Valuation and decision-making in frontal cortex: one or many serial or parallel systems?

Matthew FS Rushworth[1,2], Nils Kolling[1], Jérôme Sallet[1] and Rogier B Mars[1,2]

We evaluate the merits of different conceptualizations of frontal cortex function in value-guided decision-making. According to one view each frontal cortical region is concerned with a different aspect of the process of learning about and evaluating choices and then selecting actions. An alternative view, however, sees sets of decision-making circuits working in parallel within the frontal lobes in order to make different types of decisions. While there is a neural circuit for making choices between pairs of simultaneously presented items in the manner that is frequently assessed in the laboratory, there is also evidence that other frontal lobe circuits have evolved to make other types of choices such as those made during the course of foraging.

Addresses
[1] Department of Experimental Psychology, University of Oxford, United Kingdom
[2] Oxford Centre for Functional Magnetic Resonance Imaging of the Brain (FMRIB), University of Oxford, United Kingdom

Corresponding author: Rushworth, Matthew FS (matthew.rushworth@psy.ox.ac.uk)

Value assignment and prefrontal cortex

An emerging and influential account of frontal brain mechanisms of decision-making holds that what we and other animals do when we make a choice is to decide between two different goods on the basis of their independently computed reward values [1]. The lateral orbitofrontal cortex (lOFC) plays a central role in learning the values associated with different goods. LOFC lesions disrupt the assignment of precise values to stimuli [2••]. The representations of values in this area are appropriate for guiding choices because they exhibit a phenomenon called range adaptation that means that the same neurons can inform decisions made between stimuli associated with two small rewards of only slightly differing sizes and, on a different occasion, between stimuli linked to very large rewards [3••,4,5••].

Despite being adjacent and interconnected with one another lOFC and ventromedial prefrontal cortex/medial orbitofrontal cortex (vmPFC/mOFC) are distinct brain regions with disparate connections to other brain areas and they have different roles in learning and decision-making [6]. Although a prominent view [7] remains that lOFC and vmPFC/mOFC are, respectively, more concerned with negative and positive outcomes the balance of evidence now argues against this view. There is no spatial separation within OFC between areas encoding reward and punishment [8,9] and lesions of lOFC and vmPFC/mOFC do not have contrasting effects on sensitivity to punishment and reward [10,11]. Instead both lesion and neuroimaging studies suggest that lOFC is concerned with updating value representations on the basis of both negative and positive outcomes [2••,12••,13] (Figure 1a). Moreover, lOFC is concerned with linking stimuli not just with representations of the scalar value of a reward but also with representations of the nature, identity or type of reward that might be expected to follow choices [2••,12••,13] (Figure 2a–c).

A possible serial circuit in frontal cortex for making a value-guided choice

In contrast to lOFC, vmPFC/mOFC appears more intimately concerned with the use of reward representations to guide behaviour (Figure 1b). Little is known about neurons in vmPFC/mOFC but it is clear that while they represent rewards they differ from lOFC neurons because they encode little about the stimuli that are associated with the rewards [14]. Several human neuroimaging studies have reported vmPFC/mOFC blood oxygen level dependent (BOLD) signals that are proportional to reward expectations at the time of decision-making [9,12••,15–19] (Figure 2b).

It has been argued that these value representations are then the input to a distinct comparison process that takes place elsewhere in the brain — in dorsal anterior cingulate cortex (ACC) or adjacent dorsomedial prefrontal cortex (dmPFC) [20–22] (Figure 1b). According to this view an analogy is drawn between a vmPFC/mOFC–ACC/dmPFC circuit in which values are first represented and then compared and the circuit comprised of the visual motion area V5/MT and lateral intraparietal area (LIP) [23]. In the V5/MT-LIP circuit visual motion is first represented and then signals are compared in order to make an oculomotor decision.

Figure 1
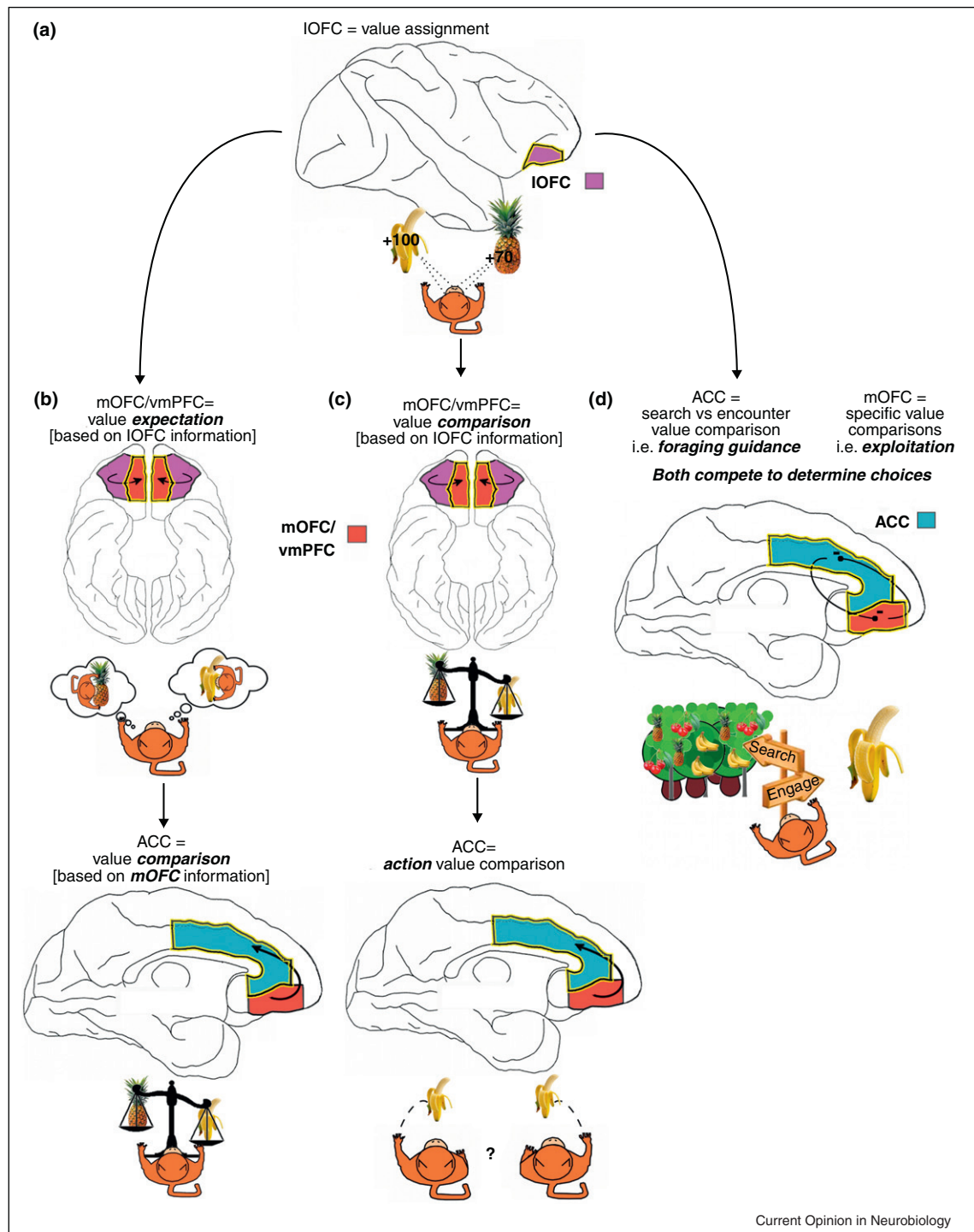
Values and reward type information are assigned to different options in lOFC **(a)**. There are then different possible ways in which valuation and decision-making mechanisms might proceed in frontal cortex. They might proceed in series with reward expectations represented in vmPFC/mOFC and the actual process of value comparison occurring in ACC **(b)**. Alternatively there may be distinct mechanisms for, first, making decisions about the rewards that should be the focus of behaviour and attention and, second, for making decisions about the actions that should be made to obtain those rewards **(c)**. According to an alternative scheme there are parallel systems for foraging and decision-making **(d)**. According to this last view the ACC, which exists albeit in different forms in all mammals, might guide foraging, a key behaviour for all mammals, while vmPFC, a primate specialization [49] might be especially important when decisions are made between simultaneously available options. Other authors have emphasized other differences between these regions, for example, emphasizing that ACC may have a greater role in reward-guided learning because of the prominence of prediction errors in this brain region [5••], the presence of reward outcome-related activity reflecting different mnemonic time constants [50••], and surprise [51,52].

