Neuron

CellPress

Review

Medial and orbital frontal cortex in decision-making and flexible behavior

Miriam C. Klein-Flügge,[1,2,3,4,*] Alessandro Bongioanni,[1] and Matthew F.S. Rushworth[1,2]
[1]Wellcome Centre for Integrative Neuroimaging (WIN), Department of Experimental Psychology, University of Oxford, Tinsley Building, Mansfield Road, Oxford OX1 3TA, UK
[2]Wellcome Centre for Integrative Neuroimaging (WIN), Centre for Functional MRI of the Brain (FMRIB), University of Oxford, Nuffield Department of Clinical Neurosciences, Level 6, West Wing, John Radcliffe Hospital, Oxford OX3 9DU, UK
[3]Department of Psychiatry, University of Oxford, Warneford Lane, Headington, Oxford OX3 7JX, UK
[4]Twitter: @MKFlugge
*Correspondence: miriam.klein-flugge@psy.ox.ac.uk
https://doi.org/10.1016/j.neuron.2022.05.022

SUMMARY

The medial frontal cortex and adjacent orbitofrontal cortex have been the focus of investigations of decision-making, behavioral flexibility, and social behavior. We review studies conducted in humans, macaques, and rodents and argue that several regions with different functional roles can be identified in the dorsal anterior cingulate cortex, perigenual anterior cingulate cortex, anterior medial frontal cortex, ventromedial prefrontal cortex, and medial and lateral parts of the orbitofrontal cortex. There is increasing evidence that the manner in which these areas represent the value of the environment and specific choices is different from subcortical brain regions and more complex than previously thought. Although activity in some regions reflects distributions of reward and opportunities across the environment, in other cases, activity reflects the structural relationships between features of the environment that animals can use to infer what decision to take even if they have not encountered identical opportunities in the past.

INTRODUCTION

It is well-established that the frontal cortex guides decision-making and flexible behavior. This conviction is based on more than half a century of investigations into how animals and people adapt their behavior so that it is appropriate to changing circumstances. For example, the most recent investigations of the orbitofrontal cortex (OFC), employing the very latest techniques (Banerjee et al., 2020), assess aspects of behavior—the ability to switch and change which choice is made—using a behavior reversal task with elements that would have been familiar to researchers investigating the OFC more than fifty years earlier (Mishkin et al., 1969). Similarly, the idea that the OFC and brain areas on the medial surface of the frontal cortex, such as the anterior cingulate cortex (ACC) and ventromedial prefrontal cortex (vmPFC), guide decision-making is bolstered by a series of investigations that show that their activity reflects the value of choices, the process of decision-making, and the value of the course of action pursued (Cai and Padoa-Schioppa, 2014; Kable and Glimcher, 2009; Rudebeck and Murray, 2014; Rushworth et al., 2011; Soltani and Koechlin, 2022; Wallis, 2011).

It is, however, clear that the frontal cortex is far from being the only brain region concerned with behavioral flexibility and reward-guided decision-making. There are other very different types of brain systems that control decision-making and behavioral flexibility, for example, the striatum, dopaminergic (DA) midbrain, and serotonergic (5HT) brainstem, and identifying the special additional contribution made by areas such as the OFC and ACC is not always straightforward. The ACC and parts of the OFC are found in many mammals, but they are especially extensive in primates. On the other hand, animals lacking the OFC and ACC still exhibit changing patterns of decision-making and behavioral flexibility. For example, larval zebrafish do not possess frontal cortical areas and yet they switch between exploiting opportunities for predation and exploration for new sources of food. In humans, the balancing of exploitation against exploration has been associated with the frontal cortex (Badre et al., 2012; Daw et al., 2006; Trudel et al., 2021; Zajkowski et al., 2017). In zebrafish, however, this pattern of reward-guided decision-making depends on the dorsal raphe nucleus (DRN) (Marques et al., 2020). In rodents and even in primates, DRN activity also tracks aspects of the reward environment and indicates when behavioral change might be necessary (Grossman et al., 2022; Hayashi et al., 2015; Khalighinejad et al., 2022; Wittmann et al., 2020). Neurons in other subcortical nuclei, for example, midbrain DA neurons, also have activity that reflects the value of potential choices and the process of decision-making (Wang et al., 2021; Yun et al., 2020). If subcortical systems carry such signals, can we specify how frontal cortical contributions differ? In order to do this, might we need to distinguish between the types of behavioral flexibility observed in zebrafish and humans?

Not only is it important to distinguish frontal cortical contributions from subcortical ones, but some recent results suggest a need to rethink the precise nature of the frontal cortical
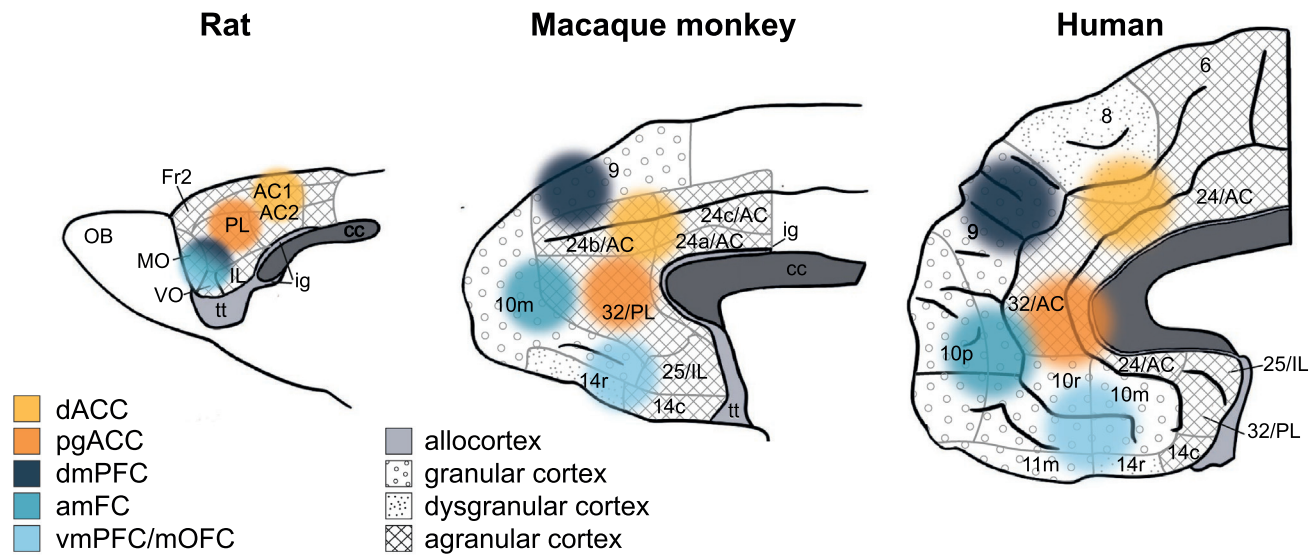
Figure 1. Medial and orbital frontal cortex in rodents, macaques, and humans
Five functional regions, dorsal anterior cingulate cortex (dACC), perigenual anterior cingulate cortex (pgACC), dorsomedial prefrontal cortex (dmPFC), anterior medial frontal cortex (amFC), and ventromedial prefrontal cortex/medial orbitofrontal cortex (vmPFC/mOFC) are shown in relation to cytoarchitectonic maps of rat (left), macaque (center), and human (right) medial and orbital frontal cortex (based on Wise, 2008). All five regions are identifiable in humans and macaques. The color scheme indicates that the orbital region of rodents has some functional features shared with primate dmPFC, amFC, and vmPFC/mOFC but that it does not correspond in a simple way to any of them. Thus, while the extent to which regions are homologous across humans and other primates, such as macaques, is relatively clear, correspondences between primates and rodents are more contentious and unlikely to be one-to-one in nature. This is important to bear in mind when evaluating evidence from different species about the regions' functional contributions.

contributions. For example, when lesions are made to the primate OFC in such a way that adjacent white matter is spared, then behavior reversal is actually uncompromised (Rudebeck et al., 2013). Not only might the frontal cortex not have the role that we thought it had in behavioral flexibility but, in addition, other scientists have argued that the OFC and other frontal areas lack the representations of value that have, for the last two decades, been thought to guide decision-making (Hayden and Niv, 2021). We may then need to conceive the specific contributions of the frontal cortex in a more differentiated way.

In this review we summarize recent evidence regarding the nature of the representations found across a set of frontal cortical regions (Figures 1 and 2). First, we discuss how dorsal ACC (dACC) represents the distribution of opportunities across the environment, computes their value at multiple timescales, decides whether to engage with a present opportunity or continue exploring, and influences information-seeking behavior. Second, we turn to the anterior medial frontal cortex (amFC), describing its role in representing the structure of the environment, even when this is not immediately relevant for reward-guided decisions. We also discuss the relationship between the rodent and primate OFC. Next, we review evidence showing that vmPFC translates values into choices. We consider work on perigenual ACC (pgACC), showing a role in integrating costs and benefits to drive behavior, and conclude by discussing how dorsomedial PFC (dmPFC) organizes relationships in social contexts. Throughout the review, we also consider whether and, if so, how such representations might differ from and complement those present in subcortical brain systems previously associated with reward-guided choices and behavioral change. We

compare the frontal cortical areas described with components of the subcortical circuit comprising DA and 5HT nuclei under the control of a pathway running through the striatum, pallidum, and habenula that controls reward-guided behavior in many vertebrates (Figure 2A). In mammals, this circuit is, itself, partly under the control of the frontal cortex as well as being influenced by dopamine and serotonin.

Anterior cingulate cortex and opportunities in the environment

An animal can manage well in many cases if it can represent the value of its current environment (whether and to what degree it is rewarding), whether the actions it takes will make its experience of the environment better or worse (positive and negative prediction errors), and its uncertainty about the estimates it is making about its environment. Such representations can guide an animal with limited environmental knowledge to find and stay in good foraging areas and to retreat from those that are not so good. For example, it is possible that a fish with such representations might manage well even if it were living in murky water that precludes remote sensing of the distribution of opportunities in the environment.

Subcortical systems provide important information in such environments (Figure 2A). For example, in addition to the DRN-centered 5HT system described above (Marques et al., 2020), DA neurons in monkeys encode prediction errors and report whether the environment is getting better or worse (Hart et al., 2014; Schultz, 2013; Schultz et al., 1997). As each opportunity is encountered, changes in DA neuron activity reflect whether or not the monkey will pursue it or wait for a better opportunity

CellPress
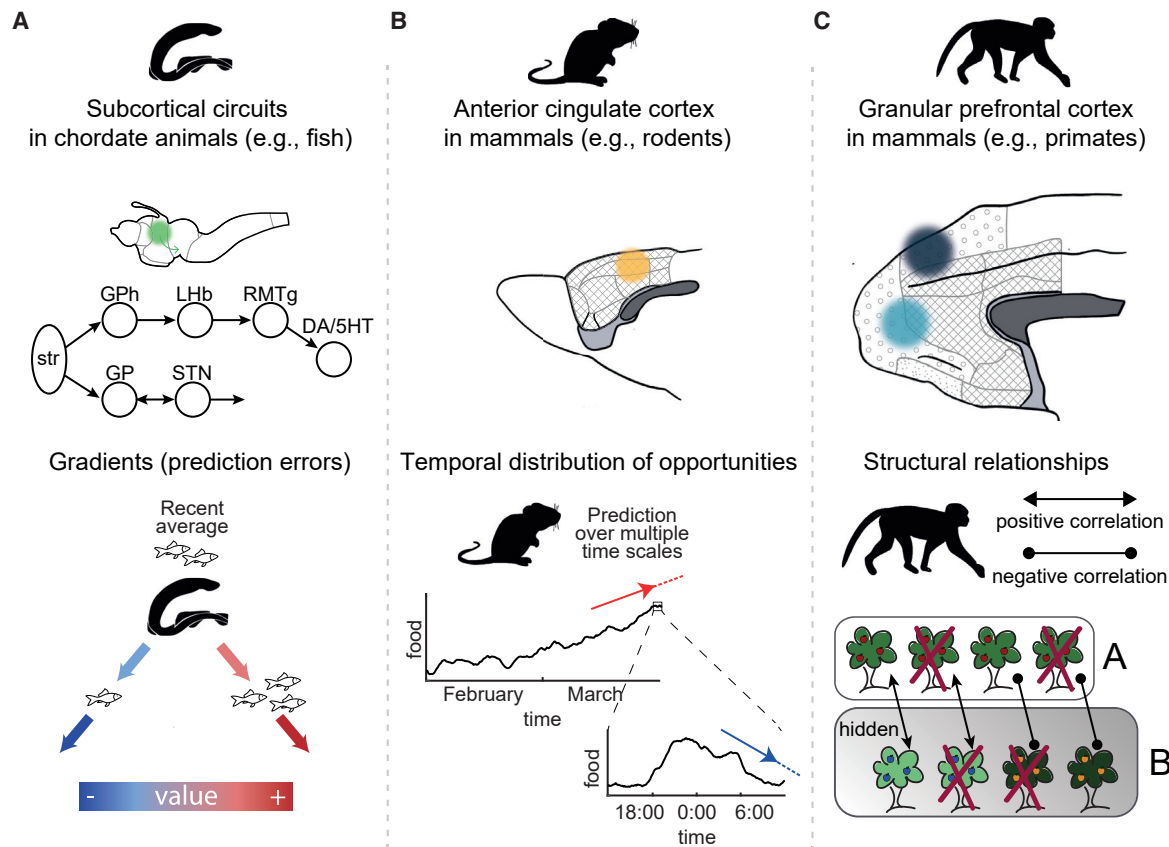
Figure 2. Multiple systems for representing the value of the environment in the vertebrate brain

(A) A subcortical circuit for identifying rewarding environments uses an estimate of the value of an animal's environment—how good it is currently—and detects prediction errors—occasions when the environment turns out to be better or worse than previously estimated. This circuit is present in rodents and primates but also in many other chordate animals (Freudenmacher et al., 2020; Hong and Hikosaka, 2008; Matsumoto and Hikosaka, 2007; Stephenson-Jones et al., 2013, 2016). Some of its component elements are identifiable even in cyclostomes—such as the lamprey—which diverged from other chordates 550 mya. This includes inhibitory GABA-ergic mediated control (rostromedial tegmental nucleus, RMTg) of dopaminergic and serotonergic regions (DA/5HT) by the lateral habenula (LHb), which is in turn innervated by a habenula-projecting pallidal region (GPh) and the striatum (str). A second pathway runs via the dorsal pallidum/globus pallidus (GP) to brainstem motor areas such as the subthalamic nucleus (STN) (adapted from Stephenson-Jones et al., 2013). Neuromodulatory systems such as the dopaminergic (DA) midbrain nuclei—the ventral tegmental area (VTA) and substantia nigra pars compacta (SNc)—and serotonergic (5HT) dorsal raphe nucleus (DRN) allow mammals and birds to follow value gradients (blue-to-red gradients indicate low-to-high reward) so that they find and stay in the best locations in an environment.

(B) Cingulate areas such as dACC (area 24 in primates and Cg1 and Cg2 in rodents) are identifiable in most mammals, including monotremes and marsupials (Ashwell et al., 2008; Mayner, 1989; Suárez et al., 2018), suggesting an origin over 200 mya. In addition to the current value of the environment and prediction errors, dACC enables mammals to represent the distribution of opportunities across the environment and over time across multiple scales. In this example, an animal might learn about food availability over an intermediate timescale—time within a day—and a long timescale—time over weeks—allowing it to make predictions into the future.
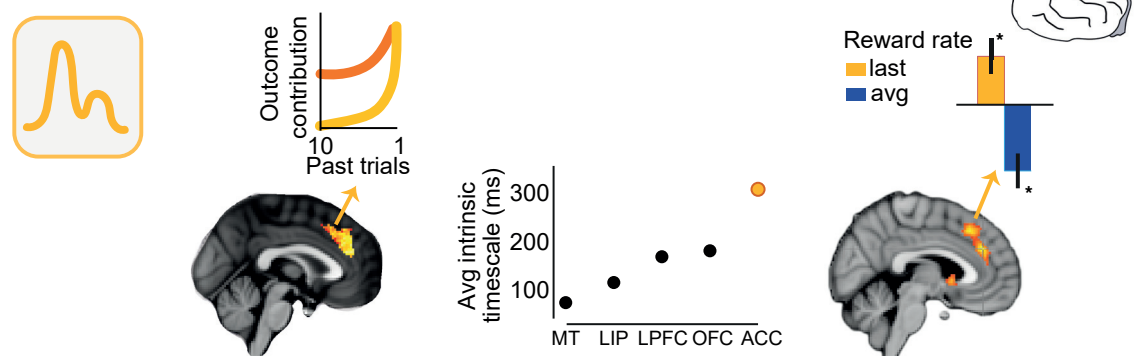
(C) Rodents and primates construct representations of the relationships between elements and features of the environment. For example, they might learn that if reward is available at one location (here indicated by one type of tree), it may be present or absent from another, and vice versa. In the four examples, illustrated by a pair of trees, a monkey has learned two types of relationships between reward in two locations. In the first two cases, a positive correlation implies that when reward is found in location A (far left), another reward is likely to be available in the location B (hidden from view), but when it is absent in A (second from left) it will also be absent in B. In the third and fourth cases (second from right and far right), the monkey has learned a negative correlation between reward in the two locations. This is the type of situation that is being investigated in reversal tasks. Such cognitive maps of environmental contingencies depend on multiple brain systems, such as MTL areas, and they are found in rodents. However, they may be an especially prominent feature of granular prefrontal areas in primates.

(Yun et al., 2020). Subcortical circuits that allow an animal to follow value gradients and find and stay in the best possible locations in their environment, or avoid dangerous ones, are present not only in rodents and monkeys but also many chordate animals such as the lamprey (Figure 2A) and other fish (Agetsuma et al., 2010; Amo et al., 2014).
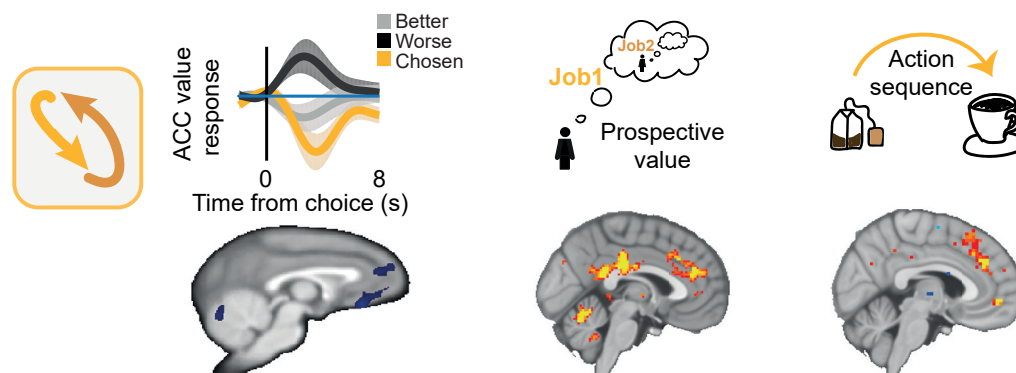
It has been pointed out, however, that not all environments are the same. For example, the terrestrial environment is a very different one, which animals with distance receptors, such as mammals, are able to survey from afar (Hunt et al., 2021; MacIver et al., 2017; Mugan and MacIver, 2020). If an animal can survey its environment, then, in addition to representing its average reward value and reward prediction errors, it becomes possible to represent the distribution of opportunities across the environment and the changes and sequences of behavior by which the opportunities might be pursued (Figure 2B). Converging experimental evidence shows that dACC activity (1) encodes the distribution of opportunities across time as well as space (Figure 3A),
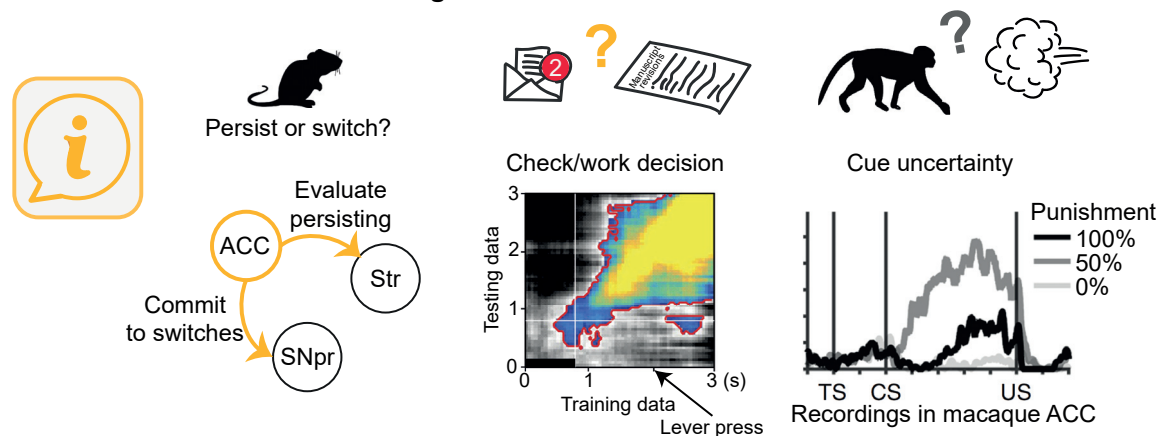
Figure 3. Exploring and navigating the distribution of opportunities in the environment via dACC

(A) (Left) Activity in dACC reflects a person's experience of success on a simple task over multiple timescales simultaneously. Activity in the lighter, yellow regions in dACC is dominated by the most recent experience, while activity in the orange regions reflects experience over a more extended timescale (adapted from Meder et al., 2017). (Center) Neurons in ACC have the longest intrinsic timescale compared with other regions (e.g., middle temporal [MT], lateral intraparietal [LIP], or lateral prefrontal or orbitofrontal cortex [LPFC/OFC] [adapted from Murray et al., 2014]). (Right) It is possible to work out whether one is on an upward or downward reward trajectory and thus make predictions into the future by comparing reward rates experienced over the short term or longer term. Activity in human dACC reflects this comparison: recent reward experience is encoded with a positive sign (shown in yellow) and reward experienced over a longer timescale is encoded with a negative sign (shown in blue) (adapted from Wittmann et al., 2016a; error bars denote SEM across participants).

