

A neuronal theory of sequential economic choice

Brain and Neuroscience Advances
Volume 2: 1–15
© The Author(s) 2018
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/2398212818766675
journals.sagepub.com/home/bna



Benjamin Y. Hayden¹  and Rubén Moreno-Bote^{2,3,4}

Abstract

Results of recent studies point towards a new framework for the neural bases of economic choice. The principles of this framework include the idea that evaluation is limited to a single option within the focus of attention and that we accept or reject that option relative to the entire set of alternatives. Rejection leads attention to a new option, although it can later switch back to a previously rejected one. The option to which a neuron's firing rate refers is determined dynamically by attention and not stably by labelled lines. Value is always computed relative to the value of rejection. Comparison results not from explicit competition between discrete populations of neurons, but indirectly, as in a horse race, from the fact that the first option whose value crosses a threshold is selected. Consequently, comparison can occur within a single pool of neurons rather than by competition between two or more neuronal populations. The computations that constitute comparison thus occur at multiple levels, including premotor levels, simultaneously (i.e. the brain uses a distributed consensus), and not in discrete stages. This framework suggests a solution to a set of otherwise unresolved neuronal binding problems that result from the need to link options to values, comparisons to actions, and choices to outcomes.

Keywords

Neuroeconomics, mutual inhibition, repetition suppression, labelled line

Received: 15 January 2018; accepted: 27 February 2018

Introduction

Economic choice, the selection of options based on their value, is a core process in the repertoire of intelligent organisms (Hayden, 2018; Pearson et al., 2014; Rangel and Hare, 2010; Rushworth et al., 2011). Neuroeconomic research has successfully identified some of the major brain regions associated with valuation and choice, especially the orbitofrontal cortex (OFC), ventromedial prefrontal cortex (vmPFC), dorsal anterior cingulate cortex (dACC), and ventral striatum (VS). These regions are activated by the values of offers and outcomes and show correlations with comparison-related processes (Bartra et al., 2013; Ebitz and Hayden, 2016; Haber and Behrens, 2014; Heilbronner and Hayden, 2016; Rushworth et al., 2011; Wallis, 2007). Lesion studies support the idea that these regions have a direct causative role in choice (e.g. Camille et al., 2011; Kennerley et al., 2006; Noonan et al., 2010). All measures point to some degree of specialisation within these regions, although their respective roles remain debated (Rushworth et al., 2011). While the locations of value-related processing are now established, the mechanisms of choice are not. Nonetheless, we believe that a series of recent studies have begun to limn something of a consensus view – if not a model, at least a framework for one.

A common proposal is that value comparison is implemented by direct competition, via mutual inhibition, between discrete sets of neurons whose responses correspond to the value of particular options (e.g. Chau et al., 2014; Hunt et al., 2012, 2015;

Louie et al., 2011; Padoa-Schioppa, 2011; Rustichini and Padoa-Schioppa, 2015; Soltani et al., 2006). From this perspective, value representations are aligned to neuron identity by a *labelled line code*: a neuron's firing rate indicates a value and its identity (its notional label) indicates which option has that value. This stable relationship makes implementing choice straightforward: the two populations compete for control of a third set of neurons and whichever set of neurons wins the competition determines the chosen option. However, this approach introduces several problems. First, it necessitates redundant reduplication of

¹Department of Neuroscience and Center for Magnetic Resonance Research, University of Minnesota, Minneapolis, MN, USA

²Department of Information and Communications Technologies, Pompeu Fabra University, Barcelona, Spain

³Center for Brain and Cognition, Pompeu Fabra University, Barcelona, Spain

⁴Serra Hünter Fellow Programme, Pompeu Fabra University, Barcelona, Spain

Corresponding author:

Benjamin Y. Hayden, Department of Neuroscience and Center for Magnetic Resonance Research, University of Minnesota, Minneapolis, MN 55455, USA.

Email: benhayden@gmail.com



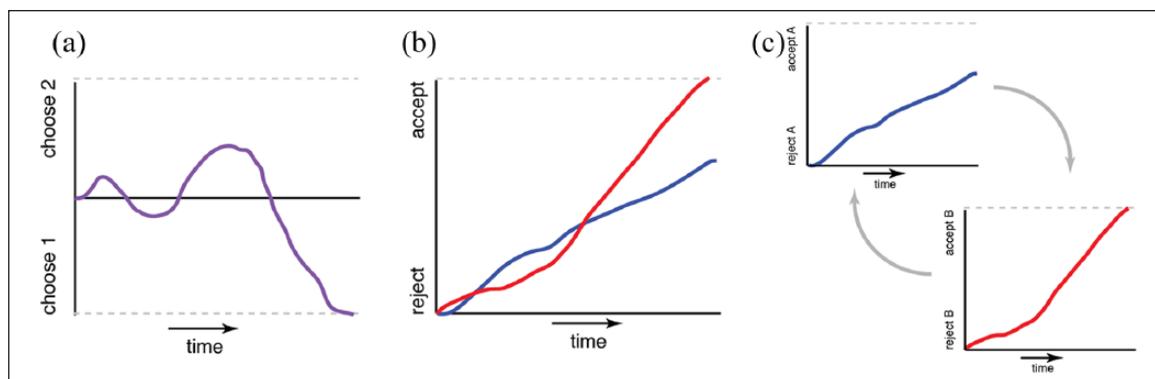


Figure 1. Illustration of basic framework of models. (a) In simultaneous choice ('tug of war'), there is one decision variable that drifts between two bounds, corresponding to choice of options 1 and 2, respectively. (b) In independent choice ('sequential choice' or horse race) models, there are two decision variables drifting between two potentially similar sets of bounds. They may interact or they may not; generally one threshold hit stops the deliberation process and the second one is halted. (c) Some recent evidence is consistent with a pair of single diffusion processes, only one of which occurs at a time, as determined by the focus of attention. The threshold in the single diffusion process is likely influenced by the background value (estimate of the value of rejecting), which in turn can be determined by the value of the other option and by the value of further exploration.

circuitry for computing value for each offer. Second, it requires precise wiring to implement it (or else a well-informed supervisory system that dynamically creates that wiring.) Third, it does not readily scale up to situations with many offers (such as choosing cereal at the grocery store) because it would require a dozens of discrete populations and precise wiring to resolve their competition. Fourth, it nor does it deal well with newly introduced novel offers because such offers would need to be rapidly added to the network which would then need to be appropriately wired. Fifth, it introduces the need to coordinate a flexible linkage between offers, values, actions, and positions in space; we call these problems the neuroeconomic binding problems. While these problems are undoubtedly surmountable, we wondered instead whether an alternative approach could provide a better framework for models of economic choice.

Several recent findings have, in our view, begun to bring into focus an alternative picture of how choice works. Here, we first review that evidence, with a focus on primate single-unit recording studies. First, we describe six major research trends that, together, point towards our integrated framework. Second, we describe that framework. This framework is also directly motivated by principles of foraging theory that, in our view, constructively interact with the principles of neuroeconomics to guide our understanding of reward-based choice (Hayden, 2018; Pearson et al., 2014). Finally, we describe how or framework can resolve the neuroeconomic binding problems.

Part I: review of empirical findings

We evaluate only one option at a time

When multiple stimuli appear in our visual world, attention selects one at a time and then moves to the next one in a serial manner, like a roving spotlight (Egeth and Yantis, 1997; Treisman and Gelade, 1980). While some features can be analysed in parallel, complex feature extraction requires unitary focal attention. Economic value seems likely to the kind of feature that requires attention. It is not surprising, then, that when two options are

presented in the visual field, our eyes naturally shift back and forth between them to evaluate them (Krajbich et al., 2010; Orquin and Mueller Loose, 2013). When gaze is held fixed by the experimenter, or when options are not presented visually, decision-makers may still covertly shift mental focus between options serially.

The idea that evaluation is serial is supported by studies of the relationship between fixation patterns and choices (Krajbich et al., 2010; Krajbich and Rangel, 2011). Neural evidence supports, or is consistent with, the idea that the core value regions, vmPFC, OFC, and VS encode the value of the single attended option (Blanchard et al., 2015a; Lim et al., 2011; McGinty et al., 2016; Rudebeck et al., 2013; Strait et al., 2014, 2015; Xie et al., 2017). In a recent study of the OFC, ensembles of neurons alternated between encoding only one of the two available options rather than encoding both at the same time (Rich and Wallis, 2016). Notably, these coding states did not track the locus of gaze, but presumably tracked the focus of attention, suggesting that it is attention, not gaze direction per se, that determines which option is evaluated.

The idea of a single-option limit is also consistent with the foraging theory-derived idea that choice naturally occurs in an accept-reject manner (Hayden, 2018; Kacelnik et al., 2011; Kolling et al., 2012; Shapiro et al., 2008; Stephens and Krebs, 1986). The underlying foraging models that support these ideas have also proven quite useful in explaining a great deal of brain activity (Blanchard and Hayden, 2014; Boorman et al., 2009, 2011, 2013; Hayden, 2018; Kolling et al., 2012, 2014).

