Review

How Reward and Aversion Shape Motivation and Decision Making: A Computational Account

Jeroen P. H. Verharen[1,2] (iD), **Roger A. H. Adan**[1,3]*, **and Louk J. M. J. Vanderschuren**[2]*

Abstract

Processing rewarding and aversive signals lies at the core of many adaptive behaviors, including value-based decision making. The brain circuits processing these signals are widespread and include the prefrontal cortex, amygdala and striatum, and their dopaminergic innervation. In this review, we integrate historic findings on the behavioral and neural mechanisms of value-based decision making with recent, groundbreaking work in this area. On the basis of this integrated view, we discuss a neuroeconomic framework of value-based decision making, use this to explain the motivation to pursue rewards and how motivation relates to the costs and benefits associated with different courses of action. As such, we consider substance addiction and overeating as states of altered value-based decision making, in which the expectation of reward chronically outweighs the costs associated with substance use and food consumption, respectively. Together, this review aims to provide a concise and accessible overview of important literature on the neural mechanisms of behavioral adaptation to reward and aversion and how these mediate motivated behaviors.

Keywords

Reward and Aversion

In order to survive and flourish in a competitive world, an organism must learn to repeat actions that have proven profitable and avoid actions that have not. In this way, one learns to adapt its behavior in a changeable environment, in order to optimally promote survival. For example, it is sensible to revisit a place that is rich in foods, but not when this same place is swarming with predators. By incorporating these positive (food) and negative (predator) experiences into a value representation of stimuli in the surrounding world, one can enjoy rewards, such as food and sex, without experiencing potentially life-threatening dangers. These value representations, and the repeated updating of these values based on each action's outcome, are important drivers of decision-making processes that living organisms encounter numerous times each day.

Adapting behavior in response to positive and negative experiences is driven by a learning process called operant conditioning or instrumental learning. First stated by Thorndike (1898), and later refined by Skinner (1938) (see also Box 1), is the notion that cats, pigeons, and rats tend to increase the frequency of a certain behavior when this behavior is reinforced—either by the delivery of something pleasant (positive reinforcement) or the removal of something aversive (negative reinforcement). Conversely, a punisher is the adverse consequence of an action that decreases the probability of that action being taken again. This punisher can be explicit, such as pain (positive punishment), or implicit, such as the omission of an expected reward (negative punishment) (for an overview of the terminology on punishment see Jean-Richard-Dit-Bressel and others 2018).

[1]Brain Center Rudolf Magnus, Department of Translational Neuroscience, University Medical Center Utrecht, Utrecht, Netherlands
[2]Department of Animals in Science and Society, Division of Behavioural Neuroscience, Faculty of Veterinary Medicine, Utrecht University, Utrecht, Netherlands
[3]Institute of Physiology and Neuroscience, Sahlgrenska Academy at the University of Gothenburg, Gothenburg, Sweden

*These authors contributed equally.

Corresponding Author:
Jeroen P. H. Verharen, Brain Center Rudolf Magnus, Department of Translational Neuroscience, University Medical Center Utrecht, Universiteitsweg 100, 3584CG Utrecht, Netherlands.
Email: jeroenverharen@gmail.com

Box 1. A Brief History of Research on Reward, Aversion, Motivation, and Decision Making.

1848	**Harlow** publishes the case report on Phineas Gage, providing the first evidence for a role of the prefrontal cortex in executive behaviors, including decision making.
1898	In his *Law of Effect*, **Thorndike** states that animals learn through trial and error, an important step in the postulation of operant conditioning theory.
1927	**Pavlov** formulates his associative learning theory on the basis of his legendary dog experiment.
1938	**Skinner** publishes *The Behavior of Organisms*, including the influential theory on operant conditioning.
1946	**Tolman** challenges earlier conditioning theories by stating that learning can also occur in the absence of reward or punishment (i.e., stimulus-stimulus learning).
1954	**Olds and Milner** discover that rats will work for electrical stimulation of certain brain areas, a phenomenon now known as intracranial self-stimulation.
1972	Publication of the influential reinforcement learning theory of **Rescorla and Wagner**.
1981	**Sutton and Barto** publish computational models that explain temporal difference learning.
1982	**Adams and Dickinson** perform a set of experiments in rats that demonstrate a distinction between goal-directed and habitual behavior.
1997	The first measurement of reward prediction error signals in dopamine neurons of monkeys by **Schultz**.
1998	**Berridge and Robinson** propose their incentive salience theory of dopamine function, introducing the dichotomy between "wanting" and "liking."
2007	**Boyden, Deisseroth, Roth** and others develop viral tools to record and manipulate brain activity with hitherto impossible precision: start of the era of neural circuit dissection.

Thorndike described his theory in his *Law of Effect* (Thorndike 1898), after observing that a cat that is restrained in a box gradually learns how to escape by trial and error. Forty years after Thorndike's experiments, Skinner set the stage for the next decades of experimental psychological research by theorizing operant conditioning in his book *The Behavior of Organisms* (Skinner 1938) and the development of the now widely used operant conditioning chambers (hence often termed "Skinner boxes"). Although his theory was more formally postulated than Thorndike's, the idea behind it remained the same: behavior that is reinforced will be repeated, and behavior that is punished will cease (for a historic overview of their definitions of punishment, see Holth 2005). The operant conditioning chambers that Skinner created became a standard laboratory tool to study how reward and aversion shape behavior of animals, and are still widely used in animal research on addiction, decision making, and learning and memory.

In more recent decades, interest in operant conditioning has sparked due to the rise of artificial intelligence and its subfield of machine learning, which studies the ability of computers to learn on the basis of data without being explicitly programmed. One form of machine learning is called reinforcement learning, which teaches computers how to ideally respond on the basis of feedback, and is essentially a quantitative approach to operant conditioning. As such, the computer uses positive and negative feedback to improve its own performance. Since its development, reinforcement learning has been applied to a wide variety of concepts, including computer-driven stock trading (Jae

Won 2001), teaching a computer how to play video games (Mnih and others 2015), and teaching robots how to move around in an environment (Peters and others 2003).

An important paper that is considered the foundation of reinforcement learning theory is work published by Rescorla and Wagner in 1972 (Rescorla and Wagner 1972), who built upon a theory that stated that "surprise", that is, a difference between expected and actually received reward, is a driving force behind learning. They proposed that the amount of expected reward was based on the pooled evidence that reward will occur from all the stimuli present in the environment. This theory was later extended by Sutton and Barto (1981) to learning from rewards that are temporally separated from its predictive cue or preceding action. The essence of a behavioral approach to reinforcement learning is that an organism makes decisions in order to maximize reward in the long term. For example, if a hungry rat performs a behavioral task in an operant cage, it tries to earn as many rewards (e.g., food pellets) as possible.

In humans, everyday value-based decision making behavior entails a complex process in which the gains and costs associated with different courses of action at any particular moment in time are compared in order to maximize reward. Such a reward can be anything, from the consumption of a delicious snack to maximizing profits during a night in the casino, to going to college in order to achieve long-term wealth and happiness. As in other organisms, reinforcement learning plays a mediating role in these decision-making processes; for each possible action, one makes a cost-benefit analysis on the basis of

previous experiences, and these costs and benefits are adjusted for their probability of occurrence and expected timing of the outcomes. For example, when you want to buy a tasty dessert, you will consider the direct reward associated with the consumption, and penalize this in some way for the direct financial costs of the purchase, as well as the long-term health consequences of the dessert. In this way, for every decision you make, the pros and cons will be weighed into a net expected value that will steer the decision of performing a certain action or not.

