

OPINION

A distributed, hierarchical and recurrent framework for reward-based choice

Laurence T. Hunt and Benjamin Y. Hayden

Abstract | Many accounts of reward-based choice argue for distinct component processes that are serial and functionally localized. In this Opinion article, we argue for an alternative viewpoint, in which choices emerge from repeated computations that are distributed across many brain regions. We emphasize how several features of neuroanatomy may support the implementation of choice, including mutual inhibition in recurrent neural networks and the hierarchical organization of timescales for information processing across the cortex. This account also suggests that certain correlates of value are emergent rather than represented explicitly in the brain.

Neurobiologists have long been interested in developing mechanistic models to explain how we evaluate options and choose the best course of action^{1–3}. Many of these models take a modular perspective. That is, they assume, even if only tacitly, that goal-directed choice can be subdivided into a set of discrete component processes and that the neural implementation of these processes is both serial and localized^{4–7}. The component processes typically include the evaluation of options, the comparison of option values in the absence of any other factors, the selection of an appropriate action plan and the monitoring of the outcome of the choice. These component processes are generally assumed to correspond to discrete neural computations that are implemented in distinct neural structures.

An alternative perspective takes features of neural circuit anatomy as a starting point and constructs circuit-based models that predict both behavioural and neural data while retaining biological plausibility at their core^{8–14}. Recent research using such an approach emphasizes three overarching principles of reward-based choice. First, decisions may be formed in a distributed manner across many brain regions that act in concert and perform similar

computations. Second, the distributed networks implementing choices are highly recurrent in nature, which affects the kinds of computation that are performed. Third, these distributed and recurrent networks are organized into functional and temporal hierarchies.

Centring behavioural models on neural circuit plausibility builds on several traditions. Its origins can be traced back to mid-century pioneers like Weiner, Hull, Hebb, and McCullough and Pitts^{15,16}. These scholars took advantage of then-new discoveries about neurophysiology and computer science to propose neurally plausible computational theories. The next generation of scientists — including Grossberg, Hopfield, McClelland, Rumelhart and Hinton — developed these ideas to model entire neural systems. The key insight of this branch of research, from McCullough and Pitts to the present day, was that neuron-like units that performed biophysically plausible computations and were connected in simple ways could perform astonishingly rich computations^{17–20}. Such systems do not have dedicated memory and processing subsystems, unlike other computing architectures. Memory and computation are instead interwoven in the system and distributed broadly throughout

it¹⁷. This is directly relevant to choice models because evaluation depends crucially on memory and comparison is a basic computation for choice.

These traditions continue to influence neuroscience to this day. Their influence on decision making has been expressed by computational neuroscientists such as Wang, Frank, O'Reilly, Rolls and Cisek, among others^{8–13}. Recently, several studies have tested key empirical predictions from circuit-based accounts of decision making. Moreover, these advances have coincided with a resurgence of interest in neural networks in machine learning and computer science²¹, and with methodological breakthroughs in neuroanatomy²² and circuit manipulation²³. These methods should allow for far more rigorous testing and refinement of circuit-level models in the near future. It is therefore particularly timely to consider the contribution that such models make to our understanding of reward-guided choice.

In this Opinion article, we first outline the key principles behind a distributed, hierarchical and recurrent account of reward-guided choice. We then discuss some empirical motivations for this account. We argue that the known and emerging neuroscience of simple economic choice is consistent with several important properties of circuit-based models.

The distributed, hierarchical and recurrent approach differs qualitatively from modular explanations of reward-guided choice in that it is eliminative²⁴. This means that the neural implementation of choice does not necessarily recapitulate the steps or modules that are often used to describe the overall process. Instead, choice algorithms may be thought of as emergent properties of network activity. This change in perspective leads to a core set of research questions that are different from those driven by the modular approach (BOX 1).

Properties of the framework

Distributed

A distributed decision is one in which separate elements perform subsidiary computations that, when combined, produce the eventual choice. Well-known examples include the actions of individual voters in a national election and the selection of hive sites by swarms of bees²⁵ (BOX 2).

Box 1 | A change of perspective

Adopting a circuit-based perspective of reward-based choice reframes the questions that we ask about how decisions are formed and alters our interpretation of the resulting neural data. Fundamentally, it changes the question from “what is represented?” to “how is the computation implemented?”

To be more specific, the modular perspective has led us to ask questions such as: how is value computed and represented in the brain? In what regions do the evaluation, comparison and selection steps occur? In what ‘space’ (for example, goods-based or action-based) does value comparison occur? Do the computations performed in a particular brain region precede or follow the decision? What are the qualitative functional differences between different regions of the reward system?

By contrast, the distributed, hierarchical and recurrent perspective of reward-based choice makes a different set of assumptions, and these in turn lead to different research questions. First, it assumes that the component processes of choice may not be localized to particular computations in discrete brain areas. Instead, the components may be distributed across many regions simultaneously — implementing fundamentally similar, canonical computations. What is the nature of these computations? And how do different brain regions interact as choices are made?

The second assumption is that a comparison may not occur in a single decision space: decision spaces may reflect the anatomical connections of a given region or even be artefacts of the experimental design. How does the hierarchical organization of decision and reward areas lead to effective choices? And how does it explain the observed haemodynamic and neuronal response patterns?

Third, most brain regions are both pre- and post-decisional. Equally, because decision formation occurs gradually, most regions may be better classified as mid-decisional. How does neuroanatomy produce a gradual transformation from offers to choices? What role do the ubiquitous recurrent and feedback connections of the brain have in that process?

Fourth, value, or at least certain correlates of value, may not be represented per se. The implementation of choice does not need to recapitulate the algorithms that can be used to describe the overall choice process. Instead, the algorithm is an emergent property of the system: certain correlates of value could therefore emerge naturally as a consequence of how neural dynamics unfold across different trials. What is the structure of these dynamics? And why might they give rise to value correlates? Finally, for which correlates of value is this true?

choice, firing rates of neurons in both the vmPFC³⁴ and the ventral striatum (VS)³⁵ encode the values of the two offers through monotonic changes in firing rates. During the comparison period of the decision, the directions of the tuning curves (positive or negative) for the two offers are opposed^{34,35} (FIG. 1c). Consequently, the ensemble activity of both areas functions as a comparator between the values of the two offers. The fact that similar effects are seen in both areas, and with largely overlapping time courses, suggests that neither is the sole site of comparison and that comparison may take place in both regions simultaneously. Nor are the vmPFC and the VS likely to be unique in this respect; the mutual inhibition model has also been used to capture the dynamics of single-unit activity in the OFC³⁶ (FIG. 1d), the dorsolateral PFC³² and the lateral intraparietal cortex^{39,42}. Data from studies in the dorsal premotor cortex are also consistent with a mutual inhibition process. In this region, neural responses during economic decisions encode the relative value of targets in their response fields and show additional sensitivity to the physical distance between those response fields^{40,41}. This finding suggests that even ostensibly motor areas are part of the distributed choice network.

Distributed transformations across cortical circuits.

The body of work described above suggests that decisions occur through repeated mutual inhibition computations occurring simultaneously in both the motor and the abstract value domains. This implies that comparison is not the unique purview of any single brain area^{43–45}. The distributed account instead suggests that multiple areas perform a similar non-specialized function: they may all perform a comparison operation like mutual inhibition on the inputs received. Importantly, however, those inputs would differ by region. The nature of the competition occurring in any given area would then depend on the interaction of the particular demands of the task⁴⁶ and of the anatomical inputs of each area^{47,48}. For example, tuning properties of neurons in the OFC may be relatively specialized for stimulus-based or gustatory comparisons⁴⁹, whereas neurons in the cingulate cortex may seem to be specialized for motoric cost-based evaluations⁵⁰.

This idea of multiple distributed comparators may help to resolve differing interpretations of imaging studies in neuroeconomics, which have attempted to localize the regions that are most important for value comparison.

In each of these cases, individual elements process a small (and often noisy) fragment of the overall input, possibly including the outcomes of other agents, to make an overarching single decision in the aggregate.

A neuron can be thought of as a small and information-limited decision maker^{26,27}. It nonlinearly transforms its dendritic inputs into its firing rate. However, this transformation of inputs into outputs does not have to be completed by a single neuron (or layer of neurons) to be useful. Neurons providing incremental change may instead contribute as part of a well-organized distributed system that includes multiple brain regions. Any neuron that influences the behaviour of the network with its activity can be thought of as participating in that behaviour. In a decision-making network, this means that individual neurons may not require a pure representation of decision parameters to contribute to the decision-making process²⁸.

Distributed transformations within cortical circuits.

Distributed decisions often involve a single, simple computation repeated in each element on different inputs. One strong candidate for that computation in economic

choice is competition through mutual inhibition. Mutual inhibition is a common motif that is found throughout the nervous system and is often considered part of a basic repertoire of neural circuits^{29,30}. Effective competition via mutual inhibition can be mediated in a biophysically realistic cortical circuit model by an appropriate choice of synaptic weights^{11,13} (FIG. 1a). Several recent results indicate that mutual inhibition may be at the core of reward-guided choice^{31–42}.

A recent study using magneto-encephalography demonstrated that the human ventromedial prefrontal cortex (vmPFC) and intraparietal sulcus (IPS) express a key signature of mutual inhibition during economic choice: a change from encoding the sum of chosen and unchosen values to the difference between them³¹ (FIG. 1b). This signature was subsequently also found in a study of local field potentials from several subregions of the macaque PFC, including the orbitofrontal cortex (OFC), the dorsolateral PFC and the anterior cingulate cortex (ACC), confirming the generality of the mutual inhibition principle across species³².

