Dopamine and reward – A view from the prefrontal cortex

Bolton K. H. Chau[1], Huw Jarvis[2], Chun-Kit Law[1], Trevor T-J Chong[2]

[1]Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hong Kong

[2]Monash Institute of Cognitive and Clinical Neurosciences, Monash University, Australia

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

Acknowledgements. BKHC is supported by the Hong Kong Research Grants Council (15603517) and The Hong Kong Polytechnic University Start-up Fund. TC is supported by the Australian Research Council (DP 180102383, DE 180100389), the Mason Foundation, the Brain Foundation, and the Society for Mental Health Research. TC and HJ are supported by the Rebecca L Cooper Medical Research Foundation. We would like to thank Michelle Wong for commenting on an earlier version of this paper.

Reward, prefrontal cortex and dopamine

The prefrontal cortex (PFC) is a heterogeneous area that is critical to reward-based decision-making. In particular, the dorsal anterior cingulate cortex (dACC), ventromedial prefrontal cortex (vmPFC), and orbitofrontal cortex (OFC) are frequently implicated in different aspects of choice behaviour. These regions receive projections from midbrain dopamine neurons, and in turn project to other key dopaminergic regions such as the striatum. However, our current understanding of the role of dopamine in reward-based processes is based mainly on studies of midbrain dopaminergic neurons and striatal dopamine release from non-human animal models. An important gap in the literature surrounds the precise functions of dopamine release in the prefrontal cortex, 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 prefrontal cortex and basal ganglia, through which dopaminergic neurotransmission guides reward-based behaviour.

Reward, prefrontal cortex and dopamine

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; and guiding flexible choices (Murray and Rudebeck, 2018). These "reward-sensitive" processes are instantiated in three key subdivisions of the PFC, including 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 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 dopamine neurons via the mesocortical pathway, and in turn project in a highly organised 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 dopamine in reward processing have focused on dopamine neurotransmission within the midbrain and striatum, and it remains largely unclear how dopamine 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 dopamine in encoding reward prediction errors. We then consider how the prefrontal cortex may interact with dopaminergic pathways to facilitate reward-based decisions. Finally, we conclude by highlighting useful approaches to studying prefrontal dopamine in humans that are based on combining currently available methodological techniques.

1. The Anatomy and Function of Reward-Sensitive PFC 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 (Figure 1).

1.1. Dorsal anterior cingulate cortex

The ACC lies on the medial surface of the frontal lobe, and consists of Brodmann areas 24, 25 and 32, which lie in and around the cingulate sulcus. The dorsal ACC (dACC) in turn encompasses regions referred to as the anterior mid-cingulate 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 pre-supplementary 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 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 frontal cortex, such as the dorsolateral prefrontal cortex (dlPFC), 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-supplementary motor area) 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 fMRI 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; Walton et al., 2003; Schweimer and Hauber, 2005; Schweimer et al., 2005; Rudebeck et al., 2006; Walton et al., 2009). Importantly, this lowered motivation is not due to a motor deficit or altered reward sensitivity (Walton et al., 2002; Walton et al., 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 are 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 energisation 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 dopamine signalling (see Section 2), 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 based on the 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 human 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 (Figure 2B) (Kolling et al., 2012). This study therefore 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 costs. 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 operationalised 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. Based on 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 (Figure 2A). Specifically, it is possible to concurrently observe an exploration signal in a relatively ventral dACC region, and a difficulty signal in a relatively dorsal dACC region (sometimes also known as pre-supplementary motor area (pre-SMA); Figure 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 specialisations of different dACC subregions remains to be clarified.

1.2. 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.

vmPFC 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 that 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; Figure 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 on decision-making, and show that focal vmPFC lesions result in specific decision-making impairments. For example, Damasio and

colleagues showed that, in a gambling-like task, patients with vmPFC lesions prefer riskier choices (Damasio, 1996; 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 performance 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.

vmPFC 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 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 (Figure 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 and colleagues recorded eye movements when human participants were choosing between two options (Lim et al., 2011). 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 (Figure 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 post-decisional, 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 non-preferred 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 and colleagues (2016) asked human participants to choose between a snack item from a preferred category and another snack item from a non-preferred category (Figure 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 non-preferred 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 vmPFC of humans versus monkeys

Although cytoarchitectonic and connectivity studies have demonstrated the homologous relationship between the vmPFC of human and non-human 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 10m, which is considered structurally homologous to 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 generalising findings across studies involving human and non-human animals. Such discrepancies are unlikely to have been simply due to experimental factors. All experiments were conducted using a similar MRI scanner. 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 reason for this discrepancy might be due to 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-dimension 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 and colleagues taught participants to recognise birds using this two-dimensional "bird space" (Constantinescu et al., 2016). 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.

