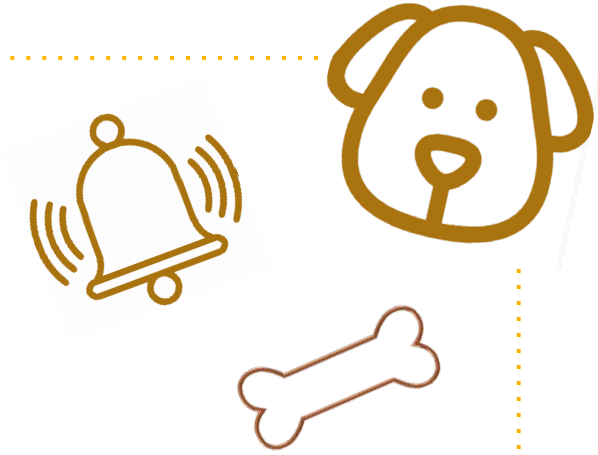
Schultz, W. (2006) *Annu. Rev. Psychol*. 57:87–115



1. Before conditioning
response
Food
Unconditioned stimulus
Salivation
Unconditioned response
2. Before conditioning
response
Tuning fork
Neutral stimulus
No salivation
No conditioned response
3. During conditioning
response
Tuning fork
Food
Salivation
Unconditioned response
4. After conditioning
response
Tuning fork
Conditioned stimulus
Salivation
Conditioned response

“ Learning occurs …

… as the previously neutral stimulus obtains predictive value for the coming reward…



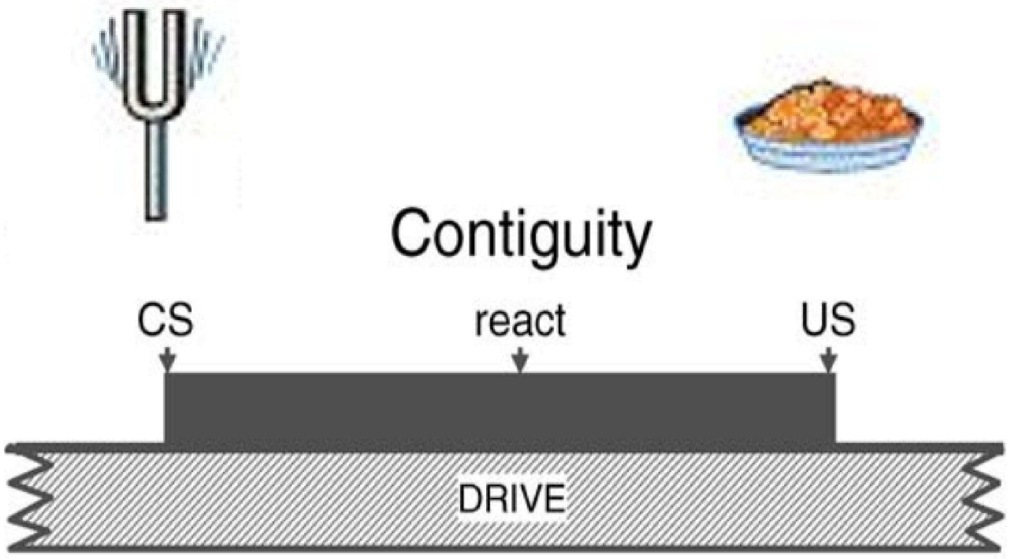
…Eventually, this novel cue is able to evoke a response that is often topographically similar to that produced by the unconditioned stimulus itself. ”

Day, J. J. and Carelli, R. M. (2007) *The Neuroscientist*, Volume 13, Number 2

Environmental cues update expectancies

Pavlovian relationships may be embedded within all operant circumstances (??)

Basic assumptions of animal learning theory **defining the behavioral functions of rewards.**

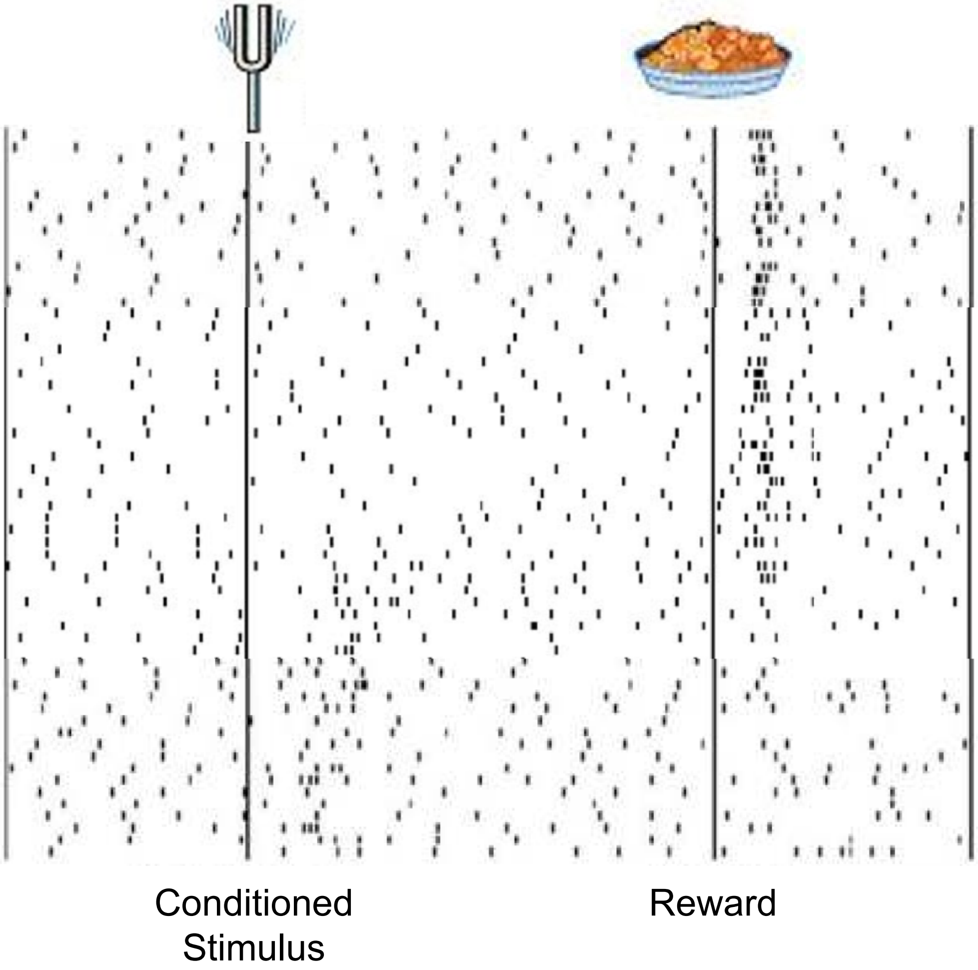


Contiguity refers to the temporal proximity of a conditioned stimulus (CS), or action, and the reward.

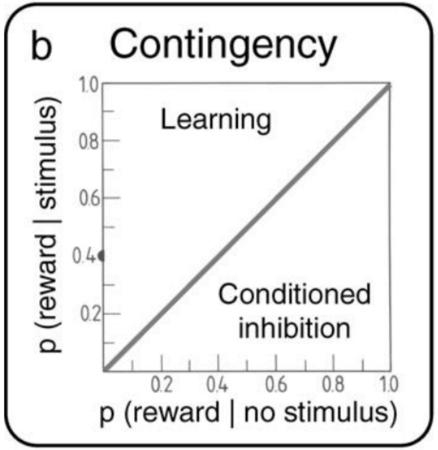
Specifically, a reward needs to follow a CS or response by an optimal interval of a few seconds, whereas rewards occurring before a stimulus or response do not contribute to learning (backward conditioning).