Figure 2
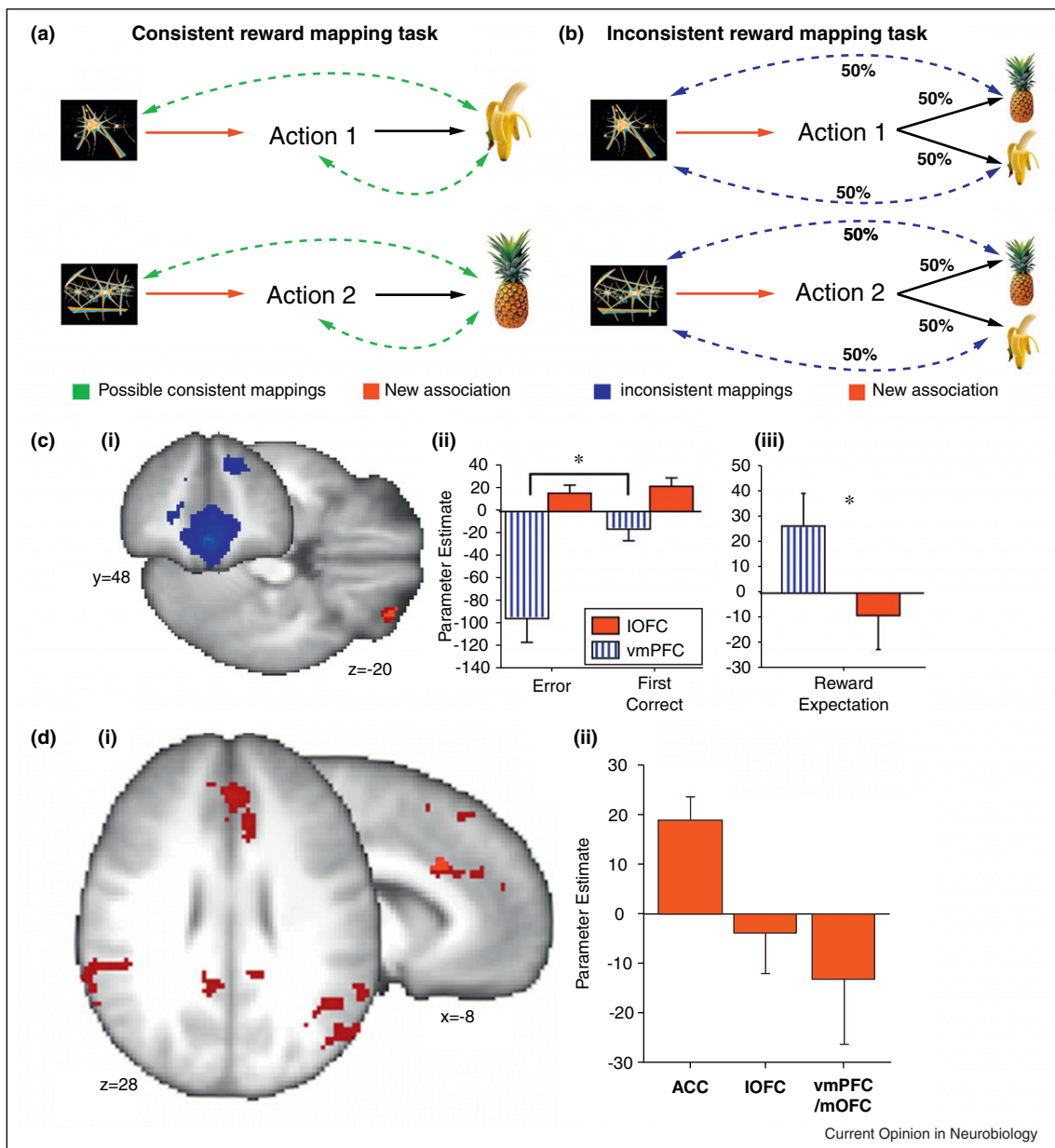
Distinct roles for lOFC, vmPFC/mOFC and ACC in, respectively, learning and updating representations of specific types of reward associated with stimuli, representations of reward value regardless of type, and linking rewards to action selection. **(a)** Schematic representation of the consistent reward mapping task. Subjects learn which action, action 1 or action 2, to make in response to seeing a stimulus, stimulus 1 or stimulus 2 in order to receive a reward. A correct response of action 1 to stimulus 1 is always reinforced with reward outcome 1. A correct response of action 2 to stimulus 2 is always reinforced with reward outcome 2. This means that subjects can select responses via S–R associations (red lines) or via associations between specific stimuli and specific reward outcomes and between specific reward outcome representations and responses (S–O and R–O associations, green lines). **(b)** Schematic representation of the inconsistent reward mapping task. As in the consistent reward mapping task subjects learn which action, action 1 or action 2, to make in response to seeing a stimulus, stimulus 1 or stimulus 2 in order to receive a reward. A correct response of action 1 to stimulus 1 is always reinforced but it may be reinforced with reward outcome 1 or 2. A correct response of action 2 to stimulus 2 is always reinforced with but again it may be reinforced with reward outcome 1 or 2. This means that, unlike in the consistent reward mapping task, subjects can only select responses via S–R associations (red) because specific S–O or R–O associations cannot be learned. **(ci)** lOFC (red) activity reflected not just error-feedback but any feedback, such as the first occurrence of positive feedback for a particular choice, that allowed updating the association between a stimulus and reward. Because feedback activity also differed between Consistent and Inconsistent groups it seemed that lOFC was learning expectations about particular types of reward rather than just reward per se. By contrast vmPFC/mOFC activity simply reflects reward value; its activity is greater when positive outcomes are received regardless of how informative they are for learning about reward associations (ci) and reward expectations **(cii)**. By contrast ACC activity varied between the two groups (greater in Consistent mapping group) as a function of how likely they were to select the correct response **(di)**. The ACC activity probably reflects the strength of reward-guided action selection mechanisms because the groups differ in how much they can use reward to guide response selection (it is possible in the Consistent group but not the Inconsistent group). Similar effects are not found in vmPFC/mOFC and lOFC **(dii)**.

