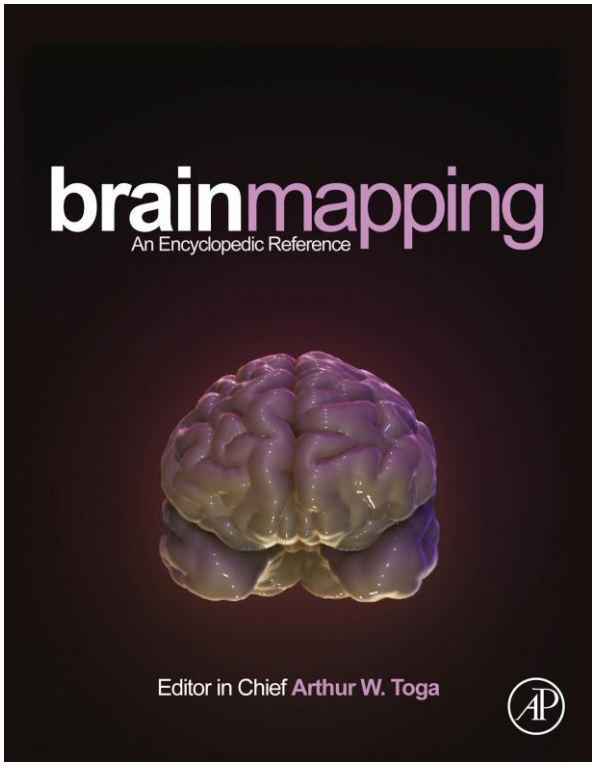

Schultz W. (2015) Reward. In: Arthur W. Toga, editor. *Brain Mapping: An Encyclopedic Reference*, vol. 2, pp. 643-651. Academic Press: Elsevier.

Reward

W Schultz, University of Cambridge, Cambridge, UK

Reward Functions

Behavioral Functions

The colloquial view regards reward as a bonus for exceptional performance, like chocolate for good school marks, or as something that makes us happy. However, the scientific investigation into reward functions requires a more complete definition, which animal learning theory and economic decision theory can provide. These theories distinguish three reward functions.

Learning

Rewards make us come back for more. This positive reinforcing function is captured by Pavlovian conditioning in which an intrinsically neutral stimulus becomes a reward predictor and elicits behavioral reactions through repeated association with reward, without requiring the subject to engage in particular behavior. The dog comes to salivate upon the ringing of a bell that had been repeatedly followed by a piece of sausage. Pavlovian conditioning establishes the essential reward predictions for making informed movements and choices. After having encountered a reward, we like to have more of it. This function is the essence of Thorndike's law of effect, which is implemented as operant conditioning in which rewards strengthen the behavior that led to the reward. Operant learning requires an active movement, like pressing a lever to obtain a food pellet. Pavlovian and operant conditioning constitute the two paradigmatic, elementary forms of associative learning about rewards.

The crucial requirement for conditioning is contingency. Pavlovian stimuli or operant actions are learned by pairing them with a reward, but only when the reward differs between the presence and absence of the event (Rescorla, 1967). The reward needs to be contingent, or dependent, on the event for learning to occur. Intuitively, we learn only events that inform us (e.g., about reward). Reward learning is driven by prediction errors, defined as the difference between experienced reward and predicted reward. Positive prediction errors lead to better reward prediction, negative prediction errors lead to less reward prediction, and no prediction errors (when nothing differs) result in no learning.

Approach and decision making

Rewards are attractive. They induce approach behavior, which brings us closer to them so that we can inspect and consume them. Rewards usually do not come alone and require choices. We choose the best reward, as if we are maximizing its value. Rewards are usually not fixed but vary a bit and are therefore formally described by probability distributions of reward values. Reward values are subjective, as they reflect our preferences that are not necessarily equal to the physical values. Subjective values are called utilities in economics. Choices described by utility theory can sometimes be inconsistent and better understood by prospect theory, which comprises different gain and loss slopes (loss aversion), subjective probability distortions, and dependency on reference points. Economic decisions involve comparisons between competing options and are captured by the winner-take-all (WTA) mechanism in which the option (object or action) with the highest subjective value wins.

Emotions

Rewards elicit positive emotions. Pleasure constitutes a transient response to a predicted or received reward and may lead to a longer-lasting state of happiness. The experience of rewards generates desire, the emotional underpinning of approach behavior. When I am thirsty and know that a glass of water helps, I desire it. Desire makes behavior purposeful and goal-directed. Thus, pleasure is the passive emotion derived from a reward, whereas desire helps to direct behavior toward a known reward. These emotions are also called liking (for pleasure) and wanting (for desire).

What makes an event rewarding?

The most basic rewards contain or predict necessary nutrients for survival, maintain our homeostatic balance, and elicit pleasure. Reproductive rewards may depend to some extent on hormones but are foremost pleasurable. Nonnutrient, nonreproductive rewards enhance our chance to obtain nutrients and to reproduce. Thus, events are rewarding because they fulfill objective homeostatic and reproductive needs and produce subjective pleasure. A few rewards, like maternal care, are neither homeostatic nor pleasurable but seem to be innate. Waking up the fifth time a night at 5 am when the baby cries neither resolves a homeostatic challenge nor is pleasurable.

Ultimate Evolutionary Function

The existing biological organisms result from evolutionary competition. The ultimate function of rewards is to increase evolutionary fitness by directing behavior toward the necessary nutrients for survival and toward successful reproduction. Only the best reward collectors will see their genes survive and propagate. Thus, the ultimate reward function is evolutionary fitness and gene propagation, which requires the behavioral reward functions of everyday life. That is why organisms have evolved that find foods, drinks, shelter, and the required economic exchanges attractive, as well as novelty, mates, and offspring.

Neuronal Reward Signals

Information processing systems use time-specific signals. In brains, the signals that propagate through the circuits are the action potentials generated by each neuron. The output of the system is the observable behavior. In between lie neurons and synapses with messenger molecules and membrane channels

Brain Mapping: An Encyclopedic Reference, (2015), vol. 2, pp. 643-651

that transmit and process the signals. Sensory and motor research shows that neurophysiological action potentials constitute crucial signals underlying behavioral processes. The postsynaptic responses to action potentials lead to blood oxygen level-dependent (BOLD) signals that are measured in humans by functional magnetic resonance imaging (fMRI). Lesions in humans and animals demonstrate the necessary involvement of specific brain structures in reward processing, and electrical and optogenetic stimulations help to dissect the behavioral role of reward signals. Thus, neurophysiological and BOLD signals transmit reward information compatible with the speed of behavior.

Rewards have several components. Their sensory component derives from stimulation of specific sensory receptors and reflects physical appearance, like color and form of an apple. It helps to differentiate between reward objects and distinguish them from all other environmental events. The attentional component enhances processing of rewards and helps to focus our behavior onto them. However, only the third, motivational component reflects the value of rewards for our well-being and survival and influences our daily economic decisions. Reward neurons process primarily the motivational component and distinguish it from the sensory and attentional components.

