Wanting Things: Action Selection and Addiction

How Your Brain Works

Prof. Jan Schnupp
wschnupp@cityu.edu.hk

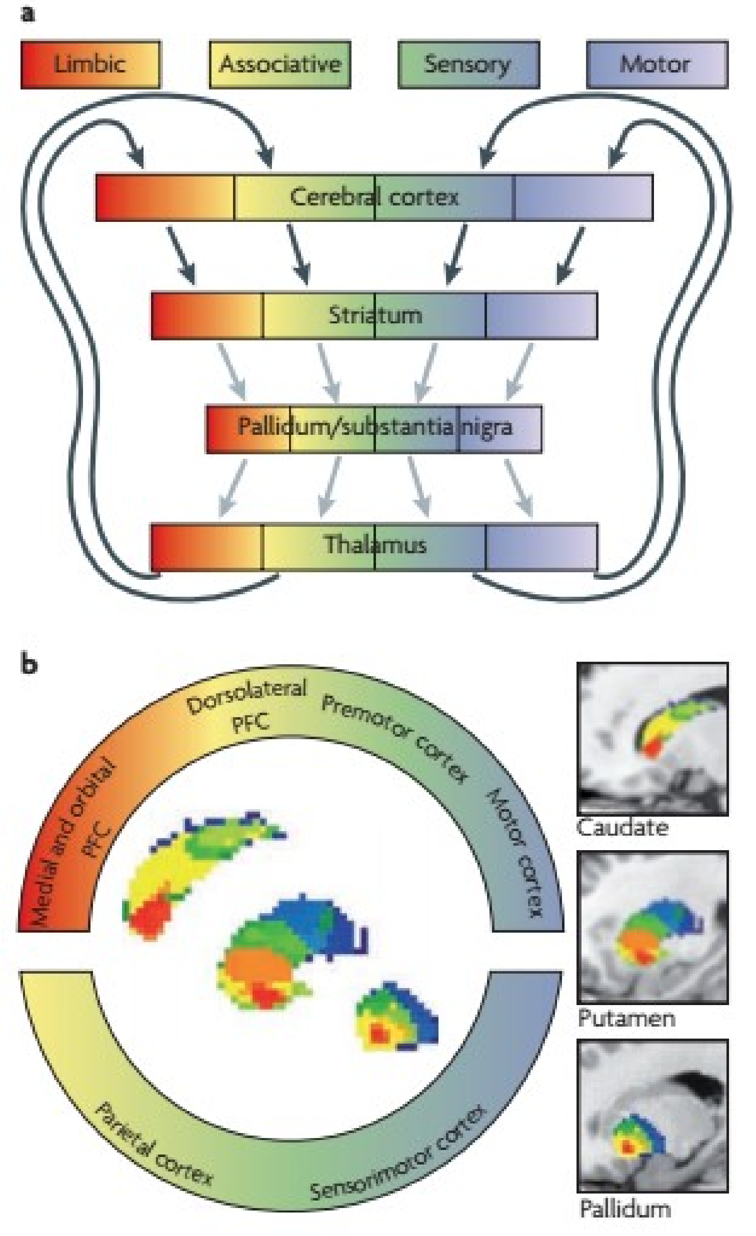
HowYourBrainWorks.net

香港城市大學
City University of Hong Kong

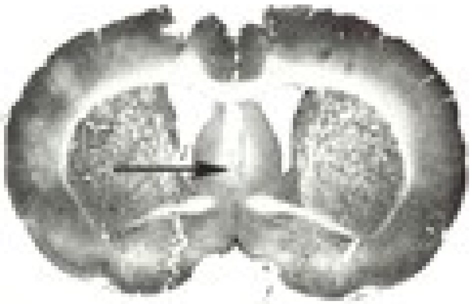
Basal Ganglia Loops

- To recap from the last lecture, loops from motor cortex through the basal ganglia regulate level of motor activity in a dopamine dependent way.
- More anterior to the motor cortex, prefrontal brain structures which are involved in “cognitive” and “limbic” processing or motor intentions, and these form analogous loops through the basal ganglia which are also modulated by dopaminergic input

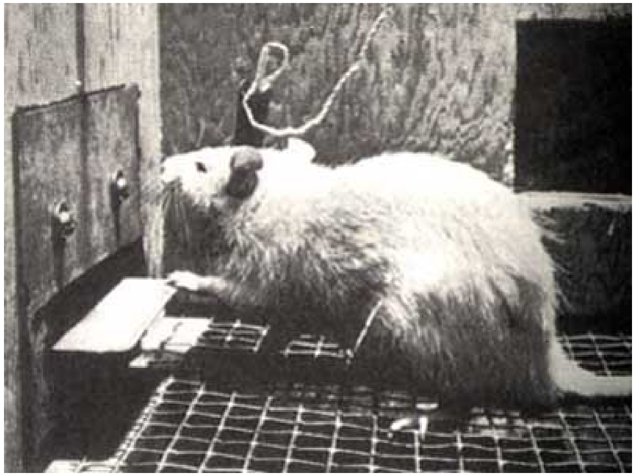
- Source: Redgrave et al 2010, Nature Neuroscience