One reason that value comparison has been thought to occur in ACC/dmPFC is that it has a prominent *inverse value difference* signal [20–22]. The ACC/dmPFC BOLD signal increases when the difference between the values of potential choices *decreases*. In several cases researchers have argued that formal models of value comparison in decision-making based on drift diffusion processes precisely predict the size of inverse value signals seen in ACC/dmPFC [20–22]. More simply the intuition is that if the options' values are close together then the comparison process will be more difficult and time consuming. According to this account it is the output of the ACC/dmPFC comparator that determines the pattern of activity seen in the motor system [22] and the response ultimately made. The response is seen as being at the end of a serial chain of processes in frontal cortex.

Different decisions in frontal cortex: decisions about rewards and about actions

An alternative interpretation is that vmPFC/mOFC is an important determinant of value-guided decisions (Figure 1c). Not only does vmPFC/mOFC activity level reflect the value of the goal that is currently being pursued [6,15] but it also represents the values of both options that are being considered during the course of a decision [24,25••,26]. There is a positive relationship between vmPFC/mOFC BOLD and the value of the choice taken, and a negative relationship with the value of the choice rejected [24]. The vmPFC/mOFC *value difference* signal is the opposite of the ACC/dmPFC inverse value signal; it increases with difference in value between choices. One interpretation of this signal is that it reflects the choice between two potential reward goals being made and the antagonistic interactions between two mutually inhibitory representations of the possible reward goals. Ultimately just a single representation comes to dominate vmPFC/mOFC as a choice is made. This is a challenge to some serial models of decision-making because it suggests that decisions are already being made by the time that information reaches vmPFC/mOFC.

Although a value difference signal is apparent in vmPFC/mOFC in several studies [20–22,25••] its interpretation is contested. An alternative view is that it reflects which options are attended rather than selected [22,25••,27]. Part of the reason a vmPFC/mOFC value signal is difficult to interpret is because it has not been clear what sort of value signal a decision-making mechanism, as opposed to some other mechanism, ought to generate. Hunt *et al.* [28••] therefore adapted a biophysical model of decision-making in the oculomotor system [29] to value-guided decision-making. Central to the model is the notion of pools of neurons representing each choice becoming more active as the evidence favouring the choice they encode increases. There are competitive, inhibitory interactions between the two pools of neurons so that ultimately only

one remains active. VmPFC/mOFC possessed signals with the same dynamics as the mean field output of the biophysical model.

Further support for a decision-making account of vmPFC/mOFC comes from vmPFC/mOFC lesions, which, unlike lOFC lesions, do not disrupt the updating of value assignments, but they do disrupt value-guided choice especially when the values of potential choices are close [10]. When choices are made between multiple stimuli then the way in which the values of any two are compared becomes less accurate if a third option is also high in value [10]. Value assignments may be maintained separately from value comparison mechanisms so that the inhibitory interactions that occur during comparison do not impact on value assignments in other contexts (see also [30••]).