Neuron
Review

CellPress

(2) assesses the value of disengaging from the present course of action (Figure 3B), and (3) regulates switching between periods of exploiting such knowledge and seeking more information (Figure 3C).

dACC and the distribution of experiences over time

One way in which resources in an environment might be unevenly distributed is across time. For example, one foraging location, such as a particular fruit tree, may have held a high value over the last couple of days since fruits have ripened, but a low average value over the last few months, when there were no fruits at all. On the other hand, another location populated by edible insects may have lower value right now, but higher long-term value, because the presence of insects is more regular across seasons. Both neuroimaging studies in humans and single-neuron recording studies in macaques demonstrate that the dACC simultaneously holds multiple representations of value with different time constants (Bernacchia et al., 2011; Cavanagh et al., 2016; Meder et al., 2017; Murray et al., 2014; Seo and Lee, 2009; Spitmaan et al., 2020; Wittmann et al., 2016a). For example, Meder et al. (2017) used fMRI to examine neural activity while people decided whether to repeat a choice or switch to an alternative. They reported that variation in dACC activity was related to variation in choice value. Importantly, however, dACC voxels carried estimates of choice value that were constructed over different timescales (Figure 3A, left). For example, activity in one ACC voxel might reflect whether a choice had been successful and delivered reward on average over the course of many previous trials. However, another voxel's activity might reflect whether the choice had been successful on just the most recent trials. This means that the ACC constructs multiple estimates of choice value over different timescales.

Several features of such representations are worth noting (Figure 3A). First, although present in the dACC, they are not ubiquitous. For example, neuroimaging suggests they are *not* prominent in the OFC, vmPFC, or most of the lateral PFC (LPFC), although they are found in the anterior LPFC and anterior insula (Fischer et al., 2017; Meder et al., 2017; Wittmann et al., 2016a). Second, there is a degree of topographic organization; choice values that depend on shorter timescales are, on average, anterior to choice-value representations reflecting longer timescales (Meder et al., 2017). Moreover, interactions between the dACC and other brain regions with the same properties are organized in a temporally structured manner: activity in dACC voxels operating on short and long timescales is, respectively, correlated with activity in voxels operating on the

same timescales in other areas, such as inferior parietal lobule. Third, there is, nevertheless, flexibility in how choice values are represented; all dACC voxels in fMRI studies (Meder et al., 2017) and neurons in single-neuron studies (Spitmaan et al., 2020) tend to encode choice value over shorter timescales when the environment is changing quickly, so that only the recent past is a good guide to the future. Conversely, the opposite happens in more stable environments when the long-term average might be more informative. This is consistent with the idea that these timescale representations might endow the dACC with a form of metaplasticity that allows an animal to integrate feedback according to the currently relevant environmental rate of change (Farashahi et al., 2017). Such patterns of activity that occur in dACC during decision-making may also underlie the way in which post-decision, reward-related dACC activity changes depending on whether weight is to be given to just recent or also longer timescales (Behrens et al., 2007).

As well as reflecting choice value over different timescales, dACC neuron activity, even at rest, shows patterns of autocorrelation over long timescales, meaning that activity fluctuations are slower compared with the rest of the brain (Figure 3A, center). Therefore, ACC neurons have both long autocorrelation time constants and they compute value with longer time constants compared with other frontal areas (Cavanagh et al., 2016; Murray et al., 2014; Soltani et al., 2021; Spitmaan et al., 2020). Moreover, the two features appear to be linked: long autocorrelation time constants exist in neurons with long reward history time constants (Spitmaan et al., 2020). When lesions are made in the ACC, macaques can only adjust their behavior in response to the most recent outcome, but the influence of the longer history of reward and choice is lost (Kennerley et al., 2006).

If an animal can represent choice value over multiple timescales, it can represent its position in the environment with respect to the distribution of opportunities within it in ways not possible for an animal that represents one instantaneous rate of reward or reward prediction errors at one timescale (compare Figures 2A and 2B). For example, by comparing longer-term and shorter-term value representations, an animal can estimate its reward trajectory, i.e., whether the environment is getting better or worse and, if it is, how quickly it is changing. For example, an insectivorous predator may estimate that the number of prey insects on a particular tree is higher today than the average of the last week but at the same time notice that this is much lower than the annual average; this indicates a short-term peak within a long-term decline. Such information can be used to guide decisions about whether to persist in the current environment or to

(B) (Left) Activity in dACC reflects the value of alternative courses of action in macaques. The better alternative and the worse alternative are associated with significantly positive and no activity (illustrated in black and gray, respectively) as opposed to the option currently pursued, which is associated with a negative activity change (shown in yellow [adapted from Fouragnan et al. (2019)]; shaded areas denote SEM across sessions). (Center) Activity in dACC reflects prospective value—the value that a course of action, such as leaving a job and looking for better employment—might lead to in the future, taking into account not just the mean value of opportunities but also their distribution and the time horizon available to explore them (adapted from Kolling et al., 2018). (Right) Activity in dACC also reflects the sequence of actions needed to acquire a goal (adapted from Holroyd et al., 2018).

(C) (Left) In rats, two anatomical projections from dACC mediate exploring and evaluating behavioral change on the one hand and commitment to adopting a behavior on the other (adapted from Tervo et al., 2021). (Center) Monkeys were taught to alternate between task performance (working) and exploratory behavior (checking) in a paradigm that captures something of how a person, for example, a scientist, might oscillate between working on a manuscript and checking their email. Activity recorded in dACC between the end of one trial (time 0) and a lever press could be decoded to predict whether monkeys would work or check: color code indicates percentage correct linear decoding (blue indicates threshold for significance at 70%, yellow indicates 90% correct decoding) (adapted from Stoll et al., 2016). (Right) Exploration of a potential choice is often guided by uncertainty about its consequences. Neurons in macaque dACC are most active when cues are associated with uncertain outcomes (50% chance of outcome) as opposed to certain outcomes (0 or 100% outcomes; adapted from Monosov, 2017).

switch to an alternative, as we discuss in the next section (Wittmann et al., 2016a), and it may also enable meta-learning (Schweighofer and Doya, 2003). Timescale-based value information is not just important for animals; value information occurring at different timescales may also be an important determinant of mood in humans (Eldar et al., 2016; Keren et al., 2021).

Careful analysis of dACC activity patterns (Figure 3A, right) suggests that subjects do indeed use the information contained in dACC representations to guide decisions about whether to keep foraging in one environment or to explore an alternative environment (Wittmann et al., 2016a). Moreover, individual variation in dACC activations reflecting short timescales predicts individual variation in the influence that value estimates constructed over short timescale will have on behavior. Similarly, individual variation in longer timescale neural representation strength predicts individual variation in the impact of longer-term reward history on behavior (Wittmann et al., 2016a).

The representation of the environment borne by the dACC is different to that found beyond the cortex, for example, in brainstem regions such as the DRN. DRN activity reflects very broad aspects of the environment, such as whether it is good or bad and what its average value might be (Cohen et al., 2015; Hayashi et al., 2015; Khalighinejad et al., 2022; Wittmann et al., 2020). A dACC-possessing animal can represent the distribution of opportunities across its environment and not just its mean value and prediction errors about that mean. This is important if an animal is to identify and pursue a distant reward goal that lies beyond what might otherwise be a barrier—a region associated with minimal rewards or even a cost. Lesions in the Cg1/Cg2 regions in the rat (areas with some similarities to primate dACC) disrupt the ability of rats to climb over a barrier to reach a more valuable outcome (Rudebeck et al., 2006; Walton et al., 2002, 2003).

Nevertheless, recent analyses of neurophysiological recordings from one subcortical region, the DA midbrain, demonstrate some distributional coding of values (Dabney et al., 2020). In neuroimaging studies, such representations have not yet been identified in the DA midbrain. This might mean that they are not as prominent as those in cortical areas such as dACC, or it may simply reflect the limits of spatial resolution in neuroimaging studies. It is also possible that representations at multiple timescales exist in both the subcortex and dACC but that their location reflects other aspects of task complexity. Directly comparing distributional coding and the degree of reward encoding over different timescales in cortical regions such as the dACC and DA midbrain and determining whether they are similar or different and the degree to which they are mutually interdependent or independent will be an important challenge.

dACC and behavioral change

How might such representations guide adaptive and flexible decision-making? We often think of decisions as being between one well-defined option and another. For example, a monkey choosing between an apple or an orange. Animals can and do make binary decisions, but this scenario may not be representative of situations that foraging animals regularly encounter (Pearson et al., 2014). It is a lucky monkey that finds itself near and equidistant to fruiting orange and apple trees, but this is the situation that many of our laboratory decision-making tasks simulate. Instead, foraging animals often encounter one opportunity at a time, and the question is whether to engage with that opportunity or whether doing so represents an opportunity cost in terms of what else they are forgoing (Charnov, 1976; Freidin and Kacelnik, 2011; Stephens and Krebs, 1986). In other words, the foraging macaque may encounter an apple tree and when it does, it needs to decide how much time and effort it should devote to foraging in the tree as opposed to continuing to explore for other opportunities. In a similar way, when one is already engaging with an option, the question is whether and when to leave it. For example, a human "foraging" in the job market might consider the value of alternative jobs to the one in which they are currently employed. Apparent idiosyncrasies in binary decision-making behavior can be explained if animals are evaluating options against the context in which they occur as would be expected of a sequential forager evaluating each opportunity encountered against the opportunity cost it entails in the current environment (Freidin and Kacelnik, 2011).

Activity across several frontal lobe areas reflects aspects of such decision variables, but it is notable that ACC activity prominently reflects the value of switching and the range of alternatives (Blanchard and Hayden, 2014; Fouragnan et al., 2019; Hayden et al., 2011; Kolling et al., 2012, 2018; Lopez-Persem et al., 2016; Mehta et al., 2019). It has been suggested that such activity might simply reflect the difficulty or response conflict when deciding which choice to take during a decision but it is now clear from careful examination of data both from early (Kolling et al., 2012) and more recent studies that such signals are strongly decorrelated from difficulty and cannot be explained by difficulty; instead, it is the value of the potential alternatives that the environment furnishes that is represented in dACC in both macaques and humans (Fouragnan et al., 2019; Kolling et al., 2016a, 2016b, 2018; Vassena et al., 2020; Wittmann et al., 2016a). Closely related activity patterns have been reported in interconnected regions of the posterior cingulate cortex (Barack and Platt, 2021; Barack et al., 2017).

For example, Fouragnan and colleagues used fMRI to examine activity across the brains of monkeys that tracked values of three possible choices. The values of the choices gradually changed over the testing session so that a choice that was good at one time was not so good at another. In addition, on any given trial only two out of three options were available for the monkey to choose between. Both task manipulations made it advantageous for the animals to switch choices from trial to trial, and this is indeed what they did. Different types of representations of the values of alternative choices were found in the dACC and hippocampus. Although hippocampal activity reflected the value of currently unavailable options, dACC activity reflected the best alternative to the current choice, regardless of whether that option was available (but rejected) on the current trial or if it was not presented at all but might reappear in a future trial (Figure 3B, left). Such a pattern of activity is not predicted by accounts of the dACC emphasizing only the difficulty of deciding. This is because alternative options that cannot be chosen during the current decision, but which might be chosen at some point in the future, should not change the difficulty of the current decision. dACC activity not only reflected the value of

Neuron
Review

CellPress

alternative choices but it also reflected how likely animals were to switch to them if they were offered on a future occasion. dACC is a crucial node in the network for changing behavior and switching between choices; when its activity was disrupted by transcranial ultrasound stimulation (TUS), there was a reduction in adaptive patterns of switching behavior; switching was no longer more determined by the value of the best alternative choice but instead, maladaptively, it was more influenced by the value of the other, less good, alternative (Fouragnan et al., 2019).

When people are in a similar situation and decide whether to engage with the most immediate opportunity or to explore alternatives, then again ACC activity reflects not just one but many aspects of the environment of alternative possibilities that contribute to improve decision-making (Kolling et al., 2018). If we reflect on how a person might be constrained to move through and explore their environment, it becomes clear that not only the best alternative opportunity should be represented, but other environmental features should be represented as well (Figure 3B, center). Imagine a scientist deciding whether to continue in their current job or to leave to invest time seeking better prospects elsewhere. The scientist should obviously consider the value of their current position and they should compare it with the job that they hope to find if they leave. However, in addition to the value of the coveted "dream job," they also need to consider how likely they are to obtain it, or any other job, given their time horizon—the time period they have in which to explore the job market. In other words, the average value of alternative options in the job market, as well as their variance in value and ease of access need to be considered. If the applicant has plenty of time to explore alternatives, then they might plan to apply and re-apply for the best jobs until they end up with one of them. However, if limited resources mean they only have limited time in which to land a new job, then they must be prepared to settle for more mediocre alternatives. All these factors determining the environment's value for the job seeker—the average value of alternatives, their variance, and the time horizon for exploration—are encoded in ACC activity (Kolling et al., 2018).

ACC activity does not just encode the opportunities available in the environment. In addition, it encodes what a person or animal might do to get to them (Figure 3B, right). Individual neuron activity patterns and the multivariate pattern of activity in dACC reflect how rats, monkeys, and people progress through sequences of actions toward a goal (Holroyd et al., 2018; Ma et al., 2014, 2016; Procyk et al., 2000; Shahnazian and Holroyd, 2018; Shidara and Richmond, 2002) and the occurrence of unexpected events as they make their progress (Ribas-Fernandes et al., 2011). Again, such knowledge may be the product of interactions between the dACC and hippocampus (Remondes and Wilson, 2013).

dACC and information seeking

So far, we have considered situations in which people and animals change and redirect their behavior to exploit opportunities they know the environment contains. However, in many cases such knowledge is absent or incomplete. When this is the case, dACC activity reflects the process of information seeking and the subsequent updating of the animal's model of the world. For example, Tervo et al. trained rats to accept or reject choices

signaled by two distinct tones. Each tone was paired with distinct reward probabilities, which could change at unpredictable and uncued times. This meant that rats usually developed a preferred option they accepted and an unpreferred option they tended to reject. Nevertheless, rats periodically changed away from making a preferred choice to find out more about the value of an alternative (Tervo et al., 2021). Although they were more likely to do this when their preferred option had recently been unrewarded, they were also guided by expectations about the duration of periods in which either option might be the better one, and they also spontaneously moved in and out of exploratory phases of behavior (Figure 3C, left).

The experiment performed by Tervo and colleagues is important not only because of the way in which it examines a naturalistic behavior in a carefully controlled setting but also because it provides a way of reconciling insights into the role of the ACC in, on the one hand, switching and exploring and, on the other hand, persistence and effort investment (Croxson et al., 2009; Kennerley et al., 2009; Klein-Flügge et al., 2016; Parvizi et al., 2013; Rudebeck et al., 2006; Walton et al., 2002, 2003). This is because dACC activity makes different contributions to behavior at different points in time during decision sequences and it does so via different microcircuits linking it to divisions of the distinct subcortical circuits. One microcircuit in rats runs from the dACC to the substantia nigra pars reticulata (SNpr) and is important for initiating exploratory behavior at the time that the choice is made. Optogenetic silencing of either the dACC (area Cg1/24b) itself or the dACC-SNpr pathway selectively reduced the frequency of exploratory choices of an alternative option as opposed to the preferred one. By contrast, another dACC microcircuit in rats runs from the dACC to the striatum, and it is important for persistence in a behavior after a decision is taken and no reward is received. Optogenetic silencing of either dACC or the dACC-striatum pathway reduced the frequency with which animals persisted with a choice after non-reward (Tervo et al., 2021).

A range of approaches have been employed to address the question of information seeking, but all have converged in highlighting the dACC (Figure 3C). Stoll and colleagues (Stoll et al., 2016) trained macaques to perform a perceptual discrimination task for food reward. The monkeys could opt to perform the task or opt out to seek information about another visual object that gradually changed in appearance until it indicated a large bonus reward would be delivered. Many neurons in the dACC changed activity at the time when the macaques sought information about the growth of the bonus indicator (Figure 3C, center). Monosov, White, and colleagues trained macaques in a very different behavioral paradigm in which macaques did not have to actively seek information themselves but nevertheless encountered cues providing information about what was to happen next and other cues confirming prior expectations (Figure 3C, right). Activity in dACC neurons ramped up when macaques expected an information-bearing cue (Monosov, 2017; White et al., 2019). Hunt and colleagues trained macaques to choose between multicomponent visual objects (Hunt et al., 2018). The component features of the objects provided information about the magnitude or probability of rewards that would be received if chosen. The features were sometimes obscured, but animals could seek the information they contained by saccading

toward them. ACC activity was linked to the usage of information that had been sought by making a saccade; ACC activity reflected the degree to which the information revealed by a feature confirmed that the monkey would be making a good decision by choosing that object. Neuroimaging studies confirm that the dACC has a preeminent role in information seeking; its activity reflects a person's uncertainty about the choice that they are taking when they are actively exploring options to obtain information rather than when they are simply randomly responding (Trudel et al., 2021). Often in these studies, as in real life, seeking information can be beneficial in the future. How this differs from curiosity—the mere desire to know—and whether curiosity relies on dACC representations remains to be fully determined (Bromberg-Martin and Hikosaka, 2009; Kidd and Hayden, 2015; Kobayashi et al., 2019; Monosov et al., 2020; Wang and Hayden, 2021).

Like Tervo and colleagues, who investigated rodents, White et al. (2019) confirm the importance in primates of a circuit spanning the dACC and striatum and in addition provide evidence that the circuit extends into the pallidum. Striatal regions near the internal capsule are strongly connected with ACC and these in turn are strongly connected with anterior globus pallidus (GP) and ventral pallidum. Neurons in all three areas show similar information-seeking-related activity.

In sum, the dACC represents the distribution of opportunities in the environment, it computes recent and long-term value, and on these bases, it determines whether a person or other animal should engage with a current option or explore the environment, including driving specific information-seeking activity (Figure 3). However, important questions remain. Uncertainty-related activity has been reported in subcortical systems, including the noradrenergic system originating in the locus coeruleus and the raphe nucleus. Our picture of how these regions interact with one another and the dACC is currently changing rapidly (Joshi and Gold, 2022; Joshi et al., 2016; Muller et al., 2019; Soltani and Izquierdo, 2019; Tervo et al., 2014), but there is much that we still do not know. Perhaps most critically, we do not know why they interact, whether they encode identical indices of uncertainty, whether they have similar influences on behavior, and how they compare with influences from other neuromodulatory systems (Danielmeier et al., 2015; Fischer et al., 2015; Khalighinejad et al., 2020a, 2022).

Learning about environmental structure versus reward distribution in the prefrontal cortex

In addition to the ability to organize behavior in time as a function of the distribution of opportunities in the environment, many mammals know about the structure and causal nature of relationships between these opportunities and other features of their environments. Such structural knowledge may suggest that the best choice to take next is not the one in which the animal has experienced the best reward distribution. For example, consider the case of a monkey that, one morning, finds a ripe fruit in a tree that yesterday's search had revealed only to contain unripe fruit. If the monkey only considered the distribution of reward in the environment that it had experienced, then the discovery of one fruit in this previously unrewarding tree might not be enough to detain the monkey from going somewhere else. However, if the

monkey knows about the way in which ripe fruit appears on a tree—in other words, if it has knowledge of the structural organization of its environment—then the presence of one ripe fruit in a tree might be sufficient to signify that all the fruit in the tree is ripening now and that the new location might now be the ideal one in which to forage. Such knowledge about the structure of the environment does not, however, depend on the dACC.

One demonstration of this, for example, was recently provided by Vertechi et al. (2020). They devised a task for mice with features reminiscent of the foraging environment discussed above. The mice learned about an environment containing two locations at which they could nose poke. Importantly, water reward was probabilistically available at just one of the two locations, but its position varied over time. If the mice understood that water was available at only one location at any given time then this meant that once they had discovered water at one location, they should no longer visit the other location, even if water had been found there on many previous occasions. After a training period, mice learned to switch quickly; once they have discovered water at one location on just a single occasion, they focused on that location regardless of whether the other was previously rewarding on many previous occasions. Vertechi and colleagues argue that knowledge of environment structure allows mice to infer that the location of the water has changed as soon as they receive a reward for the first time at a new location. Regardless of the previous distribution of reward, they should now switch and focus on this new location. In line with what we have argued about the ACC above, Vertechi and colleagues reported that ACC inactivation delayed switching regardless of the previous distribution of rewards. However, after OFC inactivation a different pattern of behavioral change was observed; mice were slower to switch but their impairment was a function of how much water had been delivered at the other location previously. In other words, after OFC inactivation the mice's behavior was guided by their experience of the distribution of reward in their environment, but it was no longer guided by their knowledge of the structure of opportunities in the environment.

This is just one of several studies emphasizing the importance of the rodent OFC for the construction of models of the relationships between features of the environment (Wikenheiser and Schoenbaum, 2016). For example, OFC neuron activity reflects the learning of associations between cues that occurs incidentally, even when the accumulation of this knowledge is not immediately reinforced (Sadacca et al., 2018), and other studies implicate the human OFC in similar inference processes (Wang et al., 2020a, 2020b).