Theories about behavioral reward function and economic decisions are based on theoretical constructs rather than hardware receptors that would detect rewards in the environment and mediate choices. There are no reward receptors, and the brain constructs reward signals itself based on behavioral requirements. Thus, the lack of hardware at the input of the brain is overcome by identifiable neuronal reward signals. Then, characteristics of such signals can be compared with the theoretical constructs and, if successfully matched, considered as expression of implementation of the constructs. Once this occurs, the premises of the theories are basically validated, and the implementation of the theoretical construct constitutes guidance and constraint for applications of the theory. If we find stronger reward signals for more likely rather than less likely food, then reward probability seems to be converted into value, just as probability theory claims. If we identify a risk signal in the brain, risk seems to be a valid construct for explaining behavior that we cannot negate and should take into account when recommending risk avoidance strategies. Or finding that addictive drugs change the reward system via neuronal plasticity means that will power may be insufficient to deal with drug addiction and that prevention is more important.

Reward Neurons

Neurons in specific brain structures process rewards and economic decisions in various forms. The following sections present significant examples of neuronal reward signals, without trying to be exhaustive and systematic.

Dopamine Neurons

Most midbrain dopamine neurons show phasic reward prediction error responses (latency <100 ms, duration <200 ms). A better reward than predicted (positive prediction error) elicits an increase in the rate of action potentials (activation), a worse reward than predicted (negative error) induces a depression, and a fully predicted reward draws no response (**Figure 1(a)**) (Cohen, Haesler, Vong, Lowell, & Uchida, 2012; Enomoto et al., 2011; Fiorillo, Tobler, & Schultz, 2003; Schultz, 1998; Schultz, Apicella, & Ljungberg, 1993). The dopamine response transfers to the earliest reward predictive stimulus (**Figure 1(b)**). Dopamine neurons code prediction errors also in cognitive tasks with proper distinction of reward components (Matsumoto & Takada, 2013; Schultz et al., 1993). In coding the crucial prediction error term for associative learning, the phasic dopamine signal is a potential teaching signal.

The dopamine prediction error response is preceded by a brief activation that codes the physical impact of the stimulus and signals its detection before having identified its reward value (**Figure 1(c)**) (Nomoto, Schultz, Watanabe, & Sakagami, 2010). This response component increases with generalization to rewarded stimuli, rewarded contexts, and novelty (Kobayashi & Schultz, 2014; Schultz, 1998). It reflects the reward nature of the main prediction error response component and is likely to boost its efficacy as teaching signal by attentional enhancement and early detection of potential reward information. Despite assumptions of limited aversive activations (Brischoux, Chakraborty, Brierley, & Ungless, 2009; Matsumoto & Hikosaka, 2009; Mirenowicz & Schultz, 1996), choice tests dissociating physical impact from aversive value reveal that dopamine activations do not reflect aversiveness (**Figure 1(d)**) (Fiorillo, 2013). Indeed, electrical and optogenetic dopamine excitations induce learning, approach, and positive choice preferences, whereas dopamine inhibitions elicit dispreferences (Corbett & Wise, 1980; Steinberg et al., 2013; Tan et al., 2012; Tsai et al., 2009).

The dopamine error response integrates reward value from reward magnitude and probability (Tobler, Fiorillo, & Schultz, 2005) and reflects the subjective reward value integrated from different reward types, risks, and delays, as assessed by the animal's choices (**Figure 1(e)** and **1(f)**) (Kobayashi & Schultz, 2008; Lak, Stauffer, & Schultz, 2014). A separate slower dopamine activation reflects the degree of variance risk in reward (**Figure 1(g)**) (Fiorillo et al., 2003). Thus, dopamine neurons code separately the first two statistical moments of reward probability distributions, expected value and variance (risk), thus providing a biological correlate for these mathematical constructs.

The dopamine error signal occurs in anatomically, physiologically, and neurochemically heterogeneous neurons. However, the dopamine reward signal itself shows graded, noncategorical differences in latency, duration, and sensitivity and thus is rather homogeneous compared with other main reward neurons of the brain. Besides this phasic function, tonic extracellular dopamine plays a separate, permissive role in a wide range of motor and cognitive processes (Robbins & Arnsten, 2009). Thus, dopamine is a pluripotent neurochemical with rather homogeneous phasic reward function and separate heterogeneous tonic functions.

The dopamine teaching function for reinforcement learning involves widespread anatomical projections and postsynaptic plasticity. Distinct populations of striatal and

