CellPress

# For Better or Worse: Reward Comparison by the Ventromedial Prefrontal Cortex

Linda M. Amarante[1] and Mark Laubach[1,*]
[1]The John B. Pierce Laboratory and Department of Neurobiology, Yale School of Medicine, New Haven, CT 06510, USA
*Correspondence: mark.laubach@yale.edu
http://dx.doi.org/10.1016/j.neuron.2014.05.046

**In this issue of *Neuron*, Strait et al. (2014) studied how neurons in the monkey ventromedial prefrontal cortex encode value-based decisions. Neurons were commonly influenced by reward magnitude and probability, showed anticorrelation for better and worse options, and covaried with choice independent of value.**

Value-based decisions are made when we choose who to date, which house to buy, or which food to eat. A large number of neuroimaging studies of these kinds of decisions have been published over the past decade. These studies have established that the ventromedial prefrontal cortex (vmPFC) is a key brain area for comparing the value of different options (Rushworth et al., 2011; Levy and Glimcher, 2012). In parallel, theoretical studies have led to the idea that choice is based on competition between neural representations of the available options (Hare et al., 2011; Hunt et al., 2012). These ideas are based on models of perceptual decision making (Gold and Shadlen, 2007), which have been very successful in unraveling the mechanisms of stimulus-based decisions.

In this issue of *Neuron*, Strait et al. (2014) examined value comparison and representation in monkey vmPFC. Two male rhesus macaques were offered options based on the size of a reward (juice) and the probability of receiving that reward in a two-option gambling task. The two offers were presented asynchronously between short delays, as to distinguish value signals from signals of comparison or choice (Figure 1A). Monkeys were given either a medium or large (and occasionally small) reward based on their eye fixation to visual stimuli of colored rectangles representing reward size and probability on a computer monitor.

Strait et al. (2014) recorded single-unit responses from the vmPFC (area 14) in two monkeys as the animals sought to maximize reward by choosing the offer with a higher expected value. Firing rates for a small but significant number of neurons encoded both reward size and probability, and there was a significant positive correlation in regression coefficients between probability and reward size. Similar to previous studies (Padoa-Schioppa, 2011), which have claimed that values are stored in a common scale (or "currency") for comparison in the vmPFC and the adjacent orbitofrontal cortex, these results strengthen the idea that value is indeed represented by single neurons in the vmPFC in a common scale so that comparison of values can occur across competing pools of neurons.

Though both values were coded simultaneously, a regression analysis of the neurons' firing rates showed evidence for "anticorrelated tuning" between offer values. That is, there was a negative correlation between the regression coefficients for neurons associated with each offer. This finding suggests that the encodings of the different offers act to mutually inhibit each other (Figure 1B). In other words, separate pools of neurons end up inhibiting each other so that only one pool of neurons fires a sufficient number of spikes to generate the animal's choice. There was also a significant correlation between neural activity and the chosen value: more neurons quickly came to signal the chosen offer value after the presentation of both offers. This finding suggests that neurons in vmPFC participate in making a choice. After regressing out effects of value, probability, and reward size, the firing rates of the vmPFC neurons still encoded the animal's choice. A choice probability signal was thus also found in the vmPFC. Based on these findings, Strait et al. (2014) suggest that vmPFC does contribute to choice through mutual inhibition of value representations (Figure 1C).

Additionally, vmPFC neurons encoded outcome values. A significant amount of neurons, more than those that encoded relative value or choice probability, had a relationship between firing rate and gamble outcome. Outcome coding was evident during the delay between trials, and the previous trial outcome had a significant influence on firing rates during both offer representations. Similar to previous studies on the dorsomedial PFC (Narayanan and Laubach, 2008; Hayden et al., 2011), vmPFC neuron responses serve as outcome monitoring signals. Inactivating this region in the rodent brain leads to a disruption of spike and field potential signals associated with performance adjustments (Narayanan et al., 2013). It would be interesting to know whether the spike correlates of value-based decisions reported by Strait et al. (2014) were associated with similar network activity that directly predicts performance adjustments in rats and humans (Narayanan et al., 2013).