Schultz, W. (2006) *Annu. Rev. Psychol*. 57:87–115

Reward and prediction of reward in N. Accumbens



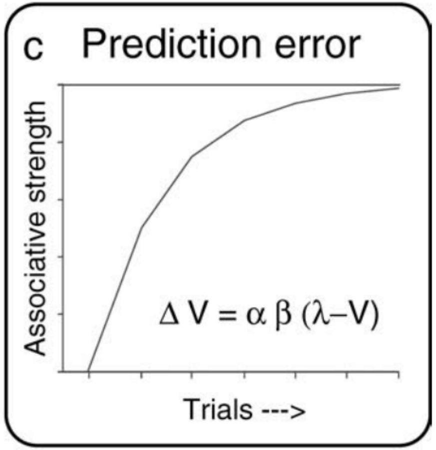
Schultz, W. (2006) *Annu. Rev. Psychol.* 57:87–115



Contingency refers to the conditional probability of reward occurring in the presence of a conditioned stimulus as opposed to its absence.

This means, that a reward needs to occur more frequently in the presence of a stimulus as compared with its absence in order to induce "excitatory" conditioning of the stimulus.

Prediction error denotes the discrepancy between 'an actually received reward' and its prediction. Learning (ΔV, associative strength) is proportional to the prediction error ($\lambda - V$) and reaches its asymptote when the prediction error approaches zero after several learning trials.



Schultz, W. (2006) *Annu. Rev. Psychol*. 57:87–115

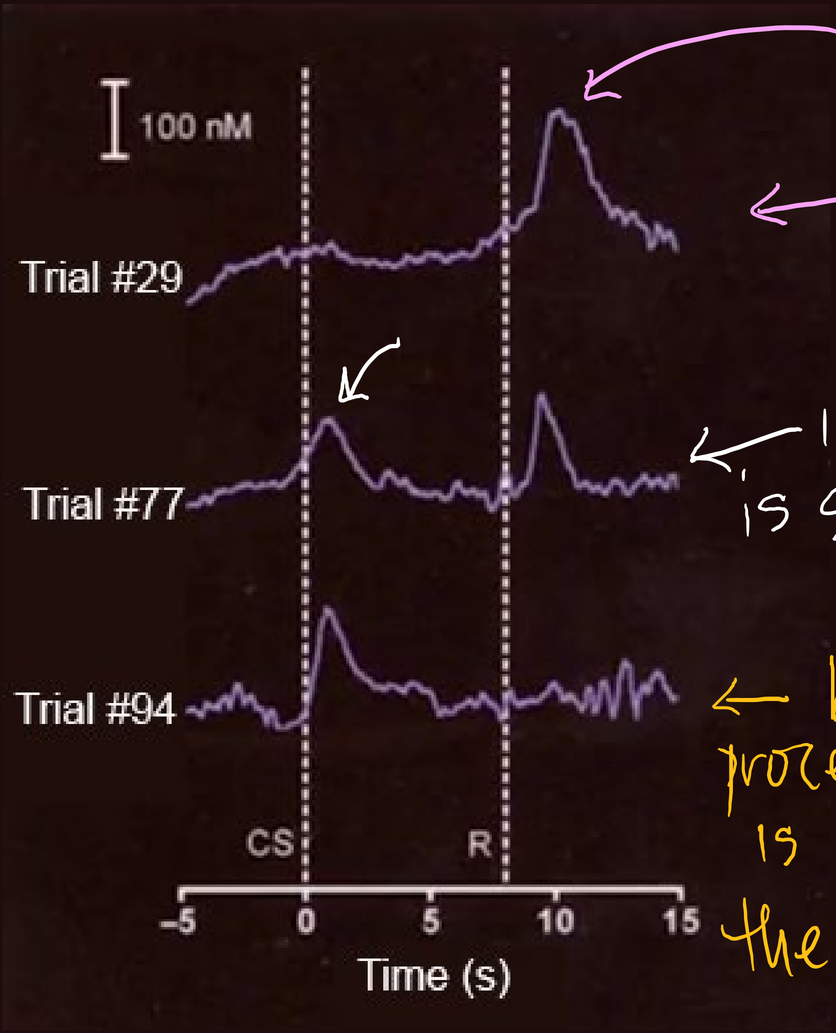
The Nucleus Accumbens and Pavlovian Reward Learning

JEREMY J. DAY and REGINA M. CARELLI
Department of Psychology
University of North Carolina at Chapel Hill

The ability to form associations between predictive environmental events and rewarding outcomes is a fundamental aspect of learned behavior. This apparently simple ability likely requires complex neural processing evolved to identify, seek, and use natural rewards and redirect these activities based on updated sensory information. Emerging evidence from both animal and human research suggests that this type of processing is mediated in part by the nucleus accumbens (NAc) and a closely associated network of brain structures. The NAc is required for a number of reward-related behaviors and processes specific information about reward availability, value, and context. In addition, this structure is critical for the acquisition and expression of most Pavlovian stimulus-reward relationships, and cues that predict rewards produce robust changes in neural activity in the NAc. Although processing within the NAc may enable or promote Pavlovian reward learning in natural situations, it has also been implicated in aspects of human drug addiction, including the ability of drug-paired cues to control behavior. This article provides a critical review of the existing animal and human literature concerning the role of the NAc in Pavlovian learning with nondrug rewards and considers some clinical implications of these findings. NEUROSCIENTIST 13(2):148–159, 2007. DOI: 10.1177/1073858406295854

KEY WORDS *Learning, Reward, Nucleus accumbens, Drug addiction, Conditioning*

Day, J. J. and Carelli, R. M. (2007) *The Neuroscientist*, Volume 13, Number 2



reward

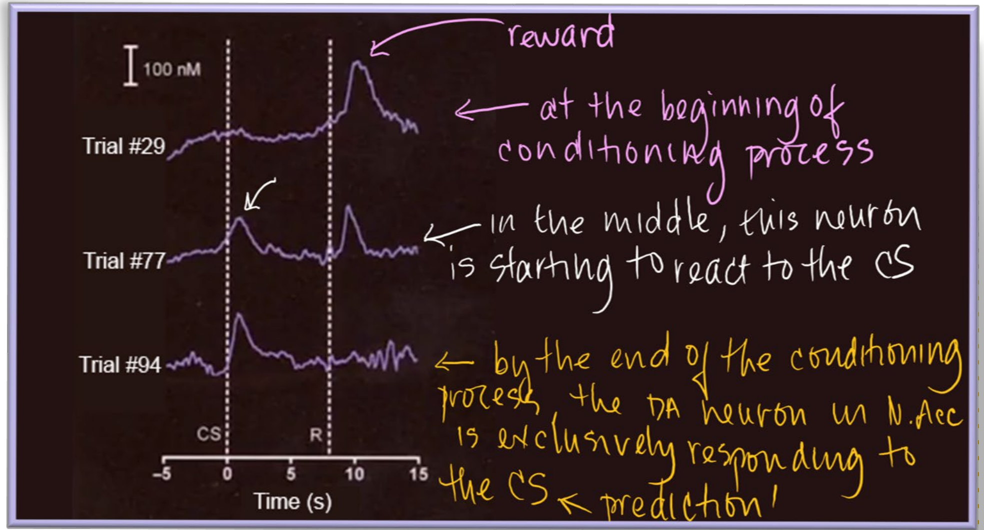
← at the beginning of conditioning process

