



HHS Public Access

Author manuscript

J Comp Neurol. Author manuscript; available in PMC 2017 June 01.

Published in final edited form as:

J Comp Neurol. 2016 June 1; 524(8): 1699–1711. doi:10.1002/cne.23880.

Components and characteristics of the dopamine reward utility signal

William R. Stauffer, Armin Lak, Shunsuke Kobayashi, and Wolfram Schultz

Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge CB2 3DY, United Kingdom

Abstract

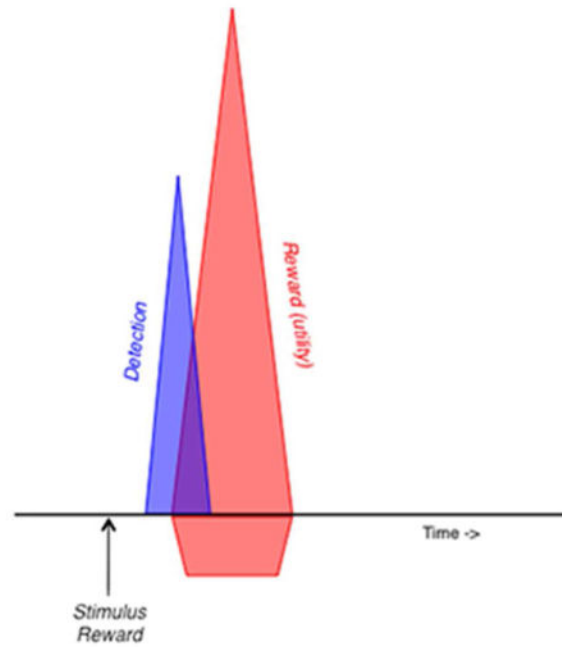
Rewards are defined by their behavioral functions in learning (positive reinforcement), approach behavior, economic choices and emotions. Dopamine neurons respond to rewards with two components, similar to higher order sensory and cognitive neurons. The initial, rapid, unselective dopamine detection component reports all salient environmental events irrespective of their reward association. It is highly sensitive to factors related to reward and thus detects a maximal number of potential rewards. It senses also aversive stimuli but reports their physical impact rather than their aversiveness. The second response component processes reward value accurately and starts early enough to prevent confusion with unrewarded stimuli and objects. It codes reward value as a numeric, quantitative utility prediction error, consistent with formal concepts of economic decision theory. Thus, the dopamine reward signal is fast, highly sensitive and appropriate for driving and updating economic decisions.

Graphical Abstract

Corresponding author: Wolfram Schultz Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge CB2 3DY, United Kingdom, ws234@cam.ac.uk Phone: +44-1223-333 779 Fax: +44-1223-333 840.

Author contributions: WRS, AL and SK did the experiments underlying this review, WS together with WRS, AL and SK wrote the paper.

Conflict of interests: The authors declare no conflict of interest.



The response of midbrain dopamine neurons to rewards and reward-predicting stimuli consists of two components. The initial, unselective component detects any salient environmental event very rapidly, whereas the subsequent component codes a formal reward utility prediction error, as conceptualized by combining animal learning theory and economic decision theory.

Keywords

Stimulus components; subjective value; temporal discounting; utility; risk; neuroeconomics

Rewards are stimuli, objects, events, situations and activities with crucial biological functions for individual survival and gene propagation. Specific behavioral learning and decision tasks serve to assess the neuronal underpinnings of reward functions. Individual neurons signaling reward information are found in the dopamine system, striatum, orbitofrontal cortex, amygdala, and their associated structures. These reward neurons process specific aspects of rewards such as amount, probability, value, utility and risk in forms suitable for economic decisions. Brain structures with reward neurons often contain also neurons that process aversive stimuli and punishment. Dopamine neurons do not seem to be activated by punishers and code reward value only in a positive monotonic manner, as results from recent well-controlled experiments suggest.

This review elaborates on the properties of one of the brain's main reward system, the dopamine neurons. We describe the recently clarified specificity of its signal to reward as opposed to other events. We then relate the dopamine signal to formal economic decision theory, which provides a stringent conceptual framework for reward function. Specifically, we demonstrate that the dopamine reward prediction error signal processes reward as a specific form of subjective value called economic utility.

BACKGROUND

Reward function

The body needs specific substances for survival, including proteins, carbohydrates, fats, vitamins, electrolytes and water, which are contained in foods and liquids. To propagate their genes, individuals must mate, reproduce and raise offspring. Food and liquid rewards subserve the alimentary needs. Different from common associations of reward with bonus or happiness, the scientific use of the term stresses three functions. Rewards are positive reinforcers, they 'make us come back for more'. This function is captured most simply by Pavlovian and operant conditioning. Further, rewards are attractive, generate approach behavior and serve as arguments for economic decisions. This function is crucially based on the stringently formalized term of economic utility. Finally, rewards are associated with positive emotions, in particular with pleasure as a reaction to something that turns out to be good, with desire of something that is already known to be good, and with happiness as a longer lasting state derived from pleasure and desire (Schroeder 2004). Thus, a reward is a stimulus, object, event, activity or situation that induces positive learning, makes us approach and select it in economic choices, and/or induces positive emotions.

The proximal reward functions of learning, approach and emotions serve the ultimate, distal function of rewards, which is to increase evolutionary fitness. To acquire and follow these primary alimentary and reproductive rewards is the reason why the brain's reward system has evolved. Thus, the proximal reward functions help the evolutionary selection of phenotypes that maximize gene propagation.

Neuronal reward signals

Reward processing requires a large array of brain functions from glia to synapses and channels. Investigations of neuronal reward processing consider that information processing systems work with identifiable signals that serve the rapid detection of important events and lead to efficient behavioral actions. Although many molecular and cellular approaches provide sufficient spatial and temporal resolution, the action potentials of individual neurons are often the most appropriate neuronal signals for investigations involving well-controlled behavioral actions. The rate of action potentials (impulses/s) provides a code for neuronal processing of reward information during behaviors defined by concepts derived from experimental psychology and experimental economics. It is thus possible to identify reward neurons by studying action potentials in relation to specific theoretical constructs (parameters) of reward functions in learning and economic choices.

Reward prediction error signal

Dopamine neurons in the pars compacta of substantia nigra and in the ventral tegmental area (VTA) show a phasic neuronal signal that codes reward relative to its prediction. A reward that is better than predicted generates a positive signal (increase in impulse rate), a fully predicted reward fails to generate a signal, and a reward that is worse than predicted generates a negative signal (decrease in impulse rate) (Schultz et al. 1997; Schultz 1998). The response occurs in the same prediction-dependent way with conditioned, reward predicting stimuli that themselves have acquired reward value (Enomoto et al. 2011). This

signal codes a reward prediction error as conceptualized in major reinforcement models, such as the Rescorla-Wagner rule (Rescorla & Wagner 1972) and temporal difference learning (Sutton & Barto 1981). It follows formal theoretical criteria for prediction error processing, namely blocking (a stimulus not associated with a prediction error is blocked from behavioral and neuronal learning; Waelti et al. 2001) and conditioned inhibition (a stimulus associated with a negative prediction error inhibits behavioral and neuronal responses; Tobler et al. 2003). A positive dopamine response would enhance coincident synaptic transmission in striatum, frontal cortex or amygdala neurons, and a negative response would reduce synaptic transmission, thus directing individuals towards better rewards and away from worse rewards and helping to maximize utility in economic choices. Such a signal is also useful for learning and updating of reward values.