If the ability to make inferences based on knowledge of the structure of the environment rather than just experience of the distribution of reward in an environment is present in a mouse, then it may be present in a number of mammals. Nevertheless, it has received special attention in primates. Primates have highly developed visual systems and spend a large part of their time engaged in visual exploration of their environments (Graziano, 2009), often appearing to survey it from a much greater distance than does a rodent (Rudebeck and Izquierdo, 2021). They may range over large areas of many thousands of hectares as they forage (Passingham and Wise, 2012; Rudebeck and Izquierdo, 2021). Despite their impressive range, their feeding is often focused on the fruit,

Neuron
Review

CellPress

young leaves, and insects found in angiosperm trees. As a result, one difficulty primates face is in identifying the best places in which to forage within their range because only a small proportion of the trees within it, perhaps as low as 4%, bear fruit at any point in time (Zuberbühler and Janmaat, 2010). The strategies that primates exploit to estimate where they might find food suggest that, in addition to using knowledge of spatial relationships such as those examined in the rodent study by Vertechi et al., they also frequently use knowledge of non-spatial, structural relationships between environmental elements to guide their choices. For example, they can learn patterns of correlation between visual aspects of the trees' appearances, which can be discerned from afar, and food; they spontaneously search for and inspect visual cues that have previously been seen near food (Menzel, 1996). When they see that one tree is in fruit, they are more likely to visit other similar trees, suggesting that they infer that similar trees might be fruiting too (Janmaat et al., 2012). When the weather has been warm, they return to trees that they have recently seen with unripe fruit, suggesting they have inferred that that fruit may now be ripe (Janmaat et al., 2006). And when persimmons are left in their home range, Japanese macaques appear to infer that persimmon trees may be fruiting and are more likely to visit persimmon trees (Menzel, 1991). Finally, primates are known to live in large social groups, and they excel in adjusting their behavior based on structural knowledge of the social dynamics and hierarchies present in their troop (a topic that we return to in more detail below in the section about dmPFC).

In summary, many animals do not just learn and exploit relationships between sensory cues and their reinforcement consequences, but they also understand the structure of relationships and contingencies between cues. Primates are particularly good at this, and in humans, such relationships may be quite abstract (Donoso et al., 2014). They often involve outcomes that will only unfold much later, or choices made on behalf of others, for example, when predicting how the choice of our children's school may impact their future prospects. We consider these ideas further in the following sections on the anterior medial and dorsal medial PFC.

Anterior medial prefrontal cortex and representation of structure

Animals possess cognitive maps of the world around them (Tolman, 1948). Importantly, such maps make it possible not just to follow previously taken routes but also to allow "vector navigation"—the ability to move directly from the current position to a goal location, even when it is not directly observable, by a novel route that might never have been taken before. Several medial temporal lobe (MTL) areas, such as the hippocampus and entorhinal cortex, and interconnected areas have long been recognized as preeminent in such computations (Hartley et al., 2003, 2014; Bush et al., 2015). Neuroimaging experiments, however, suggest that a region on the medial surface of the frontal lobe is also a component of this circuit (Doeller et al., 2010). Moreover, in both human and non-human primates, the area's role is not confined to representing spatial information; it also represents the structure of arbitrary associations between non-spatial items (Bao et al., 2019; Baram et al., 2021; Bongioanni et al., 2021; Constantinescu et al., 2016; Gerraty et al., 2018; Schuck et al.,
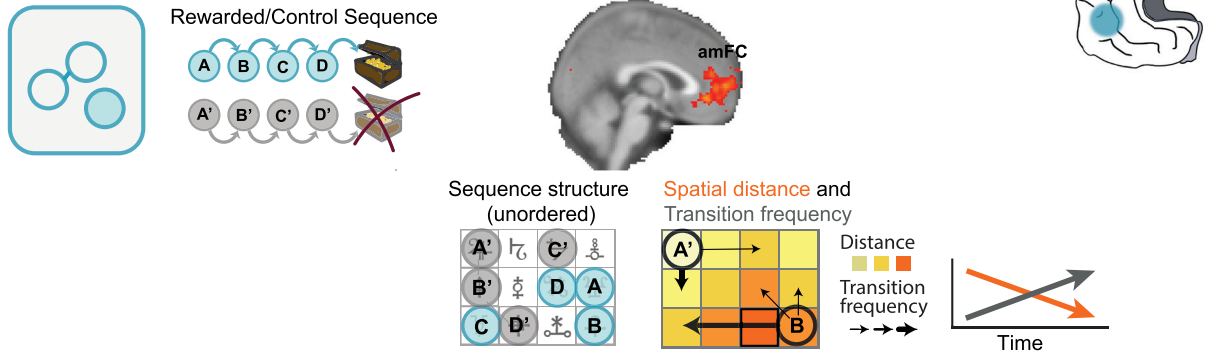
2015). It is possible that the effects of lesions of medial frontal cortex that have previously been attributed to inflexibility or perseveration are better understood in terms of an animal's beliefs about task structure (Jang et al., 2015).

It is difficult to identify the precise location of the key medial frontal region highlighted in neuroimaging experiments because our knowledge of the anatomy of this region is still evolving (Glasser et al., 2016; Neubert et al., 2015). Its posterior boundary is in or near the part of the cingulate sulcus anterior to the genu of the corpus callosum, and it extends anteriorly to reach the medial aspect of the frontal pole. It therefore seems likely to include prefrontal areas such as medial area 10 (10v and 10r), but it may also extend ventrally into prefrontal area 14 and posteriorly into anterior pgACC, p32. We refer to it here simply as amFC (Figure 4). Areas 10 and 14 have only been identified in primates such as macaques and humans, and similarities in resting state connectivity patterns suggest the areas are components of similar circuits in the two species (Neubert et al., 2014, 2015). By contrast, area 32 also bears resemblances to regions found in non-primates, such as the prelimbic region of rodents (Vogt, 2009). Even if its boundaries remain to be precisely defined, recognizing that there is a specialized subregion within this area that is different from the adjacent dACC, pgACC, and vmPFC is necessary if we are to account for the diversity of frontal cortical contributions to decision-making and behavioral flexibility.
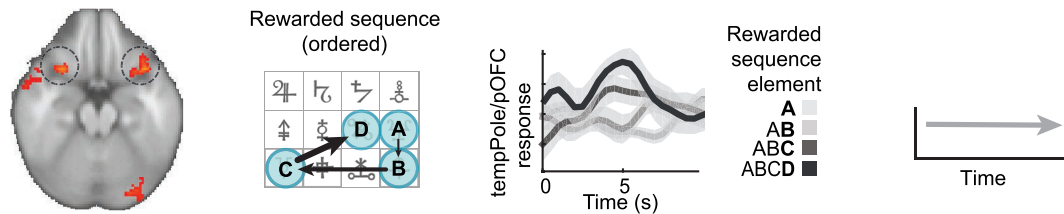
Just as maps of the spatial world allow animals to perform vector navigation, the models of abstract relationships held by the amFC mean that, even without direct prior experience of all possible states of the world, amFC can make predictions about the nature of unobserved states and the consequences that will follow for the animal if they are entered. Thus, mammals possessing amFC can flexibly adjust their behavior by simulating the consequences of potential courses of action even before experiencing them, and they can generalize and apply known relationships to new situations (Behrens et al., 2018). For example, a monkey might infer that a fruit tree in which it has never previously foraged is now likely to be of high value if it has observed that another similar tree at another location has fruited or if it has recently observed appropriate weather conditions. Similarly, humans can make informed choices between potential holiday destinations even if they have not visited any of them before, based on information gathered from places with some shared features. This requires abstract representations of relationships and the simulation of potential consequences.

The amFC and adjacent cortex are often active during value-based decision-making (Bartra et al., 2013; Clithero and Rangel, 2014; Rushworth et al., 2011), but it is currently debated whether this should be attributed to their most fundamental role being one of valuation and reward prediction or to some other cognitive process such as the identification of task structure or determination of behavioral policy (Hayden and Niv, 2021). Although some frontal areas we return to below may be especially concerned with representations of value, there is increasing evidence that the amFC is particularly important for representing the structure of the environment. Klein-Flügge et al. (2019) provided one demonstration of this in an experiment in which they taught human participants to navigate through an artificial task
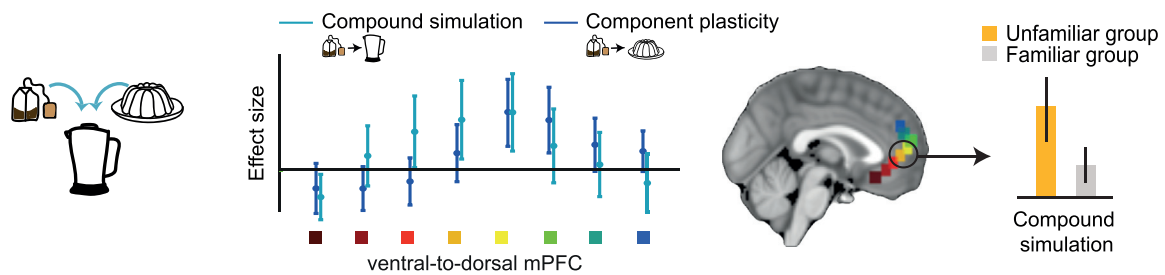
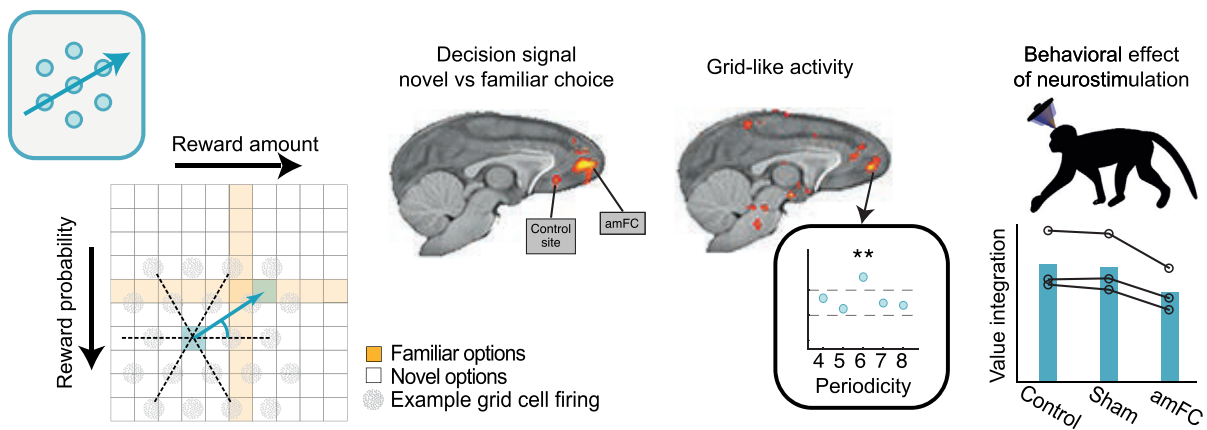**A** **amFC and flexible representation of task structure**

Rewarded/Control Sequence

Sequence structure (unordered)

Spatial distance and Transition frequency

Distance

Transition frequency

Time

B **pOFC and robust representation of reinforced structure**

Rewarded sequence (ordered)

tempPole/pOFC response

Time (s)

Rewarded sequence element
A
AB
ABC
ABCD

Time

C **amFC and simulation of new experiences**

Compound simulation — Component plasticity

Effect size

ventral-to-dorsal mPFC

Unfamiliar group
Familiar group

Compound simulation

D **amFC and grid-cell like representations of task space**

Reward amount

Reward probability

Familiar options
Novel options
Example grid cell firing

Decision signal novel vs familiar choice

Control site
amFC

Grid-like activity

**

4 5 6 7 8
Periodicity

Behavioral effect of neurostimulation

Value integration

Control   Sham   amFC

(legend on next page)

Neuron
Review

CellPress

environment comprising a 3 × 4 array of abstract shape stimuli (Figure 4A). Their goal was to learn which sequence of stimuli led to a reward, but participants also showed learning of incidental statistical relationships governing transitions through the task environment. During initial training, participants' movements through the stimulus space were constrained by spatial distances, allowing only movements between adjacent stimuli. In line with this, upon entering the scanner, amFC BOLD activity was strongly modulated by the spatial distance separating two successive stimuli encountered. However, during scanning, movements through the 3 × 4 task environment were no longer governed by spatial constraints and the amFC signature of spatial distance faded over time. Instead, participants' attention during scanning was guided by sequentially highlighted stimuli, which were no longer necessarily adjacent in space. Highlights could "jump," but their sequence followed a particular transition pattern. amFC BOLD signals at the time of a highlighted stimulus were modulated by the likelihood of the experienced transition, as tracked by a simple learning model. The ability to predict such transitions provided a behavioral advantage in the scan task, and the amFC signature of transition frequency became stronger as scanning progressed. Thus, amFC's model of the task flexibly changed in line with the most advantageous behavioral policy and reflected both spatial and non-spatial aspects of the task structure. Importantly, the structural knowledge reflected in amFC activity was present independent of whether the transitions led to reward or not. By contrast, activity in other areas, such as the posterior lateral OFC, only reflected knowledge of specific reward-reinforced stimulus sequences and, unlike in amFC, these OFC activity patterns were robust and inflexible; posterior lateral OFC activity continued to hold the same rewarded sequence representations, even when they were no longer relevant for the task in hand (Figure 4B).

Knowledge of abstract relationships can be particularly useful when we need to simulate new experiences to make predictions; several studies demonstrate that this is the case. For instance, Barron and colleagues asked participants to imagine new reward experiences based on novel combinations of previously experienced foods (Barron et al., 2013). Using fMRI in humans, they showed that the amFC held representations of the novel experience and that these amFC representations of novel food combinations were linked to representations of the previously experienced component elements (Figure 4C). In another study, Bongioanni et al. (2021) trained macaques to choose between pairs of two-dimensional stimuli for liquid rewards (Figure 4D). They created new stimuli for the monkeys by presenting new combinations of amount and probability, but importantly, the component features of the new stimuli were familiar, thus allowing the monkeys to infer the value of the new stimuli. fMRI and TUS revealed that the amFC held multidimensional representations of the new options that were needed for optimal decision-making between them. This supports the idea that the amFC performs feature integration and represents the knowledge of abstract relationships to help simulate novel experiences (Fellows, 2006; Kahnt et al., 2011; Spalding et al., 2018).

A recent proposal suggests that making novel inferences in abstract task spaces may depend on neurons with "grid"-like activity patterns (Behrens et al., 2018; Whittington et al., 2020). Such neurons, first reported in rat entorhinal cortex, have activity covering the spatial arena. Each neuron covers the space by possessing an array of place fields arranged on a triangular grid. If a rat takes some paths through the testing arena, then the neuron will fire repeatedly as the animal moves from one of the neuron's place fields to the next. If different movement trajectories are tested, the neural response will vary depending on how much any trajectory is aligned with the cells' grid field. Because there are six ways to align to a triangular grid, when examining the neural activity measured during transitions along all possible directions, the signal will increase and decrease six times per cycle, i.e., every 60°, and because the grid fields of

Figure 4. Task structure and amFC
(A) (Top left) Human participants experienced two different four-element sequences, one of which led to reward and one of which did not. (Bottom center) The four elements in each sequence corresponded to locations on a 3 × 4 grid. (Top center) Activity in amFC reflected the structure of the sequences that participants learned. (Bottom right) In addition, initially amFC activity reflected the spatial distance between any two elements on the grid. Spatial distance was important during training but became irrelevant during the scan session. Over time, amFC activity began reflecting the now relevant transition frequencies between one location and the next.
(B) (Left) By contrast, posterior OFC (pOFC) and temporal pole activity reflected the precise order of reward-reinforced sequences. (Center) Activity in pOFC and temporal pole ramped up as increasingly more rewarded sequence elements were present in the correct order. It was thus greatest for the final element, D, of the sequence but only if D was preceded by the correct stimuli—A, B, and C—in the correct order (shaded areas denote SEM across participants). (Right) These representations did not change over time—they were robust and inflexible (adapted from Klein-Flügge et al., 2019).
(C) (Left) Human participants were asked to simulate a new experience (a novel combination of tea and jelly) on the basis of past experiences (of tea and of jelly). (Center) amFC held representations of the novel simulations (light blue) and showed that representations of the previously experienced component elements became more similar to one another (dark blue). (Right) Once participants actually experienced the novel combination (familiar group versus unfamiliar group), their ability to represent it was no longer a simulation based on experience of the component elements and so amFC representations of the combination and component elements became unlinked from one another (adapted from Barron et al., 2013); error bars denote SEM across participants.
(D) (Left) Macaques chose between novel stimuli they had not previously experienced or had only limited experience of. Prior to the main part of the experiment, monkeys were, however, extensively trained but only on a subset of stimuli (yellow cross, familiar options). Different visual features of the stimuli—their color and the density of dots with which they were covered—indicated the amount and probability of juice rewards that would follow if they were chosen. However, during the critical test phase, when fMRI data were acquired, the monkeys demonstrated that they could draw on their knowledge of these visual features to accurately estimate the values of novel options (white). (Center left) During choice, the value difference reflects the key decision variable—how much better is one option than the other. When decisions were made between novel rather than familiar options, which required evaluating novel stimuli based on their magnitude and probability dimensions, the amFC decision signal was significantly stronger. (Center right) Consistent with the idea that a process of online simulation and integration of the reward probability and magnitude features was occurring, there was evidence of grid-cell-like encoding of the probability and magnitude space that defined the task; a hexagonal pattern of activity modulation when animals encountered single options on different trajectories with respect to one another. Notably, no such preferential signaling of novel choice computations nor grid-like representation formats was observed in OFC or dACC. (Right) The ability to base decisions on integrated choice values (combinations of magnitude and probability information) was disrupted by application of TUS to amFC; lines denote individual monkeys (adapted from Bongioanni et al., 2021).

CellPress

Neuron
Review

neurons are aligned within a given animal, the aggregate activity of the population can be measured with fMRI. This pattern is especially prominent in the amFC in both humans and macaques when they navigate through non-spatial task environments (Bao et al., 2019; Baram et al., 2021; Bongioanni et al., 2021; Constantinescu et al., 2016) just as when they navigate in a spatial arena (Doeller et al., 2010). Therefore, abstraction in non-physical space may rely on a similar grid cell-based coding scheme as that first discovered for physical space. In their study, Bongioanni and colleagues (Bongioanni et al., 2021) provided evidence that amFC activity is consistent with grid-like encoding of an abstract value space for novel choices. Crucially, in addition, the authors showed that disrupting this activity with TUS impaired the integration of information across this space in the guidance of decision-making (Figure 4D). This demonstrates the amFC's causal role in representing abstract task relationships and in simulating novel experiences along dimensions of this task space.

The amFC is interconnected with the MTL, where grid cells were first observed. This raises the question of whether one region's activity is driven by the other, or whether they play complementary but independent roles. A recent intracranial EEG study of spatial navigation in humans (Chen et al., 2021) suggests that the medial frontal grid signal may precede the one observed in MTL. Similarly, activity in the DA midbrain and ventral striatum has sometimes been thought to resemble activity linked to model construction and inference typically observed in the PFC (Daw et al., 2011). When directly probed, the amFC represents the full structure of participants' task model while the ventral striatum does not contain this information, even if it reflects prediction errors that are contingent on such models (Klein-Flügge et al., 2019). Nevertheless, the precise contribution and communication between subcortical and PFC representations remains to be fully determined.

Orbitofrontal cortex in rodents and primates

The previous section focused on the amFC in primates. Rodents, however, also make inferences. What frontal brain structures do they use when they do so? When Vertechi et al. (2020) investigated how mice learned that only one of two locations in an environment was associated with reward and used this knowledge to make inferences about where to forage, they focused on an area that they referred to as the OFC. The same region has also been emphasized by Liu et al. (2021), who examined how mice made inferences about auditory stimuli with respect to a shifting criterion. The mice learned to respond in one of two directions, depending on whether an auditory tone had a frequency above or below a criterion level. From time to time in the task, however, the criterion shifted upward or downward. Initially the mouse had to learn by trial and error that a given tone was now, for example, higher than the new criterion, even if it had not been higher than the previous criterion. However, such an experience should allow the mouse to infer, for example, that other even higher tones are also above the new criterion. If the mouse understands the task's structure, then it does not have to learn what to do when it hears each tone, simply by accumulating experience with that tone alone; instead, it can make inferences from one tone to another. Again, OFC disruption compromises

the mouse's ability to do so. Other aspects of task structure, such as sequential organization, are also encoded in rodent OFC (Zhou et al., 2019, 2021).

The designation of the OFC area in rodents follows from the work of Uylings and van Eden (1990), who attempted to establish rodent-primate similarities in prefrontal areas on the basis of their thalamic connection patterns rather than their intrinsic cytoarchitecture. However, whether the rodent OFC corresponds in a simple and direct way to primate OFC has been long debated (Preuss, 1995; Wise, 2008). A notable difference between the rodent and the primate OFC is that OFC lesions in macaques do not cause switching deficits in reversal tasks (Rudebeck et al., 2013). Understanding how these OFC areas in primates and rodents relate to one another is far from straightforward.
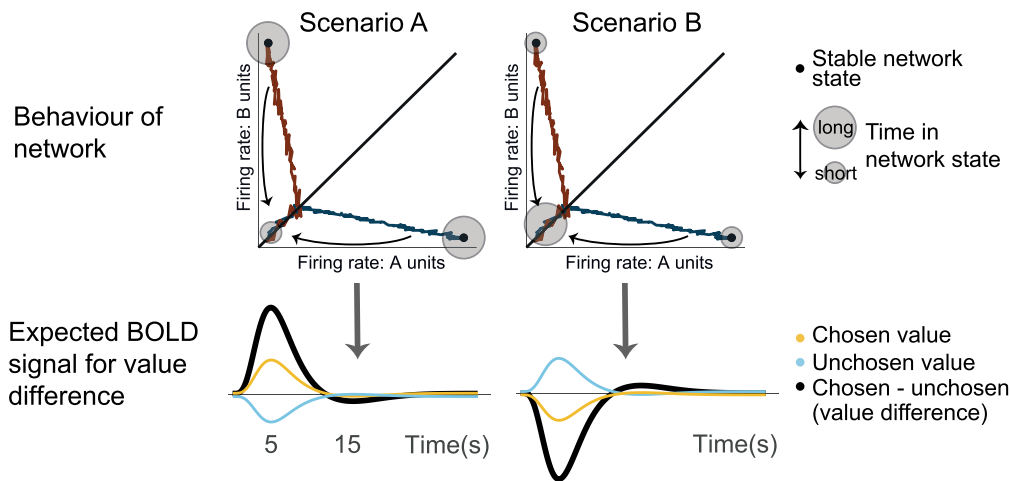
There are also functional similarities between rodent and primate OFC areas. Even if OFC lesions do not disrupt reversal task performance, Rudebeck et al. (2013, 2017) report that these lesions in primates do disrupt decision-making in reward devaluation tasks, just as they do in rodents (Lichtenberg et al., 2021; Malvaez et al., 2019; Pickens et al., 2003; Sias et al., 2021). Disruption of the OFC leads to other patterns of change in reward-guided decision-making that are similar in mice and macaques (Ballesta et al., 2020; Kuwabara et al., 2020), and there is evidence that the human OFC plays a similar role (Howard et al., 2020; Wang et al., 2020a, 2020b). In such tasks, different choices lead to different rewards. If one reward is devalued (for example, by feeding an animal to satiety on that reward prior to testing), then the animal should infer that it is no longer optimal to pick the choice that leads to the devalued option and so they should refrain from taking it and choose the alternative one. The fact that the reversal task and the devaluation task are dissociable in primates suggests that they tap into at least partially dissociable cognitive abilities. Succeeding in a reversal task involves, in addition to cognitive flexibility, an understanding of the structure of the task, for which the primate OFC may not be required, while succeeding in a devaluation test requires anticipation of future outcomes and their value for the self, which appears to be enabled by the OFC across species.