← in the middle, this neuron is starting to react to the CS

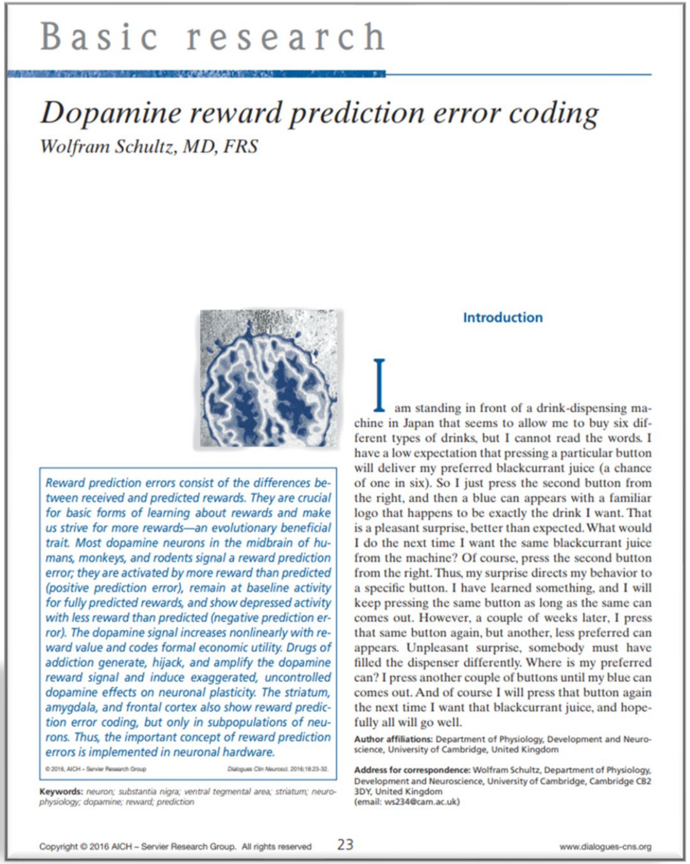
← by the end of the conditioning process, the DA neuron in N.Acc is exclusively responding to the CS ← prediction!

Clark and Phillips (unpublished data)

DA neurons within the N. Accumbens code for **expected rewards**



expected values!



Basic research

Dopamine reward prediction error coding

Wolfram Schultz, MD, FRS

Reward prediction errors consist of the differences between received and predicted rewards. They are crucial for basic forms of learning about rewards and make us strive for more rewards—an evolutionary beneficial trait. Most dopamine neurons in the midbrain of humans, monkeys, and rodents signal a reward prediction error; they are activated by more reward than predicted (positive prediction error), remain at baseline activity for fully predicted rewards, and show depressed activity with less reward than predicted (negative prediction error). The dopamine signal increases nonlinearly with reward value and codes formal economic utility. Drugs of addiction generate, hijack, and amplify the dopamine reward signal and induce exaggerated, uncontrolled dopamine effects on neuronal plasticity. The striatum, amygdala, and frontal cortex also show reward prediction error coding, but only in subpopulations of neurons. Thus, the important concept of reward prediction errors is implemented in neuronal hardware.

 Dialogues Clin Neurosci. 2016;18:23-32.

Keywords: *neuron; substantia nigra; ventral tegmental area; striatum; neurophysiology; dopamine; reward; prediction*

Introduction

I am standing in front of a drink-dispensing machine in Japan that seems to allow me to buy six different types of drinks, but I cannot read the words. I have a low expectation that pressing a particular button will deliver my preferred blackcurrant juice (a chance of one in six). So I just press the second button from the right, and then a blue can appears with a familiar logo that happens to be exactly the drink I want. That is a pleasant surprise, better than expected. What would I do the next time I want the same blackcurrant juice from the machine? Of course, press the second button from the right. Thus, my surprise directs my behavior to a specific button. I have learned something, and I will keep pressing the same button as long as the same can comes out. However, a couple of weeks later, I press that same button again, but another, less preferred can appears. Unpleasant surprise, somebody must have filled the dispenser differently. Where is my preferred can? I press another couple of buttons until my blue can comes out. And of course I will press that button again the next time I want that blackcurrant juice, and hopefully all will go well.

Author affiliations: Department of Physiology, Development and Neuroscience, University of Cambridge, United Kingdom

Address for correspondence: Wolfram Schultz, Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge CB2 3DY, United Kingdom
(email: ws234@cam.ac.uk)

"Reward prediction errors consist of the differences between received and predicted rewards.

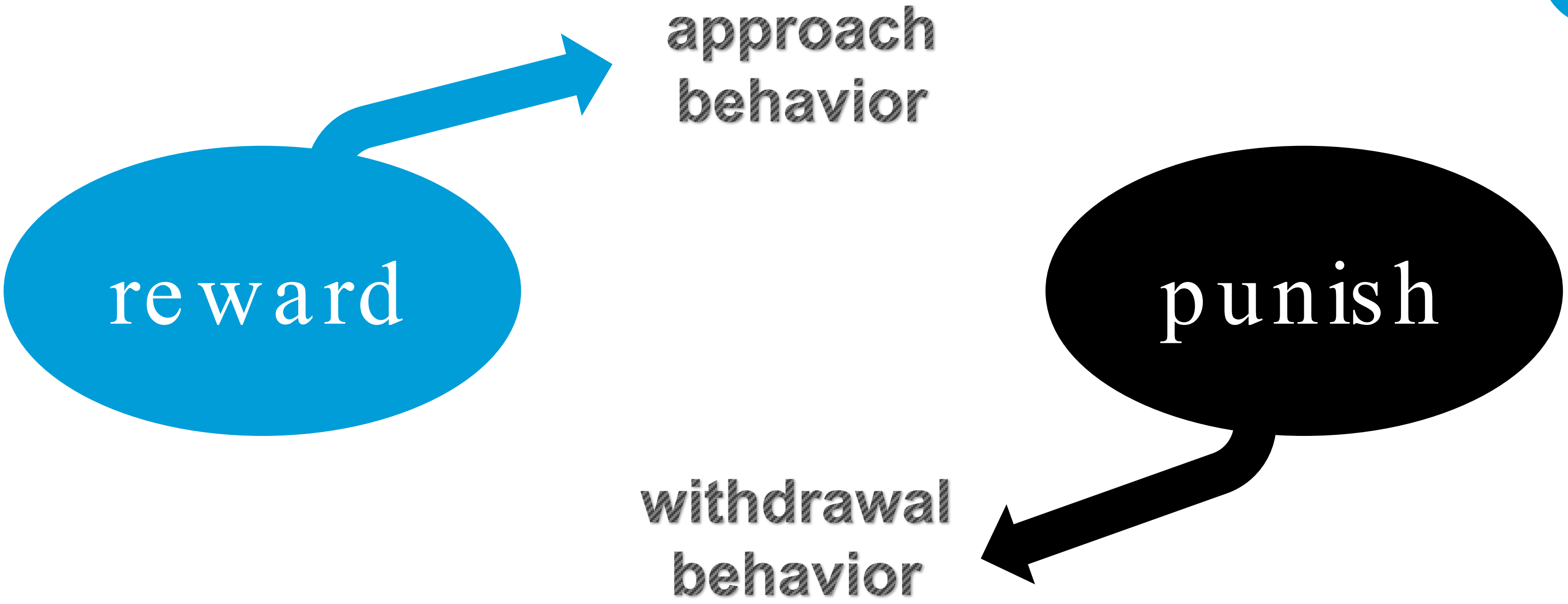
... dopamine ... signal(s) a reward prediction error; they are activated by more reward than predicted (positive prediction error), remain at baseline activity for fully predicted rewards, and show depressed activity with less reward than predicted (negative prediction error)."

Schultz, W. (2016) *Dialogues Clin Neurosci.*;18:23-32.