Accounts of vmPFC/mOFC that emphasize decision-making [10,24] or attention [25••] can be reconciled to some degree by realizing that if vmPFC/mOFC does make decisions then it does not do so in the sense of choosing an action but instead it decides the reward goal that will be the focus of behaviour, especially when there are multiple competing alternatives and attention must be directed while critical comparisons are made. Such a view explains why vmPFC/mOFC lesions lead to value comparison impairments that are augmented by the presence of additional distracting options [10].
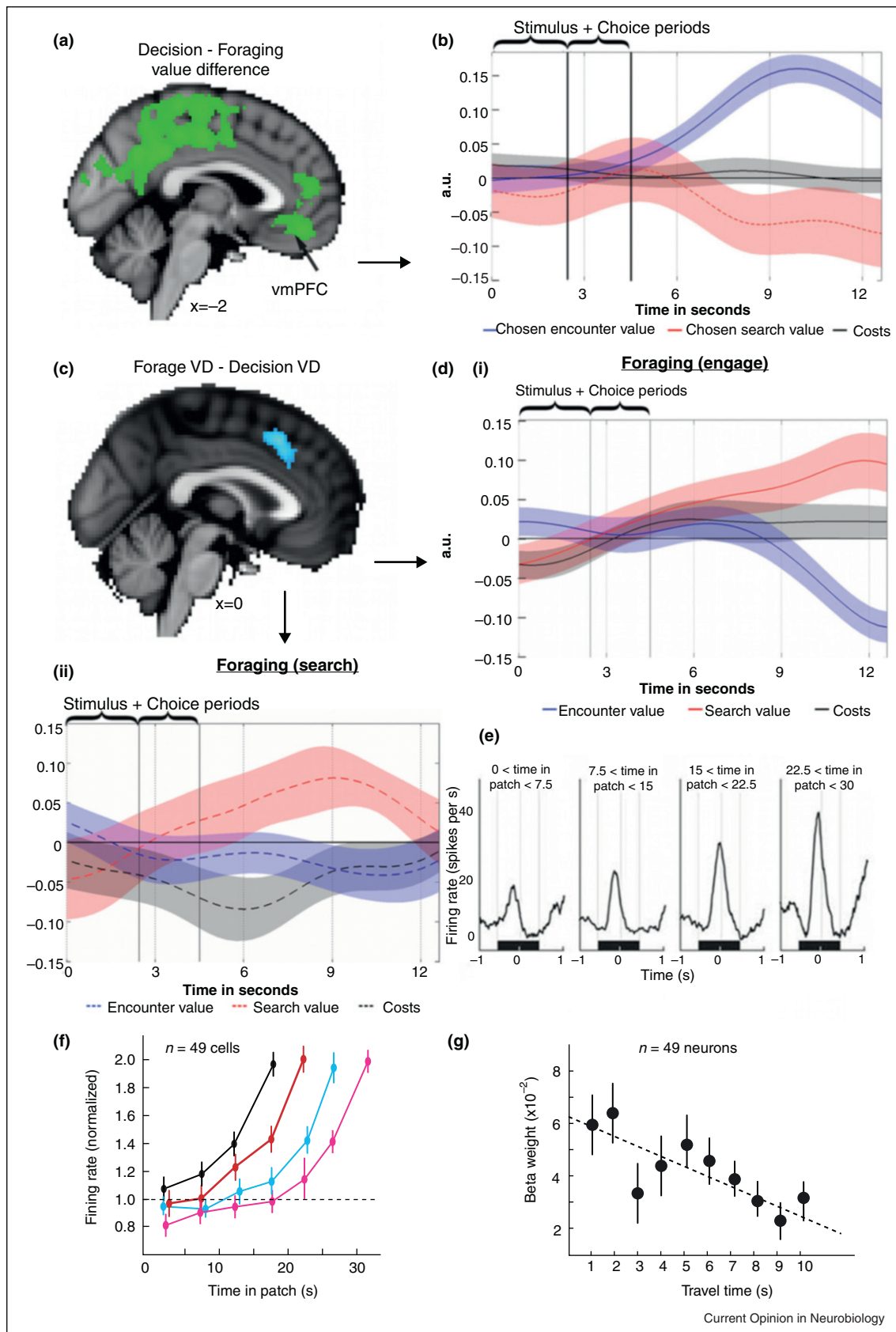
Once the reward goal is selected then ACC is important for guiding a second type of decision — a decision about which response should be made in order to obtain the reward that is the focus of behaviour [12••] (Figures 1c and 2d). The properties and timing of neuronal activity in ACC are broadly consistent with this proposal [31••].

The doubly dissociable effects of lesions provide a further reason for thinking that the decisions made by vmPFC/mOFC and ACC are distinct and independent. Lesions in lOFC or vmPFC/mOFC in monkeys [32] and humans [33] disrupt stimulus-reward but not action-reward association. By contrast ACC lesions in both species have the opposite effect. Such double dissociations are more consistent with a second view of decision-making mechanisms in which more than one frontal cortical system is making decisions albeit of different types (Figure 1c).

Parallel decision-making mechanisms

Perhaps the most intriguing question is whether there is a single valuation system or many valuation systems (Figure 1d). For example, a distinction has been drawn between habitual and goal-based decision-making [34••,35] and there is evidence for differences between social and non-social decision-making [36]. There may, however, be other types of fundamental distinctions to be drawn between types of value-guided choice.

Figure 3

(a) Decision - Foraging value difference

vmPFC

x=−2

(b) Stimulus + Choice periods

a.u.

Time in seconds

— Chosen encounter value   — Chosen search value   — Costs

(c) Forage VD - Decision VD

x=0

(d) (i) **Foraging (engage)**

Stimulus + Choice periods

a.u.

Time in seconds

— Encounter value   — Search value   — Costs

(ii) **Foraging (search)**

Stimulus + Choice periods

Time in seconds

--- Encounter value   --- Search value   --- Costs

(e)

Firing rate (spikes per s)

0 < time in patch < 7.5   7.5 < time in patch < 15   15 < time in patch < 22.5   22.5 < time in patch < 30

Time (s)

(f) n = 49 cells

Firing rate (normalized)

Time in patch (s)

(g) n = 49 neurons

Beta weight (×10⁻²)

Travel time (s)

Current Opinion in Neurobiology

It is clear that we and other primates can make the binary comparative decisions that are at the heart of most of the behavioural paradigms we investigate. Nevertheless it is not so obvious that these are the choices that we have always evolved to make. Instead we and many other animals have evolved to forage. Crucially during foraging animals rarely encounter two choice options simultaneously [37••]. Instead the critical choice to make during foraging is whether to engage with an option when it is encountered or whether better prospects are likely elsewhere in the environment. Such choices require weighing up the value of the option encountered (encounter value), the richness of the environment (search value), and the cost of searching elsewhere (search cost) which is often energetically demanding. We compared foraging-style choices, in which human subjects chose whether to engage with an option (with a known *encounter value*) or search for potentially better alternatives (*search value*), at the risk of also paying a *search cost*, with interleaved trials in which choices were made in the context of binary decisions [38]. VmPFC/mOFC and ACC were, respectively, more active during the two types of choices — decisions and forages (Figure 3). Individual differences in ACC and vmPFC/mOFC signal strength, respectively, also correlated with individual differences in foraging and decision behaviour. Moreover, ACC activity reflected search values, encounter values and search costs during forages but search value and search cost signals were absent from vmPFC/mOFC.