Single-neuron recordings in macaque monkeys provide further evidence for a mutual inhibition process. In a binary

Box 2 | Distributed decision making in bee swarms

Although modular decision systems are often intuitive — the functions map directly onto the structure — distributed ones are not. The decision process of a bee swarm provides a natural and intuitive example of a distributed decision system.

In late spring, a hive of bees will enter into a swarm state and begin the process of choosing a new hive site^{25,158}. Individual scout bees make reconnaissance flights to identify and evaluate potential sites. The ideal sites are open, dry cavities of medium volume that are located high up in the canopy, protected from wind and facing south; thus, adaptive decisions need to optimize across many dimensions. On finding a potential site, each scout returns to the swarm and signals its location and quality through a specific pattern of dancing. Dances signalling high-quality sites can recruit other bees to investigate the same site. Subsequently, bees evaluate the popularity of a potential site by counting the number of visitors there. When scouts detect a quorum of other scouts at a hive site (around 20 bees), they transmit an activation signal to the swarm. Bee swarms even show a distinct mutual inhibition signal that reduces the chance of costly ambivalence¹⁵⁹.

Beehive decision making has several notable features that make it a good analogy for distributed decision making²⁵. First, there is no localized evaluation: no individual bee has more than an extremely limited amount of information about the world, and each bee's behaviour is remarkably stochastic. Second, there is no central decision maker: no individual or subgroup makes the decision; instead, it arises in a well-understood emergent manner from the simple rules followed by individuals. Thus, removing any bee or bees would degrade performance in a graded but not all-or-none manner. Third, there is very little stable functional specialization: scouts are drawn at random and serve as site selectors, as observers of other bees' dances, and as both members and monitors of the hive site quorum. Fourth, information about the value and location of options is firmly linked at every step of the process, thus sidestepping the otherwise difficult binding problem that is associated with choice and selection⁴⁵.

As a consequence, the steps of evaluation, comparison and selection are clearly performed by bees, but at the same time they are not performed by dedicated subsets of bees or at specific times. No individual bee ever has knowledge of more than one hive site, so no bee performs a comparison¹⁵⁸; instead, the comparison step emerges as a consequence of the types of interactions that the bees are programmed to perform.

In blood-oxygen-level-dependent (BOLD) functional MRI (fMRI) studies, debate has particularly centred on different regions of the medial frontal cortex. For example, one study used the mutual inhibition model to predict how variation in levels of vmPFC GABAergic inhibition, indexed via magnetic resonance spectroscopy in humans, related to cross-subject variation in both choice stochasticity and value correlates in the vmPFC BOLD fMRI signal³³ (FIG. 1e). A related study argued that, because activity within a mutual inhibition model is highest at the end of the choice process and this activity is persistent, then the fMRI signal will be greatest for faster (easy) decisions, as is typically the case in the vmPFC⁵¹. However, other studies have argued that one should only consider accumulated activity until a decision boundary is reached. In this case, the fMRI signal would be higher for slower (difficult) decisions, as is typically found in the dorsomedial PFC⁵².

A potential reconciliation of these results would be that both regions implement a mutual inhibition process but differ in their response properties post-choice after a decision bound has been reached. This explanation is supported by the event-related profile of local field potential recordings from multiple subregions of

the PFC during reward-based choice³². It is also supported by the simultaneous emergence of single-unit choice-related signals across six simultaneously recorded cortical regions in perceptual choice⁵³ and by the demonstration that motor output (corticospinal excitability) is already biased as a decision is unfolding⁵⁴.

The question then arises of how different areas interact as choices are made. Here, whole-brain techniques (such as fMRI) come into their element⁵⁵. Although anatomical connectivity is stable, functional connectivity is more flexible. One recent study examined changes in functional connectivity during multi-attribute choices involving integration of stimulus-based and action-based attributes³⁷. A model of choice in which competition via mutual inhibition occurred at multiple levels (stimuli, actions and attributes) best explained the subjects' choices. The BOLD fMRI signal in the IPS matched the model predictions from the competition over which attribute was most relevant. In particular, IPS value difference signals exhibited opposing signs for the relevant and irrelevant attributes. Notably, IPS functional connectivity was increased to either the OFC or the putamen when stimuli or actions were respectively most relevant for the decision at hand. A related study of multidimensional

learning demonstrated that the IPS is also particularly active when subjects update their understanding of the relevance of particular dimensions for guiding future choices⁵⁶.

Recurrent

The predominant paradigm in studies of simple economic choice has been *n*-alternative forced choice, in which two or more options are presented simultaneously to the subject, who selects the most valuable. However, real-world choice is often sequential rather than simultaneous in nature^{57,58}. Even ostensibly simultaneous choices may be made sequential by virtue of limits imposed by attention or noise that needs to be integrated out across time^{2,14,59}. The relationship between simultaneous and sequential accounts of choice parallels a distinction made between the computational roles of two distinct architectures of neural networks²¹: feedforward and recurrent architectures.

Two different classes of network architecture.

Feedforward neural networks — recently popularized through the impressive performance of convolutional neural networks (ConvNets) in computer vision^{60,61} — contain units that exhibit activity that is only dependent on the currently presented input. Such networks are ideally suited to tasks involving classification of multiple, simultaneously presented input features, such as pixel intensities in an image. The design of the 'local convolution' and 'max pooling' steps in a ConvNet was based on the response properties of primary visual cortex simple and complex cells, respectively⁶⁰. Modern feedforward ConvNet models strikingly reproduce single-unit responses along the visual hierarchy^{62,63}.

However, feedforward networks are poorly adapted to tasks that require some form of persistent memory for previous states across time. By contrast, recurrent neural networks (RNNs) contain units that not only receive inputs from other network layers but also receive their own previous output at time *t*-1 (REFS 18,64). This allows RNNs to show sustained memory for inputs long after they have been removed, allowing temporally extended computations to be performed on sequential inputs^{65,66}. One of the original successes of such networks was in showing how a biologically plausible network could account for working-memory responses in the PFC^{67,68}. The recurrence of these networks recapitulates the high degree of recurrent connectivity that is observed empirically in prefrontal circuits relative to other parts of the cortex^{69,70}.