1.3. Orbitofrontal cortex

The human OFC lies on the ventral surface of the prefrontal cortex 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 based on 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 non-human 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). Due 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 homolog of only the posterior human OFC (mainly posterior part of area 13), as it consists of an agranular cortex only. Thus, findings from the monkey OFC are likely generalisable 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, 2011b, a), as well as 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 individuals 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 activity 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 on vmPFC 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 (Figure 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 non-rewarded 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, 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 (Figure 4B).

The causal role of area 47/12 in flexible decision making has been further confirmed by a recent lesion study in macaques. Rudebeck and colleagues lesioned a lateral prefrontal region that

382 includes area 47/12 (as well as the neighbouring ventrolateral PFC), and found that these animals
383 performed poorly in reversing their choices after the change in reward contingency (Figure 4C)
384 (Rudebeck et al., 2017). Interestingly, they also tested macaques with lesions in other OFC regions,
385 including central areas 11/13 and medial area 14, and found that these animals' performance was
386 comparable to controls. In summary, current data suggest a division of labour in the primate OFC,
387 with central areas 11/13 involved in value representation and stimulus-reward associations, and
388 lateral areas 47/12 in flexible decision-making.

2. 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 dopamine-rich ventral tegmental area (VTA) of the midbrain to the ventral striatum, which comprises a critical part of the mesolimbic pathway (Figure 1) (Bjorklund and Dunnett, 2007). The ventral striatum is the major input structure to the basal ganglia, and comprises the nucleus accumbens (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 provides a solid foundation for generalising findings about striatal dopamine 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).

2.1 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 based on 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 where 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

414 according to differences in magnitude of possible rewards, rather than absolute differences in
415 expected value (Tobler et al., 2005). Such experiments provided important contributions to our
416 understanding of the role of dopamine neurons in reinforcement learning.

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

432

433 **2.2 Striatal dopamine**

434 Like the VTA, extensive data across multiple species demonstrate that the ventral striatum is
435 sensitive to reward prediction errors. The magnitude of prediction errors correlates specifically with
436 dopamine release from the rodent striatum, as recorded at high temporal resolution using fast-scan
437 cyclic voltammetry (Gan et al., 2010; Papageorgiou et al., 2016; Syed et al., 2016). Human fMRI
438 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 antagonised 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-characterised role in motor control? The prevailing framework considers that phasic bursts of striatal dopamine activity are central to encoding reward prediction errors, while slower fluctuations in tonic levels of striatal dopamine 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 emphasised by separate studies showing that the expected phasic striatal dopamine 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 emphasise a close mechanistic link between learning and motor initiation, and have led to recent attempts to more parsimoniously explain the role of the striatal dopamine in both reward-based processes and motor control (Berke, 2018).

2.3. Role of other neurotransmitter systems

Although the focus of this review is on dopaminergic signalling, we emphasise that dopamine 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 dopamine 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 dopamine 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 dopamine and other neurotransmitter system in value-based decision-making is beyond the scope of this review, but will be a critical area of investigation for future studies.

3. Dopaminergic connectivity of reward-sensitive PFC regions

To summarise, 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 (Figure 1). The major dopaminergic input to the PFC is via the mesocortical route – a direct projection from the VTA. The PFC in turn sends substantial efferent output to the ventral striatum. In human and non-human primates, this output is topographically-organised along a clear connectivity gradient (Figure 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 organised within distinct cortico-basal 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 emphasise 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 dopamine-based reinforcement learning. As discussed in Section 1, 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 decrease with negative prediction errors, and increase 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 dopamine may play a role in modulating OFC activity during the generation of flexible decisions.

