

Published in final edited form as:

Neurobiol Learn Mem. 2014 October ; 0: 90–100. doi:10.1016/j.nlm.2014.05.002.

A universal role of the ventral striatum in reward-based learning: Evidence from human studies

Reka Daniel^{1,2} and Stefan Pollmann^{1,3}

¹Department of Experimental Psychology, Otto-von-Guericke-Universität Magdeburg, D-39016 Magdeburg, Germany

²Princeton Neuroscience Institute, Green Hall, Princeton University, Princeton, NJ 08540, USA

³Center for Behavioral Brain Sciences, D-39016 Magdeburg, Germany

Abstract

Reinforcement learning enables organisms to adjust their behavior in order to maximize rewards. Electrophysiological recordings of dopaminergic midbrain neurons have shown that they code the difference between actual and predicted rewards, i.e., the reward prediction error, in many species. This error signal is conveyed to both the striatum and cortical areas and is thought to play a central role in learning to optimize behavior. However, in human daily life rewards are diverse and often only indirect feedback is available. Here we explore the range of rewards that are processed by the dopaminergic system in human participants, and examine whether it is also involved in learning in the absence of explicit rewards. While results from electrophysiological recordings in humans are sparse, evidence linking dopaminergic activity to the metabolic signal recorded from the midbrain and striatum with functional magnetic resonance imaging (fMRI) is available. Results from fMRI studies suggest that the human ventral striatum (VS) receives valuation information for a diverse set of rewarding stimuli. These range from simple primary reinforcers such as juice rewards over abstract social rewards to internally generated signals on perceived correctness, suggesting that the VS is involved in learning from trial-and-error irrespective of the specific nature of provided rewards. In addition, we summarize evidence that the VS can also be implicated when learning from observing others, and in tasks that go beyond simple stimulus-action-outcome learning, indicating that the reward system is also recruited in more complex learning tasks.

Keywords

ventral striatum; reward; feedback; learning; FMRI; human

© 2014 Elsevier Inc. All rights reserved.

Corresponding author with complete address: Reka Daniel, Princeton Neuroscience Institute, Green Hall, Princeton University, Princeton, NJ 08450, USA, rekad@princeton.edu, Phone: +1 609 258 5456.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1 Introduction

As any living organism, humans are faced with the need to make decisions about how to act in response to a plethora of environmental cues every day. Often, we encounter similar situations repeatedly, which enables us to use past experiences to predict future outcomes (Cohen, 2008). Learning from trial and error, or reinforcement learning, has been extensively studied in the last decades. Normative computational models have proven to be successful in explaining learning in terms of a reward prediction error, i.e., a mismatch between predicted and actual rewards. On a neural level, it has been shown that the dopaminergic midbrain neurons, with their massive projections to the ventral striatum (VS), represent this reward prediction error and play a central role in reward-based learning (for reviews see Delgado (2007); Niv and Montague (2008); O’Doherty, Hampton, and Kim (2007); Schultz (2007)).

However, in ecologically valid settings, decision problems vary hugely in complexity: organisms may have to decide which objects in the environment to categorize as nutrition, but also which partner to choose, or, in the human case, what career path to follow, and how to provide for retirement. These problems also vary in the rewards they provide: from simple primary and secondary reinforcers such as food, the option to reproduce, and money, to more abstract rewards such as love, social approval and (financial) stability. For many of these decisions, feedback from the environment is sparse and delayed (Hogarth, 2006). In addition, learning from errors is costly, and their commission is often avoidable. In these situations we have to additionally rely on other mechanisms than learning from trial and error, such as building abstract structural representations (Ribas-Fernandes et al., 2011; Diuk, Tsai, Wallis, Botvinick, & Niv, 2013), mental models of the environment (Daw, Gershman, Seymour, Dayan, & Dolan, 2011; Gläscher, Daw, Dayan, & O’Doherty, 2010) or learning from others (Burke, Tobler, Baddeley, & Schultz, 2010). Many species are able to learn from such indirect experience, i.e., by observing the outcome of others and imitating their actions (Chamley, 2004). Humans in particular have developed a complex form of communication, allowing them to pass on learned information by providing instructions (Li, Delgado, & Phelps, 2011).

The neural substrates of learning from abstract, incomplete or absent rewards are only currently being investigated. Humans are the ideal species to study these tasks: they are willing to work without explicit rewards after each trial, which facilitates the study of observational learning, and can be verbally instructed to examine the influence of prior knowledge. In addition, experimental methodology to study higher cognitive functions is well established, and a wealth of experimental paradigms, formal models and empirical behavioral data exists on learning in humans. Using functional magnetic resonance imaging (fMRI) it is possible to non-invasively study the brain activation of human participants performing complex learning and decision making tasks.

This review focuses on how central findings on the neural underpinnings of animal learning can be replicated and elaborated upon by examining different types of reward and more complex forms of learning in human participants. To this end, we present a summary of influential results in the animal literature, and discuss how activation data obtained using

fMRI relates to more direct measures of neural firing as acquired in electrophysiological recordings, and specifically to dopaminergic activity. We then go on to summarize recent findings on the neural substrates of learning in humans, and argue that phylogenetically old pathways that mediate simple stimulus-response learning are recruited even in the absence of explicit rewards to solve complex decision making tasks.

2 Reinforcement learning and the dopaminergic system

In order to optimally behave in any given environment it is crucial to determine which actions result in rewarding events given a specific state of the environment. These rewarding events come in many different shapes and flavors: even primary rewards can be as diverse as the opportunity to mate or a drop of juice. Any agent striving to maximize positive outcomes can benefit from a system that unifies these diverse sensory inputs by encoding the motivational properties of stimuli, thereby providing a “retina of the reward system” (Schultz, 2007; Schultz, 2008).

2.1 The dopaminergic system: A system for reward and motivation

The seminal experiments of Olds and Milner (1954) were a first step towards identifying a system in the mammalian brain that is dedicated to processing motivational information. They described several brain sites in the rat where direct electrical stimulation acted as a reinforcer, inciting the animal to stimulate itself. Subsequent experiments showed that the tissue inducing the highest rates of self-stimulation is located in the medial forebrain bundle (MFB), which is connected to dopamine cell bodies in the ventral tegmental area (VTA) (Gallistel, Shizgal, & Yeomans, 1981), and that self-stimulation of the MFB is related to fluctuations in dopamine levels in one of the prominent targets of mesolimbic dopamine neurons, the nucleus accumbens (Garris, Kilpatrick, & Bunin, 1999; Hernandez et al., 2006). Gallistel et al. (1981) point out three central properties of behavior during self-stimulation: rats engage in self-stimulation with high vigor, i.e., a short latency and high intensity of responding. In addition, they pursue self-stimulation over other vital goals, and adjust their behavior flexibly to the magnitude of recent stimulation. These three properties can be viewed as hallmarks of motivation: motivation energizes the organism, directs its behavior towards a goal, and enables learning about outcomes (Gallistel et al., 1981).

2.2 The dopaminergic system in learning

Dopaminergic neurotransmission has been associated with a wide variety of functions, many of which are important for optimizing behavior, i.e., for maximizing reward and minimizing aversive outcomes. Different firing modes of dopaminergic neurons have been associated with the facilitation of a wide range of motor, cognitive, and motivational processes (Schultz, 2008; Bromberg-Martin, Matsumoto, & Hikosaka, 2010), with determining the strength and rate of responding (Niv, Daw, Joel, & Dayan, 2007), and with signaling different aspects of salience, including appetitive and aversive information (Schultz, 2007; Bromberg-Martin et al., 2010), novelty (Bunzeck & Düzel, 2006; Ljungberg, 1992; Wittmann, Bunzeck, Dolan, & Düzel, 2007), and contextual deviance (Zaehle et al., 2013). However, the majority of recent research has focused on one specific aspect of the firing of dopaminergic neurons: 75–80% of the cells convey a signal that is ideally suited to promote

learning (Schultz, 2008). They respond with short-latency phasic bursts to unpredicted rewards and reward-predicting stimuli. Importantly, when rewards are fully predicted no response is observed, while the omission of predicted rewards leads to a decrease in baseline activity (Schultz, 2002)(see Fig. 1A). This behavior corresponds to a reward prediction error signal, i.e., to a coding of reward information as the difference between received and expected reward (Montague, Dayan, & Sejnowski, 1996; Schultz, Dayan, & Montague, 1997).

A reward prediction error signal has also been postulated as a teaching signal in normative theories of reinforcement learning (Bertsekas & Tsitsiklis, 1996; Sutton & Barto, 1990). These theories focus on providing normative accounts of how agents can optimize their behavior (Niv & Montague, 2008). A wide variety of reinforcement learning algorithms exists, however most of them share some core features: they predict that whenever faced with a decision, the agent calculates a value for each available option. In order to allow for random and exploratory behavior, these values are subsequently passed through a probabilistic function before the stimulus or action with the highest value is chosen. Whenever new information becomes available, e.g. in the form of an unpredicted reward or the omission of an expected reward, the values are updated using the reward prediction error multiplied by a learning rate (Cohen, 2008).

Such a mechanism for calculating the expected subjective value of future states of the environment and updating it using a simple prediction error signal would benefit any living organism: it provides a mechanism to make predictions based on the similarity between past experiences and the current state of the environment, enabling the organism to make informed decisions. Whenever these predictions are violated, they can be updated using a single scalar signal, without having to (re-)process all available sensory information (Schultz, 2008). A correspondence between the prediction error signals postulated by normative models and phasic dopaminergic firing was shown quantitatively (Bayer & Glimcher, 2005; Tobler, Fiorillo, & Schultz, 2005) and using several different paradigms, including blocking (Waelti, Dickinson, & Schultz, 2001) and conditioned inhibition (Tobler, Dickinson, & Schultz, 2003). This postulated role of short-latency firing of dopaminergic midbrain neurons in signaling prediction errors has been termed the reward prediction error hypothesis of dopamine (for extensive reviews see Schultz (2002); Schultz (2006)).

One potential mechanism by which phasic dopamine bursts could support learning is suggested by the observation that long-term potentiation in the striatum depends not only on strong pre- and postsynaptic activation, but also on dopamine release (this is sometimes referred to as the three-factor learning rule (Ashby & Ennis, 2006)). According to this model, the same dendritic spines of striatal medium spiny cells are contacted both by glutamatergic input from cortical neurons and dopaminergic input from midbrain neurons. Synaptic transmission is strengthened only on those neurons that simultaneously receive cortical input, coding some aspect of an event in the environment, and dopaminergic input, signaling via the reward prediction error that adjustments are necessary (Schultz, 2002). Hence, on the basis of contingent feedback, dopaminergic projections to the striatum are able to modify behavior in response to salient stimuli.