Additionally, there are also similarities between primate amFC and rodent OFC. Both primate amFC and rodent OFC (Stalnaker et al., 2015) represent structural knowledge and mediate inference, and both primate amFC and rodent OFC have an approximately similar topographical relationship with other areas, such as the dACC and pgACC/prelimbic cortex. OFC activity in rats reflects goal locations and not just the current location, making it reminiscent of amFC activity in humans (Klein-Flügge et al., 2019). It is also intriguing to see the similarity between the knowledge of sequential task structure decodable from the human amFC (Klein-Flügge et al., 2019) and rat OFC (Zhou et al., 2021). In addition, just as the human amFC is concerned not only with the learning of reward-related associations but also the associations between non-rewarding task features, so is the rodent OFC (Lopatina et al., 2015, 2016; Sadacca et al., 2018). On the other hand, however, it is notable how readily humans learn task structure. For example, human participants performing a version of the Vertechi et al. (2020) inference test learn the underlying task structure an order of magnitude more quickly than mice. Such findings remind us that we should not expect learning and inference to be identical in every respect in rodents and primates.
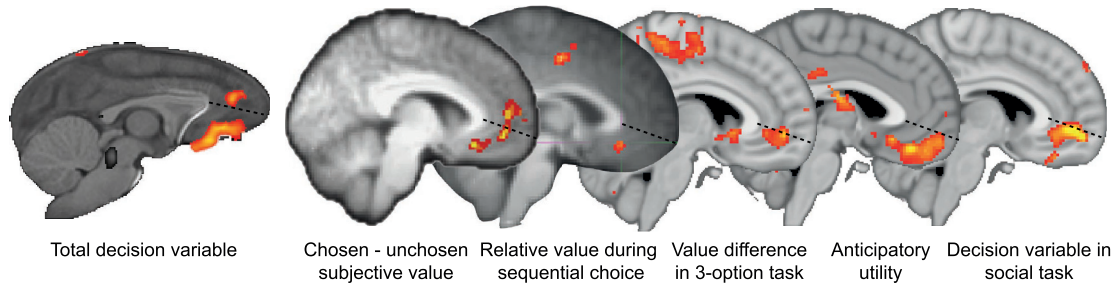
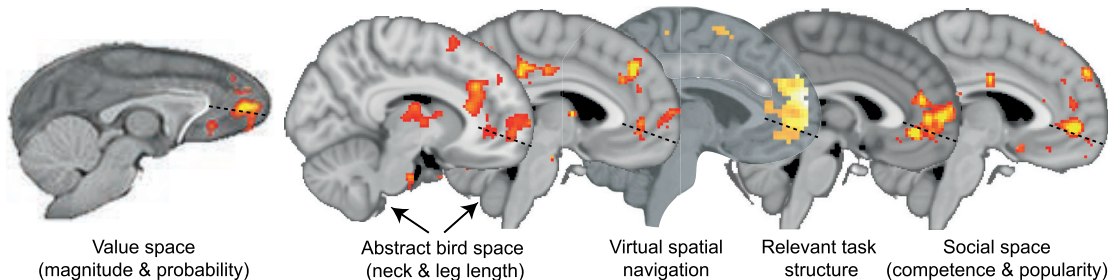**A A recurrent neural network with mutual inhibition**
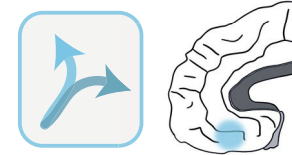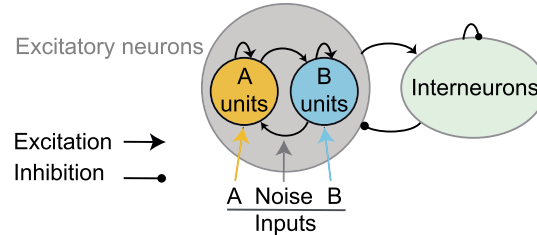
B Model characteristics and predicted effect sign

C Choice comparison: vmPFC BOLD activations of value difference

Total decision variable | Chosen - unchosen subjective value | Relative value during sequential choice | Value difference in 3-option task | Anticipatory utility | Decision variable in social task

D For comparison: amFC BOLD signals of structure/grid-like representations

Value space (magnitude & probability) | Abstract bird space (neck & leg length) | Virtual spatial navigation | Relevant task structure | Social space (competence & popularity)

Figure 5. vmPFC and decisions about not just reward
(A) Decision-making processes can be simulated in neural networks in which pools of neurons represent choices of one option or another. Neurons within each pool have recurrent excitatory connections, but inhibitory interneurons connect the two populations. This pattern of connectivity ensures that if one population

It is also important to remember that Klein-Flügge et al. (2019) found evidence of encoding of sequence structure in a second frontal area—a posterior lateral orbitofrontal area. As already noted, sequence knowledge in this area is focused on reward prediction and is updated slowly, but it is also more robust and unchanging compared with the amFC. Its position near the border between the granular and agranular cortex means that it is cytoarchitecturally more similar to the rodent OFC. The speed with which representations are updated is reminiscent of the slow speed with which rodents learn task structure.

In both macaques and humans, there is a region on the lateral border of the OFC and ventral border of the ventrolateral PFC that is important for learning specific choice-outcome contingencies (Boorman et al., 2016; Chau et al., 2015; Folloni et al., 2021; Jocham et al., 2016; Noonan et al., 2010, 2011, 2017; Rudebeck et al., 2017; Walton et al., 2010). The critical region is not in areas 11 and 13 between the lateral and medial orbitofrontal sulci (Rudebeck et al., 2017) but instead lies in and lateral to the lateral orbitofrontal sulcus in area 47/12o (Chau et al., 2015; Folloni et al., 2021). Normally, monkeys' decisions between choices reflect the history of reward received immediately after taking such choices in the past. When a lesion is made that includes 47/12o or an ultrasound is focused to disrupt 47/12o, outcomes are not correctly credited to the choices that caused them and, consequently, choices are simply repeated if they were made in the context of a high global reward state even if the choice itself was not causally responsible for a reward. Activity in other brain regions, such as the DRN and insula, reflects global reward state regardless of choice taken (Folloni et al., 2021; Wittmann et al., 2020). Therefore, animals lacking a PFC and relying on older neural circuits may learn based on this global signal as opposed to on the basis of fine-grained and specific choice-outcome contingencies, such as primates with lesions in 47/12o. In humans, adjacent but even more lateral prefrontal cortical regions mediate other related cognitive processes, for example, prospective metacognitive estimation of the impact that choices will have even before they are taken (Miyamoto et al., 2021).

In summary, the rodent OFC may hold representations that are similar to those present in both OFC and amFC in primates but not identical to either. In some regards, it might resemble very posterior OFC areas on the boundary with insula in the primate brain that often receive less attention in human and macaque investigations. Moreover, some features of activity in the rodent OFC, such as the encoding of alternating sequence elements (Zhou et al., 2021) or the relationships between task states (Bartolo and Averbeck, 2020; Tang et al., 2021), are reminiscent of patterns reported in yet other primate prefrontal regions (Shima et al., 2007). Rather than trying to link the rodent OFC with any one of these primate areas—amFC, 47/12o, or some of the areas that we discuss below, such as vmPFC—it may be better to think of rodent OFC as bearing a general similarity to all of them; we have illustrated this idea in Figure 1. Such a view suggests that while rodents have a frontal cortex that equips them to make inferences (Vertechi et al., 2020), primates are ready to employ many different and specialized circuits for inferential processes (Wise, 2008).

Ventromedial prefrontal cortex and decisions, but not just about reward

The amFC region described above is extensive and may contain different component subregions. When we focus on the decision process itself, however, it is noticeable how frequently that activity appears in or beyond the ventral border of this region. It is possible that another region here can be distinguished by both location and function. It is often referred to as the vmPFC (area 14 m) in humans, but it corresponds to areas sometimes referred to as the medial OFC (mOFC) in macaques. We also note, however, that the label vmPFC has sometimes been used to refer to activations extending dorsally beyond area 14 m into amFC and pgACC. Here, we refer to this region as vmPFC/mOFC and argue that it is particularly important for turning representations of choice options into actual decisions (Figure 5).

During reward-guided decision-making, vmPFC/mOFC acts as a choice option comparator. vmPFC/mOFC signals match predictions from a biophysically plausible network model (Wang, 2002, 2008; Wong and Wang, 2006) where two competing pools of neurons, each representing one option, mutually inhibit each other and compete for choice. Predictions from these models show that in measures of bulk activity such as those obtained with BOLD-fMRI or magnetoencephalography (MEG), the signature of a choice computation amounts to a difference between the value of the chosen and the value of the unchosen option (Hunt et al., 2012). This is precisely the signature of activity found in vmPFC/mOFC (Boorman et al., 2009; De Martino et al., 2013; FitzGerald et al., 2009; Hunt et al., 2012; Trudel et al., 2021). As already noted, other brain areas, such as the dACC, have activity that reflects key decision variables such as

---

ends up in a high-firing state, the other population's activity is suppressed; the choice associated with the first population is taken while the second is not. The population that is most likely to reach the high firing state is the population with the strongest input (the choice with the highest associated evidence).

(B) Emergence of positive and negative value differences in BOLD fMRI data may be explained by the time spent in the final attractor state before the network is reset. Top: the firing rate of A units is plotted against the firing rate of B units for a situation where A ends up being chosen (dark blue) and where B ends up being chosen (red). A key difference is the time spent in the high- and low-firing state, which is shown as larger (longer) and smaller (shorter) gray circles, and which may change how much the speed of the competition process relative to the final steady-state influences the activity measured, using fMRI (or similar techniques with limited temporal resolution). Scenario A (left): if the steady state has the strongest influence on the measured signal, it will be positively related to the value of the option chosen and negatively to the option rejected, as typically observed in human fMRI studies. Scenario B (right): by contrast, if the speed of the competition process predominates in the signal, the observed modulation will be negative for the option chosen and positive for the option rejected as typically observed in macaque fMRI experiments. Bottom: the expected BOLD signal is illustrated schematically for the two scenarios. See also Figure S1 for a more detailed description of the attractor model.

(C) vmPFC activity relates to the comparison of the value of the chosen option versus another option in macaques (left; adapted from Wittmann et al., 2020) and across a number of studies in humans (right; adapted from Boorman et al., 2009, 2013; Chau et al., 2014; Iigaya et al., 2020; Park et al., 2021).

(D) By comparison, a more dorsal region in amFC shows activity related to task structure in macaques (left; adapted from Bongioanni et al., 2021) and across multiple studies in humans (right; adapted from Constantinescu et al., 2016; Doeller et al., 2010; Klein-Flügge et al., 2019; Park et al., 2021).

Neuron
Review

CellPress

the difference in value between potential choices. However, closer inspection reveals important differences in activity patterns in the vmPFC/mOFC and dACC. As we have seen, dACC activity reflects the value of switching away from a current opportunity to explore alternatives, potentially over the course of an extended series of sequential decisions. Consistent with this observation, Boorman and colleagues (Boorman et al., 2013) reported that dACC activity tracked the longer-term value of a choice. However, vmPFC/mOFC reflected the choices' current value on the present trial. Boorman and colleagues were able to tease apart the representations because they employed a task in which choice values reflected two different features, one that remained relatively constant over several trials (tracked by dACC) and one that changed frequently every trial—vmPFC reflected the difference in value between options once both features had been integrated.

One intriguing observation pertains to the direction of the value difference signal observed in vmPFC/mOFC. Although it is consistently positive in humans (larger BOLD signal changes are associated with larger differences between the chosen and unchosen option values), it has the opposite sign in macaque monkeys (Bongioanni et al., 2021; Fouragnan et al., 2019; Papageorgiou et al., 2017; Wittmann et al., 2020). At first, these observations might seem incompatible. However, both types of patterns could reflect the output of the same biophysical attractor network, albeit with small modifications in the behavior of the neural populations. For example, allowing variation in the time spent in the high-firing attractor state before returning to baseline firing could produce two opposing predictions (Figures 5A and 5B; Figure S1). If the neural population remains in the high-firing state for some time before returning to baseline, the observed BOLD difference would be most influenced by this sustained activation, which would scale with option difference and thus lead to a positive BOLD value difference signal (Figure 5B, left). On the contrary, if the neural populations only briefly transitioned through the decisive high-attractor state (e.g., because the decision is immediately passed on to another region), then the BOLD signal would most strongly be influenced by the speed of the competition process, which would be shorter in a trial with a large value difference and longer when the value difference is small. Thus, in this situation, we would expect a negative relationship between value difference and the measured BOLD signal (Figure 5B, right). A similar argument has been put forward elsewhere (Hunt and Hayden, 2017). It is notable that slight differences in the type of stimulus material about which macaques make decisions lead to different patterns of positive and negative modulation in single-neuron-recorded patterns of activity as a result of different patterns of choice evidence encoding in a curvilinear manifold (Okazawa et al., 2021). Again, these differences might relate to the way in which the dimension of stimulus discrimination is transformed into the dimension of response selection and, therefore, in how information about choice selection is passed to subsequent brain areas.

Direct electrophysiological recordings from the macaque and human vmPFC/mOFC support its role in converting relative values into choices (Lopez-Persem et al., 2020; Strait et al., 2014). As in human neuroimaging experiments, vmPFC/mOFC activity in macaques reflects reward value integrated across di-

mensions and shows anti-correlated tuning for each of two options' values during decision-making—indicative of value comparison—followed by coding of the chosen option's value, which is indicative of encoding of just the final choice (Strait et al., 2014). Similar signals are observed in intracranial electro-encephalography recordings taken from the vmPFC/mOFC in human epilepsy patients, albeit with a positive modulation by subjective value (Lopez-Persem et al., 2020).

Not only does the vmPFC/mOFC carry signals that reflect a translation of values into choices, but these vmPFC/mOFC signals are necessary for decision-making. Lesions in the vmPFC/mOFC, both in macaques and humans (Camille et al., 2011; Noonan et al., 2010, 2017), and manipulation of human vmPFC/mOFC activity using transcranial direct current stimulation (Hämmerer et al., 2016) increase choice stochasticity and reduce the accuracy of the choice comparison process. This is in line with work showing that individual variation in the excitation-inhibition balance in the vmPFC/mOFC's activity is directly related to individual variation in choice stochasticity (Jocham et al., 2012). By contrast, ultrasonic disruption of activity in the amFC does not affect choice stochasticity (Bongioanni et al., 2021). Instead, it alters abstract value space representations, which suggests differences in the functional roles of amFC and vmPFC/mOFC. Nevertheless, vmPFC/mOFC activity may not always be required to select choices. As choices become more familiar, they rely less on online comparison processes and more on simpler heuristics and precomputed values. This in turn makes choice value signals become weaker or disappear entirely from the vmPFC/mOFC in monkeys and humans (Bongioanni et al., 2021; Hunt et al., 2012).

In macaques, perhaps the best characterized population of neurons with a role in decision-making is situated even more laterally on the orbital surface, in area 13. Here, individual neurons have been identified with activity that is selective for specific options in reward-guided decisions. In an important series of studies, Padoa-Schioppa and colleagues examined decisions between visual stimuli associated with different types of juice. For example, in one experiment, blue and yellow squares indicated water or unsweetened kool-aid. Increments in the number of stimulus elements—i.e., the number of squares—indicated increments in the amount of that juice type available. "Offer value" neurons were selective for particular stimuli/juice types, but their firing rates changed with the amount available (Cai and Padoa-Schioppa, 2014; Padoa-Schioppa and Conen, 2017). The activity distributed across the population of neurons in and near this area encodes the identities of potential choice options, and during the course of decisions, it is possible to track the relative strengths of the representations of the potential choices. The relative strengths of representations may change repeatedly during the course of a decision, especially when the options are close in value, but eventually one comes to predominate and the choice is taken (Bongioanni et al., 2021; Hunt et al., 2018; Klein-Flügge et al., 2013; Rich and Wallis, 2016). Selective and focused inactivation or stimulation of area 13 alone is sufficient to interfere with the way in which value-guided decisions are made (Ballesta et al., 2020; Murray et al., 2015). Understanding how area 13 and the more medial vmPFC/mOFC area discussed above collaborate or specialize

during decision-making remains an ongoing topic of discussion.

VmPFC/mOFC value comparison signals reflect many other influences that impact on the way that options are valued during decision-making. For instance, the presence of a third option may impact the way that two other options are valued and compared and thus affects which choice is likely to be taken (Chau et al., 2020; Dumbalska et al., 2020; Louie et al., 2015; Webb et al., 2020). In addition, the presence of a less valuable item within a compound option is known to reduce the estimated value of the compound relative to the more valuable item alone, in the "less-is-more" effect displayed by both human and non-human primates (Kralik et al., 2012; List, 2002). Both of these phenomena are reflected in the vmPFC/mOFC's choice comparison signal (Chau et al., 2014; Fouragnan et al., 2019; Lim et al., 2011; Papageorgiou et al., 2017; Suzuki et al., 2017) and both phenomena are disrupted by lesions in the vmPFC/OFC (Noonan et al., 2010, 2017; Papageorgiou et al., 2017). When satiety or background context change the way that options are valued, then vmPFC activity reflects changes in the way that the options will be valued even before any decision is made (Abitbol et al., 2015).

However, many factors other than the reward value of a choice influence whether it will be taken. Such influences include, for example, the recent reward rate regardless of which choice was taken, choice traces (the history of which choices were taken recently, regardless of whether they were rewarded), the number of offers viewed, the attended location, the sense of social controllability, the confidence and uncertainty related to a choice, or the value of upcoming information. All these variables have been shown to affect vmPFC activity at the time of decision-making (Kaanders et al., 2021; Leong et al., 2017; Mehta et al., 2019; Na et al., 2021; Trudel et al., 2021; Wittmann et al., 2020). One important consideration seems to be the policy that is currently guiding behavior. In a recent study, Trudel et al. (2021) demonstrated that vmPFC/mOFC activity displays an impressive degree of flexibility and that the same variable can be associated with either a positive or a negative change in vmPFC/mOFC activity, depending on the goal of the decision. In their study, optimal decisions required periods of exploration and periods of exploitation. Not only did vmPFC/mOFC BOLD reflect the uncertainty as well as the value of choices, but vmPFC/mOFC BOLD signatures of choice uncertainty flipped sign depending on context, with negative uncertainty coding during exploitation (when participants were selecting options that they were certain were high in value) and positive uncertainty coding during exploration (when participants were selecting options that they were uncertain about in order to find out more about their value). Such a change is consistent with the existence of not only neural circuits specialized for exploration but also neural systems mediating both reward exploration and exploitation (Costa et al., 2019).

In summary, the vmPFC/mOFC represents potential choice options, computes their comparison, and turns them into actual choices in the frame of reference currently relevant for guiding actions (Grueschow et al., 2015). More generally, and in contrast to the dACC, vmPFC/mOFC activity reflects the relative evidence for taking one choice over another, along the multiple dimensions and in the frame of reference relevant for the choice at hand. In many cases, this might mean that it reflects choice values, but other variables might be represented, depending on current action policies (Hayden and Niv, 2021).

The boundary between the amFC and vmPFC/mOFC remains to be precisely defined (Figures 5C and 5D). Many studies on reward-guided decision-making find activations in both locations (Bartra et al., 2013; Clithero and Rangel, 2014) or at the border (Schuck et al., 2016). This might be because these tasks often require simulations of novel option values or states as well as choice computations, meaning the two processes occur simultaneously and may not be easily teased apart. More evidence to support a dissociation between abstract structure representations in the amFC, and the turning of such representations into decisions in vmPFC/mOFC, will therefore be needed. However, a recent study by Park et al. (2021) provided the first compelling test for this. In their task, human participants were required to represent abstract relationships between different individuals along two dimensions (popularity and competence). When participants had been trained on these relationships, they were asked to make binary choices about which of two individuals would be a better partner for a third individual while undergoing fMRI. Making such a choice required representing the popularity and competence of each of the three individuals on the one hand and, on the other hand, a computation of the combined strength or "growth potential" of each two-person team. A hexagonal modulation of the BOLD signal in the amFC indicated grid-like coding of the growth potential in the abstract social space spanned by all possible individuals, while the BOLD signal related to the decision variable—the difference between teams' growth potential—was located more ventrally in the vmPFC/mOFC (Figures 5C and 5D). This provides compelling evidence, within the same task, that abstract grid-like representations of relevant relationships are present in the amFC and converted into decisions in the vmPFC/mOFC (Park et al., 2021).

As is the case for other frontal cortex regions, it is not always clear how the activity patterns that emerge in the vmPFC/mOFC during decision-making can be distinguished from the activity patterns seen in subcortical structures such as the ventral striatum and DA midbrain. Both of these structures show activity related to choice selection that resembles that seen in the vmPFC/mOFC and precedes it in time (Strait et al., 2015; Yun et al., 2020). So far, however, at least within the DA midbrain, such activity has been recorded in very simple situations in which monkeys are presented with a single opportunity and the decision is whether or not to engage with that option rather than a process of comparison between two or more options. It is possible that representations in the vmPFC/mOFC as opposed to the DA midbrain are especially important when the decision to be made is new and linked to inferential processes in the adjacent amFC.