The phasic reward prediction error signal occurs in 70–90% of dopamine neurons in a rather stereotyped manner and shows only graded differences in latency, duration and magnitude between the neurons (Fiorillo et al. 2013a). Some dopamine neurons show an additional, slower response that codes reward risk (Fiorillo et al. 2003) and may affect the dopamine release induced by the prediction error signal. Apparent more complex relationships to cognitive processes can be explained by reward prediction error coding with appropriate analysis (Ljungberg et al. 1992; Morris et al. 2006; de Lafuente & Romo 2011; Enomoto et al. 2011; Matsumoto & Takada 2013). There are also minor, activations or depressions before or during large reaching movements (Schultz et al. 1983; Schultz 1986; Romo & Schultz 1990) but neither with slightly different arm movements (DeLong et al. 1983; Ljungberg et al. 1992; Satoh et al. 2003) nor with licking or eye movements (Waelti et al. 2001; Cohen et al. 2012; Stauffer et al. 2015), which might inconsistently reflect general behavioral activation, attention or risk rather than robust movement relationships deficient in Parkinsonism. Distinct from these activities, dopamine neurons are well known to show similar structural and functional diversities as other neurons (e. g. Roeper 2013). The dopamine release driven by the phasic impulse responses is subject to local mechanisms, such as cholinergic activity (Threlfell et al. 2012) and glutamate co-release (Chuhma et al. 2014). Thus, the stereotyped dopamine prediction error signal may provide locally differentiated influences on postsynaptic processing. At slower time courses over minutes and hours, dopamine impulse activity and dopamine release change with a large range of behavior, including behavioral activation, forced inactivation, stress, attention, reward, punishment and movement (Schultz 2007). At the slowest time course, tonic dopamine levels are finely regulated by local presynaptic and metabolic mechanisms. Pathological and experimental deviations from these levels are associated with Parkinson's disease, schizophrenia and many other cognitive, motivational and motor dysfunctions. The implication of dopamine in these many functions is less derived from the stereotyped phasic dopamine reward signal and more from the specific functions of the postsynaptic structures affected by dopamine.

TWO-COMPONENT REWARD RESPONSES

Stimulus components and their sequential processing

Rewarding stimuli, objects, events, situations and activities are composed of sensory components (visual, auditory, somatosensory, gustatory or olfactory), attentional components (physical intensity, novelty, surprise and motivational impact) and motivational value (reward). Neurons process these components in sequential steps, which become well distinguishable with increasing stimulus complexity. The occurrence of a stimulus is initially detected by its physical impact (Fig. 1A) and gives rise to subsequent processing of its specific sensory properties, like spatial position, orientation and color that identifies the stimulus. As last step, maybe partly in parallel, neurons assess the motivational value, which determines reward and punisher function. Similar steps are assumed in models for memory retrieval and sensory decisions (Ratcliff 1978; Ratcliff & McKoon 2008).

Whereas primary sensory systems may identify simple, undemanding stimuli at once and without apparent components, higher systems may process more complex objects in the sequential steps just outlined. Thus, neurons in somatosensory barrel cortex, prefrontal cortex, frontal eye fields, lateral intraparietal cortex, visual cortex and pulvinar show step-wise processing, consisting of an initial, unselective detection response and a subsequent, selective component that identifies the specific stimulus in terms of orientation, motion, spatial frequency, visual category, target-distractor distinction, and figure-background distinction (Fig. 1B) (Thompson et al. 1996; Ringach et al. 1997; Kim & Shadlen 1999; Roelfsema et al. 2007; Lak et al. 2010). Reward-processing neurons in amygdala, primary visual cortex and inferotemporal cortex show similar distinct processing steps that comprise an initial, unselective stimulus detection and a subsequent reward valuation (Paton et al. 2006; Mogami & Tanaka 2006; Ambroggi et al. 2008; Stanisor et al. 2013). Thus, higher neuronal systems are well known to process detection, identification and valuation of stimuli in identifiable sequential steps.

Two dopamine response components

As other neurons processing complex information, dopamine neurons process reward components sequentially. Any reward with sufficient physical intensity may elicit two response components (Fig. 1C, D). The first response component consists of an unselective increase of activity ('activation') and occurs with any stimulus irrespective of being associated with reward, punishment or nothing. It reflects the detection of the event without coding its value. It occurs also as striatal dopamine release, without distinguishing between reward-predicting and non-predicting stimuli (Day et al. 2007). The subsequent, second response component reflects the reward value of a reward-predicting stimulus or a reward and codes positive and negative reward prediction errors in a graded manner. These components are difficult to separate when the reward is simple and undemanding and can be rapidly identified and valued (Fig. 1C) (Steinfels et al. 1983; Schultz & Romo 1990; Ljungberg et al. 1992; Horvitz et al. 1997; Fiorillo et al. 2013b; Fiorillo 2013). They become more distinct with more demanding reward-predicting stimuli, such as moving dots that take longer to identify and value (Fig. 1D) (Nomoto et al. 2010), and with primary rewards whose amount is determined by liquid flow duration (onset of liquid flow elicits an initial

activation that persists only with larger amounts; Lak et al. 2014). Thus, dopamine responses consist of two components with different properties, which will now be discussed.

High sensitivity of initial response component

The initial activating response is enhanced by several factors, including physical stimulus intensity (Fiorillo et al. 2013a; b). Furthermore, a mechanism called generalization induces or enhances the response to an unrewarded stimulus, even a punisher, if that stimulus resembles physically a rewarded stimulus (Waelti et al. 2001; Tobler et al. 2003). For example, visual aversive stimuli rarely elicit activations when the alternating rewarded stimulus is auditory, but the unchanged aversive stimulus is very effective in inducing substantial activations when both stimuli are visual and thus resemble each other (Fig. 2) (Mirenowicz & Schultz 1996). Such generalization is also seen with striatal dopamine release (Day et al. 2007). The initial response increases also in contexts in which rewards are known to occur (here, 'context' refers specifically to all environmental stimuli and events except the explicit, differential stimulus). Small, unrewarded stimuli elicit no or only small dopamine activations before learning (Tobler et al. 2003) or in unrewarded contexts but are very effective in well-controlled rewarded environments (Fig. 2) (Kobayashi & Schultz 2014). The initial activation increases also with stimulus novelty (Ljungberg et al. 1992), although novelty alone without sufficient stimulus intensity is ineffective (Tobler et al. 2003). Thus, the initial, unselective detection response component is highly sensitive to factors related to rewards.

Saliency, but only initially

The factors generating and enhancing the initial dopamine activation are likely endowed with stimulus-driven saliency. Intense stimuli elicit attention through their noticeable physical impact, which is called physical saliency. Rewards not only carry positive value but also induce attention due to their important biological functions, called motivational saliency. (Biologically important punishers carry negative value but induce also motivational saliency.) Thus, stimuli that resemble rewards physically (generalization) or that occur in environments in which rewards are known to occur (reward context) are associated with motivational saliency, even before they have been identified as rewards. Finally, novel stimuli, even if they require identification and comparison with previously experienced stimuli, are surprising and thus draw attention, suggesting that they are endowed with novelty/surprise saliency. Thus, the initial dopamine activation is sensitive to saliency.

Despite this activation, the phasic dopamine response is not just a saliency response, as sometimes assumed (Kapur 2003). The saliency activation is rapidly followed by the second response component, which codes reward value as a graded, bidirectional reward prediction error (Fig. 1C, D). The second component is also invisible when an unrewarded stimulus occurs unpredictably; the lack of prediction error (no reward minus no prediction) would explain the absence of the second component, and the whole dopamine response consists only of the initial, detection component. This may be the reason why initial studies that did not use rewards interpreted the whole phasic dopamine response as if it were primarily coding saliency (Steinfels et al. 1983; Horvitz et al. 1997), an interpretation strengthened by its short latency (Redgrave et al. 1999). The second component itself does not code saliency,

as the very salient negative reward prediction errors and conditioned inhibitors do not induce any activation (Schultz 1998; Tobler et al. 2003). Aversive dopamine responses (Mirenowicz & Schultz 1996; Joshua et al. 2008; Matsumoto & Hikosaka 2009) might simply constitute the first component enhanced by physical salience, reward generalization and context (see below). Thus, dopamine neurons seem to code salience only as an initial activation, which might appear as their prime response with incomplete testing.