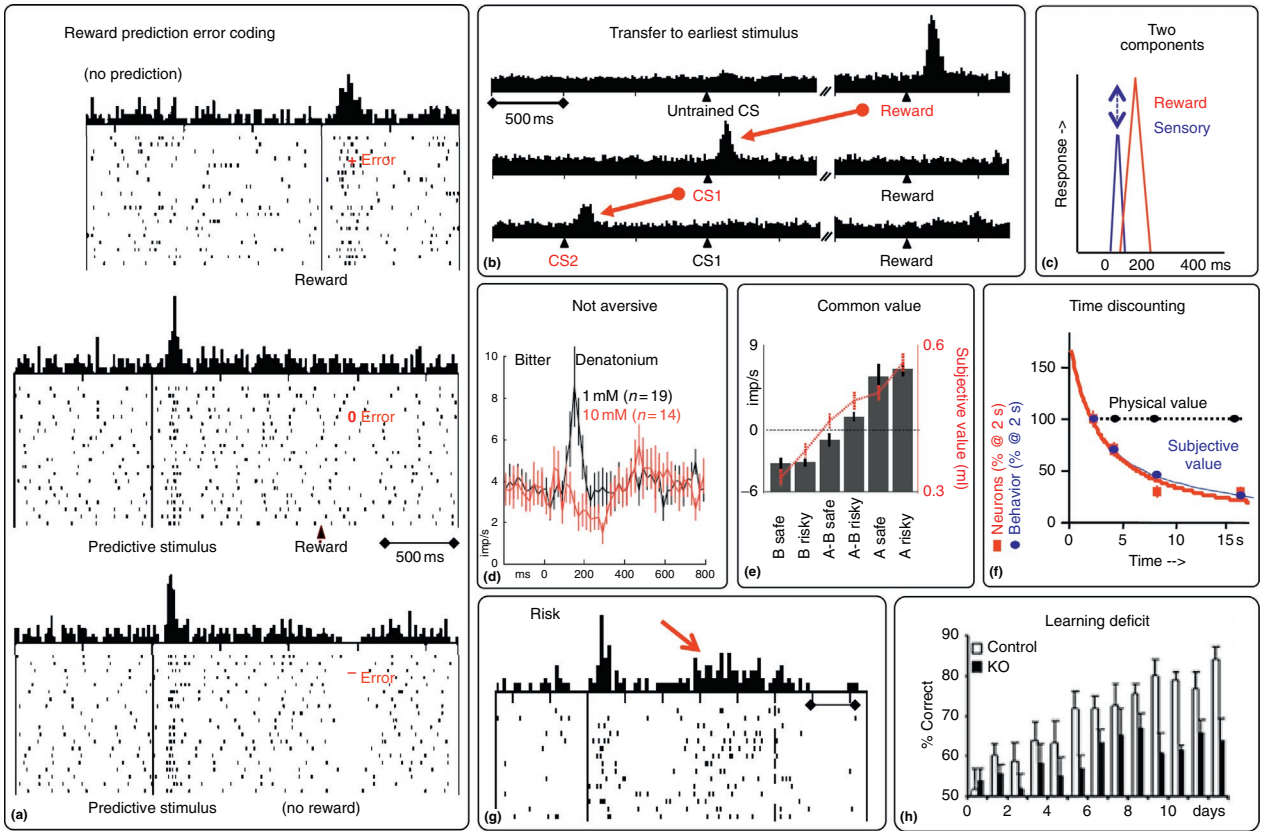

Figure 1 Neurophysiological dopamine responses to reward. (a) Reward prediction error responses at time of reward (right) and reward-predicting stimuli (left). Reproduced from Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology 80*, 1–27; Schultz et al. (1997). *Science, 275*, 1593–1599. (b) Transfer of dopamine response from reward to first reward-predicting stimulus. Reproduced from Schultz, W., Apicella, P., & Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *The Journal of Neuroscience 13*, 900–913; Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, *80*, 1–27. (c) Two components of dopamine activations to stimuli. The initial component (blue) codes sensory stimulus impact and reflects stimulus detection. It increases generalization to rewarded stimuli, rewarded contexts, and novelty (double arrow). The second component (red) codes reward value prediction error as detailed in (a). Reproduced from Kobayashi, S., & Schultz, W. (2014). Reward contexts extend dopamine signals to unrewarded stimuli. *Current Biology 24*, 56–62. (d) Dopamine activations induced by punishers do not reflect aversiveness but stimulus impact and may decrease with increasing aversive value (higher concentrated bitter solution). Reproduced from Fiorillo, C. D., Song, M. R., & Yun, S. R. (2013). Multiphasic temporal dynamics in responses of midbrain dopamine neurons to appetitive and aversive stimuli. *The Journal of Neuroscience 33,* 4710–4725. (e) Dopamine activations (red) follow subjective value integrated from different reward types and risk, as assessed by behavioral choices (black). Reward A = blackcurrant juice, B = strawberry juice, safe = 0.45 ml juice, risky = equiprobable gamble ($p = 0.5$ for 0.3 and 0.6 ml, and/or type). Reproduced from Lak, A., Stauffer, W. R., & Schultz, W. (2014). Dopamine prediction error responses integrate subjective value from different reward dimensions. *Proceedings of the National Academy of Sciences of the United States of America 111*, 2343–2348. (f) Dopamine activations (red) follow subjective value derived from temporal discounting, as assessed by intertemporal choices (blue), despite constant physical amount. Reproduced from Kobayashi, S., & Schultz, W. (2008). Influence of reward delays on responses of dopamine neurons. *The Journal of Neuroscience 28*, 7837–7846. (g) Slower dopamine risk activation. Lines in the raster display below the histogram show activity in single trials. Reproduced from Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science 299*, 1898–1902. (h) Learning deficit in T-maze with NMDA receptor knockout in midbrain dopamine neurons, impairing dopamine burst firing in mice. Reproduced from Zweifel, L. S., Parker, J. G., Lobb, C. J., Rainwater, A., Wall, V. Z., Fadok, J. P., Darvas, M., Kim, M. J., Mizumori, S. J., Paladini, C. A., Philipps, P. E. M., & Palmiter, R. (2009). Disruption of NMDAR-dependent burst firing by dopamine neurons provides selective assessment of phasic dopamine-dependent behavior. *Proceedings of the National Academy of Sciences of the United States of America 106*, 7281–7288.

cortical neurons show dopamine-dependent long-term potentiation and depression compatible with three-factor Hebbian learning (Gurden, Takita, & Jay, 2000; Pawlak & Kerr, 2008; Reynolds, Hyland, & Wickens, 2001; Shen, Flajolet, Greengard, & Surmeier, 2008). Given the rather homogeneous dopamine teaching signal, specificity for learning arises from differentially activated postsynaptic neurons. Phasic dopamine affects also immediate postsynaptic processing in the striatum, prolonging excitatory postsynaptic potentials via striatal D1 receptors and reducing striatal excitability via D2 receptors (Hernandez-Lopez, Bargas, Surmeier, Reyes, & Galarraga, 1997; Hernandez-Lopez et al., 2000).

Lesioning and psychopharmacological dopamine interventions induce learning deficits. Knockout of NMDA receptors reduces phasic dopamine activity and induces many learning deficits (**Figure 1(h)**) (Zweifel et al., 2009). Dopamine D1 receptor blockade in striatum impairs simple reward learning of stimuli that are capable of driving dopamine neurons, without affecting learning in tasks not eliciting dopamine responses (Flagel et al., 2011). Local D1 antagonist application impairs neuronal learning in monkey prefrontal cortex (Puig & Miller, 2012). The learning deficits are not explained by performance deficits (Flagel et al., 2011; Puig & Miller, 2012; Zweifel et al., 2009).

Frontal Cortex

Economic reward value is the crucial motivational component that makes rewards attractive. After detecting an event and identifying its sensory properties, coding its value constitutes the basis for reward processing. The first cortical reward neurons were discovered in the dorsolateral prefrontal cortex (**Figure 2(a)**) (Watanabe, 1996). Similar to dopamine neurons (Lak et al., 2014), some neurons in the orbitofrontal cortex code the value of the chosen reward while abstracting from its type and sensory properties, thus reporting the value from different rewards on a unified scale (Padoa-Schioppa & Assad, 2006).

The limited processing capacity of the brain contrasts with the huge number of possible rewards. Efficient reward discrimination would restrict neuronal processing to the currently available rewards. Indeed, some neuronal signals in orbitofrontal cortex reflect current reward selections. They code rewards relative to other available rewards (**Figure 2(b)**) (Tremblay & Schultz, 1999), adapt to the spread of reward distributions (Kobayashi, Pinto de Carvalho, & Schultz, 2010), or combine both mechanisms (Padoa-Schioppa, 2009). Less adaptive orbitofrontal neurons would maintain absolute value ranking to assure crucial transitivity for economic choices. Thus, efficacy combines with monotonic value coding for optimal reward processing.

The choice of goods is central to economic exchanges. The object value of a specific good is independent of the actual choice and thus does not reflect immediate reward expectation. Object value is coded by separate neurons for each good. In binary decisions about goods, two object value neurons representing the options compete against each other, and the highest valued object wins through a WTA mechanism (**Figure 2(c)**). Some orbitofrontal neurons code object value, irrespective of the animal's choice (**Figure 2(d)**) (Padoa-Schioppa & Assad, 2006). These activities suggest a neuronal implementation of a simple WTA mechanism for economic decisions and correspond to the function of orbitofrontal cortex in object rather than action processing.

Lesions of orbitofrontal cortex demonstrate its important involvement in risk processing. Indeed, some orbitofrontal neurons code the risk inherent in difficult odor choices (Kepecs, Uchida, Zariwala, & Mainen, 2008) and code variance risk, rather than value, in binary gambles between small and large rewards (**Figure 2(e)**) (O'Neill & Schultz, 2010). Due to their short latency, these responses process the risk in rewards well before a choice is being made. Thus, different from dopamine neurons, distinct orbitofrontal neurons implement the first two statistical moments of reward probability distributions.

