

**MANUSCRIPT IN PRESS**  
**PLEASE DO NOT REDISTRIBUTE WITHOUT PRIOR PERMISSION**

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

Laurence T. Hunt<sup>1</sup> and Benjamin Y. Hayden<sup>2</sup>

<sup>1</sup> Senior Research Associate, Wellcome Trust Centre for Neuroimaging,  
University College London, UK

<sup>2</sup> Associate Professor of Brain and Cognitive Sciences, University of Rochester,  
USA

## **Preface**

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

## Introduction

Neurobiologists have long been interested in developing mechanistic models to explain how we evaluate options and choose the best course of action<sup>1-3</sup>. Many accounts take a modular perspective. That is, they assume, even if only tacitly, that goal-directed choice can be subdivided into a set of component processes whose neural implementation is both *serial* and *localized*<sup>4-7</sup>. The discrete steps typically include *evaluation* of options, *comparison* of pure values absent of any other factors, *selection* of an appropriate action plan, and *monitoring* the outcome of the choice. From this perspective, the component processes are assumed to correspond to discrete neural computations implemented in distinct neural structures.

An alternative line of research takes features of neural circuit anatomy as a starting point. These approaches construct *circuit-based* models that predict both behavioural and neural data, whilst retaining biological plausibility at their core<sup>8-14</sup>. Recent research from these approaches emphasizes three overarching principles (Figure 1). First, decisions may be formed in a ***distributed*** fashion across many brain regions acting in concert performing similar computations. Second, the distributed networks implementing choices are highly ***recurrent*** in nature, which affects the kinds of computation that are performed. Finally, these distributed and recurrent networks are organised into functional and temporal ***hierarchies***.

Centring behavioural models on neural circuit plausibility builds upon several traditions. Its origins can be traced back to mid-century research programs of pioneers like Weiner, Hull, Hebb, and McCulloch and Pitts<sup>15,16</sup>. 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, capturing visual recognition and memory storage (among other topics). The key insight was that neuron-like units performing biophysically plausible computations and connected in simple ways could perform astonishingly rich computations<sup>17-20</sup>. Importantly, 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 it<sup>17</sup>.

These traditions continue to influence neuroscience to this day. Their influence on decision making has been expressed by a number of computational neuroscientists, such as Wang, Frank, O'Reilly, Rolls and Cisek, amongst others<sup>8-13</sup>. Recently, a number of studies have tested key empirical predictions from circuit-based accounts of decision-making. Moreover, these advances have recently coincided with a resurgence of interest in neural networks within machine learning and computer science<sup>21</sup>, and several breakthroughs in neuroanatomical<sup>22</sup> and circuit-breaking<sup>23</sup> methods. These should allow for far more rigorous testing and refinement of circuit-based models in the near future. It is therefore particularly timely to consider the contribution that circuit-based models make to our understanding of reward-guided choice.

In this review, we first outline the key principles behind a distributed, hierarchical and recurrent account of reward-guided choice. We then discuss some empirical motivations. 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 *eliminative*<sup>24</sup>. This means that the neural implementation of choice need not necessarily recapitulate the steps or modules 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 different core set of research questions than the modular approach (see Box 1).



## Properties of a distributed, hierarchical, and recurrent account of choice

### Distributed

A distributed decision is one in which separate elements perform subsidiary computations that, when combined, produce the ultimate decision. Well-known examples include the actions of individual voters in a national election and the selection of hive sites by swarms of bees<sup>25</sup> (see Box 2). 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, informationally limited, decision-maker<sup>26,27</sup>. It non-linearly transforms its dendritic inputs to firing rate. However, this transformation of inputs to 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-organised distributed system that includes multiple brain regions. Any neuron whose activity influences the behaviour of the network 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 process<sup>28</sup>.

*Distributed transformations within cortical circuits:* Distributed decisions often involve a single simple computation repeated in each element on different inputs. One strong candidate in economic choice is *competition through mutual inhibition*. Mutual inhibition is a common motif throughout the nervous system, and is often considered part of a basic repertoire of neural circuits<sup>29,30</sup>. Effective competition via mutual inhibition can be mediated in a biophysically realistic cortical circuit model by appropriate choice of synaptic weights<sup>11,13</sup>. Several recent results indicate that mutual inhibition may be at the core of reward-guided choice<sup>31-42</sup> (Figure 2).

Behrens and colleagues recently used magnetoencephalography (MEG) to investigate economic choice, and demonstrated that human ventromedial

prefrontal cortex (vmPFC) and intraparietal sulcus expressed a key signature of mutual inhibition: a change from encoding the sum to the difference between chosen and unchosen values<sup>31</sup>. This signature was subsequently found in a macaque study of local field potential (LFP) recorded from several subregions of macaque PFC, including orbitofrontal cortex (OFC), dorsolateral prefrontal cortex (DLPFC) and anterior cingulate cortex (ACC), confirming a cross-species generality of the mutual inhibition principle<sup>32</sup>.

Single neuron recordings in macaques provide further evidence for a mutual inhibitory process. In a binary choice, firing rates of neurons in both vmPFC<sup>34</sup> and ventral striatum (VS)<sup>35</sup> encode the values of the two offers through monotonic changes in firing rates. Critically, during the comparison period of the task, the directions of the tuning curves (positive or negative) for the two offers are opposed<sup>34,35</sup>. 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 timecourses, suggests that neither is the sole site of comparison, but that comparison may take place in both simultaneously.

Nor are the vmPFC and VS likely to be unique; the mutual inhibition model has also been used to capture the dynamics of single unit activity in OFC<sup>36</sup>, DLPFC<sup>32</sup>, and lateral intraparietal cortex (LIP)<sup>39,42</sup>. Data consistent with a mutual inhibition process are particularly clear in dorsal premotor cortex (PMd). 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<sup>40,41</sup>. This finding suggests that even ostensibly motor areas are part of the distributed choice network.

*Distributed transformations across cortical circuits:* Collectively, this body of work suggests that decisions occur through repeated mutual inhibition computations occurring simultaneously in both motor and abstract value domains. This implies that comparison is not the unique purview of any single brain area<sup>43-45</sup>. The distributed account instead suggests that multiple areas may perform a similar non-specialized function: they may all perform a comparison operation like mutual inhibition on the inputs received. However, those inputs

would differ by region. The nature of the competition occurring in any given area would depend upon the interaction of the particular task demands<sup>46</sup> and that area's anatomical inputs<sup>47,48</sup>. As a consequence of this, tuning properties of neurons in OFC may be relatively specialized for gustatory comparisons<sup>49</sup>, whilst neurons in cingulate cortex may appear specialized for motor cost evaluations<sup>50</sup>.

This idea of multiple distributed comparators may help to resolve differing interpretations of imaging studies in neuroeconomics, which have attempted to localise regions most critical to value comparison. 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 functional magnetic resonance imaging (fMRI) signal<sup>33</sup>. 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 fMRI signal will be greatest for faster (easy) decisions, as is typically the case in VMPFC<sup>51</sup>. However, other studies have argued that one should only consider accumulated activity until a decision boundary is reached. In this case fMRI signal would be higher for slower (difficult) decisions, as is typically found in dorsomedial PFC<sup>52</sup>.

A straightforward reconciliation of these results would be that both regions implement a mutual inhibition process but then differ in their response properties post-choice after the decision bound has been reached. This is supported by the event-related profile of LFP recordings from multiple subregions of PFC during reward-based choice<sup>53</sup>. It is also evidenced from the simultaneous emergence of single unit choice-related signals across six simultaneously recorded cortical regions in perceptual choice<sup>54</sup>, and the demonstration that motor output (corticospinal excitability) is biased as decisions unfold<sup>55</sup>.

The question then arises of how these different areas interact as choices are made. Here whole-brain techniques (such as fMRI) come into their element<sup>56</sup>. While anatomical connectivity is stable, functional connectivity is more flexible. One recent study examined this using multiattribute choices, involving integration of stimulus-based and action-based attributes<sup>37</sup>. A model of choice in

which competition via mutual inhibition occurred at multiple levels (stimuli, actions, and attribute) best explained subjects' choices, and blood-oxygen level dependent (BOLD) fMRI signal in the intraparietal sulcus (IPS) matched with predictions from the competition over which attribute to attend. Notably, changes in IPS functional connectivity to other brain areas depended upon which attribute was most relevant to the current choice 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<sup>57</sup>.

## **Recurrent**

The predominant paradigm in studies of simple economic choice has been N-alternative forced choice, where 2 or more options are presented simultaneously and the subject selects the most valuable. Yet real-world choice is often sequential rather than simultaneous in nature<sup>58,59</sup>. 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<sup>2,14,60,61</sup>. We believe that the relationship between simultaneous and sequential accounts of choice parallels a distinction made between the computational roles of two distinct architectures of neural network<sup>21</sup>: *feedforward* and *recurrent*.

*Different forms of network architecture:* Feedforward neural networks – recently popularised through the impressive performance of convolutional networks in computer vision (ConvNets<sup>62,63</sup>) – contain units whose activity is only dependent upon the currently presented input. They are ideally suited to tasks involving classification of static inputs, such as image recognition. Notably, features of neurophysiological and neuroanatomical data directly influenced their success. The 'local convolution' and 'max pooling' steps in a ConvNet were designed based upon the response properties of V1 simple and complex cells, respectively<sup>62</sup>. Modern feedforward ConvNet models strikingly reproduce single unit responses along the visual hierarchy<sup>64,65</sup>.