Neuronal Value Signals

Given the large number of decisions an organism has to make on a daily basis, it is reasonable to assume that value coding, feedback integration, and value comparisons are mediated through widespread neural circuits. In the past decades, many of such value-related brain signals have been identified using various neuroimaging and neuronal recording techniques. A formal distinction can be made between a *reward signal*, in which neuronal activity changes during reward delivery, and a *reward prediction error signal*, in which neuronal activity changes in response to the "surprise" associated with unexpectedly occurring reward or rewarding stimuli. A *value signal* is a type of reward signal that scales with the subjective experience of the reward. This intensity can reflect both differences in quantity (a bigger reward will yield a higher neuronal response) and quality (a better reward will yield a higher neuronal response). Moreover, these value signals could, in principle, represent a net expectation, that is, the expected value associated with a certain action after subtraction of its costs (e.g., effort and aversive consequences)—an integrated measure of value that has shown to be encoded in some parts of the human and monkey brain (Rangel and Hare 2010).

One can assume that in order to make decisions, there must be some sort of common currency, that is, a single "one size fits all" scale of value, that can be used to compare choice options of different modalities (e.g., choosing between coffee or a banana). Evidence in favor of neuronal value coding in such a common currency comes from a landmark study by Padoa-Schioppa and Assad (2006), who performed single unit recordings in the orbitofrontal cortex of rhesus monkeys. Animals could choose between two types of juices that differed in taste and were offered in different quantities on a visual screen, and the monkeys could make a choice by making eye movements. They found that during the choice process, many neurons in the orbitofrontal cortex encoded some aspect of the choices the monkeys made (Fig. 1). These neurons either encoded (1) the quantity of one of the offered juices, (2) the value (a combination of taste and amount) of the chosen juice, or (3) the taste of the chosen juice (a binary response to one of the two juices during reward delivery). In a follow-up study, these authors demonstrated that responses of a single neuron to an offered

or chosen reward did not depend on which other rewards were offered at the same time (Padoa-Schioppa and Assad 2008), suggesting absolute, rather than relative coding of value. Collectively, these data point toward orbitofrontal cortex neurons encoding aspects of choice in a single, common value measure that can be used to compare qualitatively different options. A recent study showed that during deliberation of a binary choice, orbitofrontal cortex neurons that encode the two different option values alternate in activity, providing a mechanism for these neurons to be directly involved in weighing choice options (Rich and Wallis 2016). Similar forms of economic value coding have later been found in the ventromedial region of the prefrontal cortex of monkeys (Strait and others 2014). However, despite various efforts, no direct evidence has thus far been found that individual brain cells of rodents encode value in a single, common scale.

Whether neuronal value signals are subsequently compared and courses of actions selected by distinct, downstream brain regions remains a matter of debate (Fumagalli 2013; Vlaev and others 2011). In contrast to a modular view on economic choice, in which each brain region controls one part of the chain of a choice process, some researchers have proposed that during decision making, multiple brain regions compute value components of choice independent of each other (Cisek 2012; Hunt and Hayden 2017; Rushworth and others 2012). In this regard, a parallel has been drawn with the distributed decision making of bee swarms: when looking for a potential new hive site, the bees make a choice for a new site in concert, through a distributed consensus, emerging from the information gathered by individual bees (Seeley and others 2006). Likewise, it is thought that different brain areas evaluate, compare, and/or select different choice options, and a choice emerges as a result of the interactions of these regions on a circuit level (Cisek 2012; Hunt and Hayden 2017). One paper has suggested that different brain regions have a role in disentangling the different aspects of choice from sensory cues related to the value of choice options, very similar to how the visual system delineates visual imageries (Yoo and Hayden 2018). As a result, brain regions involved in value-based decision making encode abstract decision making variables that each retain components of the value of the options. This may explain why reward signals have been observed throughout the brain (Schultz 2000), and it suggests that there is no final common pathway for choice selection, but rather that value signals converge at multiple points to eventually compete for execution in the motor system. How these ideas relate to value coding in a "common currency," that is, if and how these abstract reward signals eventually converge into value signals in a single, common scale, remains a question of outstanding interest.

There is substantial evidence that aversive stimuli are also explicitly coded in the brain. For example, lateral
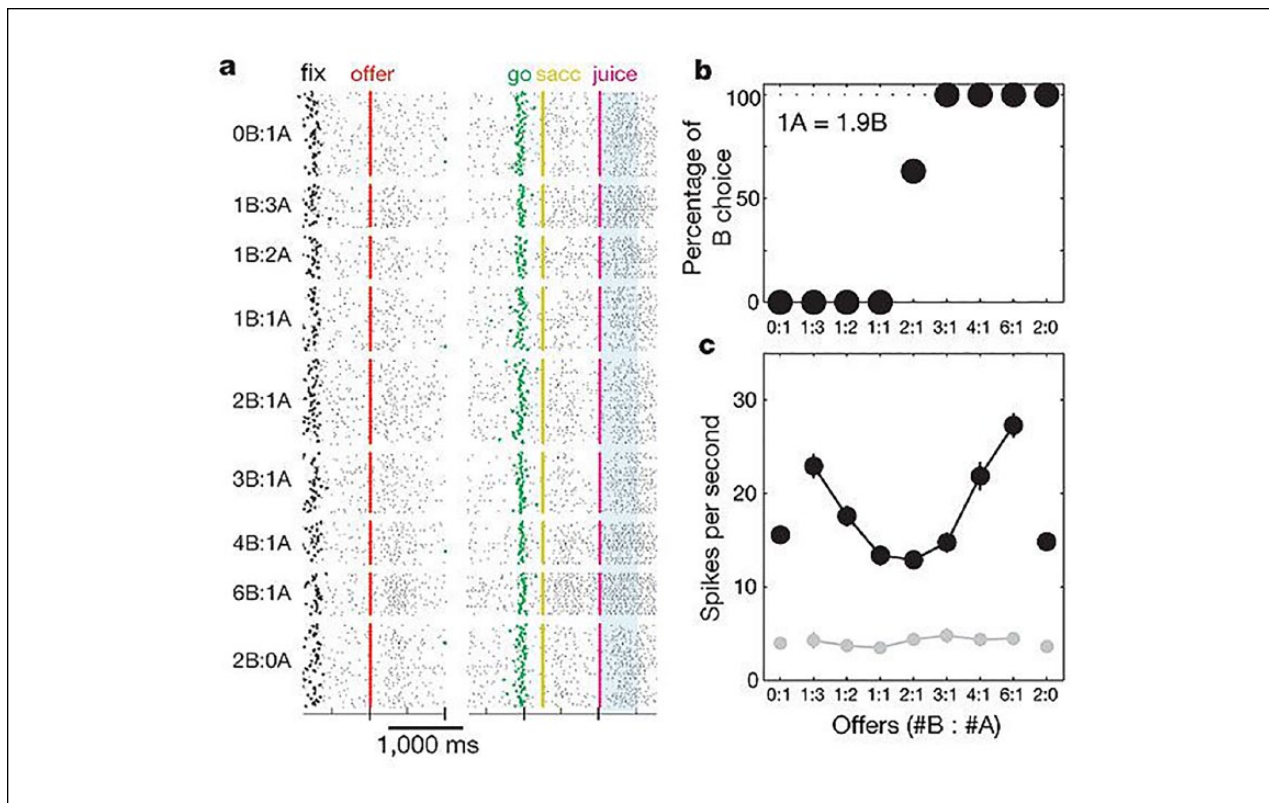
Figure 1. Responses of an example OFC neuron of a monkey, in which the animal had to choose between two different juices offered in varying quantities. The activity of the neuron dependent on the type of offer (#B : #A), but not on the choice of the animal (juice A or B), or the position of the offered juice on the screen (left or right; not shown). (a) Activity during individual trials, (b) choices of the animal, (c) average activity during trials of the same offer type. Numbers on the *x*-axes represent the quantity of an arbitrary amount of juice B and juice A offered, respectively. Error bars, SEM. Image adapted, with permission from Nature Springer, from Padoa-Schioppa and Assad (2006).