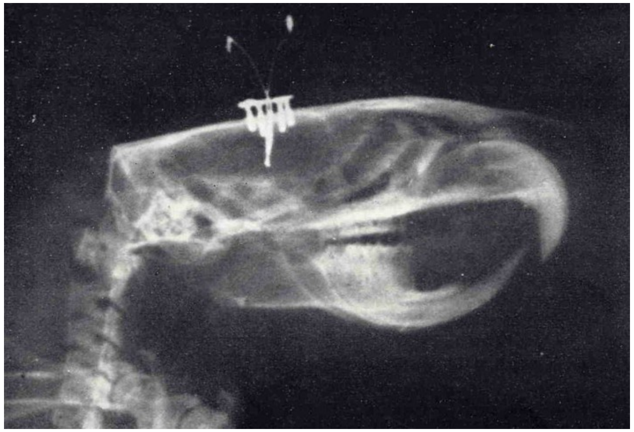


Self stimulation experiments

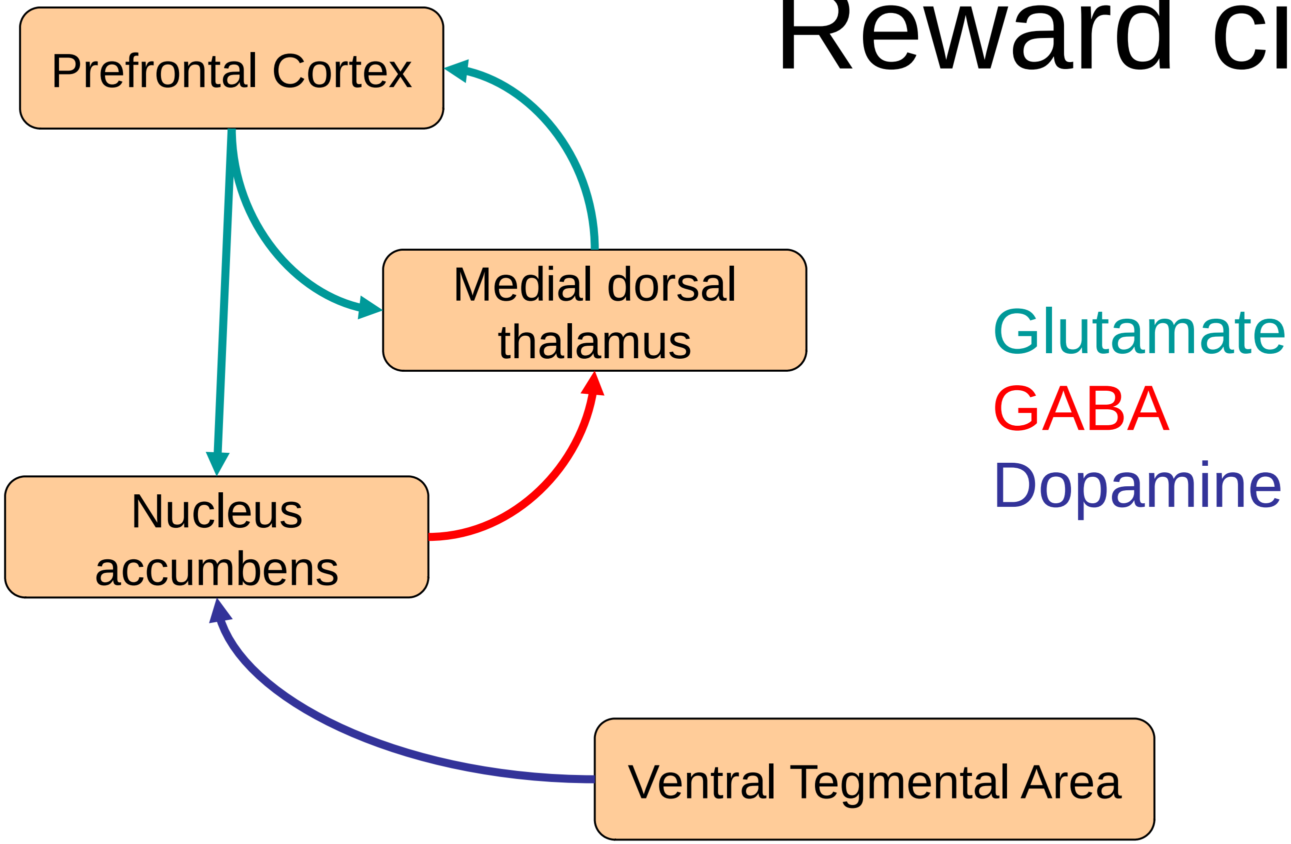


- First done by Olds & Milner in rats in 1954. They tested a wide range of stimulation sites but found that rats really only like self stimulation of the “septal” region.
- Rats may self-stimulate these areas until they are completely exhausted.
- It is now thought that “septal region” is an effective self-stimulation site because it will activate dopaminergic fibers of the medial forebrain bundle which travel from the ventral tegmental area (VTA) to the nucleus accumbens.



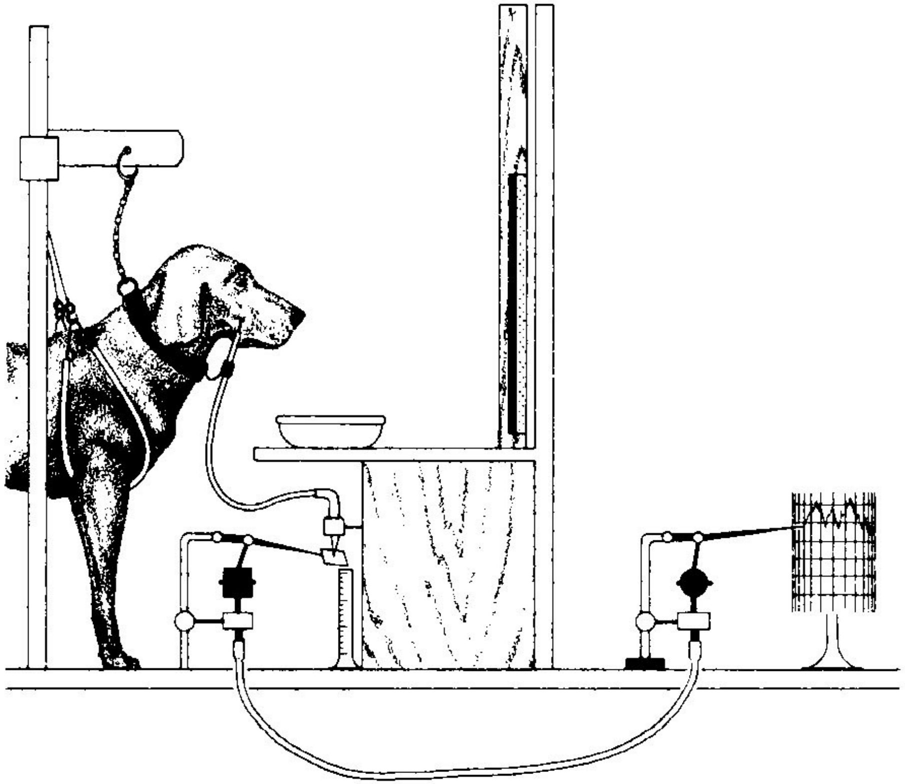


Reward circuitry



- The Nucleus Accumbens is considered part of the ventral striatum. Think of it perhaps as a “limbic” extension of putamen and caudate.