However, feedforward nets are poorly adapted to tasks that require some form of persistent memory for previous states across time. Recurrent neural

networks (RNNs), by contrast, contain units that not only receive inputs from other network layers, but also receive their own *previous output* at time  $t-1$ <sup>18,66</sup>. This allows the network to show sustained memory for inputs long after they have been removed, allowing temporally extended computations to be performed<sup>67,68</sup>. One of the original successes of such networks were in showing how a biologically plausible network could account for working memory responses in PFC<sup>69,70</sup>. Notably, the recurrence of these networks recapitulates the high degree of recurrent connectivity observed empirically in prefrontal circuits relative to other parts of cortex<sup>71,72</sup>.

*Recurrent networks as a consequence of naturalistic choices:* At first glance, the simultaneous simple economic choices between two goods often studied in the lab appear well suited to being solved with a feedforward network. The options in such choices are, after all, presented to the subject simultaneously and statically. Yet it is well known that reaction times in such tasks vary systematically as a function of value, implying that the underlying computation is dynamic rather than static<sup>73</sup>. Recent models in neuroeconomics have moved away from static economic accounts of choice towards temporally extended algorithms in which noisy estimates of value are integrated over time<sup>14,52</sup>. These algorithms can be formally related to the model of competition via mutual inhibition discussed above and in Figure 2<sup>74</sup>. This network exemplifies how a biologically plausible RNN could implement an evidence accumulation algorithm<sup>11</sup>.

But why should the brain use a recurrent network when a feedforward architecture appears sufficient? It may be because of the sequential structure of choices that are typically faced in more natural environments, such as foraging decisions. These choices are likely to have shaped the evolution of the frontal cortex most heavily<sup>43,75</sup>. 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<sup>76,77</sup>. For example, when macaque monkeys perform a foraging task, activity within 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 triggered<sup>78</sup>.

The *algorithmic* structure of economic and foraging models is certainly very different, and one possibility would be that they are implemented by completely different brain systems: the former being feedforward, and the latter being recurrent. We consider it more likely, however, that evolution will have developed a general approach to solving both types of decisions using similar temporally extended (and thus recurrent) implementations. The neural mechanisms of economic choice will have to build upon this underlying architecture to solve reward-guided choices using recurrent computations. This fits well with modern ideas in neuroeconomic accounts that argue for temporal integration of noisy value estimates across time<sup>14,52,73</sup>, with attention being sequentially allocated to different choice options<sup>79,80</sup>.

*Understanding recurrent network computations:* The majority of neuroscientific studies to date using RNNs have used networks containing hand-tuned parameters to elicit specific behaviours<sup>11</sup>. Whilst this is a valuable starting point, it produces unrealistic homogeneity of neuronal responses that do not match the empirically observed heterogeneity in cortical populations<sup>81-83</sup>. Recently introduced automatic training algorithms for RNNs have vastly increased their capacity to perform a wide variety of tasks and thus, potentially, to accurately describe neural data<sup>84,85</sup>. 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 is to perform ‘reverse engineering’ on RNNs that have been fit to data<sup>86</sup>. This was exemplified in a recent study of PFC population responses, where PFC performed both selection and integration of relevant information in a context-dependent perceptual decision task<sup>87</sup>. An RNN was developed that captured key features of both the behavioural and neural data. Reverse engineering of the RNN then 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 upon how the network’s 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 future<sup>88</sup>.

## **Hierarchical**

A hierarchical organization is a cardinal feature of the organization of the brain's reward and decision-making system. Information from one area is converted to a more abstract form that is more comprehensive as it increases in complexity. One unique feature of a hierarchy of brain areas associated with economic choice is that there are new inputs at each point. Thus, for example, OFC receives gustatory inputs, vmPFC receives limbic inputs, subgenual ACC receives hypothalamic inputs, and dACC receives motor inputs<sup>89</sup>. These inputs provide a way for different factors to enter into the distributed network subserving choice. They also allow for 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 an alternative in which all factors that influence choices must come together at a single point to create a single value scale, before they can influence choice.

*Hierarchical RNNs allow for multiple timescales:* Humans excel at tasks that demand online organisation of behaviour across multiple differing timescales. Such tasks are particularly sensitive to PFC damage<sup>90</sup>. It has been hypothesised that the presence of parallel and hierarchical architectures within PFC allow different pieces of information to be remembered and parsed at multiple timescales<sup>91,92</sup>. This appears to recapitulate recent developments in the design of RNNs within computer science. In particular, by making recurrent networks multi-layered or 'deep'<sup>21,93</sup>, it is possible to dramatically improve performance on tasks that operate over multiple timescales (such as generation of meaningful sentences to describe images<sup>94</sup>, or recognition of speech<sup>67</sup>). The relevant timescales in speech are unknown to the neural network prior to 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 empirically<sup>95</sup>. These timescales vary dependent upon the timescale of integration of the animal's current choice behaviour<sup>95,96</sup>.

*Hierarchical timescales in cortex:* A hierarchy of timescales across brain regions has recently been explicated in a large-scale network model of dynamical processing in primate cortex. Chaudhuri and colleagues combined a mean-field reduction of a RNN within each cortical area with detailed knowledge of anatomical connectivity between different areas derived from tracer studies<sup>97</sup>. 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 areas<sup>98</sup>. This work shows why early sensory areas possess an inherently transient temporal structure (changing across tens of milliseconds), whilst higher regions (e.g. in PFC) show a more sustained, longer-lasting temporal profile (across hundreds of milliseconds, or longer) (Fig 3a).

Recent decision-making paradigms have combined evidence accumulation at both slow (inter-trial) and fast (intra-trial) timescales<sup>99</sup>. It might then be predicted that evidence accumulation at slow timescales would be supported by those regions with sustained temporal structure. Interestingly, dACC emerges as having the most sustained temporal structure within the regions that have been characterised thus far<sup>98</sup>. This sits well with studies that have linked ACC activity to disengagement from a foreground option, for example in foraging tasks<sup>77,78,100,101</sup> and during exploration<sup>102</sup>. These disengagements are linked to abrupt and coordinated changes in network activity in medial prefrontal cortex across trials<sup>103</sup> (Fig 3b). Prior to disengagement, ramping activity is observed in both medial prefrontal cortex/dACC<sup>78,104,105</sup> and also another region associated with exploration<sup>102</sup>, rostralateral prefrontal cortex<sup>105</sup>.

*Relationship to cognitive hierarchies:* The observed relationship between anatomical and temporal hierarchies may be related to hierarchical accounts of PFC function in other cognitive paradigms<sup>106</sup>. One account of this suggests two



parallel streams of rostro-caudal organization within the prefrontal cortex, along the medial and lateral surfaces of PFC respectively<sup>107</sup> (Fig 3c). Along the medial surface, BOLD fMRI activity suggests a hierarchy relating to states of internal motivation<sup>107-109</sup>. More anterior portions reflect block-wise changes in reward value, whereas more posterior regions reflect trial-to-trial changes. By contrast, along the lateral surface activity show a hierarchy of cognitive control processes<sup>107,110,111</sup>. Again, anterior regions exhibit sustained changes reflecting the complexity of the current block, whereas more posterior regions reflect task complexity only for relevant individual trials. Similar rostro-caudal PFC hierarchies have also been found in a related study where activation was found to transfer to the striatum during the course of hierarchical rule learning<sup>112</sup>. This observation was well described by a hierarchically organized network model of corticostriatal interactions<sup>113</sup> based on a particularly successful form of RNN known as 'long short term memory'<sup>114,115</sup>.

## **Evidence for a distributed, hierarchical, and recurrent account of reward-guided choice**

### **No lesion causes catastrophic deficits in economic choice**

An important property of distributed computing systems is *graceful degradation*<sup>17</sup>. Because information is stored broadly, small amounts of damage to the system is 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, lesions to MT cause akinetopsia<sup>116</sup> and lesions to FFA cause prosopagnosia<sup>117</sup>, confirming functional specialization for motion perception and face processing, respectively.

It has been natural for experimenters to design lesion studies that assign different components to different brain regions, and these have indeed advocated a degree of functional specialisation within PFC<sup>118-120</sup>. Yet given the profound impact of PFC lesions on day-to-day choices<sup>121</sup>, it is often surprising how mild impairments in these studies can be. PFC lesions sometimes lead to graded deficits in choosing, with difficult decisions impaired but easier ones spared<sup>118</sup>. It is possible that these deficits are mild because the tasks seek to differentiate component processes of economic choice, but do not address the temporally extended nature of many real-world decisions<sup>90</sup>.

The variety of cognitive functions affected by lesions casts doubt on the idea that choice, or even value comparison, is a specialisation of any single brain region. This is also apparent in functional imaging data from studies of simple economic choice: these have found several different types of comparison signal in differing brain areas, depending upon the task at hand<sup>46,52,53,77</sup>. Moreover, these regions are not specific to choice: lesions to the reward system impair a variety of other tasks beyond decision making, including working memory, strategic planning, reasoning, and social cognition<sup>122</sup>. Together, this evidence suggests that many brain regions will collectively contribute to the process of comparison, and also that brain regions subserving choice are making their contribution as part of a larger supported suite of cognitive functions. By way of

analogy, reward-guided choice seems to operate more like an ability to drive a car, form a political preference, or do calculus in so far as 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 upon distributed synaptic weight changes across all parts of a network, rather than at a single site<sup>17</sup>. Value is closely related to memory: it is a feature of an option inferred from associations with reward, based on past experience<sup>123</sup>. In connectionist systems, memories are widely distributed (a feature leading to graceful degradation, see above), and traces of those memories can be observed throughout the network<sup>17</sup>; 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 OFC<sup>49</sup>, vmPFC<sup>124</sup>, and ventral striatum (VS)<sup>125</sup>, but also in amygdala, insula, dorsal striatum, midbrain, pregenual, subgenual, dorsal anterior, and posterior cingulate cortices, dorsolateral and ventrolateral prefrontal cortex, the intraparietal sulcus, and even sensory and motor cortices<sup>126-133</sup>. One recent neuroimaging study found that over 30% of the brain exhibited such signals<sup>134</sup>. While there are certainly differences between the types of value-related information represented in these regions, there are also many overlaps as well.

