Reward Signals

Source:
http://www.scholarpedia.org/article/Reward_Signals

- midbrain dopamine neurons:
  - phasic activation following primary food and liquid rewards, visual, auditory and somatosensory reward-predicting stimuli and physically intense visual and auditory stimuli
  - briefly depressed by reward omission and by stimuli predicting the absence of reward
  - respond very little to aversive stimuli and not at all to inedible objects and known neutral stimuli unless they are very intense or large
  - seemingly signal sth like difference between actual and predicted reward
The power of rewards

- nucleus accumbens (NA) is part of the ventral striatum, belonging to limbic system and basal ganglia
- rats will engage in electrical self-stimulation of their nucleus accumbens, preferring self-stimulation over eating and drinking to the point of starvation (1950s)
- addictive drugs increase dopamine levels in the NA
Midbrain structures

- Substantia nigra
- Periaqueductal gray matter
- Cerebral aqueduct
- Red nucleus
- Reticular formation
- Superior colliculus
- Ventral
- Dorsal

- Superior colliculus (receives visual input)
- Inferior colliculus (receives auditory input)

- Tectum
- Tegmentum
- Cerebellum
Response of midbrain dopamine neurons to rewards and conditioned stimuli

Tomber et al., 2005

Schultz et al., 1997
Blocking Experiment

- two phases of training:
  - first: \( \text{blocked stimulus} + \text{reward} \)
  - second: \( \text{conditioned stimulus} + \text{reward} \)
- result: \( \text{conditioned stimulus} \) doesn’t produce conditioned response

Waelti et al. (2001)
Inhibitory conditioning

- Alternate 2 types of trials
  - \( \text{inhibitory stimulus} + \text{reward} \)
  - \( \text{neutral stimulus} + \text{no reward} \)
- Result: B predicts absence of reward

Tobler et al. (2003)
Scaling of dopamine response

Response magnitude scales with range of prediction error: absolute prediction error may be very different but lead to similar neural response

Tobler et al. (2005)
Animal presses lever to turn on CS, expects reward one second after CS switches off again (habitual interval)
Response scales with expected value of reward

Tobler et al. (2005)
more than a third of dopamine neurons show a slow, sustained, moderate activation reflecting something like the variance of upcoming reward
Midbrain dopamine neurons seem to encode sth. very similar to the temporal difference error we use in RL algorithms

More subtle behaviors of these neurons (uncertainty coding) could inspire new algorithms

One of the most fruitful areas of interaction between theory and experiment in Neuroscience