How might dopamine convey the result of value computations across these corticostriatal loops? Dopamine 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 (~1-6Hz) dopamine release in the PFC maintains an extra-synaptic background concentration of dopamine, 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 dopamine was depleted locally within the marmoset OFC, elevated dopamine levels were observed at the striatum (Clarke et al., 2014). This suggests that striatal dopamine is sensitive to dopamine levels in the PFC, and that region-specific dopamine 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 Section 2.2) (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,

534 resulting in their choices becoming more stochastic. In contrast, tonic stimulation did not impair the
535 animals' ability to maintain the current rule – indeed, animals instead made perseverative errors
536 after a rule switch, indicating a failure to adapt. These findings demonstrate the dissociable roles of
537 phasic and tonic VTA-prelimbic dopamine input in maintaining and updating value representations.

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

548 Another outstanding question is how value-based representations in the prefrontal cortex
549 and basal ganglia interact computationally, and how dopamine might drive that interaction. A
550 current consensus is that the dopaminergic midbrain and striatum implement model-free
551 reinforcement learning, which is based on direct associations between stimulus and response. For
552 example, temporal difference models have been compelling in explaining the activity of
553 dopaminergic neurons in VTA (Schultz et al., 1997; Watabe-Uchida et al., 2017). In contrast, the PFC
554 is thought to implement a model-based type of reinforcement learning, which is based on internal
555 representations of task structure (Daw et al., 2005; Bromberg-Martin et al., 2010). Recently, some
556 have proposed to integrate both types of framework under a single theory of reward-based
557 decision-making, in order to more parsimoniously describe the computations underlying reward
558 valuation in the corticostriatal pathways (Wang et al., 2018). Others have proposed inter-region

559 models to describe the interactions between neurons in the frontal and parietal lobes during
560 working memory and decision-making (Murray et al., 2017). A promising path for future research
561 will be to refine such models to account for the interactions between these regions as a function of
562 dopamine release.

563

564

4. Studying prefrontal dopamine in humans

Despite the highly organised corticostriatal connectivity, surprisingly little is known about how mesocortical dopamine modulates decision-making signals in different subregions of the PFC, especially in humans. Studying the function of region-specific dopamine 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 dopamine 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 dopamine in a precise event or cognitive process. In non-human species, such investigations are often conducted using invasive methods, such as fast cyclic voltammetry, microdialysis, dopamine-selective lesion or, more recently, dopamine-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 dopamine release non-invasively, our understanding of the role of prefrontal dopamine in human decision-making relies partly on cross-species comparisons. Thus, as we have attempted to emphasise 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 dopamine 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 dopamine within different prefrontal areas in humans.

Reward, prefrontal cortex and dopamine

Another approach to elucidate the role of prefrontal dopamine in human decision-making has been through neurogenetic studies. The variability in dopamine function across individuals has been attributed to variability in a number of dopamine-specific genes. Dopamine levels in the prefrontal cortex are affected by polymorphisms in the catechol-O-methyltransferase (COMT) gene, which generates an enzyme involved in the degradation of dopamine. In contrast, dopamine levels in the striatum are affected by polymorphisms in the DRD2 gene (which generate the dopamine D2 receptor), and the DARPP-32 gene (which generates a protein for striatal synaptic plasticity). Frank and colleagues recruited healthy volunteers with different polymorphisms of these genes, and tested how genetic variability accounts for differences in decision-making (Frank et al., 2009; Doll et al., 2011; Doll et al., 2016). Their data revealed that 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 dopamine have dissociable roles in decision making, and more broadly demonstrate how genetic variability may be a useful proxy to studying regional specialisations of human dopamine function.

Further specificity can be achieved by combining such genetic approaches with neuroimaging techniques. Gao and colleagues performed a gambling task on participants with different COMT genotypes, while recording their resting-state neural activity using fMRI (Gao et al., 2016). The stimuli either emphasised 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 emphasised. 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 dopamine in the human PFC.

Reward, prefrontal cortex and dopamine

Another potentially useful approach is to image patients with dopaminergic dysfunction, such as those with idiopathic Parkinson's disease (PD). Patients with disorders of dopamine 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 which is ameliorable with dopamine replacement (Chong et al., 2015; Chong and Husain, 2016; Muhammed et al., 2016). In a two-stage reinforcement learning experiment, patients with PD underwent fMRI scanning when they were ON or OFF dopamine 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 vs 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 dopamine, presumably via the mesocortical route, in deciding between novel associations.

5. Summary and concluding remarks

The prefrontal cortex, together with its bidirectional connections with the basal ganglia, play important roles in reward-based decision-making. These areas are connected in a highly organised, 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, offer 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 dopamine contributes to the broader neural circuitry during reward-based decision-making.