habenula neurons have shown to increase activity in response to unexpected punishment and decrease activity in response to unexpected reward (Matsumoto and Hikosaka 2007). Furthermore, a subpopulation of basolateral amygdala neurons projecting to the central amygdala are primarily activated by aversive stimuli, and these have been shown to be essential for fear conditioning (Namburi and others 2015). Importantly, the brain regions involved in punishment (i.e., the negative consequence of an action that suppresses its future expression) have been shown to partially overlap with those involved in reinforcement and reward, including the nucleus accumbens, septum, prefrontal cortex, amygdala, and hippocampus (Jean-Richard-Dit-Bressel and others 2018).
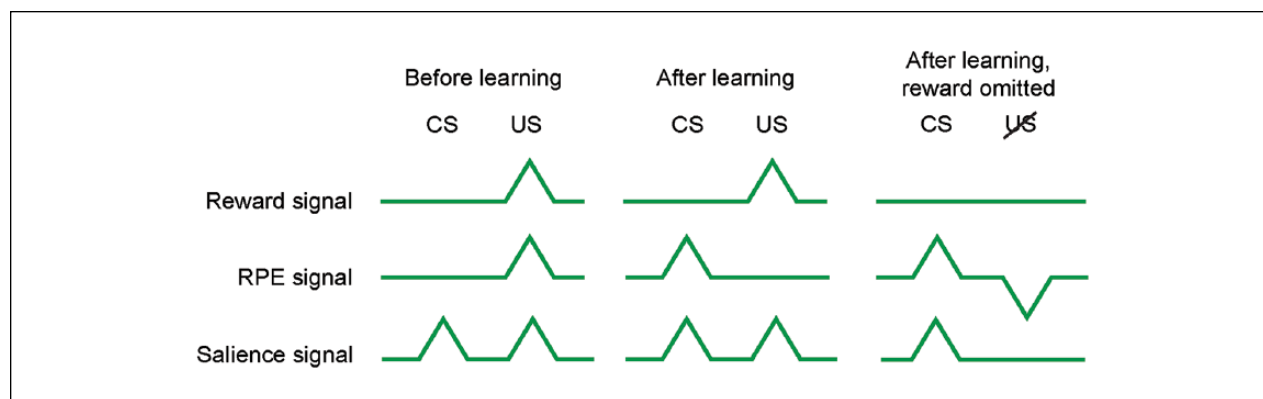
Reward Prediction Error Signals

During value-based learning, expectations of reward (and aversion) are updated on the basis of experiences, creating an up-to-date representation of the value of stimuli in the surrounding world that is necessary for making profitable decisions. As postulated by reinforcement learning theories, this updating process may be guided by prediction errors, or "surprise", computed by subtracting the received reward from the cached reward expectation:

$$
\begin{aligned}
\text{Reward prediction error} = \\
\text{Reward received} \\
- \text{Reward expected}
\end{aligned}
\tag{1}
$$

As such, when a reward is better than expected (i.e., a positive reward prediction error), the value of the action or stimulus that preceded that reward will be increased, and when a reward is worse than expected or when explicit punishment has occurred (i.e., a negative reward prediction error), the value of the preceding action or stimulus will decrease.

Thus, a reward that is fully predicted by a preceding sensory stimulus will not evoke a neuronal response during the reward itself, as the surprise (i.e., reward prediction error) associated with that reward is zero. Neurons

Figure 2. Reward and reward prediction error (RPE) signals in the brain. After extensive training, reward prediction errors signals will only emerge during the conditioned (CS; predictive cue), but not unconditioned (US; reward) stimulus.

that encode reward prediction errors will therefore, after extensive learning, only show changes in activity during the conditioned stimulus that precedes the reward or punishment, but not the unconditioned reward itself (Fig. 2). Conversely, when an expected reward is not delivered, or when explicit punishment is delivered, a negative reward prediction error occurs, resulting in a reduction in firing rate. Such positive and negative reward prediction errors are thought to be important mediators of the approach and avoidance processes that underlie instrumental learning (den Ouden and others 2012; Keiflin and Janak 2015; Schultz and others 1997). In the literature, this prediction error-based type of learning is often referred to as "model-free" reinforcement learning, as it relies on trial-and-error experience, rather than a coherent understanding of the environment (i.e., model-based learning) (Dayan and Niv 2008).

Although theoretically and physiologically distinct, it can be quite challenging to experimentally discern between reward signals, reward prediction error signals, and, for example, general responses to salient stimuli (Fig. 2). To have a full transfer of the neuronal signal from the unconditioned (i.e., reward or punishment) to the conditioned (i.e., cue) stimulus, (1) animals need to have fully learned the association (which may require a long training period), (2) the environment should be perfectly predictable, and (3) the timing of the occurrence of the unconditioned stimulus by the experimental subject should be precise. Many studies report neuronal activation during both the conditioned and unconditioned stimuli (e.g., Beyeler and others 2016; Matias and others 2017; Wang and others 2017), suggesting that these requirements have not fully been met or that mixed neuronal signals have been recorded.

The Role of Dopamine

Although neuronal signals with characteristics of reward prediction error have been found across a wide range of brain areas (den Ouden and others 2012; Watabe-Uchida

and others 2017), the neurocomputationally most pure and perhaps behaviorally most important form of prediction error coding is found in dopamine cells in the midbrain (Schultz and others 1997). A large proportion of these neurons have been shown to increase firing in response to better-than-expected reward, to decrease firing in response to worse-than-expected reward or explicit punishment, and to show no change in firing when reward is fully predictable—an observation that has been reported in a wide range of species including humans (D'Ardenne and others 2008), monkeys (Bayer and Glimcher 2005; Schultz and others 1997), and rodents (Day and others 2007; Tian and others 2016). In the last decades, dopamine neurons have therefore emerged as a prime candidate for mediating reinforcement learning.

A major line of evidence for an involvement of dopamine in reward processing was based on influential work in 1954 from Olds and Milner who showed that animals vigorously lever press in exchange for electrical stimulation of limbic brain structures (Olds and Milner 1954) (Fig. 3), a phenomenon now known as intracranial self-stimulation. This first experiment was not performed directly in the dopamine system, but follow-up studies have shown that intracranial self-stimulation was strongest for midbrain dopamine nuclei and connected regions, and that half of all the brain regions for which animals showed self-stimulation were directly connected to dopamine neurons (Wise 1996). A role for dopamine in mediating reinforcement was further suggested by a series of studies that showed that operant responding for rewards was attenuated after pharmacological blockade of dopamine receptors in the brain (Kelley 2004; Salamone and others 2003; Wise and others 1978).

The interest in dopamine further sparked when Schultz and others (1997) made an exciting discovery in the 1990s: they found neuronal correlates of reward prediction errors in midbrain dopamine neurons of monkeys, as described by Rescorla and Wagner (1972) more than two
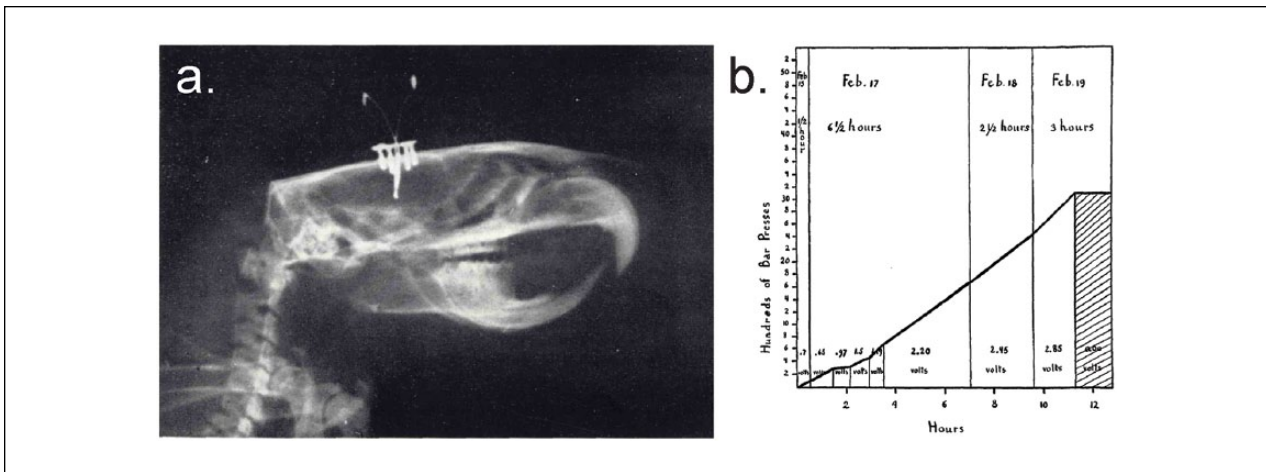
Figure 3. Images from the original Olds and Milner paper (1954; image in public domain) who for the first time demonstrated that animals will lever press for electrical stimulation of limbic brain structures. (a) X-ray image of a rat with an electrode implant. (b) Learning curve of an animal implanted with an electrode in the septal area making lever presses for electrode stimulation.

