

How We Decide: The Neuronal Reward Signal

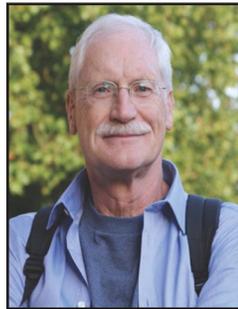
Wolfram Schultz

Department of Physiology, Development and Neuroscience, University of Cambridge

When we hear the term “neuronal reward signal” we might think of a brain signal that conveys pleasure, or a special bonus for having done something well. But the function of reward is more general and extends to learning and approach of objects that we need for survival; individuals whose brain deals best with reward have the greatest chance to live and to pass on their genes. A neuronal reward signal informs and mediates brain mechanisms for obtaining rewards and making economic decisions.

The neuroscience of reward investigates how the brain detects and responds to rewards. We define the function of rewards by their influence on the body, not by specific sensory receptors or by any external measure. We can tell from observing the behavior of a human, monkey, or mouse whether a substance, stimulus, or object is a reward or not. A subject will approach a reward and learn to do it more often if it turns out to be good and do it less if it turns out bad; a reward that makes us happy, for example, induces a desire to have it again. By contrast, the physical properties of rewards don’t offer a good measure: Think of the third steak you are eating; it is physically the same as the first two, but it won’t give you the same satisfaction because you are no longer hungry. That distinction is important, because it allows each of us to seek the reward that is best for us at the time. Getting the best, rather than just any reward may be important for daily survival.

The significance that a substance or an object has for us is called subjective reward value. Researchers cannot measure that value physically,



Wolfram Schultz

but we can infer it from observing the choices and their frequency. We assume that choosing one object over another suggests a higher subjective value of the chosen object for the subject at that moment. A similar principle holds for preferences, which are elicited at the moment a subject makes a choice. By making these choices, the subject reveals her preferences.

Economic choice theory deals with value and revealed preferences and defines the rules for how to estimate them by observing an individual’s choices. As there are many usages of the term value, including the physical amount of rewards (sips of juice for animals or dollars for humans) and their probability, economic theory has coined the term utility, defining it as the crucial value variable underlying economic choices.

Physical implementation of a theoretical construct

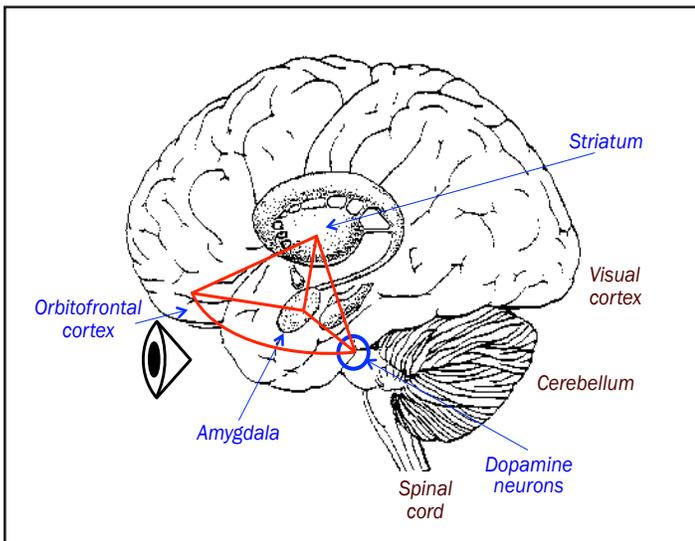
We can use a physical response such as a neuronal action potential—a momentary change in electrical charge on the surface of a neuron, caused by incoming information and influencing the activity of downstream neurons—to test our theories of reward. Patterns of neural firing rates also provide data for understanding underlying aspects of everyday life, and could influence such areas as health policy and promulgating laws. For example, if there is a brain signal for reward that drives behavior strongly toward ingesting a certain unhealthy reward—fatty food, perhaps—then it may be difficult to prevent the behavior by simply outlawing it. It might be better to find ways to reduce a reward’s attractiveness instead.

Below, I describe what we’ve learned about neuronal signals for the principal functions of reward:

learning, approach behavior, and decision-making, and how they reflect current theories of behavior. Human research also addresses the emotional and cognitive aspects of reward function (such as pleasure, happiness, and desire), but my focus here is on animal experiments, which don't easily allow us to examine these important issues.