Despite the ubiquity of value correlates, it often remains unclear whether activity in any area truly *represents* value – or even what precise definition of value we should use<sup>28,135,136</sup>. 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<sup>28,136,137</sup>. One possibility is that one of these value correlates is genuine representation of value, in the formal sense. However, in distributed systems, representation is often an emergent property, that is, driven by the specific pattern of connection between units, and not a property of any particular unit. Thus, it is also possible that value is an emergent property of the

brain's reward networks, one that is convenient for thinking about choice, but one that is not observed in particular brain regions of neurons. While this view is uncommon, it would be consistent with a subset of both classic and recent work in behavioural economics on process models of economic choice<sup>59,138,139</sup>.

One particular form of value seems likely to be 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, consider the model of competition via mutual inhibition discussed above<sup>11</sup>, and in figure 2. Although many units' activities within this network model correlate with *chosen value*, this quantity need not actually be represented in order to form a decision. Instead, it arises naturally as a consequence of the varying speed at which network *dynamics* unfold on different trials<sup>31</sup> (see Box 3). To test this idea, it was recently demonstrated that single unit correlates of chosen value can be partially explained away as a consequence of cross-trial variation in the dynamics of the local decision process<sup>32</sup>. The speed of decision formation on each trial was estimated using a principal component analysis approach to extract time-varying features of the local field potential, and then used to explain simultaneously recorded neuronal activity. Applying the same analysis to the competition via mutual inhibition RNN also elicited a similar result.

This also links to recent findings in motor control, which argue that motor cortical activity is better understood as a dynamical system than one representing movement parameters<sup>140,141</sup>. It suggests that chosen value correlates emerge as a necessary consequence of recurrent network dynamics in mediating competition<sup>142</sup> – and that chosen value may not be represented *per se*. While this work is limited to chosen value correlates, it remains possible that other ostensible value correlates (such as offer value and experienced value) are likewise by-products of the computations that underlie choice, and not reified in the activity of dedicated reward neurons or regions<sup>138</sup>.

### **Elusive *pure* value**

A modular view predicts that certain brain regions or populations of neurons should be *specialised* 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. On the other hand, 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 play other roles unrelated to valuation.

Some meta-analyses of neuroimaging data have argued that certain brain areas are central to ‘pure valuation’<sup>143</sup>. Yet even within putative core reward regions such as vmPFC, OFC and VS, there remains a wealth of information not related to value. For instance, vmPFC is engaged by several ostensibly value-neutral factors including autobiographical memory<sup>144</sup>, spatial navigation<sup>145</sup>, imagination<sup>146</sup>, and social cognition<sup>147</sup>. Likewise, factors that drive individual neurons in these areas may include ‘valueless’ changes in outcome expectancy<sup>148,149</sup>, previous outcomes<sup>95</sup>, intention to switch as well as other strategy variables<sup>150</sup>, metacognition<sup>151</sup>, spatial positions of offers and choices<sup>152</sup>, rules and task set, and even irrelevant task variables<sup>153</sup>. It is possible that these apparently value-neutral signals remain present in these tasks because the tasks nevertheless demand the computation of value. However, it is becoming 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<sup>154,155</sup>.

### **Anatomical hierarchies, reciprocity and recurrence**

Finally, 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 system<sup>15</sup>. Perceptrons, feed-forward networks, parallel distributed networks, Hebbian learning and ConvNets were all inspired from observations about brain structure and physiology.

The more specific details of this approach are also consistent as well. Within cortex, anatomical studies have indicated a reciprocal feedforward and feedback structure to connectivity<sup>156</sup>. This reciprocally connected architecture has recently been quantified via detailed tracer studies in macaques<sup>157</sup> and

diffusion imaging studies in humans<sup>22</sup>. Differences between adjacent prefrontal regions are smaller than is often appreciated, and adjacent regions generally blur into each other gradually rather than showing categorical boundaries<sup>158</sup>.

Likewise, when considering the organization of subcortical brain regions, it is now accepted that cortico-thalamo-basal ganglia loops are not segregated as once thought<sup>159</sup> but instead show strong functional convergence<sup>89</sup>.

Local intraregional connections within cortex also argue for a distributed and recurrent organization. This is particularly true of 'higher' cortical areas that subserve cognitive functions. In particular, dendritic arbors of PFC pyramidal cells are endowed with many more dendritic spines than in sensory areas, meaning that a single neuron in prefrontal areas 10/11/12 receives 16 times the number of excitatory inputs of a neuron in V1<sup>72</sup>. The majority of these cortical connections are local rather than long-distance, allowing these circuits to have a highly recurrent organization similar to that observed in RNNs.

## Discussion

We have presented a review 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 historically 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 decisions<sup>8-14</sup>. Elements of these ideas are observed in algorithmic models of choice, as well<sup>14</sup>.

According to the distributed view, the implementation of economic choice is dissimilar to a description of how it works at a more abstract level. Choice is an emergent consequence of the interactions of small computational elements<sup>160</sup>. By contrast, much recent research into the neurobiology of choice has adopted a modular framework, in which major components of choice map directly onto brain structures and discrete computations. Thus for example, this framework encourages scholars to look for the specialized sites of evaluation, comparison and action selection. The change in viewpoint we propose leads us to reframe this debate and many other central questions (Box 1).

The work we describe here brings up a philosophical question that has long influenced cognitive science, the question of mental *representation*<sup>16</sup>. 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. While 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, in regards to chosen value, it seems that while value is decodable, it is possible that it is an artefact of the way neural data are analysed<sup>53</sup>. Second, in regards to value more broadly, the case is less clear, but connectionist models suggest that is possible to construct networks that make good choices without explicit value representations; these networks at the very least have a similar flavour to neuroanatomy<sup>17</sup>.

Paul and Patricia Churchland have articulated the notion of eliminative materialism, which includes a suggestion that natural categories we use to

describe psychological phenomena do not do a good job of capturing the organization of brain processes that generate our mental lives<sup>24</sup>. 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, discrete neuron types, or to even to discrete computations. Instead, they are an emergent consequence of processing units performing simple operations on inputs, ones that are radically different from the operations of the system as a whole<sup>24</sup>.



## Inset Boxes

### Inset Box 1: A change of perspective

Adopting a circuit-based perspective 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 to us asking 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’ does value comparison occur? (e.g. goods-based or action-based)*
- *Does a particular brain region precede or follow the decision?*
- *What are the qualitative differences between different regions of the reward system?*

By contrast, the distributed, hierarchical and recurrent perspective makes a different set of assumptions. These in turn raise different research questions:

- The component processes of choice may not be localised to particular computations in discrete brain areas. They 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?*
- Comparison may not occur in a single decision space: decision spaces may reflect the anatomical connections of a given region, or even be artifacts 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 hemodynamic and neuronal response patterns?*
- 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 brain’s ubiquitous recurrent and feedback connections play in that process?*
- Value, or at least certain correlates of value, may not be represented *per se*. The implementation of choice need not 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? Which correlates of value is this true for?*

### **Inset Box 2: Distributed decision-making in beehives**

While modular decision systems are often intuitive – the functions map directly onto the structure – distributed ones are not. Here we describe the decision process of a bee swarm, which provides a natural and well understood example of a 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<sup>25,161</sup>. Individual scout bees make reconnaissance flights to identify and evaluate potential sites. Ideal sites are open, dry, cavities of medium volume located high up in the canopy, protected from wind, and facing south, so adaptive decisions will optimize across many dimensions. On finding a potential site, each scout returns to the swarm and signals its location and quality through the patterns of dancing. Dances signalling high quality sites can recruit other bees to investigate the same site. Subsequent 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. Intriguingly, bee swarms even show a distinct mutual inhibition signal that reduces the chance of costly ambivalence<sup>162</sup>.

Beehive decision-making has several notable features that make it a good analogy for distributed decision-making<sup>25</sup>. First, there is no localized evaluation: no individual bee has more than an extremely limited amount to 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 associated with choice and selection<sup>45</sup>.

As a consequence, the steps of evaluation, comparison, and selection are clearly performed by bees, but at the same time do not have clear analogues in the behaviour of individual bees, in specialized subgroups of bees, or in particular actions. No individual bee ever has knowledge of more than one hive site, so no bee performs a comparison<sup>161</sup>; instead the comparison step emerges as a consequence of the types of interactions the bees are programmed to perform.

### **Inset Box 3: Value representations as epiphenomena of decision dynamics**

One particularly ubiquitous signal during reward-guided decision tasks is a representation of *chosen value*<sup>34,35,39,49,133</sup>. This representation is isolated by correlating with some form of neural activity (firing rates, local field potential, fMRI) with the value of the option that will eventually be chosen on that trial. With time-varying measurements, this correlation can be repeated across many different timepoints, isolating the timepoint at which maximal variance is explained (green line, top panel).