decades earlier, and in accordance with Sutton and Barto's temporal difference learning model (Sutton and Barto 1981; Sutton and Barto 1998). This discovery was an important step in the understanding of dopamine function and it suggested a direct role for dopamine neurons in reinforcement and punishment learning, thereby mediating important aspects of value-based decision making (Keiflin and Janak 2015; Schultz and Dickinson 2000).

Although the importance of dopaminergic prediction errors to learning was quickly acknowledged, their necessity and sufficiency for learning has been confirmed only recently, employing optogenetic tools in rodents. In one study, Steinberg and others (2013) demonstrated that brief optogenetic activation of VTA dopamine neurons was able to drive learning of the association between a conditioned stimulus and reward (Steinberg and others 2013). They further showed that activation of dopamine neurons during the time of expected reward delivery slowed extinction learning, together suggesting that an artificial positive reward prediction error can drive appetitive learning. Conversely, Chang and others (2016) showed that brief optogenetic inhibition of VTA dopamine neurons in mice was sufficient to mimic negative reward prediction errors and thereby drive avoidance learning (Chang and others 2016). Finally, Saunders and others (2018) demonstrated that optogenetic excitation of VTA dopamine neurons during presentation of a cue was sufficient to attribute incentive motivational value to that cue, even in the absence of explicit reward. They further showed that this was mediated by dopaminergic neurotransmission in the core region of the nucleus accumbens (Saunders and others 2018).

To compute a reward prediction error, a system needs, by definition, information about the reward it expects.

Takahashi and others (2011) studied whether midbrain dopamine neurons receive this information from the orbitofrontal cortex by measuring reward prediction errors in the ventral tegmental area during a reward-learning task in rats with and without a neurotoxic lesion of the lateral orbitofrontal cortex (Takahashi and others 2011). They observed that both positive and negative reward prediction error coding in the ventral tegmental area was attenuated by the lesion. However, the pattern of observed effects did not match the hypothesis that the lesioned part of orbitofrontal cortex conveyed a pure value signal to the dopamine neurons, as the authors demonstrated by simulating electrophysiological data with reinforcement learning models. Indeed, it has later been suggested that the orbitofrontal cortex has a role in model-based, rather than model-free reinforcement learning (Jones and others 2012; Wilson and others 2014), which may explain the lack of evidence for the OFC encoding value in a way that supports prediction-error based learning. More recent work on the computations underlying dopaminergic reward prediction error suggests that dopamine neurons use value information from a wide range of areas to compute prediction errors (Tian and others 2016). Wherever these value signals arise from, electrophysiological evidence suggests that dopamine neurons use subtractions to compute the prediction error from the expected and received reward, and that inhibition through GABAergic neurotransmission in the ventral tegmental area facilitates this computation (Eshel and others 2015).

Despite the apparent homogeneity of prediction error responses in midbrain dopamine neurons in some studies (Ungless and others 2004), it must be noted that since the development of genetic tools for neural circuit dissection,

an increasing number of studies points toward heterogeneity in dopamine cells with regard to connectivity, morphology, gene expression, and function (Lammel and others 2014; Morales and Margolis 2017; Saunders and others 2018). For example, recent studies have shown that dopamine released in the prefrontal cortex biases responses of rodents to ambiguous stimuli toward avoidance (rather than approach) behavior (Vander Weele and others 2018), and a subset of mesolimbic dopamine neurons releases dopamine in response to aversive, rather than rewarding stimuli (de Jong and others 2018). Furthermore, reward-related responses of individual dopamine neurons have been shown to encode aspects of motor behavior (Howe and Dombeck 2016; Jin and Costa 2010), together suggesting that prediction errors are not encoded as mathematically pure and homogeneous as was thought before. That said, the importance of dopamine and dopaminergic reward prediction errors to value-based learning and decision making has been one of the most well-established principles in recent neuroscientific history (Hu 2016; Keiflin and Janak 2015; Schultz 2016; Watabe-Uchida and others 2017).

A Neuroeconomic Approach to Motivation

One aspect of reward-related behavior for which dopamine is critical is motivation (Cools 2008; Salamone and Correa 2012). Although different authors use slightly different definitions of this term (Salamone and Correa 2012), motivation typically refers to the willingness to invest resources (such as time or effort) in order to receive a reward or to avoid a punisher. In support of a role for dopamine in motivation, it has been found that after forebrain dopamine depletion, animals will cease to actively search for food (and eventually starve to death), but they will still consume food when it is placed in their mouth (Salamone and Correa 2012). In less extreme experiments, it was found that treatment with dopamine receptor antagonists reduced responding for food under behaviorally demanding schedules (i.e., when animals have to make a relatively large numbers of responses for food), but not when little or no effort was required to obtain it (Kelley 2004; Salamone and Correa 2012). In this context, Berridge and Robinson have proposed a useful distinction between the "liking" (i.e., the experience of pleasure) and "wanting" (i.e., the motivation to obtain it) of a reward (Berridge and Robinson 1998; Berridge and others 2009), and it is generally assumed that dopamine is mainly involved in the latter (Salamone and Correa 2012).

In neuroeconomic terms, motivation is thought of as the subjective experience that a certain action is worth pursuing. The value of such an action can be described by an economic utility function (Houthakker 1950), so that every time an organism considers a certain action, a computation is performed where the subjective experience of the costs (labor and negative consequences, corrected for the probability of occurrence) is subtracted from the expected reward that follows that action (receiving food, sex, drugs, or shelter, or avoiding punishment, corrected for probability) (Rangel and others 2008), yielding the net expected reward value associated with that action (sometimes referred to as "action value"; Rangel and Hare 2010):

$$\text{Net expected reward} = \sum \text{reward}_{\text{subjective}} - \sum \text{costs}_{\text{subjective}} \quad (2)$$

Only when this calculation has a positive outcome, an action will be pursued, as the expectation of reward is higher than its expected cost. Conversely, when the outcome of this calculation approaches 0 or becomes negative (i.e., when costs > reward), no action is taken. The subjective reward term in this equation ($\sum \text{reward}_{\text{subjective}}$) can be seen as the expectation of pleasure associated with reward ("liking"), and the outcome of the equation is proportional to motivation ("wanting"), so that

$$\text{Motivation} \propto \sum \text{reward}_{\text{subjective}} - \sum \text{costs}_{\text{subjective}} \quad (3)$$