Perigenual cingulate cortex and cost-benefit arbitration

While the amFC encodes task structure even when this does not involve value or reward, a region anterior to the genu of the corpus callosum and slightly posterior to amFC, the perigenual anterior cingulate cortex or pgACC, integrates costs and benefits to evaluate the overall value of initiating an action. Amemori

A pgACC and cost-benefit arbitration
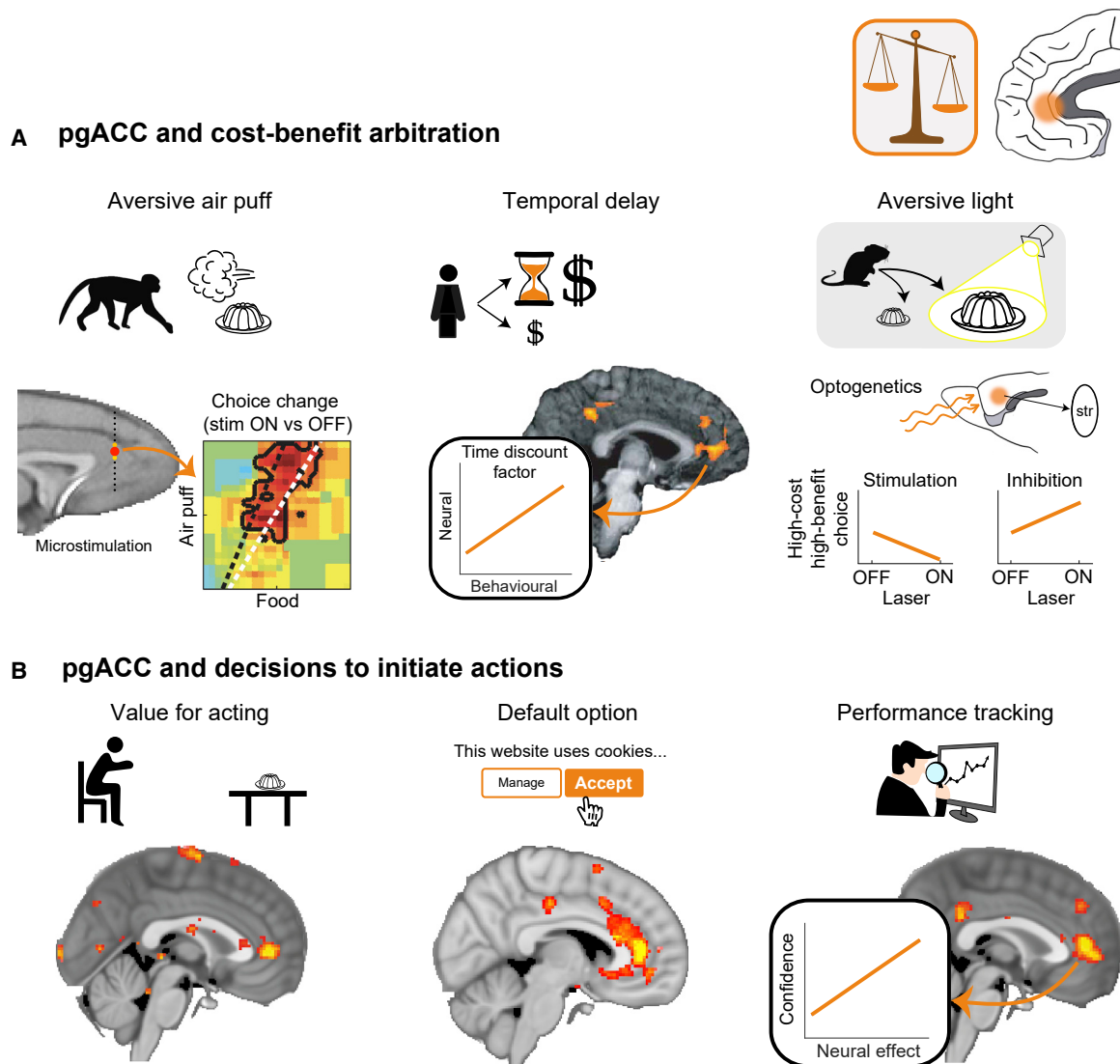
Figure 6. pgACC and the costs and benefits of initiating a course of action

(A) pgACC evaluates the cost and benefits of initiating a course of action. (Left) Neurons in macaque pgACC reflect the benefit (juice) and cost (air puff) of taking a choice. Microstimulation in the area with a predominance of aversive cost-related neurons led to a change in the decision boundary for taking the action (adapted from Amemori and Graybiel, 2012). (Center) Activity in human pgACC reflects each individual's subjective valuation of an opportunity composed of both a monetary reward and a temporal delay cost (adapted from Kable and Glimcher, 2007). (Right) Rats decided between pursuing a large food reward associated with an aversive cost (a bright light) or a small reward in a less aversive, darker environment. Optogenetic activation and inhibition of a pathway from the pgACC-like PL area to inhibitory interneurons in the striosomes of the striatum led to decrements and increments in high-value/high-cost choices, respectively (adapted from Friedman et al., 2015).

(B) pgACC is related to action initiation. (Left) In humans, individual variation in pgACC activity is predictive of whether or not a choice will be pursued (adapted from Kolling et al., 2018). (Center) Activity in human pgACC reflects the relative average preference for a default choice type; for example, people may be *a priori* more likely to accept website cookies than to manage them or to take a sweet rather than a savory snack regardless of the specific sweet and savory snacks they are offered (adapted from Lopez-Persem et al., 2016). (Right) pgACC tracks expected performance on a perceptual task, and this effect scales with subjective decision confidence (adapted from Bang and Fleming, 2018).

and colleagues (Amemori and Graybiel, 2012) recorded neural activity from the pgACC of macaque monkeys while they evaluated cues that were simultaneously associated with varying levels of air-puff (cost) and liquid food reward (benefit). When monkeys chose to approach the cue, they received both the associated air-puff and liquid reward; when they avoided the cue, they received neither outcome (Figure 6A, left). The mon-

keys' choices indicated that they were influenced by both the cost and benefit associated with an offer. Simultaneous recordings from pgACC neurons revealed different activity but, in summary, firing rates were best explained as reflecting the overall utility of the chosen outcome. In other words, firing rates showed an integration across the cost and benefit dimensions of the cue. Furthermore, microstimulation of the pgACC produced changes

in the animals' cost-benefit decisions such that they were more likely, on average, to avoid rather than approach the cue. Importantly, microstimulation was most effective on trials where the choice required a trade-off between costs and benefits, and thus when the positive and negative motivational aspects of the cue competed to drive behavior in opposite directions.

The idea that the pgACC is crucial for integrating costs and benefits to decide whether it is worth initiating an action is consistent with work in other species, including humans and rodents, all of whom share this agranular part of the PFC. In humans, the pgACC BOLD signals were shown to reflect integrated cost-benefit value in a delay-based decision-making task, where a larger delayed reward offer could be accepted or foregone for a small reward received immediately (Kable and Glimcher, 2007). In this task, the pgACC BOLD was better explained by the integrated subjective value of the delayed option than by the reward amount or the delay considered separately (Figure 6A, center). Consistent with this, causal evidence from lesion experiments in rats shows that cost-benefit integrations are impaired following lesions that include the pgACC. Walton et al. (2002) trained rats to choose between two arms of a T-maze, one of which was associated with a higher reward and a cost, climbing a barrier, while the other resulted in a smaller reward but did not require climbing a barrier. Although rewards in the high-effort arm could be adjusted such that healthy rats generally preferred the high-effort/high-reward arm, following a lesion that included pgACC, the same rats were less likely to choose the high-effort arm, even though they had no problem with climbing a barrier or with choosing the high-reward option when both arms of the T-maze included a barrier (Walton et al., 2002).

In work by Friedman et al. (2015), the trade-off between costs and benefits was shown to rely on the pgACC's projections to specialized regions in the striatum, the striosome (a component of the subcortical circuit for reward-guided behavior illustrated in Figure 2A). Friedman et al. optogenetically targeted pgACC cells projecting to the striosome in a mouse T-maze task like that used by Walton and colleagues, except that the cost in this task was the overcoming of the mice's instinctual aversion to a bright light instead of a barrier (Figure 6A, right). Friedman and colleagues found that inhibition and excitation of the pgACC-striosome pathway induced shifts in mice behavior, leading to an increase or decrease in choosing the high-cost/high-reward option, respectively. This effect was pathway-specific and specific to the cost-benefit condition and thus, situations where the net outcome entailed motivationally conflicting positive and negative components that had to be integrated to make a choice. It is notable that in primates too, the pgACC is very unusual in having a projection to the striosome; within the frontal cortex, only the pgACC and a pOFC region, on the border with the insula, project to the striosome (Eblen and Graybiel, 1995).

Whether the precise type of cost determines pgACC's involvement in decision-making remains to be clarified. In the above, we have discussed work linking the pgACC to action initiation and cost-benefit integration across a wide range of costs: aversive bright lights (Friedman et al., 2015) and air-puffs (Amemori and Graybiel, 2012), effort costs of climbing a barrier (Walton et al., 2002), as well as delays (Kable and Glimcher, 2007). By contrast, computations reflecting the direct comparison of effort- and

reward-linked choice options have been associated with a more dorsal posterior cingulate area in humans (Klein-Flügge et al., 2016), and the processing of different types of costs occurs in dissociable neural circuits, some of which are distinct from the pgACC (Bonnelle et al., 2016; Burke et al., 2013; Croxson et al., 2009; Kennerley et al., 2009; Kurniawan et al., 2013; Prévost et al., 2010; Rudebeck et al., 2006; Scholl et al., 2015; Walton et al., 2003). Although specific types of costs may be processed in separate subregions of the PFC, the pgACC seems to be crucial for integrating the motivational value of an outcome across costs and benefits to initiate or avoid initiating an approach behavior (Figure 6B). In ecological environments where opportunities typically arise sequentially, foraging animals frequently encounter this type of decision about whether to engage with a particular opportunity, given its costs and benefits. In laboratory-based tasks for humans that capture aspects of such scenarios (Figure 6B, left), individual variation in pgACC activity and pgACC connectivity with striosome-rich parts of the striatum predicts individual variation in whether behavior will be determined by the potential benefits that might ensue from the course of action, despite increasing costs, and thus how likely participants are to proceed with taking the course of action (Kolling et al., 2012, 2018). The pgACC's anatomical position and connections to regions beyond the striosome, such as with the subgenual ACC, amygdala, habenula, neuromodulatory systems, and periaqueductal gray (An et al., 1998; Chiba et al., 2001), place it in an ideal position to provide motivational regulation of the initiation of actions, perhaps especially approach and avoidance choices (Khalighinejad et al., 2020a, 2020b, 2021). In humans, pgACC activity reflects the value of default options that people are most likely to go ahead and choose (Figure 6B, center) (Lopez-Persem et al., 2016). When the pgACC is absent or not providing an input to the striosomes, then animals still initiate actions, but they do not always initiate them in the situations in which they had judged them to be worthwhile in the control condition. Although the subcortical circuitry that the pgACC projects to is sufficient for action initiation, it may not be sufficient for determining when the balance of costs and benefits suggests it is best to initiate action.

In order to decide whether to engage with an opportunity, it is important to track the success of recent engagements with opportunities or, in other words, to track one's own recent performance (Figure 6B, right). In humans, the pgACC carries signals consistent with the monitoring of one's own performance over both shorter and more extended time frames (Bang and Fleming, 2018; Wittmann et al., 2016b). For example, Wittmann and colleagues reported that pgACC activity reflected the feedback human participants received about their performance levels on a variety of arbitrary games, and it predicted the influence that the feedback would have on the participants' estimations of their ability levels. One recent hypothesis is that this type of self-awareness may be altered in mood and anxiety disorders known to implicate the pgACC (Amemori et al., 2021).

The pgACC area that we discuss in this section is close to the amFC and vmPFC; determining the precise border in relation to landmarks such as the cingulate sulcus is important but still a matter of debate. Moreover, it is possible that these regions co-activate when multidimensional features need to be

Neuron
Review

CellPress
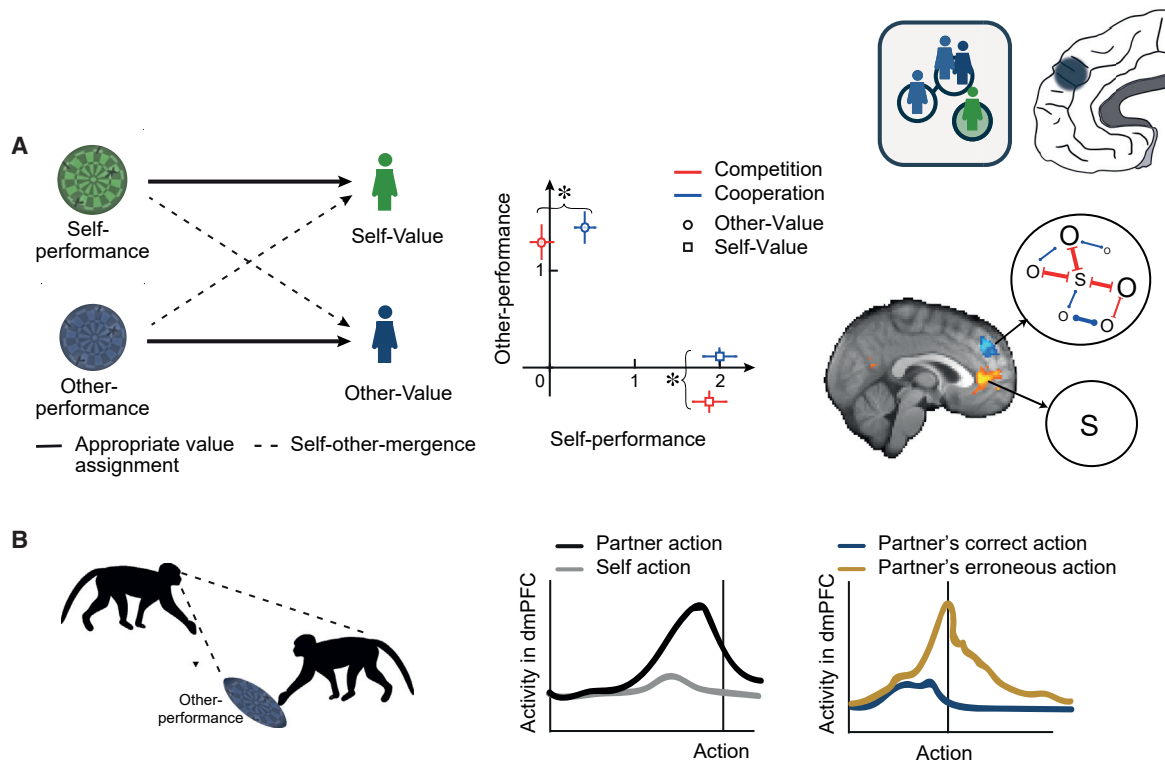
Figure 7. dmPFC and the organization of inter-individual relationships

(A) (Left) In a game of multi-player competition and cooperation, people track not just their own performances in order to form an estimate of their ability, but these estimates are also influenced by the performances of the people around them. In a complementary fashion, people's estimates of others' abilities are also influenced by their own performances. These two ways of inappropriately ascribing value to oneself or another person are referred to as self-other-mergence. (Center) The direction of self-other-mergence effects depends on context (cooperation versus competition). For example, a good partner (high other-performance) boosts self-value in cooperation but diminishes self-value in competition (error bars denote SEM across participants). (Right) Although pgACC tracks their own performance levels (see also Figure 6B), dmPFC tracks both the influence that other players have on the self-performance estimate and the estimate that the self's own performance has on estimates of the other players (adapted from Wittmann et al., 2016b, 2018, 2021).

(B) When one monkey watches another monkey in order to work out which choice is the better one to take, neural activity in dmPFC tracks the partner monkey's actions and distinguishes between situations when the partner makes erroneous versus correct actions, allowing the observing animal to learn which actions not to repeat (adapted from Yoshida et al., 2011, 2012).

integrated in order to derive a choice value estimate (Bongioanni et al., 2021), and of course all these areas and the dACC might be expected to interact with one another and with areas beyond the frontal cortex (Klein et al., 2017; Korn and Bach, 2018, 2019; Maier et al., 2015). However, when this integration concerns more abstract features and inferential processes, the peak is more anterior (in the amFC), but when the integration is between costs and benefits relating to a specific action that might or might not be taken, the peak is more posterior (in the pgACC) (Amemori and Graybiel, 2012; Khalighinejad et al., 2020a). Other studies and reviews also confirm a functional difference between the pgACC and more anterior brain areas (Grabenhorst and Rolls, 2011).

Dorsomedial frontal cortex and the organization of interpersonal relationships

So far, we have seen that activity in the amFC encodes the structure and organization of the task environment. Alongside their ability to learn about a multitude of arbitrary task environments, humans and many other primates spend a large part of their time navigating one particular type of environment—the social environment. The most dorsal part of medial frontal cortex—the dmPFC—may have a specialized role in encoding key features of the structure of the social environment and the position of the decision maker within this social environment.

Social environments share many features with other arbitrary task environments. The diverse and ever-changing patterns of the social structures in which humans dwell attest to their arbitrariness. However, at the same time, there are also features of social environments, such as competition and cooperation, which are consistently present. Competition and cooperation have an important impact on the individual animal's or person's fitness, health, and longevity. For example, in macaques, competition and collaboration are important predictors of social dominance and, in turn, these are predictors of breeding success (Schülke et al., 2010). In humans, loneliness—social isolation and absence of cooperation—has a major impact on mortality (Holt-Lunstad et al., 2010, 2015).

Patterns of competition and collaboration are rarely static for long. Despite this temporal complexity, an important feature of many social environments is that, regardless of their arbitrariness, the position of the self within the environment is a key
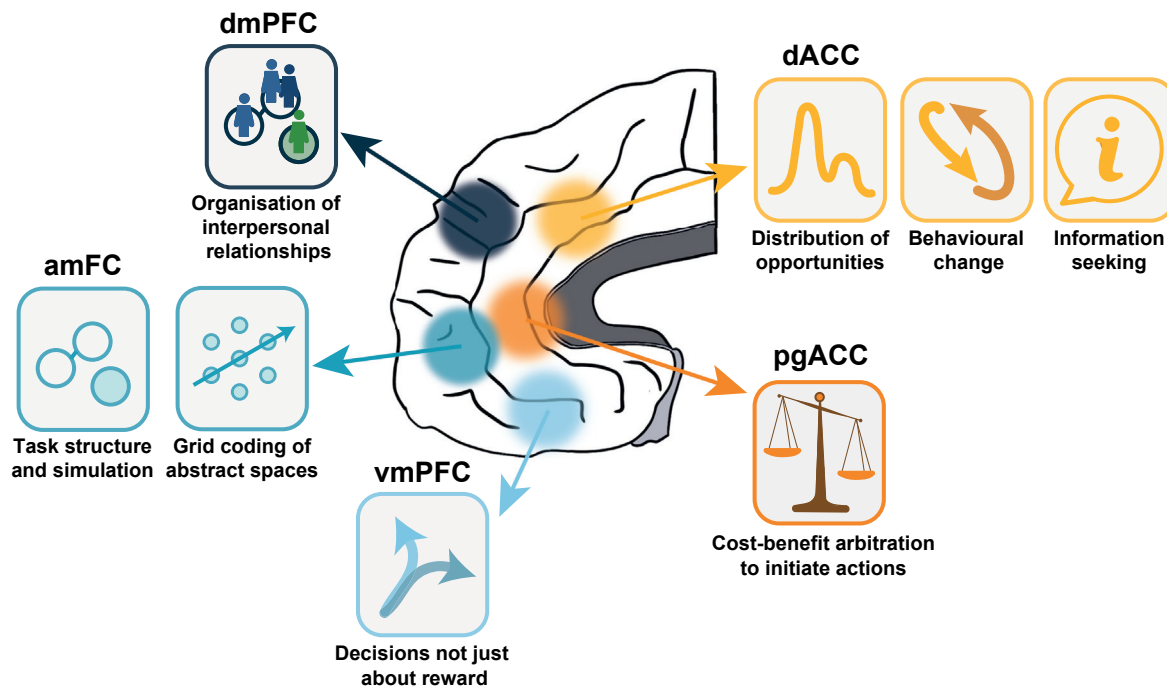
Figure 8. Summary of functional specializations
Schematic overview of the functional contributions of different subregions of medial prefrontal cortex discussed in this review.

anchor or origin for the task space. Wittmann and colleagues (Wittmann et al., 2016b, 2021) investigated how changing patterns of competition and cooperation between participant and other pre-programmed "players" are tracked over time during a series of simple games. On each trial, participants performed a simple task and received feedback about how well they and two other players had done. On some trials the participants co-operated with one of the other players—the sum of their performances determined a payoff. On other trials, however, they competed—now the differences in their performances determined payoff. Not surprisingly, the participants' assessments of their own performances reflected the feedback they received about their performances (Figure 7A, left and center). Moreover, as noted above, such feedback was associated with pgACC activity and individual variation in its impact on the pgACC was associated with individual variation in the impact it had on self-assessment.

Their evaluations of themselves, however, also varied as a function of the other players' performances. When they *cooperated* with good players, they rated their own performances as stronger and when they *competed* with good players, they rated their own performances as weaker, and vice versa for weak players (Figure 7A, left and center). Wittmann and colleagues called this phenomenon self-other mergence and found that it was dependent on the dmPFC; activity there tracked the performance of the other player and predicted its impact on self-performance estimates. In addition, it had a complementary effect; dmPFC activity also tracked self-performance and the impact that one's own performance had on the estimation of the other player (Figure 7A, right). However, self-other mergence is not a

simple consequence of dmPFC activity; when the dmPFC is disrupted by transcranial magnetic stimulation (TMS) then self-other mergence is augmented, even if its correlation with neural activity is abolished (Wittmann et al., 2021). This pattern of change suggests that the dmPFC may be critical for disentangling and tracking what each agent is doing. When this is disrupted by dmPFC TMS, then people appear to track the aggregate consequence of the competitive or cooperative interaction at the expense of the individual performances.