Neurons in the anterior cingulate cortex are involved in social processes by distinguishing between own and other's reward or sensing a conspecific's reward (Chang, Gariépy, & Platt, 2013). Medial frontal cortex neurons respond to errors of conspecifics and the resulting reward omission (Yoshida, Saito, Iriki, & Isoda, 2012). Some movement mirror neurons in the premotor cortex distinguish between reward and no reward (Caggiano et al., 2012). These neurons mediate the distinction between own and others' rewards and thus code fundamental components of competition and cooperation. Indeed, prefrontal neurons differentiate between competitive and noncompetitive video games (Hosokawa & Watanabe, 2012).

Striatum

Neurons in the striatum (caudate nucleus, putamen, and ventral striatum) respond to reward-predicting stimuli and rewards (Apicella, Scarnati, Ljungberg, & Schultz, 1992; Hikosaka, Sakamoto, & Usui, 1989). They are also activated during reward expectation, often conjointly with activity related to the preparation of movements. Indeed, most task-related striatal neurons are affected by future reward (Hollerman, Tremblay, & Schultz, 1998).

Every economic choice ultimately requires an action. Action value refers to the reward obtained by the particular action and is independent of the actual choice, without reflecting imminent reward reception. In analogy to object value, neurons coding action value serve as inputs for competitive WTA mechanisms (**Figure 2(c)**). Action values are subjective and derived from reinforcement models or logistic regressions fitted to behavioral choices (Kim, Sul, Huh, Lee, & Jung, 2009; Samejima, Ueda, Doya, & Kimura, 2005). Striatal neurons code action values irrespective of chosen arm and eye movements (**Figure 2(f)**) (Kim et al., 2009; Samejima et al., 2005). The presence of action value neurons in a major motor structure such as the striatum suggests the implementation of a WTA decision mechanism closely to motor outputs.

Observation of social partners and comparison of their rewards mediate competition and cooperation, which improve performance and give individuals access to otherwise unobtainable resources. Social factors have reward functions. Rhesus monkeys find viewing body parts of conspecifics rewarding (Deaner, Khera, & Platt, 2005). Striatal reward neurons process primarily own rewards while distinguishing between own and conspecific's reward and between the social agents whose action leads to own reward (**Figure 2(g)**) (Báez-Mendoza, Harris, & Schultz, 2013). These neurons identify the social agent whose action leads to own reward.

