Reinforcement learning

Reward (prediction error) signals by dopaminergic neurons in the primate brain

Richard Hahnloser

Institute of Neuroinformatics

University of Zurich / ETH Zurich

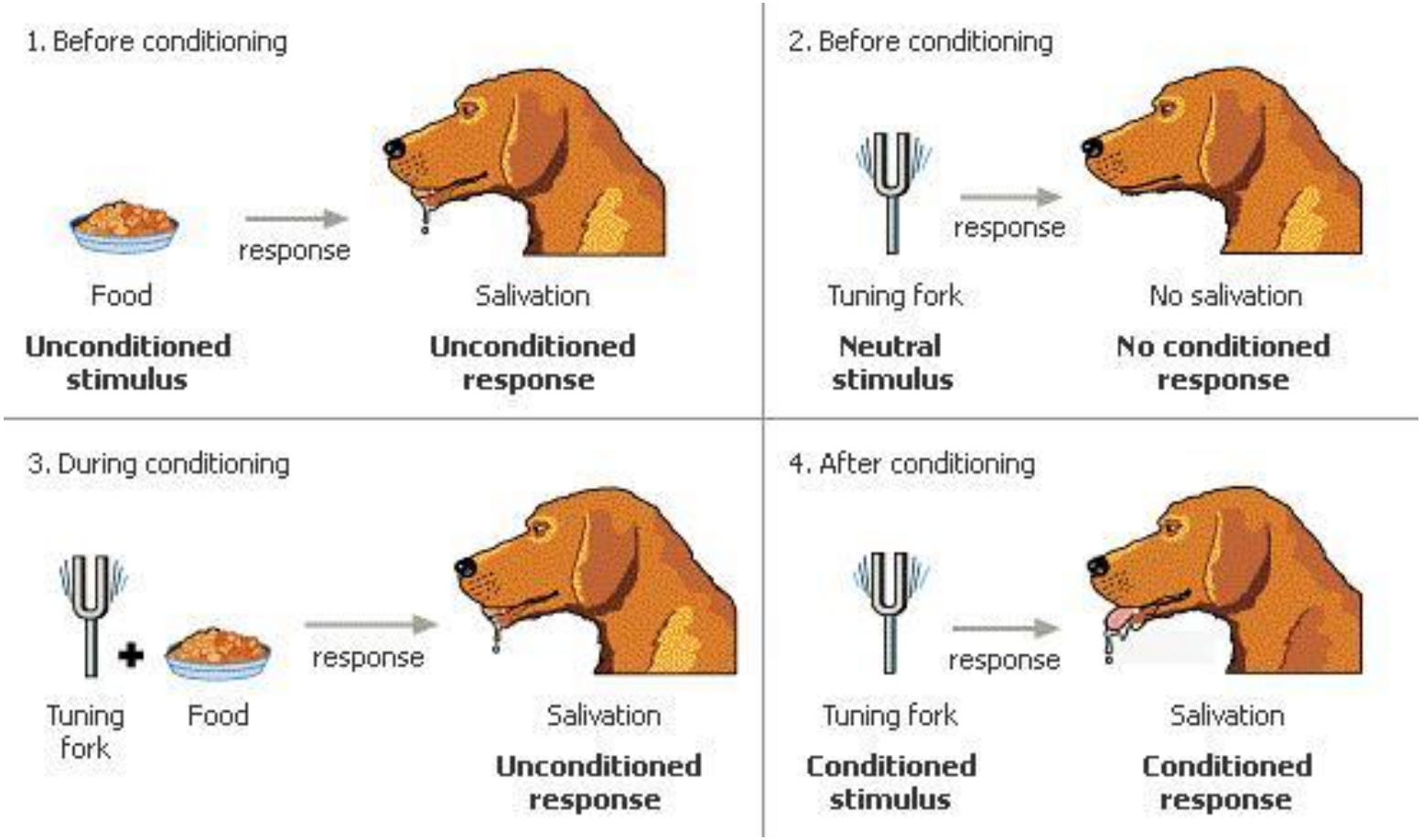
Neurophysics 2014

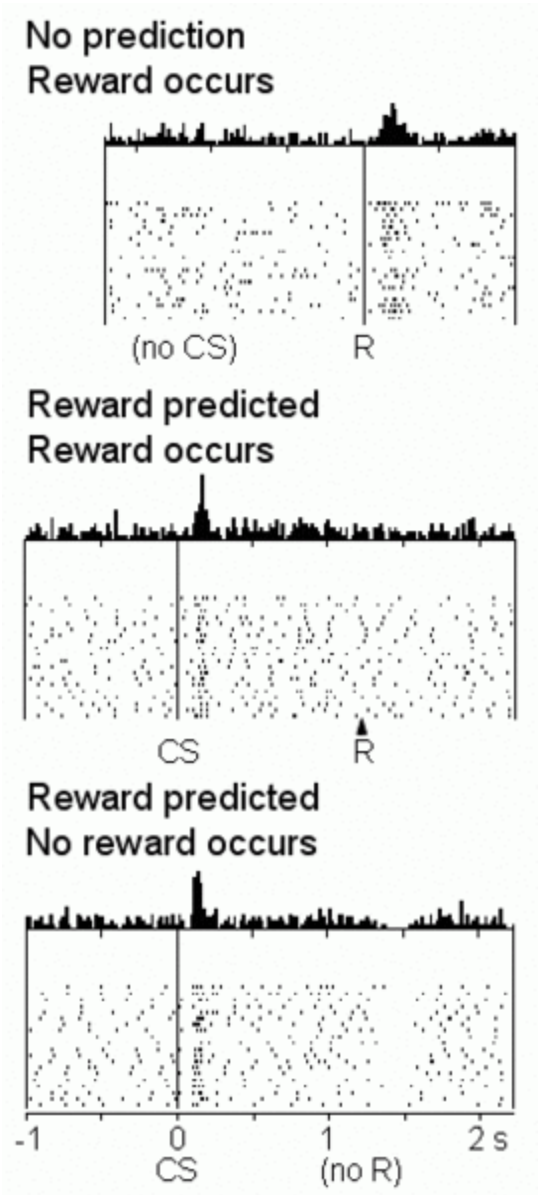
Predictive Reward Signal of Dopamine Neurons

WOLFRAM SCHULTZ
Institute of Physiology and Program in Neuroscience, University of Fribourg, CH-1700 Fribourg, Switzerland

Schultz, Wolfram. Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80: 1–27, 1998. The effects of lesions, receptor blocking, electrical self-stimulation, and drugs of abuse suggest that midbrain dopamine systems are involved in processing reward information and learning approach behavior. Most dopamine neurons show phasic activations after primary liquid and food rewards and conditioned, reward-predicting visual and auditory stimuli. They show biphasic, activation-depression responses after stimuli that resemble reward-predicting stimuli or are novel or particularly salient. However, only few phasic activations follow aversive stimuli. Thus dopamine neurons label environmental stimuli with appetitive value, predict and detect rewards and signal alerting and motivating events. By failing to discriminate between different rewards, dopamine neurons appear to emit an alerting message about the surprising presence or absence of rewards. All responses to rewards and reward-predicting stimuli depend on event predictability. Dopamine neurons are activated by rewarding events that are better than predicted, remain uninfluenced by events that are as good as predicted, and are depressed by events that are worse than predicted. By signaling