Benefits from the initial response component

The initial dopamine detection response component is very fast and occurs at latencies below 100 milliseconds. It is also very unselective, as it occurs not only with rewards but also with all kinds of unrewarded stimuli and even with punishers. Such a fast response allows brain mechanisms to start initiating behavioral reactions already before the stimulus has been fully identified and valued. If the stimulus turns out to be a reward, the behavioral initiation can proceed, rather than starting only at this moment. If it is not a reward, it is early enough to cancel the behavioral initiation and prevent errors (see below). Thus, no time is wasted by waiting for valuation, resulting in the quickest possible behavioral initiation.

The unselectivity of the initial response assures a wide, polysensory sensitivity to all possible stimuli and objects that may be rewards. The additional sensitivity to stimulus intensity, similarity, reward context and novelty further increases the chance to detect a reward and prevent missing it. It seems better to over-detect rewards than to under-detect them, as long as costly behavioral errors can be avoided (see below). Thus, dopamine neurons process potential rewards already before their reward nature has been confirmed or rejected.

The coding of physical, motivational and novelty/surprise salience would be beneficial for decision-making and behavioral learning (Pearce & Hall 1980). Attention is well known to enhance the neuronal processing of stationary and moving visual stimuli (Bushnell & Goldberg 1981; Treue & Maunsell 1996) and visual orienting (Nardo et al. 2011). The brief dopamine salience signal is likely to enhance in a similar way the neuronal processing of reward prediction error by the second component.

Thus, the fast and highly sensitive, unselective initial dopamine detection response is appropriate for very rapidly detecting a maximal number of rewards and for missing only few, thus contributing to maximal reward acquisition, which in the long run is evolutionary beneficial.

Correct reward valuation by the second response component

Once assessed by the second, reward prediction error component, which is particularly distinctive with more demanding stimuli (Fig. 1D; Nomoto et al. 2010), the value information remains present until the reward occurs. This becomes evident when a known test reward elicits a reward prediction error; a surprising reward following an unrewarded stimulus elicits a positive reward prediction error and a corresponding dopamine activation (Fig. 3; Waelti et al. 2001).

The period of valuation precedes the behavioral reactions towards the reward (Fig. 3). Postsynaptic neurons downstream to dopamine neurons are likely to 'see' both dopamine response components, but they receive the value information from the second component early enough to allow correct behavioral reactions. This mechanism may explain why animals don't often confuse rewards with non-rewards and punishers; despite neuronal generalization to unrewarded stimuli, they don't generalize in their behavioral reactions (Mirenowicz & Schultz 1996; Day et al. 2007; Joshua et al. 2008).

Thus, by occurring early enough for appropriate behavioral reactions, the second response component prevents major negative effects that might derive from the early onset and unselectivity of the first response component, thus resulting in a positive tradeoff between benefits for reward acquisition and its costs.

No aversive activation

Dopamine neurons are activated by aversive stimuli (Tsai et al. 1980; Schultz & Romo 1987; Schultz & Mirenowicz 1996; Guarraci and Kapp 1999). However, these activations reflect the physical intensities of the aversive stimuli rather than their negative value; keeping physical impact constant and increasing the bitterness of a liquid solution increases the measured behavioral aversiveness but reduces the dopamine activation (Fig. 4) (Fiorillo et al. 2013b). Purely aversive stimuli with low intensity induce only depressant responses, reflecting either negative punisher value or negative prediction error from reward omission (Fiorillo 2013). Thus, 'aversive' dopamine activations may represent the initial, unselective dopamine response component. Correspondingly, this activation is increased by reward generalization (Mirenowicz & Schultz 1996). Recent rediscoveries of 'aversive' dopamine activations have not considered these physical and generalization confounds, which may also explain the paradoxical, more frequent responses to conditioned than primary stimuli, nor have they checked for potential contribution from the rewarded context in which these tests were conducted (Joshua et al. 2008; Matsumoto & Hikosaka 2009). Stronger activations with higher punishment probability (Matsumoto & Hikosaka 2009) may derive from salience differences or anticipatory relief with punishment termination or avoidance (Budygin et al. 2012; Oleson et al. 2012) that are considered rewarding (Solomon and Corbit 1974; Gerber et al. 2014). Thus, when all confounds are accounted for, dopamine neurons are not activated by the aversive value of punishers (Fiorillo et al. 2013). Response variations among neurons may reflect graded sensitivity differences of the initial response component to physical intensity, generalization and context (Fiorillo et al. 2013a), rather than two, categorically distinct types of dopamine neurons (Matsumoto & Hikosaka 2009). Thus, aversive stimuli may induce the initial, unselective dopamine response component through their salience, but neither the first nor the second dopamine response component seems to code punishment. Nevertheless, as neighboring non-dopamine neurons code punishment (Cohen et al. 2012), a few dopamine neurons might receive aversively coding collateral afferents.

Sufficient and necessary functions in learning and approach

Electrical and optogenetic activation of dopamine neurons elicits place preference, nose poking, lever pressing, choice preferences, spatial navigation, rotation, locomotion and

unblocking of learning (Corbett & Wise 1980; Tsai et al. 2009; Witten et al. 2011; Steinberg et al. 2013; Arsenault et al. 2014). Correspondingly, optogenetically induced direct or transsynaptic inhibition of dopamine neurons induces place dispreference learning (Tan et al. 2012; van Zessen et al. 2012; Ilango et al. 2014). The activation and inhibition of dopamine neurons apparently mimics positive and negative reward prediction error signals, suggesting a sufficient role of phasic dopamine signals in learning and approach. Habenula stimulation induces place dispreference, either by activating supposedly aversive dopamine neurons (Lammel et al. 2012) or by transsynaptically inhibiting dopamine neurons (Stopper et al. 2015), although the limited specificity of the employed TH:cre mice (Lammel et al. 2015) and the known habenula inhibition of dopamine neurons (Matsumoto & Hikosaka 2007) make the latter mechanism more likely. Dopamine receptor stimulation seems also necessary for these functions. Systemic dopamine receptor antagonists reduce simple stimulus-reward learning in rats (Flagel et al. 2011), and local D1 receptor antagonist injections into frontal cortex impair behavioral and neuronal learning and memory in monkeys (Sawaguchi & Goldman-Rakic 1991; Puig & Miller 2012). NMDA receptor knock out reduces dopamine burst responses, simple learning and behavioral reactions (Zweifel et al. 2009). Thus, the phasic dopamine signal is both causal and necessary for inducing learning and approach, allowing for the possibility that natural dopamine activations influence these behaviors.

Through diverging axonal projections, the dopamine response may act as a global reinforcement signal within a triad of low affinity D1 receptors located on dendritic spines of striatal neurons that are also contacted by cortical axons (Freund et al. 1984; Schultz 1998). D1 receptor activation is necessary for long-term potentiation in striatum (Pawlak & Kerr 2008). Stimulation of dopamine D1 receptors prolongs striatal membrane depolarizations (Hernández-López et al. 1997), which may underlie the immediate focusing effect of dopamine on behavior. Thus, the synaptic dopamine actions, despite their heterogeneity (Threlfell et al. 2012; Roeper 2013; Chuhma et al. 2014), are overall consistent with the behavioral dopamine functions in learning and approach.

SUBJECTIVE VALUE AND ECONOMIC UTILITY

Objective or subjective value

The proximal function of reward is survival and reproduction of the individual gene carrier. Thus, reward value is subjective. What is important for me may not be that important for everybody. Thus, the objective, physical value a reward has for an individual, like liquid volume or money, depends on how much it enhances her well being. A millionaire walks less readily a mile to save a £ 2 bus fare than a student. Although rats normally don't drink salt solutions, they will do so after salt deprivation (Robinson & Berridge 2013). Thus, although physically bigger rewards are usually better, their subjective value increases to varying extents depending on the needs of the individual decision maker.

Subjective value signal

The phasic dopamine value response increases with increasing objective, physical reward parameters including amount, probability and statistically expected value (Fiorillo et al.