Implications for the framework. If we can only attend one offer at a time, then processing of the two offers in a binary choice must occur serially, not in parallel (Figure 1). (The same is true for choices with more than two offers, see below.) Relative to parallel models, serial processing poses a new problem and solves an old one. The new problem is that it requires a working memory buffer so that the value of a previously attended option can be maintained in order for any comparison to occur. The

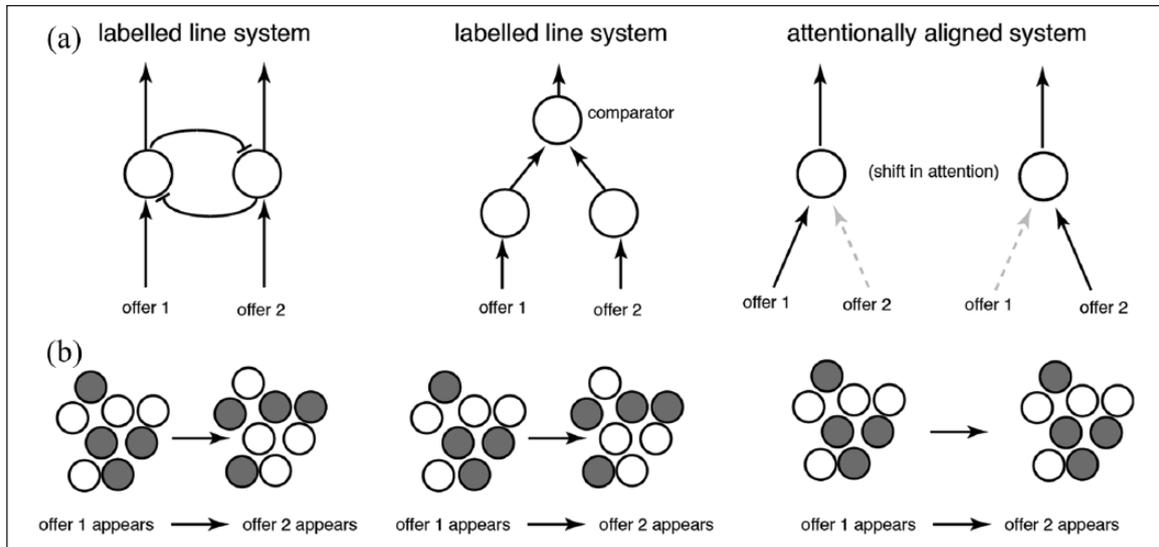


Figure 2. Illustration of core ideas of one- and two-pool models. (a) Left: One approach to modelling choice is the labelled line approach. Each neuron is associated with a specific option and these neurons compete, typically through mutual inhibition, for control of behaviour. Centre: Another take on the labelled line approach has three classes of neurons, one for each option and a third for comparison. Right: An alternative *attentionally aligned* approach (supported by recent work reviewed here) eschews labelled lines, but instead involves alternations between states corresponding to just one offer. When attention shifts, the inputs to the value neurons change to reflect the attended option. Because only one option is attended, value-sensitive neurons do not need to have information about which option their value signals. (b) When attention shifts from one option to another, a labelled line system will switch which neurons are strongly activated; an attentionally aligned system does not. Nor will tuning functions change. Thus, measuring which neurons are involved in signalling the values of the two offers, and their tuning, can test between two-pool and one-pool models.

solved problem is the option-value binding problem. Because attention is limited to one option, there is no ambiguity about the reference of value-related neural responses. As long as the decoder knows where the focus of attention is, the referent of the value signal is unambiguous (again, that focus need not be spatial; it may be abstract and conceptual).

We decide whether to accept or reject that option

If only one option is attended at a time, it is natural that the decision will be simply to accept or reject that one. Rejection would be favoured, even for very good options, when the cost of inspecting the next one is low and there is no cost to returning to the first one (as in most laboratory binary choice tasks, although not necessarily in natural contexts). In the laboratory, then, we would therefore expect a period of multiple inspections before a period of choice. As noted, foraging theory has long emphasised the idea that preys are naturally encountered alone, and thus, our brain's evolved choice strategy is to either accept or reject a single offer (Blanchard and Hayden, 2015; Charnov, 1976; Hayden, 2018; Krebs et al., 1977; Stephens and Krebs, 1986). This decision is made relative to an estimate of the value of rejection (i.e. the opportunity cost of accepting), known as the *background value*.

From this accept-reject perspective, ostensibly binary choices involve two largely distinct accept-reject decisions, one for each offer (Freidin et al., 2009; Kacelnik et al., 2011; Shapiro et al., 2008). These two decisions may be implemented by separate, possibly interacting, diffusion-to-bound processes. Another implication of this idea is that in choice,

options are given special status: default (the currently attended one) and alternative (the other one). A good deal of evidence supports the idea that cortical choice processes adopt this framing (Azab and Hayden, 2017; Boorman et al., 2009, 2013; Kolling et al., 2012, 2014).

Implications for the framework. If we attend single offers in turn and accept or reject each one, direct comparison of values per se need not occur (Kacelnik et al., 2011). Comparison may instead result indirectly from the fact that we cannot choose both options, even if both options are favoured. As a consequence, we do not need special value comparison neurons; motor gating processes perform that role. Moreover, if valuation of the attended option is performed relative to the value of the alternative, then we do not need separate pools of neurons for representing the two offers. One will suffice (Figure 2).

Getting rid of comparator neurons avoids the difficult binding problem by which the offer-selective neurons are dynamically configured to converge on specific comparator neurons. Getting rid of distinct pools for the two options lets the brain use all its resources to the difficult problem of value estimation rather than using redundant anatomically separate computational resources for every option that is (or even that can be) available. The disadvantages are that serial value estimation is slow and requires working memory.

Attention, not labelled lines, determines how value is bound to options

When attention shifts from one option to another, value-coding neurons in several regions switch from encoding the value of the

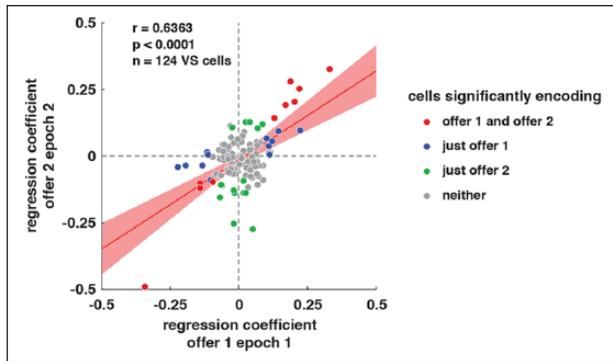


Figure 3. Neurons use a similar ensemble coding format (signed regression coefficient) for the values of offer 1 and offer 2 when they are attended in sequence. In this illustration, options were presented and thus attended asynchronously, and regression coefficients for each neuron were estimated in the appropriate epochs. Each dot indicates a single neuron; its x and y positions indicate the linear component of its tuning function for each option. The positive correlation between the two indicates a preservation of tuning, as predicted by a single-population model. A labelled line model would predict that points would cluster around the anti-diagonal and produce an anti-correlation. Data from VS are shown (Strait et al., 2015); similar patterns were observed in other core value regions (see text).

first option to encoding the value of the second. Often, neuronal responses are consistent with use of the same *format* to encode offer values across shifts in attention. That is, a neuron positively tuned for the value of the first considered offer will remain positively tuned for the value of the second and vice versa. We introduce the term *attentionally aligned coding* to refer to this response pattern, which can be distinguished from *labelled line coding*, where a neuron's firing rate refers stably to a single option regardless of the locus of attention (Azab and Hayden, 2017). The term *attentionally aligned* means that the referent of a value neuron's firing rate is not consistently aligned to a single option, but rather is aligned to the value of any option within the focus of attention. (In the form vision pathway, the analogous idea is known as biased competition (Desimone and Duncan, 1995; Pastor-Bernier and Cisek, 2011).) Attentionally aligned coding is convenient if attention is limited to a single option at a time, but becomes unwieldy if multiple options can be attended at once.

Attentional alignment is consistent with neuroimaging studies (Lim et al., 2011), and in careful studies of behaviour (Krajbich et al., 2010). It has been reported in neurons in vmPFC, VS, OFC, dACC, and subgenual anterior cingulate cortex (sgACC) (Azab and Hayden, 2017, 2018; Blanchard et al., 2015; Rich and Wallis, 2016; Rudebeck et al., 2013; Strait et al., 2014, 2015; Xie et al., 2017) and is consistent with another recent OFC study (McGinty et al., 2016). Evidence for attentional alignment is illustrated in Figure 3.

Implications for the framework. If value coding is attentionally aligned, then the framework can have a set of value-sensitive units that are ignorant of the details of the input stimuli. Specific value-sensitive units in the network have an organisational advantage: they will not need to be precisely wired with

offer-layer neurons. This arrangement gives the system much more flexibility to deal with rapidly changing options, new options, and more than two options. One disadvantage is that if an ensemble of attentionally aligned neurons uses the same format to encode the value of two different options, a decoder cannot know, without some additional information (specifically, which option is attended), to which option a neuron firing rate refers. By contrast, in a labelled line coding system, there is no ambiguity about which option a neuron's firing rate indicates: after all, the line is labelled. On the other hand, if the decoder knows the status of attention, then the referent of the neuron's firing rate is unambiguous. Thus, the option-value binding problem can be solved without need for any supervisory system other than the one that controls attention.

One pool of neurons, not two

When attention shifts, and the value code shifts with it, a good deal of evidence indicates (or at least is consistent with the idea) that it is the same neurons activated for the previous option that are activated for the next one (Azab and Hayden, 2017, 2018; Blanchard et al., 2015, 2018; Rich and Wallis, 2016; Rudebeck et al., 2013; Xie et al., 2017). In other words, the brain may use only one pool of neurons to encode the two different values at different times, not two separate ones. At least one study indicates that some of these regions use a single pool of neurons to encode offered and chosen values as well (Blanchard and Hayden, 2018).

A simple test for separate populations is to compare unsigned regression coefficients (this is similar to, but more statistically sensitive than, performing a Venn Diagram analysis, Figure 4). This method reveals evidence in favour of a single population in OFC, vmPFC, VS, and dACC (Azab and Hayden, 2018; Blanchard et al., 2015; Strait et al., 2015; Strait et al., 2015; Wang and Hayden, 2017). A more sensitive method uses Bayesian statistics to ask whether information derived from the tuning functions for the two variables can be used to support purported clusters (Blanchard and Hayden, 2018). This method rejects any option-specific clustering in any of four brain regions (VS, vmPFC, OFC, and dACC). It also rejects clustering for offer and chosen values.

Note that the case here is not definitive; there is a good deal of ostensibly contradictory empirical support for two pools, and several papers for data consistent with a two-pool model (Padoa-Schioppa, 2011). The question of how many pools there are is difficult to answer because the brain may in principle divide up the two offers in any of a number of ways, perhaps arbitrarily and perhaps randomly from trial to trial. Methods that average across multiple trials may then therefore average across the two pools making two look like one. Our analyses so far suggest that neurons do not consistently align to the first/second offer or the left/right offer in asynchronous left-right choices (Azab and Hayden, 2017; Blanchard et al., 2015, 2018). Perhaps the strongest evidence so far comes from datasets with simultaneously recorded cells, allowing robust single trial analysis, which still fail to indicate separate pools of cells (Rich and Wallis, 2016).