A limited number of brain structures show neuronal signals for reward and economic decisions. The four key structures are the midbrain dopamine system, the striatum (putamen, caudate nucleus, and ventral striatum), the amygdala (in the medial temporal lobe), and the orbitofrontal cortex (the cortex above the eyes). These structures are tightly interconnected, each sending axons of their neurons to the other three structures.

Subsidiary reward and decision signals are found in brain regions closely connected to these four structures. These regions process reward in association with sensory stimuli, which are likely involved in identifying and distinguishing among rewards based on their physical properties, or carry signals in association with movements and behavioral choices.



The interconnected, interactive and cooperative reward circuit in the brain. Only the four main reward structures are shown. *Image courtesy of Wolfram Schultz*

Dopamine neurons

The strongest and most consistent reward signal is found in the dopamine neurons in pars compacta of substantia nigra and the ventral tegmental area (VTA). The signal reflects a [reward prediction error](#), a measure of the difference between the actual reward and what an individual expected it would be (error here does not mean “mistake,” just difference). A better reward than predicted (positive prediction error) elicits an increase in the firing rate of the dopamine neurons (activation); a worse reward than predicted (negative error) reduces the firing rate (inhibition), and a fully predicted reward fails to change the rate (no response). This signal does not distinguish between rewards valued innately (the so-called unconditioned stimulus US) and those that have been trained or conditioned (conditioned stimulus CS). This response is an efficient informational signal and underlies what we term [associative learning](#), the process by which a person or animal learns an association between two events, like a stimulus and a reward. This basic type of learning includes classical and operant conditioning.

The signal also reveals the coding of formal economic utility, distinct from physical value. The dopamine reward utility signal conveys positive and negative valence; the positive error signal (activation) seems to drive animals towards better states (monkeys and rodents prefer movements and choices that lead to an activation of dopamine neurons) and away from less beneficial states (avoiding places associated with inhibition of dopamine neurons). This observation corresponds to the conditions for maximizing rewards proposed by the [Bellman equation](#) and dynamic programming that informed the development of machine learning.

Subgroups of dopamine neurons show additional slow and heterogeneous variations during general behavioral activation involving considerable sensory and motor activity, such as large reaching movements in monkeys and treadmill activity in rodents; there are no consistent dopamine neuronal activations reported during more controlled and constrained arm, hand, finger, and eye movements.

The implied dopamine informational function for [reinforcement learning](#) involves widespread

anatomical projections to the other key reward structures, including the striatum, amygdala, and orbitofrontal cortex. Dopaminergic postsynaptic effects in these regions modulate neuroplasticity through [long-term potentiation](#) and [long-term depression](#), which persistently alter the efficacy of neurotransmission at affected synapses. These neuronal adaptations are consistent with three-factor [Hebbian learning](#), in which the prediction error signal plays the third factor, affecting how well the presynaptic-postsynaptic conjunction works. In the striatum and amygdala, there are also some neurons that show a similar bidirectional reward prediction error signal, but most of their neurons exhibit other behavioral changes, some of them being involved in economic decisions. We know that dopamine influences on these structures are necessary for learning because of the deficits animals with dopamine lesions or psychopharmacological blockade of dopamine receptors have.

Striatum

Most neurons in dorsal and ventral striatum process various reward signals, including responses to reward itself (with some neurons coding reward prediction error), responses to reward-predicting stimuli, and slow activations during reward expectation. In addition to these pure reward responses, many striatal movement-related neurons are modulated in their activity by expected rewards. These activities are compatible with the notion of goal-directed action in which the reward goal is represented while the animal prepares and executes the action.

When experimenters use designs of [reinforcement learning](#), some striatal neurons code action value, the value of a reward that the animal would receive for taking a specific action, irrespective of whether the animal actually chooses that action. Thus, a left-action value-coding neuron tracks the reward value (amount or probability) expected for a left action, irrespective of taking that action, while a right-action value neuron tracks the value expected for an action to the right. In choices between left and right actions, such distinct action value neurons

would constitute competing inputs to a winner-take-all decision mechanism: The action with the highest value-related activity would win and direct the animal's choice toward that action. In this way, some striatal reward signals comply with formal notions of [reinforcement theory](#) for decision-making.