2003; Tobler et al. 2005). However, physical reward value depends also on the molecules in the reward, which differ between different reward types and substances. Therefore, the values of different reward objects are inherently difficult to assess from physical properties, and subjective reward values inferred from behavioral choices should be used as regressors for neuronal responses. Indeed, choices reveal subjective preference rankings between rewards that cannot be inferred from objective, physical parameters or do not share common dimensions, such as different or multi-component reward types. Dopamine responses follow closely the ranked preferences among liquid and food rewards (Fig. 5A) (Lak et al. 2014) and reflect the arithmetic sum of positive and negative values from rewards and punishers (Fiorillo 2013). Furthermore, risk affects subjective value. Most monkeys are risk seekers with small rewards, preferring risky over safe rewards of same expected value, whereas others are risk neutral. Dopamine responses follow closely the subjective risk-enhanced subjective value during risk seeking behavior and, correspondingly, are stronger with binary, equiprobable gambles than with safe rewards of identical expected value (Fig. 5B) (Lak et al. 2014). By contrast, dopamine responses are unaffected by risk during risk neutral behavior. Likewise, striatal voltammetric dopamine responses are higher or lower to identical gambles than to safe rewards depending on risk seeking or avoiding attitudes of rats (Sugam et al. 2012). Thus, closely corresponding to behavioral choices, dopamine neurons signal the subjective value derived from different rewards and from risk on a common currency scale. However, these behavioral methods do not allow to derive subjective value as a function of objective value in a straightforward manner.

Temporal discounting

Rewards delivered with a longer delay after a stimulus or an action lose their value. Psychophysical assessment of intertemporal choices between early and delayed rewards reveal hyperbolic or exponential decay of subjective reward value over delays of 2 to 16 s (Fig. 5C blue) (Kobayashi & Schultz 2008), thus relating subjective value to the physical measure of delay. Correspondingly, monkey dopamine responses to delay-predicting stimuli decay progressively, despite constant physical reward amount (Fig. 5C red), as do voltammetric dopamine responses in rat nucleus accumbens (Day et al. 2010). Dopamine prediction error responses at reward time increase correspondingly with longer delays. Thus, dopamine neurons code subjective value as a mathematical function of objective delay. However, temporal discounting is a specific case that does not afford a general method for deriving subjective value from objective value irrespective of time.

Formal economic utility

Expected Utility Theory makes one further and important step in the construction of subjective value. Whereas the choice preferences described so far allow only the ranking of subjective values, and temporal discounting is specific for delays, utility functions derived from choices under risk define numeric subjective value as mathematical functions of objective, physical reward. In providing numerically defined interval ratios, utility functions constitute mathematical approximations of the preference structure with unique shapes that are meaningful up to positive affine transformations ($y=a+bx$) (von Neumann and Morgenstern 1944; Savage 1954; Debreu 1959; Kagel et al. 1995). This formal definition of utility does not rely on other factors influencing subjective value such as reward type, delay

or effort cost, which are all contributors to utility but not essential in this construction. Thus, the expected utility model is a highly constrained, well-defined form of subjective value measured solely from choices under risk. The nonlinear relationship between utility, $u(x)$, and objective value, x , derives from increasing or decreasing marginal utility, which is the utility derived from one additional unit of consumption (Bernoulli, 1954). Utility functions with these numerical properties can be compared with neuronal responses to rewards in order to identify neuronal utility functions $n(x)$ that code reward in these strict quantitative economic terms.

Neuronal utility signal

Methods for estimating cardinal utility functions are based on concepts employing gambles (von Neumann and Morgenstern 1944). Binary, equiprobable gambles ($p=0.5$ each outcome) constitute the most simple and controlled risk tests identified by economic decision theory (Rothschild and Stiglitz 1970). Using the fractile method with choices between risky and safe outcomes (Caraco 1980; Machina 1987), monkeys' choices reveal utility functions that are initially convex with small rewards (risk seeking, below 0.4–0.6 ml of blackcurrant juice), then linear, and then concave with larger amounts (risk avoiding) (Fig. 6A red), amounting to initially increasing and then decreasing marginal utility (Stauffer et al. 2014). Although these functions look similar among the limited number of monkeys tested, economic theory prohibits their quantitative comparison between individuals.

The dopamine reward prediction error signal reflects the nonlinear shape of measured utility functions (Stauffer et al. 2014). With free, unpredicted rewards, utility prediction errors increase in a nonlinear fashion defined by the measured utility function, and dopamine responses show a very similar nonlinear increase with increasing amounts of unpredicted reward and thus correlate well with those utility prediction errors (Fig. 6A). Binary gambles result in positive prediction errors for the larger reward outcome and negative errors for the smaller reward. In three gambles with the same physical range (± 0.15 ml), physically identical prediction errors induce nonlinear changes in utility (due to non-monotonic marginal utility) (Fig. 6B top, grey rectangles). Dopamine prediction error responses to the higher gamble outcomes mimic the non-monotonic changes in the utility slope and correlate well with the increments on the utility axis (Fig. 6C). With all other contributions to utility kept constant, including reward type, delay and cost, these data suggest the coding of a prediction error in income utility by dopamine neurons. Thus, the dopamine reward prediction error response is, specifically, a utility prediction error signal that reflects marginal utility. Thus, the strictest and most theory-constrained definition of reward value applies to the dopamine coding of reward.

Although utility is a measure of subjective value, a neuronal utility signal reflecting a mathematical function of objective value goes well beyond subjective value coding derived from simple behavioral preferences. Thus, the identification of a dopamine utility prediction error signal suggests that the enigmatic and elusive utility, whose mere existence is often questioned by economists, is implemented in the brain as a physical, measurable signal.

SUMMARY

The description of the phasic dopamine reward prediction error signal in substantia nigra and ventral tegmental area has advanced on two major points, the component structure of the signal and the form of reward information conveyed. Dopamine neurons show two response components, analogous to neurons in higher order sensory and cognitive brain regions. The initial component consists of a fast increase in activity (activation) that is unselective and highly sensitive to factors related to rewards. It detects all environmental stimuli of sufficient physical intensity, including rewards, unrewarded stimuli and punishers, and transiently codes several forms of salience until the second response component appears. Activations by punishers reflect physical impact rather than aversiveness. Through the early onset, unselectivity, high sensitivity and beneficial salience effects on neuronal processing, the initial response enhances the speed, efficacy and accuracy of reward processing, and thus reward acquisition. The second response component codes subjective reward value derived from different rewards and delays. Further tests using behavioral tools from economic decision theory suggest that this component constitutes a utility prediction error signal. With the dual component structure, the dopamine signal detects rewards at the earliest possible moment and allows behavioral initiation mechanisms to begin already while the rewards are still being identified; the subsequent value response occurs early enough to prevent confusion with unrewarded and aversive objects. This is a very interesting computational mechanism for efficiently reacting to external stimuli. Stimulation and lesion experiments suggest that the phasic dopamine signal is sufficient and necessary for driving behavioral actions. With its formal utility coding, the dopamine signal constitutes the first known neuronal utility signal that is directly compatible with basic foundations of economic decision theory; it constitutes a biological implementation of utility, which for economists is purely theoretical.

Acknowledgments

Supported by Wellcome Trust, European Research Council (ERC) and NIH Conte Center at Caltech.