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. Annual Review of Neuroscience 9:357-381.

Amemori K, Graybiel AM (2012) Localized microstimulation of primate pregenual cingulate cortex induces negative decision making. Nature Neuroscience 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. Nature Neuroscience 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 in neurosciences 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. Cognitive, Affective, and Behavioral Neuroscience 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). The Journal of comparative neurology 514:11-29.

Carmichael ST, Price JL (1994) Architectonic subdivision of the orbital and medial prefrontal cortex in the macaque monkey. The Journal of comparative neurology 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-J, Bonnelle V, Husain M (2016) Quantifying motivation with effort-based decision-making paradigms in health and disease. Progress in Brain Research 229:71-100.

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

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

694 Chong TT, Bonnelle V, Veromann KR, Juurmaa J, Taba P, Plant O, Husain M (2018) Dissociation of
695 reward and effort sensitivity in methcathinone-induced Parkinsonism. J Neuropsychol
696 12:291-297.
697 Chong TT, Bonnelle V, Manohar S, Veromann KR, Muhammed K, Tofaris GK, Hu M, Husain M (2015)
698 Dopamine enhances willingness to exert effort for reward in Parkinson's disease. Cortex
699 69:40-46.
700 Churchland AK, Kiani R, Shadlen MN (2008) Decision-making with multiple alternatives. Nat Neurosci
701 11:693-702.
702 Clarke HF, Cardinal RN, Rygula R, Hong YT, Fryer TD, Sawiak SJ, Ferrari V, Cockcroft G, Aigbirhio FI,
703 Robbins TW, Roberts AC (2014) Orbitofrontal dopamine depletion upregulates caudate
704 dopamine and alters behavior via changes in reinforcement sensitivity. J Neurosci 34:7663-
705 7676.
706 Cohen JY, Haesler S, Vong L, Lowell BB, Uchida N (2012) Neuron-type-specific signals for reward and
707 punishment in the ventral tegmental area. Nature 482:85-88.
708 Cole MW, Yeung N, Freiwald WA, Botvinick MM (2009) Cingulate cortex: diverging data from
709 humans and monkeys. Trends in Neurosciences 32:566-574.
710 Constantinescu AO, O'Reilly JX, Behrens TEJ (2016) Organizing conceptual knowledge in humans with
711 a gridlike code. Science 352:1464-1468.
712 Croxson PL, Walton ME, O'Reilly JX, Behrens TE, Rushworth MF (2009) Effort-based cost-benefit
713 valuation and the human brain. Journal of Neuroscience 29:4531-4541.
714 Cui Q, Li Q, Geng H, Chen L, Ip NY, Ke Y, Yung WH (2018) Dopamine receptors mediate strategy
715 abandoning via modulation of a specific prelimbic cortex-nucleus accumbens pathway in
716 mice. Proc Natl Acad Sci U S A 115:E4890-E4899.
717 Damasio AR (1996) The somatic marker hypothesis and the possible functions of the prefrontal
718 cortex. Philos Trans R Soc Lond B Biol Sci 351:1413-1420.
719 Daw ND, Niv Y, Dayan P (2005) Uncertainty-based competition between prefrontal and dorsolateral
720 striatal systems for behavioral control. Nat Neurosci 8:1704-1711.
721 Debener S, Ullsperger M, Siegel M, Fiehler K, von Cramon DY, Engel AK (2005) Trial-by-trial coupling
722 of concurrent electroencephalogram and functional magnetic resonance imaging identifies
723 the dynamics of performance monitoring. Journal of neuroscience 25:11730-11737.
724 Doll BB, Hutchison KE, Frank MJ (2011) Dopaminergic genes predict individual differences in
725 susceptibility to confirmation bias. J Neurosci 31:6188-6198.
726 Doll BB, Bath KG, Daw ND, Frank MJ (2016) Variability in Dopamine Genes Dissociates Model-Based
727 and Model-Free Reinforcement Learning. J Neurosci 36:1211-1222.
728 Duzel E, Bunzeck N, Guitart-Masip M, Wittmann B, Schott BH, Tobler PN (2009) Functional imaging
729 of the human dopaminergic midbrain. Trends Neurosci 32:321-328.
730 Ellwood IT, Patel T, Wadia V, Lee AT, Liptak AT, Bender KJ, Sohal VS (2017) Tonic or Phasic
731 Stimulation of Dopaminergic Projections to Prefrontal Cortex Causes Mice to Maintain or
732 Deviate from Previously Learned Behavioral Strategies. J Neurosci 37:8315-8329.
733 Eshel N, Tian J, Bukwich M, Uchida N (2016) Dopamine neurons share common response function for
734 reward prediction error. Nat Neurosci 19:479-486.
735 Eshel N, Bukwich M, Rao V, Hemmelder V, Tian J, Uchida N (2015) Arithmetic and local circuitry
736 underlying dopamine prediction errors. Nature 525:243-246.
737 Fellows LK (2011) Orbitofrontal contributions to value-based decision making: evidence from
738 humans with frontal lobe damage. Ann N Y Acad Sci 1239:51-58.
739 Fellows LK, Farah MJ (2005) Different underlying impairments in decision-making following
740 ventromedial and dorsolateral frontal lobe damage in humans. Cereb Cortex 15:58-63.
741 Fellows LK, Farah MJ (2007) The role of ventromedial prefrontal cortex in decision making: judgment
742 under uncertainty or judgment per se? Cereb Cortex 17:2669-2674.
743 Frank MJ, Doll BB, Oas-Terpstra J, Moreno F (2009) Prefrontal and striatal dopaminergic genes
744 predict individual differences in exploration and exploitation. Nat Neurosci 12:1062-1068.