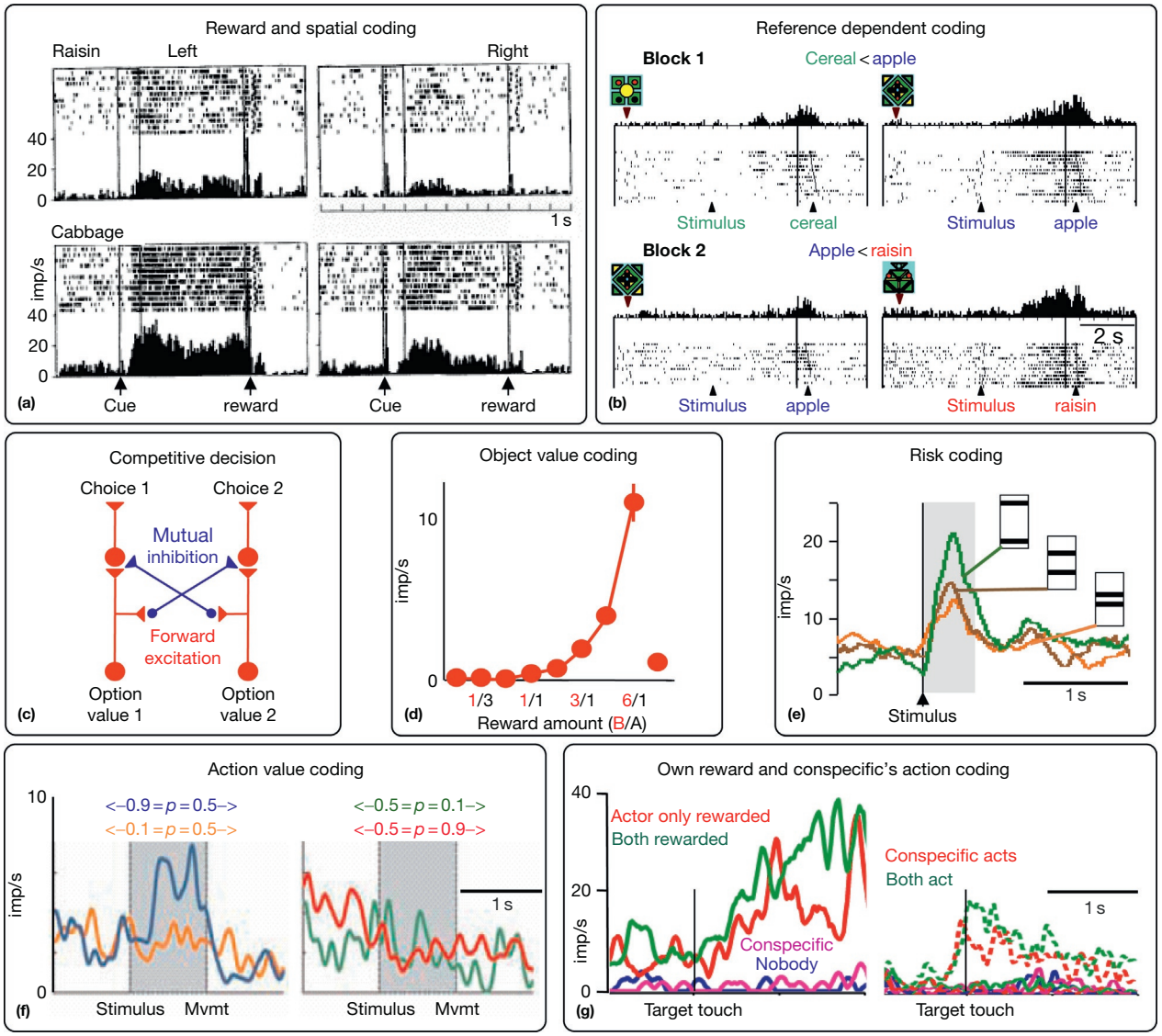

Figure 2 Reward signals in frontal cortex and striatum. (a) Conjoint coding of reward type (raisin or cabbage) and spatial information in single neuron of monkey dorsolateral prefrontal cortex during delayed response task performance. Reproduced from Watanabe, M. (1996). Reward expectancy in primate prefrontal neurons. *Nature 382*, 629–632. (b) Neuronal responses in orbitofrontal cortex adjust dynamically to currently available rewards (cereal and apple; apple and raisin). Arrows indicate behavioral preferences. The apple pieces, the predictive stimulus, and the required action were identical between top and bottom. Reproduced from Tremblay, L., & Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. *Nature 398*, 704–708. (c) Competitive winner-take-all (WTA) mechanism for decision making. Separate neurons coding the value of options (object value or action value) compete with each other. A threshold mechanism transforms the graded activities resulting from forward excitation and lateral inhibition to an all-or-none activation at the strongest output to mediate the corresponding choice. (d) Object value coding in orbitofrontal neuron. This neuron codes the value for juice B but not juice A (number of drops). Reproduced from Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature 441*, 223–226. (e) Risk coding in orbitofrontal neuron. The stimuli indicate two equiprobable reward amounts ($p = 0.5$ each) by bar heights. Vertical bar distance indicates risk. Reproduced from O'Neill, M., & Schultz, W. (2010). Coding of reward risk by orbitofrontal neurons is mostly distinct from coding of reward value. *Neuron 68*, 789–800. (f) Action value coding in striatum neuron. The premovement activity of this left action value neuron codes reward probability only in left movement trials ($p = 0.9$ compared to $p = 0.1$; left, blue vs. brown), but not in right trials (right, green vs. red). Reproduced from Samejima, K., Ueda, Y., Doya, K., & Kimura, M. (2005). Representation of action-specific reward values in the striatum. *Science 310*, 1337–1340. (g) Social reward and agent coding in striatum neurons. This neuron codes only own reward (left), and only when the conspecific acts to produce the reward (right). Reproduced from Báez-Mendoza, R., Harris, C., & Schultz, W. (2013). Activity of striatal neurons reflects social action and own reward. *Proceedings of the National Academy of Sciences of the United States of America 110*, 16634–16639.

Amygdala

The motivational function of the amygdala was long investigated with punishment and fear responses. Its neuronal reward function became recently appreciated through the identification of responses to rewards, reward-predicting stimuli, and reward prediction errors (Belova, Paton, Morrison, & Salzman, 2007; Bermudez & Schultz, 2010; Paton, Belova, Morrison, & Salzman, 2006). Amygdala reward responses are distinct from visual responses.

Many amygdala neurons are sensitive to the basic requirement of reward contingency. They respond when rewards occur more frequently during a stimulus compared to its absence. However, they fail to respond despite the same stimulus–reward pairing when the reward occurs also without the stimulus (background), in which case the stimulus is uninformative (**Figure 3(a)**) (Bermudez & Schultz, 2010). Reward reduction during stimulus absence, without changing stimulus–reward, resurrects the response. Thus, amygdala responses demonstrate the biological implementation of the theoretical construct of contingency, that is, crucial for learning.

The amygdala function in reward processing extends well beyond simple reward prediction. In economic choices, amygdala neurons process the future reward value early in the trial and switch within a few seconds to coding the abstract choice irrespective of the specific reward value chosen (**Figure 3(b)**) (Grabenhorst, Hernadi, & Schultz, 2012). These activities mediate the transition from value coding to choice and demonstrate an involvement in economic decisions.

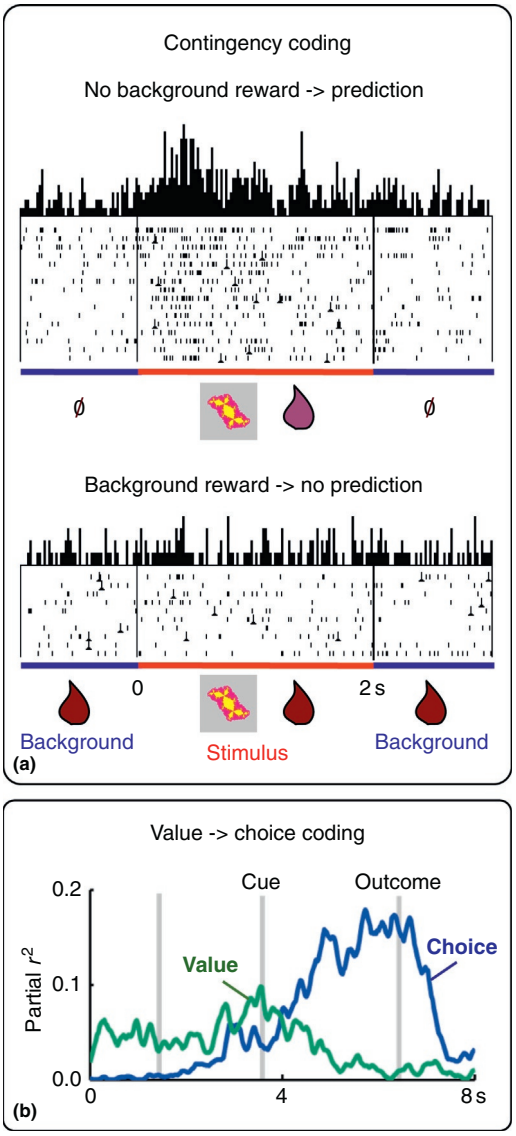

Figure 3 Reward and decision signals in primate amygdala neurons. (a) Sensitivity to reward contingency. Only stimuli associated with more reward compared to background are informative and elicit a neuronal response (top), whereas equal reward between stimulus and background renders the stimulus uninformative and the neuron unresponsive (bottom). Reproduced from Bermudez, M. A., & Schultz, W. (2010). Responses of amygdala neurons to positive reward-predicting stimuli depend on background reward (contingency) rather than stimulus–reward pairing (contiguity). *Journal of Neurophysiology 103*, 1158–1170. (b) Value to choice transition in averaged neuronal responses ($n = 12$). Ordinate shows coefficients (partial r^2) of running multiple linear regression. Reproduced from Grabenhorst, F., Hernadi, I., & Schultz, W. (2012). Prediction of economic choice by primate amygdala neurons. *Proceedings of the National Academy of Sciences of the United States of America 109*, 18950–18955.

Human Neuroimaging

Action potentials of individual neurons induce postsynaptic membrane potentials that affect metabolism and alter blood flow, which are measurable as BOLD signals from thousands of neurons. This method offers the possibility to translate the investigation of reward signals from animals to humans. Due to the postsynaptic nature of the BOLD signal, activations of dopamine neurons induce BOLD signals in the striatum and frontal cortex, whereas activations of cortical and striatal neurons lead to BOLD signals in the same structures via interneurons generating local postsynaptic potentials.

Reward Detection and Discrimination

Monetary rewards activate the human orbital and dorsolateral prefrontal cortex and the midbrain (**Figure 4(a)**) (Martin-Soelch et al., 2001; Thut et al., 1997). Monetary gains activate medial orbitofrontal cortex, whereas losses activate its lateral parts (O'Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001). Striatal BOLD responses distinguish reward predicting from neutral stimuli (**Figure 4(b)**) (Tobler, O'Doherty, Dolan, & Schultz, 2006) and code reward prediction errors reflecting dopamine input (Pessiglione, Seymour, Flandin, Dolan, & Frith, 2006).

