Dopamine and reward: a view from the prefrontal cortex

Bolton K.H. Chau[a,b], Huw Jarvis[c], Chun-Kit Law[a] and Trevor T.-J. Chong[c]

The prefrontal cortex (PFC) is a heterogeneous area that is critical to reward-based decision-making. In particular, the dorsal anterior cingulate cortex, ventromedial PFC and orbitofrontal cortex are frequently implicated in different aspects of choice behaviour. These regions receive projections from midbrain dopamine (DA) neurons and, in turn, project to other key dopaminergic regions such as the striatum. However, our current understanding of the role of DA in reward-based processes is based mainly on studies of midbrain dopaminergic neurons and striatal DA release from nonhuman animal models. An important gap in the literature surrounds the precise functions of DA release in the PFC, particularly in humans. A priority for future research will be to integrate, both computationally and biologically, the seemingly disparate value representations across different nodes within the reward-processing network. Such models should aim to define the functional interactions between the PFC and basal ganglia, through which dopaminergic neurotransmission guides reward-based behaviour. ***Behavioural Pharmacology*** **00:000–000**

Behavioural Pharmacology 2018, **00**:000–000

Keywords: anterior cingulate cortex, corticostriatal pathway, decision-making, dopamine, mesocortical pathway, orbitofrontal cortex, reward, ventromedial prefrontal cortex

[a]Department of Rehabilitation Sciences, [b]University Research Facility in behavioral and Systems Neuroscience, The Hong Kong Polytechnic University, Kowloon, Hong Kong and [c]Monash Institute of Cognitive and Clinical Neurosciences, Monash University, Clayton, Victoria, Australia

Correspondence to Bolton K.H. Chau, DPhil, Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Kowloon, Hong Kong
E-mail: boltonchau@gmail.com

Received 29 June 2018 **Accepted as revised** 13 August 2018

Introduction

Decisions are often made between options whose outcomes are represented in different, and sometimes very abstract, attributes (e.g. buying a car vs. going on holiday; choosing a relationship vs. a career). Traditional economic theories argued that such decisions are made by computing an abstract utility that allows qualitatively dissimilar options to be quantitatively comparable. Neuroeconomic studies inspired by this approach have found that rewards are represented in a distributed network of areas across the prefrontal cortex (PFC), striatum and midbrain (O'Doherty, 2004; Izuma *et al.*, 2008; Lau and Glimcher, 2008; Zink *et al.*, 2008; Peters and Buchel, 2010; Levy and Glimcher, 2012).

The PFC is a heterogeneous area that plays a broad role in multiple stages of value-based decision-making, from representing the subjective value of a reward, comparing the value difference between available rewards, motivating the decision-making process itself, to guiding flexible choices (Murray and Rudebeck, 2018). These 'reward sensitive' processes are instantiated in three key subdivisions of the PFC, the dorsal anterior cingulate cortex (dACC), ventromedial prefrontal cortex (vmPFC) and orbitofrontal cortex (OFC) (Padoa-Schioppa and Assad, 2008; Rushworth and Behrens, 2008; Grabenhorst and Rolls, 2011).

Dopamine (DA) itself has been widely implicated in reward processing (Schultz *et al.*, 2015; Hamid *et al.*, 2016; Volkow *et al.*, 2017). These prefrontal areas receive extensive projections from midbrain DA neurons by the mesocortical pathway, and they in turn project in a highly organized manner to the striatum. Together, this network of prefrontal and subcortical areas comprises the core of the brain's reward network. However, many studies on the role of DA in reward processing have focused on DA neurotransmission within the midbrain and striatum, and it remains largely unclear how DA regulates the interaction between prefrontal and midbrain/striatal activity.

In this review, we first consider the anatomy and function of the three prefrontal areas that are directly involved in reward-based decisions – the dACC, vmPFC and OFC – before discussing the key role of DA in encoding reward prediction errors. We then consider how the PFC may interact with dopaminergic pathways to facilitate reward-based decisions. Finally, we conclude by highlighting useful approaches to studying prefrontal DA in humans that are based on combining currently available methodological techniques.

The anatomy and function of reward-sensitive prefrontal cortex regions

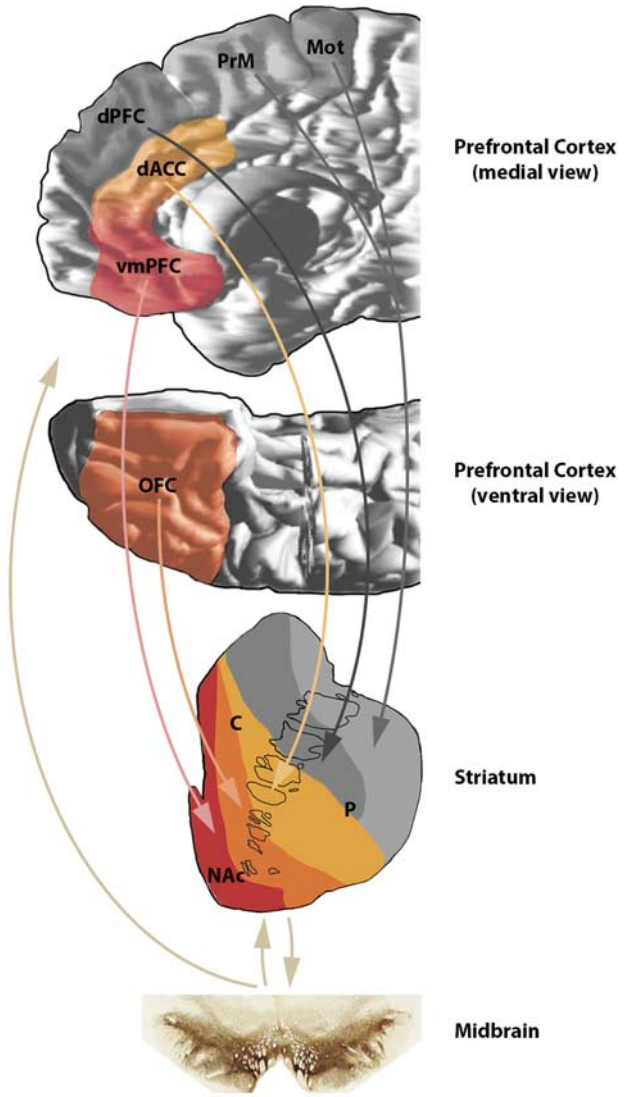
First, we survey the roles of three key PFC regions in reward-based decision-making – the dACC, vmPFC and OFC. We consider each of these regions in turn, in a dorsal-to-ventral order, reflecting their topographic striatal projections (Fig. 1).

Dorsal anterior cingulate cortex

The anterior cingulate cortex (ACC) lies on the medial surface of the frontal lobe, and consists of Brodmann

DOI: 10.1097/FBP.0000000000000424

Fig. 1



Reward-sensitive dopamine pathways. Midbrain dopaminergic neurons project directly to the striatum and prefrontal cortex. The dACC, vmPFC and OFC are the three key prefrontal areas that are directly involved in reward-based decision-making, specifically through their roles in attributing value to stimuli, associating that value with choices and adjudicating between different options. The dPFC has an important role in cognitive control (not discussed in details in this paper). These prefrontal areas in turn connect to the striatum in a highly topographically organized manner. Together, this network of corticostriatal loops comprise the core of a circuit that is central to reward-based decision-making. C, caudate; dACC, dorsal anterior cingulate cortex; dPFC, dorsal prefrontal cortex; Mot, motor cortex; NAc, nucleus accumbens; OFC, orbitofrontal cortex; P, putamen; PrM, premotore cortex; S, shell of nucleus accumbens; SN/VTA, substantia nigra/ventral tegmental area; vmPFC, ventromedial prefrontal cortex. Adapted from Haber and Knutson (2010). Adaptations are themselves works protected by copyright. So in order to publish this adaptation, authorization must be obtained both from the owner of the copyright in the original work and from the owner of copyright in the translation or adaptation.

areas 24, 25 and 32, which lie in and around the cingulate sulcus. The dACC in turn encompasses regions referred to as the anterior midcingulate cortex and rostral cingulate zone (Cole *et al.*, 2009; Shackman *et al.*, 2011; Procyk *et al.*, 2014; Heilbronner and Hayden, 2016; Vogt, 2016). Notably, it is distinct from adjacent areas such as the presupplementary motor area (pre-SMA), and is a key hub in a network of regions implicated in domain-general executive function. Some authors have suggested that the human dACC is unique, but others have argued that the dACC and its connections are relatively preserved across humans and macaques (Cole *et al.*, 2009). Similarly, cross-species comparisons between primates and rodents suggest that primate area 24 may be homologous to rodent area Cg or area 24 (Passingham and Wise, 2012; Heilbronner *et al.*, 2016). As in the primate, the rodent ACC is strongly connected with the core of the nucleus accumbens (NAc) and the basolateral amygdala. This further supports the view that ACC is preserved across rodent and primate species.

The connectivity of the dACC (and in particular area 24) positions it optimally to facilitate value-based decisions. It is tightly linked to nearby areas of the frontal cortex, such as the dorsolateral PFC, and adjacent ACC areas, such as the perigenual ACC. The dACC itself is directly connected to much of the striatum, as well as other subcortical regions such as the amygdala that encode reward and value (Haber, 2011). Through this connectivity, the dACC may therefore influence, and be influenced by, dopaminergic activity, and its direct connections to motor areas (e.g. the pre-SMA) allows it to exert direct influence over motor output (Luppino *et al.*, 1991; He *et al.*, 1995). In sum, the dACC sits at an important interface between the brain's reward valuation networks and their translation to action.

The dACC plays a central role in encoding choice value. Neuronal activity in the macaque dACC reflects reward history (Kolling *et al.*, 2016), as does functional MRI (fMRI) blood oxygen level-dependent (BOLD) activity from the human dACC, which can be used to predict future rewards and guide decisions to maintain or change behaviour (Wittmann *et al.*, 2016). Consistent with these findings are studies that have shown that dACC lesions impair the use of reward-history-dependent values to determine the balance between persistence and change (Kennerley *et al.*, 2006). Together, the value signals in the dACC may therefore reflect the recency-weighted history of previously chosen rewards.

However, the dACC has also been implicated in a multitude of cognitive processes, and its precise role remains highly controversial (Cole *et al.*, 2009; Shackman *et al.*, 2011; Kolling *et al.*, 2012; Procyk *et al.*, 2014; Shenhav *et al.*, 2014; Heilbronner and Hayden, 2016; Vogt, 2016). It has been implicated in motivation, error monitoring (Posner and Petersen, 1990; Holroyd and Coles, 2002b; Debener *et al.*, 2005), conflict detection (Carter *et al.*, 1998; Botvinick, 2007), and detecting the volatility of the reward environment (Behrens *et al.*, 2007). Across all of

these roles, two broad overarching functions for the dACC are thought to be the valuation of effort-related costs, and adaptive decision-making.