The findings of Strait et al. (2014) validate previous fMRI and MEG/EEG studies that have implicated vmPFC in value-based decisions (Rushworth et al., 2011; Levy and Glimcher, 2012) and theoretical studies of value-based decision making (Hare et al., 2011; Hunt et al., 2012). It also raises many mechanistic questions that should be addressed in future studies. First, classic methods for single-unit recording were used by Strait et al. (2014). It would be fascinating to know how simultaneously recorded groups of neurons encode the values of the options

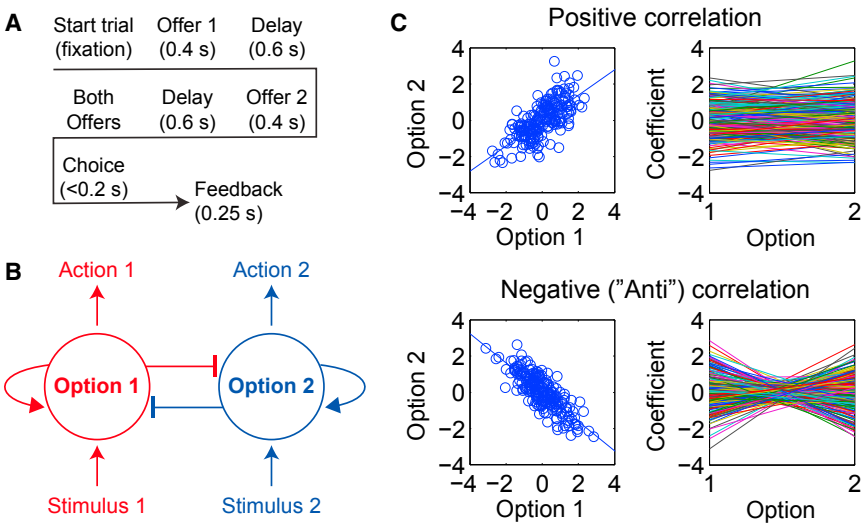


**Figure 1. Neural Mechanisms of Reward Comparison by the Ventromedial Prefrontal Cortex**

(A) Strait et al. (2014) evaluated the neural coding of value-based decisions using a delayed response task. Monkeys initiated trials by fixating near the center of a computer display. Stimuli were presented on the right or left side of the display during two "offer" epochs. The stimuli indicated the relative value of responding to the right or left when the two offers were simultaneously presented and the monkeys were asked to make a choice. Feedback (reward) was given and varied in magnitude and probability.

(B) Value-based decisions are thought to result from competition between pools of neurons that track the value of each option. A schematic for such a "mutual inhibition" model (Hare et al., 2011; Hunt et al., 2012) is shown. Monkeys choose the best option based on the level of activity in the two pools of recurrently excitatory neurons, and each pool is interconnected by inhibitory neurons.

(C) Regression analysis was used to examine how vmPFC neurons encoded the offers. Regression coefficients showed positive correlation for larger magnitude and higher probability reward. This finding is evidence for a "common currency" coding scheme in the vmPFC. Regression coefficients showed negative correlations (anticorrelation) over options during the second offer epoch, a finding that is expected by the mutual inhibition model. Examples of such positive and negative correlations (based on simulated data, with three times the correlation strength as reported by Strait et al, 2014) are shown in the figure as scatterplots (left) and parallel coordinate plots (right).

and the choices made by the participants. For example, if magnitude and probability are encoded as a common currency, then do specific groups of neurons track one or both of these measures? Could the activity of these cells be used to predict (decode) stimulus value and/or choice on a trial-by-trial basis?