References

- Ambroggi F, Ishikawa A, Fields HL, Nicola SM. Basolateral amygdala neurons facilitate reward-seeking behavior by exciting nucleus accumbens neurons. *Neuron*. 2008; 59:648–661. [PubMed: 18760700]
- Arsenault JT, Rima S, Stemmann H, Vanduffel W. Role of the primate ventral tegmental area in reinforcement and motivation. *Curr Biol*. 2014; 24:1347–1353. [PubMed: 24881876]
- Budygin EA, Park J, Bass CE, Grinevich VP, Bonin KD, Wightman RM. Aversive stimulus differentially triggers subsecond dopamine release in reward regions. *Neuroscience*. 2012; 201:331–337. [PubMed: 22108611]
- Bushnell MC, Goldberg ME, Robinson DL. Behavioral enhancement of visual responses in monkey cerebral cortex. I. Modulation in posterior parietal cortex related to selective visual attention. *J Neurophysiol*. 1981; 46:755–772. [PubMed: 7288463]
- Chuhma N, Mingote S, Moore H, Rayport S. Dopamine neurons control striatal cholinergic neurons via regionally heterogeneous dopamine and glutamate signaling. *Neuron*. 2014; 81:901–912. [PubMed: 24559678]
- Cohen JY, Haesler S, Vong L, Lowell BB, Uchida N. Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature*. 2012; 482:85–88. [PubMed: 22258508]

- Corbett D, Wise RA. Intracranial self-stimulation in relation to the ascending dopaminergic systems of the midbrain: A moveable microelectrode study. *Brain Res.* 1980; 185:1–15. [PubMed: 7353169]
- Day JJ, Jones JL, Wightman RM, Carelli RM. Phasic nucleus accumbens dopamine release encodes effort- and delay-related costs. *Biol Psychiat.* 2010; 68:306–309. [PubMed: 20452572]
- Day JJ, Roitman MF, Wightman RM, Carelli RM. Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. *Nat Neurosci.* 2007; 10:1020–1028. [PubMed: 17603481]
- de Lafuente O, Romo R. Dopamine neurons code subjective sensory experience and uncertainty of perceptual decisions. *Proc Natl Acad Sci.* 2011; 49:19767–19771. [PubMed: 22106310]
- DeLong MR, Crutcher MD, Georgopoulos AP. Relations between movement and single cell discharge in the substantia nigra of the behaving monkey. *J Neurosci.* 1983; 3:1599–1606. [PubMed: 6875659]
- Caraco T, Martindale S, Whitham TS. An empirical demonstration of risk-sensitive foraging preferences. *Anim Behav.* 1980; 28:820–830.
- Debreu G. Cardinal utility for even-chance mixtures of pairs of sure prospects. *Rev Econ Stud.* 1959; 26:174–177.
- Enomoto K, Matsumoto N, Nakai S, Satoh T, Sato TK, Ueda Y, Inokawa H, Haruno M, Kimura M. Dopamine neurons learn to encode the long-term value of multiple future rewards. *Proc Natl Acad Sci (USA).* 2011; 108:15462–15467. [PubMed: 21896766]
- Fiorillo CD. Two dimensions of value: Dopamine neurons represent reward but not aversiveness. *Science.* 2013; 341:546–549. [PubMed: 23908236]
- Fiorillo CD, Yun SR, Song MR. Diversity and homogeneity in responses of midbrain dopamine neurons. *J Neurosci.* 2013a; 33:4693–4709. [PubMed: 23486943]
- Fiorillo CD, Song MR, Yun SR. Multiphasic temporal dynamics in responses of midbrain dopamine neurons to appetitive and aversive stimuli. *J Neurosci.* 2013b; 33:4710–4725. [PubMed: 23486944]
- Fiorillo CD, Tobler PN, Schultz W. Discrete coding of reward probability and uncertainty by dopamine neurons. *Science.* 2003; 299:1898–1902. [PubMed: 12649484]
- Flagel SB, Clark JJ, Robinson TE, Mayo L, Czuj A, Willuhn I, Akers CA, Clinton SM, Phillips PE, Akil H. A selective role for dopamine in stimulus-reward learning. *Nature.* 2011; 469:53–57. [PubMed: 21150898]
- Freund TF, Powell JF, Smith AD. Tyrosine hydroxylase-immunoreactive boutons in synaptic contact with identified striatonigral neurons, with particular reference to dendritic spines. *Neuroscience.* 1984; 13:1189–1215. [PubMed: 6152036]
- Gerber B, Yarali A, Diegelmann S, Wotjak CT, Pauli P, Fendt M. Pain-relief learning in flies, rats, and man: basic research and applied perspectives. *Learn Mem.* 2014; 21:232–252. [PubMed: 24643725]
- Guarraci FA, Kapp BS. An electrophysiological characterization of ventral tegmental area dopaminergic neurons during differential pavlovian fear conditioning in the awake rabbit. *Behav Brain Res.* 1999; 99:169–179. [PubMed: 10512583]
- Hernández-López S, Bargas J, Surmeier DJ, Reyes A, Galarraga E. D1 receptor activation enhances evoked discharge in neostriatal medium spiny neurons by modulating an L-type Ca^{2+} conductance. *J Neurosci.* 1997; 17:3334–3342. [PubMed: 9096166]
- Horvitz JC, Stewart, Jacobs BL. Burst activity of ventral tegmental dopamine neurons is elicited by sensory stimuli in the awake cat. *Brain Res.* 1997; 759:251–258. [PubMed: 9221945]
- Ilango A, Kesner AJ, Keller KL, Stuber GD, Bonci A, Ikemoto S. Similar roles of substantia nigra and ventral tegmental dopamine neurons in reward and aversion. *J Neurosci.* 2014; 34:817–822. [PubMed: 24431440]
- Joshua M, Adler A, Mitelman R, Vaadia E, Bergman H. Midbrain dopaminergic neurons and striatal cholinergic interneurons encode the difference between reward and aversive events at different epochs of probabilistic classical conditioning trials. *J Neurosci.* 2008; 28:11673–11684. [PubMed: 18987203]
- Kagel, JH.; Battalio, RC.; Green, L. *Economic Choice Theory: An Experimental Analysis of Animal Behavior.* Cambridge: Cambridge University Press; 1995.

- Kapur S. Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *Am J Psychiatry*. 2003; 160:13–23. [PubMed: 12505794]
- Kim JN, Shadlen MN. Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nat Neurosci*. 1999; 2:176–185. [PubMed: 10195203]
- Kobayashi S, Schultz W. Influence of reward delays on responses of dopamine neurons. *J Neurosci*. 2008; 28:7837–7846. [PubMed: 18667616]
- Kobayashi S, Schultz W. Reward contexts extend dopamine signals to unrewarded stimuli. *Curr Biol*. 2014; 24:56–62. [PubMed: 24332545]
- Lak A, Arabzadeh E, Harris JA, Diamond ME. Correlated physiological and perceptual effects of noise in a tactile stimulus. *Proc Natl Acad Sci (USA)*. 2010; 107:7981–7986. [PubMed: 20385799]
- Lak A, Stauffer WR, Schultz W. Dopamine prediction error responses integrate subjective value from different reward dimensions. *Proc Natl Acad Sci (USA)*. 2014; 111:2343–2348. [PubMed: 24453218]
- Lammel S, Lim BK, Ran C, Huang KW, Betley MJ, Tye KM, Deisseroth K, Malenka RC. Input-specific control of reward and aversion in the ventral tegmental area. *Nature*. 2012; 491:212–217. [PubMed: 23064228]
- Lammel S, Steinberg E, Földy C, Wall NR, Beier K, Luo L, Malenka RC. Diversity of transgenic mouse models for selective targeting of midbrain dopamine neurons. *Neuron*. 2015; 85:429–438. [PubMed: 25611513]
- Ljungberg T, Apicella P, Schultz W. Responses of monkey dopamine neurons during learning of behavioral reactions. *J Neurophysiol*. 1992; 67:145–163. [PubMed: 1552316]
- Machina MJ. Choice under uncertainty: problems solved and unsolved. *J Econ Perspect*. 1987; 1:121–154.
- Matsumoto M, Hikosaka O. Lateral habenula as a source of negative reward signals in dopamine neurons. *Nature*. 2007; 447:1111–1115. [PubMed: 17522629]
- Matsumoto M, Hikosaka O. Two types of dopamine neuron distinctively convey positive and negative motivational signals. *Nature*. 2009; 459:837–841. [PubMed: 19448610]
- Matsumoto M, Takada M. Distinct representations of cognitive and motivational signals in midbrain dopamine neurons. *Neuron*. 2013; 79:1011–1024. [PubMed: 23932490]
- Mirenowicz J, Schultz W. Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature*. 1996; 379:449–451. [PubMed: 8559249]
- Mogami T, Tanaka K. Reward association affects neuronal responses to visual stimuli in macaque TE and perirhinal cortices. *J Neurosci*. 2006; 26:6761–6770. [PubMed: 16793883]
- Morris G, Nevet A, Arkadir D, Vaadia E, Bergman H. Midbrain dopamine neurons encode decisions for future action. *Nat Neurosci*. 2006; 9:1057–1063. [PubMed: 16862149]
- Nardo D, Santangelo V, Macaluso E. Stimulus-driven orienting of visuo-spatial attention in complex dynamic environments. *Neuron*. 2011; 69:1015–1028. [PubMed: 21382559]
- Nomoto K, Schultz W, Watanabe T, Sakagami M. Temporally extended dopamine responses to perceptually demanding reward-predictive stimuli. *J Neurosci*. 2010; 30:10692–10702. [PubMed: 20702700]
- Oleson EB, Gentry RN, Chioma VC, Cheer JF. Subsecond dopamine release in the nucleus accumbens predicts conditioned punishment and its successful avoidance. *J Neurosci*. 2012; 32:14804–14808. [PubMed: 23077064]
- Paton JJ, Belova MA, Morrison SE, Salzman CD. The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature*. 2006; 439:865–870. [PubMed: 16482160]
- Pearce JM, Hall G. A model for Pavlovian conditioning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol Rev*. 1980; 87:532–552. [PubMed: 7443916]
- Puig MV, Miller EK. The role of prefrontal dopamine D1 receptors in the neural mechanisms of associative learning. *Neuron*. 2012; 74:874–886. [PubMed: 22681691]
- Ratcliff R. A theory of memory retrieval. *Psychol Rev*. 1978; 83:59–108.
- Ratcliff R, McKoon G. The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Comput*. 2008; 20:873–922. [PubMed: 18085991]