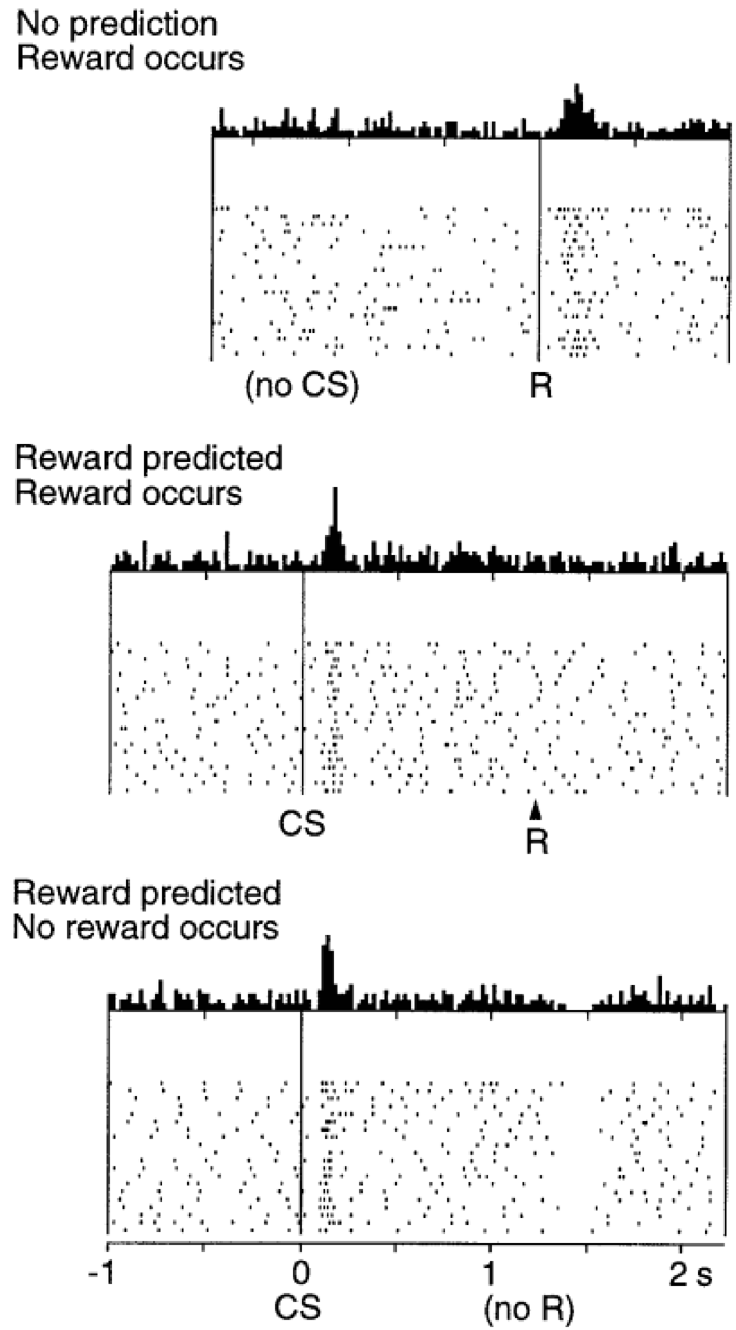
Pavlov's Dog



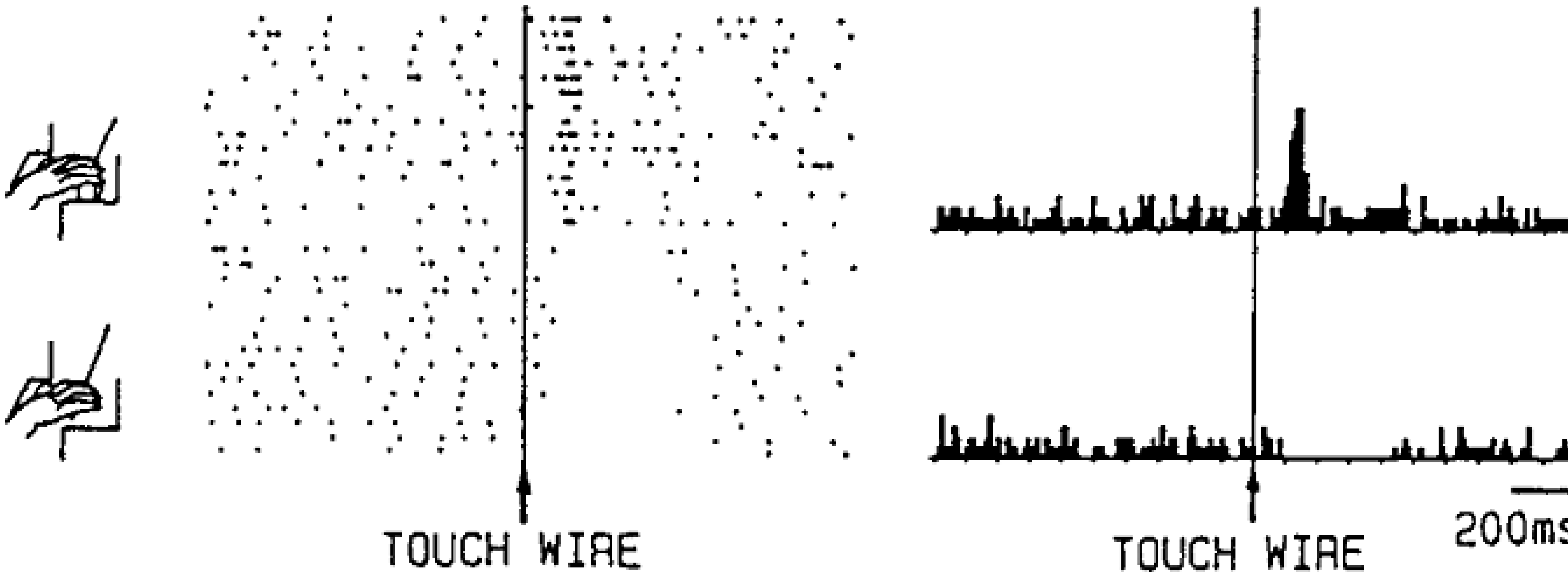
- Ivan Pavlov, Nobel prize winner in 1904, first to describe "conditioned reflexes".
- "**Unconditioned stimuli" (US)** can produce automatic responses (e.g. food -> salivating). "**Conditioned stimuli"** (**CS**, e.g. sound of the dinner bell) can produce the same responses after a period of **reinforcement learning**.

Reward prediction

- Recordings of a dopaminergic neuron by Schultz (1998) in a monkey
  - receiving unexpected rewards (top),
  - predicting and receiving rewards based on a “conditioned stimulus” (CS, middle),
  - or failing to receive predicted rewards (bottom). Note the reduced firing rate when the predicted reward did not happen.
- Phasic dopamine is a “reward prediction error signal”

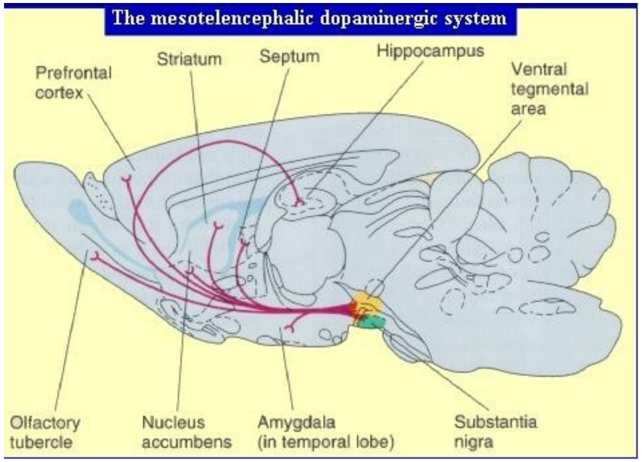


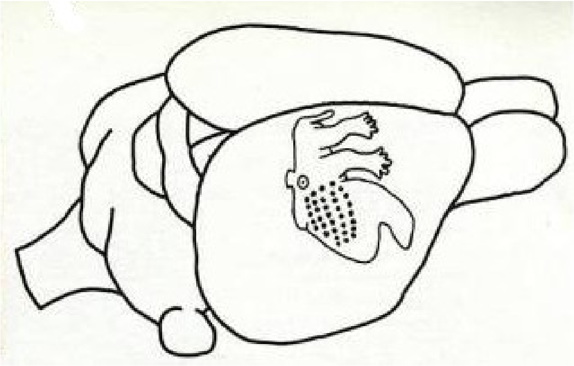
Rewarded and Unrewarded Trials



- Recordings of dopaminergic neuron by Romo & Schultz (1990). A monkey was trained to slide his fingers along a wire placed out of sight to check whether or not a food item was placed at the end of it.

Dopamine Systems and John Chapin's Robotic Rats





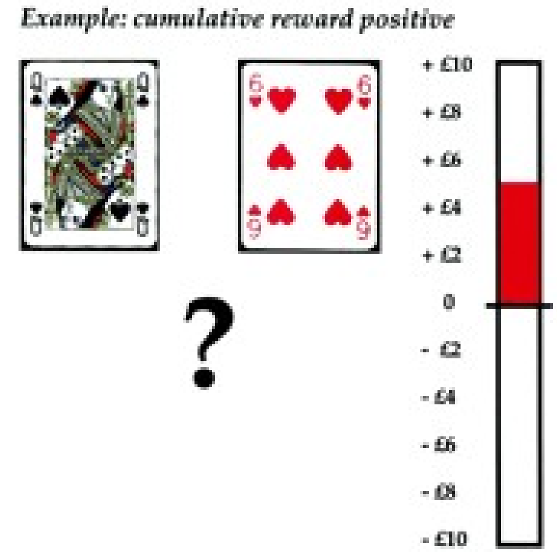
The Chapin lab remote controlled rat

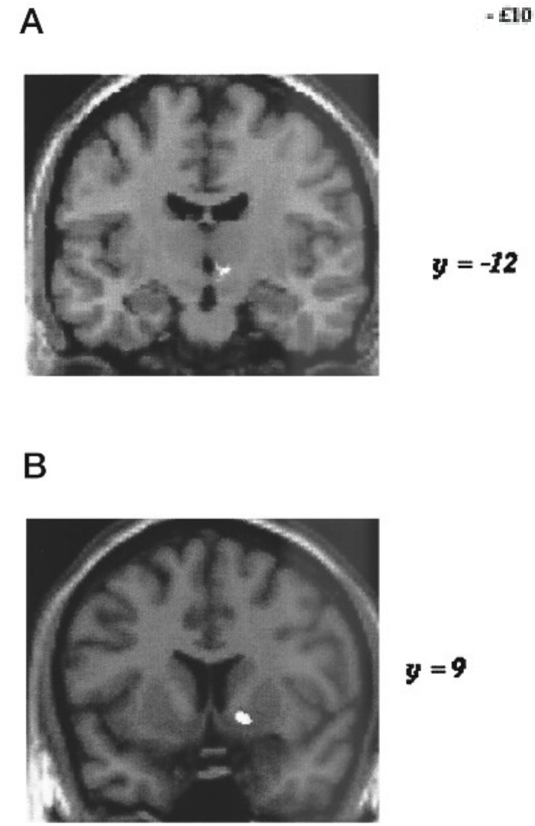
- Has one stimulating electrode implanted to stimulate the medial forebrain bundle (MFB which contains dopaminergic output fibers from the VTA) and one stimulating electrode in each somatosensory cortex.
- The somatosensory cortex electrode tells the animal which direction to turn. The MFB electrode tells the animal to predict a reward if it follows the instructions from the somatosensory electrodes.
- https://www.nature.com/articles/417037a