In monkeys, dmPFC activity also tracks the performances of other individuals; dmPFC activity reflects both the choices that other animals make and the rewards that they receive for making them (Ninomiya et al., 2020; Noritake et al., 2018; Yoshida et al., 2011, 2012; Figure 7B). In a manner reminiscent of human self-other mergence, the social context has an impact on the way in which macaques evaluate the choices that are available and the consequences that will follow. For example, anticipatory licking measures suggest macaques value stimuli more if they are associated with more reward for the macaque itself, but they value stimuli less if they are associated with more reward for other animals. Although some dmPFC neurons track a stimulus' association with reward for the individual itself, others track the stimulus' association with reward for the other monkey. Such activity in dmPFC neurons precedes activity in other brain structures, such as the DA midbrain, which reflects the monkey's evaluation of its own reward prospects given the context of the social environment (Noritake et al., 2018). In the context of competitive games, activity in the dmPFC also reflects the monkeys' adoption of response selection strategies that mean that each decision they take is difficult to predict

from previous decisions (Seo et al., 2014). As a result, the monkey's decisions are difficult for other individuals in the group to predict and pre-empt. Recording of human dmPFC neurons suggest that it not only tracks the behaviors of other individuals but also, at least in humans, the beliefs of other individuals (Jamali et al., 2021).

It is difficult to determine the degree to which the dmPFC is exclusively concerned with social structure and social decision-making. On the one hand, the amFC activity that is recorded in tasks lacking any simple social component extends into the dmPFC (Barron et al., 2013; Constantinescu et al., 2016). On the other hand, social tasks that involve tracking other individuals' thoughts or behaviors typically activate the dmPFC only (Behrens et al., 2008; Frith and Frith, 2012; Konovalov et al., 2021; Wittmann et al., 2016b, 2021).

SUMMARY

Neural systems that represent the value of the environment exist in many vertebrates. An extended subcortical circuit spanning the striatum, midbrain, and brainstem nuclei of mammals corresponds to these ancient systems. In addition, however, mammals possess several frontal cortical regions concerned with guidance of decision-making and adaptive, flexible behavior. Although these frontal systems interact extensively with these subcortical circuits, they make specific contributions to behavior and also influence behavior via other cortical routes. Some areas such as the ACC, which is present in a broad range of mammals, represent the distribution of opportunities in an environment over space and time, whereas other brain regions such as amFC and dmPFC have roles in representing structural associations and causal links between environmental features, including aspects of the social environment (Figure 8). Although the origins of these areas and their functions are traceable to rodents, they are especially prominent in primates. They make it possible not just to select choices on the basis of past experience of identical situations, but to make inferences to guide decisions in new scenarios.

SUPPLEMENTAL INFORMATION

ACKNOWLEDGMENTS

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Abitbol, R., Lebreton, M., Hollard, G., Richmond, B.J., Bouret, S., and Pessiglione, M. (2015). Neural mechanisms underlying contextual dependency of subjective values: converging evidence from monkeys and humans. J. Neurosci. 35, 2308–2320. https://doi.org/10.1523/JNEUROSCI.1878-14.2015.

Agetsuma, M., Aizawa, H., Aoki, T., Nakayama, R., Takahoko, M., Goto, M., Sassa, T., Amo, R., Shiraki, T., Kawakami, K., et al. (2010). The habenula is crucial for experience-dependent modification of fear responses in zebrafish. Nat. Neurosci. 13, 1354–1356. https://doi.org/10.1038/nn.2654.

Amemori, K., and Graybiel, A.M. (2012). Localized microstimulation of primate pregenual cingulate cortex induces negative decision-making. Nat. Neurosci. 15, 776–785. https://doi.org/10.1038/nn.3088.

Amemori, S., Graybiel, A.M., and Amemori, K.I. (2021). Causal evidence for induction of pessimistic decision-making in primates by the network of frontal cortex and striosomes. Front. Neurosci. 15, 649167. https://doi.org/10.3389/fnins.2021.649167.

Amo, R., Fredes, F., Kinoshita, M., Aoki, R., Aizawa, H., Agetsuma, M., Aoki, T., Shiraki, T., Kakinuma, H., Matsuda, M., et al. (2014). The habenulo-raphe serotonergic circuit encodes an aversive expectation value essential for adaptive active avoidance of danger. Neuron 84, 1034–1048. https://doi.org/10.1016/j.neuron.2014.10.035.

An, X., Bandler, R., Ongür, D., and Price, J.L. (1998). Prefrontal cortical projections to longitudinal columns in the midbrain periaqueductal gray in macaque monkeys. J. Comp. Neurol. 401, 455–479.

Ashwell, K.W.S., McAllan, B.M., Mai, J.K., and Paxinos, G. (2008). Cortical cyto- and chemoarchitecture in three small Australian marsupial carnivores: Sminthopsis macroura, Antechinus stuartii and Phascogale calura. Brain Behav. Evol. 72, 215–232. https://doi.org/10.1159/000165101.

Badre, D., Doll, B.B., Long, N.M., and Frank, M.J. (2012). Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration. Neuron 73, 595–607. https://doi.org/10.1016/j.neuron.2011.12.025.

Ballesta, S., Shi, W., Conen, K.E., and Padoa-Schioppa, C. (2020). Values encoded in orbitofrontal cortex are causally related to economic choices. Nature 588, 450–453. https://doi.org/10.1038/s41586-020-2880-x.

Banerjee, A., Parente, G., Teutsch, J., Lewis, C., Voigt, F.F., and Helmchen, F. (2020). Value-guided remapping of sensory cortex by lateral orbitofrontal cortex. Nature 585, 245–250. https://doi.org/10.1038/s41586-020-2704-z.

Bang, D., and Fleming, S.M. (2018). Distinct encoding of decision confidence in human medial prefrontal cortex. Proc. Natl. Acad. Sci. USA 115, 6082–6087. https://doi.org/10.1073/pnas.1800795115.

Bao, X., Gjorgieva, E., Shanahan, L.K., Howard, J.D., Kahnt, T., and Gottfried, J.A. (2019). Grid-like neural representations support olfactory navigation of a two-dimensional odor space. Neuron 102, 1066–1075.e5. https://doi.org/10.1016/j.neuron.2019.03.034.

Barack, D.L., Chang, S.W.C., and Platt, M.L. (2017). Posterior cingulate neurons dynamically signal decisions to disengage during foraging. Neuron 96, 339–347.e5. https://doi.org/10.1016/j.neuron.2017.09.048.

Barack, D.L., and Platt, M.L. (2021). Neuronal activity in the posterior cingulate cortex signals environmental information and predicts behavioral variability during trapline foraging. J. Neurosci. 41, 2703–2712. https://doi.org/10.1523/JNEUROSCI.0305-20.2020.

Baram, A.B., Muller, T.H., Nili, H., Garvert, M.M., and Behrens, T.E.J. (2021). Entorhinal and ventromedial prefrontal cortices abstract and generalize the structure of reinforcement learning problems. Neuron 109, 713–723.e7. https://doi.org/10.1016/j.neuron.2020.11.024.

Barron, H.C., Dolan, R.J., and Behrens, T.E. (2013). Online evaluation of novel choices by simultaneous representation of multiple memories. Nat. Neurosci. 16, 1492–1498. https://doi.org/10.1038/nn.3515.

Bartolo, R., and Averbeck, B.B. (2020). Prefrontal cortex predicts state switches during reversal learning. Neuron 106, 1044–1054.e4. https://doi.org/10.1016/j.neuron.2020.03.024.

Bartra, O., McGuire, J.T., and Kable, J.W. (2013). The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. Neuroimage 76, 412–427. https://doi.org/10.1016/j.neuroimage.2013.02.063.

Behrens, T.E., Hunt, L.T., Woolrich, M.W., and Rushworth, M.F. (2008). Associative learning of social value. Nature 456, 245–249. https://doi.org/10.1038/nature07538.

Behrens, T.E., Woolrich, M.W., Walton, M.E., and Rushworth, M.F. (2007). Learning the value of information in an uncertain world. Nat. Neurosci. *10*, 1214–1221. https://doi.org/10.1038/nn1954.

Behrens, T.E.J., Muller, T.H., Whittington, J.C.R., Mark, S., Baram, A.B., Stachenfeld, K.L., and Kurth-Nelson, Z. (2018). What is a cognitive map? Organizing knowledge for flexible behavior. Neuron *100*, 490–509. https://doi.org/10.1016/j.neuron.2018.10.002.

Bernacchia, A., Seo, H., Lee, D., and Wang, X.J. (2011). A reservoir of time constants for memory traces in cortical neurons. Nat. Neurosci. *14*, 366–372. https://doi.org/10.1038/nn.2752.

Blanchard, T.C., and Hayden, B.Y. (2014). Neurons in dorsal anterior cingulate cortex signal postdecisional variables in a foraging task. J. Neurosci. *34*, 646–655. https://doi.org/10.1523/JNEUROSCI.3151-13.2014.

Bongioanni, A., Folloni, D., Verhagen, L., Sallet, J., Klein-Flügge, M.C., and Rushworth, M.F.S. (2021). Activation and disruption of a neural mechanism for novel choice in monkeys. Nature *591*, 270–274. https://doi.org/10.1038/s41586-020-03115-5.

Bonnelle, V., Manohar, S., Behrens, T., and Husain, M. (2016). Individual differences in premotor brain systems underlie behavioral apathy. Cereb. Cortex *26*, 807–819. https://doi.org/10.1093/cercor/bhv247.

Boorman, E.D., Behrens, T.E., Woolrich, M.W., and Rushworth, M.F. (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. https://doi.org/10.1016/j.neuron.2009.05.014.

Boorman, E.D., Rajendran, V.G., O'Reilly, J.X., and Behrens, T.E. (2016). Two anatomically and computationally distinct learning signals predict changes to stimulus-outcome associations in hippocampus. Neuron *89*, 1343–1354. https://doi.org/10.1016/j.neuron.2016.02.014.

Boorman, E.D., Rushworth, M.F., and Behrens, T.E. (2013). Ventromedial prefrontal and anterior cingulate cortex adopt choice and default reference frames during sequential multialternative choice. J. Neurosci. *33*, 2242–2253.

Bromberg-Martin, E.S., and Hikosaka, O. (2009). Midbrain dopamine neurons signal preference for advance information about upcoming rewards. Neuron *63*, 119–126. https://doi.org/10.1016/j.neuron.2009.06.009.

Burke, C.J., Brünger, C., Kahnt, T., Park, S.Q., and Tobler, P.N. (2013). Neural integration of risk and effort costs by the frontal pole: only upon request. J. Neurosci. *33*, 1706. 113a. https://doi.org/10.1523/JNEUROSCI.3662-12.2013.

Bush, D., Barry, C., Manson, D., and Burgess, N. (2015). Using grid cells for navigation. Neuron *87*, 507–520. https://doi.org/10.1016/j.neuron.2015.07.006.

Cai, X., and Padoa-Schioppa, C. (2014). Contributions of orbitofrontal and lateral prefrontal cortices to economic choice and the good-to-action transformation. Neuron *81*, 1140–1151. https://doi.org/10.1016/j.neuron.2014.01.008.

Camille, N., Griffiths, C.A., Vo, K., Fellows, L.K., and Kable, J.W. (2011). Ventromedial frontal lobe damage disrupts value maximization in humans. J. Neurosci. *31*, 7527–7532. https://doi.org/10.1523/JNEUROSCI.6527-10.2011.

Cavanagh, S.E., Wallis, J.D., Kennerley, S.W., and Hunt, L.T. (2016). Autocorrelation structure at rest predicts value correlates of single neurons during reward-guided choice. Elife *5*, e18937. https://doi.org/10.7554/eLife.18937.

Charnov, E.L. (1976). Optimal foraging: the marginal value theorem. Theor. Popul. Biol. *9*, 129–136.

Chau, B.K., Kolling, N., Hunt, L.T., Walton, M.E., and Rushworth, M.F. (2014). A neural mechanism underlying failure of optimal choice with multiple alternatives. Nat. Neurosci. *17*, 463–470. https://doi.org/10.1038/nn.3649.

Chau, B.K., Law, C.-K., Lopez-Persem, A., Klein-Flügge, M.C., and Rushworth, M.F. (2020). Consistent patterns of distractor effects during decision making. Elife *9*, e53850. https://doi.org/10.7554/eLife.53850.

Chau, B.K., Sallet, J., Papageorgiou, G.K., Noonan, M.P., Bell, A.H., Walton, M.E., and Rushworth, M.F. (2015). Contrasting roles for orbitofrontal cortex and amygdala in credit assignment and learning in macaques. Neuron *87*, 1106–1118. https://doi.org/10.1016/j.neuron.2015.08.018.

Chen, D., Kunz, L., Lv, P., Zhang, H., Zhou, W., Liang, S., Axmacher, N., and Wang, L. (2021). Theta oscillations coordinate grid-like representations between ventromedial prefrontal and entorhinal cortex. Sci. Adv. *7*, eabj0200. https://doi.org/10.1126/sciadv.abj0200.

Chiba, T., Kayahara, T., and Nakano, K. (2001). Efferent projections of infralimbic and prelimbic areas of the medial prefrontal cortex in the Japanese monkey, Macaca fuscata. Brain Res *888*, 83–101.

Clithero, J.A., and Rangel, A. (2014). Informatic parcellation of the network involved in the computation of subjective value. Soc. Cogn. Affect. Neurosci. *9*, 1289–1302. https://doi.org/10.1093/scan/nst106.

Cohen, J.Y., Amoroso, M.W., and Uchida, N. (2015). Serotonergic neurons signal reward and punishment on multiple timescales. Elife *4*, e06346. https://doi.org/10.7554/eLife.06346.

Constantinescu, A.O., O'Reilly, J.X., and Behrens, T.E.J. (2016). Organizing conceptual knowledge in humans with a gridlike code. Science *352*, 1464–1468. https://doi.org/10.1126/science.aaf0941.

Costa, V.D., Mitz, A.R., and Averbeck, B.B. (2019). Subcortical substrates of explore-exploit decisions in primates. Neuron *103*, 533–545.e5. https://doi.org/10.1016/j.neuron.2019.05.017.

Croxson, P.L., Walton, M.E., O'Reilly, J.X., Behrens, T.E., and Rushworth, M.F. (2009). Effort-based cost-benefit valuation and the human brain. J. Neurosci. *29*, 4531–4541. https://doi.org/10.1523/JNEUROSCI.4515-08.2009.

Dabney, W., Kurth-Nelson, Z., Uchida, N., Starkweather, C.K., Hassabis, D., Munos, R., and Botvinick, M. (2020). A distributional code for value in dopamine-based reinforcement learning. Nature *577*, 671–675. https://doi.org/10.1038/s41586-019-1924-6.

Danielmeier, C., Allen, E.A., Jocham, G., Onur, O.A., Eichele, T., and Ullsperger, M. (2015). Acetylcholine mediates behavioral and neural post-error control. Curr. Biol. *25*, 1461–1468. https://doi.org/10.1016/j.cub.2015.04.022.

Daw, N.D., Gershman, S.J., Seymour, B., Dayan, P., and Dolan, R.J. (2011). Model-based influences on humans' choices and striatal prediction errors. Neuron *69*, 1204–1215. https://doi.org/10.1016/j.neuron.2011.02.027.

Daw, N.D., O'Doherty, J.P., Dayan, P., Seymour, B., and Dolan, R.J. (2006). Cortical substrates for exploratory decisions in humans. Nature *441*, 876–879.

De Martino, B., Fleming, S.M., Garrett, N., and Dolan, R.J. (2013). Confidence in value-based choice. Nat. Neurosci. *16*, 105–110.

Doeller, C.F., Barry, C., and Burgess, N. (2010). Evidence for grid cells in a human memory network. Nature *463*, 657–661. https://doi.org/10.1038/nature08704.

Donoso, M., Collins, A.G., and Koechlin, E. (2014). Human cognition. Foundations of human reasoning in the prefrontal cortex. Science *344*, 1481–1486. https://doi.org/10.1126/science.1252254.

Dumbalska, T., Li, V., Tsetsos, K., and Summerfield, C. (2020). A map of decoy influence in human multialternative choice. Proc. Natl. Acad. Sci. USA *117*, 25169–25178. https://doi.org/10.1073/pnas.2005058117.

Eblen, F., and Graybiel, A.M. (1995). Highly restricted origin of prefrontal cortical inputs to striosomes in the macaque monkey. J. Neurosci. *15*, 5999–6013.

Eldar, E., Rutledge, R.B., Dolan, R.J., and Niv, Y. (2016). Mood as representation of momentum. Trends Cogn. Sci. *20*, 15–24. https://doi.org/10.1016/j.tics.2015.07.010.

Farashahi, S., Donahue, C.H., Khorsand, P., Seo, H., Lee, D., and Soltani, A. (2017). Metaplasticity as a neural substrate for adaptive learning and choice under uncertainty. Neuron *94*, 401–414.e6. https://doi.org/10.1016/j.neuron.2017.03.044.

Fellows, L.K. (2006). Deciding how to decide: ventromedial frontal lobe damage affects information acquisition in multi-attribute decision making. Brain *129*, 944–952.

Fischer, A.G., Bourgeois-Gironde, S., and Ullsperger, M. (2017). Short-term reward experience biases inference despite dissociable neural correlates. Nat. Commun. *8*, 1690. https://doi.org/10.1038/s41467-017-01703-0.

Neuron
Review

CellPress

Fischer, A.G., Endrass, T., Reuter, M., Kubisch, C., and Ullsperger, M. (2015). Serotonin reuptake inhibitors and serotonin transporter genotype modulate performance monitoring functions but not their electrophysiological correlates. J. Neurosci. 35, 8181–8190. https://doi.org/10.1523/JNEUROSCI.5124-14.2015.

FitzGerald, T.H., Seymour, B., and Dolan, R.J. (2009). The role of human orbitofrontal cortex in value comparison for incommensurable objects. J. Neurosci. 29, 8388–8395. https://doi.org/10.1523/JNEUROSCI.0717-09.2009.

Folloni, D., Fouragnan, E., Wittmann, M.K., Roumazeilles, L., Tankelevitch, L., Verhagen, L., Attali, D., Aubry, J.-F., Sallet, J., and Rushworth, M.F.S. (2021). Ultrasound modulation of macaque prefrontal cortex selectively alters credit assignment-related activity and behavior. Sci. Adv. 7, eabg7700. https://doi.org/10.1126/sciadv.abg7700.

Fouragnan, E.F., Chau, B.K.H., Folloni, D., Kolling, N., Verhagen, L., Klein-Flügge, M., Tankelevitch, L., Papageorgiou, G.K., Aubry, J.F., Sallet, J., et al. (2019). The macaque anterior cingulate cortex translates counterfactual choice value into actual behavioral change. Nat. Neurosci. 22, 797–808. https://doi.org/10.1038/s41593-019-0375-6.

Freidin, E., and Kacelnik, A. (2011). Rational choice, context dependence, and the value of information in European Starlings (Sturnus vulgaris). Science 334, 1000–1002.

Freudenmacher, L., Schauer, M., Walkowiak, W., and von Twickel, A. (2020). Refinement of the dopaminergic system of anuran amphibians based on connectivity with habenula, basal ganglia, limbic system, pallium, and spinal cord. J. Comp. Neurol. 528, 972–988. https://doi.org/10.1002/cne.24793.

Friedman, A., Homma, D., Gibb, L.G., Amemori, K.-I., Rubin, S.J., Hood, A.S., Riad, M.H., and Graybiel, A.M. (2015). A corticostriatal path targeting striosomes controls decision-making under conflict. Cell 161, 1320–1333. https://doi.org/10.1016/j.cell.2015.04.049.

Frith, C.D., and Frith, U. (2012). Mechanisms of social cognition. Annu. Rev. Psychol. 63, 287–313. https://doi.org/10.1146/annurev-psych-120710-100449.

Gerraty, R.T., Davidow, J.Y., Foerde, K., Galvan, A., Bassett, D.S., and Shohamy, D. (2018). Dynamic flexibility in striatal-cortical circuits supports reinforcement learning. J. Neurosci. 38, 2442–2453. https://doi.org/10.1523/JNEUROSCI.2084-17.2018.

Glasser, M.F., Coalson, T.S., Robinson, E.C., Hacker, C.D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C.F., Jenkinson, M., et al. (2016). A multi-modal parcellation of human cerebral cortex. Nature 536, 171–178. https://doi.org/10.1038/nature18933.

Grabenhorst, F., and Rolls, E.T. (2011). Value, pleasure and choice in the ventral prefrontal cortex. Trends Cogn. Sci. 15, 56–67. https://doi.org/10.1016/j.tics.2010.12.004.

Graziano, M.S.A. (2009). The Intelligent Movement Machine: an Ethological Perspective on the Primate Motor System (Oxford University Press).

Grossman, C.D., Bari, B.A., and Cohen, J.Y. (2022). Serotonin neurons modulate learning rate through uncertainty. Curr. Biol. 32, 586–599.e7. https://doi.org/10.1016/j.cub.2021.12.006.

Grueschow, M., Polania, R., Hare, T.A., and Ruff, C.C. (2015). Automatic versus choice-dependent value representations in the human brain. Neuron 85, 874–885. https://doi.org/10.1016/j.neuron.2014.12.054.

Hämmerer, D., Bonaiuto, J., Klein-Flügge, M., Bikson, M., and Bestmann, S. (2016). Selective alteration of human value decisions with medial frontal tDCS is predicted by changes in attractor dynamics. Sci. Rep. 6, 25160.

Hart, A.S., Rutledge, R.B., Glimcher, P.W., and Phillips, P.E. (2014). Phasic dopamine release in the rat nucleus accumbens symmetrically encodes a reward prediction error term. J. Neurosci. 34, 698–704. https://doi.org/10.1523/JNEUROSCI.2489-13.2014.

Hartley, T., Lever, C., Burgess, N., and O'Keefe, J. (2014). Space in the brain: how the hippocampal formation supports spatial cognition. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 369, 20120510. https://doi.org/10.1098/rstb.2012.0510.