745 Gan JO, Walton ME, Phillips PE (2010) Dissociable cost and benefit encoding of future rewards by
746 mesolimbic dopamine. Nat Neurosci 13:25-27.
747 Gao X, Gong P, Liu J, Hu J, Li Y, Yu H, Gong X, Xiang Y, Jiang C, Zhou X (2016) COMT Val158Met
748 polymorphism influences the susceptibility to framing in decision-making: OFC-amygdala
749 functional connectivity as a mediator. Hum Brain Mapp 37:1880-1892.
750 Ghahremani DG, Monterosso J, Jentsch JD, Bilder RM, Poldrack RA (2010) Neural components
751 underlying behavioral flexibility in human reversal learning. Cereb Cortex 20:1843-1852.
752 Grabenhorst F, Rolls ET (2011) Value, pleasure and choice in the ventral prefrontal cortex. Trends
753 Cogn Sci 15:56-67.
754 Haber SN (2011) Neuroanatomy of Reward: A View from the Ventral Striatum. In: Neurobiology of
755 Sensation and Reward (Gottfried JA, ed). Boca Raton (FL).
756 Haber SN, Knutson B (2010) The reward circuit: linking primate anatomy and human imaging.
757 Neuropsychopharmacology 35:4-26.
758 Haber SN, Behrens TE (2014) The neural network underlying incentive-based learning: implications
759 for interpreting circuit disruptions in psychiatric disorders. Neuron 83:1019-1039.
760 Hafting T, Fyhn M, Molden S, Moser MB, Moser EI (2005) Microstructure of a spatial map in the
761 entorhinal cortex. Nature 436:801-806.
762 Hamid AA, Pettibone JR, Mabrouk OS, Hetrick VL, Schmidt R, Vander Weele CM, Kennedy RT,
763 Aragona BJ, Berke JD (2016) Mesolimbic dopamine signals the value of work. Nat Neurosci
764 19:117-126.
765 Hampshire A, Chaudhry AM, Owen AM, Roberts AC (2012) Dissociable roles for lateral orbitofrontal
766 cortex and lateral prefrontal cortex during preference driven reversal learning. Neuroimage
767 59:4102-4112.
768 Hare TA, Schultz W, Camerer CF, O'Doherty JP, Rangel A (2011) Transformation of stimulus value
769 signals into motor commands during simple choice. Proceedings of the National Academy of
770 Sciences of the USA 108:18120-18125.
771 Hayden BY, Pearson JM, Platt ML (2011) Neuronal basis of sequential foraging decisions in a patchy
772 environment. Nat Neurosci 14:933-939.
773 He SQ, Dum RP, Strick PL (1995) Topographic organization of corticospinal projections from the
774 frontal lobe: motor areas on the medial surface of the hemisphere. J Neurosci 15:3284-3306.
775 Heilbronner SR, Hayden BY (2016) Dorsal Anterior Cingulate Cortex: A Bottom-Up View. Annu Rev
776 Neurosci 39:149-170.
777 Heilbronner SR, Rodriguez-Romaguera J, Quirk GJ, Groenewegen HJ, Haber SN (2016) Circuit-Based
778 Corticostriatal Homologies Between Rat and Primate. Biol Psychiatry 80:509-521.
779 Heimer L, De Olmos JS, Alheid GF, Person J, Sakamoto N, Shinoda K, Marksteiner J, Switzer RC (1999)
780 The human basal forebrain. Part II. The primate nervous system. In: Part III. Handbook of
781 chemical neuroanatomy (Bloom FE, Bjorkland A, Hokfelt T, eds), pp 57-226. Amsterdam:
782 Elsevier.
783 Henri-Bhargava A, Simioni A, Fellows LK (2012) Ventromedial frontal lobe damage disrupts the
784 accuracy, but not the speed, of value-based preference judgments. Neuropsychologia
785 50:1536-1542.
786 Holroyd CB, Coles MG (2002a) The neural basis of human error processing: reinforcement learning,
787 dopamine, and the error-related negativity. Psychol Rev 109:679-709.
788 Holroyd CB, Coles MG (2002b) The neural basis of human error processing: reinforcement learning,
789 dopamine, and the error-related negativity. Psychological Review 109:679-709.
790 Holroyd CB, Yeung N (2012) Motivation of extended behaviors by anterior cingulate cortex. Trends in
791 Cognitive Sciences 16:122-128.
792 Hornak J, O'Doherty J, Bramham J, Rolls ET, Morris RG, Bullock PR, Polkey CE (2004) Reward-related
793 reversal learning after surgical excisions in orbito-frontal or dorsolateral prefrontal cortex in
794 humans. J Cogn Neurosci 16:463-478.