Implications for the framework. The one-pool finding, if true, goes hand in glove with the attentional alignment

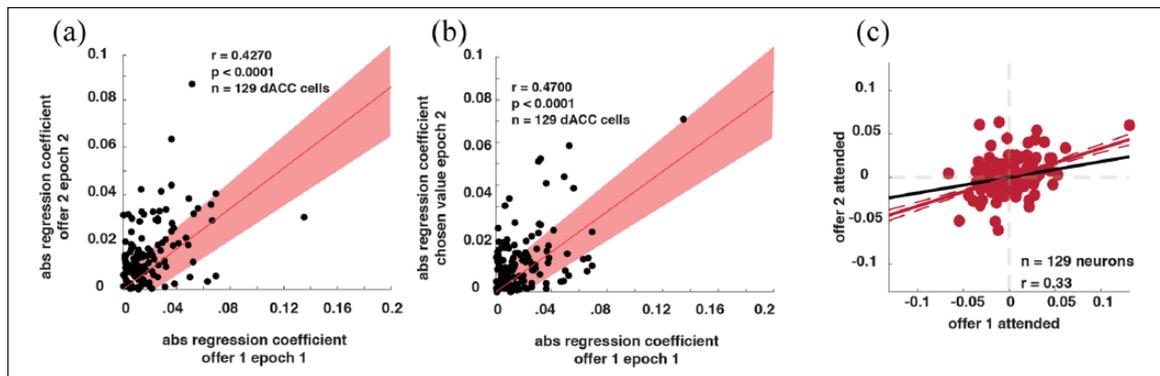


Figure 4. Some evidence for a one-pool model. (a) In a simple gambling task with two offers presented asynchronously, neurons selective for the value of the first offer are more likely to be selective for the value of the second as well. Selectivity here is measured by the absolute value of the regression coefficient of firing rate against the value of the offer. The positive correlation between the variables indicates that a neuron driven by the value of the first offer is more, not less, likely, to be driven by the value of the second one. We see no evidence of separate populations of cells, as would be predicted by a labelled line model. Illustrative data from dACC shown (Azab and Hayden, 2018); similar patterns were observed in other regions (see text). (b) Evidence for a single-population encoding offered values and chosen values. Illustrative data from dACC shown (Azab and Hayden, 2018); (c) aligned encoding of attended offer values. Values of attended offers are encoded in correlated formats across time, in contrast to two-pool model predictions of mutual inhibition through time (Azab and Hayden, 2017).

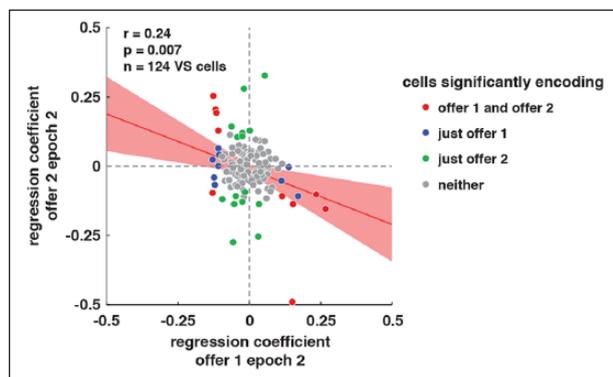


Figure 5. Neurons use inverted tuning formats (reversed regression coefficients) to encode the values of the two offers during choice. Each point indicates a single neuron. The x- and y-positions of the dots indicate the linear term of the regression coefficient for the firing rate of that neuron at the time of choice against the values of the two different (and uncorrelated) offers. The negative correlation indicates that these coefficients are anti-correlated and thus that the population encodes the difference in the values of the two offers. Illustrative data from VS shown (Strait et al., 2015); similar patterns were observed in other core reward regions.

hypothesis. Specifically, if there is a single pool of neurons, its firing rates must somehow be linked with an option. The most straightforward way would be to limit reference to a single possible option, defined by attention. That system would allow the set of neurons to flexibly encode the value of any offer and would free the system from having to have a rigid linkage for offers and values. The narrow restriction of attention to a single option would thereby resolve the option-value binding problem. This logic also works for the tentative finding that offered and chosen values are encoded by the same neurons: presumably the chosen offer is attended around and immediately after the time it

is chosen, and so it should be encoded in the same neurons that encoded its value at offer time.

Opposed tuning for offer pairs at the time of choice

At the time, the second offer is attended (and the value of the first is available in working memory), the decision-maker can begin comparing their values. At this time, firing rates of neurons in several regions encode the difference in values of the two offers. Specifically, individual neurons tend to show opposing tuning functions for their values. These regions include vmPFC (Strait et al., 2014), VS (Strait et al., Figure 5), dACC (Azab and Hayden, 2017), and sgACC (Azab and Hayden, 2018), as well as the PMd (Pastor-Bernier and Cisek, 2011) and the supplementary eye fields (SEF) (Chen and Stuphorn, 2016). These signals are observed in the same neurons that encode the values of the individual offers and not a separate class of neurons. This pattern is broadly consistent with the finding that several brain regions show coding for the difference in the values of the two offers (Basten et al., 2010; Boorman et al., 2009; FitzGerald et al., 2009; Hunt et al., 2012). The value difference is *the* key decision variable for economic choice – a simple threshold applied to value difference will produce a choice. It is thus, arguably, a signature of value comparison.

We previously argued that the value difference result is a signature of value comparison by separate populations (Strait et al., 2014). The data reviewed in the previous sections suggest an alternative interpretation that neurons encode the relative value of the attended offer. That is, they encode its value relative to the value of rejecting it, which in binary choice is equal to the value of the other offer. There is a good deal of evidence that evaluation in the brain is done in a relative manner. It has variously been labelled range adaptation (Padoa-Schioppa, 2009), mutual inhibition (Hunt et al., 2012; Jocham et al., 2012; Strait et al., 2014), and divisive normalisation (Louie et al., 2011; Yamada et al.,

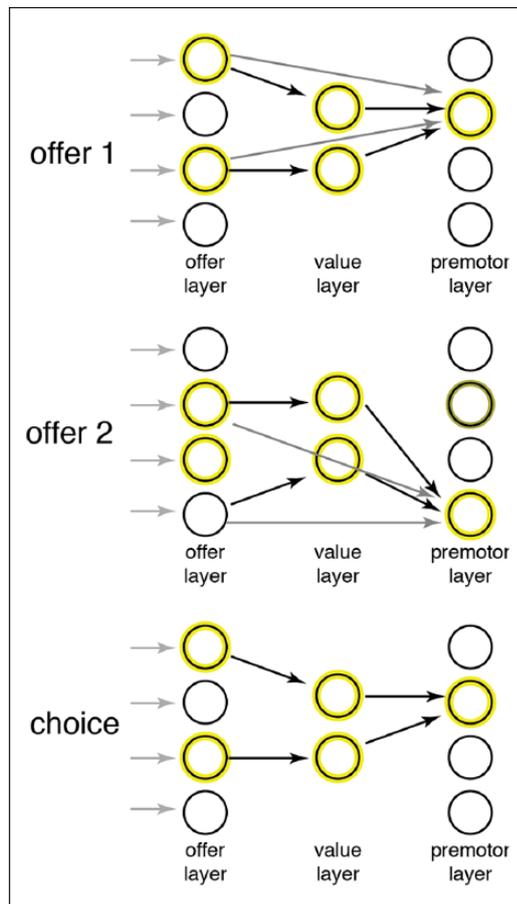


Figure 6. Our proposed framework in practice. When the *first offer* appears, its feature detectors are excited, which define the distributed response in the offer layer. These activate the appropriate value units in the appropriate manner to signal the value of the first offer. They also activate the corresponding premotor layer units (grey arrows). Those are the units that, if the action they signal is released, the animal will choose the offer. When the *second offer* appears, its feature detectors are excited. They activate the appropriate value units, which are likely the same ones that were activated by the first offer, and with the same tuning function. They also activate their appropriate premotor units. Finally, *following choice*, the chosen offer is attended and so its features, value, and action units are activated. This activation allows for credit assignment processes to know the appropriate elements to sculpt for learning.

2018). The key thing, though, is that relative value has the same benefits of value difference coding: subject only to a thresholding operation, it can serve as the basis for categorical choice.

Implications for the framework. Putative value comparison signals, in the same neurons that encode values of offers, indicate that these neurons do not specialise in encoding the value of the attended offer. Instead, they have a more sophisticated and flexible role in choice. Specifically, they can encode the value difference, or the relative value (or a function thereof) between the attended and the remembered values. Doing so requires them to have some kind of working memory store (whether active or passive) and raises the

question of what form it takes (see model below). Note also that the presence of value comparison signals in multiple regions suggests that the comparison occurs simultaneously in multiple regions (see Discussion).

Activation of motor plans during deliberation

When we evaluate options, and before we choose, the anticipated motor plans of both options are encoded in premotor and parietal cortices (Baumann et al., 2009; Cisek and Kalaska, 2005; McPeck and Keller, 2002; Scherberger and Andersen, 2007). When the action is clear and overt, that action plan is called an affordance (Cisek, 2007). We use the more generic term *action plan* here to mean pretty much the same thing as affordance, but to include contexts in which the specific motor command is not clear (imagine for example you are asked which entrée you wish to order, but there is no menu to point at). As evidence accumulates in favour of one option, its corresponding action plan gets stronger and the other one gets weaker, until the decision threshold is reached (Cisek, 2006; Cisek and Kalaska, 2005). The gradual evolution of these processes, in turn, gives rise to decisional commitment (Thura and Cisek, 2014). Together, these findings support a biased competition model for economic choice, which extends classic ideas of biased competition from the perceptual system to the motor system (Cisek, 2005; Cisek and Kalaska, 2010; Pastor-Bernier and Cisek, 2011).

Implications for the framework. These results suggest that attending to one offer will activate its action plan, and that switching to the other will suppress and enhance the action plan of the first. During deliberation, these modulations will not trigger an action, but they will be critical for the process of selection and commitment that occur when deliberation ends. These findings also raise the possibility of a solution to the action binding problem as well: if the attended offer activates its corresponding representation in the premotor system, then there is no ambiguity about which option that action code corresponds to.

Part II: a proposed framework for the neural basis of economic choice

The recent empirical findings point towards the following general framework (Figure 6). Sensory inputs activate specific populations of units that represent complex features, and thus, the activation of those features in the offer layer defines the identity and characteristics of the offer under the scrutiny of attention. These offer-layer units then activate a representation of the option's value in units in a separate value layer. The offer-layer units convey no information about the value of the offers; this information is stored in the connections and/or processes (which we do not describe) that link offer-layer units with value-layer units. Responses from the value layer convey no information about the identity of the option; they simply signal the value of the currently attended option. (Note that we use the term value for convenience. The variable could be any variable or set of variables that correlate with choosing the attended option, such as 'evidence in favour', or signals that reflect the values of individual object features). Some research indicates that the brain treats differently variables that are directly observed and those

that are inferred (e.g. Barron et al., 2013; Jones et al., 2012). In our framework, the inference process is one that occurs in the value layer. It is thus endowed with some complex computational abilities.