Indeed, using optogenetic activation of dopamine neurons, a causal role of dopaminergic signals in cue-reward learning has been established in rats (Steinberg et al., 2013). In humans, both striatal fMRI activations in healthy populations and impairments in feedback-based learning in patients with striatal dysfunctions were reported, indicating a critical role of cortico-striatal loops in feedback-based learning (Ashby & Ennis, 2006; Shohamy et al., 2004).

2.3 Outlook

The reward prediction error hypothesis of dopamine is not unchallenged (see e.g. (Berridge, 2007; Redgrave, Gurney, & Reynolds, 2008; Vitay, Fix, Beuth, Schroll, & Hamker, 2009)). However, it has inspired an extensive body of experimental research on reward-based learning, and models derived from it have proven to be able to predict behavior and neural activations across species and for a wide range of reinforcing stimuli and experimental paradigms. Even in their most simplistic form reinforcement learning models provide a parsimonious account of many behavioral phenomena and neural activation patterns. More extended models can account for complex phenomena observed in reward-based learning, such as the sensitivity to risk (Niv, Edlund, Dayan, & O'Doherty, 2012), or the ability to learn from rewards that were foregone (Fischer & Ullsperger, 2013; Lohrenz, McCabe, Camerer, & Montague, 2007), making these models an invaluable tool to derive quantitative trial-by-trial predictions about learning (Daw, 2011).

3 Examining the brain in humans using fMRI: the relationship between the BOLD signal, neuronal firing, and dopaminergic activity

In this review we present recent findings on learning-related activations in one of the most prominent targets of the dopaminergic midbrain neurons, the VS, in humans. While most experiments on dopaminergic activity in non-human animals are conducted using electrophysiological recordings, these are rarely available from human participants due to their highly invasive nature (exceptions are reported, e.g., in Münte et al. (2007); Cohen et al. (2009); Zaehle et al. (2013)). Functional magnetic resonance imaging (fMRI) can be used to safely visualize brain activation from the whole brain continuously while the participant is performing a task, providing both information about the network supporting task performance and how the involvement of that network changes over time. However, fMRI only provides an indirect measure of neuronal activity and no information about the involved neurotransmitter systems. In the present section we discuss how the fMRI signal relates to neuronal activation, and under which circumstances a relationship between ventral striatal fMRI activation and dopaminergic activity can be assumed.

3.1 Understanding the fMRI signal

In order to relate results from fMRI experiments to those obtained using electrophysiological recordings, it is crucial to understand the nature of the fMRI signal. However, despite the ever-growing popularity of fMRI to measure brain activation since its introduction in the early 1990s (Bandettini, Wong, Hinks, Tikofsky, & Hyde, 1992; Belliveau et al., 1991; Frahm, Bruhn, Merboldt, & Hänicke, 1992; Kwong, 1992; Ogawa, Lee, Kay, & Tank, 1990; Ogawa, 1992), the exact origin of the fMRI signal is still being

investigated. As fMRI technology cannot provide direct information about the electrical activity of neurons, it has to rely on an indirect index based on the metabolic changes induced by neuronal activity. The BOLD (blood oxygen level dependent) contrast relies on the fact that hemoglobin, the oxygen-carrying molecule within red blood cells, has different magnetic properties depending on its oxygenation state, which allows it to act as an endogenous contrast agent. This contrast is assumed to reflect changes in oxygen consumption in combination with changes in cerebral blood flow and volume (Buxton, Uluda, Dubowitz, & Liu, 2004). These hemodynamic responses are supposed to be elicited by local increases in neural activity (Logothetis & Wandell, 2004; Logothetis, 2010); however, the coupling mechanisms between the BOLD signal and neural activity are complex and depend on factors like the link between the neuronal response and increased energy demands, the mechanisms that signal the energy demand, and the mechanisms responsible for triggering the subsequent hemodynamic response (Arthurs & Boniface, 2002; Logothetis, 2003).

One way to empirically investigate the origin of the BOLD contrast are simultaneous cortical electrophysiological and fMRI recordings. Results from such recordings suggest that the BOLD signal is slightly better explained by local field potentials than neuronal spiking activity (Goense & Logothetis, 2008; Kayser, Kim, Ugurbil, Kim, & König, 2004; Logothetis, Pauls, Augath, Trinath, & Oeltermann, 2001; Mukamel et al., 2005; Raichle & Mintun, 2006). Local field potentials are assumed to not be driven by the activation of output neurons of a given area, but rather by input to that area, as well as local interneuronal processing, which includes both inhibitory and excitatory activations (Logothetis, 2003; Logothetis & Wandell, 2004). In addition, neuromodulatory signals as carried by e.g. acetylcholine, serotonin, and dopamine can also influence the BOLD signal directly (Attwell & Iadecola, 2002; Logothetis, 2008).

The indirect nature of the BOLD-fMRI signal has to be kept in mind whenever fMRI activity is interpreted as neuronal activation. This is especially true when it is interpreted as reflecting the activity of specific neurotransmitter / neuromodulator systems, as is often done in studies of reward-based learning, which is intimately tied to the function of the dopaminergic system.

3.2 A potential link between BOLD response and dopaminergic activation

As detailed above, fMRI is not able to measure changes in dopamine release directly. However, there are indications that the BOLD response and dopaminergic neuron firing are correlated in midbrain areas containing a large proportion of dopaminergic neurons, i.e. the substantia nigra (SN) / ventral tegmental area (VTA), and in regions receiving massive projections from these areas, like the VS.

Injections of dopamine-releasing agents can increase the BOLD signal in the nucleus accumbens, and it has been suggested that the signal has the capacity to also capture signals induced by endogenous dopaminergic firing (Knutson & Gibbs, 2007). Consistent with this assumption, fMRI studies show activations at or near dopaminergic midbrain nuclei and the VS that correlate with both reward expectation and reward prediction errors, and thereby exhibit a striking similarity to the pattern of burst firing of dopamine neurons observed in

electrophysiological recordings (D'Ardenne, McClure, Nystrom, & Cohen, 2008; Cohen, 2008; O'Doherty et al., 2007; Knutson, Taylor, & Kaufman, 2005; Niv & Montague, 2008) (see Fig. 1).

More direct evidence for the modulation of ventral striatal fMRI signals by dopamine was provided by Pessiglione, Seymour, Flandin, Dolan, and Frith (2006). In their instrumental learning paradigm, the administration of drugs that enhance or reduce dopaminergic function influenced the magnitude of BOLD-fMRI activation in the striatum in response to reward prediction errors (for similar results see van der Schaaf et al. (2012)). Subsequently, a quantitative role of dopamine in the fMRI signal was established by showing a correlation between reward-related dopamine release, as indexed by [¹¹C]raclopride PET, and fMRI activity in the SN / VTA and the VS across the same participants (Schott et al., 2008). An indication for the role of dopamine in the dorsal striatum was provided by a recent study by Schönberg et al. (2010). They showed that in patients with Parkinson's disease, which is characterized by a loss of dopaminergic input to mainly the dorsal striatum, fMRI activation in response to reward prediction errors in the dorsal putamen is significantly decreased compared to responses observed in control participants.

Although the studies outlined above provide evidence for a link between dopaminergic neurotransmission and fMRI activation during reward-learning, some caveats apply. The reviewed results do not rule out that this link is mediated by non-dopaminergic processes, like glutamatergic costimulation. In addition, this evidence was only provided for the expectation and processing of (1) rewards in (2) the SN / VTA and striatum. As fMRI cannot isolate dopamine-related activity from the effects of other afferents, the BOLD signal and dopamine release might dissociate in other situations. For example, the receipt of aversive stimuli or the omission of expected rewards might result in increased inputs to inhibitory interneurons in the SN / VTA, and thereby elicit an increased fMRI signal in these areas (Düzel et al., 2009). Indeed, positive midbrain activations in response to negative feedback and positive errors in the prediction of punishment have been observed (A. R. Aron et al., 2004; Menon et al., 2007). In a recent meta-analysis of 142 fMRI studies on reward based learning, Liu, Hairston, Schrier, and Fan (2011) report that although responses to positive outcomes tend to be stronger than responses to negative outcomes, the VS is consistently activated by both positive and negative outcomes. Similarly, dopaminergic drugs affect learning from rewards and punishments differentially, while fMRI signals in the striatum in response to reward and punishment prediction errors are altered in the same direction (van der Schaaf et al., 2012). These observations underscore that, given the indirect nature of the fMRI signal, the direction of activations in response to positive and negative outcomes needs to be interpreted with care. Additionally, the reported results on the relationship between dopaminergic activity and the fMRI signal in response to reward prediction errors can also not be transferred to other target areas of the dopaminergic system, e.g. the orbitofrontal cortex or the medial prefrontal cortex, which are less physiologically homogenous and show different pharmacokinetics (Knutson & Gibbs, 2007).

3.3 Summary and conclusions

fMRI cannot assess the exact neuronal mechanisms underlying the studied tasks. The signal changes it measures are however a response to local changes in overall neural activity, and it allows measuring these signal changes with a moderately high spatiotemporal resolution – enabling event-related signal analyses – throughout the whole brain, including deep structures that are central to reinforcement learning like the dopaminergic midbrain nuclei and the ventral striatum. For these structures evidence exists that links fMRI activation during reward processing to dopamine (Düzel et al., 2009; Knutson & Gibbs, 2007; Pessiglione et al., 2006; Schott et al., 2008; van der Schaaf et al., 2012). In combination with its non-invasive nature, these findings make fMRI the best currently available tool to study the neural basis of reward-based learning in healthy human participants.

4 Ventral striatal activation in reward learning

Using fMRI, a wide range of rewards that have been shown to activate areas of the human dopaminergic system. However, when discussing the results of this body of research, two important limitations have to be considered.

Firstly, the reward prediction error hypothesis of dopamine postulates that dopaminergic midbrain neurons and their targets respond at two timepoints during a learning trial: (1) in response to the appearance of the reward predicting stimulus, and (2) in response to the receipt or omission of the expected reward. According to this hypothesis the first response is in proportion to the (subjective) value of the reward predicted by stimuli, while the second response reflects the difference between expected and received rewards. Correlates of these responses have been observed in both electrophysiological recordings of dopaminergic midbrain neurons (Matsumoto & Hikosaka, 2009; Tobler et al., 2005) and fMRI data in the striatum (Abler, Walter, Erk, Kammerer, & Spitzer, 2006; Haruno & Kawato, 2006; Cohen, 2007) (see Fig. 1). In line with these findings, in their meta-analysis Liu et al. (2011) report that the VS is activated during both the expectation and receipt of rewards, with potentially stronger activations in the latter phase. However, signals during reward expectation and receipt are highly correlated (Niv, 2009) and not addressed separately in many of the reported studies. For the purposes of the current overview we will therefore discuss both stages of reward-processing in parallel.