795 Hosking JG, Cocker PJ, Winstanley CA (2014) Dissociable contributions of anterior cingulate cortex
796 and basolateral amygdala on a rodent cost-benefit decision-making task of cognitive effort
797 Neuropsychopharmacology 39:1558-1567.
798 Howe MW, Dombeck DA (2016) Rapid signalling in distinct dopaminergic axons during locomotion
799 and reward. Nature 535:505-510.
800 Hyman JM, Holroyd CB, Seamans JK (2017) A Novel Neural Prediction Error Found in Anterior
801 Cingulate Cortex Ensembles. Neuron 95:447-456 e443.
802 Izquierdo A, Suda RK, Murray EA (2004) Bilateral orbital prefrontal cortex lesions in rhesus monkeys
803 disrupt choices guided by both reward value and reward contingency. J Neurosci 24:7540-
804 7548.
805 Izuma K, Saito DN, Sadato N (2008) Processing of social and monetary rewards in the human
806 striatum. Neuron 58:284-294.
807 Jocham G, Furlong PM, Kroger IL, Kahn MC, Hunt LT, Behrens TE (2014) Dissociable contributions of
808 ventromedial prefrontal and posterior parietal cortex to value-guided choice. Neuroimage
809 100:498-506.
810 Kable JW, Glimcher PW (2007) The neural correlates of subjective value during intertemporal choice.
811 Nat Neurosci 10:1625-1633.
812 Kazama A, Bachevalier J (2009) Selective aspiration or neurotoxic lesions of orbital frontal areas 11
813 and 13 spared monkeys' performance on the object discrimination reversal task. J Neurosci
814 29:2794-2804.
815 Kennerley SW, Behrens TE, Wallis JD (2011) Double dissociation of value computations in
816 orbitofrontal and anterior cingulate neurons. Nat Neurosci 14:1581-1589.
817 Kennerley SW, Walton ME, Behrens TE, Buckley MJ, Rushworth MF (2006) Optimal decision making
818 and the anterior cingulate cortex. Nat Neurosci 9:940-947.
819 Kolling N, Behrens TE, Mars RB, Rushworth MF (2012) Neural mechanisms of foraging. Science
820 336:95-98.
821 Kolling N, Wittmann MK, Behrens TE, Boorman ED, Mars RB, Rushworth MF (2016) Value, search,
822 persistence and model updating in anterior cingulate cortex. Nat Neurosci 19:1280-1285.
823 Kringelbach ML, Rolls ET (2003) Neural correlates of rapid reversal learning in a simple model of
824 human social interaction. Neuroimage 20:1371-1383.
825 Lammel S, Hetzel A, Hackel O, Jones I, Liss B, Roeper J (2008) Unique properties of mesoprefrontal
826 neurons within a dual mesocorticolimbic dopamine system. Neuron 57:760-773.
827 Lau B, Glimcher PW (2008) Value representations in the primate striatum during matching behavior.
828 Neuron 58:451-463.
829 Lebreton M, Jorge S, Michel V, Thirion B, Pessiglione M (2009) An automatic valuation system in the
830 human brain: evidence from functional neuroimaging. Neuron 64:431-439.
831 Levy DJ, Glimcher PW (2012) The root of all value: a neural common currency for choice. Curr Opin
832 Neurobiol 22:1027-1038.
833 Lim SL, O'Doherty JP, Rangel A (2011) The decision value computations in the vmPFC and striatum
834 use a relative value code that is guided by visual attention. J Neurosci 31:13214-13223.
835 Lopez-Persem A, Domenech P, Pessiglione M (2016) How prior preferences determine decision-
836 making frames and biases in the human brain. eLife 5.
837 Luppino G, Matelli M, Camarda RM, Gallese V, Rizzolatti G (1991) Multiple representations of body
838 movements in mesial area 6 and the adjacent cingulate cortex: an intracortical
839 microstimulation study in the macaque monkey. The Journal of comparative neurology
840 311:463-482.
841 Matsumoto M, Matsumoto K, Abe H, Tanaka K (2007) Medial prefrontal cell activity signaling
842 prediction errors of action values. Nat Neurosci 10:647-656.
843 McClure SM, Berns GS, Montague PR (2003) Temporal prediction errors in a passive learning task
844 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. Progress in Neurobiology 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, Bogdanovic M, Hu M, Husain M (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.
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 U S A 112:E2695-2704.
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 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 (2004) Reward representations and reward-related learning in the human brain: insights from neuroimaging. Curr Opin Neurobiol 14:769-776.
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. Nature communications 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 Edition. Oxford, United Kingdom: 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. Annual Review of Neuroscience 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. Cerebral 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, 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.