The activation of the offer layer will also activate the option's action plan in a premotor layer. The action plan is the specific action that would be used to select the option and can be as specific as a reach or as abstract as the concept of 'select this option using the appropriate action when that action is later identified'. The premotor units get signals from both the offer and the value layer. The option information activates the associated action; the value information activates all output actions non-specifically (arrows are not displayed but they are understood as present), providing a general drive to act. The interaction between the value units and the offer units allows for only the attended action to be selected. Finally, the framework allows for additional non-specific modulatory inputs to all action units, which lets extraneous factors such as urgency to affect the likelihood that an action will be triggered (not shown).

First fixation: value and action plan for first option

We propose that in practice, consideration of options is nearly always asynchronous. That is, even when multiple options are presented simultaneously, attention selects one of them for scrutiny first, possibly covertly (Krajbich et al., 2010). When the first offer is attended, the units responsive to its features and/or identity in the offer layer are activated; these units proceed to activate corresponding value and action units. The action will not yet be triggered. In most cases, assuming the need to decide is not urgent, the first option is likely to be automatically rejected in order to consider the second option; this would be implemented by the global modulatory inputs. Specifically, it would be rejected because the background benefit–cost ratio is quite high: it includes the informational value of the second offer at the very low time and energetic costs of simply attending to it.

Second fixation: relative value and action plan for second option

When attention shifts to the second offer, its corresponding offer-layer units that represent their features will be activated. These units then activate the corresponding value-layer units – which will be the same ones that signalled the value of the first offer; they will also use the same format to do so (e.g. a unit with positive tuning for offer 1 will have positive tuning for offer 2). Units in the value layer take on the property of *response-dependent suppression*, meaning its response to the first attended offer attenuates its response to the second one in proportion to its response to the first (or a function thereof). Suppression is not necessary per se, as response-dependent enhancement could work as well. However, response-dependent suppression is a prominent feature of the inferotemporal cortex (IT, for example, Miller et al. 1991) and may be observed in the reward system as well (e.g. Barron et al., 2013) and dependence on previous outcomes is commonly observed in the reward system (e.g. Hayden et al., 2011a, 2011b; Kennerley et al., 2011). This response-dependent suppression will serve the purpose of a within-cell memory (i.e. does not require an

additional external memory buffer and thus can occur within a single pool) that will produce value comparisons. There are many possible neuronal mechanisms that could implement response-dependent suppression; we use a simple one for concreteness, one for which there is ample evidence in different domains: divisive normalisation (Carandini and Heeger, 2012; Reynolds and Heeger, 2009).

When the second offer is processed, the value-layer units will show response-dependent suppression for the value of the first offer. If the first offer was particularly good, the response to even a good second offer will be attenuated. If the first offer was poor, the response to the second will be less attenuated. The value-coding units will therefore exhibit simultaneous and anti-correlated tuning for the values of both offers (as in Strait et al., 2014). However, whereas in that paper we proposed that this pattern results from competing populations, response-dependent suppression only involves a single population. Notably, the value of the second offer will not be encoded per se. It will only ever be encoded relative to the value of the first. The second offer will also lead directly to the activation of its action plan, just as the first offer did. The action plan for offer 2 will be more strongly activated than the one for offer 1, because attention enhances the action plan. However, we anticipate the action plan for offer 1 will still be moderately activated, due to system hysteresis. Both action plans will in any case be activated simultaneously (as in Cisek and Kalaska, 2005).

Subsequent to the second fixation, subjects may select it or they may return to the first offer. A return to the first offer will lead the value-layer units to encode its value relative to the value of the second. (This hypothesis has not yet been tested, but follows naturally from our framework). Its action plan will also be enhanced. This process can continue back and forth until an option is selected (as in Rich and Wallis, 2016). Why would a decision-maker come back a second time rather than just decide immediately? Additional bouts of consideration may provide a more accurate estimate of the value of the offers to due accumulated response-suppression of the value unit, allowing for fine discrimination of closely valued options (Krajbich and Rangel, 2011).

Choice and outcome periods: relative value and action plan of chosen option

An option is selected when the activation in the action layer crosses some threshold. The threshold is determined by several factors that together embody the value of rejection. There is very little data on the process of threshold computation (but see Kolling et al., 2014). However, we assume that rejection has a high value following the first offer (because of the informational benefit and low cost of inspecting the second one.) The value of rejection will decrease as time increases and the opportunity cost of further deliberation rises. Once one action plan crosses a threshold, commitment occurs and the selected action inhibits other activated actions, so as to ensure only one action plan is implemented (Thura and Cisek, 2014.) The selection process leads the chosen option into the focus of attention. As such, its offer units are preferentially activated and value-layer units encode its value. Note that there are no chosen value units in our framework; the units encoding the chosen value are the same value-selective units that were involved in choice.

After selection

After selection, the reward is received, the chosen option is attended, and its corresponding offer, value, and action units are correspondingly activated (or reactivated). At this point, post-reward processes come into play. These post-reward processes include monitoring, learning, adjustment, and updating of priors, as well as possibly switching to new strategies or rules. These processes are unique to the post-reward period and will therefore create patterns that are not observed in the offer period, but that will be superimposed on the standard offer-related signals (Nogueira et al., 2017; Wang and Hayden, 2017).

Extending the framework to more than two options

Our framework deals well with binary choices, but they need an additional feature to handle choices with more than two options (which we call multi-option choices for convenience). Our model here will be more speculative since we do not have unit data from multi-option choices, although relevant lesion (Noonan et al., 2010), neuroimaging (Boorman et al., 2011, 2013), and perceptual decision-making (Churchland et al., 2008) data exist. We propose that when attention falls on the third option, the brain encodes its value relative to the value of the best of the first two options (Boorman et al., 2013). The brain could maintain a separate buffer to store the value of the best-so-far option, but we propose a simpler alternative with a single unlabelled value buffer.

Specifically, we propose that the brain maintains an active salience buffer – a representation of the entire option space (both the visual scene and some abstract set of options could be included). The buffer tracks the location of the most valued option so far – but not its value, nor its identity or action plan. The computational framework described above can also be extended to account for multi-option choices using the idea of the salience buffer. The basic idea is that only the offer with the highest value so far is actively remembered, causing divisive normalisation on the current stimulus being attended (cf. Louie et al., 2011). Finally, based on the current response, a choice needs to be made between the new stimulus and the past stimulus with the highest value. This pairwise comparison can be made in the same way as described in the previous section.

A possible neuronal implementation

The key element of our computational model is the presence of a working memory mechanism that affects the response of the value layer to the second option and depends on the value of the first option. We propose that this computation is implemented by repetition suppression, via divisive normalisation (Carandini and Heeger, 2012; Reynolds and Heeger, 2009). Note that repetition suppression per se is not necessary; similar effects can be obtained if neurons exhibit repetition enhancement. We focus here on repetition suppression because it is strongly supported empirically (e.g. firing rate adaptation).

We assume that the attended option encoded in the offer layer delivers information to the value-encoding layer (see Figure 7). In response to the first offer, with value V_1 , the firing rate of neuron i ($i = 1, \dots, N$) in the value layer is

$$r_i = \alpha_i V_1$$

where α_i is a positive coefficient. We consider for simplicity only positive values, although this framework can be naturally extended to negative values too, and to any arbitrary form of tuning curves (e.g. non-linear).

When the second offer appears, the response of the value-encoding neurons will be diminished because of repetition suppression in the value (or even in the offer) layer. We assume here that repetition suppression in the value layer is caused by divisive normalisation of the neuronal response to the second option in proportion to the strength of the first response. Specifically, the response of neuron i in the value layer to the second offer becomes

$$r_i = \frac{\alpha_i V_2}{1 + \beta_i V_1} \quad (1)$$

where β_i is a small positive number ($0 < \beta_i \ll 1$). As a consequence, responses to the second offer will be reduced even if the values of the two options are identical, consistent with the experimental data provided above.

Finally, a choice between the first and second options needs to be made based on the responses that are available in the final stage, that is, the responses to the second offer in equation (1). The choice cannot be done with a homogeneous layer of value-encoding neurons, but it can be done with a heterogeneous layer where the sensitivity parameters α_i and β_i differ across neurons. This is because setting a threshold to equation (1) with identical parameter values for all neurons causes biases in the choice by leading it to prefer the first offer over the second one, or vice versa. However, this bias can be avoided by simply linearly combining the responses of a heterogeneous population of N neurons to make it approximately equal to the difference of values of the first and second stimuli

$$\sum_{i=1}^N w_i r_i \approx V_1 - V_2 \quad (2)$$

This linear combination can deliver a very good approximation of the actual difference if neurons are heterogeneous and the population is sufficiently large, as we will see in a neuronal implementation of this basic algorithm (see Figure 8). This is because if the divisive normalisation parameter β_i in equation (1) is small, the firing rates can be expanded approximately as linear function of both values. Combining several of those approximately linear functions, it is possible to compute the value difference, which is again a linear function. Therefore, a layer of value neurons with repetition suppression has all necessary information to perform a sequential comparison of two offers, and this information can be extracted by a simple linear readout.