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<sup>32,39,133</sup> show that chosen value representations occur *between* an initial representation of offer values, and the final representation of choices (top panel). 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 trials<sup>73</sup>. Crucially, chosen value influences decision speed. At a *fixed* timepoint in the middle of the decision process, the decision may have near completed on some trials (typically with high chosen value), whereas on others it may be a long way from completion. Any part of the brain that reflects the progression of this dynamical process will thus correlate with chosen value *mid*-decision (bottom panel). Notably, these dynamics have several possible neural substrates, including ramp-to-threshold accumulation in single neurons<sup>2</sup>, neural population trajectories through a low-dimensional manifold<sup>140</sup>, or bulk activity observed at the level of LFP or MEG signals<sup>53</sup>. In each case, correlates of chosen value would naturally emerge as a consequence of varying dynamics across trials. This idea has been used to explain the origin of a commonly observed ‘unchosen minus chosen value’ signal in dorsomedial prefrontal cortex using functional MRI<sup>52</sup>. Additionally, a recent study estimated the speed at which dynamics unfolded on a trial-by-trial basis<sup>32</sup>. The authors found this explained away 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.

## Figure Captions

**Figure 1: A distributed, hierarchical and recurrent approach to reward-guided choice.** Our account of reward-guided choice centres around three principles of brain organisation. First, neural networks in brain regions such as prefrontal cortex are highly recurrent in nature, which is essential for long-term stability of neural activity in temporally extended reward-guided behaviour. Second, reward-guided behaviour is supported by distributed competitions occurring in parallel across multiple brain regions. Third, these distributed and recurrent networks are organised hierarchically; within prefrontal cortex, more anterior regions support more temporally extended and abstract aspects of decision-making.

**Figure 2: Evidence for competition via mutual inhibition during reward-guided choice.** A biophysical attractor network model of decision-making<sup>11</sup> (centre) relies upon effective competition via mutual inhibition between ‘A units’ and ‘B units’, each selective for a different option. This has been used to explain: the anticorrelation of regression coefficients between different offers in single neuron data<sup>34,35</sup> (top left); the firing properties of different neurons found in orbitofrontal cortex<sup>36,49</sup> (top right); variation in value correlates as a function of time in human magnetoencephalography and macaque single unit data<sup>31,32</sup> (bottom right); and correlation across human subjects between excitation-inhibition balance, indexed using magnetic resonance spectroscopy, and choice consistency (softmax inverse temperature)<sup>33</sup> (bottom left).

**Figure 3: Hierarchical organisation of cortical timescales and 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, whilst at rest<sup>98</sup>. This reveals a hierarchical organisation to cortical timescales, with lower areas having rapidly changing activity, and higher areas having more persistent activity (bottom panel; joined dots/lines represent different datasets). Medial prefrontal cortex, in particular ACC, has the longest intrinsic timescale. (b) Persistent network activity, and network ‘resets’ during behavioural shifts, in rat medial prefrontal cortex in a reward-guided exploration task<sup>103</sup>. Each element of the matrix reflects the trial-to-trial similarity of neuronal ensemble activity. Lighter/darker elements reflect more similar/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) Rostro-caudal hierarchy of temporal organisation prefrontal cortex<sup>107</sup>. A similar hierarchy can be seen on medial and lateral surfaces for internal motivation (reward) and task difficulty, respectively. Whereas more anterior regions reflect block-wise changes in reward/difficulty, more posterior regions reflect trial-to-trial changes. MT = middle temporal area; LIP = lateral intraparietal cortex; LPFC = lateral prefrontal cortex; OFC = orbitofrontal cortex; ACC = anterior cingulate cortex; S1/S2 = primary/secondary somatosensory cortex.

- 1 Glimcher, P. & Fehr, E. *Neuroeconomics: Decision-making and the brain*. (Academic Press, 2014).
- 2 Gold, J. I. & Shadlen, M. N. The neural basis of decision making. *Annu Rev Neurosci* **30**, 535-574, doi:10.1146/annurev.neuro.29.051605.113038 (2007).
- 3 Dolan, Ray J. & Dayan, P. Goals and Habits in the Brain. *Neuron* **80**, 312-325, doi:10.1016/j.neuron.2013.09.007 (2013).
- 4 Rangel, A., Camerer, C. & Montague, P. R. A framework for studying the neurobiology of value-based decision making. *Nature reviews. Neuroscience* **9**, 545-556, doi:10.1038/nrn2357 (2008).
- 5 Kable, J. W. & Glimcher, P. W. The neurobiology of decision: consensus and controversy. *Neuron* **63**, 733-745, doi:10.1016/j.neuron.2009.09.003 (2009).
- 6 Rangel, A. & Hare, T. Neural computations associated with goal-directed choice. *Curr Opin Neurobiol* **20**, 262-270, doi:S0959-4388(10)00037-1 [pii] 10.1016/j.conb.2010.03.001 (2010).
- 7 Padoa-Schioppa, C. Neurobiology of Economic Choice: A Good-Based Model. *Annu Rev Neurosci* **34**, 333-359, doi:10.1146/annurev-neuro-061010-113648 (2011).
- 8 Frank, M. J. in *An Introduction to Model-Based Cognitive Neuroscience* 159-177 (Springer New York, 2015).
- 9 Wang, X. J. Neural dynamics and circuit mechanisms of decision-making. *Curr Opin Neurobiol* **22**, 1039-1046, doi:10.1016/j.conb.2012.08.006 (2012).
- 10 Pezzulo, G. & Cisek, P. Navigating the Affordance Landscape: Feedback Control as a Process Model of Behavior and Cognition. *Trends in Cognitive Sciences* **20**, 414-424, doi:10.1016/j.tics.2016.03.013 (2016).
- 11 Wang, X. J. Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* **36**, 955-968 (2002).
- 12 Rolls, E. T. *Emotion and decision making explained*. (Oxford University Press, 2013).
- 13 Machens, C. K., Romo, R. & Brody, C. D. Flexible control of mutual inhibition: a neural model of two-interval discrimination. *Science* **307**, 1121-1124, doi:10.1126/science.1104171 (2005).
- 14 Krajbich, I., Armel, C. & Rangel, A. Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience* **13**, 1292-1298, doi:10.1038/nn.2635 (2010).
- 15 Medler, D. A. A brief history of connectionism. *Neural Computing Surveys* **1**, 18-72 (1998).
- 16 Gardner, H. *The mind's new science : a history of the cognitive revolution*. (Basic Books, 1985).
- 17 McClelland, J. L. & Rumelhart, D. E. *Explorations in parallel distributed processing: A handbook of models, programs, and exercises*. (MIT press, 1989).
- 18 Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. *Nature* **323**, 533-536, doi:10.1038/323533a0 (1986).

- 19 Hopfield, J. J. Neural networks and physical systems with emergent collective computational abilities. *Proc Natl Acad Sci U S A* **79**, 2554-2558 (1982).
- 20 Carpenter, G. A. & Grossberg, S. A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics, and Image Processing* **37**, 54-115, doi:10.1016/s0734-189x(87)80014-2 (1987).
- 21 LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436-444, doi:10.1038/nature14539 (2015).
- 22 Jbabdi, S., Sotiropoulos, S. N., Haber, S. N., Van Essen, D. C. & Behrens, T. E. Measuring macroscopic brain connections in vivo. *Nature Neuroscience* **18**, 1546-1555, doi:10.1038/nn.4134 (2015).
- 23 Fenno, L., Yizhar, O. & Deisseroth, K. The Development and Application of Optogenetics. *Annual Review of Neuroscience* **34**, 389-412, doi:10.1146/annurev-neuro-061010-113817 (2011).
- 24 Churchland, P. M. Eliminative Materialism and the Propositional Attitudes. *The Journal of Philosophy* **78**, 67, doi:10.2307/2025900 (1981).
- 25 Seeley, T. D. *Honeybee democracy*. (Princeton University Press, 2010).
- 26 London, M. & Häusser, M. Dendritic Computation. *Annual Review of Neuroscience* **28**, 503-532, doi:10.1146/annurev.neuro.28.061604.135703 (2005).
- 27 Couzin, I. D. Collective cognition in animal groups. *Trends Cogn Sci* **13**, 36-43, doi:10.1016/j.tics.2008.10.002 (2009).
- 28 O'Doherty, J. P. The problem with value. *Neuroscience and biobehavioral reviews* **43**, 259-268, doi:10.1016/j.neubiorev.2014.03.027 (2014).
- 29 Grillner, S. Neurobiological bases of rhythmic motor acts in vertebrates. *Science* **228**, 143-149, doi:10.1126/science.3975635 (1985).
- 30 Kovac, M. & Davis, W. Behavioral choice: neural mechanisms in Pleurobranchaea. *Science* **198**, 632-634, doi:10.1126/science.918659 (1977).
- 31 Hunt, L. T. *et al.* Mechanisms underlying cortical activity during value-guided choice. *Nat Neurosci* **15**, 470-476, doi:10.1038/nn.3017 (2012).
- 32 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**, doi:10.7554/eLife.11945 (2015).
- 33 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, doi:10.1038/nn.3140 (2012).
- 34 Strait, C. E., Blanchard, T. C. & Hayden, B. Y. Reward value comparison via mutual inhibition in ventromedial prefrontal cortex. *Neuron* **82**, 1357-1366, doi:10.1016/j.neuron.2014.04.032 (2014).
- 35 Strait, C. E., Sleezer, B. J. & Hayden, B. Y. Signatures of Value Comparison in Ventral Striatum Neurons. *PLoS biology* **13**, e1002173, doi:10.1371/journal.pbio.1002173 (2015).
- 36 Rustichini, A. & Padoa-Schioppa, C. A Neuro-Computational Model of Economic Decisions. *Journal of neurophysiology*, jn 00184 02015, doi:10.1152/jn.00184.2015 (2015).