- Redgrave P, Prescott TJ, Gurney K. Is the short-latency dopamine response too short to signal reward? *Trends Neurosci.* 1999; 22:146–151. [PubMed: 10203849]
- Rescorla, RA.; Wagner, AR. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: Black, AH.; Prokasy, WF., editors. *Classical Conditioning II: Current Research and Theory*. New York: Appleton Century Crofts; 1972. p. 64-99.
- Ringach DL, Hawken MJ, Shapley R. Dynamics of orientation tuning in macaque primary visual cortex. *Nature.* 1997; 387:281–284. [PubMed: 9153392]
- Robinson MJF, Berridge KC. Instant transformation of learned repulsion into motivational “wanting”. *Curr Biol.* 2013; 23:282–289. [PubMed: 23375893]
- Roelfsema PR, Tolboom M, Khayat PS. Different processing phases for features, figures, and selective attention in the primary visual cortex. *Neuron.* 2007; 56:785–792. [PubMed: 18054856]
- Romo R, Schultz W. Dopamine neurons of the monkey midbrain: Contingencies of responses to active touch during self-initiated arm movements. *J Neurophysiol.* 1990; 63:592–606. [PubMed: 2329363]
- Rothschild M, Stiglitz JE. Increasing risk: I. A definition. *J Econ Theory.* 1970; 2:225–243.
- Satoh T, Nakai S, Sato T, Kimura M. Correlated coding of motivation and outcome of decision by dopamine neurons. *J Neurosci.* 2003; 23:9913–9923. [PubMed: 14586021]
- Savage, LJ. *The Foundations of Statistics*. New York: Wiley; 1954.
- Sawaguchi T, Goldman-Rakic PS. D1 dopamine receptors in prefrontal cortex: involvement in working memory. *Science.* 1991; 251:947–950. [PubMed: 1825731]
- Schroeder, T. *Three Faces of Desire*. Boston: MIT Press; 2004.
- Schultz W. Responses of midbrain dopamine neurons to behavioral trigger stimuli in the monkey. *J Neurophysiol.* 1986; 56:1439–1462. [PubMed: 3794777]
- Schultz W. Predictive reward signal of dopamine neurons. *J Neurophysiol.* 1998; 80:1–27. [PubMed: 9658025]
- Schultz W. Multiple dopamine functions at different time courses. *Ann Rev Neurosci.* 2007; 30:259–288. [PubMed: 17600522]
- Schultz W, Dayan P, Montague RR. A neural substrate of prediction and reward. *Science.* 1997; 275:1593–1599. [PubMed: 9054347]
- Schultz W, Romo R. Responses of nigrostriatal dopamine neurons to high intensity somatosensory stimulation in the anesthetized monkey. *J Neurophysiol.* 1987; 57:201–217. [PubMed: 3559672]
- Schultz W, Romo R. Dopamine neurons of the monkey midbrain: Contingencies of responses to stimuli eliciting immediate behavioral reactions. *J Neurophysiol.* 1990; 63:607–624. [PubMed: 2329364]
- Schultz W, Ruffieux A, Aebischer P. The activity of pars compacta neurons of the monkey substantia nigra in relation to motor activation. *Exp Brain Res.* 1983; 51:377–387.
- Solomon RL, Corbit JD. An opponent-process theory of motivation. *Psychol Rev.* 1974; 81:119–145. [PubMed: 4817611]
- Stanisor L, van der Togt C, Cyriel MA, Pennartz CMA, Roelfsema PR. A unified selection signal for attention and reward in primary visual cortex. *Proc Natl Acad Sci (USA).* 2013; 110:9136–9141. [PubMed: 23676276]
- Stanton SJ, Mullette-Gillman O’DA, McLaurin RE, Kuhn CM, LaBar KS, Platt ML, Huettel SA. Low- and high-testosterone individuals exhibit decreased aversion to economic risk. *Psychol Sci.* 2011; 22:447–453. [PubMed: 21393575]
- Stauffer WR, Lak A, Schultz W. Dopamine reward prediction error responses reflect marginal utility. *Curr Biol.* 2014; 24:2491–2500. [PubMed: 25283778]
- Steinberg EE, Keiflin R, Boivin JR, Witten IB, Deisseroth K, Janak PH. A causal link between prediction errors, dopamine neurons and learning. *Nat Neurosci.* 2013; 16:966–973. [PubMed: 23708143]
- Steinfels GF, Heym J, Strecker RE, Jacobs BL. Behavioral correlates of dopaminergic unit activity in freely moving cats. *Brain Res.* 1983; 258:217–228. [PubMed: 6824912]