What kind of signals in the brain could carry information from the past to the present in a format that allows also comparing values of sequentially attended stimuli? One such potential candidate is synaptic depression (Abbott, 1997; Tsodyks and Markram, 1997). Synaptic depression acts on the inputs to a neuronal population in such a way that continuous stimulation causes synaptic resources, such as number of vesicles and amount of neurotransmitter, to be depleted. Due to its slow decay, depressing synapses can hold information in working memory for

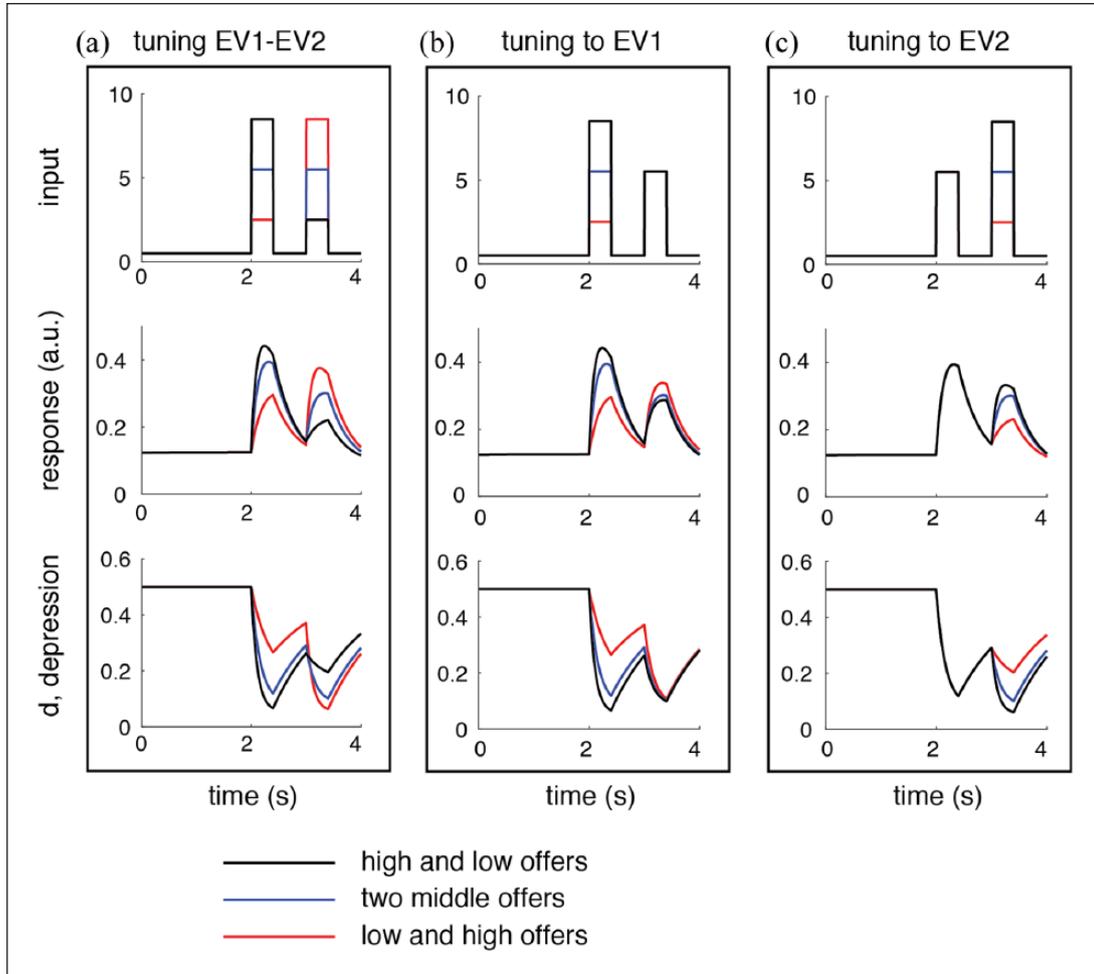


Figure 7. A neuron in the value layer has similar tuning to the values of the first and second offers and shows repetition suppression. (a) The tuning to the difference between offer values (tuning to EV1-EV2), (b) to the first value (tuning to EV1) and (c) the second value (tuning to EV2) are shown. In (b) and (c), the values of the first or second offers are fixed, respectively. The input (top row), the response (middle row), and the depression variable (bottom row) are displayed as a function of time. The input, encoded in the projections from the offer to the value layers, is proportional to the expected values of the two offers, presented at times 2 and 3s, respectively. Three different conditions are used (black, blue, and grey lines), see text

several seconds (Miller and Desimone, 1994; Miller et al., 1991; Mongillo et al., 2008). Thus, synaptic depression is a potential mechanism for facilitating the comparison between the values of two offers presented asynchronously through repetition suppression. It is possible that there are multiple mechanisms with similar effects working together, for instance, firing rate adaptation in the offer and value layers. Here we show simply that synaptic depression is a good starting point, although we acknowledge that due to its rigidity in the slow time scales involved, it will be insufficient to accommodate the large variations of timing in which decision-making can occur. This proposal then should be seen as a workable example of how these changes may occur, illustrating the viability of our framework.

We consider a value layer composed of N independent neurons described by their temporally modulated averaged firing rate, $r_i(t)$ ($i = 1, \dots, N$) receiving inputs subject to synaptic depression. Each neuron i in the value layer receives the external input $I_i(t)$, modelled as a time-varying signal weighted by

the value of the stimulus plus background activity, $I_i(t) = a_i + b_i \times V(t)$. We assume that the value of the attended stimulus $V(t)$ is computed as a weighted linear combination of activities in the offer layer (see Figure 6), although of course non-linear function beyond linear can be achieved by multi-layer networks.

The net input into each neuron is computed as $d_i(t) \times I_i(t)$, where $d_i(t)$ is the synaptic depression variable for the inputs for neuron i , with takes lower values the higher the activity was in the recent past. Therefore, the net input into each neuron is not simply the external current, but a normalised version of it with a normalisation factor that depends on the previous history of attended options and their values. Further details for the models are described in the ‘Methods’ section. The dynamics of a neuron in the value layer is shown in Figure 7 for three relevant scenarios. The external input to the neuron is shown in the top panel, while the response and the synaptic depression variables are shown in the middle and bottom panels, respectively.

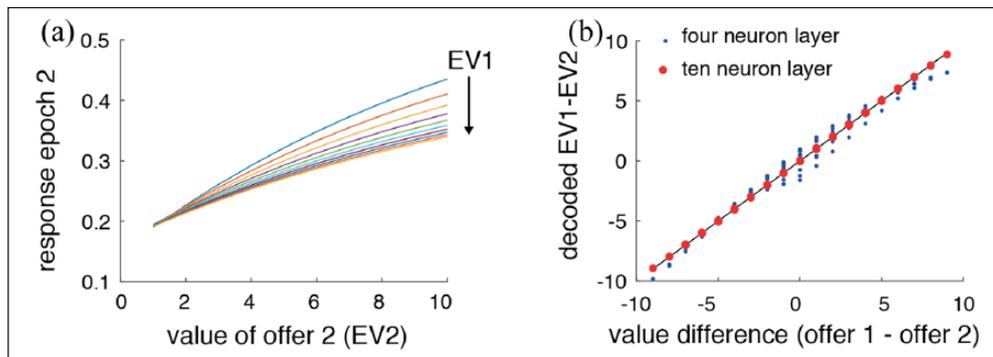


Figure 8. A heterogeneous population of neurons in the value layer can faithfully encode the value difference between first and second offers. (a) The response of a representative neuron in the value layer during the second epoch increases with the value of the offer in that epoch, EV2, but it is also negatively modulated by the value of the offer previously presented, EV1. This mixed encoding makes impossible to read the value difference from just a single neuron. (b) Decoded value difference as a function of the real value difference for a population of 4 (blue) and 10 (red) neurons. The decoded values get closer to the actual values (unit slope line, black) as the population is larger. The decoder is based on a linear readout of the population using the responses at 100 ms after second offer onset, trained using linear regression, equation (2).

In the first scenario (tuning to EV1–EV2; Figure 7(a)), the external input alternates between a high and a low value offers (black, top panel), two intermediate value offers (blue), and one low and a high value offers (red). The response of the neurons follows the same trend as the input (middle panel), while the synaptic depression variables display the reversed trend (lower panel). Note the response to the intermediate values (blue, middle panel) in the first and second epochs: the response is reduced during the second stimulation epoch compared to that in the first epoch, even though the stimulus value is identical in the two conditions. This phenomenon corresponds to repetition suppression, as implemented by divisive normalisation (Methods, equation (4)).

This can be understood by looking at the temporal evolution of synaptic depression variable. Initially, this variable has a relatively large value due to low spontaneous firing rate (around one half). However, during attention to the first option, the external input increases and as a consequence, the depression variable decays to a lower value (blue, lower panel). Right after offer offset, the depression variable starts to recover and increase towards the basal value. However, the increase is slow due to the long recovery time constant of synaptic resources, and thus, the depression variable does not have time to reach the basal value. Indeed, when the second offer is attended, the depression variable has a value that is lower than the basal value. This difference leads to a response to the second offer that depends on the value of the first offer.

In the second scenario (tuning to EV1, Figure 7(b)), the value of the first offer ranges from high (black, top panel) to intermediate (blue) and low (red), while the value of the second attended offer is fixed at an intermediate value. During the presentation of the first offer, the response of the neuron increases with its value (middle panel), indicating that this neuron has a positive encoding of the first offer value. It is interesting to observe is that during the presentation of the second offer, this cell is still tuned to the value of the offer. However, the tuning is reversed: higher responses are obtained for the lower value in the first attended offer. Also the tuning to the value of the first attended option during the second epoch is reduced compared to the tuning during

the first epoch. These two patterns reproduce the experimental results from our laboratory in vmPFC, VS, and dACC (Azab and Hayden, 2018; Strait et al., 2014, 2015) and echo those of Pastor-Bernier and Cisek (2011).

In the final scenario (tuning to EV2), the value of the first offer is fixed, while the value of the second offer varies. Consistent with experimental results (Azab and Hayden, 2018; Strait et al., 2014, 2015), the tuning of this cell to the value of the second offer is positive. Thus, the neuron tends to keep the same polarity towards stimulus value regardless of stimulus identity or presentation timing.

Decoding choices

We next asked whether a downstream decoder can make an accurate choice based on the activity of the neurons in the value layer during the second offer epoch. As noted, the response to the second attended option is inverted relative to the first. This inverted tuning allows the system to compare the values. How can this information be extracted? As with the computational framework, it is not possible to read out this information if only one type of neuron is present in the value layer. This is because the firing rate of a neuron during the second epoch depends on the values of both first and second attended offers and does not necessarily compute a value difference between the two. Our strategy is then to create a heterogeneous population of neurons in the value layer, which is a realistic feature throughout the brain architecture. Heterogeneity can be introduced by choosing neuron and depression parameters randomly (Methods).

Note that in our framework, it is not necessary that every neuron perform normalisation with respect to the previous observed value in an identical way. Rather, we postulate that the normalisation effects are heterogeneous, which allows the neurons to encode the two offer values in the same population, which in turn allows for a choice. Thus, we do not directly suggest that there are neurons for which the beta parameter is very small, although it is not possible to disregard that possibility in a small population of neurons.

With a value layer consisting of just four neurons, it is possible to estimate the value difference approximately (Figure 8, blue points; max error=0.97). With 10 neurons, it is possible to estimate value difference with high precision (Figure 8, red points; max error=0.08). Although these simulations are based on deterministic dynamics, the presence of response variability can be partially alleviated simply using larger populations if differential correlations are weak (Moreno-Bote et al., 2014). Therefore, it is possible to compare values of sequentially presented offers by linearly reading out the activity of a small neuronal population in the value layer during attention to the second offer.