- 37 Hunt, L. T., Dolan, R. J. & Behrens, T. E. Hierarchical competitions subserving multi-attribute choice. *Nat Neurosci* **17**, 1613-1622, doi:10.1038/nn.3836 (2014).
- 38 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, doi:10.1038/nn.3649 (2014).
- 39 Louie, K. & Glimcher, P. W. Separating Value from Choice: Delay Discounting Activity in the Lateral Intraparietal Area. *Journal of Neuroscience* **30**, 5498-5507, doi:10.1523/jneurosci.5742-09.2010 (2010).
- 40 Pastor-Bernier, A. & Cisek, P. Neural correlates of biased competition in premotor cortex. *J Neurosci* **31**, 7083-7088, doi:10.1523/JNEUROSCI.5681-10.2011 (2011).
- 41 Cisek, P. Integrated neural processes for defining potential actions and deciding between them: a computational model. *J Neurosci* **26**, 9761-9770, doi:26/38/9761 [pii] 10.1523/JNEUROSCI.5605-05.2006 (2006).
- 42 Soltani, A. A Biophysically Based Neural Model of Matching Law Behavior: Melioration by Stochastic Synapses. *Journal of Neuroscience* **26**, 3731-3744, doi:10.1523/jneurosci.5159-05.2006 (2006).
- 43 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, doi:10.1016/j.conb.2012.04.011 (2012).
- 44 Cisek, P. Making decisions through a distributed consensus. *Curr Opin Neurobiol* **22**, 927-936, doi:10.1016/j.conb.2012.05.007 (2012).
- 45 Strait, C. E. *et al.* Neuronal selectivity for spatial position of offers and choices in five reward regions. *Journal of Neurophysiology*, jn.00325.02015, doi:10.1152/jn.00325.2015 (2015).
- 46 Hunt, L. T., Woolrich, M. W., Rushworth, M. F. & Behrens, T. E. Trial-type dependent frames of reference for value comparison. *PLoS computational biology* **9**, e1003225, doi:10.1371/journal.pcbi.1003225 (2013).
- 47 Passingham, R. E., Stephan, K. E. & Kötter, R. The anatomical basis of functional localization in the cortex. *Nature Reviews Neuroscience* **3**, 606-616, doi:10.1038/nrn893 (2002).
- 48 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 U S A* **112**, E2695-2704, doi:10.1073/pnas.1410767112 (2015).
- 49 Padoa-Schioppa, C. & Assad, J. A. Neurons in the orbitofrontal cortex encode economic value. *Nature* **441**, 223-226, doi:10.1038/nature04676 (2006).
- 50 Hosokawa, T., Kennerley, S. W., Sloan, J. & Wallis, J. D. Single-Neuron Mechanisms Underlying Cost-Benefit Analysis in Frontal Cortex. *Journal of Neuroscience* **33**, 17385-17397, doi:10.1523/jneurosci.2221-13.2013 (2013).
- 51 Rolls, E. T., Grabenhorst, F. & Deco, G. Choice, difficulty, and confidence in the brain. *NeuroImage* **53**, 694-706, doi:10.1016/j.neuroimage.2010.06.073 (2010).

- 52 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. *Proceedings of the National Academy of Sciences* **108**, 18120-18125, doi:10.1073/pnas.1109322108 (2011).
- 53 Hunt, L. T., Behrens, T. E., Hosokawa, T., Wallis, J. D. & Kennerley, S. W. Capturing the temporal evolution of choice across prefrontal cortex. *eLife*, e11945 (2015).
- 54 Siegel, M., Buschman, T. J. & Miller, E. K. Cortical information flow during flexible sensorimotor decisions. *Science* **348**, 1352-1355, doi:10.1126/science.aab0551 (2015).
- 55 Klein-Flugge, 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, doi:10.1523/JNEUROSCI.0270-12.2012 (2012).
- 56 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. *Social Cognitive and Affective Neuroscience* **7**, 604-609, doi:10.1093/scan/nss055 (2012).
- 57 Niv, Y. *et al.* Reinforcement Learning in Multidimensional Environments Relies on Attention Mechanisms. *Journal of Neuroscience* **35**, 8145-8157, doi:10.1523/jneurosci.2978-14.2015 (2015).
- 58 Freidin, E. & Kacelnik, A. Rational Choice, Context Dependence, and the Value of Information in European Starlings (*Sturnus vulgaris*). *Science* **334**, 1000-1002, doi:10.1126/science.1209626 (2011).
- 59 Bettman, James R., Luce, Mary F. & Payne, John W. Constructive Consumer Choice Processes. *Journal of Consumer Research* **25**, 187-217, doi:10.1086/209535 (1998).
- 60 Krajbich, I., Armel, C. & Rangel, A. Visual fixations and the computation and comparison of value in simple choice. *Nat Neurosci* **13**, 1292-1298, doi:10.1038/nn.2635 (2010).
- 61 Rich, E. L. & Wallis, J. D. Decoding subjective decisions from orbitofrontal cortex. *Nature Neuroscience* **19**, 973-980, doi:10.1038/nn.4320 (2016).
- 62 LeCun, Y. *et al.* Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Computation* **1**, 541-551, doi:10.1162/neco.1989.1.4.541 (1989).
- 63 Krizhevsky, A., Sutskever, I. & Hinton, G. ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems* **25**, 1090-1098 (2012).
- 64 Yamins, D. L. *et al.* Performance-optimized hierarchical models predict neural responses in higher visual cortex. *Proc Natl Acad Sci U S A* **111**, 8619-8624, doi:10.1073/pnas.1403112111 (2014).
- 65 Kriegeskorte, N. Deep Neural Networks: A New Framework for Modeling Biological Vision and Brain Information Processing. *Annual Review of Vision Science* **1**, 417-446, doi:10.1146/annurev-vision-082114-035447 (2015).
- 66 Graves, A. *Supervised sequence labelling with recurrent neural networks*. (Springer, 2012).



- 67 Graves, A., Mohamed, A.-r. & Hinton, G. Speech Recognition with Deep Recurrent Neural Networks. *ArXiv e-prints* **1303** (2013).  
<<http://adsabs.harvard.edu/abs/2013arXiv1303.5778G>>.
- 68 Weston, J., Chopra, S. & Bordes, A. Memory Networks. *ArXiv e-prints* **1410** (2014). <<http://adsabs.harvard.edu/abs/2014arXiv1410.3916W>>.
- 69 Amari, S.-i. Dynamics of pattern formation in lateral-inhibition type neural fields. *Biological Cybernetics* **27**, 77-87, doi:10.1007/bf00337259 (1977).
- 70 Compte, A. Synaptic Mechanisms and Network Dynamics Underlying Spatial Working Memory in a Cortical Network Model. *Cerebral Cortex* **10**, 910-923, doi:10.1093/cercor/10.9.910 (2000).
- 71 Wang, X.-J. in *Principles of frontal lobe function* (eds D.T. Stuss & R.T. Knight) Ch. 15, 226-248 (Oxford University Press, 2013).
- 72 Elston, G. N. Pyramidal cells of the frontal lobe: all the more spinous to think with. *J Neurosci* **20**, RC95 (2000).
- 73 Busemeyer, J. R. & Townsend, J. T. Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review* **100**, 432-459, doi:10.1037/0033-295x.100.3.432 (1993).
- 74 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. *Psychological review* **113**, 700-765, doi:10.1037/0033-295X.113.4.700 (2006).
- 75 Passingham, R. E. & Wise, S. P. *The neurobiology of the prefrontal cortex : anatomy, evolution, and the origin of insight*. 1st edn, (Oxford University Press, 2012).
- 76 Stephens, D. W. & Krebs, J. R. *Foraging theory*. (Princeton University Press, 1986).
- 77 Kolling, N., Behrens, T. E., Mars, R. B. & Rushworth, M. F. Neural mechanisms of foraging. *Science* **336**, 95-98, doi:10.1126/science.1216930 (2012).
- 78 Hayden, B. Y., Pearson, J. M. & Platt, M. L. Neuronal basis of sequential foraging decisions in a patchy environment. *Nat Neurosci* **14**, 933-939, doi:10.1038/nn.2856 (2011).
- 79 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, doi:10.1523/JNEUROSCI.1246-11.2011 (2011).
- 80 Rich, E. L. & Wallis, J. D. Decoding subjective decisions from orbitofrontal cortex. *Nat Neurosci* **19**, 973-980, doi:10.1038/nn.4320 (2016).
- 81 Harvey, C. D., Coen, P. & Tank, D. W. Choice-specific sequences in parietal cortex during a virtual-navigation decision task. *Nature* **484**, 62-68, doi:10.1038/nature10918 (2012).
- 82 Stokes, M. G. 'Activity-silent' working memory in prefrontal cortex: a dynamic coding framework. *Trends Cogn Sci* **19**, 394-405, doi:10.1016/j.tics.2015.05.004 (2015).
- 83 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**, doi:10.7554/eLife.18937 (2016).