Reward Circuits in People

- One piece of evidence that dopaminergic reward prediction operates in humans too: Volunteers were invited to play a card guessing game for money while their brains were scanned with fMRI.
- Winning money was associated with increased activity in SN/VTA region (A) and the ventral striatum (B).

- Elliot, Friston and Dolan, J Neurosci 2000



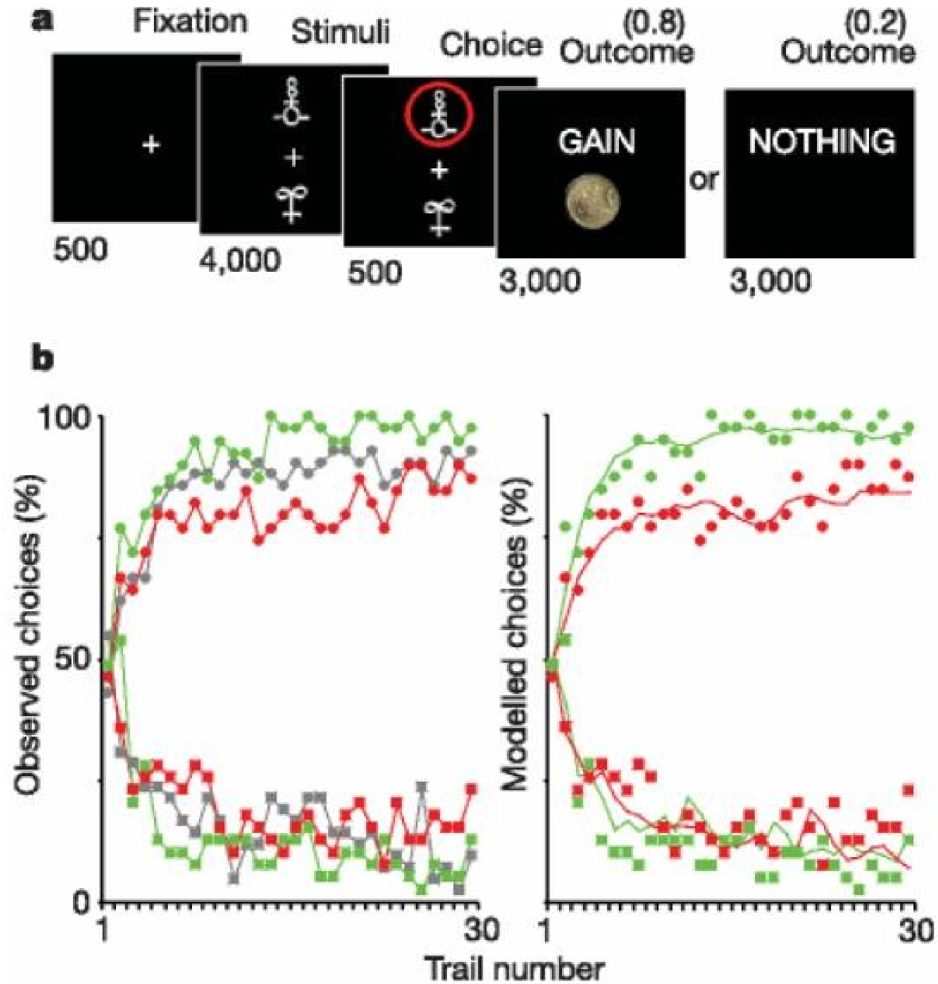


Dopamine and Synaptic Plasticity

- Cortico-striatal glutamatergic synapses are highly plastic, and can undergo long term depression (LTD) or potentiation (LTP).
- Dopaminergic input regulates this plasticity.
- The molecular details are complicated (see Calabresi et al Trends in Neuroscience 2007)

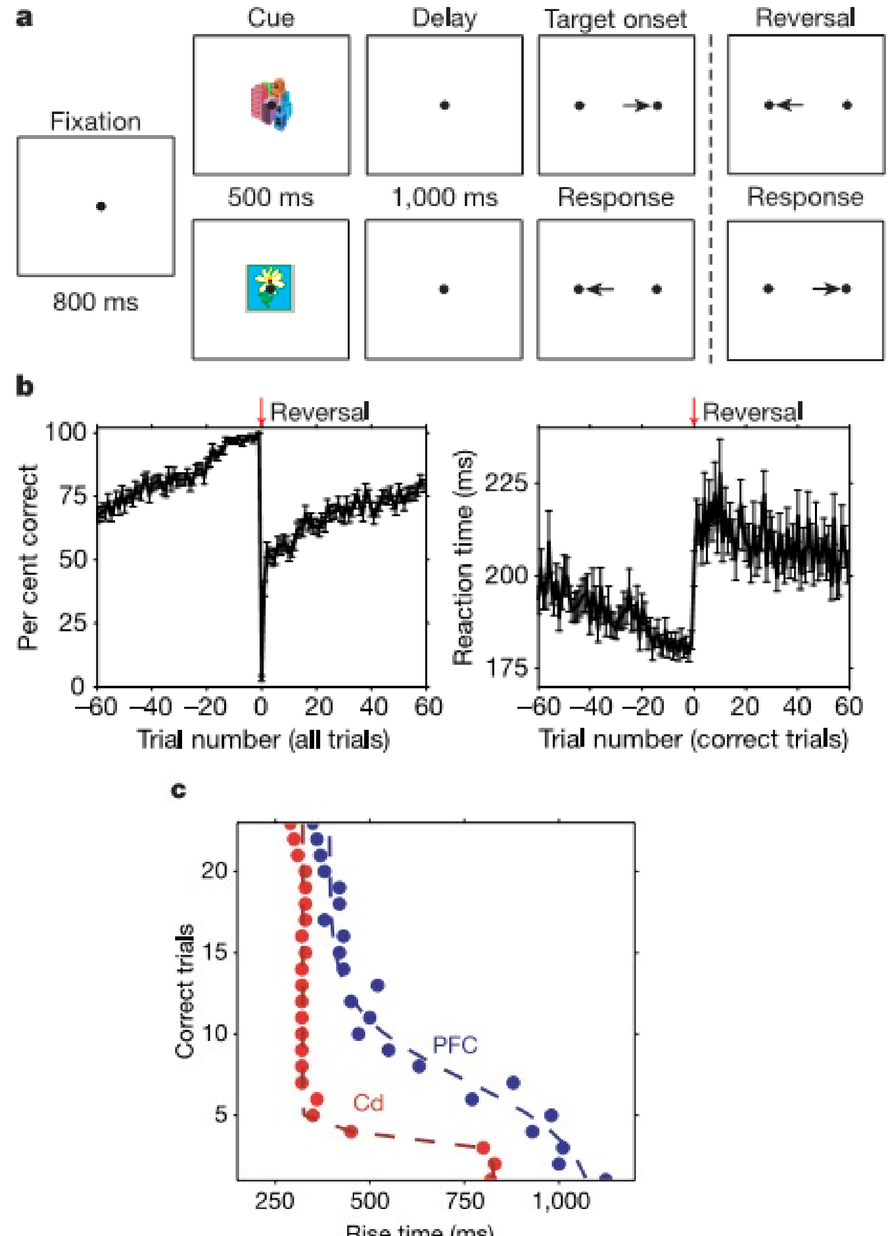
Manipulating the size of DA reward prediction signals affects learning rates