For example, whether an animal will start foraging for food depends on several factors. First, it depends on the amount of food it expects to receive in that environment ($\sum \text{reward}$). Second, it depends on to what extent the food is appreciated; a satiated animal will appreciate food less than a hungry animal, and palatable food is appreciated more than plain food. Hence, the objective reward expectation $\sum \text{reward}$ should be multiplied with a subjectivity factor that reflects the metabolic and hedonic state of the animal, leading to a subjective reward value $\sum \text{reward}_{\text{subjective}}$. Conversely, the costs of foraging depend on the effort the animal has to exert to seek for food and the dangers associated with food seeking (i.e., the probability of explicit negative consequences, like a predator attack). Again, this factor should be corrected for subjectivity, leading to a subjective cost factor $\sum \text{costs}_{\text{subjective}}$. When the expected reward outweighs the costs, the animal will start foraging. Logically, subjective reward value increases with hunger (a meal tastes much better when you are hungry), so that even in a dangerous environment, reward will at some point outweigh costs, and the motivation to start to seek for food will increase. Furthermore, influential economic and psychological theories state that rewards and costs that are further away in the future or that are less
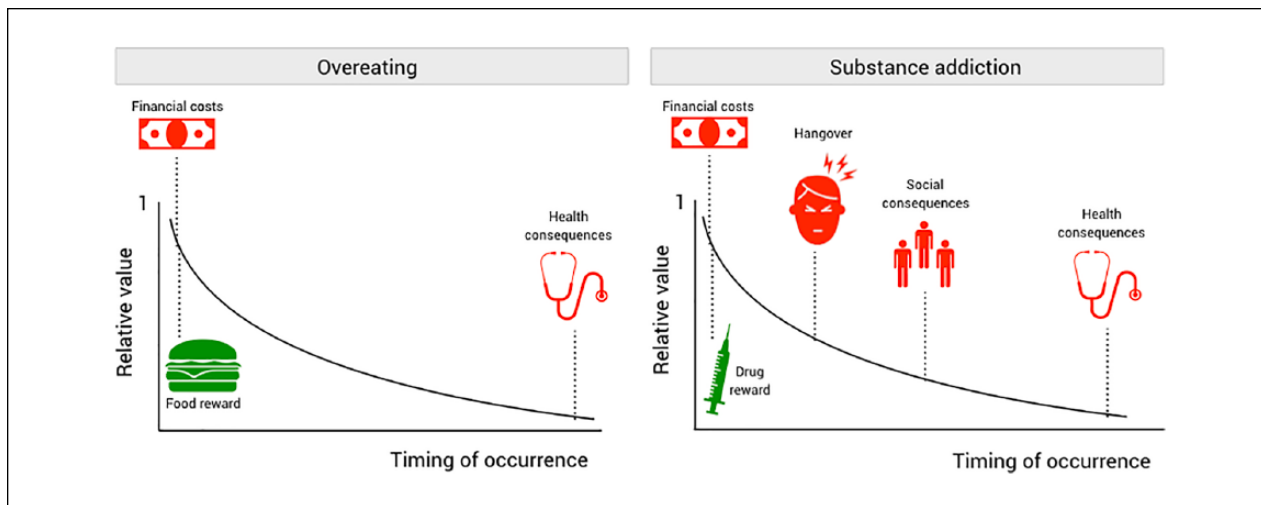
Figure 4. For every decision a person makes, the pros and cons will be weighed into a net expected value that will steer the decision of performing a certain action or not. In this regard, overeating and substance addiction can be viewed as motivational states in which the reward associated with the action (green; eating unhealthy foods or taking substances, respectively) continuously outweigh the costs (red). Costs are usually further away in the future (and sometimes probabilistic), which discounts them to sometimes negligible levels, thereby making it ostensibly profitable to pursue reward. Note that for visualization purposes, both rewards and costs are depicted in the standard exponential discounting curve, but the shape of these curves may differ between factors; see MacKeigan and others (1993), Mischel and others (1969), and Petry (2003).

likely to be received are discounted, that is, its subjective value is reduced with time and probability—a process known as temporal or probability discounting, respectively (Critchfield and Kollins 2001; Green and Myerson 2004). Hence, Equation 3 can be rewritten as

$$\text{Motivation} \propto$$
$$\sum s_{\text{reward}} * \gamma_{\text{reward}} * \text{reward}$$
$$- \sum s_{\text{costs}} * \gamma_{\text{costs}} * \text{costs} \tag{4}$$

in which s represents a subjectivity factor that scales the reward/cost on the basis of the animal's intrinsic state and desires, and γ a discounting factor that is low when the rewards or costs are further away in the future or are less likely to occur.

This simple framework of motivation may help structuring our understanding of phenomena that are associated with reward seeking and motivation (Fig. 4). For example, the vast increase in the prevalence of obesity in the Western world (World Health Organization 2000) is thought to arise from the abundance of cheap and palatable calorie-rich food, the difficulty to make healthy food choices, and the fact that it is hard to lose weight (Rangel and Hare 2010). In our society, the costs associated with food intake are radically different than they have been for the past millennia and different than for animals in the wild. For animals and premodern man, the costs mainly comprised the physical effort and the dangers that were associated with hunting and other forms of foraging. For

modern man, given the abundance of food, the costs comprise the financial costs of the food and the negative health consequences that are associated with food intake. Given that food is usually directly available, Equation (4) can be given by

$$\text{Motivation} \propto s * \big[\text{food reward}\big]$$
$$- s * \gamma * \big[\text{health consequences}\big]$$
$$- s * \big[\text{financial costs}\big] \tag{5}$$

Despite the potential severity of the health consequences of palatable foods, they often develop over a longer period of time and they are thus not immediately noticed. This may discount the subjective experience of the negative health consequences to a negligible level, except perhaps when someone has low temporal discounting characteristics. Indeed, trait impulse control, of which temporal discounting capacity is an important component, is predictive for the maintenance of overweight in children (Nederkoorn and others 2006) and adults (Nederkoorn and others 2010). An additional point is that unhealthy foods, high in carbohydrates and fat, are often cheaper than healthy foods, adding an extra costs factor to the equation, thereby decreasing the motivation to make healthy food choices—a factor that may especially play a role in people with a low income (Steptoe and others 1995). Thus, the direct reward of palatable food and the absence of any direct costs associated with its intake makes it ostensibly unprofitable to

make healthy food choices. Limiting palatable food intake is especially hard during dieting, as this in fact increases sensitivity to food reward (Laeng and others 1993; van der Plasse and others 2015), making the left side of this equation more dominant.

A second useful application of this framework is to understand substance addiction and the fact that some people are more prone to develop this mental disorder than others. Every time a user gets reminded of the substance (by, e.g., cravings, cues, or social pressure), this person will make a decision to use them or not. Considering the expectation of reward from the "high" of the substance and the negative consequences of its use (financial costs, hangovers, long-term health consequences, and consequences for social obligations), Equation (4) can be written as

$$
\begin{aligned}
\text{Motivation} \propto s \;&* \; \big[\text{high}\big] \\
&- s * \gamma * \big[\text{financial costs}\big] \\
&- s * \gamma * \big[\text{hangover}\big] \qquad (6) \\
&- s * \gamma * \big[\text{social consequences}\big] \\
&- s * \gamma \; * \; \big[\text{health consequences}\big]
\end{aligned}
$$

In recreational substance users, the expected reward of substance intake only occasionally outweighs its cost, while in addicted individuals, the left side of this equation is chronically dominant. Given this list of costs associated with prolonged substance use, it is not surprising that only a minority of recreational drug users eventually develops addiction (Warner and others 1995). Based on this equation, however, several risk factors can be identified for the development of addiction. First, increased expectation of substance-induced euphoria (note that this is different from the actually experienced pleasure) would strengthen the left side of Equation (6). Second, low baseline levels of the costs factors—that is, a poor social life, no job or study, and bad health—make the costs of substance use relatively low. Third, a low value of temporal discounting factor γ (i.e., discounting of subjective value over time is stronger) also reduces the weight of the costs of substance abuse. Indeed, several studies have demonstrated that increased expectation of drug effects (Volkow and others 2010), a low socioeconomic status (Jordan and Andersen 2017; Nesse and Berridge 1997), and high temporal discounting levels (Fineberg and others 2014) increase the risk for the development of addiction. In this context, it is important to realize that the value of the discounting factor γ is likely different for the various cost factors. For example, it has been shown that money is discounted at a faster rate than freedom, which is discounted at a faster rate than health, whereby discounting rates were higher in substance addicts compared to healthy controls (Petry 2003). The negative consequences

of repeated substance use also decrease the baseline levels of health and social life, essentially decreasing the cost factors in this equation, thus making future use more likely. Furthermore, both animal and human studies have shown that repeated exposure to substances of abuse increase temporal discounting (Fineberg and others 2014). Importantly, in contrast to the intake of (palatable) nutrients, substances of abuse directly bind to proteins (such as receptors and transporters) and result in plastic changes of neural circuits that mediate value-based decision making (Lüscher and Ungless 2006), thus hitting the hardware of neural computation in the brain.