Discussion

Review recent discoveries about the neuronal correlates of economic choices suggest a novel framework for future models of how choice occurs. In this framework, only one option is attended at a time and processing of that option leads to either acceptance or rejection. This occurs regardless of whether the options are presented asynchronously or whether they are presented at the same time. Rejection often leads to consideration of the next option. During deliberation, attention to an option activates a representation of its value and of the action plan associated with choosing it. This action plan can be specific or it can be abstract, that is, it can in principle be a commitment to a proposition (Shadlen and Kiani, 2013). Our framework requires a single pool of value-sensitive units whose responses encode the value of the attended option relative to the value of previously attended options. It does not involve two pools or more of cells that use labelled line coding and that compete for control of action. Comparison is accomplished through a value normalisation process that can occur simultaneously at multiple levels and that may involve a response-dependent suppression of future responding. As such, the evaluation, comparison, and selection are performed by the same pool of neurons.

Our proposal is preliminary and is not meant to serve as a formal model. Indeed, we believe that there is no currently sufficient data in the literature to make one. The innovative aspects of our theory are (1) that choice is serial, not parallel, and (2) that this allows for even ostensibly multi-option decisions to be implemented using accept-reject principles. In particular, it (3) requires the narrow focus on a single option (attention), and that this new requirement, while cumbersome, allows for (4) choice to occur within a single pool of neurons rather than between competing pools. That in turn allows for (5) an abandonment of a labelled line organisation. Giving up labelled lines carries several benefits, most notably produces networks that are more efficient to build and are more flexible; it also leads to a solution to important and often ignored binding problems.

Where is this one pool of neurons that is critical to our proposal? We suspect that it is not in any single brain region, but instead that pool really reflects processes occurring in parallel in multiple regions at the same time. The function of this distributed group of neurons is likely to be quite broad, but in the context of choice, that function can be thought of as simply the representation and the sequential comparison of values.

The proposal is not meant to be a formal model for choice, but is, rather, to be a general framework that can guide the development of such models in the future. One particular

limitation of the framework is that it does not correspond to the unit level. For example, the strict division into an offer layer, a value layer, and an action layer is not supported by current data. Instead, individual cells are likely to have multiple contributions in multiple layers simultaneously. These multiple functions in turn are reflected in the multiplexed and mixed selectivity patterns that are characteristic of the key regions (Blanchard and Hayden, 2018; Fusi et al., 2016; Raposo et al., 2014; Rigotti et al., 2013). These functions may even change and adjust with task context (e.g. Hunt et al., 2013). Another example is that value-sensitive neurons, such as those in OFC may be stimulus specific, and thus not directly analogous to our value layer (e.g. Schoenbaum et al., 1998).

A third example of a limitation of our framework is that value comparison is likely to occur not within a single region, but rather through a distributed consensus process that includes ostensibly motor and association regions both (Chen and Stuphorn, 2016; Cisek, 2012; Hunt and Hayden, 2017). Ultimately, we propose that our framework may be a description of the algorithmic level, but not the implementation level, of choice.

An important future direction for our theory is to deal with its clumsy handling of both the object and the action layer. That is, in our framework, these layers still have labelled lines. Instead, the focus of our proposal is showing that separate value populations are not needed. We suspect that it is possible to simplify these other layers as well and hope that future work, guided by to-be-collected data in these regions, will do so. We suspect that greater research on the relationship between memory and economic choice will provide grounding for these ideas (Shadlen and Shohamy, 2016). In particular, this research will need to shed light on how different dimensions that constitute value are combined to form an abstract value signal (e.g. Barron et al., 2013; Hayden, 2016; Kolling et al., 2014; Padoa-Schioppa and Assad, 2006).

Another important direction for future research will be understanding the process by which attention is withdrawn from a rejected option and shifted towards a new one. This in turn involves two component decisions: (1) the decision to commit to rejecting a currently attended option, so that attention can be withdrawn, and (2) the decision for where to move attention next. That decision has been treated as random (Krajbich and Rangel, 2011). However, it is likely that there is some prioritisation of information going on, perhaps in something like a salience map that is maintained to guide future decisions (Itti and Koch, 2001). Indeed, recent research linking ostensibly value-related regions to switching processes provides suggestive evidence that the ideas of attentional switching and choice have a deep linkage (Birrell and Brown, 2000; Sleezer et al., 2017; Sleezer and Hayden, 2016; Blanchard et al., 2014)

What are the critical tests of our proposal? The first question that must be addressed is whether the representation of the values of different offers occurs within a single population of neurons or within two discrete sets. As noted, our data support a one-pool model, but other studies support a two-pool model (most importantly, Padoa-Schioppa, 2011). Second, we need a greater understanding of the dynamics of choice within the neurons that implement it. Third, the generality of the ‘mutual inhibition’ finding must be established. Fourth, the causal role of the mutual inhibition signals, and, indeed, the other signals we review above

must be demonstrated. Finally, although theoretical work has addressed the dynamics of accept–reject decisions, its neural basis is poorly understood, and its proposed overlap with binary choice must be demonstrated.

Relation to models of sensory memory-guided decisions

Our framework is partially inspired by well-known models of (working) memory-guided perceptual decisions (Hayden and Gallant, 2013; Lui and Pasternak, 2011; Machens et al., 2005; Miller and Desimone, 1993; Miller et al., 1991; Mirabella et al., 2007; Romo et al., 2002; Romo and Salinas, 2001, 2003). Typically, in memory-guided perceptual decisions, the subject is presented with a memorandum, and then, following the delay, a probe. The subject is then asked to perform a perceptual discrimination on the relationship between the memorandum and the probe (e.g. do they match? Which has higher frequency?). One approach to modelling these decisions is to allow the memorandum to modify the sensory tuning properties of neurons so that their response to the probe makes the correct classification automatically (Machens et al., 2005; Miller and Desimone, 1994; Pagan et al., 2013). This general approach has been successful in modelling mid-/high-level and prefrontal responses (with visual memoranda) and prefrontal responses (with somatosensory memoranda.) Indeed, we propose that binary economic and mnemonic decisions may function through similar brain mechanisms.

The overlap between our proposed framework and the framework used for perceptual decisions is not limited to its relationship with memory-guided decisions. The attentionally aligned coding scheme we propose is shared with perceptual systems. For example, neurons in the ventral visual system have large receptive fields that often contain multiple stimuli competing for attention. Focusing attention on a particular stimulus causes the neurons to behave as if the attended option was the only one present. Thus, the identity of the attended option is identified only by the status of attention. When attention shifts (within the receptive field), the tuning stays the same but the response changes to one that is based on the newly attended stimulus.

This principle, known as *biased competition*, has proven successful at explaining responses of neurons in the ventral visual stream and offers a basis for theorising about memory, attention, and learning (Chelazzi et al., 1998; Desimone and Duncan, 1995; Moran and Desimone, 1985). Our framework is, in essence, an extension of these ideas past the temporal pole, along the uncinate fasciculus, and into the orbital and medial regions of the prefrontal cortex. We are not the first to make this analogy. From the motor side, Cisek and colleagues have argued that biased competition principles also apply to representation and choice signals in motor and premotor regions (Cisek, 2007; Cisek and Kalaska, 2010; Pastor-Bernier and Cisek, 2011). We agree with this idea and propose that it extends backwards. One appealing feature of this idea is that it allows the brain to make use of a single principle to make both perceptual and economic decisions, rather than use a completely different architecture for the two types of choices.

The neuroeconomic binding problems. One virtue of our framework is that it offers a solution to three important neuroeconomic binding problems that are difficult to avoid with labelled line models. They concern how values are bound to options, to actions, and to choices (Akaishi and Hayden, 2016; Cai and Padoa-Schioppa, 2014; Hare et al., 2011; Strait et al., 2016). Values must be linked to their corresponding options (the *value binding* problem, (Akaishi and Hayden, 2016; Strait et al., 2016). Then, to select that option, we need to link the result of a comparison with the action that will be used to select it (the *action binding* problem, Cai and Padoa-Schioppa, 2014; Hare et al., 2011; Strait et al., 2016; Yoo et al., 2018). Finally, once the choice is resolved, we need to link the experienced value with the choice that produced it (the *outcome binding* problem). This is one example of the broader class known as credit assignment problems (Schultz, 2006; Sutton and Barto, 1998).

These binding problems can be understood by analogy to the feature binding problem (Engel and Singer, 2001; Shadlen and Movshon, 1999; Treisman and Gelade, 1980). Imagine seeing a red square and a blue circle; how does your brain know that it is not seeing a red circle and a blue square? Neuronal activity encoding each option dimension must be somehow coordinated. This coordination is unlikely to come through specialised neurons that are sensitive to any combination of multiple features – this would lead to a problem of a combinatorial explosion (Plaut and McClelland, 2010; Shadlen and Movshon, 1999; Von der Malsburg, 1981). One possibility is that this problem is solved by the degeneracy introduced by attention: if only one option is attended at a time, then the dimensions can be assumed to be related to the same single object. The same principle may apply for value as well.

These binding problems are difficult to solve in a labelled line system. If line labels are stable, our brains would need neurons for all possible options; this is unrealistic. If a new option is added to the mix, new neurons would have to be added. Would they be kept in reserve just in case a new option appears? What if 10 new options appear at once? This approach would require complex and specific wiring, ready to go for any possible choice. If the labels are assigned dynamically, then in situations with dozens of choices – such as when choosing cereals at the grocery store – we would need competition between dozens of neuron types. This approach would also require an as yet unidentified supervisory system to assign labels and implement the assignment. More importantly, it would not solve the binding problem: how would the decoders know which options had been assigned to which neurons? How could they know which action to perform to select the option? The costs of such coordination are daunting.

In our model, option identity and value/action/choice can be decoded by the principle of degeneracy: if there are only one option, value, and action within the focus of attention, then they can be assumed to correspond. Thus, in our framework, binding is implemented by attention and is determined solely by temporal context, not by stable labelled lines. By doing so, the potential combinatorial explosion is contained (Shadlen and Movshon, 1999; Von der Malsburg, 1981). Thus, in our view, the strict bottleneck imposed by limited attentional capacity is a feature, not a bug.

Methods

Description of the neuronal model

The dynamics of the firing rate and synaptic depression variables follow, respectively, the equations

$$\frac{d}{dt}r_i = -\frac{r_i}{\tau_r} + f(d_i I_i(t)) \quad (3)$$

$$\frac{d}{dt}d_i = \frac{1-d_i}{\tau_d} - \gamma_i d_i I_i(t) \quad (4)$$

where $f(x)$ is the rectified linear function (i.e. $f(x) = x$ if $x > 0$ and $f(x) = 0$ if $x \leq 0$) (Abbott, 1997; Tsodyks and Markram, 1997). The time constants for the firing rate and synaptic depression dynamics are chosen to be long, $\tau_r = 500$ ms and $\tau_d = 2$ s, to allow keeping a memory of the previous stimulus over the interval between stimuli presentation (Mongillo et al., 2008). Long firing rate effective time constants can also be obtained through recurrent dynamics. The firing rate of the neuron in equation (3) tracks the total input current with the time constant τ_r . The synaptic depression variable in equation (4) depresses whenever the synapse is strongly stimulated and recovers to its maximum value of one with time constant τ_d if there is no stimulation. How fast the synapse depresses depends on the value of the parameter γ_i .