Perhaps most strikingly the ACC signal was revealed *not* to be an *inverse value* signal as had previously been thought (Figure 1b and section: *A possible serial circuit in frontal cortex for making a value-guided choice*); ACC BOLD was *not* always positively correlated with the value of the unchosen option and negatively correlated with the value of the chosen option. Instead, during forages, ACC BOLD correlated positively with search value and negatively with the encounter value regardless of which choice subjects made (Figure 3d). This has two implications. First, if ACC does not carry an inverse value signal then it argues against it having a simple and ubiquitous role in value comparison (Figure 1b). Second it suggests instead that ACC signals promote a particular behaviour: searching or exploring the environment for better alternatives to the course of action currently being pursued (Figure 1d).

Although ACC activity was always positively correlated with the value of searching, regardless of the choice taken, the search signal increased more rapidly when subjects chose to search [38] (Figure 3d). Similarly single neurons in ACC that respond to reward receipt do so with higher firing rates as monkeys move towards leaving a known patch to search for a new one in a foraging task [39••]. Moreover, the signal gain was inversely proportional to search costs (Figure 3e–g).

If ACC is a mechanism for valuation and promotion of behavioural change and search then this may explain a number of findings. The action-reward learning tasks that are impaired by ACC lesions [32,33] typically involve alternation between actions but no informative stimuli. Repetitive selection of an action interleaved with periods of exploration of alternative actions may be just the sort of behaviour that is normally under the control of a foraging system. It may explain why, despite neuroimaging findings, it has proven difficult to identify response conflict or value comparison neurons in ACC [31••,40,41••] but why ACC activity is prominent when monkeys explore a new situation [42]. It may also explain why ACC encodes counterfactual feedback — information that is given about what consequences a choice would have had had it been taken — which can also promote behavioural change [43•,44] and some features of ACC activity in dual-task experiments [45].

In human imaging experiments it is not just ACC that is implicated in foraging but also anterior prefrontal cortex (aPFC). Differences in the signals carried by aPFC and ACC are beginning to emerge [24,43•] (Figure 4a–d). While aPFC also encodes information about alternative courses of action it does so in a different way to ACC; instead of representing the average value of alternative options it represents the value of the best alternative option. Moreover, while the coding of value in ACC occurs in a fixed framework (there is always a positive relationship between activity and the value of searching regardless of the choice taken) by contrast aPFC encoding is more flexible. At the instant that a new course of action is about to be pursued aPFC activity is already encoding the value of the best alternative to the new course of action. Whether a region similar to aPFC exists in other primates is unclear; human aPFC is near the frontal pole but it has a pattern of connectivity with

---

(Figure 3 Legend) While vmPFC/mOFC is more active during decision-making than foraging **(a, b)**, an ACC region is more active during foraging than decision-making **(c)** [38]. During foraging, when subjects choose between engaging with particular options or searching for better alternatives ACC BOLD is positively correlated with the search value and negatively correlated with the encounter value regardless of whether subjects choose to stick with the option encountered **(di)** or to search for potential alternatives **(dii)**. The search cost is also represented when subjects choose to search. By contrast vmPFC/mOFC BOLD is positively correlated with the encounter value, when it is chosen but there is no representation of search values or costs regardless of the choice ultimately made **(b)**. The activity of neurons in ACC in monkeys also indicate a role in foraging. Outcome-related activity increases with each outcome that is received for foraging in a given patch in an example neuron **(e)** and, on average, the gain of the outcome-related response with time in the patch co-varied with the speed of departure from the patch **(f)**. The rate of gain was a function of the search costs that were to be paid; here rate of gain is indicated by a regression slope (beta weight) relating time spent foraging in a patch with firing rate and it can be seen that it decreases with the search cost that will have to be paid in order to travel to a new patch [39••].