As a second limitation it has to be noted that although several fMRI studies report reward-related activity in the vicinity of the bodies of dopaminergic neurons in the SN/VTA (see e.g. Adcock, Thangavel, Whitfield-Gabrieli, Knutson, and Gabrieli (2006); D'Ardenne et al. (2008); Dreher, Kohn, and Berman (2006); O'Doherty, Deichmann, Critchley, and Dolan (2002); O'Doherty, Buchanan, Seymour, and Dolan (2006); Schott et al. (2007); Wittmann et al. (2005)), fMRI activations in response to reward are more consistently observed in the VS, one of the most prominent target areas of dopaminergic projections. Potential reasons for this observation include the nature of the fMRI BOLD signal (i.e., the signal might reflect the input to an area more than its output; see previous section) as well as several imaging difficulties that are specific to the midbrain (D'Ardenne et al., 2008). In the present section we therefore focus on recent findings on the different types of reinforcing stimuli that have been shown to evoke fMRI activations in the VS.

4.1 From primary to secondary reinforcers

Different types of reinforcers are traditionally categorized according to the way they have acquired their ability to strengthen the behavior they were consequent to. Primary or unconditioned reinforcers such as food, drink, and sex have a biological/evolutionary basis, and act as positive reinforcers without any antecedent learning. On the other hand, secondary or conditioned reinforcers acquire their ability to strengthen behavior by being associated with unconditioned reinforcers. The simplest example of a conditioned reinforcer is a previously neutral stimulus, such as a light, which is consistently paired with a primary reinforcer, such as juice, within the context of an experiment.

Although the categorization into primary and secondary reinforcers is a useful theoretical concept, in many settings the distinction between innate and learned responses is more gradual (see Fig. 2). For example, strictly speaking the flavor of most foods has to be considered a conditioned reinforcer which has acquired its value by being associated with the food's nutritional postingestive effects (Sclafani, 1999). Additionally, during the lifetime of a person some stimuli are paired with a many other reinforcers, under a wide variety of conditions. They are therefore reinforcing in many contexts, and therefore have properties that are similar to those of primary reinforcers. These stimuli are sometimes referred to as generalized conditioned reinforcers, the most frequently cited examples of which are tokens such as money or school grades; however, many social stimuli such as attention, affection and approval, which consistently precede the receipt of primary reinforcers, have to be considered generalized conditioned reinforcers, too (Skinner, 1953). Partially conditioned stimuli fall in-between the conditioned/unconditioned categories. For example, the reinforcing effect of the ability to control the physical world around us through acquiring skills or information seems to be largely unconditioned, whereas the reinforcing effects of beauty in music or art have been postulated to be largely conditioned (Skinner, 1953).

Generalized conditioned reinforcers, and specifically monetary rewards, are administered in many fMRI studies of reward-based learning. Monetary rewards have several advantages, including that they are easy to administer and quantify, that they are likely not subject to devaluation over the course of the experiment, and that participants do not have to be liquid- or food-deprived for the experiment. Consequently, several studies and meta-analyses are available that compare the effects of monetary rewards to other classes of reinforcers. In the present section, we will discuss three frequently administered classes of reinforcers (primary, social and cognitive) and will assess their relationship with monetary rewards.

Primary rewards—In line with electrophysiological recordings from dopaminergic midbrain neurons, the VS is activated by primary rewards such as juice or glucose solution both during classical and instrumental conditioning (O'Doherty et al., 2002; Pagnoni, Zink, Montague, & Berns, 2002). In a direct comparison with monetary rewards, fMRI activation in the nucleus accumbens was observed to correlate more strongly with errors in the prediction of monetary than juice rewards (Valentin & O'Doherty, 2009). A recent meta-analysis of 206 fMRI studies on reward reports overlapping activation foci for primary and monetary rewards, however with more densely clustered foci in the nucleus accumbens for monetary rewards (Bartra, McGuire, & Kable, 2013), again suggesting stronger or more

robust effects in response to monetary rewards. Similarly, in their meta-analysis of 87 studies with monetary, erotic and food rewards, Sescousse, Caldú, Segura, and Dreher (2013) report stronger activations for monetary rewards than for both primary rewards.

Taken together, these results might indicate that the nucleus accumbens is more involved in the processing of secondary rewards. However, Sescousse et al. (2013) also report two other important differences between studies administering primary vs. secondary rewards. While monetary rewards are often employed in studies that use probabilistic rewards and instrumental conditioning, primary rewards are often administered according a fully predictable schedule and in a passive observation setting. These differences might indicate that the VS is especially activated in response to surprising rewards, as postulated by the reward prediction error hypothesis, and in settings that require learning the optimal response to maximize rewards. Additionally, it has to be assumed that human participants are aware that after a short delay after the experiment they will have nearly unlimited access to primary rewards such as food, drink or sexually arousing images. Therefore, the actual subjective value of primary rewards offered in an experimental setting is unclear and might be lower than that of monetary rewards, indicating that the observed lower activations in the VS in response to primary than to monetary rewards might merely reflect a difference in subjective values.

Social rewards—Although they are sometimes treated as single class of reinforcers, social rewards span the whole continuum from unconditioned to conditioned reinforcers. While erotic images are often viewed as primary reinforcers, smiling faces and a gaze that is directed at the observer also carry information about approval and attention and therefore have properties of a conditioned reinforcer.

Studies investigating social rewards have consistently reported fMRI activation in the striatum or dopaminergic midbrain nuclei in response to as diverse stimuli as facial attractiveness in the opposite sex (Aharon et al., 2001; Cloutier, Heatherton, Whalen, & Kelley, 2008), gaze direction (Kampe, Frith, Dolan, & Frith, 2001), or images of romantic partners (Acevedo, Aron, Fisher, & Brown, 2012; A. Aron et al., 2005). In a direct comparison between monetary and social rewards, images of smiling faces engaged the VS at a similar locus as monetary rewards, but only at a lower threshold (Lin, Adolphs, & Rangel, 2012).

Interestingly, VS activation in response to social situations is not restricted to the processing of positive social cues. It also correlates with behavior in complex social tasks involving cooperation, norm-abiding and altruism. For example, the decision to donate to charitable organizations activates the VS (Harbaugh, Mayr, & Burghart, 2007; Moll, Krueger, & Zahn, 2006), especially when this donation is observed by others (Izuma, Saito, & Sadato, 2010). For a more complete coverage of the literature on social decision-making we refer to recent reviews by Rilling and Sanfey (2011) and Bhanji and Delgado (2013).

Cognitive feedback—Potentially because of both its effectiveness in guiding learning and the ease with which it is administered, cognitive feedback, or giving information about the correctness of an answer, is also often employed as a reinforcer in fMRI experiments

with human participants. Given the importance of prediction errors for learning (Schönberg, Daw, Joel, & O'Doherty, 2007), it has also been suggested that the processing of information (Bromberg-Martin & Hikosaka, 2011; Niv & Chan, 2011) and cognitive feedback (A. R. Aron et al., 2004) also involve the reward circuitry. Cognitive feedback is often interpreted as a form of social approval and therefore a generalized conditioned reinforcer comparable to monetary rewards, but it also has a partially biological component of guiding the acquisition of a skill that can help the participant control the environment of the experiment. Psychologically, monetary rewards can be viewed as a form of extrinsic motivation and are known to undermine intrinsic motivation, which is not necessarily true for cognitive feedback (Murayama, Matsumoto, Izuma, & Matsumoto, 2010). We directly compared the neural correlates of monetary rewards with cognitive feedback using two parallel learning tasks, providing participants with either (1) information about the correctness of their decision, or (2) a monetary reward for every correct answer. While several dopaminergic projection sites were activated in both task versions, the nucleus accumbens as only structure was activated more strongly in response to cues predicting monetary reward compared to cues predicting cognitive feedback (see Fig. 3A). While this effect can be explained by the higher subjective value of monetary rewards, a closer examination of the nucleus accumbens activation indicated that it correlated with questionnaire items measuring intrinsic motivation in the cognitive feedback task, while the activation correlated with items measuring extrinsic motivation in the monetary reward task (Daniel & Pollmann, 2010). This suggests that reward-related processing in the ventral striatum is dependent on the motivational states induced by the specific type of reward. Indeed, when intrinsic motivation is undermined by providing extrinsic rewards, striatal activation to successful task performance is lower (Murayama et al., 2010).

Summary and caveats—FMRI results show that the human VS responds at the same locus to the anticipation and receipt of a wide range of rewards, spanning the whole spectrum from unconditioned to conditioned reinforcers, including abstract concepts such as beauty and information about the correctness of an answer (for meta-analyses see Bartra et al. (2013); Clithero and Rangel (2013); Kühn and Gallinat (2012)). Yet, not all reinforcers are equivalent, and different types/strengths of reinforcers can induce different motivational states. In general, monetary reward seems to elicit the strongest activations across the reviewed categories. In addition, monetary rewards have also been shown to shift participants to an extrinsically motivated state, which alters both behavior and ventral striatal activation patterns (Murayama et al., 2010; Daniel & Pollmann, 2010). Further research is needed to determine whether these two effects are qualitatively inherent to monetary reinforcement, or whether they can be explained by the high subjective value of monetary rewards.

Given the reliable but varied response of the VS to a wide range of reinforcers, it is tempting to use VS activation strength as an index for the value of a reinforcer. However, as fMRI activation in the VS has also been shown to respond to non-rewarding events such as negative outcomes or saliency (Jensen et al., 2007; Menon et al., 2007; Litt, Plassmann, Shiv, & Rangel, 2011; Metereau & Dreher, 2013; Robinson, Overstreet, Charney, Vytal, & Grillon, 2013; Seymour et al., 2004; Zink, Pagnoni, Chappelow, Martin-Skurski, & Berns,

2006), this approach has to be interpreted with extreme care (Poldrack, 2006) and should always be accompanied by behavioral measures of reinforcer value.

The presented results indicate that the dopaminergic system plays a central role in processing a wide range of stimuli that can act as reinforcers in instrumental conditioning tasks. Recently the dopaminergic system has also been shown to be involved in learning outside the context of instrumental conditioning, e.g. when learning from observation. These results are discussed in the next section.