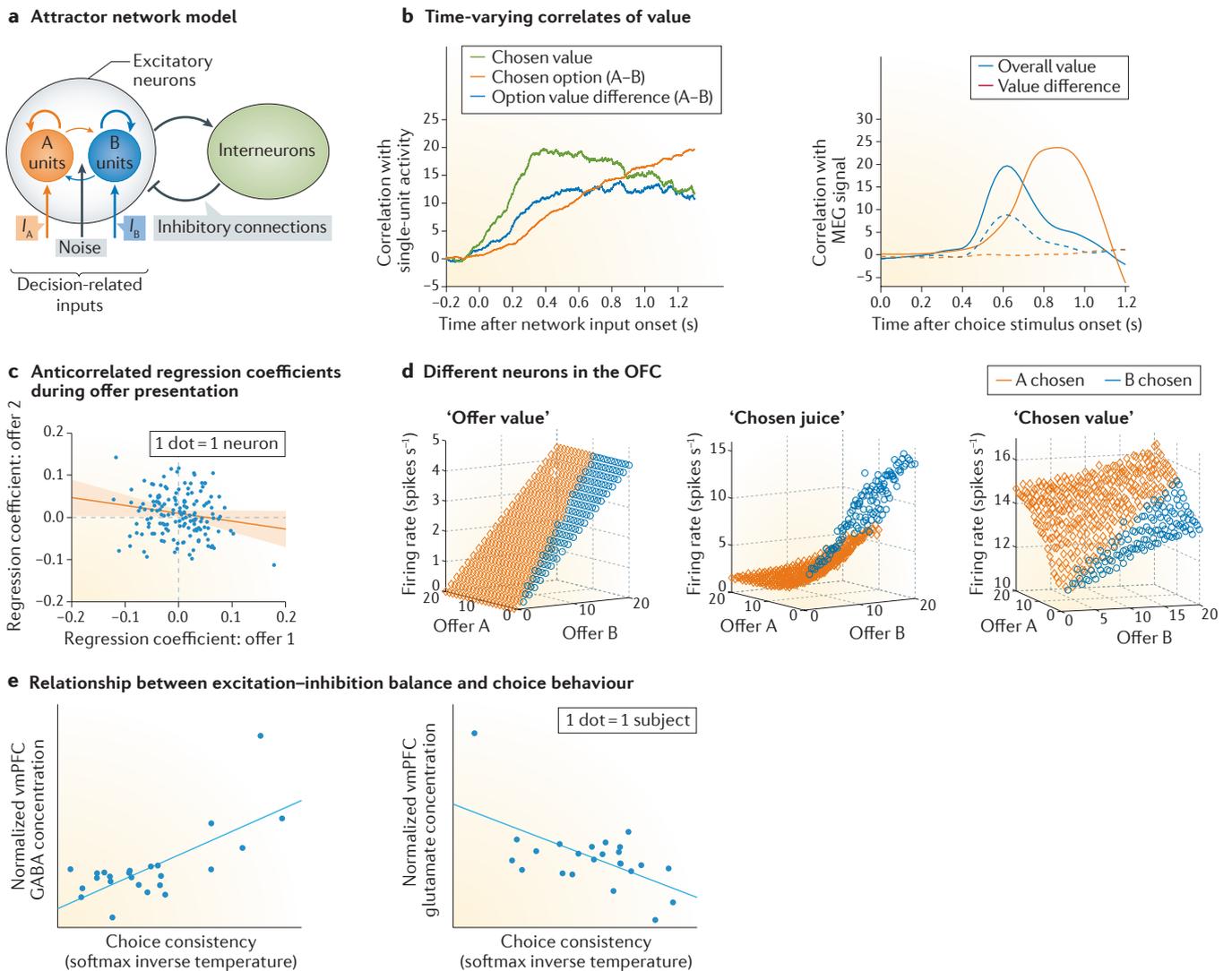


Figure 1 | Evidence for competition via mutual inhibition during reward-guided choice. **a** | A biophysical attractor network model of decision making¹¹ relies on effective competition via mutual inhibition between ‘A units’ and ‘B units’, which are neurons selective for a different option. Noise is added as an input to all excitatory units in the network, causing choice behaviour to vary stochastically on different simulations. This model has been used to explain various findings. **b** | This model predicts variation in correlates of value as a function of time^{31,32}. In single units recorded from the dorsolateral prefrontal cortex (PFC) of macaque monkeys³², the model predicts a sequence of value correlates from the original difference in options available to the eventual categorical choice (left panel). In human magnetoencephalography (MEG) data³¹, the model predicts a transition from ‘overall value’ correlates to ‘value difference’ correlates (right panel). Solid lines indicate ‘correct’ trials; dashed lines indicate ‘error’ trials (in which the network chooses a lower-valued option). **c** | This model predicts an anticorrelation of regression coefficients between different presented offers. This is observed in single-neuron data recorded from the ventromedial PFC

(vmPFC) and ventral striatum during sequential choice^{34,35} **d** | In the orbito-frontal cortex (OFC), different neurons are found that respond to the values of different juice offers, to the value of the chosen offer or to the identity of the chosen juice⁴⁹. It has been proposed that these neuronal classes correspond to different identities of neurons in the network model³⁶, the response properties of which are shown here. **e** | This model predicts variation in choice behaviour as the balance between excitation and inhibition is altered. This matches with evidence from human subjects, whose choice consistency is found to correlate with resting concentrations of glutamate and GABA in the vmPFC, indexed using magnetic resonance spectroscopy³³. Part **a** is adapted from REF. 11, Cell Press/Elsevier. Part **b** is adapted from REF. 31, Macmillan Publishers Limited, and from REF. 32. Part **c** is adapted from REF. 34, Cell Press/Elsevier. Part **d** is adapted from: A neuro-computational model of economic decisions. *Journal of Neurophysiology*, Rustichini, A. & Padoa-Schioppa, C., 113, 2015, 1382–1398. Part **e** is adapted from REF. 33, Macmillan Publishers Limited.

RNNs as a naturalistic substrate for sequential choices. At first glance, the simultaneous simple economic choices between two goods that are often studied in the laboratory seem to be well suited to being solved with a feedforward network.

The options in such choices are, after all, presented to the subject simultaneously and statically. However, reaction times in such tasks vary systematically as a function of value, and this implies that the underlying computation is dynamic rather than

static⁷¹. Recent models in neuroeconomics have therefore moved away from static economic accounts of choice towards temporally extended algorithms, in which noisy estimates of value are integrated sequentially over time^{14,52}. These algorithms

can be formally related to the model of competition via mutual inhibition, which is discussed above and in FIG. 1 (REF. 72). This network exemplifies how a biologically plausible RNN could implement an evidence accumulation algorithm¹¹.

But why should the brain use a recurrent network if a feedforward architecture seems to be sufficient? It may be because of the sequential structure of choices that are typically faced in natural environments, such as foraging decisions. These choices probably have shaped the evolution of the frontal cortex most heavily^{43,73}. Foraging theory emphasizes the fact that nearly all decisions must be made in a strategic manner, and doing so requires a telescoping representation of rewards and their future allocation distribution^{74,75}. For example, when macaque monkeys perform a foraging task, activity within the ACC accumulates slowly over time in a depleting resource environment until a fixed firing threshold is reached, at which point a change in behaviour is triggered⁷⁶.

The algorithmic structures of economic and foraging models are certainly very different, and one possibility would be that they are implemented by different brain systems, with the former being feedforward and the latter recurrent. However, we consider it more plausible that evolution will have developed a general approach to solving both types of decisions using temporally extended, recurrent implementations. The neural mechanisms of economic choice may even build on the underlying architecture laid down to solve foraging choices using recurrent computations. This fits well with modern ideas in neuroeconomic accounts that argue for temporal integration of noisy value estimates across time^{14,52,71}, with attention being sequentially allocated to different choice options^{59,77}.

Understanding recurrent network computations. To date, most neuroscientific studies using RNNs have used networks containing hand-tuned synaptic weights to elicit specific behaviours¹¹ (FIG. 1a). Although this is a valuable starting point, it produces an unrealistic homogeneity of neuronal responses that does not match the empirically observed heterogeneity in cortical populations^{78–80}. Recently introduced automatic training algorithms for RNNs have vastly increased their capacity to perform a wide range of tasks and also, potentially, to accurately describe

neural data^{81,82}. A major challenge is then to understand the nature of the computations that an RNN is performing after it has been automatically trained.

An elegant solution to this challenge is to perform ‘reverse engineering’ on RNNs that have been fit to data⁸³. This was exemplified in a recent study of PFC population responses, in which the PFC performed both the selection and the integration of relevant information in a context-dependent perceptual decision task⁸⁴. The authors trained a randomly connected RNN to capture key features of the behavioural data; activity in the trained RNN then matched with several key features of the neural population data. Reverse engineering of the RNN revealed the mechanisms whereby such a computation could be achieved within a single cortical circuit. Two stable line attractors were present within the trained RNN; the selection of relevant inputs depended on how the network activity relaxed towards these line attractors under different contexts. Similar analyses of population recordings should yield important insights into the recurrent computations supporting reward-based choice in the near future⁸⁵.

Hierarchical

A hierarchical organization is a cardinal feature of the organization of the reward and decision-making system of the brain. Information from one area is converted to a more abstract form that becomes more comprehensive as it increases in complexity. One feature of the hierarchy of brain areas that are associated with economic choice is that there are new inputs at each point. Thus, for example, the OFC receives gustatory inputs, the vmPFC receives limbic inputs, the subgenual ACC receives hypothalamic inputs, and the dorsal ACC receives motor inputs⁸⁶. These inputs provide a way for different factors to enter into the distributed network subserving choice. They also allow each area to have a different specific contribution to choice, even if the general role of incorporating information into ongoing decisions is similar. This viewpoint differs from one in which all factors must be integrated at a single locus to create a unitary value scale before they can influence choice.

Hierarchical RNNs allow for multiple timescales. Humans excel at tasks that demand online organization of behaviour across multiple differing timescales. Such tasks are particularly sensitive to PFC

damage⁸⁷. It has been hypothesized that the presence of parallel and hierarchical architectures within the PFC allows different pieces of information to be remembered and parsed at multiple timescales^{88,89}. This idea recapitulates recent developments in the design of RNNs in computer science. In particular, by making RNNs multilayered or ‘deep’ (REFS 21,90), it is possible to dramatically improve performance on tasks that operate over multiple timescales (such as the generation of meaningful sentences to describe images⁹¹ and the recognition of speech⁶⁵). The relevant timescales in speech are unknown to the neural network before training, but using automated training of multiple connected RNNs, the relevant temporal structure is extracted to successfully generate the requisite output. In decision making, the ability of a population of cortical neurons to exhibit multiple timescales of integration over a previous history of rewards has recently been shown empirically⁹². These timescales vary in a manner that is dependent on the timescale of integration of the animal’s current choice behaviour^{92,93}.

Hierarchical timescales for information processing in the cortex. A hierarchy of timescales across brain regions has recently been explicated in a large-scale network model of dynamical processing in the non-human primate neocortex. One study combined a mean-field reduction of a RNN within each cortical area with detailed knowledge of anatomical connectivity between different areas derived from tracer studies⁹⁴. This established a mechanism whereby multiple timescales could coexist within a single anatomical network, explaining the temporal structure of neurophysiological data recorded at rest across different cortical areas⁹⁵. This work suggests an explanation for why early sensory areas possess an inherently transient temporal structure (changing across tens of milliseconds), whereas higher regions (for example, the PFC) show a more sustained temporal profile (changing across hundreds of milliseconds or longer) (FIG. 2a).