Reward Value

Striatal BOLD signals increase with the magnitude and probability of monetary rewards (Tobler, O'Doherty, Dolan, & Schultz, 2007). Without distinguishing between these two variables, these signals code the expected value of probability distributions. BOLD signals in the striatum and ventromedial frontal cortex code the subjective value of monetary and food rewards (Levy & Glimcher, 2011) and reflect the subjective

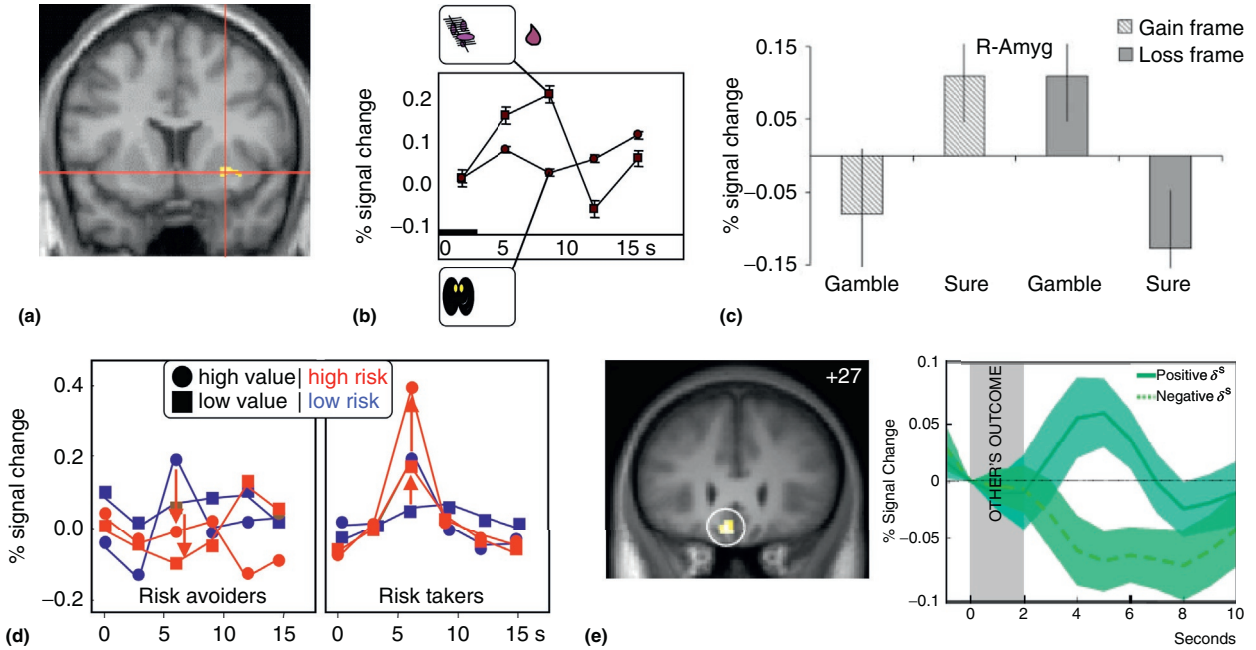

Figure 4 Human reward signals. (a) Activation of human ventral striatum by monetary reward (red cross, Positron Emission Tomography, PET). Reproduced from Martin-Soelch, C., Leenders, K. L., Chevalley, A. F., Missimer, J., Kunig, G., Magyar, S., Mino, A., & Schultz, W. (2001). Reward mechanisms in the brain and their role in dependence: Evidence from neurophysiological and neuroimaging studies. *Brain Research Reviews 36*, 139–149. (b) Differential BOLD response in putamen to juice reward-predicting and control stimuli. Reproduced from Tobler, P. N., O'Doherty, J. P., Dolan, R., & Schultz, W. (2006). Human neural learning depends on reward prediction errors in the blocking paradigm. *The Journal of Neurophysiology 95*, 301–310. (c) Amygdala BOLD response reflecting frame of reference. A sure gain of £20 out of a potential gain of £50 is valued positively, whereas a loss of £30 out of a potential gain of £50 is valued negatively, although both cases result in obtaining £20. Reproduced from De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. (2006). Frames, biases, and rational decision-making in the human brain. *Science 313*, 684–687. (d) Differential influence of risk on dorsolateral prefrontal value signal. Risk decreases BOLD value responses in risk avoiders and increases value responses in risk takers. Reproduced from Tobler, P. N., Christopoulos, G. I., O'Doherty, J. P., Dolan, R. J., & Schultz W. (2009). Risk-dependent reward value signal in human prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America 106*, 7185–7190. (e) Bidirectional BOLD responses in ventromedial frontal cortex to positive and negative reward prediction errors observed in another player during learning. Reproduced from Burke, C. J., Tobler, P. N., Baddeley, M., & Schultz, W. (2010) Neuronal mechanisms of observational learning. *Proceedings of the National Academy of Sciences of the United States of America 107*, 14431–14436.

value of food assessed by willingness to pay (Plassmann, O'Doherty, & Rangel, 2007). Temporal discounting, which decreases subjective but not objective value, reduces monetary BOLD signals in the striatum and ventromedial frontal cortex across widely varying delays (Gregorios-Pippas, Tobler, & Schultz, 2009; Kable & Glimcher, 2007). Monetary value coding depends on reference rewards in the striatum and prefrontal cortex (Nieuwenhuis et al., 2005) and varies according to win or loss frames of choice in the amygdala (**Figure 4(c)**) (De Martino, Kumaran, Seymour, & Dolan, 2006). Thus, reward signals incorporate basic notions of reward value and exemptions conceptualized by utility and prospect theories.

Reward Risk

BOLD signals in the striatum, orbital, and dorsolateral prefrontal cortex code variance risk distinct from striatal value signals (Preuschoff, Bossaerts, & Quartz, 2006; Tobler et al., 2007). In different parts of the dorsolateral prefrontal cortex, risk signals correspond closely to individual risk attitudes. Ambiguity, which elicits more pronounced behavioral attitudes than risk, induces stronger BOLD signals than risk in the amygdala and orbitofrontal cortex (Hsu, Bhatt, Adolphs, Tranel, & Camerer, 2005). Besides being coded as own variable, risk reduces prefrontal value signals in risk avoiders and increases value responses in risk seekers (**Figure 4(d)**) (Tobler, Christopoulos, O'Doherty, Dolan, & Schultz, 2009), thus reflecting the influence of risk on subjective value conceptualized by utility theory.

The observation and comparison of reward between individuals are particularly important and well developed in humans. We usually hate to receive less reward than others, all other factors being equal, called disadvantage inequity aversion, and we often feel guilty for getting more than others, called advantageous inequity aversion, unless we are in competition. Reward signals in human striatum decrease with disadvantageous inequity. The response to the same amount of money decreases when a conspecific receives twice as much (Fliessbach et al., 2007). Also, we learn from watching others. While participants benefit from observing probabilistic learning in conspecifics, BOLD signals in ventromedial cortex code the other's reward prediction errors (in addition to striatal signals for own reward prediction errors) (**Figure 4(e)**) (Burke, Tobler, Baddeley, & Schultz, 2010). These examples show only a selection of the many human neuroimaging studies on widely ranging social processes.