Implications for Psychiatry

Within this proposed framework, overeating or substance addiction can be viewed as a state of altered value-based decision making. Importantly, however, for an individual within these states, the decision to take unhealthy foods or substances can be perfectly rational. The negative consequences of overeating and substance use are diffuse and delayed in time and thus shift the weight in Equations (5) and (6) dramatically to the left. This may explain why these mental conditions are among the most difficult to treat, as the failure rates of dieting and relapse rates of substance abuse are notoriously high (Brandon and others 2007; Kärkkäinen and others 2018).

Abnormalities in the brain circuits involved in value processing, motivation, and decision making have been implicated in overeating, substance addiction, as well as a wide variety of other neuropsychiatric behaviors. For example, dysfunctions in the dopamine system have been associated with obesity (Wang and others 2001), addiction (Volkow and Morales 2015), depression (Russo and Nestler 2013), bipolar disorder (Cousins and others 2009), attention-deficit hyperactivity disorder (Volkow and others 2009), and schizophrenia (Weinstein and others 2017)—not the least because most of the effective pharmacotherapies for some of these diseases target the dopamine system. Moreover, dysfunction of the prefrontal cortex, another important region for value-based decision making, has been implicated in a partially overlapping set of disorders, including addiction (Volkow and Morales 2015), impulse control disorders (Bechara and Van Der Linden 2005), depression (Han and Nestler 2017), and schizophrenia (Barch and others 2001). Besides dysfunctions in these brain circuits, altered value-based decision making has been observed in all of these patient groups (Fineberg and others 2014; Garon and others 2006; Grant and others 2000; Murphy and others 2001; Noel and others 2013; Shurman and others 2005), an indication that changes in value processes might be involved in the etiology of neuropsychiatric behaviors. Whether this is indeed the case, and whether changes in value-based decision
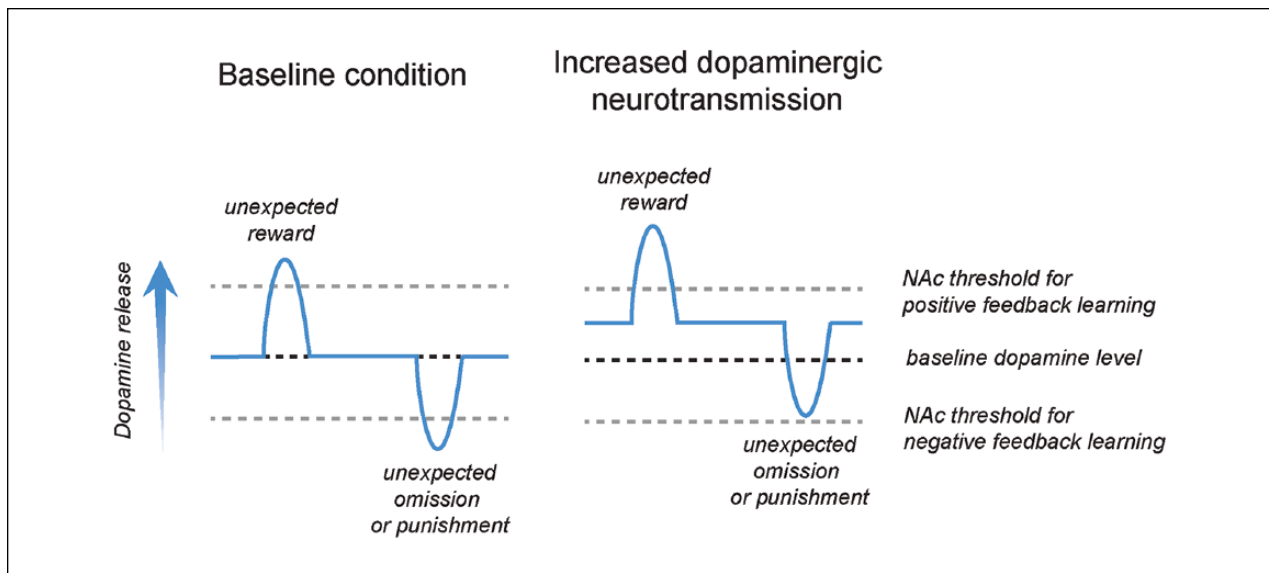
Figure 5. Increased dopamine release in the nucleus accumbens (NAc) specifically interferes with negative reward prediction error learning, while leaving positive reward prediction error learning intact. This mechanism may explain the overoptimistic decision making behavior that is observed in states of increased dopaminergic neurotransmission, such as during substance use, mania, and dopamine replacement therapy in Parkinson's disease. Image adapted from Verharen and others (2018); published under a Creative Commons License.

making directly mediate disease progression, remains a challenging question, although some important theories have been postulated in recent years.

For example, it has been suggested that depression at least partially arises from unrealistically low reward expectations, mainly due to pessimistically set priors (i.e., assumptions) in model-based (but not model-free) reasoning (Huys and others 2015). Furthermore, neurocomputational models predicted that the reckless and overoptimistic decision-making behavior after levodopa treatment in Parkinson's disease patients is induced by impaired prediction error learning due to overstimulation of striatal dopamine receptors (Frank and others 2004). This hypothesis has been supported by several clinical studies (Cools 2006) and by a recent rodent study (Verharen and others 2018), and this may also be of importance for the understanding of mania, as this mental state is also associated with elevated dopamine levels (Cousins and others 2009) (Fig. 5). A third example is anxiety disorders, which have been suggested to result from increased threat avoidance due to an overestimation of the probability and magnitude of aversive outcomes (Bishop and Gagne 2018). This mechanism may arise from alterations in brain areas involved in learning and value-based decision making, like the amygdala and anterior cingulate cortex (Bishop and Gagne 2018).

The recent emergence of several new methods for computational analyses, large-scale neuronal recordings and neuronal manipulations with unprecedented precision now allow for a detailed investigation of the neural circuits involved in reward and aversion processing that contribute to value-based decision making and motivation. These developments hold great promise to increase our understanding of these processes, and may ultimately contribute to the development of improved treatment strategies for a wide array of mental disorders.

Declaration of Conflicting Interests

Funding

ORCID iD

Jeroen P. H. Verharen https://orcid.org/0000-0001-7582 -802X

References

Barch DM, Carter CS, Braver TS, Sabb FW, MacDonald A, Noll DC, and others. 2001. Selective deficits in prefrontal

cortex function in medication-naive patients with schizophrenia. Arch Gen Psychiatry 58(3):280–8.

Bayer HM, Glimcher PW. 2005. Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron 47(1):129–41.

Bechara A, Van Der Linden M. 2005. Decision-making and impulse control after frontal lobe injuries. Curr Opin Neurol 18(6):734–9.

Berridge KC, Robinson TE. 1998. What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? Brain Res Brain Res Rev 28(3):309–69.

Berridge KC, Robinson TE, Aldridge JW. 2009. Dissecting components of reward: "liking", "wanting", and learning. Curr Opin Pharmacol 9(1):65–73.

Beyeler A, Namburi P, Glober GF, Simonnet C, Calhoon GG, Conyers GF, and others. 2016. Divergent routing of positive and negative information from the amygdala during memory retrieval. Neuron 90(2):348–61.

Bishop SJ, Gagne C. 2018. Anxiety, depression, and decision making: a computational perspective. Annu Rev Neurosci 41:371–88.

Brandon TH, Vidrine JI, Litvin EB. 2007. Relapse and relapse prevention. Annu Rev Clin Psychol 3:257–84.