Motivating effortful actions

Motivation involves a cost-benefit analysis, in which the costs of an action are weighed against its potential rewards (Chong *et al.*, 2016). The dACC, together with the OFC and striatum, are key structures in the valuation of effort costs. Lesions encompassing the dACC disrupt the willingness of rats to invest effort in pursuit of rewards (Walton *et al.*, 2002, 2003, 2009; Schweimer and Hauber, 2005; Schweimer *et al.*, 2005; Rudebeck *et al.*, 2006). Importantly, this lowered motivation is not due to a motor deficit or altered reward sensitivity (Walton *et al.*, 2002, 2003; Rudebeck *et al.*, 2006). Rather, it is due to an impairment in the ability to integrate effort and reward information, suggesting a particularly important role for the dACC in effort-based decision-making, in both the physical (Shidara and Richmond, 2002; Amemori and Graybiel, 2012) and cognitive domains (Hosking *et al.*, 2014).

Similarly, human studies have shown that the dACC encodes the subjective value of effortful actions (Croxson *et al.*, 2009; Chong *et al.*, 2017). Recent work has shown that the subjective value of rewards discounted by effort is encoded in the dACC, regardless of the specific domain of effort involved (i.e. for both cognitive and physical effort) (Chong *et al.*, 2017). The causative role of the dACC in energization and motivated behaviour is evidenced by lesion studies, which have shown that dACC lesions have been associated with general slowing of response time (Stuss *et al.*, 2005), and a higher threshold for overcoming effortful obstacles (Holroyd and Yeung, 2012). Lesions to areas encompassing the human dACC result in clinically severe impairments in motivation, such as akinetic mutism. Conversely, dACC stimulation produces experiences of a 'willingness to persevere' through impending challenges (Parvizi *et al.*, 2013).

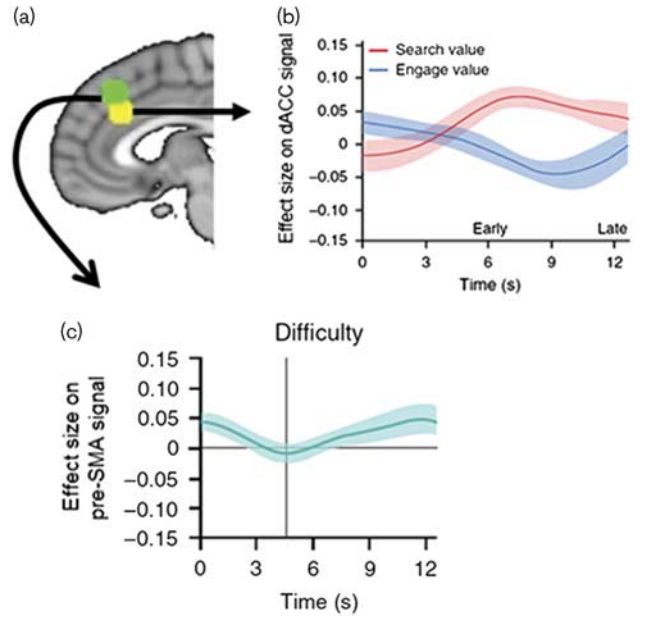
Adaptive decision-making

Another influential set of theories has linked the dACC to 'conflict monitoring' – the process of monitoring action outcomes, and detecting when two competing choices might be made during a difficult task (Botvinick *et al.*, 2004; Botvinick, 2007). By these accounts, the dACC underlies our ability to flexibly adjust behaviour to accord with internally maintained goals, and away from behaviours that may distract from those goals, especially in response to unexpected events (Holroyd and Coles, 2002a). A possible mechanism for this conflict-monitoring process is the encoding of prediction errors within the dACC. Although prediction errors are often discussed in the context of striatal DA signalling (see below), several studies have shown that prediction error signals are also encoded at the cellular level within single dACC neurons (Matsumoto *et al.*, 2007; Bryden *et al.*, 2011; Hayden *et al.*, 2011). However, the types of prediction error that are signalled by dopaminergic and dACC neurons are fundamentally different. Dopaminergic neurons characteristically signal a signed difference between the predicted and actual outcomes (Schultz *et al.*, 1997). In contrast, dACC neurons rarely generate signed prediction errors (although see Kennerley *et al.*, 2011), but instead generate representations of expected outcomes on the basis of accumulation of previous outcomes (Hyman *et al.*, 2017). This comparison process that takes into account previous trial history may then be used to detect violations of expected outcomes.

In humans, a topical alternative approach to determining the role of the dACC in adaptive decision-making has been to examine foraging behaviour with fMRI. A recent study examined how humans decide whether to explore a set of alternative choices, or stick with the opportunity to make a 'default' choice (Fig. 2b) (Kolling *et al.*, 2012). This study required individuals to weigh the value of the encountered option (the default 'encounter value'), against the richness of the environment ('search value'), and the effort cost of searching elsewhere ('search cost'). The value of exploring was encoded by a positive 'search value' signal in dACC, which indexed the average value of the set of alternative actions. Conversely, dACC activity was negatively influenced by both the encounter value and search cost. However, dACC activity was not modulated by the choice participants subsequently made. This pattern of positive and negative modulations may represent an inverse value difference signal, as activity increases when the difference between the value of the chosen option and the value of the option that is foregone decreases (Hare *et al.*, 2011). Overall, this pattern of activity is suggestive of a comparison process in the dACC that could inform decisions about whether to continue exploiting the current reward patch, or to explore the environment for superior alternatives (Kolling *et al.*, 2012).

However, decisions close to the subjective indifference point between searching and engaging also tend to be more difficult. Thus, an alternative interpretation suggests that the dACC does not necessarily encode search value, but the difficulty of a decision in general (Shenhav *et al.*, 2014). In the context of the foraging experiment above, difficulty can be operationalized as the absolute difference between the search and engage values, as opposed to the relative exploration value that is the signed difference between the two values. On the basis of connectivity patterns (Beckmann *et al.*, 2009; Neubert *et al.*, 2015), the subregions within the dACC that encode 'exploration' and 'difficulty' appear to be anatomically segregated (Fig. 2a). Specifically, it is possible to concurrently observe an exploration signal in a relatively ventral dACC region, and a difficulty signal in a relatively

Fig. 2



Multiple decision signals are found in dACC. (a) A more ventral dACC region (yellow) and a more dorsal pre-SMA region showed different signals associated with the decision. (b) The dACC activity was modulated as a function of relative search value – opposite value signals for 'engaging' versus 'searching' were observed. (c) The pre-SMA encoded the difficulty of the trial. dACC, dorsal anterior cingulate cortex; pre-SMA, presupplementary motor area. Adapted from Kolling *et al.* (2016). Adaptations are themselves works protected by copyright. So in order to publish this adaptation, authorization must be obtained both from the owner of the copyright in the original work and from the owner of copyright in the translation or adaptation.

dorsal dACC region (sometimes also known as pre-SMA; Fig. 2c) (Kolling *et al.*, 2016). These data suggest that different subregions of the dACC may play separate roles in adaptive decision-making, although the broader functional specializations of different dACC subregions remain to be clarified.

Ventromedial prefrontal cortex

The vmPFC is a poorly defined anatomical region in the PFC, with its precise location and boundaries varying widely across different studies. For example, the part of the medial PFC adjacent to the genu of the corpus callosum has been variously labelled the 'vmPFC' or 'ACC'. The nominal 'vmPFC' is large, with cytoarchitectonic studies parcellating the 'ventromedial' part of the human PFC into areas 10m, 10r, 11m, 14c and 14r (Carmichael and Price, 1994; Ongur and Price, 2000; Price, 2007). Despite this heterogeneity, research in the last two decades has provided strong evidence that parts of the ventromedial PFC are important to reward-based decisions, by representing subjective reward value, as well as by implementing value-based comparisons between available options.

Ventromedial prefrontal cortex encodes reward value

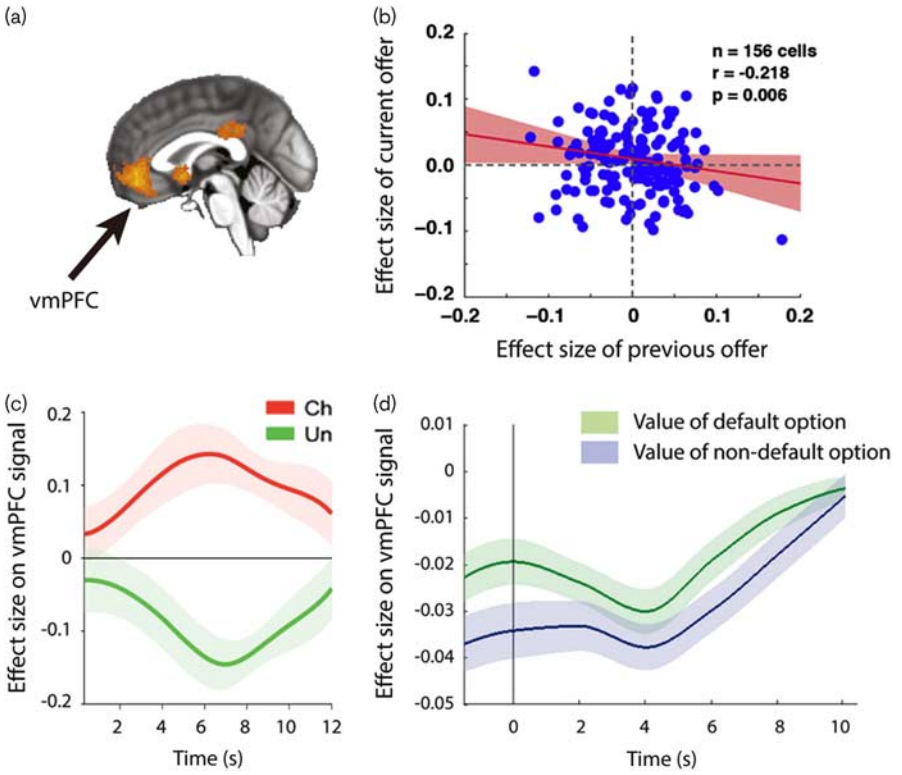
A large volume of data has shown that the vmPFC encodes the value of a presented reward. Importantly, however, the activity of this region does not merely correlate with the objective value of a reward, but in fact is better explained by how subjectively rewarding that option is to the individual (Kable and Glimcher, 2007; Lebreton *et al.*, 2009). Neuroeconomic theories posit a central role for subjective value in guiding individuals' decisions. An important characteristic of neural signal that reflects value is that it should be greater when an option is more rewarding, as well as when an option is less aversive (i.e. the relationship between the signal and value should be linear throughout the positive and negative sides of the valence spectrum). A recent meta-analysis on 206 fMRI studies on subjective value found just such a value signal in a cluster of vmPFC regions that peaked at area 10r (standard Montreal Neurological Institute coordinates of 2, 46, −8; Fig. 3a) (Bartra *et al.*, 2013). The subjective value signal in the vmPFC is therefore thought to provide an important biophysical substrate for value-based decisions.