Hartley, T., Maguire, E.A., Spiers, H.J., and Burgess, N. (2003). The well-worn route and the path less traveled: distinct neural bases of route following and

wayfinding in humans. Neuron 37, 877–888. https://doi.org/10.1016/s0896-6273(03)00095-3.

Hayashi, K., Nakao, K., and Nakamura, K. (2015). Appetitive and aversive information coding in the primate dorsal raphé nucleus. J. Neurosci. 35, 6195–6208. https://doi.org/10.1523/JNEUROSCI.2860-14.2015.

Hayden, B.Y., and Niv, Y. (2021). The case against economic values in the orbitofrontal cortex (or anywhere else in the brain). Behav. Neurosci. 135, 192–201. https://doi.org/10.1037/bne0000448.

Hayden, B.Y., Pearson, J.M., and Platt, M.L. (2011). Neuronal basis of sequential foraging decisions in a patchy environment. Nat. Neurosci. 14, 933–939. https://doi.org/10.1038/nn.2856.

Holroyd, C.B., Ribas-Fernandes, J.J.F., Shahnazian, D., Silvetti, M., and Verguts, T. (2018). Human midcingulate cortex encodes distributed representations of task progress. Proc. Natl. Acad. Sci. USA 115, 6398–6403. https://doi.org/10.1073/pnas.1803650115.

Holt-Lunstad, J., Smith, T.B., Baker, M., Harris, T., and Stephenson, D. (2015). Loneliness and social isolation as risk factors for mortality: a meta-analytic review. Perspect. Psychol. Sci. 10, 227–237. https://doi.org/10.1177/1745691614568352.

Holt-Lunstad, J., Smith, T.B., and Layton, J.B. (2010). Social relationships and mortality risk: a meta-analytic review. PLoS Med 7, e1000316. https://doi.org/10.1371/journal.pmed.1000316.

Hong, S., and Hikosaka, O. (2008). The globus pallidus sends reward-related signals to the lateral habenula. Neuron 60, 720–729. https://doi.org/10.1016/j.neuron.2008.09.035.

Howard, J.D., Reynolds, R., Smith, D.E., Voss, J.L., Schoenbaum, G., and Kahnt, T. (2020). Targeted stimulation of human orbitofrontal networks disrupts outcome-guided behavior. Curr. Biol. 30, 490–498.e4. https://doi.org/10.1016/j.cub.2019.12.007.

Hunt, L.T., Daw, N.D., Kaanders, P., MacIver, M.A., Mugan, U., Procyk, E., Redish, A.D., Russo, E., Scholl, J., Stachenfeld, K., et al. (2021). Formalizing planning and information search in naturalistic decision-making. Nat. Neurosci. 24, 1051–1064. https://doi.org/10.1038/s41593-021-00866-w.

Hunt, L.T., and Hayden, B.Y. (2017). A distributed, hierarchical and recurrent framework for reward-based choice. Nat. Rev. Neurosci. 18, 172–182. https://doi.org/10.1038/nrn.2017.7.

Hunt, L.T., Kolling, N., Soltani, A., Woolrich, M.W., Rushworth, M.F., and Behrens, T.E. (2012). Mechanisms underlying cortical activity during value-guided choice. Nat. Neurosci. 15, 470–476. https://doi.org/10.1038/nn.3017.

Hunt, L.T., Malalasekera, W.M.N., de Berker, A.O., Miranda, B., Farmer, S.F., Behrens, T.E.J., and Kennerley, S.W. (2018). Triple dissociation of attention and decision computations across prefrontal cortex. Nat. Neurosci. 21, 1471–1481. https://doi.org/10.1038/s41593-018-0239-5.

Iigaya, K., Hauser, T.U., Kurth-Nelson, Z., O'Doherty, J.P., Dayan, P., and Dolan, R.J. (2020). The value of what's to come: neural mechanisms coupling prediction error and the utility of anticipation. Sci. Adv. 6, eaba3828. https://doi.org/10.1126/sciadv.aba3828.

Jamali, M., Grannan, B.L., Fedorenko, E., Saxe, R., Báez-Mendoza, R., and Williams, Z.M. (2021). Single-neuronal predictions of others' beliefs in humans. Nature 591, 610–614. https://doi.org/10.1038/s41586-021-03184-0.

Jang, A.I., Costa, V.D., Rudebeck, P.H., Chudasama, Y., Murray, E.A., and Averbeck, B.B. (2015). The role of frontal cortical and medial-temporal lobe brain areas in learning a bayesian prior belief on reversals. J. Neurosci. 35, 11751–11760. https://doi.org/10.1523/JNEUROSCI.1594-15.2015.

Janmaat, K.R.L., Byrne, R.W., and Zuberbühler, K. (2006). Primates take weather into account when searching for fruits. Curr. Biol. 16, 1232–1237. https://doi.org/10.1016/j.cub.2006.04.031.

Janmaat, K.R.L., Chapman, C.A., Meijer, R., and Zuberbühler, K. (2012). The use of fruiting synchrony by foraging mangabey monkeys: a "simple tool" to find fruit. Anim. Cogn. 15, 83–96. https://doi.org/10.1007/s10071-011-0435-0.

Jocham, G., Brodersen, K.H., Constantinescu, A.O., Kahn, M.C., Ianni, A.M., Walton, M.E., Rushworth, M.F., and Behrens, T.E. (2016). Reward-guided

learning with and without causal attribution. Neuron 90, 177–190. https://doi.org/10.1016/j.neuron.2016.02.018.

Jocham, G., Hunt, L.T., Near, J., and Behrens, T.E. (2012). A mechanism for value-guided choice based on the excitation-inhibition balance in prefrontal cortex. Nat. Neurosci. 15, 960–961. https://doi.org/10.1038/nn.3140.

Joshi, S., and Gold, J.I. (2022). Context-dependent relationships between locus coeruleus firing patterns and coordinated neural activity in the anterior cingulate cortex. Elife 11, e63490. https://doi.org/10.7554/eLife.63490.

Joshi, S., Li, Y., Kalwani, R.M., and Gold, J.I. (2016). Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. Neuron 89, 221–234. https://doi.org/10.1016/j.neuron.2015.11.028.

Kaanders, P., Nili, H., O'Reilly, J.X., and Hunt, L. (2021). Medial frontal cortex activity predicts information sampling in economic choice. J. Neurosci. 41, 8403–8413. https://doi.org/10.1523/JNEUROSCI.0392-21.2021.

Kable, J.W., and Glimcher, P.W. (2007). The neural correlates of subjective value during intertemporal choice. Nat. Neurosci. 10, 1625–1633. https://doi.org/10.1038/nn2007.

Kable, J.W., and Glimcher, P.W. (2009). The neurobiology of decision: consensus and controversy. Neuron 63, 733–745. https://doi.org/10.1016/j.neuron.2009.09.003.

Kahnt, T., Heinzle, J., Park, S.Q., and Haynes, J.-D. (2011). Decoding different roles for vmPFC and dlPFC in multi-attribute decision making. NeuroImage 56, 709–715. https://doi.org/10.1016/j.neuroimage.2010.05.058.

Kennerley, S.W., Dahmubed, A.F., Lara, A.H., and Wallis, J.D. (2009). Neurons in the frontal lobe encode the value of multiple decision variables. J. Cogn. Neurosci. 21, 1162–1178. https://doi.org/10.1162/jocn.2009.21100.

Kennerley, S.W., Walton, M.E., Behrens, T.E., Buckley, M.J., and Rushworth, M.F. (2006). Optimal decision making and the anterior cingulate cortex. Nat. Neurosci. 9, 940–947.

Keren, H., Zheng, C., Jangraw, D.C., Chang, K., Vitale, A., Rutledge, R.B., Pereira, F., Nielson, D.M., and Stringaris, A. (2021). The temporal representation of experience in subjective mood. Elife 10, e62051. https://doi.org/10.7554/eLife.62051.

Khalighinejad, N., Bongioanni, A., Verhagen, L., Folloni, D., Attali, D., Aubry, J.F., Sallet, J., and Rushworth, M.F.S. (2020a). A basal forebrain-cingulate circuit in macaques decides it is time to act. Neuron 105, 370–384.e8. https://doi.org/10.1016/j.neuron.2019.10.030.

Khalighinejad, N., Garrett, N., Priestley, L., Lockwood, P., and Rushworth, M.F.S. (2021). A habenula-insular circuit encodes the willingness to act. Nat. Commun. 12, 6329. https://doi.org/10.1038/s41467-021-26569-1.

Khalighinejad, N., Manohar, S., Husain, M., and Rushworth, M.F.S. (2022). Complementary roles of serotonergic and cholinergic systems in decisions about when to act. Curr. Biol. CB 32, 1150–1162.e7. https://doi.org/10.1016/j.cub.2022.01.042.

Khalighinejad, N., Priestley, L., Jbabdi, S., and Rushworth, M.F.S. (2020b). Human decisions about when to act originate within a basal forebrain-nigral circuit. Proc. Natl. Acad. Sci. USA 117, 11799–11810. https://doi.org/10.1073/pnas.1921211117.

Kidd, C., and Hayden, B.Y. (2015). The psychology and neuroscience of curiosity. Neuron 88, 449–460. https://doi.org/10.1016/j.neuron.2015.09.010.

Klein, T.A., Ullsperger, M., and Jocham, G. (2017). Learning relative values in the striatum induces violations of normative decision making. Nat. Commun. 8, 16033. https://doi.org/10.1038/ncomms16033.

Klein-Flügge, M.C., Barron, H.C., Brodersen, K.H., Dolan, R.J., and Behrens, T.E. (2013). Segregated encoding of reward-identity and stimulus-reward associations in human orbitofrontal cortex. J. Neurosci. 33, 3202–3211. https://doi.org/10.1523/JNEUROSCI.2532-12.2013.

Klein-Flügge, M.C., Kennerley, S.W., Friston, K., and Bestmann, S. (2016). Neural signatures of value comparison in human cingulate cortex during decisions requiring an effort-reward trade-off. J. Neurosci. 36, 10002–10015. https://doi.org/10.1523/JNEUROSCI.0292-16.2016.

Klein-Flügge, M.C., Wittmann, M.K., Shpektor, A., Jensen, D.E.A., and Rushworth, M.F.S. (2019). Multiple associative structures created by reinforcement and incidental statistical learning mechanisms. Nat. Commun. 10, 4835. https://doi.org/10.1038/s41467-019-12557-z.

Kobayashi, K., Ravaioli, S., Baranès, A., Woodford, M., and Gottlieb, J. (2019). Diverse motives for human curiosity. Nat. Hum. Behav. 3, 587–595. https://doi.org/10.1038/s41562-019-0589-3.

Kolling, N., Behrens, T., Wittmann, M.K., and Rushworth, M. (2016a). Multiple signals in anterior cingulate cortex. Curr. Opin. Neurobiol. 37, 36–43. https://doi.org/10.1016/j.conb.2015.12.007.

Kolling, N., Behrens, T.E., Mars, R.B., and Rushworth, M.F. (2012). Neural mechanisms of foraging. Science 336, 95–98. https://doi.org/10.1126/science.1216930.

Kolling, N., Scholl, J., Chekroud, A., Trier, H.A., and Rushworth, M.F.S. (2018). Prospection, perseverance, and insight in sequential behavior. Neuron 99, 1069–1082.e7. https://doi.org/10.1016/j.neuron.2018.08.018.

Kolling, N., Wittmann, M.K., Behrens, T.E., Boorman, E.D., Mars, R.B., and Rushworth, M.F. (2016). Value, search, persistence and model updating in anterior cingulate cortex. Nat Neurosci. 19, 1280–1285. https://doi.org/10.1038/nn.4382.

Konovalov, A., Hill, C., Daunizeau, J., and Ruff, C.C. (2021). Dissecting functional contributions of the social brain to strategic behavior. Neuron 109, 3323–3337.e5. https://doi.org/10.1016/j.neuron.2021.07.025.

Korn, C.W., and Bach, D.R. (2018). Heuristic and optimal policy computations in the human brain during sequential decision-making. Nat. Commun. 9, 325. https://doi.org/10.1038/s41467-017-02750-3.

Korn, C.W., and Bach, D.R. (2019). Minimizing threat via heuristic and optimal policies recruits hippocampus and medial prefrontal cortex. Nat. Hum. Behav. 3, 733–745. https://doi.org/10.1038/s41562-019-0603-9.

Kralik, J.D., Xu, E.R., Knight, E.J., Khan, S.A., and Levine, W.J. (2012). When less is more: evolutionary origins of the affect heuristic. PloS One 7, e46240. https://doi.org/10.1371/journal.pone.0046240.

Kurniawan, I.T., Guitart-Masip, M., Dayan, P., and Dolan, R.J. (2013). Effort and valuation in the brain: the effects of anticipation and execution. J. Neurosci. 33, 6160–6169. https://doi.org/10.1523/JNEUROSCI.4777-12.2013.

Kuwabara, M., Kang, N., Holy, T.E., and Padoa-Schioppa, C. (2020). Neural mechanisms of economic choices in mice. Elife 9, e49669. https://doi.org/10.7554/eLife.49669.

Leong, Y.C., Radulescu, A., Daniel, R., DeWoskin, V., and Niv, Y. (2017). Dynamic interaction between reinforcement learning and attention in multidimensional environments. Neuron 93, 451–463. https://doi.org/10.1016/j.neuron.2016.12.040.

Lichtenberg, N.T., Sepe-Forrest, L., Pennington, Z.T., Lamparelli, A.C., Greenfield, V.Y., and Wassum, K.M. (2021). The medial orbitofrontal cortex-basolateral amygdala circuit regulates the influence of reward cues on adaptive behavior and choice. J. Neurosci. 41, 7267–7277. https://doi.org/10.1523/JNEUROSCI.0901-21.2021.

Lim, S.L., O'Doherty, J.P., and 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. https://doi.org/10.1523/JNEUROSCI.1246-11.2011.

List, J.A. (2002). Preference reversals of a different kind: the "more is less" phenomenon. Am. Econ. Rev. 92, 1636–1643.

Liu, Y., Xin, Y., and Xu, N.L. (2021). A cortical circuit mechanism for structural knowledge-based flexible sensorimotor decision-making. Neuron 109, 2009–2024.e6. https://doi.org/10.1016/j.neuron.2021.04.014.

Lopatina, N., McDannald, M.A., Styer, C.V., Peterson, J.F., Sadacca, B.F., Cheer, J.F., and Schoenbaum, G. (2016). Medial orbitofrontal neurons preferentially signal cues predicting changes in reward during unblocking. J. Neurosci. 36, 8416–8424. https://doi.org/10.1523/JNEUROSCI.1101-16.2016.

Lopatina, N., McDannald, M.A., Styer, C.V., Sadacca, B.F., Cheer, J.F., and Schoenbaum, G. (2015). Lateral orbitofrontal neurons acquire responses to

upshifted, downshifted, or blocked cues during unblocking. Elife 4, e11299. https://doi.org/10.7554/eLife.11299.

Lopez-Persem, A., Bastin, J., Petton, M., Abitbol, R., Lehongre, K., Adam, C., Navarro, V., Rheims, S., Kahane, P., Domenech, P., et al. (2020). Four core properties of the human brain valuation system demonstrated in intracranial signals. Nat. Neurosci. 23, 664–675. https://doi.org/10.1038/s41593-020-0615-9.

Lopez-Persem, A., Domenech, P., and Pessiglione, M. (2016). How prior preferences determine decision-making frames and biases in the human brain. Elife 5. https://doi.org/10.7554/eLife.20317.

Louie, K., Glimcher, P.W., and Webb, R. (2015). Adaptive neural coding: from biological to behavioral decision-making. Curr. Opin. Behav. Sci. 5, 91–99. https://doi.org/10.1016/j.cobeha.2015.08.008.

Ma, L., Hyman, J.M., Durstewitz, D., Phillips, A.G., and Seamans, J.K. (2016). A quantitative analysis of context-dependent remapping of medial frontal cortex neurons and ensembles. J. Neurosci. 36, 8258–8272. https://doi.org/10.1523/JNEUROSCI.3176-15.2016.

Ma, L., Hyman, J.M., Phillips, A.G., and Seamans, J.K. (2014). Tracking progress toward a goal in corticostriatal ensembles. J. Neurosci. 34, 2244–2253. https://doi.org/10.1523/JNEUROSCI.3834-13.2014.

MacIver, M.A., Schmitz, L., Mugan, U., Murphey, T.D., and Mobley, C.D. (2017). Massive increase in visual range preceded the origin of terrestrial vertebrates. Proc. Natl. Acad. Sci. USA 114, E2375–E2384. https://doi.org/10.1073/pnas.1615563114.

Maier, S.U., Makwana, A.B., and Hare, T.A. (2015). Acute stress impairs self-control in goal-directed choice by altering multiple functional connections within the Brain's decision circuits. Neuron 87, 621–631. https://doi.org/10.1016/j.neuron.2015.07.005.

Malvaez, M., Shieh, C., Murphy, M.D., Greenfield, V.Y., and Wassum, K.M. (2019). Distinct cortical-amygdala projections drive reward value encoding and retrieval. Nat. Neurosci. 22, 762–769. https://doi.org/10.1038/s41593-019-0374-7.

Marques, J.C., Li, M., Schaak, D., Robson, D.N., and Li, J.M. (2020). Internal state dynamics shape brainwide activity and foraging behaviour. Nature 577, 239–243. https://doi.org/10.1038/s41586-019-1858-z.

Matsumoto, M., and Hikosaka, O. (2007). Lateral habenula as a source of negative reward signals in dopamine neurons. Nature 447, 1111–1115. https://doi.org/10.1038/nature05860.

Mayner, L. (1989). A cyto-architectonic study of the cortex of the tammar wallaby, Macropus eugenii. Brain Behav. Evol. 33, 303–316. https://doi.org/10.1159/000115938.

Meder, D., Kolling, N., Verhagen, L., Wittmann, M.K., Scholl, J., Madsen, K.H., Hulme, O.J., Behrens, T.E.J., and Rushworth, M.F.S. (2017). Simultaneous representation of a spectrum of dynamically changing value estimates during decision making. Nat. Commun. 8, 1942. https://doi.org/10.1038/s41467-017-02169-w.

Mehta, P.S., Tu, J.C., LoConte, G.A., Pesce, M.C., and Hayden, B.Y. (2019). Ventromedial prefrontal cortex tracks multiple environmental variables during search. J. Neurosci. 39, 5336–5350. https://doi.org/10.1523/JNEUROSCI.2365-18.2019.

Menzel, C.R. (1991). Cognitive aspects of foraging in Japanese monkeys. Anim. Behav. 41, 397–402. https://doi.org/10.1016/S0003-3472(05)80840-1.

Menzel, C.R. (1996). Spontaneous use of matching visual cues during foraging by long-tailed macaques (Macaca fascicularis). J. Comp. Psychol. 110, 370–376. https://doi.org/10.1037/0735-7036.110.4.370.

Mishkin, M., Vest, B., Waxler, M., and Rosvold, H.E. (1969). A re-examination of the effects of frontal lesions on object alternation. Neuropsychologia 7, 357–363. https://doi.org/10.1016/0028-3932(69)90060-8.

Miyamoto, K., Trudel, N., Kamermans, K., Lim, M.C., Lazari, A., Verhagen, L., Wittmann, M.K., and Rushworth, M.F.S. (2021). Identification and disruption of a neural mechanism for accumulating prospective metacognitive information prior to decision-making. Neuron 109, 1396–1408.e7. https://doi.org/10.1016/j.neuron.2021.02.024.

Monosov, I.E. (2017). Anterior cingulate is a source of valence-specific information about value and uncertainty. Nat. Commun. 8, 134. https://doi.org/10.1038/s41467-017-00072-y.

Monosov, I.E., Haber, S.N., Leuthardt, E.C., and Jezzini, A. (2020). Anterior cingulate cortex and the control of dynamic behavior in primates. Curr. Biol. 30, R1442–R1454. https://doi.org/10.1016/j.cub.2020.10.009.

Mugan, U., and MacIver, M.A. (2020). Spatial planning with long visual range benefits escape from visual predators in complex naturalistic environments. Nat. Commun. 11, 3057. https://doi.org/10.1038/s41467-020-16102-1.

Muller, T.H., Mars, R.B., Behrens, T.E., and O'Reilly, J.X. (2019). Control of entropy in neural models of environmental state. Elife 8, e39404. https://doi.org/10.7554/eLife.39404.

Murray, E.A., Moylan, E.J., Saleem, K.S., Basile, B.M., and Turchi, J. (2015). Specialized areas for value updating and goal selection in the primate orbitofrontal cortex. Elife 4, e11695. https://doi.org/10.7554/eLife.11695.

Murray, J.D., Bernacchia, A., Freedman, D.J., Romo, R., Wallis, J.D., Cai, X., Padoa-Schioppa, C., Pasternak, T., Seo, H., Lee, D., et al. (2014). A hierarchy of intrinsic timescales across primate cortex. Nat. Neurosci. 17, 1661–1663. https://doi.org/10.1038/nn.3862.

Na, S., Chung, D., Hula, A., Perl, O., Jung, J., Heflin, M., Blackmore, S., Fiore, V.G., Dayan, P., and Gu, X. (2021). Humans use forward thinking to exploit social controllability. eLife 10, e64983. https://doi.org/10.7554/eLife.64983.

Neubert, F.X., Mars, R.B., Sallet, J., and Rushworth, M.F. (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. https://doi.org/10.1073/pnas.1410767112.

Neubert, F.X., Mars, R.B., Thomas, A.G., Sallet, J., and Rushworth, M.F. (2014). Comparison of human ventral frontal cortex areas for cognitive control and language with areas in monkey frontal cortex. Neuron 81, 700–713. https://doi.org/10.1016/j.neuron.2013.11.012.

Ninomiya, T., Noritake, A., Kobayashi, K., and Isoda, M. (2020). A causal role for frontal cortico-cortical coordination in social action monitoring. Nat. Commun. 11, 5233. https://doi.org/10.1038/s41467-020-19026-y.

Noonan, M.P., Chau, B.K.H., Rushworth, M.F.S., and Fellows, L.K. (2017). Contrasting effects of medial and lateral orbitofrontal cortex lesions on credit assignment and decision-making in humans. J. Neurosci. 37, 7023–7035. https://doi.org/10.1523/JNEUROSCI.0692-17.2017.