Chang CY, Esber GR, Marrero-Garcia Y, Yau HJ, Bonci A, Schoenbaum G. 2016. Brief optogenetic inhibition of dopamine neurons mimics endogenous negative reward prediction errors. Nat Neurosci 19(1):111–6.

Cisek P. 2012. Making decisions through a distributed consensus. Curr Opin Neurobiol 22(6):927–36.

Cools R. 2006. Dopaminergic modulation of cognitive function-implications for L-DOPA treatment in Parkinson's disease. Neurosci Biobehav Rev 30(1):1–23.

Cools R. 2008. Role of dopamine in the motivational and cognitive control of behavior. Neuroscientist 14(4):381–95.

Cousins DA, Butts K, Young AH. 2009. The role of dopamine in bipolar disorder. Bipolar Disord 11(8):787–806.

Critchfield TS, Kollins SH. 2001. Temporal discounting: basic research and the analysis of socially important behavior. J Appl Behav Anal 34(1):101–22.

D'Ardenne K, McClure SM, Nystrom LE, Cohen JD. 2008. BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. Science 319:1264–8.

Day JJ, Roitman MF, Wightman RM, Carelli RM. 2007. Associative learning mediates dynamic shifts in dopamine signaling in the nucleus accumbens. Nat Neurosci 10(8):1020–8.

Dayan P, Niv Y. 2008. Reinforcement learning: the good, the bad and the ugly. Curr Opin Neurobiol 18(2):185–96.

de Jong JW, Afjei SA, Pollak Dorocic I, Peck JR, Liu C, Kim CK, and others. 2018. A neural circuit mechanism for encoding aversive stimuli in the mesolimbic dopamine system. Neuron 101(1):133–51.e7.

den Ouden HE, Kok P, de Lange FP. 2012. How prediction errors shape perception, attention, and motivation. Front Psychol 3:548.

Eshel N, Bukwich M, Rao V, Hemmelder V, Tian J, Uchida N. 2015. Arithmetic and local circuitry underlying dopamine prediction errors. Nature 525:243–6.

Fineberg NA, Chamberlain SR, Goudriaan AE, Stein DJ, Vanderschuren LJ, Gillan CM, and others. 2014. New developments in human neurocognition: clinical, genetic, and brain imaging correlates of impulsivity and compulsivity. CNS Spectr 19(1):69–89.

Frank MJ, Seeberger LC, O'Reilly RC. 2004. By carrot or by stick: cognitive reinforcement learning in parkinsonism. Science 306(5703):1940–3.

Fumagalli R. 2013. The futile search for true utility. Econ Philos 29(3):325–47.

Garon N, Moore C, Waschbusch DA. 2006. Decision making in children with ADHD only, ADHD-anxious/depressed, and control children using a child version of the Iowa Gambling Task. J Atten Disord 9(4):607–19.

Grant S, Contoreggi C, London ED. 2000. Drug abusers show impaired performance in a laboratory test of decision making. Neuropsychologia 38:1180–7.

Green L, Myerson J. 2004. A discounting framework for choice with delayed and probabilistic rewards. Psychol Bull 130(5):769–92.

Han MH, Nestler EJ. 2017. Neural substrates of depression and resilience. Neurotherapeutics 14(3):677–86.

Holth P. 2005. Two definitions of punishment. Behav Anal Today 6(1):43.

Houthakker HS. 1950. Revealed preference and the utility function. Economica 17(66):159–74.

Howe MW, Dombeck DA. 2016. Rapid signalling in distinct dopaminergic axons during locomotion and reward. Nature 535(7613):505–10.

Hu H. 2016. Reward and aversion. Annu Rev Neurosci 39:297–324.

Hunt LT, Hayden BY. 2017. A distributed, hierarchical and recurrent framework for reward-based choice. Nat Rev Neurosci 18:172.

Huys QJ, Daw ND, Dayan P. 2015. Depression: a decision-theoretic analysis. Annu Rev Neurosci 38:1–23.

Jae Won L. 2001. Stock price prediction using reinforcement learning. Paper presented at: ISIE 2001; Pusan, Korea.

Jean-Richard-Dit-Bressel P, Killcross S, McNally GP. 2018. Behavioral and neurobiological mechanisms of punishment: implications for psychiatric disorders. Neuropsychopharmacology 43:1639–50.

Jin X, Costa RM. 2010. Start/stop signals emerge in nigrostriatal circuits during sequence learning. Nature 466(7305):457–62.

Jones JL, Esber GR, McDannald MA, Gruber AJ, Hernandez A, Mirenzi A, and others. 2012. Orbitofrontal cortex supports behavior and learning using inferred but not cached values. Science 338(6109):953–6.

Jordan CJ, Andersen SL. 2017. Sensitive periods of substance abuse: early risk for the transition to dependence. Dev Cogn Neurosci 25:29–44.

Kärkkäinen U, Mustelin L, Raevuori A, Kaprio J, Keski-Rahkonen A. 2018. Successful weight maintainers among young adults—a ten-year prospective population study. Eat Behav 29:91–8.

Keiflin R, Janak PH. 2015. Dopamine prediction errors in reward learning and addiction: from theory to neural circuitry. Neuron 88(2):247–63.

Kelley AE. 2004. Ventral striatal control of appetitive motivation: role in ingestive behavior and reward-related learning. Neurosci Biobehav Rev 27(8):765–76.

Laeng B, Berridge KC, Butter CM. 1993. Pleasantness of a sweet taste during hunger and satiety: effects of gender and "sweet tooth." Appetite 21(3):247–54.

Lammel S, Lim BK, Malenka RC. 2014. Reward and aversion in a heterogeneous midbrain dopamine system. Neuropharmacology 76:351–9.

Lüscher C, Ungless MA. 2006. The mechanistic classification of addictive drugs. PLoS Med 3(11):e437.

MacKeigan LD, Larson LN, Draugalis JR, Bootman JL, Burns LR. 1993. Time preference for health gains versus health losses. Pharmacoeconomics 3(5):374–86.

Matias S, Lottem E, Dugue GP, Mainen ZF. 2017. Activity patterns of serotonin neurons underlying cognitive flexibility. Elife 6.

Matsumoto M, Hikosaka O. 2007. Lateral habenula as a source of negative reward signals in dopamine neurons. Nature 447(7148):1111–5.

Mischel W, Grusec J, Masters JC. 1969. Effects of expected delay time on the subjective value of rewards and punishments. J Pers Soc Psychol 11(4):363–73.

Mnih V, Kavukcuoglu K, Silver D, Rusu A, Veness J, Bellemare M, and others. 2015. Human-level control through deep reinforcement learning. Nature 518:529–33.

Morales M, Margolis EB. 2017. Ventral tegmental area: cellular heterogeneity, connectivity and behaviour. Nat Rev Neurosci 18(2):73–85.

Murphy FC, Rubinsztein JS, Michael A, Rogers RD, Robbins TW, Paykel ES, and others. 2001. Decision-making cognition in mania and depression. Psychol Med 31:679–93.

Namburi P, Beyeler A, Yorozu S, Calhoon GG, Halbert SA, Wichmann R, and others. 2015. A circuit mechanism for differentiating positive and negative associations. Nature 520:675–8.

Nederkoorn C, Braet C, Van Eijs Y, Tanghe A, Jansen A. 2006. Why obese children cannot resist food: the role of impulsivity. Eat Behav 7(4):315–22.

Nederkoorn C, Houben K, Hofmann W, Roefs A, Jansen A. 2010. Control yourself or just eat what you like? Weight gain over a year is predicted by an interactive effect of response inhibition and implicit preference for snack foods. Health Psychol 29(4):389–93.

Nesse RM, Berridge KC. 1997. Psychoactive drug use in evolutionary perspective. Science 278(5335):63–6.

Noel X, Brevers D, Bechara A. 2013. A neurocognitive approach to understanding the neurobiology of addiction. Curr Opin Neurobiol 23(4):632–8.