Second, the idea that value-based decisions are embodied by a neural circuit in which there is competition between pools of neurons selective to each option should lead to functional interactions among groups of simultaneously recorded neurons. Would neurons that track the options exhibit sufficient covariance such that they would be revealed on a specific measure of population activity, such as a principal component? Are the cells that take part in these computations pyramidal neurons or interneurons? What are the dynamics of these cell classes during a typical trial? Could perturbation methods be used to disrupt processing by different classes of cells and would such perturbations impair behavioral performance?

Third, understanding the mechanisms of value-based decisions in the vmPFC will require resolving the key inputs and outputs of this cortical region. For example, value-based decisions are thought to depend on encoding the subjective values of the task stimuli (options) and the context in which the stimuli are presented. Several inputs to the vmPFC that might mediate these encodings would seem to be the hippocampus, amygdala, and insular cortex (Ongür and Price, 2000). On the output side, vmPFC, across species, projects to dopamine neurons in the midbrain, feeding related centers in the hypothalamus, and the ventral striatum (Ongür and Price, 2000). We suggest that future studies should examine the roles of these input and output structures in the encoding of subjective value, behavioral context, and action selection.

Resolving these issues would require the use of methods such as optogenetics, which have not been fully worked out for routine use in primates. Rapid progress on addressing the neuronal mechanisms of value-based decisions might thus require the use of rodent models, and a major outstanding question is whether the vmPFC area studied by Strait et al. (2014) is comparable to medial frontal areas in rodents. These medial areas in rodents and primates share many anatomical connections (Ongür and Price, 2000; Hoover and Vertes, 2007, 2011) and might be a common cortical region found in all mammals (Laubach, 2011).

A final issue that we would like to raise is that an alternative perspective on value-based decision making has been developed in the field of behavioral ecology (Kacelnik et al., 2011). A model of these decisions by birds (starlings) was proposed, the sequential choice model, that does not involve inhibition between cells encoding the available options. If these models were applied to the study by Strait et al. (2014), then would they explain the results as well as the competition ("tug-of-war") models that have been prominent in human studies on value-based decisions? Unfortunately, the behavioral design used in the Strait study precluded measurement of precise RTs for each option. It would be interesting to run the task with the stimuli presented at a common time point before the choice to determine whether the simpler "independent horse race" models could account for value-based choice.

## REFERENCES

Gold, J.I., and Shadlen, M.N. (2007). Annu. Rev. Neurosci. *30*, 535–574.

Hare, T.A., Schultz, W., Camerer, C.F., O'Doherty, J.P., and Rangel, A. (2011). Proc. Natl. Acad. Sci. USA *108*, 18120–18125.

Hayden, B.Y., Heilbronner, S.R., Pearson, J.M., and Platt, M.L. (2011). J. Neurosci. *31*, 4178–4187.

Hoover, W.B., and Vertes, R.P. (2007). Brain Struct. Funct. *212*, 149–179.

Hoover, W.B., and Vertes, R.P. (2011). J. Comp. Neurol. *519*, 3766–3801.

Hunt, L.T., Kolling, N., Soltani, A., Woolrich, M.W., Rushworth, M.F., and Behrens, T.E. (2012). Nat. Neurosci. *15*, 470–476, S1–S3.

Kacelnik, A., Vasconcelos, M., Monteiro, T., and Aw, J. (2011). Behav. Ecol. Sociobiol. *65*, 547–558.

Laubach, M. (2011). A comparative perspective on executive and motivational control by the medial prefrontal cortex. In Neural Basis of Motivational and Cognitive Control, R.B. Mars, J. Sallet, M.F.S. Rushworth, and N. Yeung, eds. (Cambridge: MIT Press), pp. 95–110.

Levy, D.J., and Glimcher, P.W. (2012). Curr. Opin. Neurobiol. *22*, 1027–1038.

Narayanan, N.S., and Laubach, M. (2008). J. Neurophysiol. *100*, 520–525.