Human lesion studies support the causal role of the vmPFC in decision-making, and show that focal vmPFC lesions result in specific decision-making impairments. For example, Damasio (1996) showed that, in a gambling-like task, patients with vmPFC lesions prefer riskier choices (Bechara *et al.*, 2005). However, although such patients are more stochastic in reward-based decisions, the speed of their decisions is not necessarily impaired, and their performances in perceptual-based decision-making tasks are comparable to controls (Fellows and Farah, 2005, 2007; Henri-Bhargava *et al.*, 2012; Noonan *et al.*, 2017). Thus, the vmPFC should not be considered a 'primary decision cortex' for general value computations and decision-making; rather, it is involved specifically in decisions driven by subjective preferences. To understand the exact role of vmPFC in decision making, it is important to consider the nature of the signal in this region.

Ventromedial prefrontal cortex encodes a value difference signal

A key property of any area that is purported to be involved in the process of reward-based decision-making is its capacity to represent the relative values of available options, in order to be able to compare the difference between them. In the vmPFC, a 'value difference' signal has been broadly reported in human fMRI studies. When a person is choosing between two options, vmPFC activity is both positively correlated with the value of one option, and negatively correlated with the value of the other, such that the difference in value between the two options is compared. Similar findings have been observed during neurophysiological recordings from vmPFC neurons, while macaques were making decisions between

Fig. 3



Value signals in the ventromedial prefrontal cortex (vmPFC). (a) A meta-analysis showed that the vmPFC signal is modulated linearly as a function of the option value – it becomes more active as the value increases from negative to positive (adapted from Bartra *et al.*, 2013). (b) Neurophysiology data showed that the firing of vmPFC neurons was modulated by the value of two options in an opposite manner, suggesting that vmPFC neurons compared the value between the two options. There are multiple hypotheses on the framework of the value comparison in vmPFC (adapted from Strait *et al.*, 2014). (c) One framework suggests that vmPFC compares the value between the chosen and unchosen option (adapted from Papageorgiou *et al.*, 2017). (d) Another framework suggests that the vmPFC compares the value between a default and a nondefault option (adapted from Lopez-Persem *et al.*, 2016). Adaptations are themselves works protected by copyright. So in order to publish this adaptation, authorization must be obtained both from the owner of the copyright in the original work and from the owner of copyright in the translation or adaptation.

two sequentially-presented options (Strait *et al.*, 2014). When the second option was presented, the activity of vmPFC neurons was modulated by the value of that option, and in the opposite direction by the value of the option presented earlier. In other words, the vmPFC neurons encoded a signal that was related to the value difference between the current offer and the alternative option (Fig. 3b). A value difference signal is an important neural signature of decision making, and understanding the nature of this signal is important to revealing the specific role of vmPFC in value-based decisions.

There are multiple frameworks through which the values of two options can be compared to reach a decision. For example, a neural network can use a space-based framework to compare the value difference between two options located in physically different locations (e.g. left vs. right). In more posterior regions such as the lateral intraparietal area, each neuron has a receptive field that corresponds to a small proportion of the visual field. Their activity is modulated positively as a function of the value of the option presented spatially within their response field, and negatively as a function of other options outside their receptive field (Platt and Glimcher, 1999; Churchland *et al.*, 2008). This neuronal signal is particularly useful to evaluate the value of an option at a given location, relative to options elsewhere. However, unlike posterior visual regions, vmPFC neurons lack the spatial tuning required for a space-based framework.

An alternative to the space-based approach suggests that the vmPFC uses an attention-based framework, which compares attended versus unattended options. Lim *et al.* (2011) recorded eye movements when human participants were choosing between two options. When they attended to an option by gazing at it, the vmPFC signal was positively related to the value of the attended option, and negatively related to the value of an unattended option, which suggested that the vmPFC encoded a value difference between both alternatives. Importantly, however, the attentional modulation of the vmPFC signal was independent of the option that was eventually chosen (Lim *et al.*, 2011). Collectively, these data suggest that, even though the vmPFC signals the difference in

value between options, it is not involved in choice selection per se. The causal role of the vmPFC in guiding attention during reward-based decisions is further supported by patients with vmPFC lesions, who show less attention to information relevant to the decision itself (Vaidya and Fellows, 2015, 2016).

In contrast to the spatial/attentional frameworks, vmPFC signals have also been proposed to encode the value difference between an option that is about to be chosen and an alternative that is about to be foregone. Several human fMRI studies have shown that the vmPFC encodes a value difference signal between the chosen and unchosen options (Fig. 3c) (Boorman *et al.*, 2009; Kolling *et al.*, 2012; Jocham *et al.*, 2014; Papageorgiou *et al.*, 2017). This framework is appealing because it suggests that the vmPFC is not only critical to value comparison, but is also involved in the choice selection process by encoding the value of the impending choice. Note that this contrasts with the attentional framework, in which the vmPFC is not critical to the selection of an option. Neurophysiological data support this idea by showing that the firing rate of a large proportion of neurons is modulated by the value of the chosen option before the decision is made (Strait *et al.*, 2014). However, critics argue that the signal difference between the chosen and unchosen options is postdecisional, and is not critical to the choice selection process.

Finally, a more recent proposal has been that the vmPFC encodes value in a preference-based framework. Such theories propose that the vmPFC compares options in a preferred category with an alternative in a nonpreferred category. For example, one might in general prefer chocolate to cookies, but the exact decision would depend on the actual choices offered (e.g. one might dislike particular types of chocolate). Lopez-Persem *et al.* (2016) asked human participants to choose between a snack item from a preferred category and another snack item from a nonpreferred category (Fig. 3d). The vmPFC signal was modulated positively as a function of the snack of the preferred category, and negatively as a function of the snack of the nonpreferred category, regardless of which option was then chosen. They also ran a computational model to explain participants' choices, which suggested that both category preference and visual attention are important factors that explain choice. Further investigations could test whether the vmPFC simultaneously encodes both preferred versus nonpreferred value difference, and attended versus unattended value difference.

Value difference signals in ventromedial prefrontal cortex of humans versus those in ventromedial prefrontal cortex of monkeys

Although cytoarchitectonic and connectivity studies have shown the homologous relationship between the vmPFC of human and nonhuman primates, a direct comparison using the same measurement and decision-making task provides the best test to assess whether the vmPFC is functionally comparable across primate species. A recent study applied fMRI in one human experiment and two monkey experiments that involved binary choice decision-making tasks (Papageorgiou *et al.*, 2017). In humans, a classical value difference signal was reported at vmPFC area 10r – activity in this region was correlated with the value difference between the two options. This accords with the results from two monkey experiments, which also showed a value difference signal in area 10 m, which is considered structurally homologous to the human area 10r (Price, 2007; Neubert *et al.*, 2015). Interestingly, however, the sign of the value difference signals differed across species, such that it was positive in humans (consistent with previous studies), but negative in both macaque experiments.

The reason for the reversed value difference signal across species is unclear, and is a further illustration of the complexities of generalizing findings across studies involving human and nonhuman animals. Such discrepancies are unlikely to have been simply due to experimental factors. All experiments were conducted using a similar MRI scanner, and, although there were some task differences between the human and macaque experiments (humans were explicitly presented the reward probabilities of each option, but monkeys had to learn these probabilities trial-by-trial), these alone should not have reversed the sign of the value difference signal. One possible explanation for this discrepancy is that it could result from even minor differences between the neural networks across the two species. For example, a single inhibitory connection would be sufficient to reverse the positivity or negativity of a signal, and it may be that the direction of a signal may be of less functional consequence than its magnitude. Nevertheless, it remains for future studies to clarify whether this discrepancy in the sign of the value signal reflects divergent evolutionary decision processes across primate species.

Value signals and cognitive maps

Apart from computing value difference, recent evidence suggests that the vmPFC also encodes a 'cognitive map', which provides insights into how value signals emerge in this region. In spatial perception, physical space can be represented by a two-dimensional Cartesian map, and grid cells in the entorhinal cortex use a hexagonally symmetric code to represent this two-dimensional space (Hafting *et al.*, 2005). Similar to physical space, concepts can also be represented by continuous dimensions. For example, the identity of bird species can be represented by continuous dimensions of leg length and neck length, and different bird species can be located at different positions of the two-dimensional leg-and-neck space. Constantinescu *et al.* (2016) taught participants to recognize birds using this two-dimensional 'bird space'. Similar

to the representation of physical space, both the entorhinal cortex and the vmPFC used a hexagonally symmetric code to represent 'bird space'. In reward-based decision-making, integrating decision attributes (e.g. reward magnitude and probability) is an important computation for representations of value. Such a two-dimensional cognitive map in the vmPFC could be useful in value-based computations during choice behaviour.

Orbitofrontal cortex

The human OFC lies on the ventral surface of the PFC adjacent to the orbits. It can be divided into medial area 14, central-anterior area 11, central-posterior area 13 and lateral area 47/12 (Carmichael and Price, 1994; Wallis, 2007). These areas are separated by three major sulci, namely the medial orbital sulcus, lateral orbital sulcus and transverse orbital sulcus. In terms of cytoarchitecture, the human OFC comprises an anterior granular cortex and a posterior agranular cortex, which are distinguished on the basis of the presence or absence of small and round neurons in layer IV (Wise, 2008; Wallis, 2012). This anterior-to-posterior gradient in cytoarchitecture of OFC is shared by other nonhuman primates, including macaques and marmosets (a more distant relative to humans than macaques) (Burman and Rosa, 2009). In addition, OFC connectivity in humans and monkeys are similar – for example, area 47/12 in both species are strongly connected to regions such as area 44v, anterior temporal regions, striatum, hypothalamus, hippocampus and amygdala (Neubert *et al.*, 2015). Owing to the similarities in cytoarchitecture and connectivity profiles, it is widely accepted that human and monkey OFCs are homologous. In contrast, the rodent OFC is arguably a homologue of only the posterior human OFC (mainly the posterior part of area 13), as it consists of an agranular cortex only. Thus, findings from the monkey OFC are likely generalizable to humans, but caution should be exercised in extrapolating rodent OFC data to humans.

Stimulus–reward associations

Like the vmPFC and striatum, the OFC has been shown to encode reward value. More specifically, a major function of central OFC area 11/13 is to encode stimulus–reward associations – the value of a stimulus based on past experiences with it (Thorpe *et al.*, 1983; Tremblay and Schultz, 1999; Padoa-Schioppa and Assad, 2006, 2008; Bouret and Richmond, 2010). For example, if an animal has learnt that objects A and B are associated with a reward of an apple or a grape, respectively, a population of central OFC neurons will then encode the value of object A, and a separate population will encode the value of object B (Padoa-Schioppa and Assad, 2006, 2008). Importantly, the neuronal activity is independent of visuospatial features of the stimuli and the motor response required to obtain the object, suggesting that the signal is related specifically to the value of the object itself.