- 84 Sussillo, D. & Abbott, L. F. Generating coherent patterns of activity from chaotic neural networks. *Neuron* **63**, 544-557, doi:10.1016/j.neuron.2009.07.018 (2009).
- 85 Hennequin, G., Vogels, Tim P. & Gerstner, W. Optimal Control of Transient Dynamics in Balanced Networks Supports Generation of Complex Movements. *Neuron* **82**, 1394-1406, doi:10.1016/j.neuron.2014.04.045 (2014).
- 86 Sussillo, D. & Barak, O. Opening the Black Box: Low-Dimensional Dynamics in High-Dimensional Recurrent Neural Networks. *Neural Computation* **25**, 626-649, doi:10.1162/NECO\_a\_00409 (2013).
- 87 Mante, V., Sussillo, D., Shenoy, K. V. & Newsome, W. T. Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* **503**, 78-84, doi:10.1038/nature12742 (2013).
- 88 Song, H. F., Yang, G. R. & Wang, X. Reward-based training of recurrent neural networks for cognitive and value-based tasks. *bioRxiv*, doi:10.1101/070375 (2016).
- 89 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, doi:10.1016/j.neuron.2014.08.031 (2014).
- 90 Shallice, T. & Burgess, P. W. Deficits in strategy application following frontal lobe damage in man. *Brain* **114** ( Pt 2), 727-741 (1991).
- 91 Fuster, J. M. The prefrontal cortex--an update: time is of the essence. *Neuron* **30**, 319-333 (2001).
- 92 Holroyd, C. B. & Yeung, N. Motivation of extended behaviors by anterior cingulate cortex. *Trends Cogn Sci* **16**, 122-128, doi:10.1016/j.tics.2011.12.008 (2012).
- 93 Pascanu, R., Gulcehre, C., Cho, K. & Bengio, Y. How to Construct Deep Recurrent Neural Networks. *ArXiv e-prints* **1312** (2013). <<http://adsabs.harvard.edu/abs/2013arXiv1312.6026P>>.
- 94 Karpathy, A. & Fei-Fei, L. Deep Visual-Semantic Alignments for Generating Image Descriptions. (2015).
- 95 Bernacchia, A., Seo, H., Lee, D. & Wang, X.-J. A reservoir of time constants for memory traces in cortical neurons. *Nature Neuroscience* **14**, 366-372, doi:10.1038/nn.2752 (2011).
- 96 Behrens, T. E. J., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. S. Learning the value of information in an uncertain world. *Nature Neuroscience* **10**, 1214-1221, doi:10.1038/nn1954 (2007).
- 97 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, doi:10.1016/j.neuron.2015.09.008 (2015).
- 98 Murray, J. D. *et al.* A hierarchy of intrinsic timescales across primate cortex. *Nature Neuroscience* **17**, 1661-1663, doi:10.1038/nn.3862 (2014).
- 99 Purcell, B. A. & Kiani, R. Hierarchical decision processes that operate over distinct timescales underlie choice and changes in strategy. *Proc Natl Acad Sci U S A* **113**, E4531-4540, doi:10.1073/pnas.1524685113 (2016).
- 100 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, doi:10.1523/JNEUROSCI.3022-12.2013 (2013).
- 101 Wittmann, M. K. *et al.* Predictive decision making driven by multiple time-linked reward representations in the anterior cingulate cortex. *Nature Communications* **7**, 12327, doi:10.1038/ncomms12327 (2016).
- 102 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, doi:10.1038/nature04766 (2006).
- 103 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, doi:10.1126/science.1226518 (2012).
- 104 Blanchard, T. C., Strait, C. E. & Hayden, B. Y. Ramping ensemble activity in dorsal anterior cingulate neurons during persistent commitment to a decision. *Journal of Neurophysiology* **114**, 2439-2449, doi:10.1152/jn.00711.2015 (2015).
- 105 Desrochers, Theresa M., Chatham, Christopher H. & Badre, D. The Necessity of Rostrolateral Prefrontal Cortex for Higher-Level Sequential Behavior. *Neuron* **87**, 1357-1368, doi:10.1016/j.neuron.2015.08.026 (2015).
- 106 Badre, D. & D'Esposito, M. Is the rostro-caudal axis of the frontal lobe hierarchical? *Nature reviews. Neuroscience* **10**, 659-669, doi:10.1038/nrn2667 (2009).
- 107 Kouneiher, F., Charron, S. & Koechlin, E. Motivation and cognitive control in the human prefrontal cortex. *Nat Neurosci* **12**, 939-945, doi:10.1038/nn.2321 (2009).
- 108 O'Reilly, R. C. The What and How of prefrontal cortical organization. *Trends in Neurosciences* **33**, 355-361, doi:10.1016/j.tins.2010.05.002 (2010).
- 109 Zarr, N. & Brown, J. W. Hierarchical error representation in medial prefrontal cortex. *Neuroimage* **124**, 238-247, doi:10.1016/j.neuroimage.2015.08.063 (2016).
- 110 Nee, D. E. & Brown, J. W. Rostral-caudal gradients of abstraction revealed by multi-variate pattern analysis of working memory. *NeuroImage* **63**, 1285-1294, doi:10.1016/j.neuroimage.2012.08.034 (2012).
- 111 Nee, D. E. & D'Esposito, M. The hierarchical organization of the lateral prefrontal cortex. *eLife* **5**, doi:10.7554/eLife.12112 (2016).
- 112 Badre, D., Kayser, A. S. & D'Esposito, M. Frontal cortex and the discovery of abstract action rules. *Neuron* **66**, 315-326, doi:10.1016/j.neuron.2010.03.025 (2010).
- 113 Frank, M. J. & Badre, D. Mechanisms of hierarchical reinforcement learning in corticostriatal circuits 1: computational analysis. *Cereb Cortex* **22**, 509-526, doi:10.1093/cercor/bhr114 (2012).
- 114 Hochreiter, S. & Schmidhuber, J. Long Short-Term Memory. *Neural Computation* **9**, 1735-1780, doi:10.1162/neco.1997.9.8.1735 (1997).
- 115 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, doi:10.1162/089976606775093909 (2006).

- 116 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. *Journal of neurophysiology* **57**, 1262-1287 (1987).
- 117 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).
- 118 Noonan, M. P. *et al.* Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. *Proc Natl Acad Sci U S A* **107**, 20547-20552, doi:10.1073/pnas.1012246107 (2010).
- 119 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, doi:10.1016/j.neuron.2010.02.027 (2010).
- 120 Vaidya, A. R. & Fellows, L. K. Testing necessary regional frontal contributions to value assessment and fixation-based updating. *Nature Communications* **6**, 10120, doi:10.1038/ncomms10120 (2015).
- 121 Eslinger, P. J. & Damasio, A. R. Severe disturbance of higher cognition after bilateral frontal lobe ablation: patient EVR. *Neurology* **35**, 1731-1741 (1985).
- 122 Szczepanski, Sara M. & Knight, Robert T. Insights into Human Behavior from Lesions to the Prefrontal Cortex. *Neuron* **83**, 1002-1018, doi:10.1016/j.neuron.2014.08.011 (2014).
- 123 Shadlen, Michael N. & Shohamy, D. Decision Making and Sequential Sampling from Memory. *Neuron* **90**, 927-939, doi:10.1016/j.neuron.2016.04.036 (2016).
- 124 Plassmann, H., O'Doherty, J. & Rangel, A. Orbitofrontal cortex encodes willingness to pay in everyday economic transactions. *J Neurosci* **27**, 9984-9988, doi:10.1523/JNEUROSCI.2131-07.2007 (2007).
- 125 Knutson, B., Taylor, J., Kaufman, M., Peterson, R. & Glover, G. Distributed neural representation of expected value. *J Neurosci* **25**, 4806-4812, doi:10.1523/JNEUROSCI.0642-05.2005 (2005).
- 126 McCoy, A. N. & Platt, M. L. Risk-sensitive neurons in macaque posterior cingulate cortex. *Nature neuroscience* **8**, 1220-1227, doi:10.1038/nn1523 (2005).
- 127 Kim, S., Hwang, J. & Lee, D. Prefrontal coding of temporally discounted values during intertemporal choice. *Neuron* **59**, 161-172, doi:10.1016/j.neuron.2008.05.010 (2008).
- 128 Louie, K., Gratton, L. E. & Glimcher, P. W. Reward value-based gain control: divisive normalization in parietal cortex. *J Neurosci* **31**, 10627-10639, doi:10.1523/JNEUROSCI.1237-11.2011 (2011).
- 129 Shuler, M. G. Reward Timing in the Primary Visual Cortex. *Science* **311**, 1606-1609, doi:10.1126/science.1123513 (2006).
- 130 Pastor-Bernier, A. & Cisek, P. Neural Correlates of Biased Competition in Premotor Cortex. *Journal of Neuroscience* **31**, 7083-7088, doi:10.1523/jneurosci.5681-10.2011 (2011).
- 131 Philiastides, M. G., Biele, G. & Heekeren, H. R. A mechanistic account of value computation in the human brain. *Proceedings of the National*