rewards according to a prediction error, dopamine responses have the formal characteristics of a teaching signal postulated by reinforcement learning theories. Dopamine responses transfer during learning from primary rewards to reward-predicting stimuli. This may contribute to neuronal mechanisms underlying the retrograde action of rewards, one of the main puzzles in reinforcement learning. The impulse response releases a short pulse of dopamine onto many dendrites, thus broadcasting a rather global reinforcement signal to postsynaptic neurons. This signal may improve approach behavior by providing advance reward information before the behavior occurs, and may contribute to learning by modifying synaptic transmission. The dopamine reward signal is supplemented by activity in neurons in striatum, frontal cortex, and amygdala, which process specific reward information but do not emit a global reward prediction error signal. A cooperation between the different reward signals may assure the use of specific rewards for selectively reinforcing behaviors. Among the other projection systems, noradrenaline neurons predominantly serve attentional mechanisms and nucleus basalis neurons code rewards heterogeneously. Cerebellar climbing fibers signal errors in motor performance or errors in the prediction of aversive events to cerebellar Purkinje cells. Most deficits following dopamine-depleting lesions are not easily explained by a defective reward signal but may reflect the absence of a general enabling function of tonic levels of extracellular dopamine. Thus dopamine systems may have two functions, the phasic transmission of reward information and the tonic enabling of postsynaptic neurons.