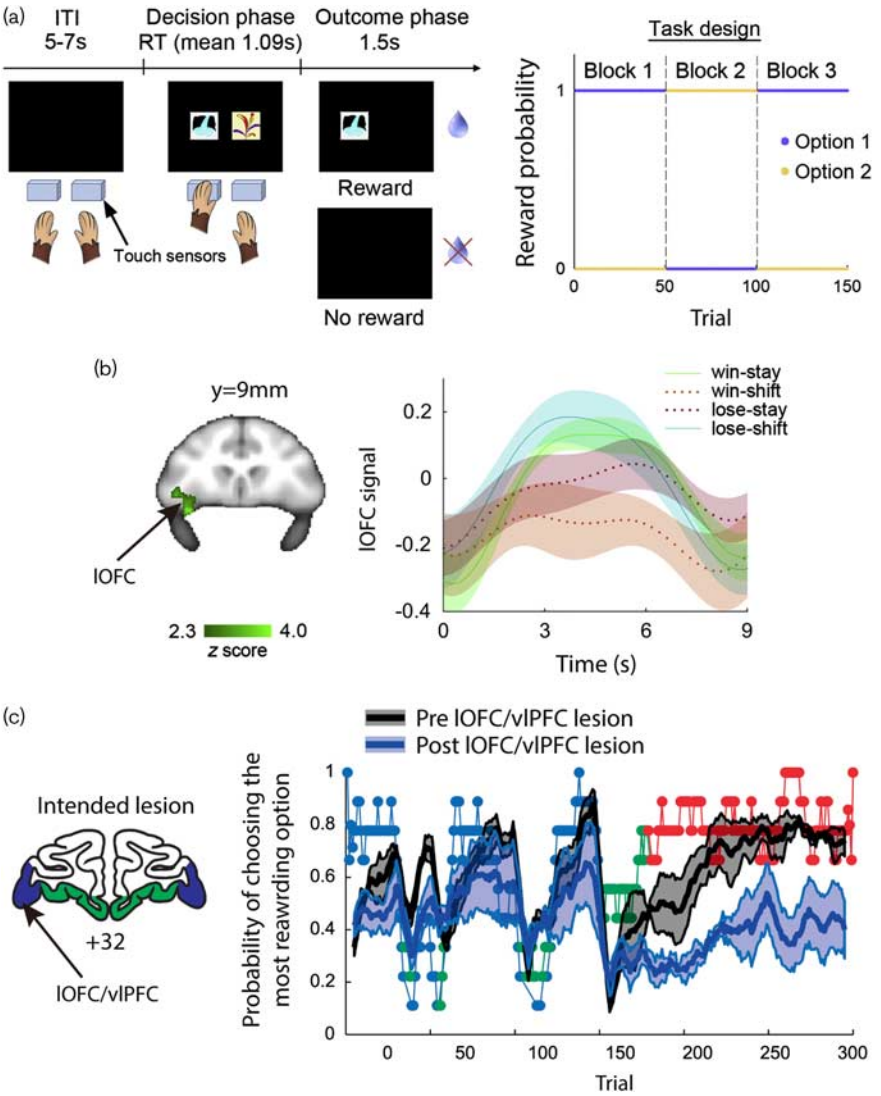
The notion that the central OFC area 11/13 is important for learning stimulus–reward associations fits well with findings from reinforcement devaluation studies. In a typical study, subjects must choose between two objects that are associated with different rewards (e.g. a grape and an apple). These choices are assessed at baseline, and after a devaluation session in which they are fed with one of the rewards to satiety. Usually, subjects avoid the sated reward after the devaluation session. However, this devaluation effect is weaker in monkeys with bilateral central OFC lesions (Izquierdo *et al.*, 2004; Murray and Izquierdo, 2007; Rudebeck and Murray, 2011a, 2011b), as well as in monkeys with smaller central OFCs (Burke *et al.*, 2014). This suggests an important role for the central OFC in updating stimulus–reward associations.

In addition, some have argued that the central OFC is involved in the choice selection process itself. This is based on the aforementioned findings that the firing of individual neurons captures the value of a presented option, while the firing of other neurons within the same region captures the value of the chosen option. Importantly, however, the activity of individual neurons in OFC reflect only the value of a single option, and is independent of the value of the alternative (Padoa-Schioppa and Assad, 2006, 2008). Thus, unlike vmPFC neurons, the activities of central OFC neurons do not show any evidence of comparison or competition between the available options. If one accepts that an important signature for the choice selection process is value comparison (see section 'Ventromedial prefrontal cortex' above), separate populations of OFC neurons are more likely to provide an input to this process, rather than be central to the decision-making process itself.

Flexible decision-making

In addition to encoding stimulus–reward associations, a second major function of the OFC is to guide flexible decisions. A typical paradigm to assess flexible decision-making is the reversal learning task. Such tasks require participants to choose between one of two stimuli, one of which is associated with a reward, and the other an omission (Fig. 4a). The key manipulation is that the reward contingency is reversed once there is a high probability of the individual choosing the rewarded stimulus – the previously rewarded stimulus becomes nonrewarded and vice versa. Human fMRI studies of reinforcement learning have consistently reported strong activity at the OFC when participants reverse their choices (Monchi *et al.*, 2001; O'Doherty *et al.*, 2001; Kringelbach and Rolls, 2003; Ghahremani *et al.*, 2010; Hampshire *et al.*, 2012). In addition, patients with OFC lesions show deficits in choice reversal, suggesting that the OFC plays a causal role in generating flexible decisions (Hornak *et al.*, 2004; Fellows, 2011). However,

Fig. 4



The role of orbitofrontal cortex (OFC) in flexible decision-making. (a) An example of an object discrimination reversal task (left). Participants choose repeatedly between two objects (sometimes three in other studies). Each object is associated with a certain probability of gaining a reward (usually a primary reinforcer for animals, such as food, or a secondary reward for humans). Initially, one option is associated with a higher reward probability than the other (right). After a while, the reward contingency will be reversed – the more rewarding option becomes less rewarding and vice versa. (b) fMRI data showed that the signal in the lateral OFC (area 47/12) was stronger when individuals were about to repeat the choice of a rewarded option (win-stay; green line), or switch to the alternative after choosing a nonrewarded option (lose-shift; blue line) (a, b) (adapted from Chau *et al.*, 2015). (c) After the lateral OFC (as well as the ventrolateral PFC; blue lines) was lesioned, individuals were poorer at choosing the more rewarding option after the reversal in reward contingency (trials labelled by red dots). The color of the dot on each trial (red, blue or green) indicates which option (a, b or c) was the most rewarding option on that trial. ITI, intertrial interval; lOFC, lateral orbitofrontal cortex; RT, response time; vlPFC, ventrolateral prefrontal cortex (adapted from Rudebeck *et al.*, 2017). Adaptations are themselves works protected by copyright. So in order to publish this adaptation, authorization must be obtained both from the owner of the copyright in the original work and from the owner of copyright in the translation or adaptation.

given that human OFC lesions are rarely focal, such studies are limited in revealing the precise OFC subdivision that contributes to flexible decision-making.

Studies on animals with homologous OFC areas, such as macaques and marmosets, have been able to provide further insights. Traditionally, deficits in flexible decision-making have been attributed to lesions of central OFC areas 11/13. However, some of these earlier findings may have been attributable to damage in neighbouring regions. Recent studies that have specifically and precisely lesioned areas 11/13 in macaques using neurotoxin have failed to observe any impaired performance in reversal learning tasks (Kazama and Bachevalier, 2009; Rudebeck *et al.*, 2013). In our recent study, we trained macaques to perform such a task while undergoing fMRI (Chau *et al.*, 2015). We found that area 47/12, rather than area 11/13, was particularly active when

the animals reversed their choices according to a change in reward contingencies. In addition, area 47/12 was also more active when animals repeated their choice of a rewarding option – in other words, the signal in this area was related to the implementation of a win-stay/lose-shift strategy, an optimal strategy for guiding flexible decisions (Fig. 4b).

The causal role of area 47/12 in flexible decision making has been further confirmed by a recent lesion study in macaques. Rudebeck *et al.* (2017) lesioned a lateral prefrontal region that includes area 47/12 (as well as the neighbouring ventrolateral PFC), and found that these animals performed poorly in reversing their choices after the change in reward contingency (Fig. 4c). Interestingly, they also tested macaques with lesions in other OFC regions, including central areas 11/13 and medial area 14, and found that these animals' performance was comparable to controls. In summary, current data suggest a division of labour in the primate OFC, with central areas 11/13 involved in value representation and stimulus–reward associations, and lateral areas 47/12 in flexible decision-making.

The roles of mesolimbic dopamine in reward-based signalling

Turning now to the basal ganglia, a key reward pathway is the subcortical projection from the DA-rich ventral tegmental area (VTA) of the midbrain to the ventral striatum, which comprises a critical part of the mesolimbic pathway (Fig. 1) (Bjorklund and Dunnett, 2007). The ventral striatum is the major input structure to the basal ganglia, and comprises the following: the NAc; the caudate nucleus and putamen ventral to the rostral internal capsule; the olfactory tubercle and the rostrolateral portion of the anterior perforated space adjacent to the lateral olfactory tract in primates (Heimer *et al.*, 1999). The striatum is broadly preserved across commonly studied animals, including humans, monkeys and rodents, which provide a solid foundation for generalizing findings about striatal DA across species. In addition to the striatum, the VTA projects to limbic structures including the amygdala and hippocampus. This mesolimbic pathway is central to reward-based learning and motivation, and provides a crucial link between emotion and action (Mogenson *et al.*, 1980; Salamone and Correa, 2012; Chong and Husain, 2016).

Midbrain dopaminergic neurons

A well-described function of dopaminergic neurons in the VTA is in signalling a reward prediction error – the difference between expected and actual reward outcomes (Schultz, 1986). Early studies measured the firing rates of midbrain DA neurons in monkeys while they performed a Pavlovian behavioural conditioning task. The recorded neurons were identified as dopaminergic on the basis of their location and firing pattern. The animals were trained to respond to auditory and visual cues that indicated the presence of a food reward, and these responses corresponded to spikes in DA firing rates that represented expected reward. In trials wherein reward was omitted, there was a marked reduction in firing rate following the initial spike. These results were later modelled using temporal difference learning algorithms, which confirmed that changes in DA firing rates corresponded to reward prediction errors (Schultz *et al.*, 1997). These neural responses scale according to differences in magnitude of possible rewards, rather than absolute differences in expected value (Tobler *et al.*, 2005). Such experiments provided important contributions to our understanding of the role of DA neurons in reinforcement learning.

Recent advances in optogenetics have provided even more direct evidence of the role of DA neurons in reinforcement learning. Traditionally, neurons have been presumptively labelled as dopaminergic on the basis of their location and activity, but this approach has recently been criticized (Lammel *et al.*, 2008). In contrast, state-of-the-art optogenetic techniques allow researchers to definitively identify midbrain dopaminergic neurons. For example, one study used light-sensitive channelrhodopsin to tag dopaminergic neurons in the rodent VTA, and recorded neuronal activity in the same region (Cohen *et al.*, 2012). By testing these mice in an association learning task, the data definitively confirmed that reward prediction errors were signalled by specific dopaminergic neurons within the VTA. Subsequent studies have also confirmed that VTA dopaminergic neurons compute reward prediction errors by an output subtraction mechanism, in keeping with previously suggested models of reinforcement learning (e.g. temporal difference models) (Eshel *et al.*, 2015, 2016). Finally, an impressive series of optogenetic experiments has shown that prediction error signals are not unique to the VTA; rather, partial components of those signals are encoded in a redundant manner across a distributed network of subcortical areas, which ultimately converge onto DA neurons (Tian *et al.*, 2016).