- Stopper CM, Tse MTL, Montes DR, Wiedman CR, Floresco SB. Overriding phasic dopamine signals redirects action selection during risk/reward decision making. *Neuron*. 2014; 84:177–189. [PubMed: 25220811]
- Sugam JA, Day JJ, Wightman RM, Carelli RM. Phasic nucleus accumbens dopamine encodes risk-based decision-making behavior. *Biol Psychiat*. 2012; 71:199–2015. [PubMed: 22055017]
- Sutton RS, Barto AG. Toward a modern theory of adaptive networks: expectation and prediction. *Psychol Rev*. 1981; 88:135–170. [PubMed: 7291377]
- Thompson KG, Hanes DP, Bichot NP, Schall JD. Perceptual and motor processing stages identified in the activity of macaque frontal eye field neurons during visual search. *J Neurophysiol*. 1996; 76:4040–4055. [PubMed: 8985899]
- Threlfell S, Lalic T, Platt NJ, Jennings KA, Deisseroth K, Cragg SJ. Striatal dopamine release is triggered by synchronized activity in cholinergic interneurons. *Neuron*. 2012; 75:58–64. [PubMed: 22794260]
- Tobler PN, Dickinson A, Schultz W. Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. *J Neurosci*. 2003; 23:10402–10410. [PubMed: 14614099]
- Tobler PN, Fiorillo CD, Schultz W. Adaptive coding of reward value by dopamine neurons. *Science*. 2005; 307:1642–1645. [PubMed: 15761155]
- Treue S, Maunsell JHR. Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature*. 1996; 382:539–541. [PubMed: 8700227]
- Tsai CT, Nakamura S, Iwama K. Inhibition of neuronal activity of the substantia nigra by noxious stimuli and its modification by the caudate nucleus. *Brain Res*. 1980; 195:299–311. [PubMed: 7397503]
- Tsai H-C, Zhang F, Adamantidis A, Stuber GD, Bonci A, de Lecea L, Deisseroth K. Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science*. 2009; 324:1080–1084. [PubMed: 19389999]
- van Zessen R, Phillips JL, Budygin EA, Stuber GD. Activation of VTA GABA neurons disrupts reward consumption. *Neuron*. 2012; 73:1184–1194. [PubMed: 22445345]
- von Neumann, J.; Morgenstern, O. *The Theory of Games and Economic Behavior*. Princeton: Princeton University Press; 1944.
- Waelti P, Dickinson A, Schultz W. Dopamine responses comply with basic assumptions of formal learning theory. *Nature*. 2001; 412:43–48. [PubMed: 11452299]
- Witten IB, Steinberg EE, Lee SY, Davidson TJ, Zalocusky KA, Brodsky M, Yizhar O, Cho SL, Gong S, Ramakrishnan C, Stuber GD, Tye KM, Janak PH, Deisseroth K. Recombinase-driver rat lines: tools, techniques, and optogenetic application to dopamine-mediated reinforcement. *Neuron*. 2011; 72:721–733. [PubMed: 22153370]
- Zweifel LS, Parker JG, Lobb CJ, Rainwater A, Wall VZ, Fadok JP, Darvas M, Kim MJ, Mizumori SJ, Paladini CA, Philipps PEM, Palmiter R. Disruption of NMDAR-dependent burst firing by dopamine neurons provides selective assessment of phasic dopamine-dependent behavior. *Proc Natl Acad Sci (USA)*. 2009; 106:7281–7288. [PubMed: 19342487]

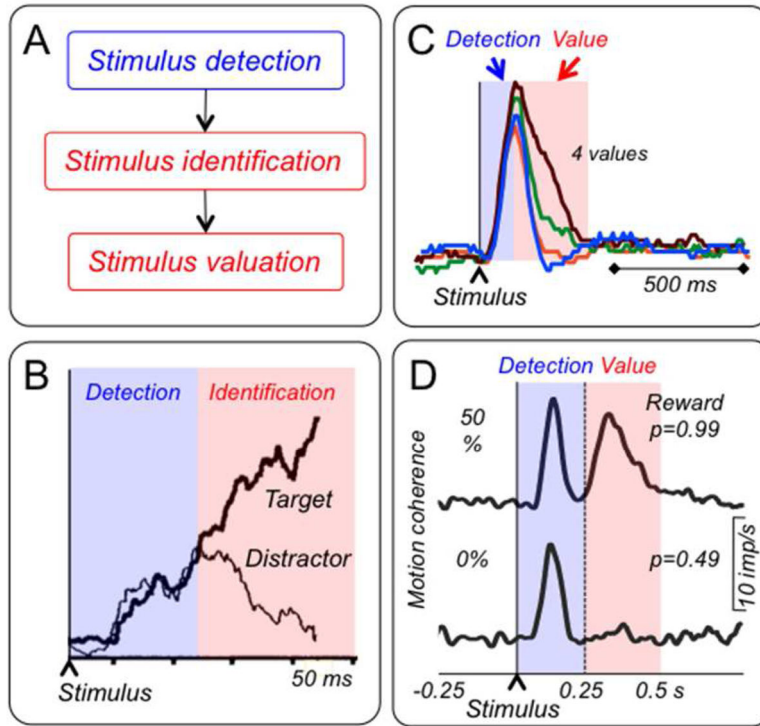


Figure 1. Stimulus components and their neuronal processing

A: Scheme of sequential processing steps of individual stimulus components.

B: Time course of target discrimination during visual search in monkey frontal eye fields neuron. The response initially detects the stimulus indiscriminately (blue zone) and only later differentiates between target and distractor (red). From Thompson et al. (1996).

C: Distinction of initial indiscriminate detection response (blue) from main response component coding reward prediction error (red) in monkey dopamine neurons during temporal discounting. Reward value increases from blue via orange and green to red, inversely with delays of 2, 4, 8 and 16 s. From Kobayashi & Schultz (2008).

D: Better distinction of the two dopamine response components in more demanding random dot motion discrimination task. Better dot motion discrimination with increasing motion coherence (0%, 50%) results in increasing reward probability (from $p=0.49$ to $p=0.99$). Neuronal activity shows an initial, non-differential increase (blue), a decrease back to baseline, and then a second, graded increase reflecting reward value (due to increasing reward probability, red). Vertical dotted line marks onset of discriminating ocular saccade and indicates that assessment of the reward value of the identified motion direction requires several hundred milliseconds. From Nomoto et al. (2010).

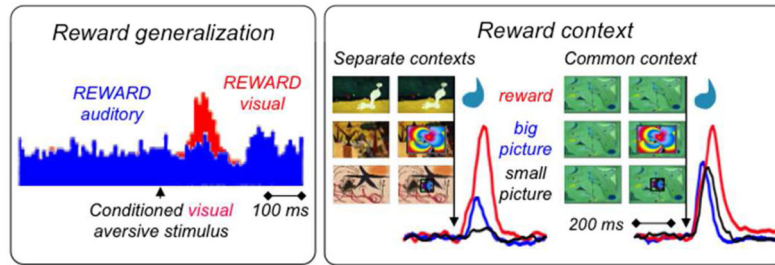


Figure 2. High sensitivity of initial dopamine detection response component

Left: Enhancement by reward generalization. In the red trials, both rewarded and aversive conditioned stimuli are visual. In the blue trials (covering large parts of red trials except the peak), the conditioned aversive stimulus remains visual, but the rewarded conditioned stimulus is auditory. The activating response to the identical visual aversive stimulus is higher when the rewarded stimulus is also visual (red peak) rather than auditory (blue), demonstrating response enhancement by sensory similarity with rewarded stimulus. The blue activity depression reflects the second component. From Mirenowicz & Schultz (1996).

Right: Enhancement by reward context. **Left:** in an experiment that separates unrewarded from rewarded contexts, dopamine neurons show only small activations to unrewarded large and small pictures (blue and black; red: response to liquid reward). Three distinct contexts are achieved by three well separated trial blocks, three different background pictures and removal of liquid spout in the picture trial types (center and bottom, blue and black). **Right:** by contrast, in an experiment using a common reward context without these separations, dopamine neurons show substantial activations to unrewarded large and small pictures. Each of the six picture pairs shows the trial background on a large computer monitor (left) and the continuing background together with the specific reward or superimposed picture (right). Each of the six neuronal traces shows the average population response from 31–33 monkey dopamine neurons. From Kobayashi & Schultz (2014).

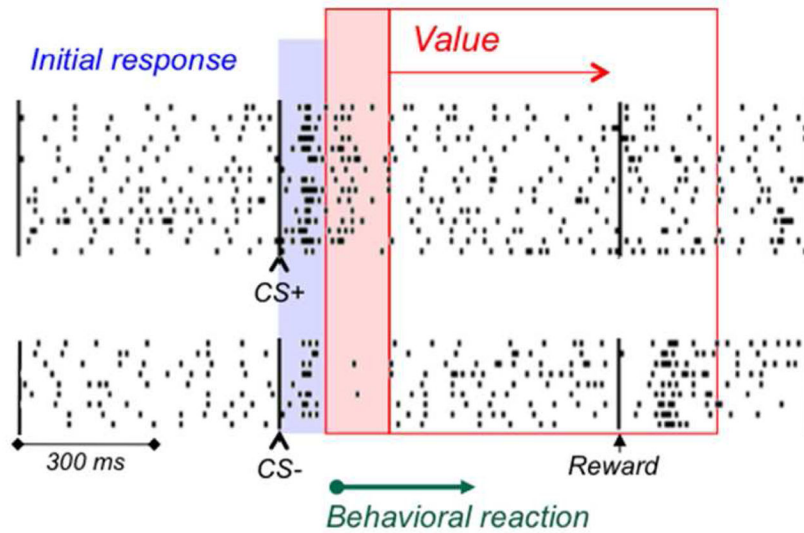


Figure 3. Accurate dopamine value coding after initial detection response. Both rewarded and unrewarded conditioned stimuli (CS+, CS-) elicit a common initial increase of neuronal activity (blue). This activation continues after the CS+ (top, red), but turns into a depression after the CS- (bottom, red). In CS+ trials, the fully predicted reward elicits no response (no prediction error, right), whereas in CS- trials, a surprising (identical) test reward induces an activation (positive prediction error). Thus, correct, positive or negative reward value coding begins immediately after the common initial response and early enough for initiating corresponding behavioral reactions (green arrow); correct value coding continues until the time of reward (red arrow). From Waelti et al. (2001).