Heterogeneity of the neuronal populations used in Figure 8 was generated by selecting at random from a normal distribution the tuning parameters a_i, b_i and the depression parameter γ_i (means: [0.5, 1, 1]; standard deviations: [0.05 0.1 0.1], respectively).

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

This work was supported by an R01 from NIH to BYH (DA037229) and grants BFU2017-85936-P and FLAGERA-PCIN-2015-162-C02-02 from MINECO (Spain) to RMB.

ORCID iD

Benjamin Y. Hayden  <https://orcid.org/0000-0002-7678-4281>

References

- Abbott LF (1997) Synaptic depression and cortical gain control. *Science* 275(5297): 221–224.
- Akaishi R and Hayden BY (2016) A spotlight on reward. *Neuron* 90(6): 1148–1150.
- Azab H and Hayden BY (2017) Correlates of decisional dynamics in the dorsal anterior cingulate cortex. *PLoS Biology*. Epub ahead of print 15 November. DOI: 10.1371/journal.pbio.2003091.
- Azab H and Hayden BY (2018) Correlates of economic decisions in the dorsal and subgenual anterior cingulate cortices. *European Journal of Neuroscience*. Epub ahead of print 12 February. DOI: 10.1111/ejn.13865.
- Barron HC, Dolan RJ and Behrens TEJ (2013) Online evaluation of novel choices by simultaneous representation of multiple memories. *Nature Neuroscience* 16(10): 1492–1498.
- Bartra O, McGuire JT and Kable JW (2013) The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage* 76: 412–427.
- Basten U, Biele G, Heekeren HR, et al. (2010) How the brain integrates costs and benefits during decision making. *Proceedings of the National Academy of Sciences of the United States of America* 107(50): 21767–21772.
- Baumann MA, Fluet MC and Scherberger H (2009) Context-specific grasp movement representation in the macaque anterior intraparietal area. *Journal of Neuroscience* 29(20): 6436–6448.
- Birrell JM and Brown VJ (2000) Medial frontal cortex mediates perceptual attentional set shifting in the rat. *Journal of Neuroscience* 20(11): 4320–4324.
- Blanchard TC and Hayden BY (2014) Neurons in dorsal anterior cingulate cortex signal postdecisional variables in a foraging task. *Journal of Neuroscience* 34(2): 646–655.
- Blanchard TC and Hayden BY (2015) Monkeys are more patient in a foraging task than in a standard intertemporal choice task. *PLoS ONE* 10(2): 1–11.
- Blanchard TC, Piantadosi ST and Hayden BY (2018) Robust mixture modeling reveals category-free selectivity in reward region neuronal ensembles. *Journal of Neurophysiology*. Epub ahead of print 6 December. DOI: 10.1152/jn.00808.2017.
- Blanchard TC, Hayden BY and Bromberg-Martin ES (2015a) Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. *Neuron* 85(3): 602–614.
- Blanchard TC, Strait CE and Hayden BY (2015b) Ramping ensemble activity in dorsal anterior cingulate neurons during persistent commitment to a decision. *Journal of Neurophysiology* 114(4): 2439–2449.
- Blanchard TC, Wilke A and Hayden BY (2014) Hot hand bias in rhesus monkeys. *Journal of Experimental Psychology: Animal Learning and Cognition* 40(3): 280–286.
- Boorman ED, Behrens TE and Rushworth MF (2011) Counterfactual choice and learning in a neural network centered on human lateral frontopolar cortex. *PLoS Biology* 9(6): e1001093.
- Boorman ED, Behrens TE, Woolrich MW, et al. (2009) How green is the grass on the other side? Frontopolar cortex and the evidence in favor of alternative courses of action. *Neuron* 62(5): 733–743.
- Boorman ED, Rushworth MF and Behrens TE (2013) Ventromedial prefrontal and anterior cingulate cortex adopt choice and default reference frames during sequential multi-alternative choice. *Journal of Neuroscience* 33(6): 2242–2253.
- Cai X and Padoa-Schioppa C (2014) Contributions of orbitofrontal and lateral prefrontal cortices to economic choice and the good-to-action transformation. *Neuron* 81(5): 1140–1151.
- Camille N, Griffiths CA, Vo K, et al. (2011) Ventromedial frontal lobe damage disrupts value maximization in humans. *Journal of Neuroscience* 31(20): 7527–7532.
- Carandini M and Heeger DJ (2012) Normalization as a canonical neural computation. *Nature Reviews. Neuroscience* 13(1): 51–62.
- Charnov EL (1976) Optimal foraging, the marginal value theorem. *Theoretical Population Biology* 9(2): 129–136.
- Chau BKH, Kolling N, Hunt LT, et al. (2014) A neural mechanism underlying failure of optimal choice with multiple alternatives. *Nature Neuroscience* 17(3): 463–470.
- Chelazzi L, Duncan J, Miller EK, et al. (1998) Responses of neurons in inferior temporal cortex during memory-guided visual search. *Journal of Neurophysiology* 80(6): 2918–2940.
- Chen X and Stuphorn V (2016) Sequential selection of economic good and action in medial frontal cortex of macaques during value-based decisions. *eLife* 4: e09418.
- Churchland AK, Kiani R and Shadlen MN (2008) Decision-making with multiple alternatives. *Nature Neuroscience* 11(6): 693–702.
- Cisek P (2006) Integrated neural processes for defining potential actions and deciding between them: A computational model. *Journal Neuroscience* 26(38): 9761–9770.

- Cisek P (2007) Cortical mechanisms of action selection: The affordance competition hypothesis. *Philosophical Transactions of the Royal Society of London B, Biological Sciences* 362(1485): 1585–1599.
- Cisek P (2012) Making decisions through a distributed consensus. *Current Opinion in Neurobiology* 22(6): 927–936.
- Cisek P and Kalaska JF (2005) Neural correlates of reaching decisions in dorsal premotor cortex: Specification of multiple direction choices and final selection of action. *Neuron* 45(5): 801–814.
- Cisek P and Kalaska JF (2010) Neural mechanisms for interacting with a world full of action choices. *Annual Review of Neuroscience* 33: 269–298.
- Desimone R and Duncan J (1995) Neural mechanisms of selective visual attention. *Annual Review of Neuroscience* 18: 193–222.
- Ebitz RB and Hayden BY (2016) Dorsal anterior cingulate: A Rorschach test for cognitive neuroscience. *Nature Neuroscience* 19(10): 1278–1279.
- Egeth HE and Yantis S (1997) Visual attention: Control, representation, and time course. *Annual Review of Psychology* 48: 269–297.
- Engel AK and Singer W (2001) Temporal binding and the neural correlates of sensory awareness. *Trends in Cognitive Sciences* 5(1): 16–25.
- FitzGerald TH, Seymour B and Dolan RJ (2009) The role of human orbitofrontal cortex in value comparison for incommensurable objects. *Journal of Neuroscience* 29: 8388–8395.
- Freidin E, Aw J and Kacelnik A (2009) Sequential and simultaneous choices: Testing the diet selection and sequential choice models. *Behavioural Processes* 80(3): 218–223.
- Fusi S, Miller EK and Rigotti M (2016) Why neurons mix: High dimensionality for higher cognition. *Current Opinion in Neurobiology* 37: 66–74.
- Haber SN and Behrens TEJ (2014) The neural network underlying incentive-based learning: Implications for interpreting circuit disruptions in psychiatric disorders. *Neuron* 83(5): 1019–1039.
- Hare TA, Schultz W, Camerer CF, et al. (2011) Transformation of stimulus value signals into motor commands during simple choice. *Proceedings of the National Academy of Sciences of the United States of America* 108(44): 18120–18125.
- Hayden BY (2016) Time discounting and time preference in animals: A critical review. *Psychonomic Bulletin & Review* 23: 39–53.
- Hayden BY (2018) Economic choice: The foraging perspective. *Current Opinion in Behavioral Sciences*; 24: 1–6.
- Hayden BY and Gallant JL (2013) Working memory and decision processes in visual area v4. *Frontiers in Neuroscience* 7: 18.
- Hayden BY, Heilbronner SR, Pearson JM, et al. (2011a) Surprise signals in anterior cingulate cortex: Neuronal encoding of unsigned reward prediction errors driving adjustment in behavior. *Journal of Neuroscience* 31(11): 4178–4187.
- Hayden BY, Pearson JM and Platt ML (2011b) Neuronal basis of sequential foraging decisions in a patchy environment. *Nature Neuroscience* 14(7): 933.
- Heilbronner SR and Hayden BY (2016) Dorsal anterior cingulate cortex: A bottom-up view. *Annual Review of Neuroscience* 39: 149–170.
- Hunt LT and Hayden BY (2017) A distributed, hierarchical and recurrent framework for reward-based choice. *Nature Reviews. Neuroscience* 18(3): 172–182.
- Hunt LT, Behrens TEJ, Hosokawa T, et al. (2015) Capturing the temporal evolution of choice across prefrontal cortex. *eLife* 4: 1–25.
- Hunt LT, Kolling N, Soltani A, et al. (2012) Mechanisms underlying cortical activity during value-guided choice. *Nature Neuroscience* 15(3): 470–476.
- Hunt LT, Woolrich MW, Rushworth MFS, et al. (2013) Trial-type dependent frames of reference for value comparison. *PLoS Computational Biology* 9(9): e1003225.
- Itti L and Koch C (2001) Computational modelling of visual attention. *Nature Reviews. Neuroscience* 2(2): 1–11.
- Jocham G, Hunt LT, Near J, et al. (2012) A mechanism for value-guided choice based on the excitation-inhibition balance in prefrontal cortex. *Nature Neuroscience* 15(7): 960–961.
- Jones JL, Esber GR, McDannald MA, et al. (2012) Orbitofrontal cortex supports behavior and learning using inferred but not cached values. *Science* 338(6109): 953–956.
- Kacelnik A, Vasconcelos M, Monteiro T, et al. (2011) Darwin's 'tug-of-war' vs. starlings' 'horse-racing': How adaptations for sequential encounters drive simultaneous choice. *Behavioral Ecology and Sociobiology* 56(3): 547–558.
- Kennerley SW, Behrens TE and Wallis JD (2011) Double dissociation of value computations in orbitofrontal and anterior cingulate neurons. *Nature Neuroscience* 14(12): 1581.
- Kennerley SW, Walton ME, Behrens TE, et al. (2006) Optimal decision making and the anterior cingulate cortex. *Nature Neuroscience* 9(7): 940.
- Kolling N, Behrens TE, Mars RB, et al. (2012) Neural mechanisms of foraging. *Science* 336(6077): 95.
- Kolling N, Wittmann M and Rushworth MF (2014) Multiple neural mechanisms of decision making and their competition under changing risk pressure. *Neuron* 81(5): 811190–811202.
- Krajbich I and Rangel A (2011) Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences of the United States of America* 108(33): 13852.
- Krajbich I, Armel C and Rangel A (2010) Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience* 13(10): 1292.
- Krebs JR, Erichsen JT, Webber MI, et al. (1977) Optimal prey selection in the great tit (*Parus major*). *Animal Behaviour* 25(1): 30–38.
- Lim SL, O'Doherty JP and Rangel A (2011) The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. *Journal of Neuroscience* 31(37): 13214–13223.
- Louie K, Gratton LE and Glimcher PW (2011) Reward value-based gain control: Divisive normalization in parietal cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 31(29): 10627–10639.
- Lui LL and Pasternak T (2011) Representation of comparison signals in cortical area MT during a delayed direction discrimination task. *Journal of Neurophysiology* 106(3): 1260–1273.
- McGinty VB, Rangel A and Newsome WT (2016). Orbitofrontal cortex value signals depend on fixation location during free viewing. *Neuron* 90(6): 1299–1311.
- Machens CK, Romo R and Brody CD (2005) Flexible control of mutual inhibition: A neural model of two-interval discrimination. *Science* 307(5712): 1121–1124.
- McPeck RM and Keller EL (2002) Superior colliculus activity related to concurrent processing of saccade goals in a visual search task. *Journal of Neurophysiology* 87(4): 1805–1815.
- Miller EK and Desimone R (1994) Parallel neuronal mechanisms for short-term memory. *Science* 263(5146): 520–522.
- Miller EK, Li L and Desimone R (1991) A neural mechanism for working and recognition memory in inferior temporal cortex. *Science* 29(5036): 1377–1379.
- Mirabella G, Bertini G, Samengo I, et al. (2007) Neurons in area V4 of the macaque translate attended visual features into behaviorally relevant categories. *Neuron* 54(2): 303–318.
- Mongillo G, Barak O and Tsodyks M (2008) Synaptic theory of working memory. *Science* 319(5869): 1543–1546.
- Moran J and Desimone R (1985) Selective attention gates visual processing in the extrastriate cortex. *Science* 229(4715): 782–784.
- Moreno-Bote R, Beck J, Kanitscheider I, et al. (2014) Information-limiting correlations. *Nature Neuroscience* 17(10): 1410–1417.
- Nogueira R, Abolafia JM, Drugowitsch J, et al. (2017) Lateral orbitofrontal cortex anticipates choices and integrates prior with current information. *Nature Communications* 8: 14823.