4.2 Indirect learning: learning from observation and instruction

When learning from a direct interaction with the environment, an agent experiences events and outcomes that are relevant to himself, and, in the case of instrumental learning, takes actions to influence them. However, this learning from trial-and-error can be costly: if one approach provides unsatisfactory results, a new course of action has to be initiated, and sometimes the opportunity to obtain reward or avoid punishment has been permanently lost. Making matters worse, in many environments immediate feedback is not available after each action, or is non-deterministic, so that the exploration behavior has to be repeated many times.

One way to reduce the cost of exploration is to gather information by observing how others act, and what the consequences of their actions are. While this indirect learning will not result in immediate rewards, it can help to learn about the structure of the environment and therefore inform future actions without having to directly engage in costly or dangerous behavior (Rendell et al., 2010). Observational learning has been shown in many species, ranging from penguins to monkeys (Chamley, 2004; Chang, Winecoff, & Platt, 2011; Biederman, Robertson, & Vanayan, 1986; Heyes & Dawson, 1990; Meunier, Monfardini, & Boussaoud, 2007; Myers, 1970; Range, Viranyi, & Huber, 2007; Subiaul, Cantlon, Holloway, & Terrace, 2004). In humans, learning from observing others, i.e. social learning, is considered one of the most central and basic forms of cognition (Bandura, 1977; Frith & Frith, 2012). In addition, as humans are capable of symbolic communication, decision tasks can also be mastered by following the instructions of other, more experienced individuals.

Behaviorally, it is known that observational learning can be less effective than trial-and-error learning for complex multidimensional tasks (Ashby & Maddox, 2005). However, it does reliably increase performance in a wide variety of tasks, including in tasks that are traditionally studied in the context of trial-and-error learning: in a two-armed bandit task observing the actions of a confederate leads to an improvement in the number of optimal choices; additionally observing the confederates' outcomes increases the effect (Burke et al., 2010). Similarly, participants are able to utilize verbally instructed information on probabilities in a betting game (Li et al., 2011), and can learn to integrate information across different dimensions in a categorization task simply by observing stimulus-stimulus associations (Daniel & Pollmann, 2012). Recently it has been suggested that even behavioral automaticity can be induced by observational learning (Liljeholm, Molloy, & O'Doherty, 2012).

The neural networks that underlie observational learning are similar to those that support learning from direct experience (Dunne & O'Doherty, 2013), possibly reflecting an attempt to simulate the observed person's mental states (Suzuki et al., 2012). However, mixed results have been reported on prediction-error related activity in the ventral striatum during indirect learning: while some studies show normal or slightly weakened responses to observational prediction errors (Bellebaum, Jokisch, Gizewski, Forsting, & Daum, 2012; Daniel & Pollmann, 2012; Kätsyri, Hari, Ravaja, & Nummenmaa, 2013), others show no (Cooper, Dunne, Furey, & O'Doherty, 2012; Suzuki et al., 2012) or inverted (Burke et al., 2010) prediction error related activity.

One possible explanation for these inconsistent results is that the VS responds to errors in reward prediction rather than to the errors themselves: activation in the VS was shown to increase when the agent receives unpredicted reward, irrespective of whether this is caused by erroneous or correct responses of another person (de Bruijn, de Lange, von Cramon, & Ullsperger, 2009). Therefore, it is possible that VS activation in response to positive errors of reward prediction is stronger if the observer is able to benefit from the rewards of the observed agent. In line with this argumentation, ventral striatal reward-related activation was reported when the observer's rewards were coupled to the observed rewards of another person (Bellebaum et al., 2012; Kätsyri et al., 2013). The usefulness for future own actions of the observed information may also influence VS activation: no activation related to prediction-errors of an observed person has been shown when the task of the observed person is different from the observer's task (Suzuki et al., 2012). In contrast, studies in which directly relevant information can be gathered via observation show VS responses even in the absence of rewards during observation (Daniel & Pollmann, 2012; Burke et al., 2010). The suggestion that the nucleus accumbens is only involved in observation whenever the observer is learning is also well in line with results showing that electrical stimulation of the nucleus accumbens in mice enhances the acquisition of behavior when learning from observation (Jurado-Parras, Gruart, & Delgado-García, 2012). The results of Li et al. (2011) provide a direct test for this account: fully reliable instructions on reward probabilities decreased VS responses to rewards, potentially because the reward carried no further useful information for learning. This decrease in responsiveness was paralleled by learning rates close to zero and increased negative functional connectivity of the VS with a dorsolateral prefrontal area. Other potential factors that might modulate the usefulness of observed information, and therefore VS implication, include social factors like the trustworthiness of the confederate, and his similarity to the agent himself (Delgado, Frank, & Phelps, 2005; Mobbs et al., 2009).

To investigate whether the VS is involved in task performance after information has been acquired observationally, we trained participants on a categorization task (Daniel & Pollmann, 2012). During training they passively observed category label-stimulus pairs; in this phase VS activation decreased with time, supporting the assumption that the striatum plays a role in encoding novelty (Bunzeck & Düzzel, 2006) and stimulus-stimulus associations (den Ouden, Friston, Daw, McIntosh, & Stephan, 2009). In the subsequent retrieval phase, no rewards were obtainable, and no behavioral learning effects were observed. Nevertheless, activation in the VS followed a prediction error on confidence: it increased when the participant was more confident than expected based on the confidence in

previous decisions, and decreased when the participant was less confident than expected (see Fig. 2B). This indicates that even in the absence of external feedback, activation in the VS tracks confidence in decisions, which is inherently related to the ability to successfully navigate the task environment to obtain potential rewards.

In summary, the VS is not unconditionally activated during observational learning by simulating the rewards of confederates; rather, activations in response to the confederate's rewards are modulated by several factors such as implications for the observer himself (Daniel & Pollmann, 2012) or the similarity between the confederate and observer (Mobbs et al., 2009). In contrast, the VS is activated during indirect learning when task-relevant information can be acquired from observing the outcomes of others (Bellebaum et al., 2012; Burke et al., 2010), or even when observing task-relevant stimulus-stimulus associations (Daniel & Pollmann, 2012). During subsequent task performance, the VS might be involved in simulating or anticipating positive outcomes (Daniel & Pollmann, 2012), which are critical processes in supporting correct task performance and goal-attainment (Han, Huettel, Raposo, Adcock, & Dobbins, 2010). Additionally, VS activation in response to outcome information is attenuated when reliable instructed information is available (Li et al., 2011), further supporting the interpretation that the VS is mainly involved in processing task-relevant trial-by-trial information.

5 Summary and future directions

Traditionally, the dopaminergic system with its strong projections to the VS has been assumed to support the acquisition of new behaviors based on trial-by-trial rewards (Ashby & Maddox, 2005; Shohamy, Myers, Kalanithi, & Gluck, 2008). However, recent fMRI studies on human participants show that the VS is activated not only when learning from explicit rewards, but also from a wide range of stimuli that can be considered as feedback, including cognitive information (Daniel & Pollmann, 2010) or highly abstract concepts such as beauty (Salimpoor, Benovoy, Larcher, Dagher, & Zatorre, 2011). In addition, the VS is also implicated when observing other's actions and other's feedback, as well as in response to internally generated error signals, indicating that in human learning "reinforcement" has to be considered a very broad term.

Alternatively, the striatum may play a more general role in learning than just supporting the gradual acquisition of stimulus-response contingencies depending on reinforcement. Structures of the dopaminergic system, including the VS, have also been implicated in other forms of learning (Shohamy, 2011), including model-based learning (Daw et al., 2011), the formation and retrieval of episodic memories (Adcock et al., 2006; Han et al., 2010; Tricomi & Fiez, 2008) and complex social behaviors (Izuma et al., 2010; Phan, Sripada, Angstadt, & McCabe, 2010; Tricomi, Rangel, Camerer, & O'Doherty, 2010). Interestingly, recent results indicate that the VS might not only be involved in tracking errors in the prediction of the value of a reward (as would be expected if the VS were merely representing rewards in a common neural currency), but also in processing more complex surprises about the reward's unique identity (McDannald, 2012).

The summarized results are very suggestive of a central role of the human dopaminergic system in a wide range of learning and decision making tasks. However, for these results to be conclusive, further research has to be conducted on the causal role and nature of the observed fMRI activations. As has been done for the gradual acquisition of stimulus-response associations based on explicit rewards (Ashby & Maddox, 2005; Shohamy et al., 2008), fMRI studies exploring the role of the VS in the complex, partially uniquely human domains summarized here have to be accompanied by PET, patient, and pharmacological studies. This will allow to confirm whether the observed activations in these broader domains are indeed caused by dopaminergic input to the VS, and whether VS involvement is essential for the acquisition of the reported behaviors (Shohamy, 2011). In addition, characterizing the tasks which recruit the striatum is not sufficient to gain a full understanding of the role of the dopaminergic system in learning; to achieve this goal it is also crucial to determine the actual process by which the striatum supports learning, and to explore its interactions with other structures like the hippocampus (Wimmer & Shohamy, 2012) or the orbitofrontal cortex (Klein-Flügge, 2013; Takahashi et al., 2011).

This future research will provide a holistic understanding of how reward-related and reward-independent processes interact to produce the rich behavior that allows us to function in the complex environment we face every day, and of the processes that are disrupted in diseases which impair this functioning.

Acknowledgments

This work was supported by Deutsche Forschungsgemeinschaft, SFB 779/A4, and NIH grant R01MH098861.