Figure 4
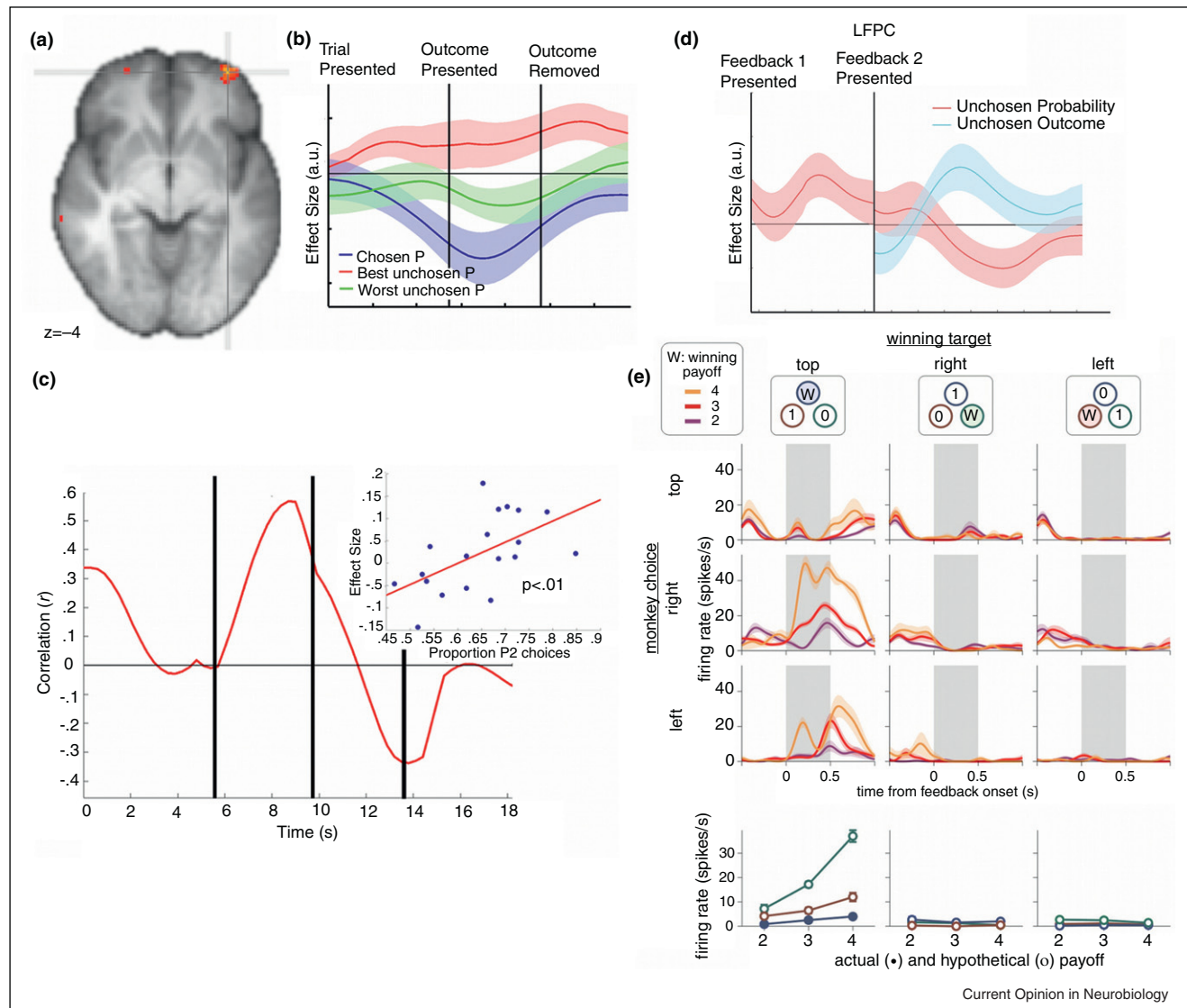
An aPFC region **(a)** encodes the value of switching to the best alternative action [43•]. BOLD activity is positively correlated with the value of switching to the best alternative or currently unchosen action but it is negatively correlated with the value of the option that is being chosen and with value of the worst possible alternative **(b)**. One interpretation of such a signal is that it is representing the value of switching to the best alternative course of action together with the opportunity that would be lost of not taking the action is currently being pursued or the other, worse, alternative. The difference between the signals encoding the value of switching to the best alternative and the worst alternative varied between subjects and the variation was related to individual differences in how frequently subjects would go on to the pick the best alternative on the next trial. The correlation between the BOLD effect (the standardized regression coefficient relating BOLD to the difference in value between the best unchosen action and the worse unchosen action) and the probability of switching to the best unchosen option on the next trial is shown throughout the trial. The inset shows the relationship in detail at the time of the peak size of the BOLD effect (note because it is the peak size of the BOLD effect, rather than the correlation plotted in the main graph, the selection of the scatter plot is unbiased with respect to the correlation analysis). The same region also encodes a counterfactual prediction error on unchosen options **(c)**. Before feedback there is a positive relationship between BOLD and the prior expectation of reward for the unchosen option. After the counterfactual feedback is delivered there is a negative relationship between BOLD and the prior expectation of reward for the unchosen option but a positive relationship between BOLD and counterfactual reward delivery **(d)**. Spike density function of a neuron in dorsolateral prefrontal cortex encoding a counterfactual or fictive prediction error signal [47] **(e)**. The three columns correspond to trials in which the top (left column), right (middle column), or left (right column) stimulus position was the winning choice for a monkey to make. The three rows correspond to the monkey's actual choice being made to the top position (top row), right position (middle row), left position (bottom row). The neuron's activity was greatest when the best outcome would have been delivered for making the 'top' response (left hand column) but not when the top response was actually the one that was made (top left hand corner).

parietal cortex that makes it unlike any frontal polar area in the monkey [46]. Instead, despite differences in its gross anatomical position, human aPFC may have features that resemble dorsolateral prefrontal cortex in monkeys; in monkeys counterfactual outcome-related activity has been reported in dorsolateral prefrontal cortex [47] but not in the frontal pole [48] (Figure 4e).

Conclusions

In summary, a goods-based account of decision-making provides a persuasive description of several frontal cortical brain regions as well as of the decisions of greatest interest to economists. We may not, however, have evolved to make only such decisions. Other regions such as aPFC and ACC, that encode the best alternative action to the one that is currently being taken or the average richness of the foraging environment, also influence choice (Figures 1d, 3 and 4). Their existence suggests that we do not always evaluate every possible choice that we might pursue in the same way but that there are representations of particular alternative courses of action that are especially privileged, such as the best alternative to what we are now doing, and that sometimes what we do is simply to decide whether it is better to stay with the current course of action or if the richness of the environment suggests that searching elsewhere is merited.

Acknowledgements

References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as:

- •• of outstanding interest

1. Padoa-Schioppa C: **Neurobiology of economic choice: a good-based model**. *Annu Rev Neurosci* 2011, **34**:333-359.

2. Walton ME, Behrens TE, Buckley MJ, Rudebeck PH,
•• Rushworth MF: **Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning**. *Neuron* 2010, **65**:927-939.
LOFC lesions disrupt the assignment of precise values to stimuli on the basis of the conjoint history of choices that have been made and rewards that were received contingent on those choices. Instead, in the absence of lOFC, macaques assign a value to a stimulus that corresponds to the recency-weighted history of all rewards received for all choices, both of the particular stimulus in question and all other stimuli chosen shortly beforehand or afterwards.

3. Padoa-Schioppa C: **Range-adapting representation of**
•• **economic value in the orbitofrontal cortex**. *J Neurosci* 2009, **29**:14004-14014.
Although a given neuron may always encode an object's value with changes in firing rate the gain of such encoding, in other words the slope relating value to the neuron's firing rate, changes with circumstances. On days when choices were made between only low value stimuli neurons still exhibited a large range of firing rates when encountering different choice options as if the range of possible firing rates had adapted to the range of possible values that were being encountered.

4. Kobayashi S, Pinto de Carvalho O, Schultz W: **Adaptation of reward sensitivity in orbitofrontal neurons**. *J Neurosci* 2010, **30**:534-544.

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

Kennerley *et al.* reported that OFC neuron firing rates for a given choice stimulus were high when recent reward levels had been low and low if recent reward levels had been high. Such a phenomenon might be expected if OFC neurons are to show range adaptation [3••]. Similar activity changes were, however, absent in ACC. By contrast ACC neurons were more likely to multiplex several aspects of value in a systematic way (e.g. activity in a given neuron might increase with expected reward probability and decrease with expected effort) and to encode reward prediction errors.