- Subjects learn from experience that one visual stimulus is associated with more likely reward than another.
- They learn faster when given L-DOPA (green), slower when given halperidol (red). Gray is placebo control.
- Activity levels in ventral striatum and posterior putamen reflect reward prediction errors.
- *Pessiglione et al*, Nature 442 (2006)

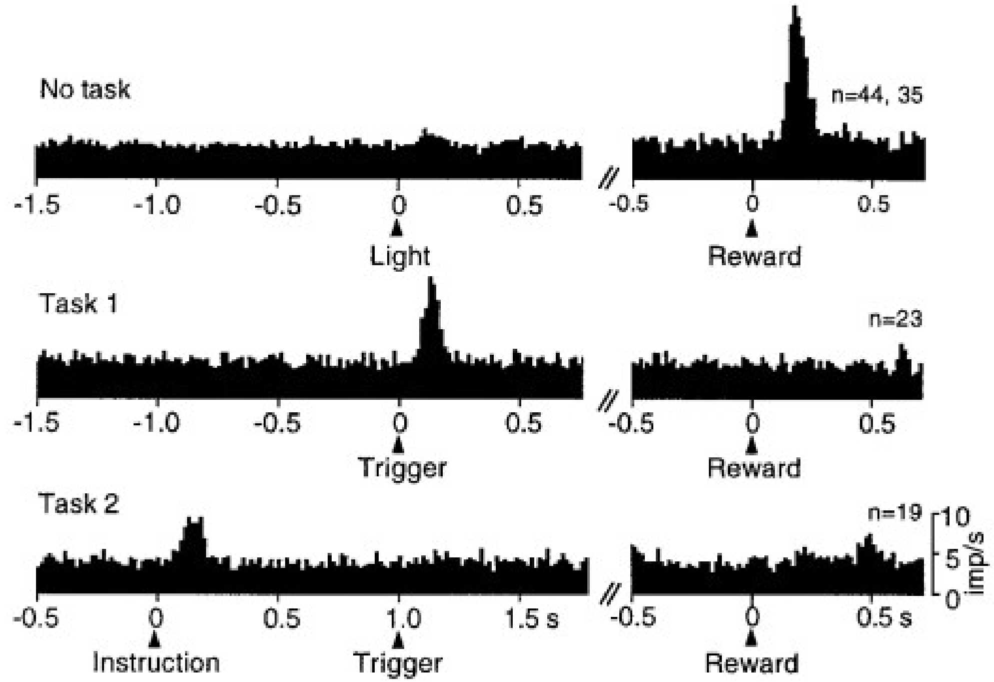


“Reversal Learning” and the PFC – basal ganglia network

- Pasupathy & Miller, Nature 2005, trained monkeys to move their eyes from a fixation spot to one of two possible targets. Cue images informed the monkeys about which of the two targets would be rewarded.
- Once the animals had learned the correct association of image to target direction, the association was suddenly reversed (i.e. the flower picture which used to cue for left now cued for right) and the animals had to relearn the association.
- Decoding of neural activity in Caudate and Prefrontal Cortex showed that Caudate neurons learned the new association much quicker than cortex, after only 5 or 6 trials.



Dopamine response transfer to earliest predictive stimulus



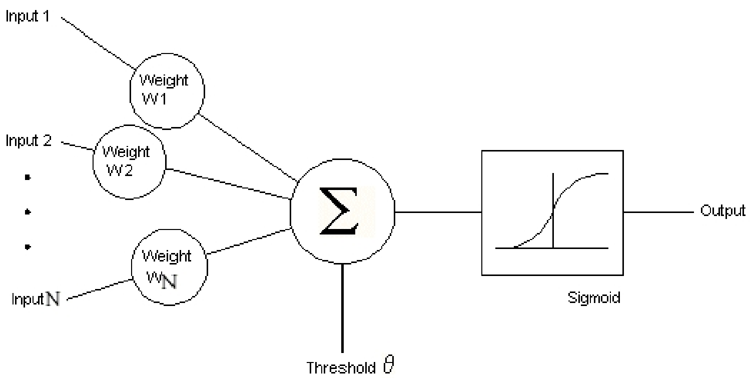
- Top: outside of any behavioral task there is no response in neurons tested with a small light, but response occurs to a drop of liquid delivered at a spout in front of the thirsty animal's mouth.
- Middle: response to a reward-predicting those following reward-predicting trigger stimulus in a 2-choice spatial reaching task, but absence of response to reward delivered during established task
- Bottom: response to an instruction cue preceding the reward-predicting trigger stimulus by a fixed interval of 1 s

Temporal Difference (TD) Learning

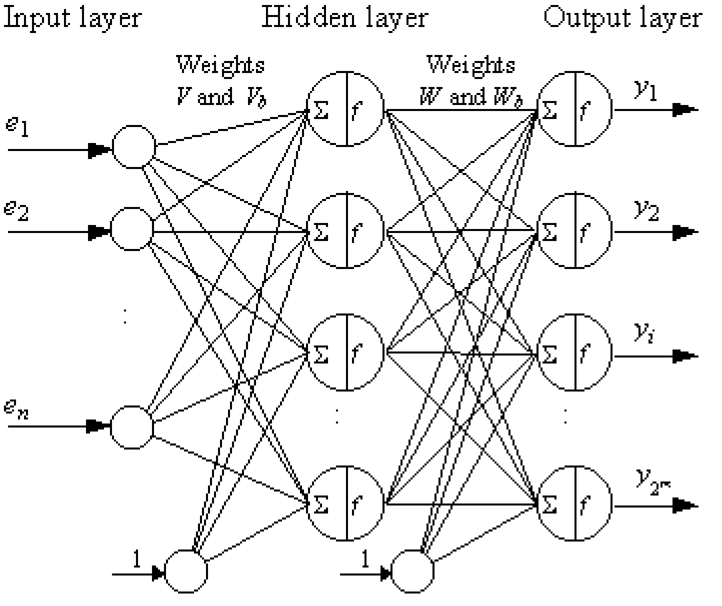
A clever dog might ask itself:

- Which is the best predictor of an upcoming meal: the sound of the bell or the sound of the footsteps of an approaching carer?
- Which of my actions makes the carer more likely to approach: barking or tail wagging?
- What about tail wagging when the carer is looking vs tail wagging when he isn’t?