References

- Abler B, Walter H, Erk S, Kammerer H, Spitzer M. Prediction error as a linear function of reward probability is coded in human nucleus accumbens. *NeuroImage*. 2006; 31(2):790–795. [PubMed: 16487726]
- Acevedo BP, Aron A, Fisher HE, Brown LL. Neural correlates of long-term intense romantic love. *Social Cognitive and Affective Neuroscience*. 2012; 7(2):145–159. [PubMed: 21208991]
- Adcock RA, Thangavel A, Whitfield-Gabrieli S, Knutson B, Gabrieli JDE. Reward-motivated learning: mesolimbic activation precedes memory formation. *Neuron*. 2006; 50(3):507–517. [PubMed: 16675403]
- Aharon I, Etcoff N, Ariely D, Chabris CF, O'Connor E, Breiter HC. Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron*. 2001; 32(3):537–351. [PubMed: 11709163]
- Aron A, Fisher H, Mashek DJ, Strong G, Li H, Brown LL. Reward, motivation, and emotion systems associated with early-stage intense romantic love. *Journal of Neurophysiology*. 2005; 94(1):327–37. [PubMed: 15928068]
- Aron AR, Shohamy D, Clark J, Myers CE, Gluck MA, Poldrack RA. Human midbrain sensitivity to cognitive feedback and uncertainty during classification learning. *Journal of Neurophysiology*. 2004; 92(2):1144–1152. [PubMed: 15014103]
- Arthurs OJ, Boniface S. How well do we understand the neural origins of the fMRI BOLD signal? *Trends in Neurosciences*. 2002; 25(1):27–31. [PubMed: 11801335]
- Ashby, FG.; Ennis, JM. The role of the basal ganglia in category learning. In: Ross, BH., editor. *The psychology of learning and motivation*. Vol. Vol. 46. New York, NY: Elsevier; 2006. p. 1-36.
- Ashby FG, Maddox W. Human category learning. *Annual Review of Psychology*. 2005; 56:149–178.
- Attwell D, Iadecola C. The neural basis of functional brain imaging signals. *Trends in Neurosciences*. 2002; 25(12):621–625. [PubMed: 12446129]

- Bandettini PA, Wong EC, Hinks RS, Tikofsky RS, Hyde JS. Time course EPI of human brain function during task activation. *Magnetic Resonance in Medicine*. 1992; 25(2):390–397. [PubMed: 1614324]
- Bandura, A. *Social learning theory*. Englewood Cliffs (NJ: Prentice-Hall; 1977.
- Bartra O, McGuire JT, Kable JW. The valuation system: A coordinate-based metaanalysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage*. 2013; 76:412–427. [PubMed: 23507394]
- Bayer HM, Glimcher PW. Midbrain Dopamine Neurons Encode a Quantitative Reward Prediction Error Signal. *Neuron*. 2005; 47(1):129–141. [PubMed: 15996553]
- Bellebaum C, Jokisch D, Gizewski ER, Forsting M, Daum I. The neural coding of expected and unexpected monetary performance outcomes: dissociations between active and observational learning. *Behavioural brain research*. 2012; 227(1):241–251. [PubMed: 22074898]
- Belliveau JW, Kennedy DN Jr, McKinstry RC, Buchbinder BR, Weisskoff RM, Cohen MS, ... Rosen BR. Functional mapping of the human visual cortex by magnetic resonance imaging. *Science*. 1991; 254(5032):716–719. [PubMed: 1948051]
- Berridge KC. The debate over dopamine's role in reward: the case for incentive salience. *Psychopharmacology*. 2007; 191(3):391–431. [PubMed: 17072591]
- Bertsekas, DP.; Tsitsiklis, JN. *Neuro-dynamic Programming*. London, UK: Athena; 1996.
- Bhanji JP, Delgado MR. The social brain and reward: social information processing in the human striatum. *Wiley Interdisciplinary Reviews: Cognitive Science*. 2013
- Biederman GB, Robertson HA, Vanayan M. Observational learning of two visual discriminations by pigeons: a within-subjects design. *Journal of the Experimental Analysis of Behavior*. 1986; 46(1): 45–49. [PubMed: 3746187]
- Bromberg-Martin ES, Hikosaka O. Lateral habenula neurons signal errors in the prediction of reward information. *Nature Neuroscience*. 2011; 14(9):1209–1216.
- Bromberg-Martin ES, Matsumoto M, Hikosaka O. Dopamine in motivational control: rewarding, aversive, and alerting. *Neuron*. 2010; 68(5):815–834. [PubMed: 21144997]
- Bunzeck N, Düzal E. Absolute coding of stimulus novelty in the human substantia nigra/VTA. *Neuron*. 2006; 51(3):369–379. [PubMed: 16880131]
- Burke CJ, Tobler PN, Baddeley M, Schultz W. Neural mechanisms of observational learning. *Proceedings of the National Academy of Sciences of the United States of America*. 2010; 107(32): 14431–14436. [PubMed: 20660717]
- Buxton RB, Uluda K, Dubowitz DJ, Liu TT. Modeling the hemodynamic response to brain activation. *NeuroImage*. 2004; 23(Suppl 1):S220–s233. [PubMed: 15501093]
- Chamley, CP. *Rational Herds: Economic Models of Social Learning*. New York, NY, USA: Cambridge University Press; 2004.
- Chang SWC, Wincoff AA, Platt ML. Vicarious reinforcement in rhesus macaques (*macaca mulatta*). *Frontiers in Neuroscience*. 2011; 5(27)
- Clithero J, Rangel A. Informatic parcellation of the network involved in the computation of subjective value. *Social Cognitive and Affective Neuroscience*. 2013
- Cloutier J, Heatherton TF, Whalen PJ, Kelley WM. Are attractive people rewarding? Sex differences in the neural substrates of facial attractiveness. *Journal of Cognitive Neuroscience*. 2008; 20(6): 941–951. [PubMed: 18211242]
- Cohen MX. Individual differences and the neural representations of reward expectation and reward prediction error. *Social Cognitive and Affective Neuroscience*. 2007; 2(1):20–30. [PubMed: 17710118]
- Cohen MX. Neurocomputational mechanisms of reinforcement-guided learning in humans: A review. *Cognitive, Affective, & Behavioral Neuroscience*. 2008; 8(2):113–125.
- Cohen MX, Axmacher N, Lenartz D, Elger CE, Sturm V, Schlaepfer TE. Neuroelectric signatures of reward learning and decision-making in the human nucleus accumbens. *Neuropsychopharmacology*. 2009; 34(7):1649–1658. [PubMed: 19092783]

- Cooper JC, Dunne S, Furey T, O'Doherty JP. Human dorsal striatum encodes prediction errors during observational learning of instrumental actions. *Journal of Cognitive Neuroscience*. 2012; 24(1): 106–118. [PubMed: 21812568]
- Daniel R, Pollmann S. Comparing the neural basis of monetary reward and cognitive feedback during information-integration category learning. *The Journal of Neuroscience*. 2010; 30(1):47–55. [PubMed: 20053886]
- Daniel R, Pollmann S. Striatal activations signal prediction errors on confidence in the absence of external feedback. *NeuroImage*. 2012; 59(4):3457–3467. [PubMed: 22146752]
- D'Ardenne K, McClure SM, Nystrom LE, Cohen JD. BOLD responses reflecting dopaminergic signals in the human ventral tegmental area. *Science*. 2008; 319(5867):1264–1267. [PubMed: 18309087]
- Daw, ND. Trial-by-trial data analysis using computational models. In: Delgado, MR.; Phelps, EQ.; Robbins, TW., editors. *Decision making, affect, and learning: Attention and performance*. Oxford University Press; 2011. p. 1-26.
- Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ. Model-Based Influences on Humans' Choices and Striatal Prediction Errors. *Neuron*. 2011; 69(6):1204–1215. [PubMed: 21435563]
- de Bruijn, ERa; de Lange, FP.; von Cramon, DY.; Ullsperger, M. When errors are rewarding. *The Journal of Neuroscience*. 2009; 29(39):12183–12186. [PubMed: 19793976]
- Delgado MR. Reward-Related Responses in the Human Striatum. *Annals Of The New York Academy Of Sciences*. 2007; 1104(1):70–88. [PubMed: 17344522]
- Delgado MR, Frank RH, Phelps EA. Perceptions of moral character modulate the neural systems of reward during the trust game. *Nature Neuroscience*. 2005; 8(11):1611–1618.
- den Ouden HEM, Friston KJ, Daw ND, McIntosh AR, Stephan KE. A dual role for prediction error in associative learning. *Cerebral Cortex*. 2009; 19(5):1175–1185. [PubMed: 18820290]
- Diuk C, Tsai K, Wallis J, Botvinick M, Niv Y. Hierarchical learning induces two simultaneous, but separable, prediction errors in human basal ganglia. *The Journal of Neuroscience*. 2013; 33(13): 5797–5805. [PubMed: 23536092]
- Dreher J-C, Kohn P, Berman KF. Neural coding of distinct statistical properties of reward information in humans. *Cerebral Cortex*. 2006; 16(4):561–573. [PubMed: 16033924]
- Dunne S, O'Doherty JP. Insights from the application of computational neuroimaging to social neuroscience. *Current Opinion in Neurobiology*. 2013; 23(3):387–392. [PubMed: 23518140]
- Düzel E, Bunzeck N, Guitart-Masip M, Wittmann BC, Schott BH, Tobler PN. Functional imaging of the human dopaminergic midbrain. *Trends in Neurosciences*. 2009; 32(6):321–328. [PubMed: 19446348]
- Fischer AG, Ullsperger M. Real and fictive outcomes are processed differently but converge on a common adaptive mechanism. *Neuron*. 2013; 79(6):1243–1255. [PubMed: 24050408]
- Frahm J, Bruhn H, Merboldt KD, Hänicke W. Dynamic MR imaging of human brain oxygenation during rest and photic stimulation. *Journal of Magnetic Resonance Imaging*. 1992; 2(5):501–505. [PubMed: 1392241]
- Frith CD, Frith U. Mechanisms of social cognition. *Annual Review of Psychology*. 2012; 63:287–313.
- Gallistel CR, Shizgal P, Yeomans JS. A portrait of the substrate for self-stimulation. *Psychological Review*. 1981; 88(3):228–273. [PubMed: 6264530]
- Garris P, Kilpatrick M, Bunin M. Dissociation of dopamine release in the nucleus accumbens from intracranial self-stimulation. *Nature*. 1999; 8569(1985):67–69. [PubMed: 10078530]
- Gläscher J, Daw N, Dayan P, O'Doherty J. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*. 2010; 66(4):585–595. [PubMed: 20510862]
- Goense JBM, Logothetis NK. Neurophysiology of the BOLD fMRI signal in awake monkeys. *Current Biology*. 2008; 18(9):631–640. [PubMed: 18439825]
- Han S, Huettel SA, Raposo A, Adcock RA, Dobbins IG. Functional significance of striatal responses during episodic decisions: recovery or goal attainment? *The Journal of Neuroscience*. 2010; 30(13):4767–4775. [PubMed: 20357127]