Striatal dopamine

Like the VTA, extensive data across multiple species show that the ventral striatum is sensitive to reward prediction errors. The magnitude of prediction errors correlates specifically with DA release from the rodent striatum, as recorded at high temporal resolution using fast-scan cyclic voltammetry (Gan *et al.*, 2010; Papageorgiou *et al.*, 2016; Syed *et al.*, 2016). Human fMRI studies provide convergent evidence, showing that the ventral striatum encodes reward prediction error (Pagnoni *et al.*, 2002; McClure *et al.*, 2003; Abler *et al.*, 2006). Subsequent work showed that prediction error signals from these areas are processed in the ventral putamen to learn stimulus–reward associations (Tobler *et al.*, 2006). Interestingly, these reward prediction error signals in the human striatum could be modulated by exogenous administration of levodopa or haloperidol, which enhanced or

antagonized dopaminergic function, respectively (Pessiglione *et al.*, 2006). Together, these data indicate that striatal synaptic plasticity is important in representing prediction errors, and translating action–reward associations into optimum behavioural policies.

How can the role of the striatum in reward-based learning be reconciled with its other well-characterized role in motor control? The prevailing framework considers that phasic bursts of striatal DA activity are central to encoding reward prediction errors, while slower fluctuations in tonic levels of striatal DA are more closely related to locomotor activity. However, this traditional view has been challenged by emerging optogenetic data showing that phasic signalling in striatum-targeting dopaminergic axons is capable of triggering locomotion in mice (Howe and Dombeck, 2016). This close relationship between reward processing and motor execution has been emphasized by separate studies showing that the expected phasic striatal DA release that follows a reward-predicting cue is present only when the required action is correctly initiated, but is otherwise attenuated (Syed *et al.*, 2016). Such findings emphasize a close mechanistic link between learning and motor initiation, and have led to recent attempts to more parsimoniously explain the role of the striatal DA in both reward-based processes and motor control (Berke, 2018).

Role of other neurotransmitter systems

Although the focus of this review is on dopaminergic signalling, we emphasize that DA has complex interactions with other neurotransmitter systems (e.g. GABA, acetylcholine, noradrenaline) in guiding reward-based decisions. For example, GABAergic signalling in the VTA facilitates the rapid reduction in firing rates of dopaminergic neurons associated with a negative prediction error (Eshel *et al.*, 2015). Some have also proposed that the switch between reward-based learning and motor control may be driven by cholinergic interneurons, which modulate the firing rate of DA terminals in the striatum (Berke, 2018). In addition, noradrenergic neurons in the locus coeruleus also have extensive projections to the PFC, and the separate roles of noradrenaline and DA in decision-making are only just coming into focus. For example, a recent study required rhesus monkeys to decide whether to accept or reject different amounts of juice that were associated with varying levels of physical effort (Varazzani *et al.*, 2015). When the monkeys were presented with an option, dopaminergic neurons (specifically within the substantia nigra) encoded both the reward and effort cost associated with that option. In contrast, noradrenergic neurons increased mainly with the production of the effortful response. Together, these results suggest that dopaminergic neurons mainly encode the subjective value of an option (which integrates an action's costs and benefits), whereas noradrenergic neurons reflect the energisation of behaviour. The interactions between DA and other neurotransmitter systems in value-based decision-making are beyond the scope of this review, but will be a critical area of investigation for future studies.

Dopaminergic connectivity of reward-sensitive prefrontal cortex regions

To summarize, a large volume of data indicates that regions within the PFC and basal ganglia are broadly involved in encoding value. Importantly, these areas are heavily interconnected (Fig. 1). The major dopaminergic input to the PFC is by the mesocortical route – a direct projection from the VTA. The PFC in turn sends substantial efferent output to the ventral striatum. In human and nonhuman primates, this output is topographically organized along a clear connectivity gradient (Fig. 1; red to yellow arrows) (Haber and Knutson, 2010; Haber and Behrens, 2014). Specifically, the posterior PFC (including the dACC) is strongly connected to the dorsal striatum, and the anterior PFC (including vmPFC and OFC) is strongly connected to the ventral striatum. Together, therefore, the PFC, striatum and midbrain are organized within distinct corticobasal ganglia loops that form the core of the brain's reward pathway (Alexander *et al.*, 1986; Sesack and Pickel, 1992).

A key challenge for the field is to reconcile the two seemingly separate systems of value-based representation in the striatum and PFC. As discussed above, traditional accounts emphasize the importance of midbrain tegmental and striatal reward prediction errors in learning action–reward associations. However, accumulating data clearly indicate that the PFC implements multiple mechanisms for reward-based learning, some of which very closely resemble those traditionally attributed to DA-based reinforcement learning. As discussed above, regions of the PFC represent the value of actions, objects and states (Padoa-Schioppa and Assad, 2006; Rushworth and Behrens, 2008), and encode, not only the recent history of actions and rewards (Seo and Lee, 2008; Seo *et al.*, 2012; Tsutsui *et al.*, 2016), but also reward prediction errors themselves.

In humans, for example, BOLD activity in both the striatum and OFC decreases with negative prediction errors, and increases with positive prediction errors in appetitive learning tasks (McClure *et al.*, 2003; O'Doherty *et al.*, 2003). Similarly, disrupting the dopaminergic innervation of the marmoset OFC results in more stochastic choices (relative to sham lesions) in a reversal learning task (Walker *et al.*, 2009; Clarke *et al.*, 2014). In addition, the OFC-lesioned animals showed greater persistence in choosing a previously rewarding option (i.e. slower extinction). Such findings provide important evidence that mesocortical DA may play a role in modulating OFC activity during the generation of flexible decisions.

How might DA convey the result of value computations across these corticostriatal loops? DA is likely to modulate activity within this pathway in a bidirectional manner. Intra-VTA stimulation leads to dopaminergic release, and measurable physiological effects, on PFC neurons. It is thought that tonic (~1–6 Hz) DA release in the PFC maintains an extrasynaptic background concentration of DA, while phasic signalling occurs in response to behaviourally relevant stimuli. Indeed, just such a mechanism is understood to play a role in working memory processes. Conversely, when DA was depleted locally within the marmoset OFC, elevated DA levels were observed at the striatum (Clarke *et al.*, 2014). This suggests that striatal DA is sensitive to DA levels in the PFC, and that region-specific DA can interact dynamically with the corticostriatal pathways to drive reward-based decisions. Exciting refinements to this framework are undoubtedly poised to occur given the recent conceptual shifts in the role of phasic/tonic signalling to reward and motor control at the level of the striatum (see above) (Berke, 2018).

Indeed, optogenetic studies in rodents are beginning to elucidate the functional mechanisms underlying reward-based dopaminergic signalling in the corticostriatal pathways. In two recent studies, rodents received optogenetic stimulation while performing reversal learning tasks that required flexible switching between two rules. One study tested the contributions of the specific pathway between VTA and the prelimbic cortex (which is arguably homologous to human dACC; Heilbronner and Hayden, 2016) to flexible behaviour (Ellwood *et al.*, 2017). Once animals started to respond reliably by one rule, the VTA–prelimbic pathway was either tonically or phasically stimulated, and this stimulation then continued throughout the rest of the task. The results showed that phasic stimulation resulted in animals being unable to maintain the previously established rule, resulting in their choices becoming more stochastic. In contrast, tonic stimulation did not impair the animals' ability to maintain the current rule – indeed, animals instead made perseverative errors after a rule switch, indicating a failure to adapt. These findings show the dissociable roles of phasic and tonic VTA–prelimbic DA input in maintaining and updating value representations.

A separate study applied excitatory and inhibitory optogenetics to test the prelimbic–NAc pathway (Cui *et al.*, 2018). The results indicated that animals were slower to adjust to a new rule after a rule switch when the prelimbic–NAc pathway was inhibited. In contrast, they were faster to adapt their behaviour when the pathway was excited – note that this was an opposite effect to that observed after stimulation of the VTA–prelimbic pathway (Ellwood *et al.*, 2017). Interestingly, such stimulation was even able to counteract the impaired behavioural adaptation caused by local depletion of striatal DA. Taken together, the studies by Ellwood *et al.* (2017) and Cui *et al.* (2018) show that the VTA, prelimbic cortex and NAc interact to guide behavioural flexibility in a changing environment. Further studies should be conducted to test the subtle functional differences of these pathways.

Another outstanding question is how value-based representations in the PFC and basal ganglia interact computationally, and how DA might drive that interaction. A current consensus is that the dopaminergic midbrain and striatum implement model-free reinforcement learning, which is based on direct associations between stimulus and response. For example, temporal difference models have been compelling in explaining the activity of dopaminergic neurons in VTA (Schultz *et al.*, 1997; Watabe-Uchida *et al.*, 2017). In contrast, the PFC is thought to implement a model-based type of reinforcement learning, which is based on internal representations of task structure (Daw *et al.*, 2005; Bromberg-Martin *et al.*, 2010). Recently, some have proposed to integrate both types of framework under a single theory of reward-based decision-making, in order to more parsimoniously describe the computations underlying reward valuation in the corticostriatal pathways (Wang *et al.*, 2018). Others have proposed inter-region models to describe the interactions between neurons in the frontal and parietal lobes during working memory and decision-making (Murray *et al.*, 2017). A promising path for future research will be to refine such models to account for the interactions between these regions as a function of DA release.

Studying prefrontal dopamine in humans

Despite the highly organized corticostriatal connectivity, surprisingly little is known about how mesocortical DA modulates decision-making signals in different subregions of the PFC, especially in humans. Studying the function of region-specific DA is challenging, because it requires a high degree of spatiotemporal specificity. It requires anatomical specificity to focus on a defined brain region (e.g. dACC, vmPFC or OFC), and/or a defined neural circuit (e.g. the VTA–dACC pathway). It requires neurochemical specificity to focus on DA and its specific receptors, rather than the general function of a neural region or circuit. It also requires temporal specificity to test the role of DA in a precise event or cognitive process. In nonhuman species, such investigations are often conducted using invasive methods, such as fast cyclic voltammetry, microdialysis, DA-selective lesion or, more recently, DA-selective optogenetic stimulation, all of which are not feasible to apply in humans.

Given that human research is necessarily limited by our inability to measure DA release noninvasively, our understanding of the role of prefrontal DA in human decision-making relies partly on cross-species comparisons. Thus, as we have attempted to emphasize in this review, it is essential to be mindful of the differences in cross-species homologies and experimental paradigms

that might limit our interpretation of cross-species data. However, other effective methodologies exist to examine region-specific DA function in humans less invasively. For instance, although fMRI only captures surrogate markers of neuronal activity (the BOLD response), and lacks the specificity to isolate the effect of individual neurotransmitters, previous studies have suggested that the BOLD signal can capture dopaminergic responses reasonably well (Duzel *et al.*, 2009). Combining fMRI with dopaminergic manipulations in healthy individuals or patient populations may therefore provide a useful approach to test the function of DA within different prefrontal areas in humans.