Ribena
Ribena
Ribena
Ribena
Ribena
Ribena


COLD

TERMS:



reward and punish(ment)

BUT: involves information about the future.

different than what was predicted

prediction error

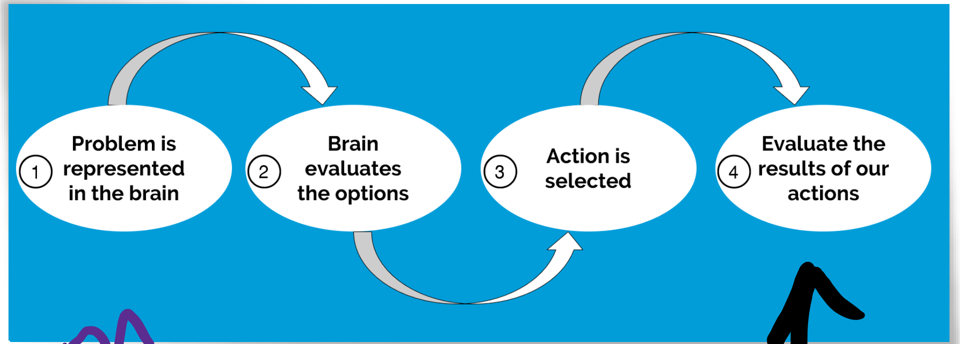
How is the error calculated?

Error = what is currently happening − what was predicted

THE REWARD PREDICTION ERROR IS THEREFORE:

Reward Prediction Error = current reward − predicted reward

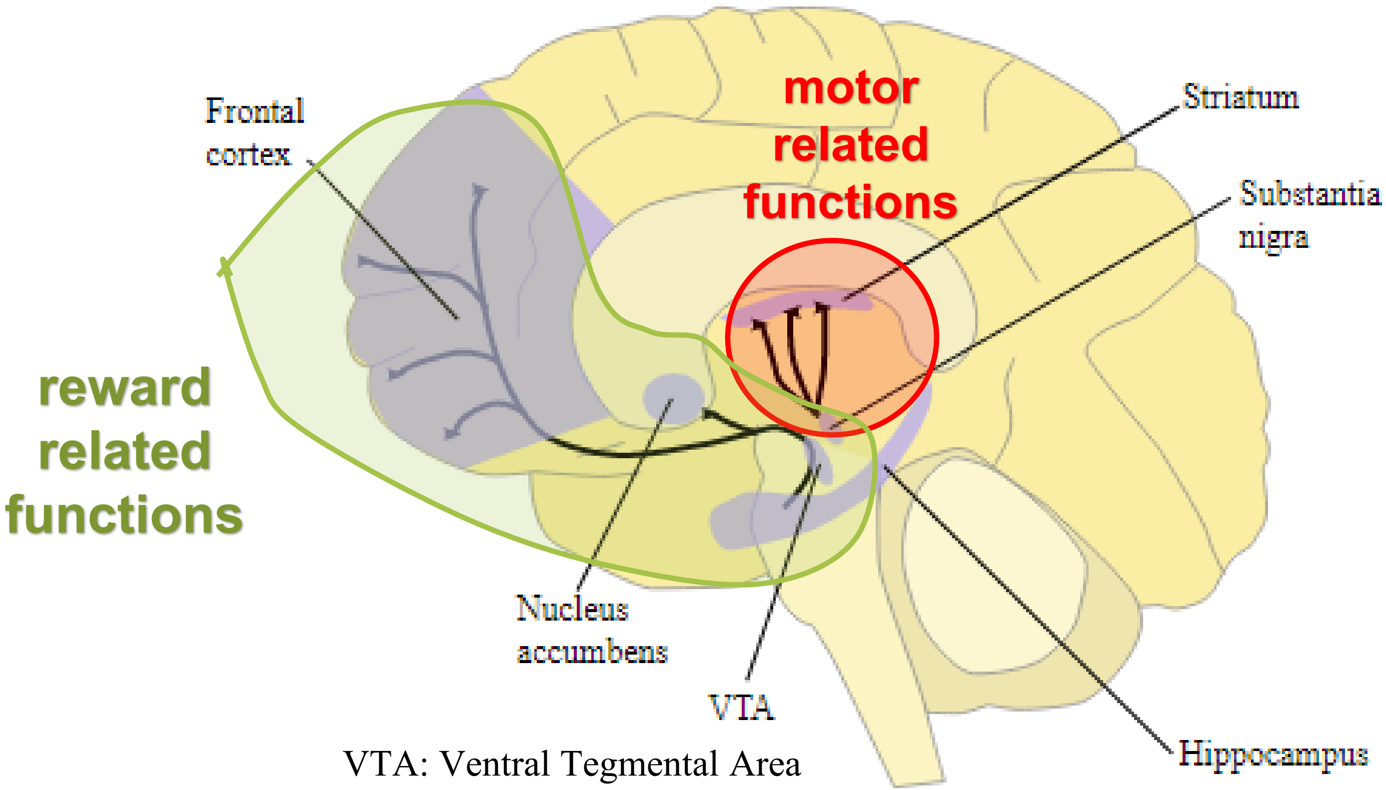
Reward prediction error is the difference between a reward that is being received and the reward that is predicted to be received.



Based on the evaluation Learning occurs

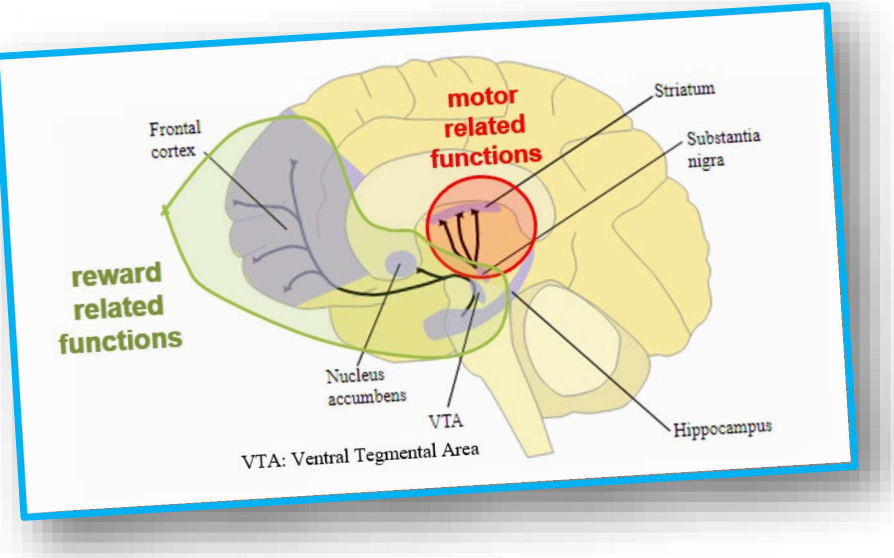
Learning updates the representation, the valuation and the action-selection processes.