- Harbaugh WT, Mayr U, Burghart DR. Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science*. 2007; 316(5831):1622–1625. [PubMed: 17569866]
- Haruno M, Kawato M. Different neural correlates of reward expectation and reward expectation error in the putamen and caudate nucleus during stimulus-action-reward association learning. *Journal of Neurophysiology*. 2006; 95(2):948–959. [PubMed: 16192338]
- Hernandez G, Hamdani S, Rajabi H, Conover K, Stewart J, Arvanitogiannis A, Shizgal P. Prolonged rewarding stimulation of the rat medial forebrain bundle: neurochemical and behavioral consequences. *Behavioral neuroscience*. 2006; 120(4):888–904. [PubMed: 16893295]
- Heyes CM, Dawson GR. A demonstration of observational learning in rats using a bidirectional control. *The Quarterly journal of experimental psychology. B, Comparative and physiological psychology*. 1990; 42(1):59–71.
- Hogarth, RM. Is confidence in decisions related to feedback? Evidence from random samples of real-world behavior. In: Fiedler, K.; Juslin, P., editors. *Information sampling and adaptive cognition*. New York, NY: Cambridge University Press; 2006. p. 456-484.
- Izuma K, Saito DN, Sadato N. Processing of the incentive for social approval in the ventral striatum during charitable donation. *Journal of Cognitive Neuroscience*. 2010; 22(4):621–631. [PubMed: 19320552]
- Jensen J, Smith AJ, Willeit M, Crawley AP, Mikulis DJ, Vitcu I, Kapur S. Separate brain regions code for salience vs. valence during reward prediction in humans. *Human Brain Mapping*. 2007; 28(4): 294–302. [PubMed: 16779798]
- Jurado-Parras MT, Gruart A, Delgado-García JM. Observational learning in mice can be prevented by medial prefrontal cortex stimulation and enhanced by nucleus accumbens stimulation. *Learning & memory*. 2012; 19(3):99–106. [PubMed: 22354947]
- Kampe KK, Frith CD, Dolan RJ, Frith U. Reward value of attractiveness and gaze. *Nature*. 2001; 413(6856):589. [PubMed: 11595937]
- Kätsyri J, Hari R, Ravaja N, Nummenmaa L. Just watching the game ain't enough: striatal fMRI reward responses to successes and failures in a video game during active and vicarious playing. *Frontiers in Human Neuroscience*. 2013; 7(278)
- Kayser C, Kim M, Ugurbil K, Kim D-S, König P. A comparison of hemodynamic and neural responses in cat visual cortex using complex stimuli. *Cerebral Cortex*. 2004; 14(8):881–891. [PubMed: 15084493]
- Klein-Flügge M. Segregated Encoding of RewardIdentity and StimulusReward Associations in Human Orbitofrontal Cortex. *The Journal of Neuroscience*. 2013; 33(7):3202–3211. [PubMed: 23407973]
- Knutson B, Gibbs SEB. Linking nucleus accumbens dopamine and blood oxygenation. *Psychopharmacology*. 2007; 191(3):813–822. [PubMed: 17279377]
- Knutson B, Taylor J, Kaufman M. Distributed neural representation of expected value. *The Journal of Neuroscience*. 2005; 25(19):4806–4812. [PubMed: 15888656]
- Kühn S, Gallinat J. The neural correlates of subjective pleasantness. *NeuroImage*. 2012; 61(1):289–294. [PubMed: 22406357]
- Kwong KK. Dynamic Magnetic Resonance Imaging of Human Brain Activity During Primary Sensory Stimulation. *Proceedings of the National Academy of Sciences of the United States of America*. 1992; 89(12):5675–5679. [PubMed: 1608978]
- Li J, Delgado MR, Phelps Ea. How instructed knowledge modulates the neural systems of reward learning. *Proceedings of the National Academy of Sciences of the United States of America*. 2011; 108(1):55–60. [PubMed: 21173266]
- Liljeholm M, Molloy CJ, O'Doherty JP. Dissociable brain systems mediate vicarious learning of stimulus-response and action-outcome contingencies. *The Journal of Neuroscience*. 2012; 32(29): 9878–9886. [PubMed: 22815503]
- Lin A, Adolphs R, Rangel A. Social and monetary reward learning engage overlapping neural substrates. *Social Cognitive and Affective Neuroscience*. 2012; 7:274–281. [PubMed: 21427193]
- Litt A, Plassmann H, Shiv B, Rangel A. Dissociating valuation and saliency signals during decision-making. *Cerebral Cortex*. 2011; 21(1):95–102.

- Liu X, Hairston J, Schrier M, Fan J. Common and distinct networks underlying reward valence and processing stages: a meta-analysis of functional neuroimaging studies. *Neuroscience and biobehavioral reviews*. 2011; 35(5):1219–36. [PubMed: 21185861]
- Ljungberg T. Responses of monkey dopamine neurons during learning of behavioral reactions. *Journal of Neurophysiology*. 1992; 67(1):145–163. [PubMed: 1552316]
- Logothetis NK. The underpinnings of the BOLD functional magnetic resonance imaging signal. *The Journal of Neuroscience*. 2003; 23(10):3963–3971. [PubMed: 12764080]
- Logothetis NK. What we can do and what we cannot do with fMRI. *Nature*. 2008; 453(7197):869–878. [PubMed: 18548064]
- Logothetis NK. Neurovascular uncoupling: much ado about nothing. *Frontiers in Neuroenergetics*. 2010; 2:1–4. [PubMed: 20162100]
- Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A. Neurophysiological investigation of the basis of the fMRI signal. *Nature*. 2001; 412(6843):150–157. [PubMed: 11449264]
- Logothetis NK, Wandell BA. Interpreting the BOLD signal. *Annual Review of Physiology*. 2004; 66:735–769.
- Lohrenz T, McCabe K, Camerer CF, Montague PR. Neural signature of fictive learning signals in a sequential investment task. *Proceedings of the National Academy of Sciences of the United States of America*. 2007; 104(22):9493–9498. [PubMed: 17519340]
- Matsumoto M, Hikosaka O. Two types of dopamine neuron distinctly convey positive and negative motivational signals. *Nature*. 2009; 459(7248):837–841. [PubMed: 19448610]
- McDannald M. Model-based learning and the contribution of the orbitofrontal cortex to the model-free world. *European Journal of ...* 2012; 35(7):991–996.
- Menon M, Jensen J, Vitcu I, Graff-Guerrero A, Crawley AP, Smith MA, Kapur S. Temporal difference modeling of the blood-oxygen level dependent response during aversive conditioning in humans: effects of dopaminergic modulation. *Biological Psychiatry*. 2007; 62(7):765–772. [PubMed: 17224134]
- Metereau E, Dreher J. Cerebral correlates of salient prediction error for different rewards and punishments. *Cerebral Cortex*. 2013; 23(2):477–487. [PubMed: 22368086]
- Meunier M, Monfardini E, Boussaoud D. Learning by observation in rhesus monkeys. *Neurobiology of Learning and Memory*. 2007; 88(2):243–248. [PubMed: 17572114]
- Mobbs D, Yu R, Meyer M, Passamonti L, Seymour B, Calder AJ, ... Dalgleish T. A key role for similarity in vicarious reward. *Science*. 2009; 324(5929):900. [PubMed: 19443777]
- Moll J, Krueger F, Zahn R. Human frontomesolimbic networks guide decisions about charitable donation. *Proceedings of the National Academy of Sciences of the United States of America*. 2006; 103(42):15623–15628. [PubMed: 17030808]
- Montague P, Dayan P, Sejnowski T. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *The Journal of Neuroscience*. 1996; 16(5):1936–1947. [PubMed: 8774460]
- Mukamel R, Gelbard H, Arieli A, Hasson U, Fried I, Malach R. Coupling between neuronal firing, field potentials, and FMRI in human auditory cortex. *Science*. 2005; 309(5736):951–954. [PubMed: 16081741]
- Münte TF, Heldmann M, Hinrichs H, Marco-Pallares J, Krämer UM, Sturm V, Heinze HJ. Nucleus Accumbens is Involved in Human Action Monitoring: Evidence from Invasive Electrophysiological Recordings. *Frontiers in Human Neuroscience*. 2007; 1(11)
- Murayama K, Matsumoto M, Izuma K, Matsumoto K. Neural basis of the undermining effect of monetary reward on intrinsic motivation. *Proceedings of the National Academy of Sciences of the United States of America*. 2010; 107(49):20911–20916. [PubMed: 21078974]
- Myers WA. Observational learning in monkeys. *Journal of the Experimental Analysis of behavior*. 1970; 14(2):225–235. [PubMed: 16811470]
- Niv Y. Reinforcement learning in the brain. *Journal of Mathematical Psychology*. 2009; 53(3):139–154.
- Niv Y, Chan S. On the value of information and other rewards. *Nature Neuroscience*. 2011; 14(9):1095–1097.

- Niv Y, Daw ND, Joel D, Dayan P. Tonic dopamine: opportunity costs and the control of response vigor. *Psychopharmacology*. 2007; 191(3):507–520. [PubMed: 17031711]
- Niv Y, Edlund JA, Dayan P, O'Doherty JP. Neural prediction errors reveal a risk-sensitive reinforcement-learning process in the human brain. *The Journal of Neuroscience*. 2012; 32(2): 551–562. [PubMed: 22238090]
- Niv, Y.; Montague, PR. Theoretical and Empirical Studies of Learning. In: Glimcher, PW.; Camerer, CF.; Fehr, E.; Poldrack, RA., editors. *Neuroeconomics: Decision making and the brain*. New York, NY: Academic Press; 2008. p. 329-350.
- O'Doherty JP, Buchanan TW, Seymour B, Dolan RJ. Predictive neural coding of reward preference involves dissociable responses in human ventral midbrain and ventral striatum. *Neuron*. 2006; 49(1):157–166. [PubMed: 16387647]
- O'Doherty JP, Deichmann R, Critchley H, Dolan R. Neural responses during anticipation of a primary taste reward. *Neuron*. 2002; 33(5):815–826. [PubMed: 11879657]
- O'Doherty JP, Hampton A, Kim H. Model-based fMRI and its application to reward learning and decision making. *Annals of the New York Academy of Sciences*. 2007; 1104:35–53. [PubMed: 17416921]
- Ogawa S. Intrinsic Signal Changes Accompanying Sensory Stimulation: Functional Brain Mapping with Magnetic Resonance Imaging. *Proceedings of the National Academy of Sciences of the United States of America*. 1992; 89(13):5951–5955. [PubMed: 1631079]
- Ogawa S, Lee TM, Kay AR, Tank DW. Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Sciences of the United States of America*. 1990; 87(24):9868–9872. [PubMed: 2124706]
- Olds J, Milner P. Positive reinforcement produced by electrical stimulation of septal area and other regions of rat brain. *Journal of Comparative and Physiological Psychology*. 1954; 47(6):419–427. [PubMed: 13233369]
- Pagnoni G, Zink CF, Montague PR, Berns GS. Activity in human ventral striatum locked to errors of reward prediction. *Nature Neuroscience*. 2002; 5(2):97–98.
- Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD. Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature*. 2006; 442(7106):1042–1045. [PubMed: 16929307]
- Phan KL, Sripada CS, Angstadt M, McCabe K. Reputation for reciprocity engages the brain reward center. *Proceedings of the National Academy of Sciences of the United States of America*. 2010; 107(29):13099–13104. [PubMed: 20615982]
- Poldrack RA. Can cognitive processes be inferred from neuroimaging data? *Trends in Cognitive Sciences*. 2006; 10(2):59–63. [PubMed: 16406760]
- Raichle ME, Mintun MA. Brain work and brain imaging. *Annual Review of Neuroscience*. 2006; 29:449–476.
- Range F, Viranyi Z, Huber L. Selective imitation in domestic dogs. *Current Biology*. 2007; 17(10): 868–72. [PubMed: 17462893]
- Redgrave P, Gurney K, Reynolds J. What is reinforced by phasic dopamine signals? *Brain Research Reviews*. 2008; 58(2):322–339. [PubMed: 18055018]
- Rendell L, Boyd R, Cownden D, Enquist M, Eriksson K, Feldman MW, ... Laland KN. Why copy others? Insights from the social learning strategies tournament. *Science*. 2010; 328(5975):208–13. [PubMed: 20378813]
- Ribas-Fernandes JJ, Solway A, Diuk C, McGuire JT, Barto AG, Niv Y, Botvinick MM. A Neural Signature of Hierarchical Reinforcement Learning. *Neuron*. 2011; 71(2):370–379. [PubMed: 21791294]
- Rilling JK, Sanfey AG. The neuroscience of social decision-making. *Annual Review of Psychology*. 2011; 62:23–48.
- Robinson OJ, Overstreet C, Charney DR, Vytal K, Grillon C. Stress increases aversive prediction error signal in the ventral striatum. *Proceedings of the National Academy of Sciences of the United States of America*. 2013; 110(10):4129–4133. [PubMed: 23401511]