Noonan, M.P., Mars, R.B., and Rushworth, M.F. (2011). Distinct roles of three frontal cortical areas in reward-guided behavior. J. Neurosci. 31, 14399–14412.

Noonan, M.P., Walton, M.E., Behrens, T.E., Sallet, J., Buckley, M.J., and Rushworth, M.F. (2010). Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. Proc. Natl. Acad. Sci. USA 107, 20547–20552. https://doi.org/10.1073/pnas.1012246107.

Noritake, A., Ninomiya, T., and Isoda, M. (2018). Social reward monitoring and valuation in the macaque brain. Nat. Neurosci. 21, 1452–1462. https://doi.org/10.1038/s41593-018-0229-7.

Okazawa, G., Hatch, C.E., Mancoo, A., Machens, C.K., and Kiani, R. (2021). Representational geometry of perceptual decisions in the monkey parietal cortex. Cell 184, 3748–3761.e18. https://doi.org/10.1016/j.cell.2021.05.022.

Padoa-Schioppa, C., and Conen, K.E. (2017). Orbitofrontal cortex: a neural circuit for economic decisions. Neuron 96, 736–754. https://doi.org/10.1016/j.neuron.2017.09.031.

Papageorgiou, G.K., Sallet, J., Wittmann, M.K., Chau, B.K.H., Schüffelgen, U., Buckley, M.J., and Rushworth, M.F.S. (2017). Inverted activity patterns in ventromedial prefrontal cortex during value-guided decision-making in a less-is-more task. Nat. Commun. 8, 1886. https://doi.org/10.1038/s41467-017-01833-5.

Park, S.A., Miller, D.S., and Boorman, E.D. (2021). Inferences on a multidimensional social hierarchy use a grid-like code. Nat. Neurosci. 24, 1292–1301. https://doi.org/10.1038/s41593-021-00916-3.

Parvizi, J., Rangarajan, V., Shirer, W.R., Desai, N., and Greicius, M.D. (2013). The will to persevere induced by electrical stimulation of the human cingulate gyrus. Neuron 80, 1359–1367. https://doi.org/10.1016/j.neuron.2013.10.057.

Passingham, R.E., and Wise, S.P. (2012). The Neurobiology of the Prefrontal Cortex: Anatomy, Evolution, and the Origin of Insight (Oxford University Press).

Pearson, J.M., Watson, K.K., and Platt, M.L. (2014). Decision making: the neuroethological turn. Neuron 82, 950–965. https://doi.org/10.1016/j.neuron.2014.04.037.

Pickens, C.L., Saddoris, M.P., Setlow, B., Gallagher, M., Holland, P.C., and Schoenbaum, G. (2003). Different roles for orbitofrontal cortex and basolateral amygdala in a reinforcer devaluation task. J. Neurosci. 23, 11078–11084.

Preuss, T.M. (1995). Do rats have prefrontal cortex? The Rose-Woolsey-Akert program reconsidered. J. Cogn. Neurosci. 7, 1–24.

Prévost, C., Pessiglione, M., Météreau, E., Cléry-Melin, M.L., and Dreher, J.C. (2010). Separate valuation subsystems for delay and effort decision costs. J. Neurosci. 30, 14080–14090. https://doi.org/10.1523/JNEUROSCI.2752-10.2010.

Procyk, E., Tanaka, Y.L., and Joseph, J.P. (2000). Anterior cingulate activity during routine and non-routine sequential behaviors in macaques. Nat. Neurosci. 3, 502–508.

Remondes, M., and Wilson, M.A. (2013). Cingulate-hippocampus coherence and trajectory coding in a sequential choice task. Neuron 80, 1277–1289. https://doi.org/10.1016/j.neuron.2013.08.037.

Ribas-Fernandes, J.J.F., Solway, A., Diuk, C., McGuire, J.T., Barto, A.G., Niv, Y., and Botvinick, M.M. (2011). A neural signature of hierarchical reinforcement learning. Neuron 71, 370–379. https://doi.org/10.1016/j.neuron.2011.05.042.

Rich, E.L., and Wallis, J.D. (2016). Decoding subjective decisions from orbitofrontal cortex. Nat. Neurosci. 19, 973–980. https://doi.org/10.1038/nn.4320.

Rudebeck, P.H., and Izquierdo, A. (2021). Foraging with the frontal cortex: a cross-species evaluation of reward-guided behavior. Neuropsychopharmacology 47, 134–146. https://doi.org/10.1038/s41386-021-01140-0.

Rudebeck, P.H., and Murray, E.A. (2014). The orbitofrontal oracle: cortical mechanisms for the prediction and evaluation of specific behavioral outcomes. Neuron 84, 1143–1156. https://doi.org/10.1016/j.neuron.2014.10.049.

Rudebeck, P.H., Saunders, R.C., Lundgren, D.A., and Murray, E.A. (2017). Specialized representations of value in the orbital and ventrolateral prefrontal cortex: desirability versus availability of outcomes. Neuron 95, 1208–1220.e5. https://doi.org/10.1016/j.neuron.2017.07.042.

Rudebeck, P.H., Saunders, R.C., Prescott, A.T., Chau, L.S., and Murray, E.A. (2013). Prefrontal mechanisms of behavioral flexibility, emotion regulation and value updating. Nat. Neurosci. 16, 1140–1145. https://doi.org/10.1038/nn.3440.

Rudebeck, P.H., Walton, M.E., Smyth, A.N., Bannerman, D.M., and Rushworth, M.F. (2006). Separate neural pathways process different decision costs. Nat. Neurosci. 9, 1161–1168.

Rushworth, M.F., Noonan, M.P., Boorman, E.D., Walton, M.E., and Behrens, T.E. (2011). Frontal cortex and reward-guided learning and decision-making. Neuron 70, 1054–1069. https://doi.org/10.1016/j.neuron.2011.05.014.

Sadacca, B.F., Wied, H.M., Lopatina, N., Saini, G.K., Nemirovsky, D., and Schoenbaum, G. (2018). Orbitofrontal neurons signal sensory associations underlying model-based inference in a sensory preconditioning task. Elife 7, e30373. https://doi.org/10.7554/eLife.30373.

Scholl, J., Kolling, N., Nelissen, N., Wittmann, M.K., Harmer, C.J., and Rushworth, M.F. (2015). The good, the bad, and the irrelevant: neural mechanisms of learning real and hypothetical rewards and effort. J. Neurosci. 35, 11233–11251. https://doi.org/10.1523/JNEUROSCI.0396-15.2015.

Schuck, N.W., Cai, M.B., Wilson, R.C., and Niv, Y. (2016). Human orbitofrontal cortex represents a cognitive map of state space. Neuron 91, 1402–1412. https://doi.org/10.1016/j.neuron.2016.08.019.

Schuck, N.W., Gaschler, R., Wenke, D., Heinzle, J., Frensch, P.A., Haynes, J.-D., and Reverberi, C. (2015). Medial prefrontal cortex predicts internally driven strategy shifts. Neuron 86, 331–340. https://doi.org/10.1016/j.neuron.2015.03.015.

Schülke, O., Bhagavatula, J., Vigilant, L., and Ostner, J. (2010). Social bonds enhance reproductive success in male macaques. Curr. Biol. 20, 2207–2210. https://doi.org/10.1016/j.cub.2010.10.058.

Schultz, W. (2013). Updating dopamine reward signals. Curr. Opin. Neurobiol. 23, 229–238. https://doi.org/10.1016/j.conb.2012.11.012.

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

Schweighofer, N., and Doya, K. (2003). Meta-learning in reinforcement learning. Neural Netw 16, 5–9. https://doi.org/10.1016/s0893-6080(02)00228-9.

Seo, H., Cai, X., Donahue, C.H., and Lee, D. (2014). Neural correlates of strategic reasoning during competitive games. Science 346, 340–343. https://doi.org/10.1126/science.1256254.

Seo, H., and Lee, D. (2009). Behavioral and neural changes after gains and losses of conditioned reinforcers. J. Neurosci. 29, 3627–3641. https://doi.org/10.1523/JNEUROSCI.4726-08.2009.

Shahnazian, D., and Holroyd, C.B. (2018). Distributed representations of action sequences in anterior cingulate cortex: a recurrent neural network approach. Psychon. Bull. Rev. 25, 302–321. https://doi.org/10.3758/s13423-017-1280-1.

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

Shima, K., Isoda, M., Mushiake, H., and Tanji, J. (2007). Categorization of behavioural sequences in the prefrontal cortex. Nature 445, 315–318. https://doi.org/10.1038/nature05470.

Sias, A.C., Morse, A.K., Wang, S., Greenfield, V.Y., Goodpaster, C.M., Wrenn, T.M., Wikenheiser, A.M., Holley, S.M., Cepeda, C., Levine, M.S., et al. (2021). A bidirectional corticoamygdala circuit for the encoding and retrieval of detailed reward memories. Elife 10, e68617. https://doi.org/10.7554/eLife.68617.

Soltani, A., and Izquierdo, A. (2019). Adaptive learning under expected and unexpected uncertainty. Nat. Rev. Neurosci. 20, 635–644. https://doi.org/10.1038/s41583-019-0180-y.

Soltani, A., and Koechlin, E. (2022). Computational models of adaptive behavior and prefrontal cortex. Neuropsychopharmacology 47, 58–71. https://doi.org/10.1038/s41386-021-01123-1.

Soltani, A., Murray, J.D., Seo, H., and Lee, D. (2021). Timescales of cognition in the brain. Curr. Opin. Behav. Sci. 41, 30–37. https://doi.org/10.1016/j.cobeha.2021.03.003.

Spalding, K.N., Schlichting, M.L., Zeithamova, D., Preston, A.R., Tranel, D., Duff, M.C., and Warren, D.E. (2018). Ventromedial prefrontal cortex is necessary for normal associative inference and memory integration. J. Neurosci. 38, 3767–3775. https://doi.org/10.1523/JNEUROSCI.2501-17.2018.

Spitmaan, M., Seo, H., Lee, D., and Soltani, A. (2020). Multiple timescales of neural dynamics and integration of task-relevant signals across cortex. Proc. Natl. Acad. Sci. USA 117, 22522–22531. https://doi.org/10.1073/pnas.2005993117.

Stalnaker, T.A., Cooch, N.K., and Schoenbaum, G. (2015). What the orbitofrontal cortex does not do. Nat. Neurosci. 18, 620–627. https://doi.org/10.1038/nn.3982.

Stephens, D.W., and Krebs, J.R. (1986). Foraging Theory (Princeton University Press).

Stephenson-Jones, M., Kardamakis, A.A., Robertson, B., and Grillner, S. (2013). Independent circuits in the basal ganglia for the evaluation and selection of actions. Proc. Natl. Acad. Sci. USA 110, E3670–E3679. https://doi.org/10.1073/pnas.1314511110.

Stephenson-Jones, M., Yu, K., Ahrens, S., Tucciarone, J.M., van Huijstee, A.N., Mejia, L.A., Penzo, M.A., Tai, L.-H., Wilbrecht, L., and Li, B. (2016). A basal ganglia circuit for evaluating action outcomes. Nature 539, 289–293. https://doi.org/10.1038/nature19845.

Neuron
Review

CellPress

Stoll, F.M., Fontanier, V., and Procyk, E. (2016). Specific frontal neural dynamics contribute to decisions to check. Nat. Commun. *7*, 11990. https://doi.org/10.1038/ncomms11990.

Strait, C.E., Blanchard, T.C., and Hayden, B.Y. (2014). Reward value comparison via mutual inhibition in ventromedial prefrontal cortex. Neuron *82*, 1357–1366. https://doi.org/10.1016/j.neuron.2014.04.032.

Strait, C.E., Sleezer, B.J., and Hayden, B.Y. (2015). Signatures of value comparison in ventral striatum neurons. PLoS Biol *13*, e1002173. https://doi.org/10.1371/journal.pbio.1002173.

Suárez, R., Paolino, A., Fenlon, L.R., Morcom, L.R., Kozulin, P., Kurniawan, N.D., and Richards, L.J. (2018). A pan-mammalian map of interhemispheric brain connections predates the evolution of the corpus callosum. Proc. Natl. Acad. Sci. USA *115*, 9622–9627. https://doi.org/10.1073/pnas.1808262115.

Suzuki, S., Cross, L., and O'Doherty, J.P. (2017). Elucidating the underlying components of food valuation in the human orbitofrontal cortex. Nat. Neurosci. *20*, 1780–1786. https://doi.org/10.1038/s41593-017-0008-x.

Tang, H., Bartolo, R., and Averbeck, B.B. (2021). Reward-related choices determine information timing and flow across macaque lateral prefrontal cortex. Nat. Commun. *12*, 894. https://doi.org/10.1038/s41467-021-20943-9.

Tervo, D.G.R., Kuleshova, E., Manakov, M., Proskurin, M., Karlsson, M., Lustig, A., Behnam, R., and Karpova, A.Y. (2021). The anterior cingulate cortex directs exploration of alternative strategies. Neuron *109*, 1876–1887.e6. https://doi.org/10.1016/j.neuron.2021.03.028.

Tervo, D.G.R., Proskurin, M., Manakov, M., Kabra, M., Vollmer, A., Branson, K., and Karpova, A.Y. (2014). Behavioral variability through stochastic choice and its gating by anterior cingulate cortex. Cell *159*, 21–32. https://doi.org/10.1016/j.cell.2014.08.037.

Tolman, E.C. (1948). Cognitive maps in rats and men. Psychol. Rev. *55*, 189–208. https://doi.org/10.1037/h0061626.

Trudel, N., Scholl, J., Klein-Flügge, M.C., Fouragnan, E., Tankelevitch, L., Wittmann, M.K., and Rushworth, M.F.S. (2021). Polarity of uncertainty representation during exploration and exploitation in ventromedial prefrontal cortex. Nat. Hum. Behav. *5*, 83–98. https://doi.org/10.1038/s41562-020-0929-3.

Uylings, H.B., and van Eden, C.G. (1990). Qualitative and quantitative comparison of the prefrontal cortex in rat and in primates, including humans. Prog. Brain Res. *85*, 31–62.

Vassena, E., Deraeve, J., and Alexander, W.H. (2020). Surprise, value and control in anterior cingulate cortex during speeded decision-making. Nat. Hum. Behav. *4*, 412–422. https://doi.org/10.1038/s41562-019-0801-5.

Vertechi, P., Lottem, E., Sarra, D., Godinho, B., Treves, I., Quendera, T., Oude Lohuis, M.N., and Mainen, Z.F. (2020). Inference-based decisions in a hidden state foraging task: differential contributions of prefrontal cortical areas. Neuron *106*, 166–176.e6. https://doi.org/10.1016/j.neuron.2020.01.017.

Vogt, B.A. (2009). Architecture, neurocytology, and comparative organization of monkey and human cingulate cortices. In Cingulate Neurobiology and Disease, B.A. Vogt, ed. (Oxford University Press).

Wallis, J.D. (2011). Cross-species studies of orbitofrontal cortex and value-based decision-making. Nat. Neurosci. *15*, 13–19. https://doi.org/10.1038/nn.2956.

Walton, M.E., Bannerman, D.M., Alterescu, K., and Rushworth, M.F.S. (2003). Functional specialization within medial frontal cortex of the anterior cingulate for evaluating effort-related decisions. J. Neurosci. *23*, 6475–6479.

Walton, M.E., Bannerman, D.M., and Rushworth, M.F.S. (2002). The role of rat medial frontal cortex in effort-based decision making. J. Neurosci. *22*, 10996–11003.

Walton, M.E., Behrens, T.E., Buckley, M.J., Rudebeck, P.H., and Rushworth, M.F. (2010). Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. Neuron *65*, 927–939. https://doi.org/10.1016/j.neuron.2010.02.027.

Wang, M.Z., and Hayden, B.Y. (2021). Latent learning, cognitive maps, and curiosity. Curr. Opin. Behav. Sci. *38*, 1–7. https://doi.org/10.1016/j.cobeha.2020.06.003.

Wang, X.J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. Neuron *36*, 955–968.

Wang, X.J. (2008). Decision making in recurrent neuronal circuits. Neuron *60*, 215–234. https://doi.org/10.1016/j.neuron.2008.09.034.

Wang, F., Howard, J.D., Voss, J.L., Schoenbaum, G., and Kahnt, T. (2020). Targeted Stimulation of an Orbitofrontal Network Disrupts Decisions Based on Inferred, Not Experienced Outcomes. J. Neurosci. Off. J. Soc. Neurosci. *40*, 8726–8733. https://doi.org/10.1523/JNEUROSCI.1680-20.2020.

Wang, F., Schoenbaum, G., and Kahnt, T. (2020). Interactions between human orbitofrontal cortex and hippocampus support model-based inference. PLoS Biol. *18*, e3000578. https://doi.org/10.1371/journal.pbio.3000578.

Wang, Y., Toyoshima, O., Kunimatsu, J., Yamada, H., and Matsumoto, M. (2021). Tonic firing mode of midbrain dopamine neurons continuously tracks reward values changing moment-by-moment. Elife *10*, e63166. https://doi.org/10.7554/eLife.63166.

Webb, R., Glimcher, P.W., and Louie, K. (2020). Divisive normalization does influence decisions with multiple alternatives. Nat. Hum. Behav. *4*, 1118–1120. https://doi.org/10.1038/s41562-020-00941-5.

White, J.K., Bromberg-Martin, E.S., Heilbronner, S.R., Zhang, K., Pai, J., Haber, S.N., and Monosov, I.E. (2019). A neural network for information seeking. Nat. Commun. *10*, 5168. https://doi.org/10.1038/s41467-019-13135-z.

Whittington, J.C.R., Muller, T.H., Mark, S., Chen, G., Barry, C., Burgess, N., and Behrens, T.E.J. (2020). The Tolman-Eichenbaum machine: unifying space and relational memory through generalization in the hippocampal formation. Cell *183*, 1249–1263.e23. https://doi.org/10.1016/j.cell.2020.10.024.

Wikenheiser, A.M., and Schoenbaum, G. (2016). Over the river, through the woods: cognitive maps in the hippocampus and orbitofrontal cortex. Nat. Rev. Neurosci. *17*, 513–523. https://doi.org/10.1038/nrn.2016.56.

Wise, S.P. (2008). Forward frontal fields: phylogeny and fundamental function. Trends Neurosci *31*, 599–608. https://doi.org/10.1016/j.tins.2008.08.008.

Wittmann, M.K., Fouragnan, E., Folloni, D., Klein-Flügge, M.C., Chau, B.K.H., Khamassi, M., and Rushworth, M.F.S. (2020). Global reward state affects learning and activity in raphe nucleus and anterior insula in monkeys. Nat. Commun. *11*, 3771. https://doi.org/10.1038/s41467-020-17343-w.

Wittmann, M.K., Kolling, N., Akaishi, R., Chau, B.K., Brown, J.W., Nelissen, N., and Rushworth, M.F. (2016a). Predictive decision making driven by multiple time-linked reward representations in the anterior cingulate cortex. Nat. Commun. *7*, 12327. https://doi.org/10.1038/ncomms12327.

Wittmann, M.K., Kolling, N., Faber, N.S., Scholl, J., Nelissen, N., and Rushworth, M.F. (2016b). Self-Other Mergence in the Frontal Cortex during Cooperation and Competition. Neuron *91*, 482–493. https://doi.org/10.1016/j.neuron.2016.06.022.

Wittmann, M.K., Lockwood, P.L., and Rushworth, M.F.S. (2018). Neural mechanisms of social cognition in primates. Annu. Rev. Neurosci. *41*, 99–118. https://doi.org/10.1146/annurev-neuro-080317-061450.

Wittmann, M.K., Trudel, N., Trier, H.A., Klein-Flügge, M.C., Sel, A., Verhagen, L., and Rushworth, M.F.S. (2021). Causal manipulation of self-other mergence in the dorsomedial prefrontal cortex. Neuron *109*, 2353–2361.e11. https://doi.org/10.1016/j.neuron.2021.05.027.

Wong, K.F., and Wang, X.J. (2006). A recurrent network mechanism of time integration in perceptual decisions. J. Neurosci. *26*, 1314–1328. https://doi.org/10.1523/JNEUROSCI.3733-05.2006.

Yoshida, K., Saito, N., Iriki, A., and Isoda, M. (2011). Representation of others' action by neurons in monkey medial frontal cortex. Curr. Biol. *21*, 249–253. https://doi.org/10.1016/j.cub.2011.01.004.

Yoshida, K., Saito, N., Iriki, A., and Isoda, M. (2012). Social error monitoring in macaque frontal cortex. Nat. Neurosci. *15*, 1307–1312. https://doi.org/10.1038/nn.3180.

Yun, M., Kawai, T., Nejime, M., Yamada, H., and Matsumoto, M. (2020). Signal dynamics of midbrain dopamine neurons during economic decision-making in monkeys. Sci. Adv. *6*, eaba4962. https://doi.org/10.1126/sciadv.aba4962.

Zajkowski, W.K., Kossut, M., and Wilson, R.C. (2017). A causal role for right frontopolar cortex in directed, but not random, exploration. Elife *6*, e27430. https://doi.org/10.7554/eLife.27430.

Zhou, J., Gardner, M.P.H., Stalnaker, T.A., Ramus, S.J., Wikenheiser, A.M., Niv, Y., and Schoenbaum, G. (2019). Rat orbitofrontal ensemble activity contains multiplexed but dissociable representations of value and task structure in an odor sequence task. Curr. Biol. *29*, 897–907.e3. https://doi.org/10.1016/j.cub.2019.01.048.

Zhou, J., Jia, C., Montesinos-Cartagena, M., Gardner, M.P.H., Zong, W., and Schoenbaum, G. (2021). Evolving schema representations in orbitofrontal ensembles during learning. Nature *590*, 606–611. https://doi.org/10.1038/s41586-020-03061-2.

Zuberbühler, K., and Janmaat, K. (2010). Foraging cognition in nonhuman primates. In Primate Neuroethology, M. Platt and A. Ghazanfar, eds. (Oxford University Press), pp. 64–83.