prediction Error will improve learning

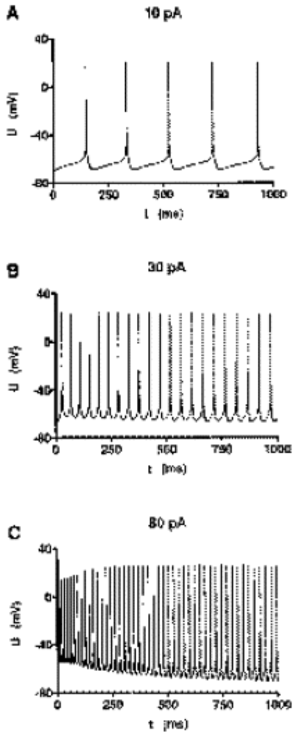


Dopaminergic Pathways

image: wikipedia.org

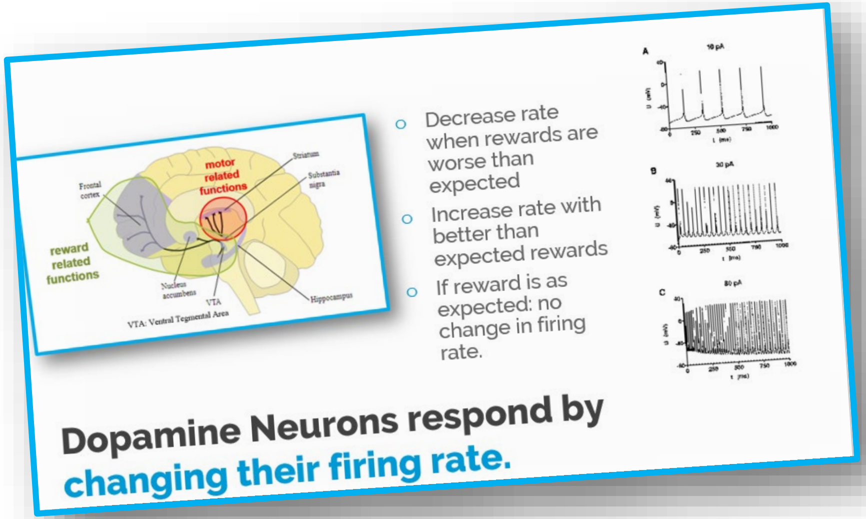


- Decrease rate when rewards are worse than expected
- If reward is as expected: no change in firing rate.
- Increase rate with better than expected rewards



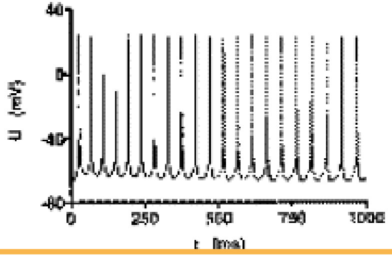
Dopamine Neurons respond by **changing their firing rate.**

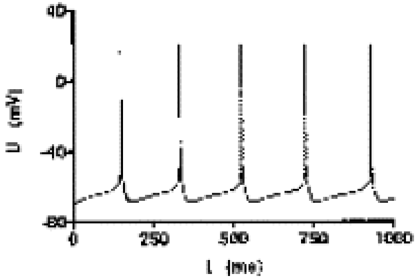
Reward Prediction Error



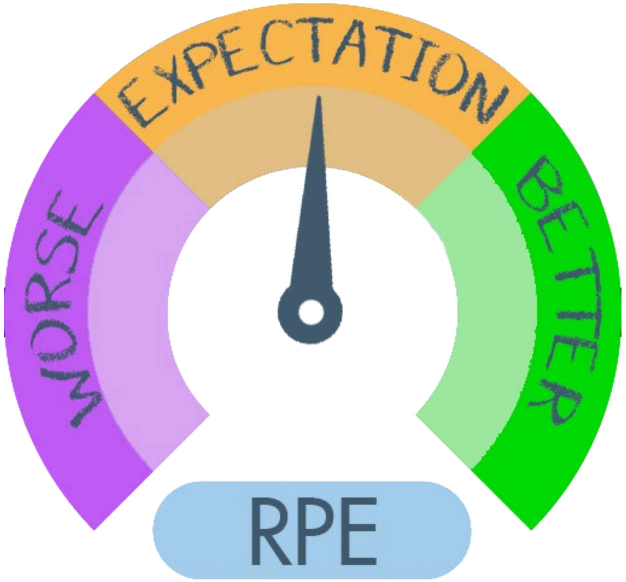
- Brain creates an expectation reference point about the reward received.

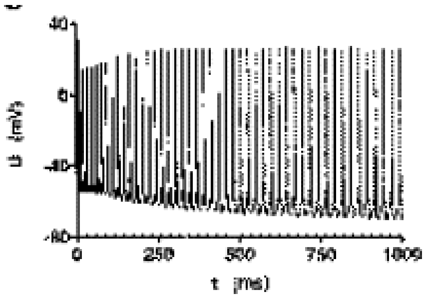
No change in firing rate when reward is exactly as expected





If rewards are worse than expected, then dopamine neurons decrease their firing rate.





When rewards are better than expected, then dopamine neurons increase their firing rate.

Note: DA neurons do no respond to the rewards themselves; they respond to whether a reward was better or worse than the current reference point – the reward prediction error.

Walking along and find
a Scratch & Win ticket.

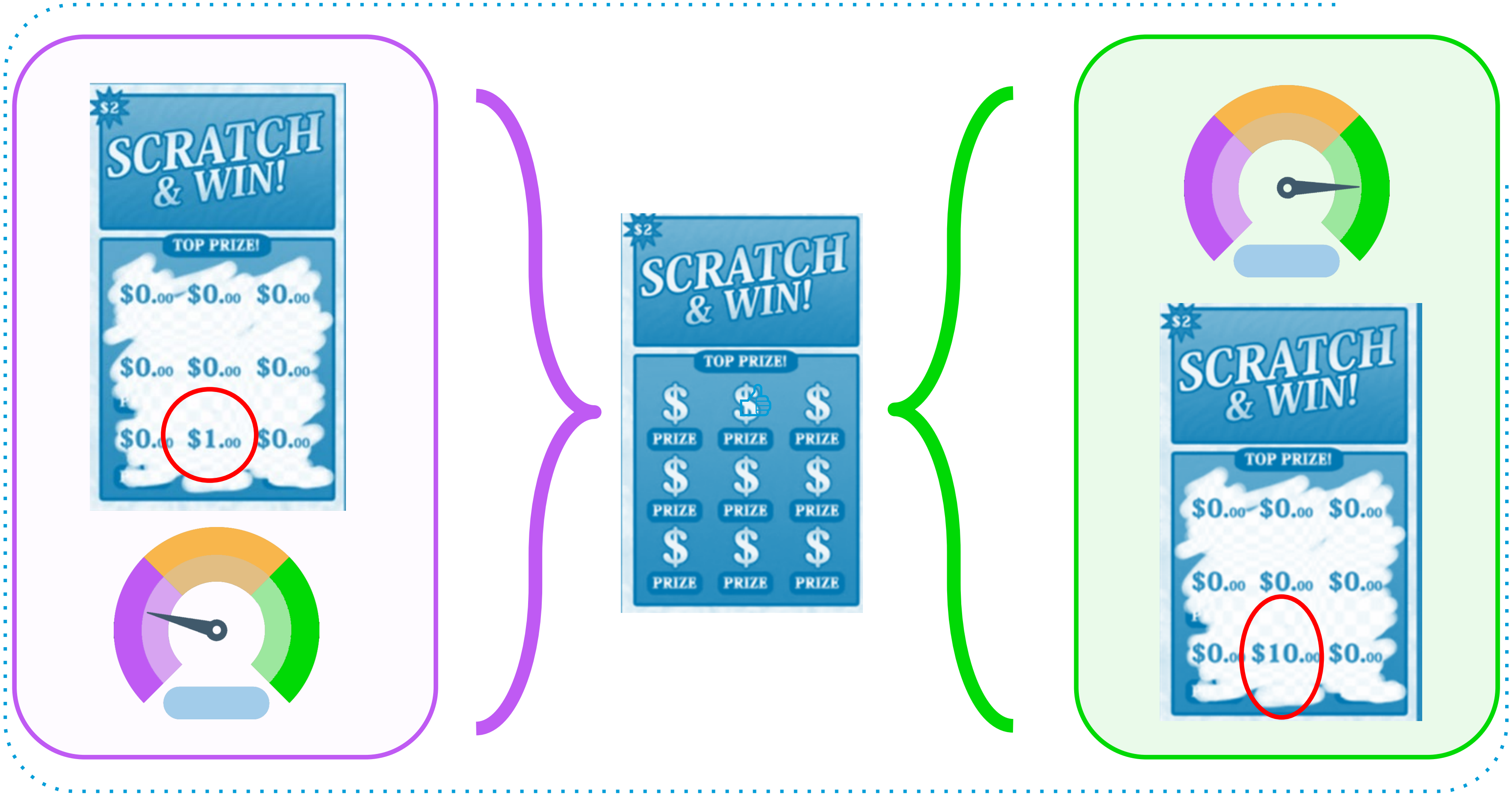
Your DA neurons would be firing because you found the ticket – unexpected reward.