Recent decision-making paradigms have combined evidence accumulation at both slow (inter-trial) and fast (intra-trial) timescales⁹⁶. It might then be predicted that evidence accumulation at slow timescales would be supported by cortical regions with sustained temporal structure. Interestingly, the dorsal ACC emerges as having the most sustained temporal structure within the regions that have been

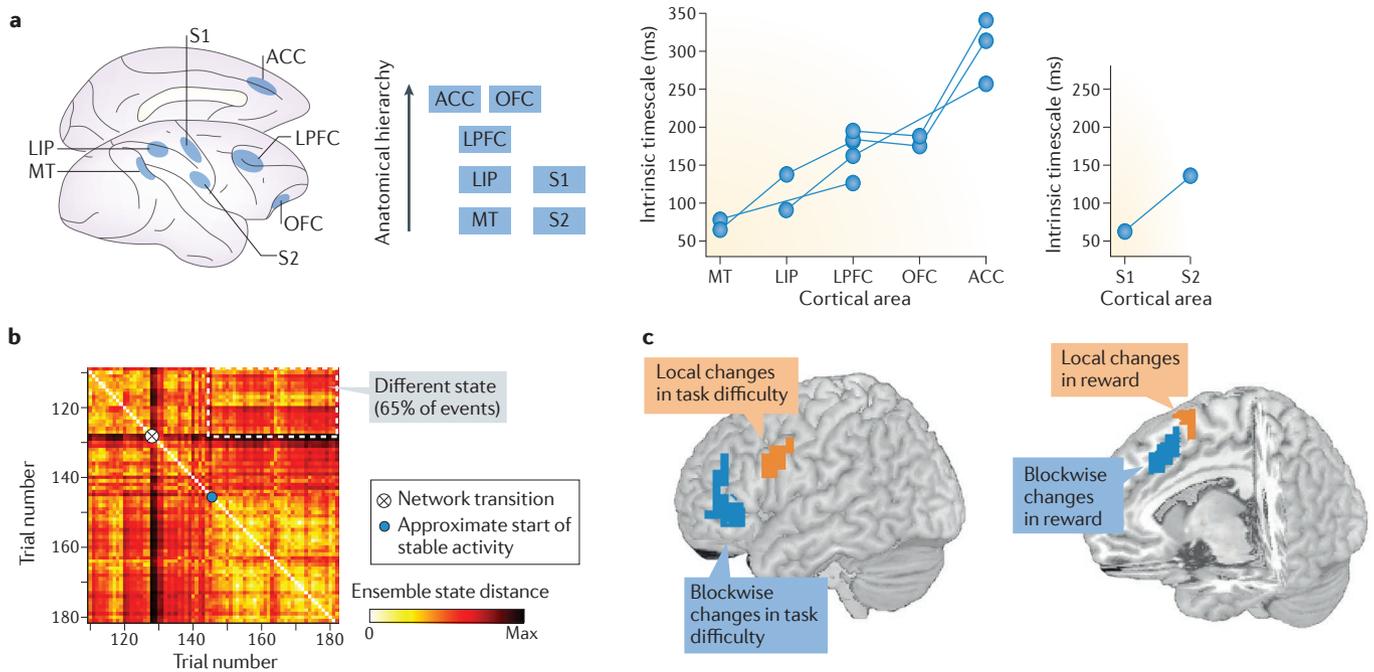


Figure 2 | Hierarchical organization of cortical timescales and its relationship with reward-guided choice. **a** | The ‘intrinsic timescale’ of different cortical regions can be indexed by examining the rate of decay of their spike rate autocorrelation while the subject is at rest⁹⁵. This reveals a hierarchical organization to cortical timescales, with lower areas having rapidly changing activity and higher areas having more persistent activity (data points from the same data set are joined). The medial prefrontal cortex (PFC), in particular the anterior cingulate cortex (ACC), has the longest intrinsic timescale. **b** | This part shows persistent network activity, and network ‘resets’ during behavioural shifts, in the rat medial PFC in a reward-guided exploration task¹⁰⁰. Each element of the matrix reflects the trial-to-trial similarity of neuronal ensemble activity. The lighter and darker elements reflect more similar and more dissimilar network activity, respectively. Large, bright blocks reflect persistent, stable states across trials. Abrupt network transitions, and

relaxations into different stable states, can be readily detected within network activity and are related to behavioural shifts. **c** | This part depicts the rostrocaudal hierarchy of temporal organization in the PFC¹⁰⁴. A similar hierarchy can be seen on medial and lateral surfaces for internal motivation (reward) and task difficulty, respectively. Whereas more-anterior regions reflect blockwise changes in reward or in task difficulty, more-posterior regions reflect trial-to-trial changes. LIP, lateral intraparietal area; LPFC, lateral PFC; MT, middle temporal area; OFC, orbitofrontal cortex; S1, primary somatosensory cortex; S2, secondary somatosensory cortex. Part **a** is adapted from REF. 95, Macmillan Publishers Limited. Part **b** is reprinted with permission from Karlsson, M. P., Tervo, D. G. R. & Karpova, A. Y., Network resets in medial prefrontal cortex mark the onset of behavioral uncertainty, *Science*, 338, 135–139 (2012), American Association for the Advancement of Science (AAAS). Part **c** is adapted from REF. 104, Macmillan Publishers Limited.

characterized thus far⁹⁵. This observation sits well with studies that have linked ACC activity to disengagement from a foreground option in foraging tasks^{75,76,97,98} and during exploration⁹⁹. A foreground option may be held in mind over many trials, but disengagements are linked to abrupt and coordinated changes in network activity¹⁰⁰ (FIG. 2b). Before disengagement, ramping activity is observed in both the dorsal ACC^{76,101,102} and another region that is associated with exploration⁹⁹, the rostralateral PFC¹⁰².

Relationship to cognitive hierarchies.

The observed relationship between anatomical and temporal hierarchies may be related to hierarchical accounts of PFC function in other cognitive paradigms¹⁰³. One such account suggests the existence of two parallel streams of rostrocaudal organization within the PFC, along the medial and lateral surfaces of this region,

respectively¹⁰⁴ (FIG. 2c). Data from BOLD fMRI studies suggest that a hierarchy relating to states of internal motivation exists along the medial surface^{104–106}, with more-anterior portions reflecting blockwise changes in reward value and more-posterior regions reflecting trial-to-trial changes. By contrast, a hierarchy of cognitive control processes exists along the lateral surface^{104,107,108}. Again, anterior regions of the lateral surface exhibit sustained changes reflecting the complexity of the current block, whereas more-posterior regions of this surface reflect task complexity only for relevant individual trials. Similar rostrocaudal PFC hierarchies have also been found to transfer to the striatum during the course of hierarchical rule learning¹⁰⁹. This observation can be described by a hierarchically organized network model of corticostriatal interactions¹¹⁰, inspired by a particularly successful form of RNN known as ‘long short-term memory’ (REFS 111,112).

Evidence for the framework

Lesions and component processes

An important property of distributed computing systems is graceful degradation¹⁷. Because information is stored broadly, small amounts of damage to the system are seldom catastrophic; major impairments only come with large amounts of damage. Damage selectively impairs difficult retrieval processes and spares easier ones. The analogue of damage in connectionist networks is brain lesions. Brain lesions have long provided an important source of evidence for functional specialization. In the visual system, for example, lesions to the middle temporal area cause akinetopsia¹¹³, and lesions to the fusiform face area cause prosopagnosia¹¹⁴, confirming functional specialization for motion perception and face processing, respectively.

It has been natural for experimenters to design lesion studies that allow conceptually different components of

choice to be assigned to different brain regions, and results of these studies have advocated a degree of functional specialization within the PFC^{115–117}. However, given how profound the impact of PFC lesions on day-to-day choices can be¹¹⁸, the impairments on choice tasks often appear to be surprisingly mild. PFC lesions sometimes lead to graded deficits in choosing, with difficult decisions impaired but easier ones spared¹¹⁵. It is possible that these deficits are mild because the tasks seek to differentiate component processes of economic choice rather than addressing the temporally extended nature of real-world economic decisions⁸⁷.

The range of cognitive functions that are affected by PFC lesions also casts doubt on the idea that choice, or even value comparison, is a specialization of any single brain region. This doubt is supported by functional imaging data from studies of simple economic choice: different types of comparison signal are observed in different brain areas, depending on the task at hand^{32,46,52,75}. Moreover, regions that are implicated in choice in lesion studies and fMRI studies are not specific to choice: they are also associated with cognitive processes that include working memory, strategic planning, executive control, reasoning and social cognition¹¹⁹. Together, this evidence suggests that many brain regions will collectively contribute to the process of comparison and that brain regions subserving choice make their contribution as part of a larger supporting suite of cognitive functions. By way of analogy, reward-guided choice seems to operate less like motion- or face-processing and more like the ability to drive a car, form a political preference or do calculus: such abilities rely on coordinated computations across many brain structures and systems.

Ubiquitous value correlates

Another feature of connectionist networks is that their storage of memories depends on distributed synaptic weight changes across all parts of a network rather than at a single site¹⁷. Value is closely related to memory: it is a feature of an option that is inferred from associations with reward, which are based on past experiences¹²⁰. In connectionist systems, memories are widely distributed (a feature leading to graceful degradation, as discussed above), and traces of those memories can be observed throughout the network¹⁷; one might then predict that the same could also be true of value.

It is indeed the case that value correlates can be found in multiple brain regions. Correlates of value are seen in core reward regions such as the OFC⁴⁹, the vmPFC¹²¹ and the VS¹²², but they are also observed in the amygdala, the insula, the dorsal striatum, the midbrain, the pregenual, subgenual, dorsal anterior and posterior cingulate cortices, the dorsolateral and ventrolateral PFC, the IPS, and even in the sensory and motor cortices^{40,123–130}. One recent neuroimaging study found that more than 30% of the brain exhibited such signals¹³¹. Although there are certainly differences between the types of value-related information that are represented in these regions, there are also many overlaps.

Despite the ubiquity of value correlates, it often remains unclear whether activity in any area truly represents value — or what precise definition of value we should use^{28,132,133}. One common criticism of neuroeconomic studies is that putative value correlates often reflect other correlated variables, such as attention, salience or, even, subthreshold premotor activation. There have certainly been some important efforts to disambiguate value correlates from alternative explanations^{28,133,134}. Of course, it is possible that one of these correlates genuinely represents value in the formal sense. However, in distributed systems, representation is often an emergent property; that is, it is driven by the specific pattern of connection between units and is not a property of any particular unit. In other words, value may be a network routing principle rather than a quantity that is represented. This eliminative view of value would be consistent with a subset of both classic studies and recent work in behavioural economics on process models of economic choice^{58,135,136}.