- Salimpoor VN, Benovoy M, Larcher K, Dagher A, Zatorre RJ. Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nature Neuroscience*. 2011; 14(2): 257–262.
- Schönberg T, Daw ND, Joel D, O’Doherty JP. Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. *The Journal of Neuroscience*. 2007; 27(47):12860–12867. [PubMed: 18032658]
- Schönberg T, O’Doherty JP, Joel D, Inzelberg R, Segev Y, Daw ND. Selective impairment of prediction error signaling in human dorsolateral but not ventral striatum in Parkinson’s disease patients: evidence from a model-based fMRI study. *NeuroImage*. 2010; 49(1):772–781. [PubMed: 19682583]
- Schott BH, Minuzzi L, Krebs RM, Elmenhorst D, Lang M, Winz OH, ... Bauer A. Mesolimbic functional magnetic resonance imaging activations during reward anticipation correlate with reward-related ventral striatal dopamine release. *The Journal of Neuroscience*. 2008; 28(52): 14311–14319. [PubMed: 19109512]
- Schott BH, Niehaus L, Wittmann BC, Schütze H, Seidenbecher CI, Heinze HJ, Düzel E. Ageing and early-stage Parkinson’s disease affect separable neural mechanisms of mesolimbic reward processing. *Brain*. 2007; 130(9):2412–2424. [PubMed: 17626038]
- Schultz W. Getting Formal with Dopamine and Reward. *Neuron*. 2002; 36(2):241–263. [PubMed: 12383780]
- Schultz W. Behavioral theories and the neurophysiology of reward. *Annual Review of Psychology*. 2006; 57:87–115.
- Schultz W. Behavioral dopamine signals. *Trends in Neurosciences*. 2007; 30(5):203–210. [PubMed: 17400301]
- Schultz, W. Midbrain Dopamine Neurons: a Retina of the Reward System?. In: Glimcher, PW.; Camerer, CF.; Fehr, E.; Poldrack, RA., editors. *Neuroeconomics: Decision making and the brain* (pp. New York, NY: Academic Press; 2008. p. 319-328.
- Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science*. 1997; 275(5306):1593–1599. [PubMed: 9054347]
- Sclafani, A. Macronutrient-conditioned flavor preferences. In: Berthoud, H-R.; Seeley, R., editors. *Neural and metabolic control of macronutrient intake*. Boca Raton, FL: CRC Press; 1999. p. 93-107.
- Sescousse G, Caldú X, Segura B, Dreher J-C. Processing of primary and secondary rewards: a quantitative meta-analysis and review of human functional neuroimaging studies. *Neuroscience and biobehavioral reviews*. 2013; 37(4):681–96. [PubMed: 23415703]
- Seymour B, O’Doherty JP, Dayan P, Koltzenburg M, Jones AK, Dolan RJ, ... Frackowiak RS. Temporal difference models describe higher-order learning in humans. *Nature*. 2004; 429(10): 664–667. [PubMed: 15190354]
- Shohamy D. Learning and motivation in the human striatum. *Current Opinion in Neurobiology*. 2011; 21(3):408–414. [PubMed: 21658933]
- Shohamy D, Myers CE, Grossman S, Sage J, Gluck MA, Poldrack RA. Cortico-striatal contributions to feedback-based learning: converging data from neuroimaging and neuropsychology. *Brain*. 2004; 127(4):851–859. [PubMed: 15013954]
- Shohamy D, Myers CE, Kalanithi J, Gluck MA. Basal ganglia and dopamine contributions to probabilistic category learning. *Neuroscience & Biobehavioral Reviews*. 2008; 32(2):219–236. [PubMed: 18061261]
- Skinner, BF. New York, NY, USA: Simon and Schuster; 1953. *Science and human behavior*.
- Steinberg EE, Keiflin R, Boivin JR, Witten IB, Deisseroth K, Janak PH. A causal link between prediction errors, dopamine neurons and learning. *Nature Neuroscience*. 2013; 16(7):966–973.
- Subiaul F, Cantlon JF, Holloway RL, Terrace HS. Cognitive imitation in rhesus macaques. *Science*. 2004; 305(5682):407–410. [PubMed: 15256673]
- Sutton, RS.; Barto, AG. Time-derivative models of Pavlovian reinforcement. In: Gabriel, M.; Moore, J., editors. *Learning and computational neuroscience: Foundations of adaptive networks*. Cambridge, MA: MIT Press; 1990. p. 497-537.

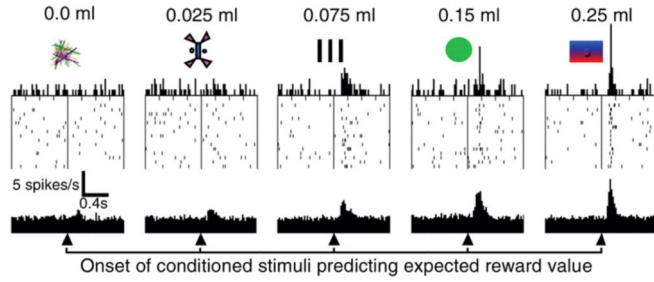
- Suzuki S, Harasawa N, Ueno K, Gardner JL, Ichinohe N, Haruno M, ... Nakahara H. Learning to simulate others' decisions. *Neuron*. 2012; 74(6):1125–1137. [PubMed: 22726841]
- Takahashi YK, Roesch MR, Wilson RC, Toreson K, O'Donnell P, Niv Y, Schoenbaum G. Expectancy-related changes in firing of dopamine neurons depend on orbitofrontal cortex. *Nature Neuroscience*. 2011; 14(12):1590–1597.
- Tobler PN, Dickinson A, Schultz W. Coding of predicted reward omission by dopamine neurons in a conditioned inhibition paradigm. *The Journal of Neuroscience*. 2003; 23(32):10402–10410. [PubMed: 14614099]
- Tobler PN, Fiorillo CD, Schultz W. Adaptive coding of reward value by dopamine neurons. *Science*. 2005; 307(5715):1642–1645. [PubMed: 15761155]
- Tricomi E, Fiez JA. Feedback signals in the caudate reflect goal achievement on a declarative memory task. *NeuroImage*. 2008; 41(3):1154–1167. [PubMed: 18445531]
- Tricomi E, Rangel A, Camerer C, O'Doherty J. Neural evidence for inequality-averse social preferences. *Nature*. 2010; 463(7284):1089–1091. [PubMed: 20182511]
- Valentin VV, O'Doherty JP. Overlapping prediction errors in dorsal striatum during instrumental learning with juice and money reward in the human brain. *Journal of Neurophysiology*. 2009; 102(6):3384–3391. [PubMed: 19793875]
- van der Schaaf ME, van Schouwenburg MR, Geurts DEM, Schellekens AFA, Buitelaar JK, Verkes RJ, Cools R. Establishing the Dopamine Dependency of Human Striatal Signals During Reward and Punishment Reversal Learning. *Cerebral Cortex*. 2012
- Vitay J, Fix J, Beuth F, Schroll H, Hamker FH. Biological models of reinforcement learning. *Künstliche Intelligenz*. 2009; 3:12–18.
- Waelti P, Dickinson A, Schultz W. Dopamine responses comply with basic assumptions of formal learning theory. *Nature*. 2001; 412(6842)
- Wimmer GE, Shohamy D. Preference by association: how memory mechanisms in the hippocampus bias decisions. *Science*. 2012; 338(6104):270–273. [PubMed: 23066083]
- Wittmann BC, Bunzeck N, Dolan RJ, Düzel E. Anticipation of novelty recruits reward system and hippocampus while promoting recollection. *NeuroImage*. 2007; 38(1):194–202. [PubMed: 17764976]
- Wittmann BC, Schott BH, Guderian S, Frey JU, Heinze HJ, Düzel E. Reward-related FMRI activation of dopaminergic midbrain is associated with enhanced hippocampus-dependent long-term memory formation. *Neuron*. 2005; 45(3):459–467. [PubMed: 15694331]
- Zaehle T, Bauch EM, Hinrichs H, Schmitt FC, Voges J, Heinze H-J, Bunzeck N. Nucleus accumbens activity dissociates different forms of saliency: evidence from human intracranial recordings. *The Journal of Neuroscience*. 2013; 33(20):8764–8771. [PubMed: 23678119]
- Zink CF, Pagnoni G, Chappelow J, Martin-Skurski M, Berns GS. Human striatal activation reflects degree of stimulus saliency. *NeuroImage*. 2006; 29(3):977–983. [PubMed: 16153860]