The expected value of the ticket is $2.00 .


$2
SCRATCH & WIN!
TOP PRIZE!
$0.00 $0.00 $0.00
$0.00 $0.00 $0.00
$0.00 $10.00 $0.00


$2
SCRATCH & WIN!
TOP PRIZE!
$ PRIZE $ PRIZE $ PRIZE
$ PRIZE $ PRIZE $ PRIZE
$ PRIZE $ PRIZE $ PRIZE

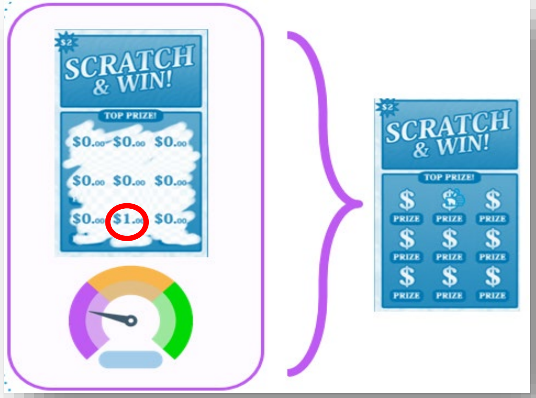

$2
SCRATCH & WIN!
TOP PRIZE!
$0.00 $0.00 $0.00
$0.00 $0.00 $0.00
$0.00 $1.00 $0.00

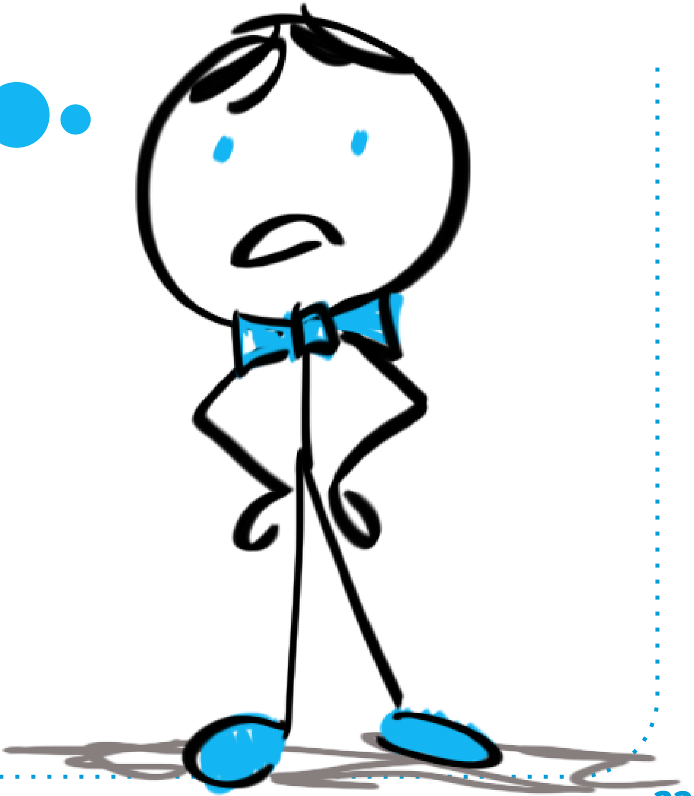


$2
SCRATCH & WIN!
TOP PRIZE!
$0.00 $0.00 $0.00
$0.00 $0.00 $0.00
$0.00 $1.00 $0.00
$2
SCRATCH & WIN!
TOP PRIZE!
PRIZE PRIZE PRIZE
PRIZE PRIZE PRIZE
PRIZE PRIZE PRIZE
$2
SCRATCH & WIN!
TOP PRIZE!
$0.00 $0.00 $0.00
$0.00 $0.00 $0.00
$0.00 $10.00 $0.00

I won money.
That should be a 'reward'.
My dopamine neurons are not firing.... because

$1 is **less** than the expected amount.

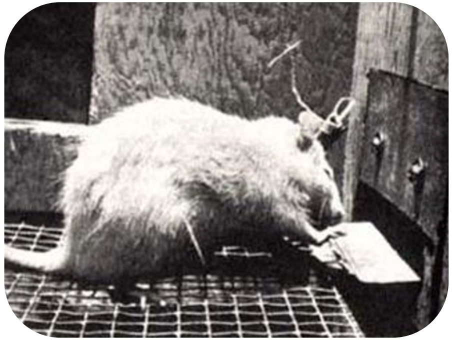




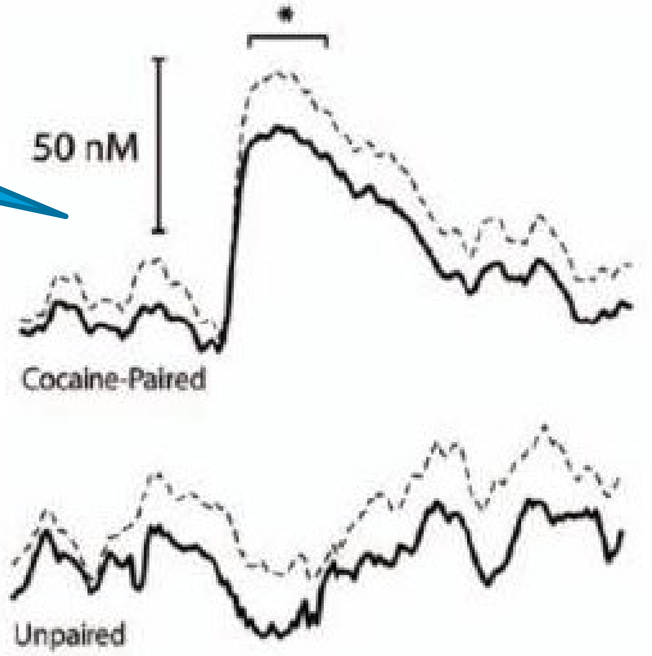
I won money.
My dopamine neurons are firing!!!

$10 is **more** than the expected amount.

Animal is expecting to get cocaine because of the pairing associated with cocaine. There is a strong DA response in the N. Accumbens.