One particular value signal is probably an emergent property. Recent work indicates that the ‘chosen-value’ variable is a by-product of variation in decision dynamics across trials. To understand why, one should consider the model of competition via mutual inhibition that was discussed above¹¹ (FIG. 1). Although the activities of many units in this network model correlate with chosen value, this quantity does need not to be represented to form a decision. Instead, it may arise naturally as a consequence of the varying speed at which network dynamics unfold on different trials³¹ (BOX 3).

The above-described findings are in line with those in recent studies of motor control, which indicate that motor cortical

activity is better understood as a dynamical system rather than as one that represents movement parameters^{137,138}. This line of work suggests that chosen-value correlates emerge as a necessary consequence of recurrent network dynamics in mediating competition¹³⁹ and, again, that chosen value may not be represented per se. It remains possible that other ostensible value correlates (such as offer value and experienced value) are also by-products of the computations that underlie choice and are not reified in the activity of dedicated reward neurons or regions¹³⁵.

Elusive pure value

A modular view of reward-based choice predicts that certain brain regions or populations of neurons should be specialized for value, meaning that they respond primarily to values of options. Because value computation is a key intermediate stage in economic choice, the existence of specialized value regions or neurons is an important prediction of modular theories. By contrast, distributed theories do not demand any specialized value computation. Instead, in these theories, value is distributed broadly across a large number of regions and is predicted in neurons that have other roles unrelated to valuation.

Some meta-analyses of neuroimaging data have argued that certain brain areas are central to ‘pure valuation’ (REF. 140). However, even in putative core reward regions such as the vmPFC, OFC and VS, a wealth of information processing occurs that is not related to value. For instance, the vmPFC is engaged by several ostensibly value-neutral factors, including autobiographical memory¹⁴¹, spatial navigation¹⁴², imagination¹⁴³ and social cognition¹⁴⁴; the OFC is engaged by non-reward processes like conflict, working memory and rule encoding^{145,146}. Likewise, factors that modulate individual neurons in these areas may include ‘valueless’ changes in outcome expectancy^{147,148}, previous outcomes⁹², intention to switch, as well as other strategy variables¹⁴⁹, metacognition¹⁵⁰, spatial positions of offers and choices⁴⁵, rules and task set¹⁴⁶, and even irrelevant task variables¹⁵¹. It is possible that these apparently value-neutral signals are observed in these tasks because the tasks nevertheless demand the computation of value. However, it is becoming more clear that neurons in most, if not all, value-relevant regions encode a large number of task-relevant variables simultaneously, a property known as mixed selectivity^{152,153}.

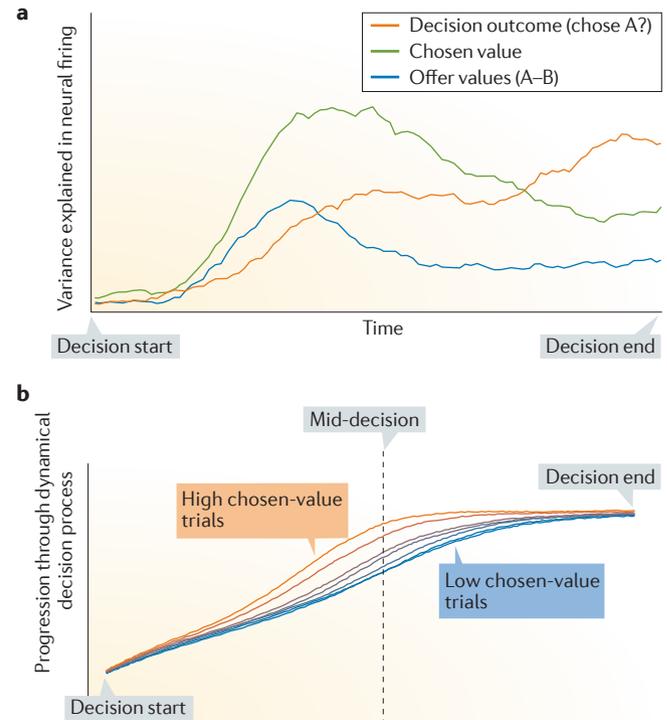
Box 3 | Value representations as epiphenomena of decision dynamics

One particularly ubiquitous signal during reward-guided decision tasks is a representation of chosen value^{34,35,39,49,129}. This representation is isolated by correlating some measure of neural activity (firing rates, local field potentials (LFPs) or functional MRI signals) with the value of the option that will eventually be chosen on that trial. In time-varying measurements, this correlation can be repeated across many different time points, isolating the time point at which maximal variance is explained (see the figure, part a; green line).

Why might this signal occur so commonly? Is it expressly represented by the brain, or is it an artefact? A clue comes from examining brain areas that also carry representations of other decision variables. These include the offer values of the options available and the eventual categorical choice that the subject will make. Several studies^{32,39,129} show that the maximal variance explained by chosen value occurs between an initial representation of offer values and the final representation of choices (see the figure, part a). Chosen-value representations emerge as the decision is being formed rather than after the choice is completed.

Decision formation is, of course, a dynamical process occurring at different rates on different trials⁷¹. Crucially, chosen value influences decision speed. At a fixed time point in the 'middle' of the decision process, the decision may have neared completion on some trials (typically with high chosen value), whereas on others it may be a long way from completion. Activity in any part of the brain that reflects the progression of this dynamical process will thus correlate with the chosen value at the mid-decision point (see the figure, part b). Notably, these dynamics have several possible neural substrates, including ramp-to-threshold accumulation in single neurons², neural population trajectories through a low-dimensional manifold¹³⁷, or bulk activity observed at the level of LFP or magnetoencephalography (MEG) signals³². In each case, correlates of chosen value would naturally emerge as a consequence of varying dynamics across trials. This idea has, for example, been used to explain the origin of a commonly observed 'unchosen-minus-chosen-value' signal in the dorsomedial prefrontal cortex using functional MRI⁵². In addition, a recent study estimated the speed at which dynamics unfolded on

a trial-by-trial basis using LFP and MEG data³². The authors found that this single-trial estimate of decision dynamics explained some of the variance that was previously explained by chosen value — implying that the underlying cause of chosen-value correlates was indeed linked to decision speed. Part a is adapted from REF. 32. Part b is adapted from REF. 31, Macmillan Publishers Limited.



Neuroanatomy

The strongest evidence for a distributed, hierarchical and recurrent approach to choice comes from neuroanatomy. Of course, the original work that provides the foundation for this approach developed side by side with progress in understanding the anatomy (and physiology) of the nervous system¹⁵. Perceptrons, feedforward networks, parallel distributed networks, Hebbian learning and ConvNets were all inspired by observations about brain structure and physiology.

In the cortex, anatomical studies have indicated a reciprocal feedforward and feedback structure¹⁵⁴. This architecture was quantified via detailed tracer studies in macaque monkeys¹⁵⁵ and diffusion imaging studies in humans²². Connectivity differences between adjacent prefrontal regions are smaller than is often appreciated, and adjacent regions generally blur into each other gradually rather than showing categorical boundaries¹⁵⁶. Likewise, when considering the organization of subcortical brain regions, it is now accepted that cortico-thalamo-basal

ganglia loops are not segregated as once thought¹⁵⁷ but instead show strong functional convergence⁸⁶.

Local intraregional connections in the cortex also suggest a distributed and recurrent organization. This is particularly true of 'higher' cortical areas that subservise cognitive functions. In particular, dendritic arbors of PFC pyramidal cells are endowed with many more dendritic spines than pyramidal cells in sensory areas, meaning that a single neuron in prefrontal areas 10, 11 or 12 receives 16 times the number of excitatory inputs of a neuron in the primary visual cortex⁷⁰. Most of these cortical connections are local rather than long-distance, allowing these circuits to have a highly recurrent organization that is similar to that observed in RNNs.

Discussion

In this Opinion article, we have presented an overview of recent work suggesting that reward-based decisions reflect the outcome of a distributed, hierarchical and recurrent computational process. These ideas have

their genesis in connectionist and neural network models that have also been used to understand perception and memory, among other processes. More recently, these ideas have been integrated to form detailed models of economic or reward-based decisions⁸⁻¹⁴.

The work that we described here brings up a philosophical issue that has long influenced cognitive science: mental representation¹⁶. It has long been unclear whether we 'represent' mental concepts. That is, whether specific patterns of brain activity serve to recapitulate a mental version of some external event or object. Although neural correlates of important events and objects are observed, these correlations may be consequences of internal processes. Recent work provides two reasons to doubt that value, at least, is explicitly represented in the brain. First, with regard to chosen value, it seems that, although value is decodable, it is possible that it is an artefact of the way neural data are analysed³². Second, with regard to value more broadly, the case is less clear, but connectionist models suggest that it is possible to construct networks

that make good choices without explicit value representations; these networks at the very least have a similar flavour to neuroanatomy¹⁷.

Paul and Patricia Churchland have articulated the notion of eliminative materialism, which includes a suggestion that the natural categories that we use to describe psychological phenomena do not do a good job of capturing the organization of brain processes that generate our mental lives²⁴. Economic choice may be one such case in point. Choice as a whole and steps like evaluation, comparison, selection and monitoring stages do not necessarily correspond to discrete anatomical substrates, to discrete neuron types or even to discrete computations. Instead, they may be emergent consequences of processing units performing simple operations on inputs, ones that are radically different from the operations of the system as a whole²⁴.