See also: **INTRODUCTION TO ANATOMY AND PHYSIOLOGY**: Amygdala; Basal Ganglia; Lateral and Dorsomedial Prefrontal Cortex and the Control of Cognition; **INTRODUCTION TO CLINICAL BRAIN MAPPING**: Emotion and Stress; **INTRODUCTION TO COGNITIVE NEUROSCIENCE**: Interactions between Attention and Emotion; Neuroimaging of Economic Decision-Making; Neuroimaging Studies of Reinforcement-Learning; Prediction and Expectation; Reward Processing; Salience/Bottom-Up Attention; Substantia Nigra; Uncertainty; Value Representation; **INTRODUCTION TO SOCIAL COGNITIVE NEUROSCIENCE**: Cooperation and Fairness; Emotion Perception and Elicitation; Emotion Regulation; Fairness and Inequity Aversion; Prosocial Motivation; Social Reward; The Amygdala and Social Perception; **INTRODUCTION TO SYSTEMS**: Emotion; Salience Network; Taste, Flavor, and Appetite.

References

Apicella, P., Scarnati, E., Ljungberg, T., & Schultz, W. (1992). Neuronal activity in monkey striatum related to the expectation of predictable environmental events. *Journal of Neurophysiology*, *68*, 945–960.

Báez-Mendoza, R., Harris, C., & Schultz, W. (2013). Activity of striatal neurons reflects social action and own reward. *Proceedings of the National Academy of Sciences of the United States of America*, *110*, 16634–16639.

Belova, M. A., Paton, J. J., Morrison, S. E., & Salzman, C. D. (2007). Expectation modulates neural responses to pleasant and aversive stimuli in primate amygdala. *Neuron*, *55*, 970–984.

Bermudez, M. A., & Schultz, W. (2010). Responses of amygdala neurons to positive reward predicting stimuli depend on background reward (contingency) rather than stimulus-reward pairing (contiguity). *Journal of Neurophysiology*, *103*, 1158–1170.

Brischoux, F., Chakraborty, S., Brierley, D. I., & Ungless, M. A. (2009). Phasic excitation of dopamine neurons in ventral VTA by noxious stimuli. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 4894–4899.

Burke, C. J., Tobler, P. N., Baddeley, M., & Schultz, W. (2010). Neuronal mechanisms of observational learning. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 14431–14436.

Caggiano, V., Fogassi, L., Rizzolatti, G., Casile, A., Giese, M. A., & Thier, P. (2012). Mirror neurons encode the subjective value of an observed action. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 11848–11853.

Chang, S. W. C., Gariépy, J.-F., & Platt, M. L. (2013). Neuronal reference frames for social decisions in primate frontal cortex. *Nature Neuroscience*, *16*, 243–250.

Cohen, J. Y., Haesler, S., Vong, L., Lowell, B. B., & Uchida, N. (2012). Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature*, *482*, 85–88.

Corbett, D., & Wise, R. A. (1980). Intracranial self-stimulation in relation to the ascending dopaminergic systems of the midbrain: A moveable microelectrode study. *Brain Research*, *185*, 1–15.

De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. (2006). Frames, biases, and rational decision-making in the human brain. *Science*, *313*, 684–687.

Deaner, R. O., Khera, A. V., & Platt, M. L. (2005). Monkeys pay per view: Adaptive valuation of social images by rhesus monkeys. *Current Biology*, *15*, 543–548.

Enomoto, K., Matsumoto, N., Nakai, S., Satoh, T., Sato, T. K., Ueda, Y., et al. (2011). Dopamine neurons learn to encode the long-term value of multiple future rewards. *Proceedings of the National Academy of Sciences of the United States of America*, *108*, 15462–15467.

Fiorillo, C. D. (2013). Two dimensions of value: Dopamine neurons represent reward but not aversiveness. *Science*, *341*, 546–549.

Fiorillo, C. D., Song, M. R., & Yun, S. R. (2013). Multiphasic temporal dynamics in responses of midbrain dopamine neurons to appetitive and aversive stimuli. *The Journal of Neuroscience*, *33*, 4710–4725.

Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty by dopamine neurons. *Science*, *299*, 1898–1902.

Flagel, S. B., Clark, J. J., Robinson, T. E., Mayo, L., Czuj, A., Willuhn, I., et al. (2011). A selective role for dopamine in stimulus-reward learning. *Nature*, *469*, 53–57.

Fliessbach, K., Weber, B., Trautner, P., Dohmen, T., Sunde, U., Elger, C. E., et al. (2007). Social comparison affects reward-related brain activity in the human ventral striatum. *Science*, *318*, 1305–1308.

Grabenhorst, F., Hernadi, I., & Schultz, W. (2012). Prediction of economic choice by primate amygdala neurons. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 18950–18955.

Gregorios-Pippas, L., Tobler, P. N., & Schultz, W. (2009). Short term temporal discounting of reward value in human ventral striatum. *Journal of Neurophysiology*, *101*, 1507–1523.

Gurden, H., Takita, M., & Jay, T. M. (2000). Essential role of D1 but not D2 receptors in the NMDA receptor-dependent long-term potentiation at hippocampal-prefrontal cortex synapses in vivo. *Journal of Neuroscience*, *20*, RC106, 1–5.

Hernandez-Lopez, S., Bargas, J., Surmeier, D. J., Reyes, A., & Galarraga, E. (1997). D1 receptor activation enhances evoked discharge in neostriatal medium spiny neurons by modulating an L-type Ca^{2+} conductance. *Journal of Neuroscience*, *17*, 3334–3342.

Hernandez-Lopez, S., Tkatch, T., Perez-Garci, E., Galarraga, E., Bargas, J., Hamm, H., et al. (2000). D2 dopamine receptors in striatal medium spiny neurons reduce L-type Ca^{2+} currents and excitability via a novel PLCβ1-IP_3-calcineurin-signaling cascade. *Journal of Neuroscience*, *20*, 8987–8995.

Hikosaka, O., Sakamoto, M., & Usui, S. (1989). Functional properties of monkey caudate neurons. III. Activities related to expectation of target and reward. *Journal of Neurophysiology*, *61*, 814–832.

Hollerman, J. R., Tremblay, L., & Schultz, W. (1998). Influence of reward expectation on behavior-related neuronal activity in primate striatum. *Journal of Neurophysiology*, *80*, 947–963.

Hosokawa, T., & Watanabe, M. (2012). Prefrontal neurons represent winning and losing during competitive video shooting games between monkeys. *Journal of Neuroscience*, *32*, 7662–7671.

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., & Camerer, C. F. (2005). Neural systems responding to degrees of uncertainty in human decision-making. *Science*, *310*, 1680–1683.

Kable, J. W., & Glimcher, P. W. (2007). The neural correlates of subjective value during intertemporal choice. *Nature Neuroscience*, *10*, 1625–1633.

Kepecs, A., Uchida, N., Zariwala, H., & Mainen, Z. F. (2008). Neural correlates, computation and behavioural impact of decision confidence. *Nature*, *455*, 227–231.

Kim, H., Sul, J. H., Huh, N., Lee, D., & Jung, M. W. (2009). Role of striatum in updating values of chosen actions. *Journal of Neuroscience*, *29*, 14701–14712.

Kobayashi, S., Pinto de Carvalho, O., & Schultz, W. (2010). Adaptation of reward sensitivity in orbitofrontal neurons. *Journal of Neuroscience*, *30*, 534–544.