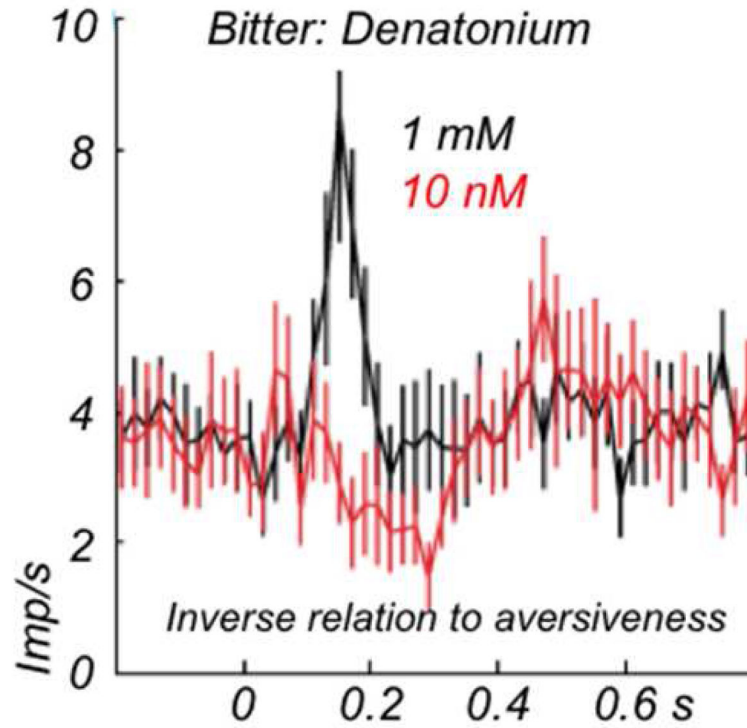


Figure 4.

Phasic activation of dopamine neurons to aversive stimulus reflects physical salience rather than aversiveness. The increased aversiveness generated by the more concentrated bitter deca-tonium solution decreases the dopamine response (physical impact of liquid delivery remains constant), suggesting an inverse relationship between aversiveness and dopamine activation. The increased depression from the higher aversiveness reduces the activation generated by the physical stimulation from the liquid drops. Average population responses from 19 and 14 monkey dopamine neurons, respectively. From Forillo et al. (2013b).

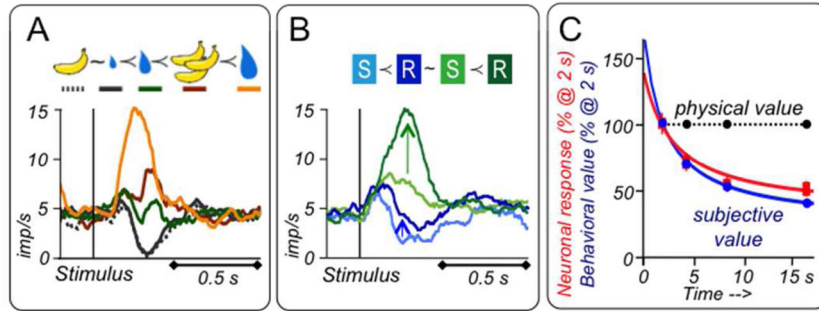


Figure 5. Dopamine neurons code subjective rather than objective reward value

A: Neuronal coding of common currency subjective value. Stimulus responses follow preferences among different liquid and food rewards. Rewards were different quantities of blackcurrant juice (top: blue) and liquified mixture of banana, chocolate and hazelnut food (yellow banana), color bars below rewards at top refer to color of neuronal responses, curved arrows indicate behavioral preferences assessed in binary behavioral choices between the indicated rewards.

B: Increase of stimulus responses with risky compared to safe rewards (vertical arrows). Blue and green colors indicate blackcurrant juice (more preferred = higher value) and orange juice (less preferred = lower value), respectively, S and R indicate safe reward amounts and binary, equiprobable gambles between two reward amounts of same reward juice with identical expected value, respectively. A and B from Lak et al. (2014).

C: Temporal discounting: decreasing responses of dopamine neurons to stimuli predicting increasing reward delays of 2–16 s (red), corresponding to subjective value decrements measured by intertemporal choices (blue), contrasted with constant physical amount (black). Y-axis shows behavioral value and neuronal responses in % of reward amount at 2 s delay (0.56 ml). From Kobayashi & Schultz (2008).

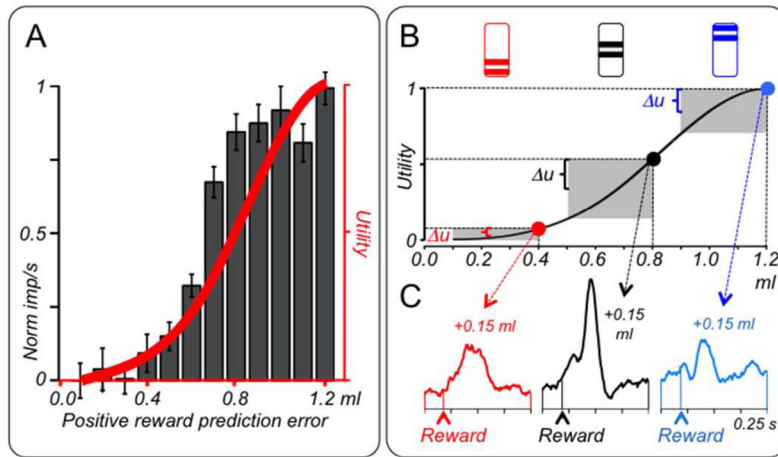


Figure 6. Dopamine neurons code formal economic utility

A: Positive utility prediction error responses to unpredicted juice rewards (black), superimposed on nonlinear utility function in same monkey (red). Psychophysically varied behavioral choices between a variable safe reward and a specific binary, equiprobable gamble ($p=0.5$ each outcome) served to assess its certainty equivalent (subjective value of gamble indicated by amount of safe reward at choice indifference); the certainty equivalents of specifically placed gambles served to estimate the utility function according to the structured 'fractile' procedure (Caraco 1980; Machina 1987).

B: Top: three conditioned stimuli indicating three binary, equiprobable gambles (0.1–0.4 ml; 0.5–0.8 ml; 0.9–1.2 ml juice); bar height specifies juice volume. In pseudorandom alternation, one of these stimuli is shown to the animal, followed 1.5 s later by one of the two specified juice volumes. Bottom: nonlinear utility function (same as in A). Delivery of higher reward in each gamble generates identical positive physical prediction error (0.15 ml, red, black and blue dots). However, due to different positions on the utility function, the prediction errors vary non-monotonically in utility (Δu). Shaded areas indicate physical volumes (horizontal) and utilities (vertical) of gambles.

C: Dopamine coding of utility prediction error (same animal as in B). The red, black and blue traces indicate responses to the higher outcomes of the three gambles shown as colored fat dots in B (0.4 ml, 0.8 ml, 1.2 ml). These responses reflect the positive utility prediction errors that vary according to the slope of the utility function (Δu in B), rather than the identical positive physical prediction errors of +0.15. A–C from Stauffer et al. (2014).