- Academy of Sciences* **107**, 9430-9435, doi:10.1073/pnas.1001732107 (2010).
- 132 Peck, C. J., Lau, B. & Salzman, C. D. The primate amygdala combines information about space and value. *Nature Neuroscience* **16**, 340-348, doi:10.1038/nn.3328 (2013).
- 133 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, doi:10.1016/j.neuron.2010.11.041 (2011).
- 134 Vickery, Timothy J., Chun, Marvin M. & Lee, D. Ubiquity and Specificity of Reinforcement Signals throughout the Human Brain. *Neuron* **72**, 166-177, doi:10.1016/j.neuron.2011.08.011 (2011).
- 135 Maunsell, J. H. R. Neuronal representations of cognitive state: reward or attention? *Trends in cognitive sciences* **8**, 261-265, doi:10.1016/j.tics.2004.04.003 (2004).
- 136 Leathers, M. L. & Olson, C. R. In monkeys making value-based decisions, LIP neurons encode cue salience and not action value. *Science (New York, NY)* **338**, 132-135, doi:10.1126/science.1226405 (2012).
- 137 Litt, A., Plassmann, H., Shiv, B. & Rangel, A. Dissociating Valuation and Saliency Signals during Decision-Making. *Cerebral Cortex* **21**, 95-102, doi:10.1093/cercor/bhq065 (2010).
- 138 Vlaev, I., Chater, N., Stewart, N. & Brown, G. D. Does the brain calculate value? *Trends in cognitive sciences* **15**, 546-554 (2011).
- 139 Lichtenstein, S. & Slovic, P. Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology* **89**, 46-55, doi:10.1037/h0031207 (1971).
- 140 Churchland, M. M. *et al.* Neural population dynamics during reaching. *Nature* **487**, 51-56, doi:10.1038/nature11129 (2012).
- 141 Shenoy, K. V., Sahani, M. & Churchland, M. M. Cortical control of arm movements: a dynamical systems perspective. *Annu Rev Neurosci* **36**, 337-359, doi:10.1146/annurev-neuro-062111-150509 (2013).
- 142 Hunt, L. T. What are the neural origins of choice variability? *Trends in cognitive sciences* **18**, 222-224, doi:10.1016/j.tics.2014.01.004 (2014).
- 143 Levy, D. J. & Glimcher, P. W. The root of all value: a neural common currency for choice. *Curr Opin Neurobiol* **22**, 1027-1038, doi:10.1016/j.conb.2012.06.001 (2012).
- 144 Svoboda, E., McKinnon, M. C. & Levine, B. The functional neuroanatomy of autobiographical memory: A meta-analysis. *Neuropsychologia* **44**, 2189-2208, doi:10.1016/j.neuropsychologia.2006.05.023 (2006).
- 145 Doeller, C. F., Barry, C. & Burgess, N. Evidence for grid cells in a human memory network. *Nature* **463**, 657-661, doi:10.1038/nature08704 (2010).
- 146 Schacter, Daniel L. *et al.* The Future of Memory: Remembering, Imagining, and the Brain. *Neuron* **76**, 677-694, doi:10.1016/j.neuron.2012.11.001 (2012).
- 147 Van Overwalle, F. Social cognition and the brain: A meta-analysis. *Human brain mapping* **30**, 829-858, doi:10.1002/hbm.20547 (2009).
- 148 Lopatina, N. *et al.* Lateral orbitofrontal neurons acquire responses to upshifted, downshifted, or blocked cues during unblocking. *eLife* **4**, doi:10.7554/eLife.11299 (2015).

- 149 McDannald, M. A. *et al.* Orbitofrontal neurons acquire responses to 'valueless' Pavlovian cues during unblocking. *eLife* **3**, doi:10.7554/eLife.02653 (2014).
- 150 Tsujimoto, S., Genovesio, A. & Wise, S. P. Neuronal Activity during a Cued Strategy Task: Comparison of Dorsolateral, Orbital, and Polar Prefrontal Cortex. *Journal of Neuroscience* **32**, 11017-11031, doi:10.1523/jneurosci.1230-12.2012 (2012).
- 151 Kepecs, A., Uchida, N., Zariwala, H. A. & Mainen, Z. F. Neural correlates, computation and behavioural impact of decision confidence. *Nature* **455**, 227-231, doi:10.1038/nature07200 (2008).
- 152 Strait, C. E. *et al.* Neuronal selectivity for spatial positions of offers and choices in five reward regions. *Journal of neurophysiology* **115**, 1098-1111, doi:10.1152/jn.00325.2015 (2016).
- 153 Genovesio, A., Tsujimoto, S., Navarra, G., Falcone, R. & Wise, S. P. Autonomous Encoding of Irrelevant Goals and Outcomes by Prefrontal Cortex Neurons. *Journal of Neuroscience* **34**, 1970-1978, doi:10.1523/jneurosci.3228-13.2014 (2014).
- 154 Rigotti, M. *et al.* The importance of mixed selectivity in complex cognitive tasks. *Nature* **497**, 585-590, doi:10.1038/nature12160 (2013).
- 155 Raposo, D., Kaufman, M. T. & Churchland, A. K. A category-free neural population supports evolving demands during decision-making. *Nature Neuroscience* **17**, 1784-1792, doi:10.1038/nn.3865 (2014).
- 156 Felleman, D. J. & Van Essen, D. C. Distributed Hierarchical Processing in the Primate Cerebral Cortex. *Cerebral Cortex* **1**, 1-47, doi:10.1093/cercor/1.1.1 (1991).
- 157 Markov, N. T. *et al.* A Weighted and Directed Interareal Connectivity Matrix for Macaque Cerebral Cortex. *Cerebral Cortex* **24**, 17-36, doi:10.1093/cercor/bhs270 (2012).
- 158 Dombrowski, S. M., Hilgetag, C. C. & Barbas, H. Quantitative architecture distinguishes prefrontal cortical systems in the rhesus monkey. *Cereb Cortex* **11**, 975-988 (2001).
- 159 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, doi:10.1146/annurev.ne.09.030186.002041 (1986).
- 160 McClelland, J. L. Emergence in Cognitive Science. *Topics in Cognitive Science* **2**, 751-770, doi:10.1111/j.1756-8765.2010.01116.x (2010).
- 161 Seeley, T. D. & Buhrman, S. C. Group decision making in swarms of honey bees. *Behavioral Ecology and Sociobiology* **45**, 19-31, doi:10.1007/s002650050536 (1999).
- 162 Seeley, T. D. *et al.* Stop Signals Provide Cross Inhibition in Collective Decision-Making by Honeybee Swarms. *Science* **335**, 108-111, doi:10.1126/science.1210361 (2011).

## Acknowledgments

We are grateful to many colleagues with whom interactions have shaped ideas within this review, in particular Tim Behrens, Tommy Blanchard, Steve Kennerley, John Pearson, Steve Piantadosi, Michael Platt, Matthew Rushworth and Tom Seeley. We thank Helen Barron, Archy de Berker, Ray Dolan, MaryAnn Noonan and Tricia 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 an NIH award to B.Y.H. (DA037229).

Figure 1

**Recurrent** neural networks support naturalistic, temporally extended reward-guided behaviours such as foraging

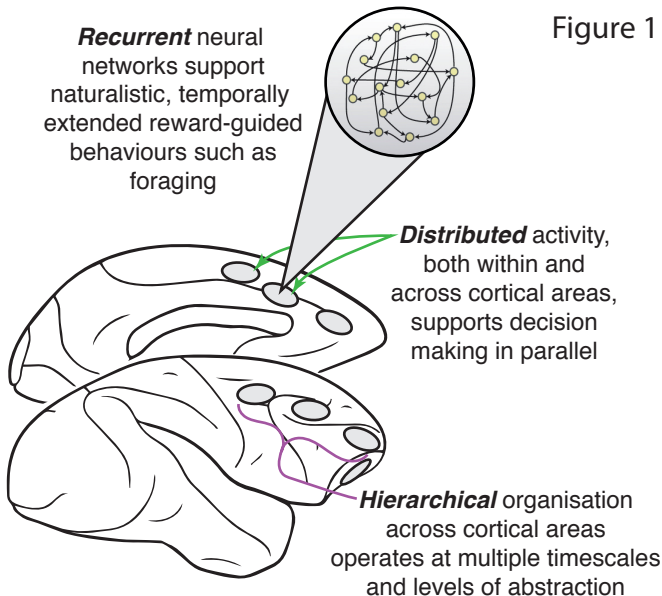
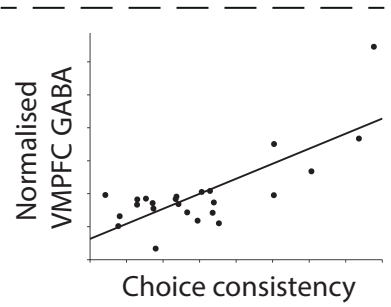
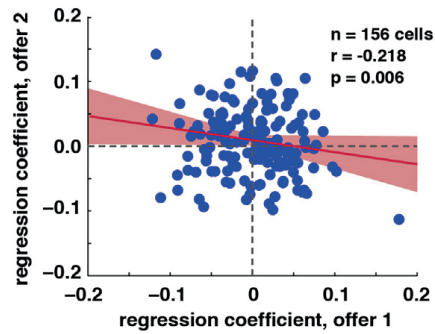