Laurence T. Hunt is at the Wellcome Trust Centre for Neuroimaging, University College London, 12 Queen Square, London WC1N 3BG, UK.

Benjamin Y. Hayden is at the Department of Brain and Cognitive Sciences, University of Rochester, 309 Meliora Hall, Rochester, New York 14618, USA.

laurence.hunt@ucl.ac.uk;
benhayden@gmail.com

doi:10.1038/nrn.2017.7

Published online 17 Feb 2017

- Glimcher, P. & Fehr, E. *Neuroeconomics: Decision Making and the Brain* (Academic Press, 2014).
- Gold, J. I. & Shadlen, M. N. The neural basis of decision making. *Annu. Rev. Neurosci.* **30**, 535–574 (2007).
- Dolan, R. J. & Dayan, P. Goals and habits in the brain. *Neuron* **80**, 312–325 (2013).
- Rangel, A., Camerer, C. & Montague, P. R. A framework for studying the neurobiology of value-based decision making. *Nat. Rev. Neurosci.* **9**, 545–556 (2008).
- Kable, J. W. & Glimcher, P. W. The neurobiology of decision: consensus and controversy. *Neuron* **63**, 733–745 (2009).
- Rangel, A. & Hare, T. Neural computations associated with goal-directed choice. *Curr. Opin. Neurobiol.* **20**, 262–270 (2010).
- Padoa-Schioppa, C. Neurobiology of economic choice: a good-based model. *Annu. Rev. Neurosci.* **34**, 333–359 (2011).
- Frank, M. J. in *An Introduction to Model-Based Cognitive Neuroscience* (eds Forstmann, B. & Wagenmakers, E.-J.) 159–177 (Springer, 2015).
- Wang, X. J. Neural dynamics and circuit mechanisms of decision-making. *Curr. Opin. Neurobiol.* **22**, 1039–1046 (2012).
- Pezzulo, G. & Cisek, P. Navigating the affordance landscape: feedback control as a process model of behavior and cognition. *Trends Cogn. Sci.* **20**, 414–424 (2016).
- Wang, X. J. Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* **36**, 955–968 (2002).
- Rolls, E. T. *Emotion and Decision Making Explained* (Oxford Univ. Press, 2013).
- Machens, C. K., Romo, R. & Brody, C. D. Flexible control of mutual inhibition: a neural model of two-interval discrimination. *Science* **307**, 1121–1124 (2005).
- Krajbich, I., Armel, C. & Rangel, A. Visual fixations and the computation and comparison of value in simple choice. *Nat. Neurosci.* **13**, 1292–1298 (2010).
- Medler, D. A. A brief history of connectionism. *Neural Comput. Surveys* **1**, 18–72 (1998).
- Gardner, H. *The Mind's New Science: A History of the Cognitive Revolution* (Basic Books, 1985).
- McClelland, J. L. & Rumelhart, D. E. *Explorations in Parallel Distributed Processing: A Handbook of Models, Programs, and Exercises* (MIT Press, 1989).
- Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. *Nature* **323**, 533–536 (1986).
- Hopfield, J. J. Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl Acad. Sci. USA* **79**, 2554–2558 (1982).
- Carpenter, G. A. & Grossberg, S. A massively parallel architecture for a self-organizing neural pattern recognition machine. *Comput. Vision Graph.* **37**, 54–115 (1987).
- LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
- Jbabdi, S., Sotiropoulos, S. N., Haber, S. N., Van Essen, D. C. & Behrens, T. E. Measuring macroscopic brain connections in vivo. *Nat. Neurosci.* **18**, 1546–1555 (2015).
- Fenko, L., Yizhar, O. & Deisseroth, K. The development and application of optogenetics. *Annu. Rev. Neurosci.* **34**, 389–412 (2011).
- Churchland, P. M. Eliminative materialism and the propositional attitudes. *J. Philos.* **78**, 67 (1981).
- Seeley, T. D. *Honeybee Democracy* (Princeton Univ. Press, 2010).
- London, M. & Häusser, M. Dendritic computation. *Annu. Rev. Neurosci.* **28**, 503–532 (2005).
- Couzin, I. D. Collective cognition in animal groups. *Trends Cogn. Sci.* **13**, 36–43 (2009).
- O'Doherty, J. P. The problem with value. *Neurosci. Biobehav. Rev.* **43**, 259–268 (2014).
- Grillner, S. Neurobiological bases of rhythmic motor acts in vertebrates. *Science* **228**, 143–149 (1985).
- Kovacs, M. & Davis, W. Behavioral choice: neural mechanisms in Pleurobranchaea. *Science* **198**, 632–634 (1977).
- Hunt, L. T. *et al.* Mechanisms underlying cortical activity during value-guided choice. *Nat. Neurosci.* **15**, 470–476 (2012).
- Hunt, L. T., Behrens, T. E., Hosokawa, T., Wallis, J. D. & Kennerley, S. W. Capturing the temporal evolution of choice across prefrontal cortex. *eLife* **4**, e11945 (2015).
- Jocham, G., Hunt, L. T., Near, J. & Behrens, T. E. A mechanism for value-guided choice based on the excitation-inhibition balance in prefrontal cortex. *Nat. Neurosci.* **15**, 960–961 (2012).
- Strait, C. E., Blanchard, T. C. & Hayden, B. Y. Reward value comparison via mutual inhibition in ventromedial prefrontal cortex. *Neuron* **82**, 1357–1366 (2014).
- Strait, C. E., Sleezer, B. J. & Hayden, B. Y. Signatures of value comparison in ventral striatum neurons. *PLoS Biol.* **13**, e1002173 (2015).
- Rustichini, A. & Padoa-Schioppa, C. A neuro-computational model of economic decisions. *J. Neurophysiol.* **114**, 1382–1398 (2015).
- Hunt, L. T., Dolan, R. J. & Behrens, T. E. Hierarchical competitions subserving multi-attribute choice. *Nat. Neurosci.* **17**, 1613–1622 (2014).
- Chau, B. K., Kolling, N., Hunt, L. T., Walton, M. E. & Rushworth, M. F. A neural mechanism underlying failure of optimal choice with multiple alternatives. *Nat. Neurosci.* **17**, 463–470 (2014).
- Louie, K. & Glimcher, P. W. Separating value from choice: delay discounting activity in the lateral intraparietal area. *J. Neurosci.* **30**, 5498–5507 (2010).
- Pastor-Bernier, A. & Cisek, P. Neural correlates of biased competition in premotor cortex. *J. Neurosci.* **31**, 7083–7088 (2011).
- Cisek, P. Integrated neural processes for defining potential actions and deciding between them: a computational model. *J. Neurosci.* **26**, 9761–9770 (2006).
- Soltani, A. A biophysically based neural model of matching law behavior: melioration by stochastic synapses. *J. Neurosci.* **26**, 3731–3744 (2006).
- Rushworth, M. F., Kolling, N., Sallet, J. & Mars, R. B. Valuation and decision-making in frontal cortex: one or many serial or parallel systems? *Curr. Opin. Neurobiol.* **22**, 946–955 (2012).
- Cisek, P. Making decisions through a distributed consensus. *Curr. Opin. Neurobiol.* **22**, 927–936 (2012).
- Strait, C. E. *et al.* Neuronal selectivity for spatial positions of offers and choices in five reward regions. *J. Neurophysiol.* **115**, 1098–1111 (2016).
- Hunt, L. T., Woolrich, M. W., Rushworth, M. F. & Behrens, T. E. Trial-type dependent frames of reference for value comparison. *PLoS Comput. Biol.* **9**, e1003225 (2013).
- Passingham, R. E., Stephan, K. E. & Kötter, R. The anatomical basis of functional localization in the cortex. *Nat. Rev. Neurosci.* **3**, 606–616 (2002).
- Neubert, F. X., Mars, R. B., Sallet, J. & Rushworth, M. F. 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 (2015).
- Padoa-Schioppa, C. & Assad, J. A. Neurons in the orbitofrontal cortex encode economic value. *Nature* **441**, 223–226 (2006).
- Hosokawa, T., Kennerley, S. W., Sloan, J. & Wallis, J. D. Single-neuron mechanisms underlying cost-benefit analysis in frontal cortex. *J. Neurosci.* **33**, 17385–17397 (2013).
- Rolls, E. T., Grabenhorst, F. & Deco, G. Choice, difficulty, and confidence in the brain. *Neuroimage* **53**, 694–706 (2010).
- Hare, T. A., Schultz, W., Camerer, C. F., O'Doherty, J. P. & Rangel, A. Transformation of stimulus value signals into motor commands during simple choice. *Proc. Natl Acad. Sci. USA* **108**, 18120–18125 (2011).
- Siegel, M., Buschman, T. J. & Miller, E. K. Cortical information flow during flexible sensorimotor decisions. *Science* **348**, 1352–1355 (2015).
- Klein-Flügge, M. C. & Bestmann, S. Time-dependent changes in human corticospinal excitability reveal value-based competition for action during decision processing. *J. Neurosci.* **32**, 8373–8382 (2012).
- O'Reilly, J. X., Woolrich, M. W., Behrens, T. E. J., Smith, S. M. & Johansen-Berg, H. Tools of the trade: psychophysiological interactions and functional connectivity. *Soc. Cogn. Affect. Neurosci.* **7**, 604–609 (2012).
- Niv, Y. *et al.* Reinforcement learning in multidimensional environments relies on attention mechanisms. *J. Neurosci.* **35**, 8145–8157 (2015).
- Freidin, E. & Kacelnik, A. Rational choice, context dependence, and the value of information in European starlings (*Sturnus vulgaris*). *Science* **334**, 1000–1002 (2011).
- Bettman, J. R., Luce, M. F. & Payne, J. W. Constructive consumer choice processes. *J. Consum. Res.* **25**, 187–217 (1998).
- Rich, E. L. & Wallis, J. D. Decoding subjective decisions from orbitofrontal cortex. *Nat. Neurosci.* **19**, 973–980 (2016).
- LeCun, Y. *et al.* Backpropagation applied to handwritten ZIP code recognition. *Neural Comput.* **1**, 541–551 (1989).
- Krizhevsky, A., Sutskever, I. & Hinton, G. ImageNet classification with deep convolutional neural networks. *Adv. Neural Inform. Process. Systems* **25**, 1090–1098 (2012).
- Yamins, D. L. *et al.* Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proc. Natl Acad. Sci. USA* **111**, 8619–8624 (2014).
- Kriegeskorte, N. Deep neural networks: a new framework for modeling biological vision and brain information processing. *Annu. Rev. Vision Sci.* **1**, 417–446 (2015).
- Graves, A. *Supervised Sequence Labelling with Recurrent Neural Networks* (Springer, 2012).
- Graves, A., Mohamed, A.-r. & Hinton, G. Speech recognition with deep recurrent neural networks. eprint at [arXiv http://arXiv.org/abs/1303.5778](http://arXiv.org/abs/1303.5778) (2013).
- Weston, J., Chopra, S. & Borde, A. Memory networks. eprint at [arXiv http://arXiv.org/abs/1410.3916](http://arXiv.org/abs/1410.3916) (2014).
- Amari, S.-i. Dynamics of pattern formation in lateral-inhibition type neural fields. *Biol. Cybern.* **27**, 77–87 (1977).
- Compte, A. Synaptic mechanisms and network dynamics underlying spatial working memory in a cortical network model. *Cereb. Cortex* **10**, 910–923 (2000).
- Wang, X.-J. in *Principles of Frontal Lobe Function* (eds Stuss, D. T. & Knight, R. T.) 226–248 (Oxford Univ. Press, 2013).
- Elston, G. N. Pyramidal cells of the frontal lobe: all the more spinous to think with. *J. Neurosci.* **20**, RC95 (2000).