Networks of Artificial Model Neurons



$Y=f(\Sigma\, x_i \cdot w_i)$



- Multi-layer perceptrons and error back-propagation

TD-Gammon

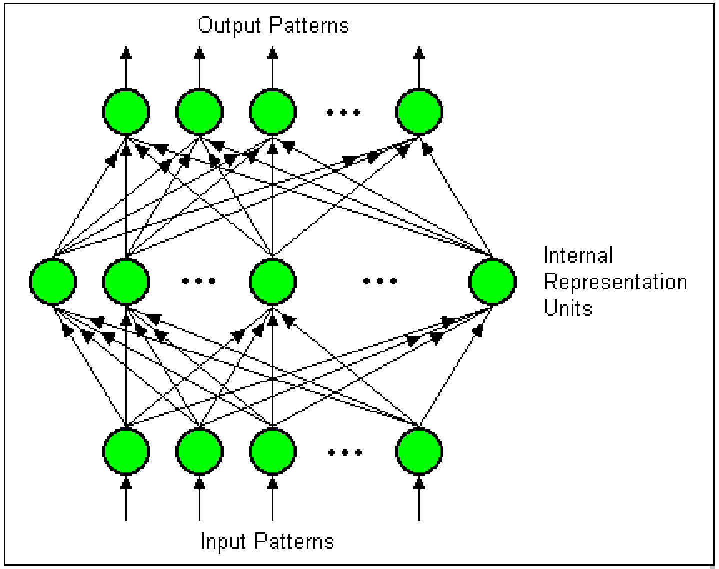


Figure 1. An illustration of the multilayer perception architecture used in TD-Gammon's neural network. This architecture is also used in the popular backpropagation learning procedure. Figure reproduced from [9].

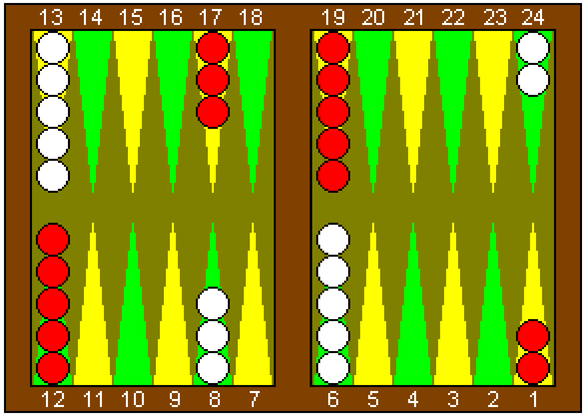
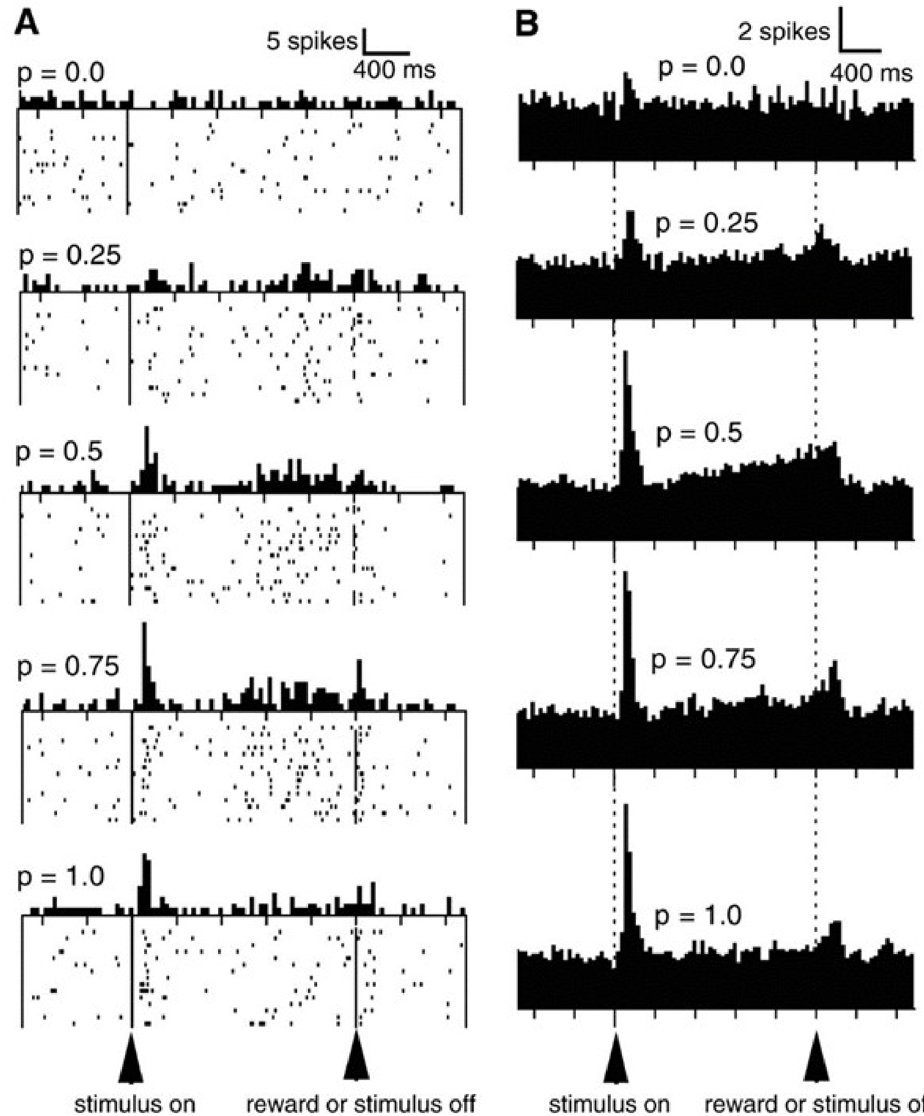


Figure 2. An illustration of the normal opening position in backgammon. TD-Gammon has sparked a near-universal conversion in the way experts play certain opening rolls. For example, with an opening roll of 4-1, most players have now switched from the traditional move of 13-9, 6-5, to TD-Gammon's preference, 13-9, 24-23. TD-Gammon's analysis is given in Table 2.

- An artificial neural network with 198 “input neurons”, (number of white or black pieces on various board positions and home positions, and whether white or black plays next), ca 40 “hidden” neurons and 4 “output” neurons.
- Trained to play backgammon at master level by adjusting connections between artificial neurons depending on changes in predicted outcome as games unfold (temporal difference learning).

Reward Neurons may respond more strongly to uncertain rewards

- Monkeys underwent Pavlovian conditioning to expect water reward when a light stimulus is switched off, but in different sessions the probability of a reward at the end varied from always (p=1) through 50/50 (p=0.5) to never (p=0).
- When reward was unpredictable, dopamine neurons ramped up their activity “in expectation”.
- From: Fiorillo et al, Science 2003



Break

Let’s recap

- DA neurons fire tonically to set overall activity levels, and phasically to signal “reward prediction errors”. That means DA levels spike when the environment is more rewarding than predicted and they drop briefly when the environment is less rewarding than expected.
- Triggering phasic DA activity spikes with electrical stimulation creates artificial “reward predictions” that can manipulate an animal’s choices (roborat)
- Phasic dopamine also enables plastic changes in the cortex – basal ganglia loops, presumably to implement reinforcement learning of stimulus-action associations that are deemed predictive of rewarding outcomes.
- The system is set up to try to predict rewarding outcomes as far ahead if possible, and to be robust in the face of uncertainty.
- The cortex-basal ganglia circuits are thus designed to engrain habits of action selection which, according to our experiences, increase our chances of “rewarding” outcomes.

Drug Self-administration experiments

- Rats will self-inject of amphetamine into the Nucleus Accumbens.
- D-amphetamine stimulates DA release by messing with transporter proteins in dopaminergic terminals of afferents from the VTA.

Addictive drugs

- **Cocaine** (“crack” cocaine in particular): potentiates DA action by blocking DA reuptake
- **Amphetamine** (“speed”) and derivatives releases DA (and noradrenaline) by interfering with cell internal catecholamine transporters
- **Opiates**: dopaminergric neurons in the brainstem and midbrain have opiate receptors.
- Striatal neurons carry cholinergic receptors that are activated by **nicotine**.
- These four substances are chemically quite different, but they **all activate dopamine sensitive neurons in the striatum**.

Addiction Potential

- Why does smoking cocaine bring a much greater addiction risk than drinking cocaine in tea?
- Speed of absorption matters hugely. Only very rapidly absorbed drugs can give a phasic DA spike.