Olds J, Milner P. 1954. Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. J Comp Physiol Psychiatry 47(6):419–27.

Padoa-Schioppa C, Assad JA. 2006. Neurons in the orbitofrontal cortex encode economic value. Nature 441(7090):223–6.

Padoa-Schioppa C, Assad JA. 2008. The representation of economic value in the orbitofrontal cortex is invariant for changes of menu. Nat Neurosci 11(1):95–102.

Peters J, Vijayakumar S, Schaal S. 2003. Reinforcement learning for humanoid robotics. Paper presented at: Humanoids2003, Third IEEE-RAS International Conference on Humanoid Robots; Karlsruhe, Germany.

Petry NM. 2003. Discounting of money, health, and freedom in substance abusers and controls. Drug Alcohol Depend 71(2):133–41.

Rangel A, Camerer C, Montague PR. 2008. A framework for studying the neurobiology of value-based decision making. Nat Rev Neurosci 9(7):545.

Rangel A, Hare T. 2010. Neural computations associated with goal-directed choice. Curr Opin Neurobiol 20(2):262–70.

Rescorla RA, Wagner AR. 1972. A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. Classical Conditioning II. Curr Res Theory 2:64–99.

Rich EL, Wallis JD. 2016. Decoding subjective decisions from orbitofrontal cortex. Nat Neurosci 19(7):973–80.

Rushworth MF, Kolling N, Sallet J, Mars RB. 2012. Valuation and decision-making in frontal cortex: one or many serial or parallel systems? Curr Opin Neurobiol 22(6):946–55.

Russo SJ, Nestler EJ. 2013. The brain reward circuitry in mood disorders. Nat Rev Neurosci 14(9):609–25.

Salamone JD, Correa M. 2012. The mysterious motivational functions of mesolimbic dopamine. Neuron 76(3):470–85.

Salamone JD, Correa M, Mingote SM, Weber SM. 2003. Nucleus accumbens dopamine and the regulation of effort in food-seeking behavior: implications for studies of natural motivation, psychiatry, and drug abuse. J Pharmacol Exp Ther 305(1):1–8.

Saunders BT, Richard JM, Margolis EB, Janak PH. 2018. Dopamine neurons create Pavlovian conditioned stimuli with circuit-defined motivational properties. Nat Neurosci 21:1072–83.

Schultz W. 2000. Multiple reward signals in the brain. Nat Rev Neurosci 1(3):199–207.

Schultz W. 2016. Dopamine reward prediction-error signalling: a two-component response. Nat Rev Neurosci 17(3):183–95.

Schultz W, Dayan P, Montague PR. 1997. A neural substrate of prediction and reward. Science 275:1593–601.

Schultz W, Dickinson A. 2000. Neuronal coding of prediction errors. Annu Rev Neurosci 23:473–500.

Seeley TD, Visscher PK, Passino KM. 2006. Group decision making in honey bee swarms: when 10,000 bees go house hunting, how do they cooperatively choose their new nesting site? American Scientist 94(3):220–9.

Shurman B, Horan WP, Nuechterlein KH. 2005. Schizophrenia patients demonstrate a distinctive pattern of decision-making impairment on the Iowa Gambling Task. Schizophr Res 72(2–3):215–24.

Skinner BF. 1938. The Behavior of Organisms: An Experimental Analysis. New York: Appleton Century Crofts.

Steinberg EE, Keiflin R, Boivin JR, Witten IB, Deisseroth K, Janak PH. 2013. A causal link between prediction errors, dopamine neurons and learning. Nat Neurosci 16(7):966–73.

Steptoe A, Pollard TM, Wardle J. 1995. Development of a measure of the motives underlying the selection of food: the food choice questionnaire. Appetite 25(3):267–284.

Strait CE, Blanchard TC, Hayden BY. 2014. Reward value comparison via mutual inhibition in ventromedial prefrontal cortex. Neuron 82(6):1357–66.

Sutton RS, Barto AG. 1981. Towards a modern theory of adaptive networks: expectation and prediction. Psychol Rev 88(2):135–70.

Sutton RS, Barto AG. 1998. Reinforcement Learning: An Introduction. Cambridge: MIT Press.

Takahashi YK, Roesch MR, Wilson RC, Toreson K, O'Donnell P, Niv Y, and others. 2011. Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. Nat Neurosci 14(12):1590–7.

Thorndike EL. 1898. Animal intelligence: an experimental study of the associative processes in animals. Psychol Rev 5(5):551–3.

Tian J, Huang R, Cohen JY, Osakada F, Kobak D, Machens CK, and others. 2016. Distributed and mixed information in monosynaptic inputs to dopamine neurons. Neuron 91(6):1374–89.

Ungless MA, Magill PJ, Bolam JP. 2004. Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. Science 303(5666):2040–2.

van der Plasse G, van Zessen R, Luijendijk MC, Erkan H, Stuber GD, Ramakers GM, and others. 2015. Modulation of cue-induced firing of ventral tegmental area dopamine neurons by leptin and ghrelin. Int J Obes (Lond) 39(12):1742–9.

Vander Weele CM, Siciliano CA, Matthews GA, Namburi P, Izadmehr EM, Espinel IC, and others. 2018. Dopamine enhances signal-to-noise ratio in cortical-brainstem encoding of aversive stimuli. Nature 563:397–401.

Verharen JPH, de Jong JW, Roelofs TJ, Huffels CFM, van Zessen R, Luijendijk MC, and others. 2018. A neuronal mechanism underlying decision-making deficits during hyperdopaminergic states. Nat Commun 9(1):731.

Vlaev I, Chater N, Stewart N, Brown GDA. 2011. Does the brain calculate value? Trends Cogn Sci 15(11):546–54.

Volkow ND, Morales M. 2015. The brain on drugs: from reward to addiction. Cell 162(4):712–25.

Volkow ND, Wang GJ, Fowler JS, Tomasi D, Telang F, Baler RD. 2010. Addiction: decreased reward sensitivity and increased expectation sensitivity conspire to overwhelm the brain's control circuit. Bioessays 32(9):748–55.

Volkow ND, Wang GJ, Kollins SH, Wigal TL, Newcorn JH, Telang F, and others. 2009. Evaluating dopamine reward pathway in ADHD: clinical implications. JAMA 302(10):1084–91.

Wang D, Li Y, Feng Q, Guo Q, Zhou J, Luo M. 2017. Learning shapes the aversion and reward responses of lateral habenula neurons. Elife 6.

Wang GJ, Volkow ND, Logan J, Pappas NR, Wong CT, Zhu W, and others. 2001. Brain dopamine and obesity. Lancet 357(9253):354–7.

Warner LA, Kessler RC, Hughes M, Anthony JC, Nelson CB. 1995. Prevalence and correlates of drug use and dependence in the United States. Results from the National Comorbidity Survey. Arch Gen Psychiatry 52(3):219–29.

Watabe-Uchida M, Eshel N, Uchida N. 2017. Neural circuitry of reward prediction error. Annu Rev Neurosci 40:373–94.

Weinstein JJ, Chohan MO, Slifstein M, Kegeles LS, Moore H, Abi-Dargham A. 2017. Pathway-specific dopamine abnormalities in schizophrenia. Biol Psychiatry 81(1):31–42.

Wilson RC, Takahashi YK, Schoenbaum G, Niv Y. 2014. Orbitofrontal cortex as a cognitive map of task space. Neuron 81(2):267–79.

Wise RA. 1996. Addictive drugs and brain stimulation reward. Annu Rev Neurosci 19:319–40.

Wise RA, Spindler J, deWit H, Gerberg G. 1978. Neuroleptic-induced "anhedonia" in rats: pimozide blocks reward quality of food. Science 201(4352):262–4.

World Health Organization. 2000. Obesity: preventing and managing the global epidemic. Report of WHO consultation. World Health Organ Tech Rep Ser 894:i–xii, 1–253.

Yoo SBM, Hayden BY. 2018. Economic choice as an untangling of options into actions. Neuron 99(3):434–47.