Highlights

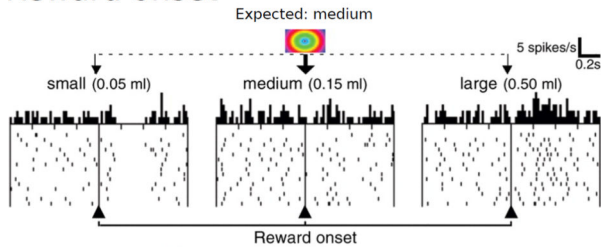
- The ventral striatum processes reward expectation and reward prediction errors
- This pattern has been shown in many species and can be measured using fMRI
- In humans, simple reinforcers and abstract concepts elicit similar activations
- The ventral striatum is also activated when learning from observing others
- It therefore plays a universal role in processing reinforcement

A. Spiking

Onset of reward predicting stimulus

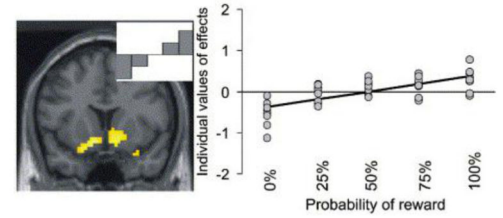


Reward onset



B. BOLD signal

Onset of reward predicting stimulus



Reward onset

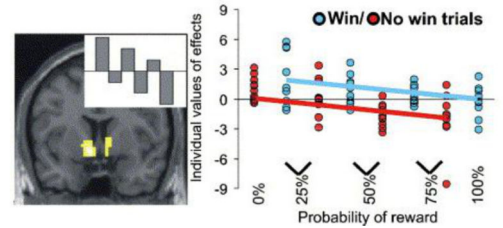


Figure 1. Brain responses to reward expectation and receipt

A. In Macaque monkeys, spiking activity in dopaminergic midbrain neurons increases with the expected value (probability \times magnitude) of the reward that is predicted by a conditioned stimulus, while the response during the receipt of the reward scales with the prediction error, i.e. the received minus the expected reward (Panel adapted from Tobler et al. (2005)). **B.** A similar activation pattern is observed when measuring fMRI activation in the ventral striatum (results only shown for the right VS; Panel adapted from Abler et al. (2006)). These responses are in line with the predictions of normative models of reinforcement learning (Montague et al., 1996; Schultz et al., 1997).

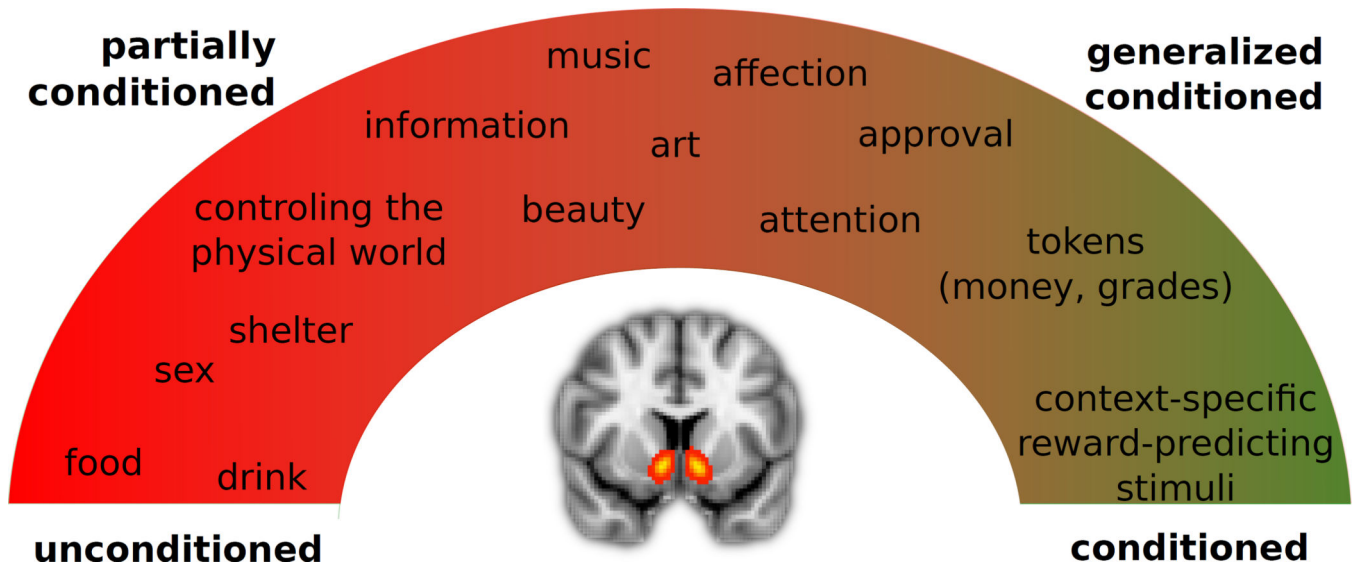


Figure 2. Ventral striatal processing of a wide range of rewards

The ventral striatum has been shown to respond to unexpected rewards from the whole spectrum of reinforcers, including largely unconditioned reinforcers such as food or drink, generalized conditioned reinforcers such as money, and stimuli that are associated with reward only in the context of the specific experiment.

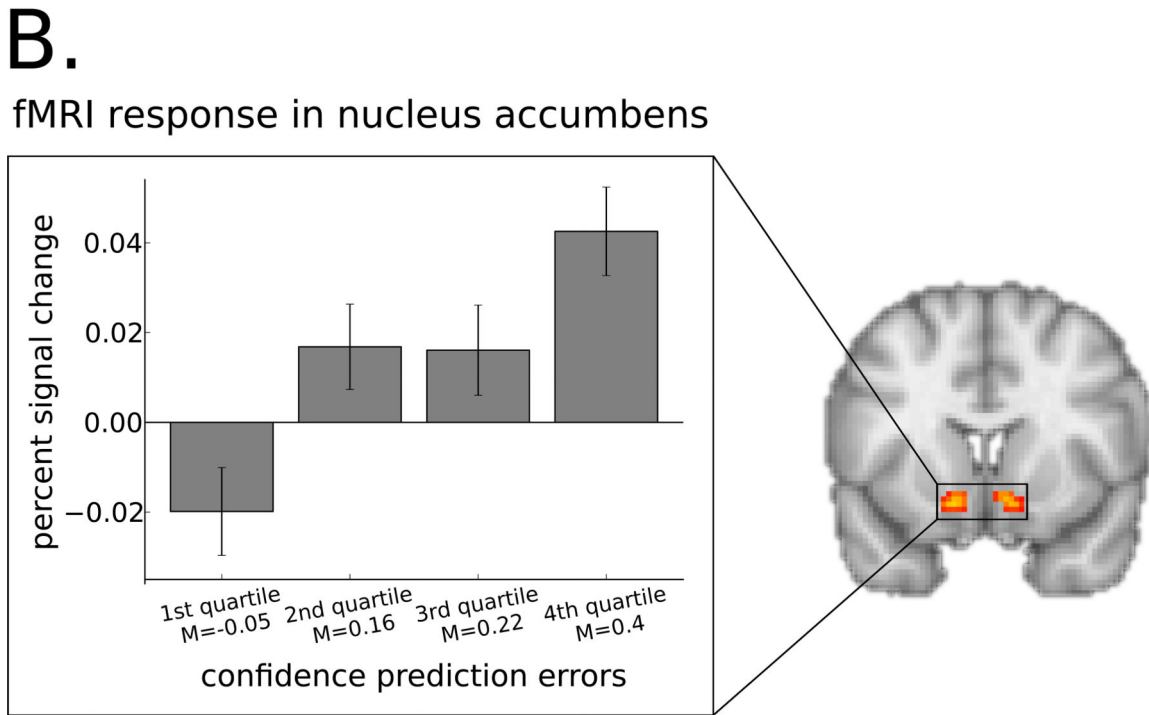
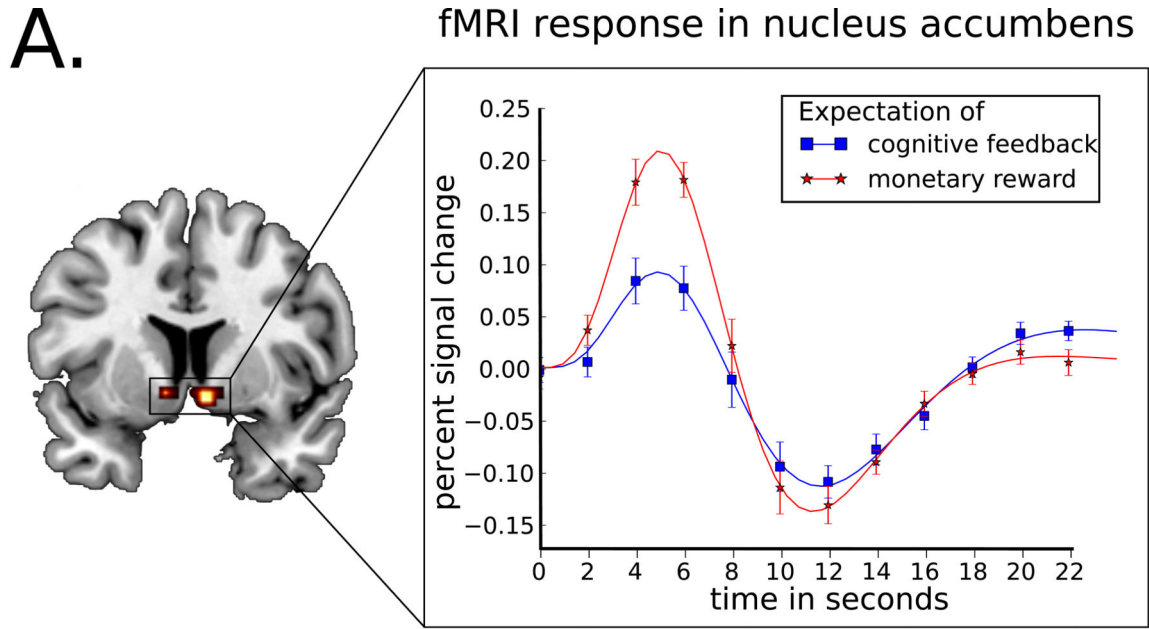


Figure 3. fMRI activation in VS in response to different rewards

A. With similar reward probabilities, the expectation of monetary reward (€0.2) leads to significantly higher activation in the VS than the expectation of cognitive feedback (a green circle signaling the correctness of the response) (Panel adapted from Daniel and Pollmann (2010)). **B.** After observational training in the complete absence of feedback, the VS responds to a prediction error on confidence, i.e. activation is higher when the participant is

more confident than expected that his or her response is correct (Panel adapted from Daniel and Pollmann (2012)).