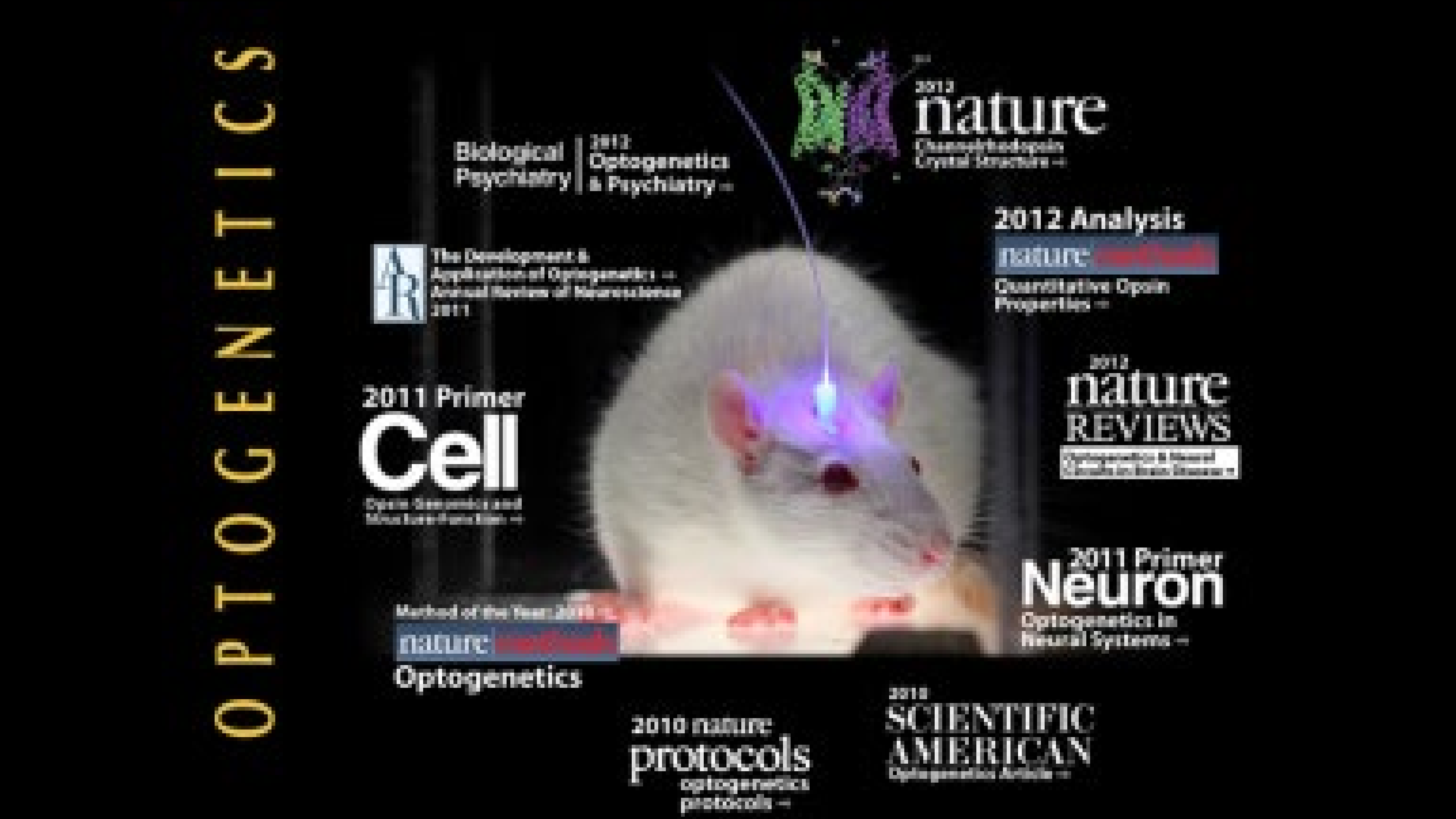


If the stimulus is not associated with cocaine, dopamine is not released.

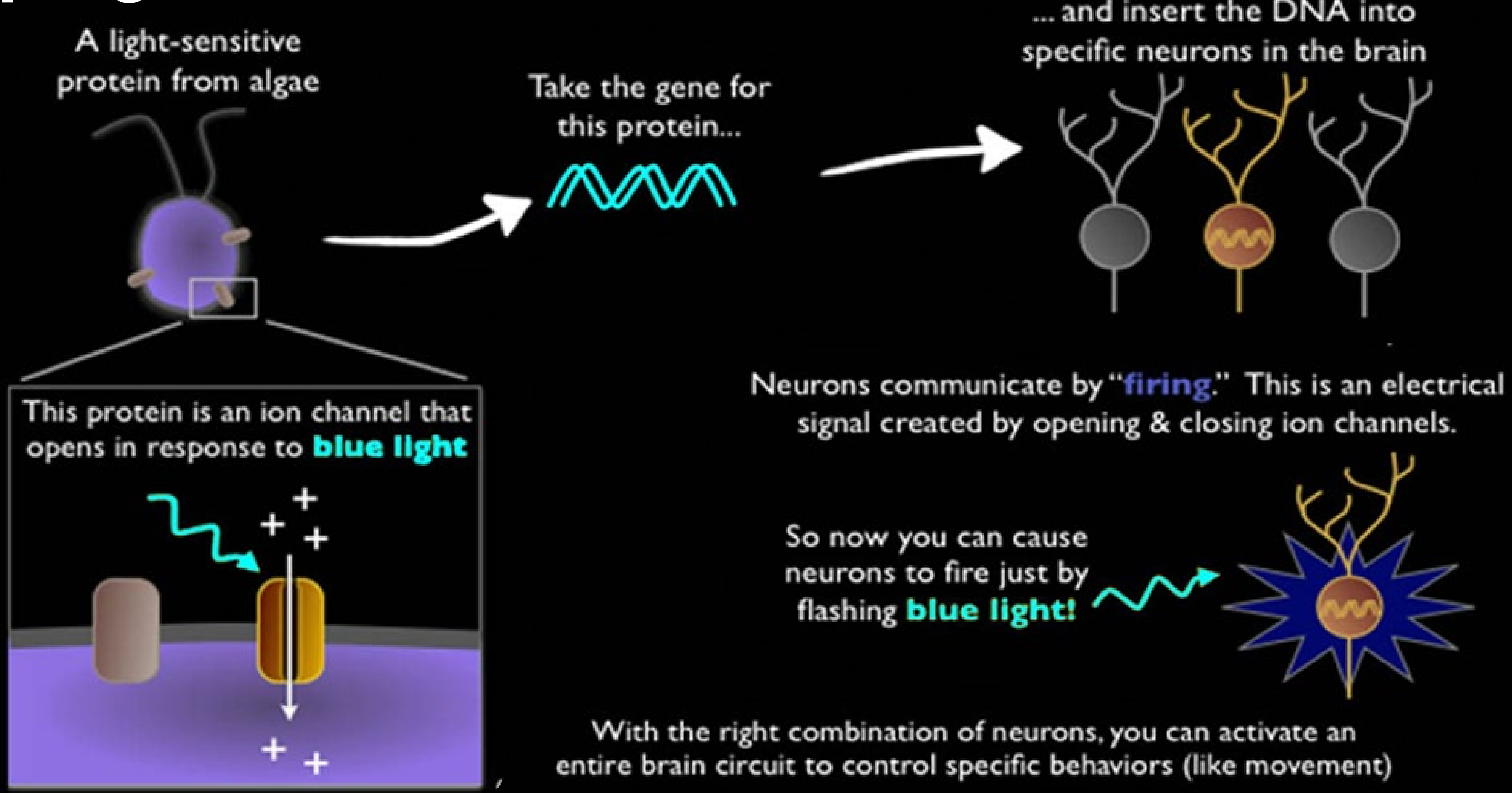


Expectation of reward:
N. Accumbens & Dopamine

Day, J. J. and Carelli, R. M. (2007) *The Neuroscientist*, Volume 13, Number 2


OPTOGENETICS
Biological Psychiatry
2012 Optogenetics & Psychiatry
2012 nature
Channelrhodopsin Crystal Structure
The Development & Application of Optogenetics
Annual Review of Neuroscience 2011
2012 Analysis
nature methods
Quantitative Opsin Properties
2011 Primer
Cell
Opsin Genomics and Structure-Function
2012 nature REVIEWS
2011 Primer
Neuron
Optogenetics in Neural Systems
Method of the Year 2010
nature methods
Optogenetics
2010 nature protocols
optogenetics protocols
2010 SCIENTIFIC AMERICAN
Optogenetics Article

Optogenetics:



https://gadgtecs.com/2018/04/03/researchers-can-now-mind-control-mice-using-optogenetics/

TIME

How Winning Can Mean Losing in Poker and Life

BY Jeffrey Kluger Wednesday, Jan. 13, 2010

You can learn a lot about gambling if you're willing to analyze 27 million hands of online poker. Don't have time for that? No worries; sociology doctoral student Kyle Siler of Cornell University has done it for you. His counterintuitive message: the more hands you win, the more money you're likely to lose — and this has implications that go well beyond a hand of cards.