6. Rushworth MF, Noonan MP, Boorman ED, Walton ME, Behrens TE: **Frontal cortex and reward-guided learning and decision-making**. *Neuron* 2011, **70**:1054-1069.

7. Milad MR, Rauch SL: **Obsessive–compulsive disorder: beyond segregated cortico-striatal pathways**. *Trends Cogn Sci* 2012, **16**:43-51.

8. Morrison SE, Salzman CD: **The convergence of information about rewarding and aversive stimuli in single neurons**. *J Neurosci* 2009, **29**:11471-11483.

9. Plassmann H, O'Doherty JP, Rangel A: **Appetitive and aversive goal values are encoded in the medial orbitofrontal cortex at the time of decision making**. *J Neurosci* 2010, **30**:10799-10808.

10. Noonan MP, Walton ME, Behrens TE, Sallet J, Buckley MJ, Rushworth MF: **Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex**. *Proc Natl Acad Sci U S A* 2010, **107**:20547-20552.

11. Noonan M, Kolling N, Walton M, Rushworth M: **Re-evaluating the role of orbitofrontal cortex in reward and reinforcement**. *Eur J Neurosci* 2012, **35**:997-1010.

12. Noonan MP, Mars RB, Rushworth MF: **Distinct roles of three**
•• **frontal cortical areas in reward-guided behavior**. *J Neurosci* 2011, **31**:14399-14412.
Noonan *et al.* compared activity across the whole brain in human subjects who could use the prospect of particular rewards to guide response selection with the activity seen in a second group of subjects unable to use reward expectations to guide response selection. There was significantly greater activity in the ACC in the first group of subjects and the increase was greater as a function of how likely the subjects were to make their action choices correctly. If anything ACC activation in the group that could not use reward to guide response selection was associated with slightly worse performance.

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

14. Bouret S, Richmond BJ: **Ventromedial and orbital prefrontal neurons differentially encode internally and externally driven motivational values in monkeys**. *J Neurosci* 2010, **30**:8591-8601.

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

16. Smith DV, Hayden BY, Truong TK, Song AW, Platt ML, Huettel SA: **Distinct value signals in anterior and posterior ventromedial prefrontal cortex**. *J Neurosci* 2010, **30**:2490-2495.

17. Litt A, Plassmann H, Shiv B, Rangel A: **Dissociating valuation and saliency signals during decision-making**. *Cereb Cortex* 2011, **21**:95-102.

18. Gershman SJ, Pesaran B, Daw ND: **Human reinforcement learning subdivides structured action spaces by learning effector-specific values**. *J Neurosci* 2009, **29**:13524-13531.

19. Palminteri S, Boraud T, Lafargue G, Dubois B, Pessiglione M: **Brain hemispheres selectively track the expected value of contralateral options**. *J Neurosci* 2009, **29**:13465-13472.

20. Basten U, Biele G, Heekeren HR, Fiebach CJ: **How the brain integrates costs and benefits during decision making**. *Proc Natl Acad Sci U S A* 2010, **107**:21767-21772.

21. Philiastides MG, Biele G, Heekeren HR: **A mechanistic account of value computation in the human brain**. *Proc Natl Acad Sci U S A* 2010, **107**:9430-9435.

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

23. Gold JI, Shadlen MN: **The neural basis of decision making**. *Annu Rev Neurosci* 2007, **30**:535-574.

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

25. Lim SL, O'Doherty JP, Rangel A: **The decision value**
•• **computations in the vmPFC and striatum use a relative value code that is guided by visual attention**. *J Neurosci* 2011, **31**:13214-13223.
Lim *et al*. demonstrated that which options' values are encoded positively and which negatively within vmPFC/mOFC depends on the direction of attention; the value of the attended choice has a positive relationship with vmPFC/mOFC BOLD. Rather than conceiving of vmPFC/mOFC value difference signals as guiding choice the intriguing interpretation of this result is that the direction of attention determines the rate at which the value signal is integrated into the ACC/dmPFC value comparison process.

26. Wunderlich K, Rangel A, O'Doherty J: **Economic choices can be made using only stimulus values**. *Proc Natl Acad Sci U S A* 2010, **107**:15005-15010.

27. Krajbich I, Armel C, Rangel A: **Visual fixations and the computation and comparison of value in simple choice**. *Nat Neurosci* 2010, **13**:1292-1298.

28. Hunt LT, Kolling N, Soltani A, Woolrich MW, Rushworth MF,
•• Behrens TE: **Mechanisms underlying cortical activity during value-guided choice**. *Nat Neurosci* 2012, **15**:470-476.
A biophysical model of neuronal activity during decision-making predicted low frequency signals in the 2–10 Hz range that initially corresponded to the overall sum of the values of the potential choices on a trial but then reflected the value difference between choices. Using magnetoencephalography (MEG) Hunt *et al*. found that such a sequence of signals was found in vmPFC/mOFC and also superior parietal cortex near the intraparietal sulcus. The presence of overall value [26] and value difference signals [24] in vmPFC/mOFC that have been reported in different studies can be understood as being the outputs at different points in time from the same value comparison process. The decision-related vmPFC/mOFC activity was most apparent when the two facets defining the rewards they used, reward probability and reward magnitude, were in conflict and it was not clear which was the more important.

29. Wang XJ: **Probabilistic decision making by slow reverberation in cortical circuits**. *Neuron* 2002, **36**:955-968.

30. Kable JW, Glimcher PW: **The neurobiology of decision:**
•• **consensus and controversy**. *Neuron* 2009, **63**:733-745.
Kable and Glimcher have also advanced an argument along similar lines to the one we present for separation between valuation and comparison mechanisms although they placed a particular emphasis on oculomotor decision-making mechanisms. Their model treats lOFC and vmPFC/mOFC as a unitary valuation mechanism and argues that value comparisons between possible choices occur in an action space anchored to the parietal and posterior frontal cortical mechanisms for movement selection.

31. Cai X, Padoa-Schioppa C: **Neuronal encoding of subjective**
•• **value in dorsal and ventral anterior cingulate cortex**. *J Neurosci* 2012, **32**:3791-3808.
Neurons in the dorsal bank of the cingulate sulcus are selective for different directions of response and so may encode different actions. At least in some situations, however, they encode values with a longer latency than do neurons in OFC, as if the value information were only present at the time of response selection. Neurons in the ventral bank of ACC formed a distinct population. Their activity was less modulated by movement direction, less apparent at the time of movement initiation, and most apparent after reward delivery.