Another approach to elucidate the role of prefrontal DA in human decision-making has been through neurogenetic studies. The variability in DA function across individuals has been attributed to variability in a number of DA-specific genes. DA levels in the PFC are affected by polymorphisms in the catechol-O-methyltransferase (*COMT*) gene, which generates an enzyme involved in the degradation of DA. In contrast, DA levels in the striatum are affected by polymorphisms in the *DRD2* gene (which generate the DA D2 receptor), and in the *DARPP-32* gene (which generates a protein for striatal synaptic plasticity). Frank *et al.* (2009) recruited healthy volunteers with different polymorphisms of these genes, and tested how genetic variability accounts for differences in decision-making (Doll *et al.*, 2011, 2016). Their data revealed that the *COMT* genotype predicted exploratory decisions, susceptibility to confirmation bias, and model-based learning. In contrast, *DRD2* or *DARPP-32* genotype predicted exploitative decisions, and model-free learning. These findings provide evidence that prefrontal and striatal DA have dissociable roles in decision-making, and more broadly show how genetic variability may be a useful proxy to studying regional specializations of human DA function.

Further specificity can be achieved by combining such genetic approaches with neuroimaging techniques. Gao *et al.* (2016) performed a gambling task on participants with different *COMT* genotypes, while recording their resting-state neural activity using fMRI. The stimuli either emphasized the gains or the losses of identical gambles, and participants demonstrated a typical 'framing effect', such that, in general, they tended to avoid risky choices when losses were emphasized. Importantly, the magnitude of this framing effect was associated with variability in the *COMT* gene, and this relationship was mediated by the resting-state connectivity strength between the OFC and amygdala. These results illustrate the potential of combined genetic/neuroimaging approaches in understanding regional modulation of DA in the human PFC.

Another potentially useful approach is to image patients with dopaminergic dysfunction, such as those with idiopathic Parkinson's disease. Patients with disorders of DA function typically have high rates of motivational impairments, such as apathy (Chong *et al.*, 2018). In addition, their sensitivity to reward is typically impaired – a deficit that is ameliorable with DA replacement (Chong *et al.*, 2015; Chong and Husain, 2016; Muhammed *et al.*, 2016). In a two-stage reinforcement learning experiment, patients with Parkinson's disease underwent fMRI scanning when they were ON or OFF DA medication (Shiner *et al.*, 2012). In the initial learning stage, patients were presented on each trial with pairs of stimuli, and were asked to learn which of the two was more often associated with a correct outcome. In a subsequent test phase, they were presented with the same stimuli, but in different combinations, and were again asked to choose the more correct option. The key result was that drug state had no effect on the initial learning of stimulus values. Instead, patients performed more accurately in the ON versus OFF state only in the test phase, when they had to perform novel associations. Interestingly, fMRI data showed that the vmPFC and the NAc encoded a signal related to the value of the chosen option, but only in the ON state, and not when patients were OFF. These results suggest that value signals in vmPFC are modulated by DA, presumably by the mesocortical route, in deciding between novel associations.

Summary and concluding remarks

The PFC, together with its bidirectional connections with the basal ganglia, plays important roles in reward-based decision-making. These areas are connected in a highly organized, topographic manner, with each node of this network having distinct, yet partially overlapping, roles in the representation of value, and in the decision-making process itself (Izuma *et al.*, 2008; Zink *et al.*, 2008; Levy and Glimcher, 2012). With current advances in neurophysiological techniques, we are well positioned to elucidate the spatiotemporal properties of dopaminergic neurons in facilitating cortical value representations. In humans, the application of a convergence of techniques, such as neuroimaging, genetics, patient studies and pharmacological manipulations, offers complementary approaches to understanding the properties of the mesocorticolimbic and corticostriatal pathways. These data should be integrated with novel computational models that can provide a more holistic understanding of how region-specific DA contributes to the broader neural circuitry during reward-based decision-making.

Acknowledgements

B.K.H.C. is supported by the Hong Kong Research Grants Council (15603517) and the Hong Kong Polytechnic University Start-up Fund. T.C. is supported by the Australian Research Council (DP 180102383, DE 180100389), the Mason Foundation, the Brain Foundation and the Society for Mental Health Research. T.C. and H.J. are supported by the Rebecca L. Cooper

Medical Research Foundation. The authors thank Michelle Wong for commenting on an earlier version of this paper.

Conflicts of interest

There are no conflicts of interest.

References

Abler B, Walter H, Erk S, Kammerer H, Spitzer M (2006). Prediction error as a linear function of reward probability is coded in human nucleus accumbens. *Neuroimage* **31**:790–795.

Alexander GE, De Long MR, Strick PL (1986). Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu Rev Neurosci* **9**:357–381.

Amemori K, Graybiel AM (2012). Localized microstimulation of primate pregenual cingulate cortex induces negative decision making. *Nat Neurosci* **15**:776–785.

Bartra O, McGuire JT, Kable JW (2013). The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage* **76**:412–427.

Bechara A, Damasio H, Tranel D, Damasio AR (2005). The Iowa Gambling Task and the somatic marker hypothesis: some questions and answers. *Trends Cogn Sci* **9**:159–162.

Beckmann M, Johansen-Berg H, Rushworth MF (2009). Connectivity-based parcellation of human cingulate cortex and its relation to functional specialization. *J Neurosci* **29**:1175–1190.

Behrens TE, Woolrich MW, Walton ME, Rushworth MF (2007). Learning the value of information in an uncertain world. *Nat Neurosci* **10**:1214–1221.

Berke JD (2018). What does dopamine mean? *Nat Neurosci* **21**:787–793.

Bjorklund A, Dunnett SB (2007). Dopamine neuron systems in the brain: an update. *Trends Neurosci* **30**:194–202.

Boorman ED, Behrens TE, Woolrich MW, Rushworth MF (2009). How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron* **62**:733–743.

Botvinick MM (2007). Conflict monitoring and decision making: reconciling two perspectives on anterior cingulate function. *Cogn Affect Behav Neurosci* **7**:356–366.

Botvinick MM, Cohen JD, Carter CS (2004). Conflict monitoring and anterior cingulate cortex: an update. *Trends Cogn Sci* **8**:539–546.

Bouret S, Richmond BJ (2010). ventromedial and orbital prefrontal neurons differentially encode internally and externally driven motivational values in monkeys. *J Neurosci* **30**:8591–8601.

Bromberg-Martin ES, Matsumoto M, Nakahara H, Hikosaka O (2010). Multiple timescales of memory in lateral habenula and dopamine neurons. *Neuron* **67**:499–510.

Bryden DW, Johnson EE, Tobia SC, Kashtelyan V, Roesch MR (2011). Attention for learning signals in anterior cingulate cortex. *J Neurosci* **31**:18266–18274.

Burke SN, Thome A, Plange K, Engle JR, Trouard TP, Gothard KM, Barnes CA (2014). Orbitofrontal cortex volume in area 11/13 predicts reward devaluation, but not reversal learning performance, in young and aged monkeys. *J Neurosci* **34**:9905–9916.

Burman KJ, Rosa MG (2009). Architectural subdivisions of medial and orbital frontal cortices in the marmoset monkey (*Callithrix jacchus*). *J Comp Neurol* **514**:11–29.

Carmichael ST, Price JL (1994). Architectonic subdivision of the orbital and medial prefrontal cortex in the macaque monkey. *J Comp Neurol* **346**:366–402.

Carter CS, Braver TS, Barch DM, Botvinick MM, Noll D, Cohen JD (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science* **280**:747–749.

Chau BK, Sallet J, Papageorgiou GK, Noonan MP, Bell AH, Walton ME, Rushworth MF (2015). Contrasting roles for orbitofrontal cortex and amygdala in credit assignment and learning in macaques. *Neuron* **87**:1106–1118.

Chong TT, Husain M (2016). The role of dopamine in the pathophysiology and treatment of apathy. *Prog Brain Res* **229**:389–426.

Chong TT, Bonnelle V, Manohar S, Veromann KR, Muhammed K, Tofaris GK, *et al.* (2015). Dopamine enhances willingness to exert effort for reward in Parkinson's disease. *Cortex* **69**:40–46.

Chong TT-J, Bonnelle V, Husain M (2016). Quantifying motivation with effort-based decision-making paradigms in health and disease. *Prog Brain Res* **229**:71–100.

Chong TT, Apps M, Giehl K, Sillence A, Grima LL, Husain M (2017). Neurocomputational mechanisms underlying subjective valuation of effort costs. *PLoS Biol* **15**:e1002598.

Chong TT, Bonnelle V, Veromann KR, Juurmaa J, Taba P, Plant O, Husain M (2018). Dissociation of reward and effort sensitivity in methcathinone-induced Parkinsonism. *J Neuropsychol* **12**:291–297.

Churchland AK, Kiani R, Shadlen MN (2008). Decision-making with multiple alternatives. *Nat Neurosci* **11**:693–702.

Clarke HF, Cardinal RN, Rygula R, Hong YT, Fryer TD, Sawiak SJ, *et al.* (2014). Orbitofrontal dopamine depletion upregulates caudate dopamine and alters behavior via changes in reinforcement sensitivity. *J Neurosci* **34**:7663–7676.

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

Cole MW, Yeung N, Freiwald WA, Botvinick MM (2009). Cingulate cortex: diverging data from humans and monkeys. *Trends Neurosci* **32**:566–574.

Constantinescu AO, O'Reilly JX, Behrens TEJ (2016). Organizing conceptual knowledge in humans with a gridlike code. *Science* **352**:1464–1468.

Croxson PL, Walton ME, O'Reilly JX, Behrens TE, Rushworth MF (2009). Effort-based cost-benefit valuation and the human brain. *J Neurosci* **29**:4531–4541.

Cui Q, Li Q, Geng H, Chen L, Ip NY, Ke Y, Yung WH (2018). Dopamine receptors mediate strategy abandoning via modulation of a specific prelimbic cortex–nucleus accumbens pathway in mice. *Proc Natl Acad Sci USA* **115**:E4890–E4899.

Damasio AR (1996). The somatic marker hypothesis and the possible functions of the prefrontal cortex. *Philos Trans R Soc Lond B Biol Sci* **351**:1413–1420.

Daw ND, Niv Y, Dayan P (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci* **8**:1704–1711.

Debener S, Ullsperger M, Siegel M, Fiehler K, von Cramon DY, Engel AK (2005). Trial-by-trial coupling of concurrent electroencephalogram and functional magnetic resonance imaging identifies the dynamics of performance monitoring. *J Neurosci* **25**:11730–11737.

Doll BB, Hutchison KE, Frank MJ (2011). Dopaminergic genes predict individual differences in susceptibility to confirmation bias. *J Neurosci* **31**:6188–6198.