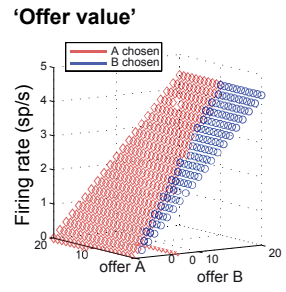
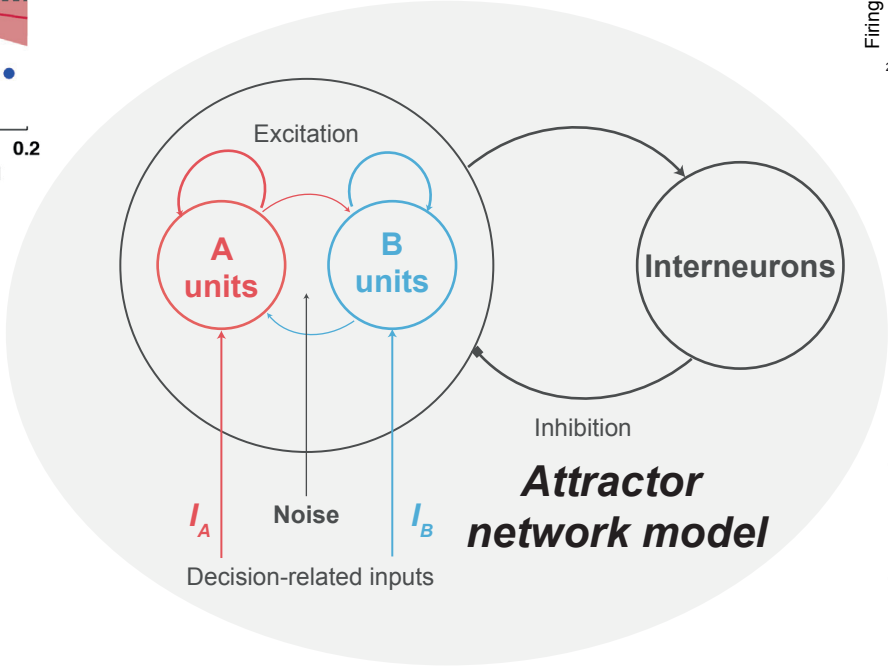
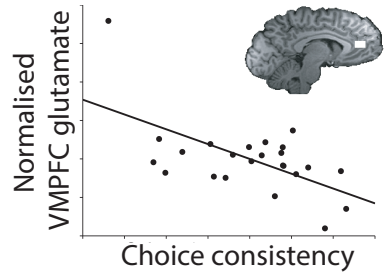
Figure 2

# **Anticorrelated regression coefficients during offer presentation**

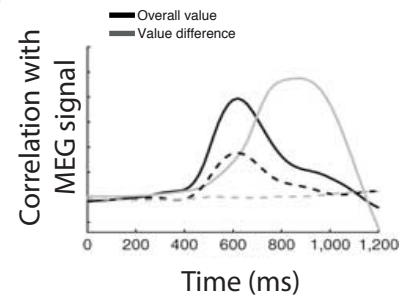
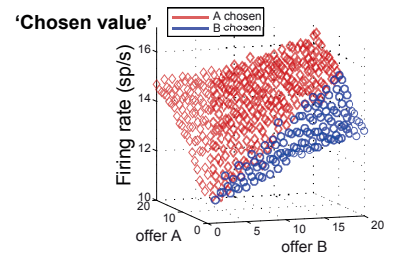
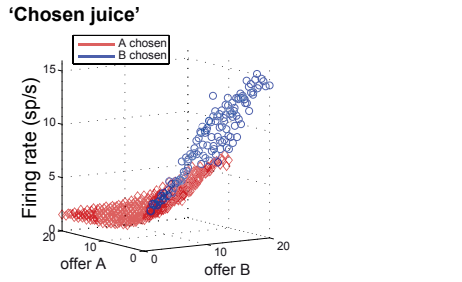
Strait et al., 2014, 2015



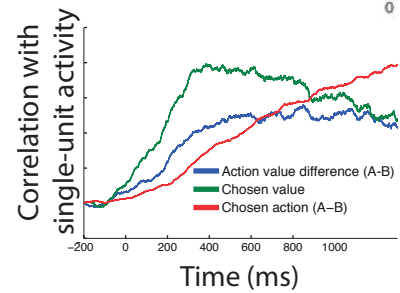
**Relationship between excitation-inhibition balance and choice behaviour**  
 Jochem et al., 2012



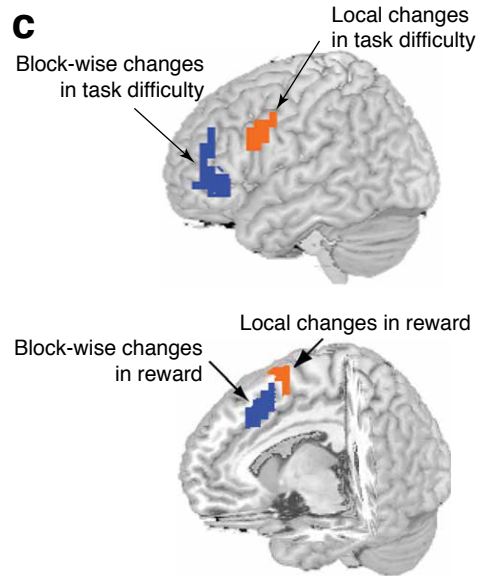
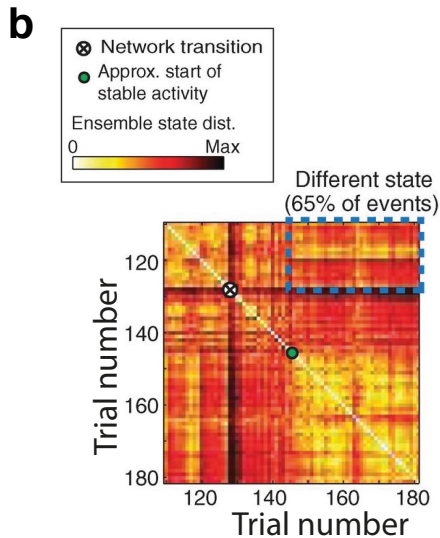
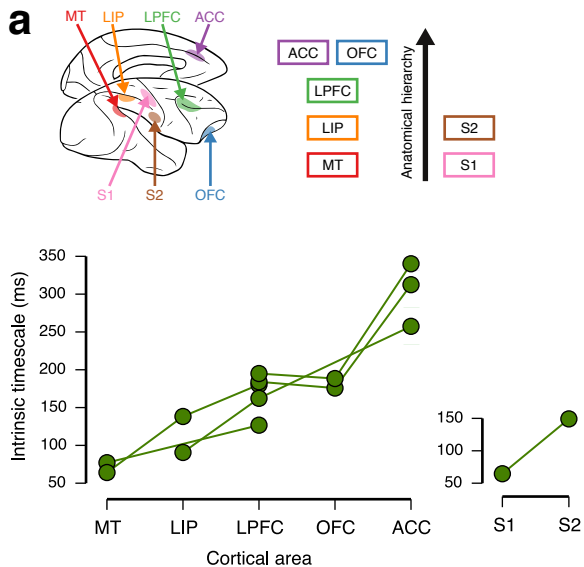
# **Different neurons in orbitofrontal cortex** Rustichini and Padoa-Schioppa, 2015

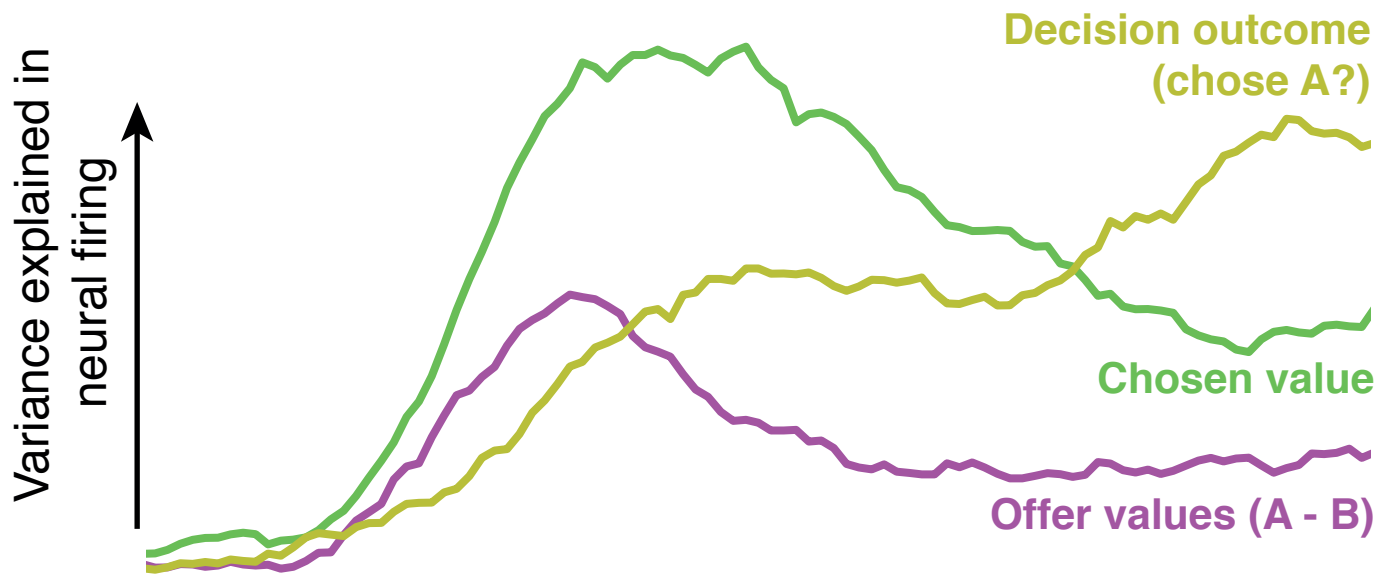


**Time-varying correlates of value**  
 Hunt et al., 2012;  
 Hunt et al., 2015



# Figure 3





Strongest *chosen value* correlates across trials occur here: mid-decision