Classical conditioning

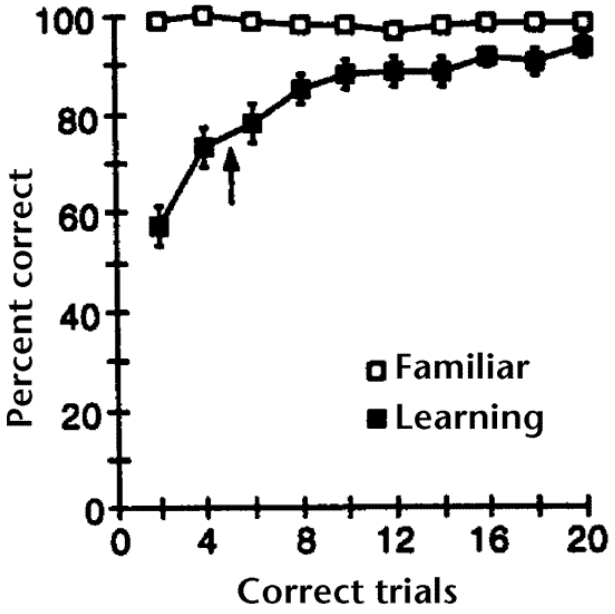


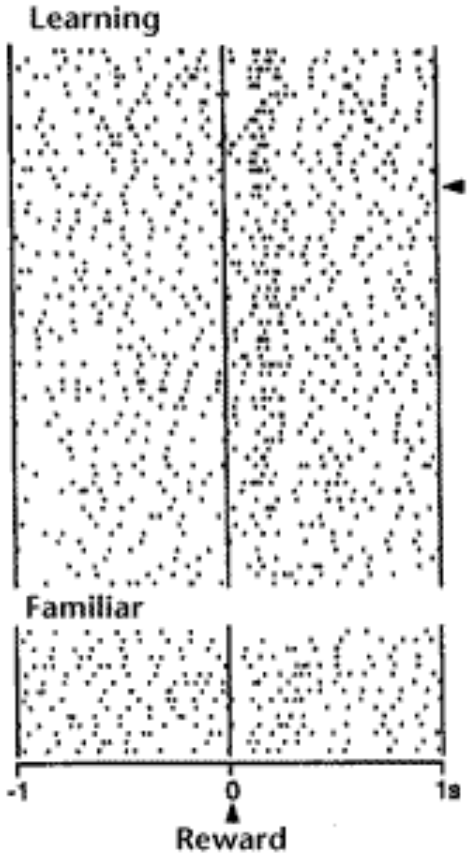


Midbrain dopamine neuron (Ventral tegmental area, VTA)