Dopamine Dysregulation Syndrome (DDS)

- Parkinsonian patients receiving prolonged courses of dopamine-enhancing medication are at some risk of addiction-like side effects including:
- Craving
- Gambling
- Hypersexuality
- Hypomania
- “Punding” (stereotypic behaviours)

Developing Addiction

- Rats can be trained to self-administer virtually all addictive drugs, including opiates, amphetamines, cocaine, alcohol and nicotine.
- They usually need repeated exposure / prolonged access to intravenous drugs to develop an addiction.
- Addictions are then marked by:
  - Escalation in dose
  - Habituation / Tolerance, and finally
  - “Resistance to extinction” (i.e. behaviours persist even if they no longer deliver drugs, but instead deliver punishment such as foot-shock)

Recreational Drug Use vs Addiction

- As long as potentially addictive drugs “only” reinforce behaviour patterns to seek and take substances which are experienced as pleasant or rewarding they remain just one desirable thing among many. For disruptive addictions to emerge, additional things must occur.
- We have already mentioned “tolerance”. A homeostatic mechanism: prolonged exposure to dopaminergic inputs makes the brain less sensitive to such inputs. Greater doses are needed. At the same time, other (healthy) positive reinforcers may become less effective.
- Withdrawal/cravings: not having access to the drug is perceived as stressful. Stress hormone levels rise in response to withdrawal.

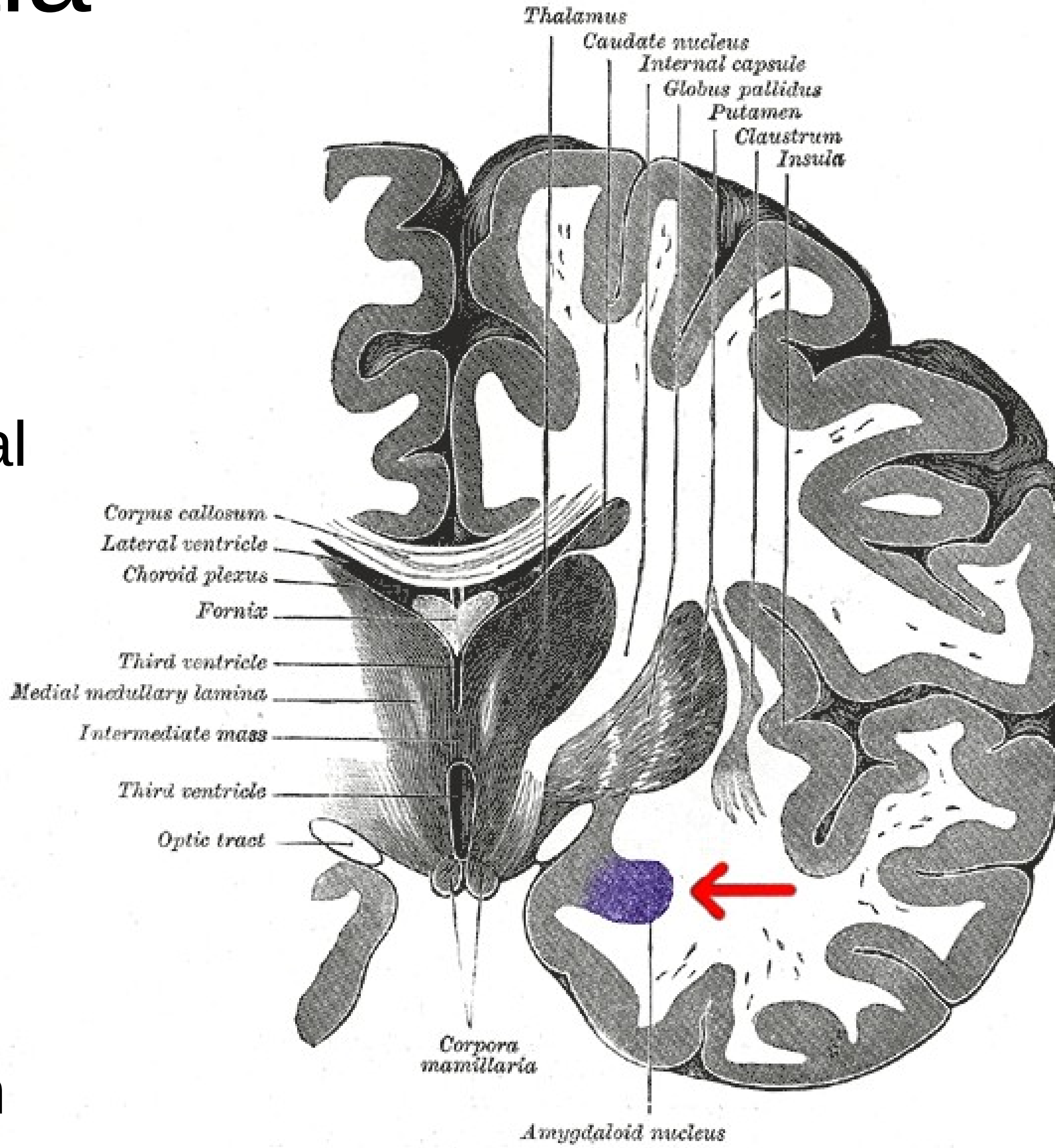
Stages in Addiction

1. Explorative drug taking is found to be pleasurable.
2. Positive reinforcement learning encourages repeated drug taking.
3. Habituation / Tolerance develops.
4. Withdrawal causes stress.

Note that stages 3 and 4 are not inevitable. For example, many people are capable of enjoying occasional alcoholic drinks for many years without becoming alcoholics. But some develop severe alcohol dependency.

The Amygdala

- A major component of the limbic system.
- Known to be involved in fear conditioning responses, but may also mediate positive emotional responses.
- One of the major sources of input to the VTA
- Contains neurons which produce “corticotropin releasing factor” (CRF), a stress hormone.
- Abrupt withdrawal of addictive substances triggers CRF release from those neurons.

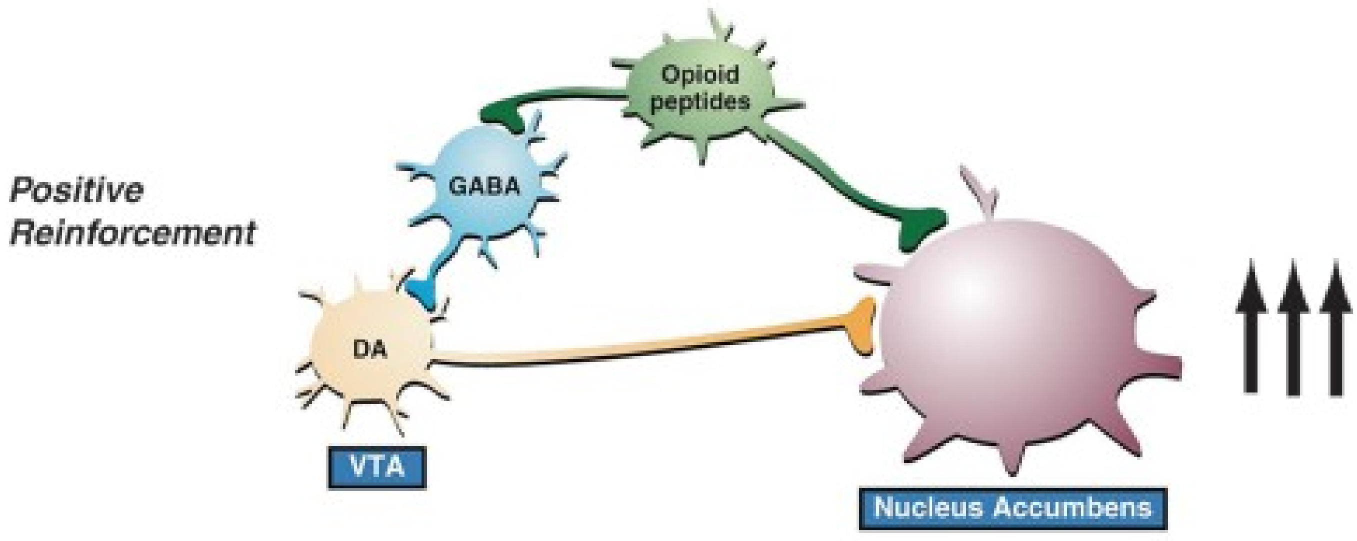


Some addictive drugs also have acute “stress relieving” GABAergic effect on the Amygdala. When the drugs wear off, the stress may come back worse.