Rudebeck PH, Walton ME, Smyth AN, Bannerman DM, Rushworth MF (2006) Separate neural pathways process different decision costs. Nature Neuroscience 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.

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. Learning and Memory 12:334-342.

Schweimer J, Saft S, Hauber W, 2005. . Behav. Neurosci. 119, 1687 1692. (2005) Involvement of catecholamine neurotransmission in the rat anterior cingulate in effortrelated decision-making. Behavioral Neuroscience 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. Journal of Comparative Neurology 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. Nature Reviews Neuroscience 12:154.

Shenhav A, Straccia MA, Cohen JD, Botvinick MM (2014) Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. Nature Neuroscience 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.

946 Stuss DT, Alexander MP, Shallice T, Picton TW, Binns MA, Macdonald R, Borowiec A, Katz DI (2005)
947 Multiple frontal systems controlling response speed. Neuropsychologia 43:396-417.
948 Syed EC, Grima LL, Magill PJ, Bogacz R, Brown P, Walton ME (2016) Action initiation shapes
949 mesolimbic dopamine encoding of future rewards. Nat Neurosci 19:34-36.
950 Thorpe SJ, Rolls ET, Maddison S (1983) The orbitofrontal cortex: neuronal activity in the behaving
951 monkey. Exp Brain Res 49:93-115.
952 Tian J, Huang R, Cohen JY, Osakada F, Kobak D, Machens CK, Callaway EM, Uchida N, Watabe-Uchida
953 M (2016) Distributed and Mixed Information in Monosynaptic Inputs to Dopamine Neurons.
954 Neuron 91:1374-1389.
955 Tobler PN, Fiorillo CD, Schultz W (2005) Adaptive coding of reward value by dopamine neurons.
956 Science 307:1642-1645.
957 Tobler PN, O'Doherty J P, Dolan RJ, Schultz W (2006) Human neural learning depends on reward
958 prediction errors in the blocking paradigm. J Neurophysiol 95:301-310.
959 Tremblay L, Schultz W (1999) Relative reward preference in primate orbitofrontal cortex. Nature
960 398:704-708.
961 Tsutsui K, Grabenhorst F, Kobayashi S, Schultz W (2016) A dynamic code for economic object
962 valuation in prefrontal cortex neurons. Nature communications 7:12554.
963 Vaidya AR, Fellows LK (2015) Ventromedial Frontal Cortex Is Critical for Guiding Attention to Reward-
964 Predictive Visual Features in Humans. J Neurosci 35:12813-12823.
965 Vaidya AR, Fellows LK (2016) Necessary Contributions of Human Frontal Lobe Subregions to Reward
966 Learning in a Dynamic, Multidimensional Environment. J Neurosci 36:9843-9858.
967 Varazzani C, San-Galli A, Gilardeau S, Bouret S (2015) Noradrenaline and dopamine neurons in the
968 reward/effort trade-off: a direct electrophysiological comparison in behaving monkeys. J
969 Neurosci 35:7866-7877.
970 Vogt BA (2016) Midcingulate cortex: Structure, connections, homologies, functions and diseases.
971 Journal of Chemical Neuroanatomy 74:28-46.
972 Volkow ND, Wise RA, Baler R (2017) The dopamine motive system: implications for drug and food
973 addiction. Nat Rev Neurosci 18:741-752.