Kobayashi, S., & Schultz, W. (2008). Influence of reward delays on responses of dopamine neurons. *Journal of Neuroscience*, *28*, 7837–7846.

Kobayashi, S., & Schultz, W. (2014). Reward contexts extend dopamine signals to unrewarded stimuli. *Current Biology*, *24*, 56–62.

Lak, A., Stauffer, W. R., & Schultz, W. (2014). Dopamine prediction error responses integrate subjective value from different reward dimensions. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 2343–2348.

Levy, D., & Glimcher, P. W. (2011). Comparing apples and oranges: Using reward-specific and reward-general subjective value representation in the brain. *Journal of Neuroscience*, *31*, 14693–14707.

Martin-Soelch, C., Leenders, K. L., Chevalley, A. F., Missimer, J., Kunig, G., Magyar, S., et al. (2001). Reward mechanisms in the brain and their role in dependence: Evidence from neurophysiological and neuroimaging studies. *Brain Research Reviews*, *36*, 139–149.

Matsumoto, M., & Hikosaka, O. (2009). Two types of dopamine neuron distinctively convey positive and negative motivational signals. *Nature*, *459*, 837–841.

Matsumoto, M., & Takada, M. (2013). Distinct representations of cognitive and motivational signals in midbrain dopamine neurons. *Neuron*, *79*, 1011–1024.

Mirenowicz, J., & Schultz, W. (1996). Preferential activation of midbrain dopamine neurons by appetitive rather than aversive stimuli. *Nature*, *379*, 449–451.

Nieuwenhuis, S., Heslenfeld, D. J., Alting van Geusau, N., Mars, R. B., Holroyd, C. B., & Yeung, N. (2005). Activity in human reward-sensitive brain areas is strongly context dependent. *NeuroImage*, *25*, 1302–1309.

Nomoto, K., Schultz, W., Watanabe, T., & Sakagami, M. (2010). Temporally extended dopamine response to perceptually demanding reward-predictive stimuli. *Journal of Neuroscience*, *30*, 10692–10702.

O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience*, *4*, 95–102.

O'Neill, M., & Schultz, W. (2010). Coding of reward risk by orbitofrontal neurons is mostly distinct from coding of reward value. *Neuron*, *68*, 789–800.

Padoa-Schioppa, C. (2009). Range-adapting representation of economic value in the orbitofrontal cortex. *Journal of Neuroscience*, *29*, 14004–14014.
Padoa-Schioppa, C., & Assad, J. A. (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature*, *441*, 223–226.
Paton, J. J., Belova, M. A., Morrison, S. E., & Salzman, C. D. (2006). The primate amygdala represents the positive and negative value of visual stimuli during learning. *Nature*, *439*, 865–870.
Pawlak, V., & Kerr, J. N. D. (2008). Dopamine receptor activation is required for corticostriatal spike-timing-dependent plasticity. *Journal of Neuroscience*, *28*, 2435–2446.
Pessiglione, M., Seymour, B., Flandin, G., Dolan, R. J., & Frith, C. D. (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*, *442*, 1042–1045.
Plassmann, H., O'Doherty, J., & Rangel, A. (2007). Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *Journal of Neuroscience*, *27*, 9984–9988.
Preuschoff, K., Bossaerts, P., & Quartz, S. R. (2006). Neural differentiation of expected reward and risk in human subcortical structures. *Neuron*, *51*, 381–390.
Puig, M. V., & Miller, E. K. (2012). The role of prefrontal dopamine D1 receptors in the neural mechanisms of associative learning. *Neuron*, *74*, 874–886.
Rescorla, R. A. (1967). Pavlovian conditioning and its proper control procedures. *Psychological Review*, *74*, 71–80.
Reynolds, J. N. J., Hyland, B. I., & Wickens, J. R. (2001). A cellular mechanism of reward-related learning. *Nature*, *413*, 67–70.
Robbins, T. W., & Arnsten, A. F. T. (2009). The neuropsychopharmacology of fronto-executive function: Monoaminergic modulation. *Annual Review of Neuroscience*, *32*, 267–287.
Samejima, K., Ueda, Y., Doya, K., & Kimura, M. (2005). Representation of action-specific reward values in the striatum. *Science*, *310*, 1337–1340.
Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology*, *80*, 1–27.
Schultz, W., Apicella, P., & Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *Journal of Neuroscience*, *13*, 900–913.
Shen, W., Flajolet, M., Greengard, P., & Surmeier, D. J. (2008). Dichotomous dopaminergic control of striatal synaptic plasticity. *Science*, *321*, 848–851.
Steinberg, E. E., Keiflin, R., Boivin, J. R., Witten, I. B., Deisseroth, K., & Janak, P. H. (2013). A causal link between prediction errors, dopamine neurons and learning. *Nature Neuroscience*, *16*, 966–973.
Tan, K. R., Yvon, C., Turiault, M., Mirzabekov, J. J., Doehner, J., Labouèbe, G., et al. (2012). GABA neurons of the VTA drive conditioned place aversion. *Neuron*, *73*, 1173–1183.
Thut, G., Schultz, W., Roelcke, U., Nienhusmeier, M., Missimer, J., Maguire, R. P., et al. (1997). Monetary reward activates human prefrontal cortex. *Neuroreport*, *8*, 1225–1228.
Tobler, P. N., Christopoulos, G. I., O'Doherty, J. P., Dolan, R. J., & Schultz, W. (2009). Risk-dependent reward value signal in human prefrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 7185–7190.
Tobler, P. N., Fiorillo, C. D., & Schultz, W. (2005). Adaptive coding of reward value by dopamine neurons. *Science*, *307*, 1642–1645.
Tobler, P. N., O'Doherty, J. P., Dolan, R., & Schultz, W. (2006). Human neural learning depends on reward prediction errors in the blocking paradigm. *Journal of Neurophysiology*, *95*, 301–310.
Tobler, P. N., O'Doherty, J. P., Dolan, R., & Schultz, W. (2007). Reward value coding distinct from risk attitude-related uncertainty coding in human reward systems. *Journal of Neurophysiology*, *97*, 1621–1632.
Tremblay, L., & Schultz, W. (1999). Relative reward preference in primate orbitofrontal cortex. *Nature*, *398*, 704–708.
Tsai, H.-C., Zhang, F., Adamantidis, A., Stuber, G. D., Bonci, A., de Lecea, L., et al. (2009). Phasic firing in dopaminergic neurons is sufficient for behavioral conditioning. *Science*, *324*, 1080–1084.
Watanabe, M. (1996). Reward expectancy in primate prefrontal neurons. *Nature*, *382*, 629–632.
Yoshida, K., Saito, N., Iriki, A., & Isoda, M. (2012). Social error monitoring in macaque frontal cortex. *Nature Neuroscience*, *15*, 1307–1312.
Zweifel, L. S., Parker, J. G., Lobb, C. J., Rainwater, A., Wall, V. Z., Fadok, J. P., et al. (2009). Disruption of NMDAR-dependent burst firing by dopamine neurons provides selective assessment of phasic dopamine-dependent behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 7281–7288.