- Noonan MP, Walton ME, Behrens TE, et al. (2010.) Separate value comparison and learning mechanisms in macaque medial and lateral orbitofrontal cortex. *Proceedings of the National Academy of Sciences of the United States of America* 107(47): 20547–20552.
- Orquin JL and Mueller Loose S (2013) Attention and choice: A review on eye movements in decision making. *Acta Psychologica* 144(1): 190–206.
- Padoa-Schioppa C (2009) Range-adapting representation of economic value in the orbitofrontal cortex. *Journal of Neuroscience* 29(44): 14004–14014.
- Padoa-Schioppa C (2011) Neurobiology of economic choice: A good-based model. *Annual Review of Neuroscience* 34: 333–359.
- Pagan M, Urban LS, Wohl MP, et al. (2013) Signals in inferotemporal and perirhinal cortex suggest an untangling of visual target information. *Natural Neuroscience* 16(8): 1132–1139.
- Pastor-Bernier A and Cisek P (2011) Neural correlates of biased competition in premotor cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 31(19): 7083–7088.
- Pearson JM, Watson KK and Platt ML (2014) Decision making: The neurotheological turn. *Neuron* 82(5): 950–965.
- Plaut DC and McClelland JL (2010) Locating object knowledge in the brain: Comment on Bowers's (2009) attempt to revive the grandmother cell hypothesis. *Psychological Review* 117(1): 284–288.
- Rangel A and Hare T (2010) Neural computations associated with goal-directed choice. *Current Opinion in Neurobiology* 20(2): 262–270.
- Raposo D, Kaufman MT and Churchland AK (2014) A category-free neural population supports evolving demands during decision-making. *Nature Neuroscience* 17(12): 1784–1792.
- Reynolds JH and Heeger DJ (2009) The normalization model of attention. *Neuron* 61(2): 168–185.
- Rich EL and Wallis JD (2016) Decoding subjective decisions from orbitofrontal cortex. *Nature Neuroscience* 19(7): 973–980.
- Rigotti M, Barak O, Warden MR, et al. (2013) The importance of mixed selectivity in complex cognitive tasks. *Nature* 497(7451): 585–590.
- Romo R and Salinas E (2001) Touch and go: Decision-making mechanisms in somatosensation. *Annual Review of Neuroscience* 24: 107–137.
- Romo R and Salinas E (2003) Flutter discrimination: Neural codes, perception, memory and decision making. *Nature Reviews. Neuroscience* 4(3): 203–218.
- Romo R, Hernandez A, Zainos A, et al. (2002) Neuronal correlates of decision-making in secondary somatosensory cortex. *Nature Neuroscience* 5(9): 1217–1225.
- Rudebeck PH, Mitz AR, Chacko RV, et al. (2013) Effects of amygdala lesions on reward-value coding in orbital and medial prefrontal cortex. *Neuron* 80(6): 1519–1531.
- Rushworth MF, Noonan MP, Boorman ED, et al. (2011) Frontal cortex and reward-guided learning and decision-making. *Neuron* 70(6): 1054–1069.
- Rustichini A and Padoa-Schioppa C (2015) A neuro-computational model of economic decisions. *Journal of Neurophysiology* 114(3): 1382–1398.
- Scherberger H and Andersen RA (2007) Target selection signals for arm reaching in the posterior parietal cortex. *Journal of Neuroscience* 27(8): 2001–2012.
- Schoenbaum G, Chiba AA and Gallagher M (1998) Orbitofrontal cortex and basolateral amygdala encode expected outcomes during learning. *Nature Neuroscience* 1(2): 155–159.
- Schultz W (2006) Behavioral theories and the neurophysiology of reward. *Annual Review of Psychology* 57: 87–115.
- Shadlen MN and Kiani R (2013) Decision making as a window on cognition. *Neuron* 80(3): 791–806.
- Shadlen MN and Movshon JA (1999) Synchrony unbound: A critical evaluation of the temporal binding hypothesis. *Neuron* 24(1): 67–77, 111.
- Shadlen MN and Shohamy D (2016) Decision making and sequential sampling from memory. *Neuron* 90(5): 927–939.
- Shapiro MS, Siller S and Kacelnik A (2008) Simultaneous and sequential choice as a function of reward delay and magnitude: Normative, descriptive and process-based models tested in the European starling (*Sturnus vulgaris*). *Journal of Experimental Psychology. Animal Behavior Processes* 34(1): 75–93.
- Sleezer BJ and Hayden BY (2016.) Differential contributions of ventral and dorsal striatum to early and late phases of cognitive set reconfiguration. *Journal of Cognitive Neuroscience* 28(12): 1–16.
- Sleezer BJ, Loconte G, Castagno MD, et al. (2017.) Neuronal responses support a role for orbitofrontal cortex in cognitive set reconfiguration. *The European Journal of Neuroscience* 45(7): 42–49.
- Soltani A, Lee D and Wang X-J (2006) Neural mechanism for stochastic behaviour during a competitive game. *Neural Networks: The Official Journal of the International Neural Network Society* 19(8): 1075–1090.
- Stephens DW and Krebs JR (1986) *Foraging Theory*. Princeton, NJ: Princeton University Press.
- Strait CE, Blanchard TC and Hayden BY (2014) Reward value comparison via mutual inhibition in ventromedial prefrontal cortex. *Neuron* 82(6): 1–10.
- Strait CE, Sleezer BJ and Hayden BY (2015) Signatures of value comparison in ventral striatum neurons. *PLoS Biology* 13(6): e1002173.
- Strait CE, Sleezer BJ, Blanchard TC, et al. (2016) Neuronal selectivity for spatial positions of offers and choices in five reward regions. *Journal of Neurophysiology* 115(3): 1098–1111.
- Sutton RS and Barto AG (1998) *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.
- Thura D and Cisek P (2014.) Deliberation and commitment in the premotor and primary motor cortex during dynamic decision making. *Neuron* 81(6): 1401–1416.
- Treisman AM and Gelade G (1980) A feature-integration theory of attention. *Cognitive Psychology* 12(1): 97–136.
- Tsodyks M and Markram H (1997) The neural code between neocortical pyramidal neurons depends. *Proceedings of the National Academy of Sciences* 94(2): 719–723.
- Von Der Malsburg C (1981) The correlation theory of brain function: MPI biophysical chemistry, Internal Report 81-2. In: Domany E, Van Hemmen JL and Schulten K (eds) *Models of Neural Networks II (1994)* (pp. 95–119). Berlin: Springer.
- Wallis JD (2007) Orbitofrontal cortex and its contribution to decision-making. *Annual Review of Neuroscience* 30: 31–56.
- Wang MZ and Hayden BY (2017) Reactivation of associative structure specific outcome responses during prospective evaluation in reward-based choices. *Nature Communications* 8: 15821.
- Xie Y, Nie C and Yang T (2017) Covert shift of attention modulates the value encoding in the orbitofrontal cortex. *BioRxiv*. Epub ahead of print 29 August. DOI: 10.1101/181784.
- Yamada H, Louie K, Tymula A, et al. (2018) Free choice shapes normalized value signals in medial orbitofrontal cortex. *Nature Communications* 9: 162.
- Yoo SBM, Sleezer BJ and Hayden BY (2018) Robust encoding of spatial information in orbitofrontal cortex and striatum. *Journal of Cognitive Neuroscience* (in press).