32. Rudebeck PH, Behrens TE, Kennerley SW, Baxter MG, Buckley MJ, Walton ME, Rushworth MF: **Frontal cortex subregions play distinct roles in choices between actions and stimuli**. *J Neurosci* 2008, **28**:13775-13785.

33. Camille N, Tsuchida A, Fellows LK: **Double dissociation of stimulus-value and action-value learning in humans with orbitofrontal or anterior cingulate cortex damage**. *J Neurosci* 2011, **31**:15048-15052.

34. Wunderlich K, Dayan P, Dolan RJ: **Mapping value based**
•• **planning and extensively trained choice in the human brain**. *Nat Neurosci* 2012, **15**:786-791.
A distinction has been proposed between neural systems that employ values learned through reinforcement learning and model-based, or goal-based decision-making systems that search through a model of the environment and the actions that may be made upon it in order to identify the best choice to make. Wunderlich *et al*. show that model-based and habitual value representations are respectively linked to activity in the caudate and putamen divisions of the striatum; however, comparison of both types of values depends on vmPFC/mOFC.

35. Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ: **Model-based influences on humans' choices and striatal prediction errors**. *Neuron* 2011, **69**:1204-1215.

36. Behrens TE, Hunt LT, Rushworth MF: **The computation of social behavior**. *Science* 2009, **324**:1160-1164.

37. Freidin E, Kacelnik A: **Rational choice, context dependence, and**
•• **the value of information in European Starlings (*Sturnus vulgaris*)**. *Science* 2011, **334**:1000-1002.
Birds were given two types of task in which choice options were either presented simultaneously, as in most decision-making experiments, or sequentially as is thought to occur in most real-life foraging situations. On some trials information about the choice context informed the animals about the other options that were potentially available. Such context information led to improved performance in the sequential foraging-style scenario but depressed performance in the simultaneous decision-making task.

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

39. Hayden BY, Pearson JM, Platt ML: **Neuronal basis of sequential**
•• **foraging decisions in a patchy environment**. *Nat Neurosci* 2011, **14**:933-939.
Hayden taught monkeys a simple computerized foraging task in which they could chose to continue foraging in the same patch for diminishing returns or seek an alternative patch at the expense of paying a cost of a travel delay before foraging could resume. Single neurons in ACC that fire to reward receipt did so at an increasing rate as monkeys moved towards leaving a known patch to search for a new one. Patch moving was initiated when the ACC activity reached a threshold. The threshold firing rate that had to be reached was proportional to the search costs that were to be incurred by switching away from the current foraging patch. The firing to threshold-based mechanism may, therefore, share features with the one proposed in the oculomotor system [23].

40. Hayden BY, Heilbronner SR, Pearson JM, Platt ML: **Surprise signals in anterior cingulate cortex: neuronal encoding of unsigned reward prediction errors driving adjustment in behavior**. *J Neurosci* 2011, **31**:4178-4187.

41. Hayden BY, Platt ML: **Neurons in anterior cingulate cortex**
•• **multiplex information about reward and action**. *J Neurosci* 2010, **30**:3339-3346.
When monkeys were exploring a new situation, oscillatory activity in the gamma range was prominent and correlated in ACC and dorsolateral prefrontal cortex when informative feedback, both positive and negative, indicated what choices should be made in that environment in the future. When monkeys remained in the same situation on trial after trial then correlated oscillatory activity in the two areas was more prominent in the beta range.

42. Rothe M, Quilodran R, Sallet J, Procyk E: **Coordination of high gamma activity in anterior cingulate and lateral prefrontal cortical areas during adaptation**. *J Neurosci* 2011, **31**:11110-11117.

43. Boorman ED, Behrens TE, Rushworth MF: **Counterfactual choice**
• **and learning in a neural network centered on human lateral frontopolar cortex**. *PLoS Biol* 2011, **9**:e1001093.
In monkeys, counterfactual feedback-related activity was proportional to the size of the reward that had been missed and predicted the likelihood of the monkey's changing to pursue that alternative course of action on future trials.

44. Hayden BY, Pearson JM, Platt ML: **Fictive reward signals in the anterior cingulate cortex**. *Science* 2009, **324**:948-950.
During dual-task performance human ACC regions in each of the two hemispheres track the value of responding in each of the two different tasks.

45. Charron S, Koechlin E: **Divided representation of concurrent goals in the human frontal lobes**. *Science* 2010, **328**:360-363.

46. Mars RB, Jbabdi S, Sallet J, O'Reilly JX, Croxson PL, Olivier E, Noonan MP, Bergmann C, Mitchell AS, Baxter MG *et al*.: **Diffusion-weighted imaging tractography-based parcellation of the human parietal cortex and comparison with human and**

macaque resting-state functional connectivity. *J Neurosci* 2011, **31**:4087-4100.

47. Abe H, Lee D: **Distributed coding of actual and hypothetical outcomes in the orbital and dorsolateral prefrontal cortex**. *Neuron* 2011, **70**:731-741.

48. Tsujimoto S, Genovesio A, Wise SP: **Evaluating self-generated decisions in frontal pole cortex of monkeys**. *Nat Neurosci* 2010, **13**:120-126.

49. Wise SP: **Forward frontal fields: phylogeny and fundamental function**. *Trends Neurosci* 2008, **31**:599-608.

50. Bernacchia A, Seo H, Lee D, Wang XJ: **A reservoir of time**
•• **constants for memory traces in cortical neurons**. *Nat Neurosci* 2011, **14**:366-372.

Reward feedback may be used to guide learning but the significance of any reward for updating future expectations changes depends on the current learning rate. If the learning rate is low then the reward's presence on a trial is considered in the light of many previous trials' outcomes but if the rate is high then the reward is considered in the light of just the most recent trials' outcomes. In this study neurons with a range of different time constants reflecting reward memory — over either a small or large number of recent trials were found in ACC and other areas.

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

52. Alexander WH, Brown JW: **Medial prefrontal cortex as an action-outcome predictor**. *Nat Neurosci* 2011, **14**: 1338-1344.