The striatal reward signals that incorporate movement information and action value are compatible with the general motor function of the striatum: The stimulation of striatal neurons carrying D1 or D2 dopamine receptors leads to respectively facilitatory (D1) or inhibitory (D2) effects on economic choices. In this way, striatal neurons seem to link the dopamine reward signal to economic decision-making.

Amygdala

Although the function of the amygdala has long been primarily associated with fear, recent experiments demonstrate substantial reward and decision signals. A landmark study by Paton and collaborators showed that amygdala neurons respond to reward and conditioned reward-predicting stimuli in well-controlled [Pavlovian experiments](#) (classical conditioning). The likely underlying behavioral mechanism relies on the information provided by a reward predictor, over and above simple stimulus-reward pairing; a stimulus during which the animal receives more (or less) reward than during stimulus absence is more informative compared with a stimulus that does not change reward. To test this hypothesis, we elevated the reward outside the stimulus to the same level as during the stimulus, which eliminated the specific reward information and thus the reward prediction. We then reduced the outside reward to below the level of the stimulus; this made the stimulus informative and predictive again. Importantly, all of these changes concerned only the reward outside the stimulus, while leaving stimulus-reward pairing unchanged. During these tests, about half of the tested stimulus-sensitive neurons in the amygdala correspondingly lost and then regained their reward prediction response.

Besides these basic reward functions, amygdala neurons process decision variables for more demanding economic choices. They signal reward value early in a trial, and switch within a second or two to coding the choice irrespective of the specific reward chosen, thus covering the transition from value to choice. Other types of amygdala neurons show increasing activity while the animal performs sophisticated save-spend choices; this activity does not occur with imperative trials in which the animal cannot make choices. Taken together, amygdala neurons follow fundamental conditions for reward learning and are actively involved in sophisticated economic decision-making processes. These experiments have established the amygdala as a main component of the brain's reward system.

Frontal cortex

Several frontal cortical areas that signal various forms of reward information are involved in processing diverse aspects of reward. Different groups of neurons in orbitofrontal cortex code (1) the value of specific reward objects irrespective of being chosen (object value coding analogous to action value coding), (2) the value of the chosen reward irrespective of its type and sensory properties, (3) reward relative to the current distribution of rewards, and (4) the risk in reward gambles. Other, more dorsolateral parts of the frontal cortex integrate reward information into movement-related activity.

Neurons in the frontal cortex are also involved in social processes, including the sensing of a conspecific's reward delivery and omission, the distinction between own and other's reward, and the

animal's level of engagement in competitive vs. non-competitive video games. The further investigation of reward processes in cognitive brain structures such as frontal cortex may become a fruitful approach to studying social processes at the neuronal level.

Brief conclusions

These electrophysiologic studies have revealed patterns of neuronal activity that comport, sometimes in mathematically predictable ways, with the economic principles of decision-making. The neuronal signals occur in several principal brain structures that are highly interconnected and contribute in different ways to detecting rewards, using reward information for learning and economic decisions, and driving behavior towards obtaining more and better rewards.

Further reading

- Schultz W, Dayan P, Montague RR. [A neural substrate of prediction and reward](#) (PDF). *Science* 275: 1593-1599, 1997.
- Samejima K, Ueda Y, Doya K, Kimura M. [Representation of action-specific reward values in the striatum](#). *Science* 310: 1337-1340, 2005.
- Paton JJ, Belova MA, Morrison SE, Salzman CD. [The primate amygdala represents the positive and negative value of visual stimuli during learning](#). *Nature* 439: 865-870, 2006.
- Padoa-Schioppa C, Assad JA. [Neurons in the orbitofrontal cortex encode economic value](#). *Nature* 441: 223-226, 2006.
- Bermudez MA, Schultz W. [Responses of amygdala neurons to positive reward predicting stimuli depend on background reward \(contingency\) rather than stimulus-reward pairing \(contiguity\)](#) (PDF). *J Neurophysiol* 103: 1158-1170, 2010.
- Stauffer WR, Lak A, Schultz W. [Dopamine reward prediction error responses reflect marginal utility](#). *Curr Biol* 24: 2491-2500, 2014.
- Schultz W. [Neuronal Reward and Decision Signals: From Theories to Data](#). *Physiological Reviews* 95: 853-951, 2015