974 Walker SC, Robbins TW, Roberts AC (2009) Differential contributions of dopamine and serotonin to
975 orbitofrontal cortex function in the marmoset. Cereb Cortex 19:889-898.
976 Wallis JD (2007) Orbitofrontal cortex and its contribution to decision-making. Annu Rev Neurosci
977 30:31-56.
978 Wallis JD (2012) Cross-species studies of orbitofrontal cortex and value-based decision-making. Nat
979 Neurosci 15:13-19.
980 Walton ME, Bannerman DM, Rushworth MFS (2002) The role of rat medial frontal cortex in effort-
981 based decision-making. Journal of Neuroscience 22:10996-11003.
982 Walton ME, Bannerman DM, Alterescu K, Rushworth MFS (2003) Functional specialization within
983 medial frontal cortex of the anterior cingulate for evaluating effort-related decisions. Journal
984 of Neuroscience 23:6475-6479.
985 Walton ME, Groves J, Jennings KA, Croxson PL, Sharp T, Rushworth MFS, Bannerman DM (2009)
986 Comparing the role of the anterior cingulate cortex and 6-hydroxydopamine nucleus
987 accumbens lesions on operant effort-based decision-making. European Journal of
988 Neuroscience 29:1678-1691.
989 Wang JX, Kurth-Nelson Z, Kumaran D, Tirumala D, Soyer H, Leibo JZ, Hassabis D, Botvinick M (2018)
990 Prefrontal cortex as a meta-reinforcement learning system. Nat Neurosci 21:860-868.
991 Watabe-Uchida M, Eshel N, Uchida N (2017) Neural Circuitry of Reward Prediction Error. Annu Rev
992 Neurosci 40:373-394.
993 Wise SP (2008) Forward frontal fields: phylogeny and fundamental function. Trends Neurosci 31:599-
994 608.

995 Wittmann MK, Kolling N, Akaishi R, Chau BK, Brown JW, Nelissen N, Rushworth MF (2016) Predictive
996 decision making driven by multiple time-linked reward representations in the anterior
997 cingulate cortex. Nature communications 7:12327.
998 Zink CF, Tong Y, Chen Q, Bassett DS, Stein JL, Meyer-Lindenberg A (2008) Know your place: neural
999 processing of social hierarchy in humans. Neuron 58:273-283.

1000

1001

Figure legend

Figure 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-organised manner. Together, this network of corticostriatal loops comprise the core of a circuit that is central to reward-based decision-making. dACC dorsal anterior cingulate cortex; DPFC dorsal prefrontal cortex; OFC orbitofrontal cortex; S shell of nucleus accumbens; SN/VTA substantia nigra/ventral tegmental area; vmPFC ventromedial prefrontal cortex. Adapted from (Haber and Knutson, 2010).

Figure 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' vs 'searching' were observed. **(C)** The pre-SMA encoded the difficulty of the trial. Adapted from (Kolling et al., 2016).

Figure 3. Value signals in 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 non-default option (adapted from (Lopez-Persem et al., 2016)).

Figure 4. The role of 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 non-rewarded 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 blue, red and green dots indicate that Option A, B and C was the most rewarding option on a given trial. lOFC lateral orbitofrontal cortex; vlPFC ventrolateral PFC (adapted from (Rudebeck et al., 2017)).