71. Busemeyer, J. R. & Townsend, J. T. Decision field theory: a dynamic-cognitive approach to decision making in an uncertain environment. *Psychol. Rev.* **100**, 432–459 (1993).
72. Bogacz, R., Brown, E., Moehlis, J., Holmes, P. & Cohen, J. D. The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol. Rev.* **113**, 700–765 (2006).
73. Passingham, R. E. & Wise, S. P. *The Neurobiology of the Prefrontal Cortex: Anatomy, Evolution, and the Origin of Insight* 1st edn (Oxford Univ. Press, 2012).
74. Stephens, D. W. & Krebs, J. R. *Foraging Theory* (Princeton Univ. Press, 1986).
75. Kolling, N., Behrens, T. E., Mars, R. B. & Rushworth, M. F. Neural mechanisms of foraging. *Science* **336**, 95–98 (2012).
76. Hayden, B. Y., Pearson, J. M. & Platt, M. L. Neuronal basis of sequential foraging decisions in a patchy environment. *Nat. Neurosci.* **14**, 933–939 (2011).
77. Lim, S. L., O'Doherty, J. P. & Rangel, A. 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 (2011).
78. Harvey, C. D., Coen, P. & Tank, D. W. Choice-specific sequences in parietal cortex during a virtual-navigation decision task. *Nature* **484**, 62–68 (2012).
79. Stokes, M. G. 'Activity-silent' working memory in prefrontal cortex: a dynamic coding framework. *Trends Cogn. Sci.* **19**, 394–405 (2015).
80. Cavanagh, S. E., Wallis, J. D., Kennerley, S. W. & Hunt, L. T. Autocorrelation structure at rest predicts value correlates of single neurons during reward-guided choice. *eLife* **5**, e18937 (2016).
81. Sussillo, D. & Abbott, L. F. Generating coherent patterns of activity from chaotic neural networks. *Neuron* **63**, 544–557 (2009).
82. Hennequin, G., Vogels, T. P. & Gerstner, W. Optimal control of transient dynamics in balanced networks supports generation of complex movements. *Neuron* **82**, 1394–1406 (2014).
83. Sussillo, D. & Barak, O. Opening the black box: low-dimensional dynamics in high-dimensional recurrent neural networks. *Neural Comput.* **25**, 626–649 (2013).
84. Mante, V., Sussillo, D., Shenoy, K. V. & Newsome, W. T. Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* **503**, 78–84 (2013).
85. Song, H. F., Yang, G. R. & Wang, X. Reward-based training of recurrent neural networks for cognitive and value-based tasks. *eLife* **6**, e21492 (2017).
86. Haber, S. N. & Behrens, T. E. The neural network underlying incentive-based learning: implications for interpreting circuit disruptions in psychiatric disorders. *Neuron* **83**, 1019–1039 (2014).
87. Shallice, T. & Burgess, P. W. Deficits in strategy application following frontal lobe damage in man. *Brain* **114**, 727–741 (1991).
88. Fuster, J. M. The prefrontal cortex — an update: time is of the essence. *Neuron* **30**, 319–353 (2001).
89. Holroyd, C. B. & Yeung, N. Motivation of extended behaviors by anterior cingulate cortex. *Trends Cogn. Sci.* **16**, 122–128 (2012).
90. Pascanu, R., Gulcehre, C., Cho, K. & Bengio, Y. How to construct deep recurrent neural networks. eprint at arXiv <http://adsabs.harvard.edu/abs/2013arXiv1312.6026P> (2013).
91. Karpathy, A. & Fei-Fei, L. Deep visual-semantic alignments for generating image descriptions. *IEEE Trans. Pattern Anal. Mach. Intell.* <http://dx.doi.org/10.1109/TPAMI.2016.2598339> (2016).
92. Bernacchia, A., Seo, H., Lee, D. & Wang, X.-J. A reservoir of time constants for memory traces in cortical neurons. *Nat. Neurosci.* **14**, 366–372 (2011).
93. Behrens, T. E. J., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. S. Learning the value of information in an uncertain world. *Nat. Neurosci.* **10**, 1214–1221 (2007).
94. Chaudhuri, R., Knoblauch, K., Gariel, M.-A., Kennedy, H. & Wang, X.-J. A large-scale circuit mechanism for hierarchical dynamical processing in the primate cortex. *Neuron* **88**, 419–431 (2015).
95. Murray, J. D. *et al.* A hierarchy of intrinsic timescales across primate cortex. *Nat. Neurosci.* **17**, 1661–1663 (2014).
96. Purcell, B. A. & Kiani, R. Hierarchical decision processes that operate over distinct timescales underlie choice and changes in strategy. *Proc. Natl Acad. Sci. USA* **113**, E4531–E4540 (2016).
97. Boorman, E. D., Rushworth, M. F. & Behrens, T. E. Ventromedial prefrontal and anterior cingulate cortex adopt choice and default reference frames during sequential multi-alternative choice. *J. Neurosci.* **33**, 2242–2253 (2013).
98. Wittmann, M. K. *et al.* Predictive decision making driven by multiple time-linked reward representations in the anterior cingulate cortex. *Nat. Commun.* **7**, 12327 (2016).
99. Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B. & Dolan, R. J. Cortical substrates for exploratory decisions in humans. *Nature* **441**, 876–879 (2006).
100. Karlsson, M. P., Tervo, D. G. R. & Karpova, A. Y. Network resets in medial prefrontal cortex mark the onset of behavioral uncertainty. *Science* **338**, 135–139 (2012).
101. Blanchard, T. C., Strait, C. E. & Hayden, B. Y. Ramping ensemble activity in dorsal anterior cingulate neurons during persistent commitment to a decision. *J. Neurophysiol.* **114**, 2439–2449 (2015).
102. Desrochers, T. M., Chatham, C. H. & Badre, D. The necessity of rostralateral prefrontal cortex for higher-level sequential behavior. *Neuron* **87**, 1357–1368 (2015).
103. Badre, D. & D'Esposito, M. Is the rostro-caudal axis of the frontal lobe hierarchical? *Nat. Rev. Neurosci.* **10**, 659–669 (2009).
104. Kounieher, F., Charron, S. & Koehlin, E. Motivation and cognitive control in the human prefrontal cortex. *Nat. Neurosci.* **12**, 939–945 (2009).
105. O'Reilly, R. C. The what and how of prefrontal cortical organization. *Trends Neurosci.* **33**, 355–361 (2010).
106. Zarr, N. & Brown, J. W. Hierarchical error representation in medial prefrontal cortex. *Neuroimage* **124**, 238–247 (2016).
107. Nee, D. E. & Brown, J. W. Rostral-caudal gradients of abstraction revealed by multi-variate pattern analysis of working memory. *Neuroimage* **63**, 1285–1294 (2012).
108. Nee, D. E. & D'Esposito, M. The hierarchical organization of the lateral prefrontal cortex. *eLife* **5**, e12112 (2016).
109. Badre, D., Kayser, A. S. & D'Esposito, M. Frontal cortex and the discovery of abstract action rules. *Neuron* **66**, 315–326 (2010).
110. Frank, M. J. & Badre, D. Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: computational analysis. *Cereb. Cortex* **22**, 509–526 (2012).
111. Hochreiter, S. & Schmidhuber, J. Long short-term memory. *Neural Comput.* **9**, 1735–1780 (1997).
112. O'Reilly, R. C. & Frank, M. J. Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural Comput.* **18**, 283–328 (2006).
113. Dursteler, M. R., Wurtz, R. H. & Newsome, W. T. Directional pursuit deficits following lesions of the foveal representation within the superior temporal sulcus of the macaque monkey. *J. Neurophysiol.* **57**, 1262–1287 (1987).
114. Barton, J. J., Press, D. Z., Keenan, J. P. & O'Connor, M. Lesions of the fusiform face area impair perception of facial configuration in prosopagnosia. *Neurology* **58**, 71–78 (2002).
115. Noonan, M. P. *et al.* Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. *Proc. Natl Acad. Sci. USA* **107**, 20547–20552 (2010).
116. Walton, M. E., Behrens, T. E., Buckley, M. J., Rudebeck, P. H. & Rushworth, M. F. Separable learning systems in the macaque brain and the role of orbitofrontal cortex in contingent learning. *Neuron* **65**, 927–939 (2010).
117. Vaidya, A. R. & Fellows, L. K. Testing necessary regional frontal contributions to value assessment and fixation-based updating. *Nat. Commun.* **6**, 10120 (2015).
118. Eslinger, P. J. & Damasio, A. R. Severe disturbance of higher cognition after bilateral frontal lobe ablation: patient EVR. *Neurology* **35**, 1731–1741 (1985).
119. Szczepanski, S. M. & Knight, R. T. Insights into human behavior from lesions to the prefrontal cortex. *Neuron* **83**, 1002–1018 (2014).
120. Shadlen, M. N. & Shohamy, D. Decision making and sequential sampling from memory. *Neuron* **90**, 927–939 (2016).
121. Plassmann, H., O'Doherty, J. & Rangel, A. Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J. Neurosci.* **27**, 9984–9988 (2007).
122. Knutson, B., Taylor, J., Kaufman, M., Peterson, R. & Glover, G. Distributed neural representation of expected value. *J. Neurosci.* **25**, 4806–4812 (2005).
123. McCoy, A. N. & Platt, M. L. Risk-sensitive neurons in macaque posterior cingulate cortex. *Nat. Neurosci.* **8**, 1220–1227 (2005).
124. Kim, S., Hwang, J. & Lee, D. Prefrontal coding of temporally discounted values during intertemporal choice. *Neuron* **59**, 161–172 (2008).
125. Louie, K., Grattan, L. E. & Glimcher, P. W. Reward value-based gain control: divisive normalization in parietal cortex. *J. Neurosci.* **31**, 10627–10639 (2011).
126. Shuler, M. G. Reward timing in the primary visual cortex. *Science* **311**, 1606–1609 (2006).
127. Philiastrides, M. G., Biele, G. & Heekeren, H. R. A mechanistic account of value computation in the human brain. *Proc. Natl Acad. Sci. USA* **107**, 9430–9435 (2010).
128. Peck, C. J., Lau, B. & Salzman, C. D. The primate amygdala combines information about space and value. *Nat. Neurosci.* **16**, 340–348 (2013).
129. Cai, X., Kim, S. & Lee, D. Heterogeneous coding of temporally discounted values in the dorsal and ventral striatum during intertemporal choice. *Neuron* **69**, 170–182 (2011).
130. Heilbronner, S. R. & Hayden, B. Y. Dorsal anterior cingulate cortex: a bottom-up view. *Annu. Rev. Neurosci.* **39**, 149–170 (2016).
131. Vickery, T. J., Chun, M. M. & Lee, D. Ubiquity and specificity of reinforcement signals throughout the human brain. *Neuron* **72**, 166–177 (2011).
132. Maunsell, J. H. R. Neuronal representations of cognitive state: reward or attention? *Trends Cogn. Sci.* **8**, 261–265 (2004).
133. Leathers, M. L. & Olson, C. R. In monkeys making value-based decisions, LIP neurons encode cue salience and not action value. *Science* **338**, 132–135 (2012).
134. Litt, A., Plassmann, H., Shiv, B. & Rangel, A. Dissociating valuation and saliency signals during decision-making. *Cereb. Cortex* **21**, 95–102 (2010).
135. Vlaev, I., Chater, N., Stewart, N. & Brown, G. D. Does the brain calculate value? *Trends Cogn. Sci.* **15**, 546–554 (2011).
136. Lichtenstein, S. & Slovic, P. Reversals of preference between bids and choices in gambling decisions. *J. Exp. Psychol.* **89**, 46–55 (1971).
137. Churchland, M. M. *et al.* Neural population dynamics during reaching. *Nature* **487**, 51–56 (2012).
138. Shenoy, K. V., Sahani, M. & Churchland, M. M. Cortical control of arm movements: a dynamical systems perspective. *Annu. Rev. Neurosci.* **36**, 337–359 (2013).
139. Hunt, L. T. What are the neural origins of choice variability? *Trends Cogn. Sci.* **18**, 222–224 (2014).
140. Levy, D. J. & Glimcher, P. W. The root of all value: a neural common currency for choice. *Curr. Opin. Neurobiol.* **22**, 1027–1038 (2012).
141. Svoboda, E., McKinnon, M. C. & Levine, B. The functional neuroanatomy of autobiographical memory: a meta-analysis. *Neuropsychologia* **44**, 2189–2208 (2006).
142. Doeller, C. F., Barry, C. & Burgess, N. Evidence for grid cells in a human memory network. *Nature* **463**, 657–661 (2010).
143. Schacter, D. L. *et al.* The future of memory: remembering, imagining, and the brain. *Neuron* **76**, 677–694 (2012).
144. Van Overwalle, F. Social cognition and the brain: a meta-analysis. *Hum. Brain Mapp.* **30**, 829–858 (2009).
145. Mansouri, F. A., Buckley, M. J. & Tanaka, K. The essential role of primate orbitofrontal cortex in conflict-induced executive control adjustment. *J. Neurosci.* **34**, 11016–11031 (2014).
146. Slezeev, B. J., Castagno, M. D. & Hayden, B. Y. Rule encoding in orbitofrontal cortex and striatum guides selection. *J. Neurosci.* **36**, 11223–11237 (2016).
147. Lopatina, N. *et al.* Lateral orbitofrontal neurons acquire responses to upshifted, downshifted, or blocked cues during unblocking. *eLife* **4**, e11299 (2015).
148. McDannald, M. A. *et al.* Orbitofrontal neurons acquire responses to 'valueless' Pavlovian cues during unblocking. *eLife* **3**, e02653 (2014).