Karen Bleier / AFP / Getty

Siler, whose work was published in December in the online edition of the *Journal of Gambling Studies* and will appear later this year in the print edition, was not interested in poker alone but in the larger idea of how humans handle risk, reward and variable payoffs. Few things offer a better way of quantifying that than gambling — and few gambling dens offer a richer pool of data than the Internet, where millions of people can play at once and transactions are easy to observe and record.

To gather his data, Siler used a software system called PokerTracker and directed it to collect and collate information on small- medium- and large-stakes games. He limited the games to no-limit Texas Hold 'Em with six players in order to eliminate at least some extraneous variables. It was in the course of crunching all that information that he found the strangely inverse relationship between the number of hands won and the amount of money lost. He also noticed that it was novice players who lost the most.

The reason for the paradoxical results was straightforward enough: the majority of the wins the players tallied were for relatively small stakes. But the longer they played — and the more confident they got — the likelier they were to get blown out on one or a few very big hands. Win a dozen $50 pots and you're still going to wind up far behind if you lose a single $1,000 one. "People overweigh their frequent small gains vis-à-vis occasional large losses," Siler says.

Small-stakes players also tend to do better with small-denomination cards. A pair of jacks may easily beat a pair of fours, but people who don't gamble much tend to win more with the fours — or with any cards from twos to sevens. That's because the cards' modest numerical worth is easy to understand: they're valuable but not that valuable. When you get into the more rarefied air of eights to aces, you may start losing perspective and putting up more money. "Small pairs have a less ambiguous value," Siler says.

So what does this have to do with you if you don't gamble? It's the wrong question because, actually, you do. Investing, driving, buying a house and merely crossing the street are all acts that involve discernible risks and uncertain rewards. The more small returns you get from your small investments in stocks, the likelier you are to make — and lose — a big investment. The more times you get behind the wheel and speed a little bit, the likelier you are to speed a lot — with deadlier consequences.

"These kinds of calculations are made every day," says Siler. "Adultery is another good example. People get away with it countless times but they get caught just once and they lose everything."

And unlike the risks at the poker table, where your losses are just yours, in the larger world, you can take down a lot of other people with you. "Organizational malfeasance in general depends on this kind of risk analysis," says Siler. "Look at a place like Enron. People took a lot of small chances and won, then took big chances and lost big." Indeed, Siler points out, during the recent financial crisis, an entire nation — Iceland — went bankrupt in a similar way, trusting high-risk, high-reward investments that quit paying off.

While walking away from the poker table can be easy, walking away from life — and all the risks and rewards it presents you — is not an option. But in both venues, the rule should be the same: gamble only what you can afford to lose — and know when you're approaching those stakes.

http://content.time.com/time/health/article/0,8599,1953205,00.html

J Gambl Stud (2010) 26:401–420
DOI 10.1007/s10899-009-9168-2

ORIGINAL PAPER

Social and Psychological Challenges of Poker

Kyle Siler

Published online: 25 December 2009

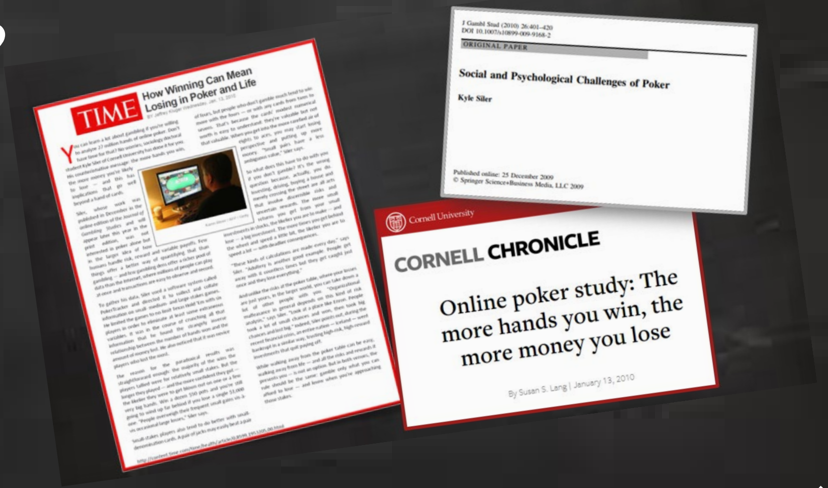
Cornell University

CORNELL CHRONICLE

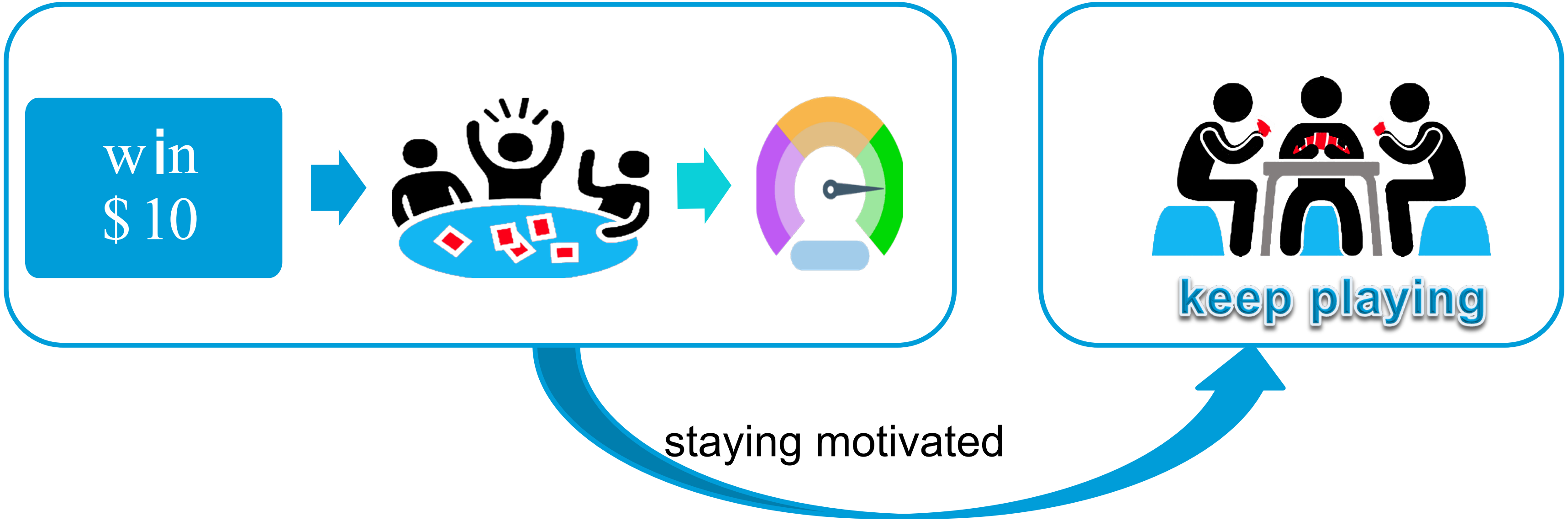
Online poker study: The more hands you win, the more money you lose

By Susan S. Lang | January 13, 2010

How is it that the poker players who won most often were also the ones who lost the most money?



reference dependence, poker & dopamine



Dopamine cares about wins and losses – not about the amount!

losing a large amount of money

for example: $1,000 loss

brain codes event as bad

but it does not appear to be 100x worse than the good events.

Dopamine neurons are sensitive to whether you win or lose – but less sensitive to the amount!