Doll BB, Bath KG, Daw ND, Frank MJ (2016). Variability in dopamine genes dissociates model-based and model-free reinforcement learning. *J Neurosci* **36**:1211–1222.

Duzel E, Bunzeck N, Guitart-Masip M, Wittmann B, Schott BH, Tobler PN (2009). Functional imaging of the human dopaminergic midbrain. *Trends Neurosci* **32**:321–328.

Ellwood IT, Patel T, Wadia V, Lee AT, Liptak AT, Bender KJ, Sohal VS (2017). Tonic or Phasic stimulation of dopaminergic projections to prefrontal cortex causes mice to maintain or deviate from previously learned behavioral strategies. *J Neurosci* **37**:8315–8329.

Eshel N, Bukwich M, Rao V, Hemmelder V, Tian J, Uchida N (2015). Arithmetic and local circuitry underlying dopamine prediction errors. *Nature* **525**:243–246.

Eshel N, Tian J, Bukwich M, Uchida N (2016). Dopamine neurons share common response function for reward prediction error. *Nat Neurosci* **19**:479–486.

Fellows LK (2011). Orbitofrontal contributions to value-based decision making: evidence from humans with frontal lobe damage. *Ann N Y Acad Sci* **1239**:51–58.

Fellows LK, Farah MJ (2005). Different underlying impairments in decision-making following ventromedial and dorsolateral frontal lobe damage in humans. *Cereb Cortex* **15**:58–63.

Fellows LK, Farah MJ (2007). The role of ventromedial prefrontal cortex in decision making: judgment under uncertainty or judgment per se? *Cereb Cortex* **17**:2669–2674.

Frank MJ, Doll BB, Oas-Terpstra J, Moreno F (2009). Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nat Neurosci* **12**:1062–1068.

Gan JO, Walton ME, Phillips PE (2010). Dissociable cost and benefit encoding of future rewards by mesolimbic dopamine. *Nat Neurosci* **13**:25–27.

Gao X, Gong P, Liu J, Hu J, Li Y, Yu H, *et al.* (2016). COMT Val158Met polymorphism influences the susceptibility to framing in decision-making: OFC–amygdala functional connectivity as a mediator. *Hum Brain Mapp* **37**:1880–1892.

Ghahremani DG, Monterosso J, Jentsch JD, Bilder RM, Poldrack RA (2010). Neural components underlying behavioral flexibility in human reversal learning. *Cereb Cortex* **20**:1843–1852.

Grabenhorst F, Rolls ET (2011). Value, pleasure and choice in the ventral prefrontal cortex. *Trends Cogn Sci* **15**:56–67.

Haber SN (2011). Neuroanatomy of reward: a view from the ventral striatum. In: Gottfried JA, editor. *Neurobiology of sensation and reward*. Boca Raton, FL: CRC Press.

Haber SN, Knutson B (2010). The reward circuit: linking primate anatomy and human imaging. *Neuropsychopharmacology* **35**:4–26.

Haber SN, Behrens TE (2014). The neural network underlying incentive-based learning: implications for interpreting circuit disruptions in psychiatric disorders. *Neuron* **83**:1019–1039.

Hafting T, Fyhn M, Molden S, Moser MB, Moser EI (2005). Microstructure of a spatial map in the entorhinal cortex. *Nature* **436**:801–806.

Hamid AA, Pettibone JR, Mabrouk OS, Hetrick VL, Schmidt R, Vander Weele CM, *et al.* (2016). Mesolimbic dopamine signals the value of work. *Nat Neurosci* **19**:117–126.

Hampshire A, Chaudhry AM, Owen AM, Roberts AC (2012). Dissociable roles for lateral orbitofrontal cortex and lateral prefrontal cortex during preference driven reversal learning. *Neuroimage* **59**:4102–4112.

Hare TA, Schultz W, Camerer CF, O'Doherty JP, Rangel A (2011). Transformation of stimulus value signals into motor commands during simple choice. *Proc Natl Acad Sci USA* **108**:18120–18125.

Hayden BY, Pearson JM, Platt ML (2011). Neuronal basis of sequential foraging decisions in a patchy environment. *Nat Neurosci* **14**:933–939.

He SQ, Dum RP, Strick PL (1995). Topographic organization of corticospinal projections from the frontal lobe: motor areas on the medial surface of the hemisphere. *J Neurosci* **15**:3284–3306.

Heilbronner SR, Hayden BY (2016). Dorsal anterior cingulate cortex: a bottom-up view. *Annu Rev Neurosci* **39**:149–170.

Heilbronner SR, Rodriguez-Romaguera J, Quirk GJ, Groenewegen HJ, Haber SN (2016). Circuit-based corticostriatal homologies between rat and primate. *Biol Psychiatry* **80**:509–521.

Heimer L, De Olmos JS, Alheid GF, Person J, Sakamoto N, Shinoda K, *et al.* (1999). The human basal forebrain. Part II. The primate nervous system. In: Bloom FE, Bjorkland A, Hokfelt T, editors. *Part III Handbook of chemical neuroanatomy*. Amsterdam: Elsevier. pp. 57–226.

Henri-Bhargava A, Simioni A, Fellows LK (2012). Ventromedial frontal lobe damage disrupts the accuracy, but not the speed, of value-based preference judgments. *Neuropsychologia* **50**:1536–1542.

Holroyd CB, Coles MG (2002a). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol Rev* **109**:679–709.

Holroyd CB, Coles MG (2002b). The neural basis of human error processing: reinforcement learning, dopamine, and the error-related negativity. *Psychol Rev* **109**:679–709.

Holroyd CB, Yeung N (2012). Motivation of extended behaviors by anterior cingulate cortex. *Trends Cogn Sci* **16**:122–128.

Hornak J, O'Doherty J, Bramham J, Rolls ET, Morris RG, Bullock PR, Polkey CE (2004). Reward-related reversal learning after surgical excisions in orbitofrontal or dorsolateral prefrontal cortex in humans. *J Cogn Neurosci* **16**:463–478.

Hosking JG, Cocker PJ, Winstanley CA (2014). Dissociable contributions of anterior cingulate cortex and basolateral amygdala on a rodent cost-benefit decision-making task of cognitive effort. *Neuropsychopharmacology* **39**:1558–1567.

Howe MW, Dombeck DA (2016). Rapid signalling in distinct dopaminergic axons during locomotion and reward. *Nature* **535**:505–510.

Hyman JM, Holroyd CB, Seamans JK (2017). A novel neural prediction error found in anterior cingulate cortex ensembles. *Neuron* **95**:447.e3–456.e3.

Izquierdo A, Suda RK, Murray EA (2004). Bilateral orbital prefrontal cortex lesions in rhesus monkeys disrupt choices guided by both reward value and reward contingency. *J Neurosci* **24**:7540–7548.

Izuma K, Saito DN, Sadato N (2008). Processing of social and monetary rewards in the human striatum. *Neuron* **58**:284–294.

Jocham G, Furlong PM, Kroger IL, Kahn MC, Hunt LT, Behrens TE (2014). Dissociable contributions of ventromedial prefrontal and posterior parietal cortex to value-guided choice. *Neuroimage* **100**:498–506.

Kable JW, Glimcher PW (2007). The neural correlates of subjective value during intertemporal choice. *Nat Neurosci* **10**:1625–1633.

Kazama A, Bachevalier J (2009). Selective aspiration or neurotoxic lesions of orbital frontal areas 11 and 13 spared monkeys' performance on the object discrimination reversal task. *J Neurosci* **29**:2794–2804.

Kennerley SW, Walton ME, Behrens TE, Buckley MJ, Rushworth MF (2006). Optimal decision making and the anterior cingulate cortex. *Nat Neurosci* **9**:940–947.

Kennerley SW, Behrens TE, Wallis JD (2011). Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nat Neurosci* **14**:1581–1589.

Kolling N, Behrens TE, Mars RB, Rushworth MF (2012). Neural mechanisms of foraging. *Science* **336**:95–98.

Kolling N, Wittmann MK, Behrens TE, Boorman ED, Mars RB, Rushworth MF (2016). Value, search, persistence and model updating in anterior cingulate cortex. *Nat Neurosci* **19**:1280–1285.

Kringelbach ML, Rolls ET (2003). Neural correlates of rapid reversal learning in a simple model of human social interaction. *Neuroimage* **20**:1371–1383.

Lammel S, Hetzel A, Hackel O, Jones I, Liss B, Roeper J (2008). Unique properties of mesoprefrontal neurons within a dual mesocorticolimbic dopamine system. *Neuron* **57**:760–773.

Lau B, Glimcher PW (2008). Value representations in the primate striatum during matching behavior. *Neuron* **58**:451–463.

Lebreton M, Jorge S, Michel V, Thirion B, Pessiglione M (2009). An automatic valuation system in the human brain: evidence from functional neuroimaging. *Neuron* **64**:431–439.

Levy DJ, Glimcher PW (2012). The root of all value: a neural common currency for choice. *Curr Opin Neurobiol* **22**:1027–1038.

Lim SL, O'Doherty JP, Rangel A (2011). The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. *J Neurosci* **31**:13214–13223.

Lopez-Persem A, Domenech P, Pessiglione M (2016). How prior preferences determine decision-making frames and biases in the human brain. *eLife* **5**: e20317.

Luppino G, Matelli M, Camarda RM, Gallese V, Rizzolatti G (1991). Multiple representations of body movements in mesial area 6 and the adjacent cingulate cortex: an intracortical microstimulation study in the macaque monkey. *J Comp Neurol* **311**:463–482.

Matsumoto M, Matsumoto K, Abe H, Tanaka K (2007). Medial prefrontal cell activity signaling prediction errors of action values. *Nat Neurosci* **10**:647–656.

McClure SM, Berns GS, Montague PR (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron* **38**:339–346.

Mogenson GJ, Jones DL, Yim CY (1980). From motivation to action: functional interface between the limbic system and the motor system. *Prog Neurobiol* **14**:69–97.

Monchi O, Petrides M, Petre V, Worsley K, Dagher A (2001). Wisconsin Card Sorting revisited: distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. *J Neurosci* **21**:7733–7741.

Muhammed K, Manohar S, Ben Yehuda M, Chong TT, Tofaris G, Lennox G, *et al.* (2016). Reward sensitivity deficits modulated by dopamine are associated with apathy in Parkinson's disease. *Brain* **139**:2706–2721.

Murray EA, Izquierdo A (2007). Orbitofrontal cortex and amygdala contributions to affect and action in primates. *Ann N Y Acad Sci* **1121**:273–296.

Murray EA, Rudebeck PH (2018). Specializations for reward-guided decision-making in the primate ventral prefrontal cortex. *Nat Rev Neurosci* **19**:404–417.

Murray JD, Jaramillo J, Wang XJ (2017). Working memory and decision-making in a frontoparietal circuit model. *J Neurosci* **37**:12167–12186.