Narayanan, N.S., Cavanagh, J.F., Frank, M.J., and Laubach, M. (2013). Nat. Neurosci. *16*, 1888–1895.

Ongür, D., and Price, J.L. (2000). Cereb. Cortex *10*, 206–219.

Padoa-Schioppa, C. (2011). Annu. Rev. Neurosci. *34*, 333–359.

Rushworth, M.F., Noonan, M.P., Boorman, E.D., Walton, M.E., and Behrens, T.E. (2011). Neuron *70*, 1054–1069.

Strait, C.E., Blanchard, T.C., and Hayden, B.Y. (2014). Neuron *82*, this issue, 1357–1366.

# Bringing the Dynamics of Movement under Control

**Alfonso Renart**[1,*]
[1]Champalimaud Neuroscience Programme, Champalimaud Centre for the Unknown, 1400-038 Lisbon, Portugal
*Correspondence: alfonso.renart@neuro.fchampalimaud.org
http://dx.doi.org/10.1016/j.neuron.2014.06.002

**The mechanisms underlying the dynamics of movement-related neural activity are not known. In this issue of *Neuron*, Hennequin et al. (2014) show that a recurrent network whose spontaneous activity is stabilized by learning reproduces many aspects of preparatory and movement-related activity.**

Recurrent neural networks (RNNs), such as the ones generally used to model cortical circuits, are complex dynamical systems expected to display a rich repertoire of dynamical behaviors. However, designing networks capable of generating dynamics useful for temporal computation has proven to be surprisingly difficult.

A generically important example of temporal computation is the production of specific smooth temporal sequences of moderate dimensionality, such as the set of muscle activations giving rise to movements. In this issue of *Neuron*, Hennequin et al. (2014) provide a new and elegant solution to this problem, which also introduces new ideas about stabilization of network activity and about the nature of evoked and spontaneous network activity.

Most early work on recurrent networks focused on static phenomena, such as fixed-point or line attractors or the generation of different types of selectivity to sensory inputs. Over the last 10 years or so, research on how RNNs could be used to generate interesting time-varying activity has flourished. Interest in this question has been motivated in part by recent results suggesting that the temporal dynamics of neural activity in motor and premotor areas might be key to understanding how movement is generated (Shenoy et al., 2013).

Recurrent networks can be broadly construed as selective amplifiers, which constantly amplify or suppress spatiotemporal activity patterns, either externally driven or internally generated. RNNs display a tradeoff between amplification and stability, because loops of excitatory and inhibitory pathways within the network can be a source of positive feedback. In general, the potential for selective amplification increases with the magnitude of the synaptic connections in the network, but large synaptic weights also increase the potential for runaway excitation and other forms of instability. When the typical magnitude of the synaptic interactions in randomly connected networks goes beyond a certain value, the resulting instability turns these networks chaotic. In the chaotic state, networks are extremely sensitive, and minute perturbations get amplified leading to wildly diverging patterns of network activity.

Recent studies have used chaotic states as the starting point for generating structured time-varying activity patterns, a connection that has been most thoroughly established in networks of firing-rate units. In seminal studies, it was shown that chaos in these networks can be suppressed by temporally structured external input (Rajan et al., 2010) and that this feature can be exploited by supervised learning algorithms to train RNNs to produce nonchaotic patterns of activity, which can be linearly read out to produce a wide variety of desired time-varying outputs (Sussillo and Abbott, 2009; Laje and Buonomano, 2013). In these studies, chaotic ongoing activity in the RNN was stabilized, leading to temporally structured ongoing activity. The activity of single neurons after learning becomes more reliable, but it is otherwise similar as before learning, with strong sustained temporal fluctuations of similar magnitude. This is in contrast to what we typically think of as an evoked response, which is transient. For instance, the magnitude of the temporal fluctuations in short-term firing rate in a movement-responsive neuron is expected to increase during movement compared to baseline. How can these transient time-varying patterns be generated?