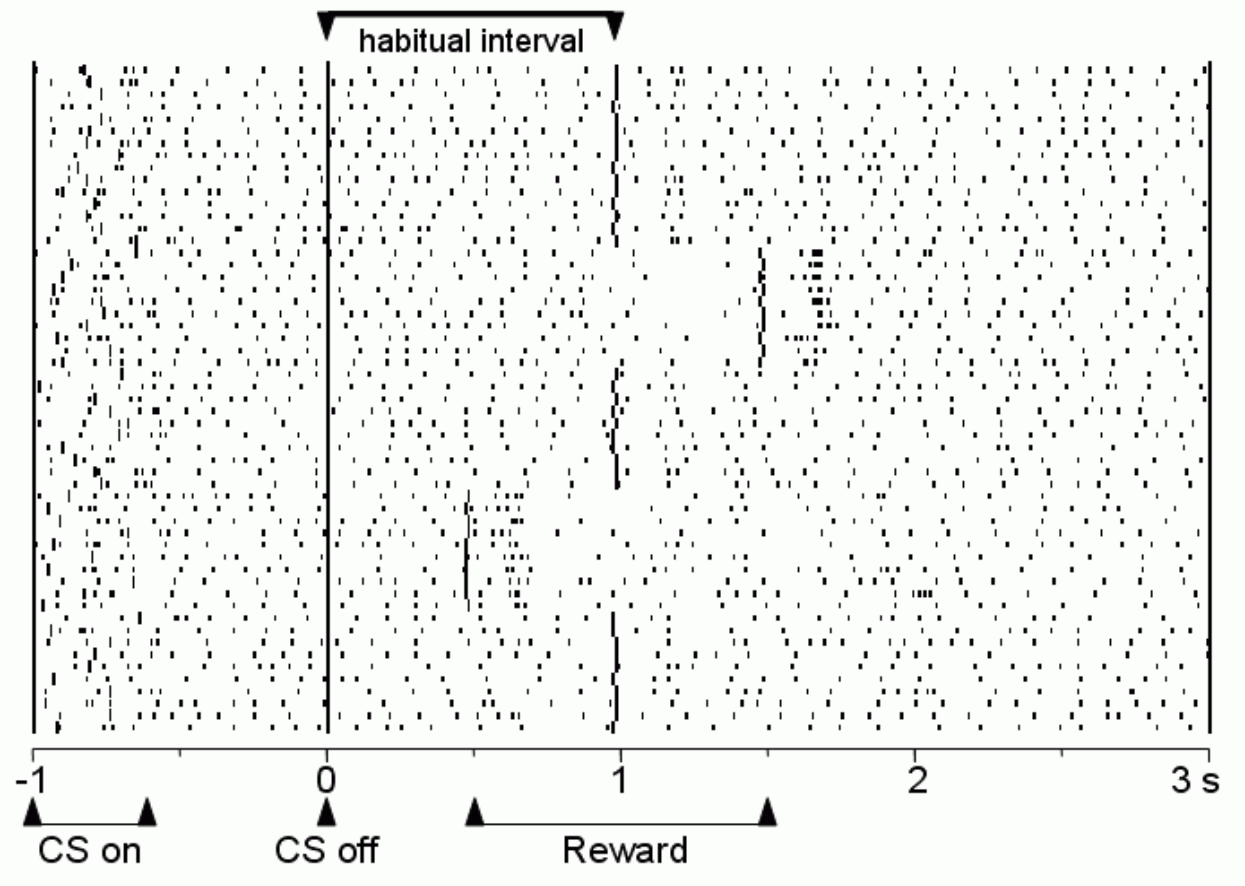
Reward prediction error response

Single neuron and behavior





Dopamine neurons code errors in the prediction of both the occurrence and the time of rewards



Temporal sensitivity of prediction error response of dopamine neuron. From top to bottom: Reward delay by 0.5 s leads to considerable depression at the habitual time of reward and activation at the new time. Earlier reward leads to activation at new time but not to major depression at the habitual time. The habitual time of reward is at 1.0 s after touch of an operant key and simultaneous offset of conditioned stimulus (CS). From Hollerman & Schultz (1998).