Drug of abuse	Neurotransmitter	Site
Cocaine and amphetamines	Dopamine	Nucleus accumbens
	γ-aminobutyric acid	Amygdala
Opiates	Opioid peptides	Nucleus accumbens
	Dopamine	Ventral tegmental area
	Endocannabinoids	
Nicotine	Nicotinic acetylcholine	Nucleus accumbens
	Dopamine	Ventral tegmental area
	γ-aminobutyric acid	Amygdala
	Opioid peptides	
Δ^9-Tetrahydrocannabinol	Endocannabinoids	Nucleus accumbens
	Opioid peptides	Ventral tegmental area
	Dopamine	
Alcohol	Dopamine	Nucleus accumbens
	Opioid peptides	Ventral tegmental area
	γ-aminobutyric acid	Amygdala
	Endocannabinoids	

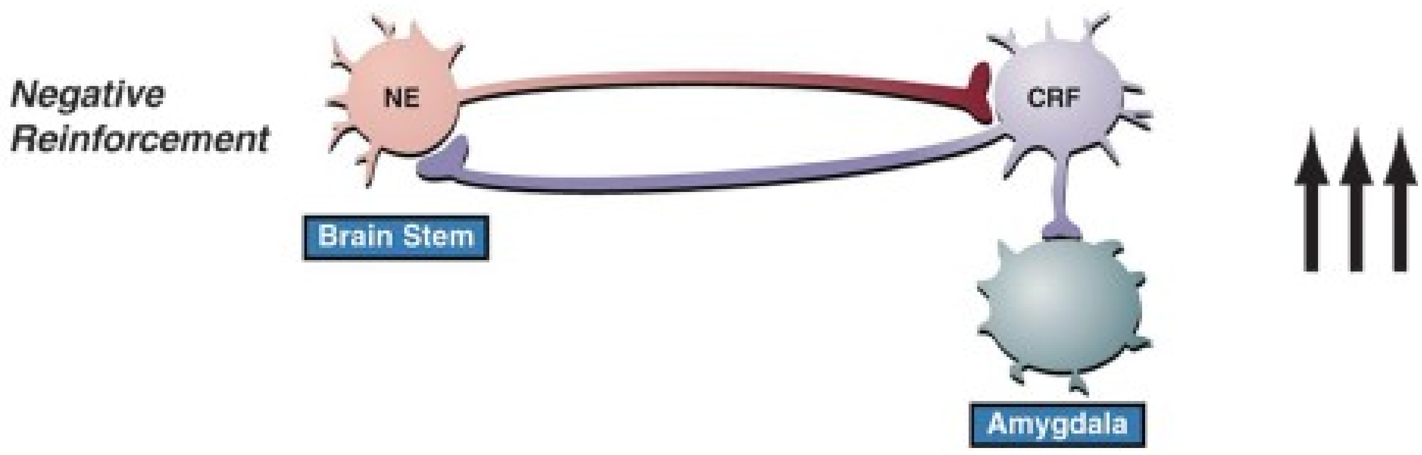
- Koob & Le Moal 2008 “Addiction and the Brain Antireward System”

Non-addict on a “High”



- During recreational drug use, activation of dopaminergic inputs into the NAc is experienced as rewarding and (positively) reinforced.
- The effectiveness of that reinforcer declines with habituation.
- Koob & Le Moal 2008

Addict



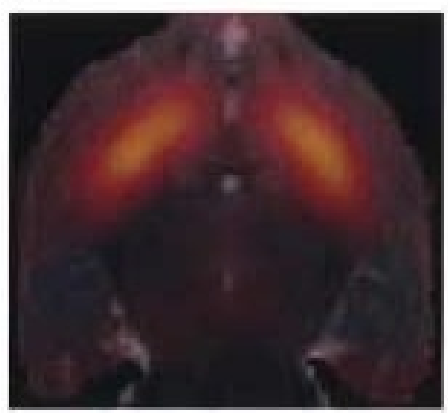
- In the addict, the key motivation is no longer (just) pleasure seeking.
- Enhanced DA input to the amygdala through drug use inhibits stress signal release, but these circuits too can adapt to prolonged elevation of DA signals, and when they do, withdrawal of drug is stressful.
- In the addict, drugs have become a “negative reinforcer” (they “take away the punishment” of withdrawal.)
- Koob & Le Moal 2008

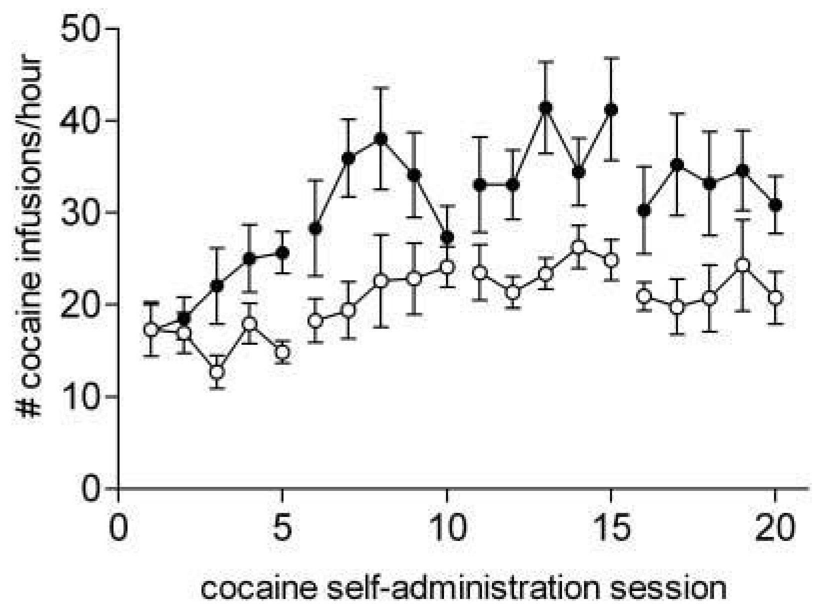
Individual Differences

- Dalley et al (Science 2007) tested rats for “impulsiveness”.
- Impulsive rats have less effective D2/3 receptors in the ventral striatum. They also escalate cocaine doses more quickly in self administration experiments.
- Escalating doses make habituation, and hence addiction, more likely.
- Similar innate differences are likely to exist in humans.

C

D





Summary 1

- Frontal (motor) and prefrontal cortices select actions in collaboration with feedback loops through the basal ganglia.
- Dopamine input to the basal ganglia changes the likelihood of actions being selected.
- Phasic dopamine is thought to provide a “reward prediction error signal” which causes plasticity in the BG loops and increases likelihood of certain actions being selected again in the future.

Summary 2

- Temporal difference learning makes it possible to learn which actions are more likely to be rewarded many steps into the future.
- Drugs which potentiate DA activity can hijack these mechanisms to reinforce drug seeking and drug taking behaviours.
- However, liking a substance and finding it positively reinforcing is usually not enough to trigger full-fledged addiction.
- Full-fledged addiction usually also requires tolerance and withdrawal symptoms which arise from interactions with stress signalling pathways.