Neubert FX, Mars RB, Sallet J, Rushworth MF (2015). Connectivity reveals relationship of brain areas for reward-guided learning and decision making in human and monkey frontal cortex. *Proc Natl Acad Sci USA* **112**: E2695–E2704.

Noonan MP, Chau BKH, Rushworth MFS, Fellows LK (2017). Contrasting effects of medial and lateral orbitofrontal cortex lesions on credit assignment and decision-making in humans. *J Neurosci* **37**:7023–7035.

O'Doherty JP (2004). Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr Opin Neurobiol* **14**:769–776.

O'Doherty J, Kringelbach ML, Rolls ET, Hornak J, Andrews C (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat Neurosci* **4**:95–102.

O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ (2003). Temporal difference models and reward-related learning in the human brain. *Neuron* **38**:329–337.

Ongur D, Price JL (2000). The organization of networks within the orbital and medial prefrontal cortex of rats, monkeys and humans. *Cereb Cortex* **10**:206–219.

Padoa-Schioppa C, Assad JA (2006). Neurons in the orbitofrontal cortex encode economic value. *Nature* **441**:223–226.

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

Pagnoni G, Zink CF, Montague PR, Berns GS (2002). Activity in human ventral striatum locked to errors of reward prediction. *Nat Neurosci* **5**:97–98.

Papageorgiou GK, Baudonnat M, Cucca F, Walton ME (2016). Mesolimbic dopamine encodes prediction errors in a state-dependent manner. *Cell Rep* **15**:221–228.

Papageorgiou GK, Sallet J, Wittmann MK, Chau BKH, Schuffelgen U, Buckley MJ, Rushworth MFS (2017). Inverted activity patterns in ventromedial prefrontal cortex during value-guided decision-making in a less-is-more task. *Nat Commun* **8**:1886.

Parvizi J, Rangarajan V, Shirer WR, Desai N, Greicius MD (2013). The will to persevere induced by electrical stimulation of the human cingulate gyrus. *Neuron* **80**:1359–1367.

Passingham RE, Wise SP (2012). *The neurobiology of the prefrontal cortex : anatomy, evolution, and the origin of insight*, 1st ed. Oxford, UK: Oxford University Press.

Pessiglione M, Seymour B, Flandin G, Dolan RJ, Frith CD (2006). Dopamine-dependent prediction errors underpin reward-seeking behaviour in humans. *Nature* **442**:1042–1045.

Peters J, Buchel C (2010). Neural representations of subjective reward value. *Behav Brain Res* **213**:135–141.

Platt ML, Glimcher PW (1999). Neural correlates of decision variables in parietal cortex. *Nature* **400**:233–238.

Posner MI, Petersen SE (1990). The attention system of the human brain. *Annu Rev Neurosci* **13**:25–42.

Price JL (2007). Definition of the orbital cortex in relation to specific connections with limbic and visceral structures and other cortical regions. *Ann N Y Acad Sci* **1121**:54–71.

Procyk E, Wilson CR, Stoll FM, Faraut MC, Petrides M, Amiez C (2014). Midcingulate motor map and feedback detection: converging data from humans and monkeys. *Cereb Cortex* **26**:467–476.

Rudebeck PH, Murray EA (2011a). Balkanizing the primate orbitofrontal cortex: distinct subregions for comparing and contrasting values. *Ann N Y Acad Sci* **1239**:1–13.

Rudebeck PH, Murray EA (2011b). Dissociable effects of subtotal lesions within the macaque orbital prefrontal cortex on reward-guided behavior. *J Neurosci* **31**:10569–10578.

Rudebeck PH, Walton ME, Smyth AN, Bannerman DM, Rushworth MF (2006). Separate neural pathways process different decision costs. *Nat Neurosci* **9**:1161–1168.

Rudebeck PH, Saunders RC, Prescott AT, Chau LS, Murray EA (2013). Prefrontal mechanisms of behavioral flexibility, emotion regulation and value updating. *Nat Neurosci* **16**:1140–1145.

Rudebeck PH, Saunders RC, Lundgren DA, Murray EA (2017). Specialized representations of value in the orbital and ventrolateral prefrontal cortex: desirability versus availability of outcomes. *Neuron* **95**:1208–1220 e1205.

Rushworth MF, Behrens TE (2008). Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat Neurosci* **11**:389–397.

Salamone JD, Correa M (2012). The mysterious motivational functions of mesolimbic dopamine. *Neuron* **76**:470–485.

Schultz W (1986). Responses of midbrain dopamine neurons to behavioral trigger stimuli in the monkey. *J Neurophysiol* **56**:1439–1461.

Schultz W, Dayan P, Montague PR (1997). A neural substrate of prediction and reward. *Science* **275**:1593–1599.

Schultz W, Carelli RM, Wightman RM (2015). Phasic dopamine signals: from subjective reward value to formal economic utility. *Curr Opin Behav Sci* **5**:147–154.

Schweimer J, Hauber W (2005). Involvement of the rat anterior cingulate cortex in control of instrumental responses guided by reward expectancy. *Learn Mem* **12**:334–342.

Schweimer J, Saft S, Hauber W (2005). Involvement of catecholamine neurotransmission in the rat anterior cingulate in effortrelated decision-making. *Behav Neurosci* **119**:1687–1692.

Seo H, Lee D (2008). Cortical mechanisms for reinforcement learning in competitive games. *Philos Trans R Soc Lond B Biol Sci* **363**:3845–3857.

Seo M, Lee E, Averbeck BB (2012). Action selection and action value in frontal-striatal circuits. *Neuron* **74**:947–960.

Sesack SR, Pickel VM (1992). Prefrontal cortical efferents in the rat synapse on unlabeled neuronal targets of catecholamine terminals in the nucleus accumbens septi and on dopamine neurons in the ventral tegmental area. *J Comp Neurol* **320**:145–160.

Shackman AJ, Salomons TV, Slagter HA, Fox AS, Winter JJ, Davidson RJ (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nat Rev Neurosci* **12**:154.

Shenhav A, Straccia MA, Cohen JD, Botvinick MM (2014). Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. *Nat Neurosci* **17**:1249–1254.

Shidara M, Richmond BJ (2002). Anterior cingulate: single neuronal signals related to degree of reward expectancy. *Science* **296**:1709–1711.

Shiner T, Seymour B, Wunderlich K, Hill C, Bhatia KP, Dayan P, Dolan RJ (2012). Dopamine and performance in a reinforcement learning task: evidence from Parkinson's disease. *Brain* **135**:1871–1883.

Strait CE, Blanchard TC, Hayden BY (2014). Reward value comparison via mutual inhibition in ventromedial prefrontal cortex. *Neuron* **82**:1357–1366.

Stuss DT, Alexander MP, Shallice T, Picton TW, Binns MA, Macdonald R, *et al.* (2005). Multiple frontal systems controlling response speed. *Neuropsychologia* **43**:396–417.

Syed EC, Grima LL, Magill PJ, Bogacz R, Brown P, Walton ME (2016). Action initiation shapes mesolimbic dopamine encoding of future rewards. *Nat Neurosci* **19**:34–36.

Thorpe SJ, Rolls ET, Maddison S (1983). The orbitofrontal cortex: neuronal activity in the behaving monkey. *Exp Brain Res* **49**:93–115.

Tian J, Huang R, Cohen JY, Osakada F, Kobak D, Machens CK, *et al.* (2016). Distributed and mixed information in monosynaptic inputs to dopamine. *Neurons Neuron* **91**:1374–1389.

Tobler PN, Fiorillo CD, Schultz W (2005). Adaptive coding of reward value by dopamine neurons. *Science* **307**:1642–1645.

Tobler PN, O'Doherty JP, Dolan RJ, Schultz W (2006). Human neural learning depends on reward prediction errors in the blocking paradigm. *J Neurophysiol* **95**:301–310.

Tremblay L, Schultz W (1999). Relative reward preference in primate orbitofrontal cortex. *Nature* **398**:704–708.

Tsutsui K, Grabenhorst F, Kobayashi S, Schultz W (2016). A dynamic code for economic object valuation in prefrontal cortex neurons. *Nat Commun* **7**:12554.

Vaidya AR, Fellows LK (2015). Ventromedial frontal cortex is critical for guiding attention to reward-predictive visual features in humans. *J Neurosci* **35**:12813–12823.

Vaidya AR, Fellows LK (2016). Necessary contributions of human frontal lobe subregions to reward learning in a dynamic, multidimensional environment. *J Neurosci* **36**:9843–9858.

Varazzani C, San-Galli A, Gilardeau S, Bouret S (2015). Noradrenaline and dopamine neurons in the reward/effort trade-off: a direct electrophysiological comparison in behaving monkeys. *J Neurosci* **35**:7866–7877.

Vogt BA (2016). Midcingulate cortex: structure, connections, homologies, functions and diseases. *J Chem Neuroanat* **74**:28–46.

Volkow ND, Wise RA, Baler R (2017). The dopamine motive system: implications for drug and food addiction. *Nat Rev Neurosci* **18**:741–752.

Walker SC, Robbins TW, Roberts AC (2009). Differential contributions of dopamine and serotonin to orbitofrontal cortex function in the marmoset. *Cereb Cortex* **19**:889–898.

Wallis JD (2007). Orbitofrontal cortex and its contribution to decision-making. *Annu Rev Neurosci* **30**:31–56.

Wallis JD (2012). Cross-species studies of orbitofrontal cortex and value-based decision-making. *Nat Neurosci* **15**:13–19.

Walton ME, Bannerman DM, Rushworth MFS (2002). The role of rat medial frontal cortex in effort-based decision-making. *J Neurosci* **22**:10996–11003.

Walton ME, Bannerman DM, Alterescu K, Rushworth MFS (2003). Functional specialization within medial frontal cortex of the anterior cingulate for evaluating effort-related decisions. *J Neurosci* **23**:6475–6479.

Walton ME, Groves J, Jennings KA, Croxson PL, Sharp T, Rushworth MFS, Bannerman DM (2009). Comparing the role of the anterior cingulate cortex and 6-hydroxydopamine nucleus accumbens lesions on operant effort-based decision-making. *Eur J Neurosci* **29**:1678–1691.

Wang JX, Kurth-Nelson Z, Kumaran D, Tirumala D, Soyer H, Leibo JZ, *et al.* (2018). Prefrontal cortex as a meta-reinforcement learning system. *Nat Neurosci* **21**:860–868.

Watabe-Uchida M, Eshel N, Uchida N (2017). Neural circuitry of reward prediction error. *Annu Rev Neurosci* **40**:373–394.

Wise SP (2008). Forward frontal fields: phylogeny and fundamental function. *Trends Neurosci* **31**:599–608.

Wittmann MK, Kolling N, Akaishi R, Chau BK, Brown JW, Nelissen N, Rushworth MF (2016). Predictive decision making driven by multiple time-linked reward representations in the anterior cingulate cortex. *Nat Commun* **7**:12327.

Zink CF, Tong Y, Chen Q, Bassett DS, Stein JL, Meyer-Lindenberg A (2008). Know your place: neural processing of social hierarchy in humans. *Neuron* **58**:273–283.