149. Tsujimoto, S., Genovesio, A. & Wise, S. P. Neuronal activity during a cued strategy task: comparison of dorsolateral, orbital, and polar prefrontal cortex. *J. Neurosci.* **32**, 11017–11031 (2012).
150. Kepecs, A., Uchida, N., Zariwala, H. A. & Mainen, Z. F. Neural correlates, computation and behavioural impact of decision confidence. *Nature* **455**, 227–231 (2008).
151. Genovesio, A., Tsujimoto, S., Navarra, G., Falcone, R. & Wise, S. P. Autonomous encoding of irrelevant goals and outcomes by prefrontal cortex neurons. *J. Neurosci.* **34**, 1970–1978 (2014).
152. Rigotti, M. *et al.* The importance of mixed selectivity in complex cognitive tasks. *Nature* **497**, 585–590 (2013).
153. Raposo, D., Kaufman, M. T. & Churchland, A. K. A category-free neural population supports evolving demands during decision-making. *Nat. Neurosci.* **17**, 1784–1792 (2014).
154. Felleman, D. J. & Van Essen, D. C. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* **1**, 1–47 (1991).
155. Markov, N. T. *et al.* A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cereb. Cortex* **24**, 17–36 (2012).
156. Dombrowski, S. M., Hilgetag, C. C. & Barbas, H. Quantitative architecture distinguishes prefrontal cortical systems in the rhesus monkey. *Cereb. Cortex* **11**, 975–988 (2001).
157. Alexander, G. E., DeLong, M. R. & Strick, P. L. Parallel organization of functionally segregated circuits linking basal ganglia and cortex. *Annu. Rev. Neurosci.* **9**, 357–381 (1986).
158. Seeley, T. D. & Buhrman, S. C. Group decision making in swarms of honey bees. *Behav. Ecol. Sociobiol.* **45**, 19–31 (1999).
159. Seeley, T. D. *et al.* Stop signals provide cross inhibition in collective decision-making by honeybee swarms. *Science* **335**, 108–111 (2011).

Acknowledgements

The authors are grateful to many colleagues with whom interactions have shaped ideas within this article, in particular T. Behrens, T. Blanchard, S. Kennerley, J. Pearson, S. Piantadosi, M. Platt, M. Rushworth and T. Seeley. The authors thank H. Barron, A. de Berker, R. Dolan, M. A. Noonan and T. Seow for comments on a previous draft of the manuscript. L.T.H. is supported by a Sir Henry Wellcome Fellowship from the Wellcome Trust (098830/Z/12/Z). B.Y.H. is supported by a US National Institutes of Health award (DA037229).

Competing interests statement

The authors declare no competing interests.