A Spotlight on Reward

Rei Akaishi[1],* and Benjamin Y. Hayden[1]
[1]Department of Brain and Cognitive Sciences and Center for Visual Science, University of Rochester, Rochester, NY 14627, USA
*Correspondence: akaishirei@gmail.com
http://dx.doi.org/10.1016/j.neuron.2016.06.008

Research published in this issue of *Neuron* from McGinty et al. (2016) suggests that attention may help bind information about value to specific options in economic choice. Responses of orbitofrontal neurons are strongly modulated by the distance from gaze to the position of a reward-predicting target.

Choosing between options based on their values, a mental operation that is known as economic or reward-based choice, is a core element of psychology, economics, and ethology. Economic choice is also important for the etiology of mental disorders: aberrant evaluation and comparison are hallmarks of addiction, depression, obsessive-compulsive disorder, and other diseases. Consequently, understanding the neural computations that determine the outcome of economic choice is one of the most beguiling but important outstanding problems in decision neuroscience.

Over the past 20 years, neuroscientists have successfully identified the core brain regions involved in economic choice and have delineated their anatomical connections (Rushworth et al., 2011; Padoa-Schioppa, 2011). Among these reward regions, the orbitofrontal cortex (OFC) has attracted perhaps the greatest scholarly interest (Padoa-Schioppa, 2011; Rushworth et al., 2011; Wallis, 2007). Neurons throughout both primate and rodent OFC respond to reward and offers of reward and signal aspects of comparison.

Scientists have begun to tentatively piece together a basic mechanistic understanding of how choice occurs. Across many studies, it appears that neurons in several reward-sensitive regions compute and signal, through modulations in their firing rates, the values of available options in an *evaluation stage*. These value signals are then fed into a *comparison stage* (which may involve the same neurons and may occur simultaneously) that implements competition through mutual inhibition; this competition probably occurs in multiple areas simultaneously (Strait et al., 2016; Hunt et al., 2015).

The Selectivity Problem

Choice models must deal with a binding problem, which we will call the Selectivity Problem (Strait et al., 2016). If a neuron's firing rate signals the value of one option, how does the comparator know which option its firing rate refers to? The neuron's signal must somehow be *bound* to a particular option. Likewise, once the preferred option is identified, the brain must bind the result of the comparison with a specific action that will select it. Binding is also required after choice, during learning, to link a choice with its rewarding result. The selectivity problem is an example of a psychological binding problem—akin to the famous feature-binding problem of attention (Treisman and Gelade, 1980).

One simple solution to the Selectivity Problem would be to use a labeled line—each neuron could be designated for a particular option (Padoa-Schioppa, 2011; Hunt et al., 2015). Then comparing values would involve a mutual inhibition between discrete sets of neurons, each set representing the values of its corresponding option. There are some problems with a labeled line approach, however. First, each neuron would have to be stably assigned to a particular option and maintain that relationship. The relationship between the neuron and its corresponding option would have to be known by the brain region reading out these signals. In the case that there are more than two options, such as when choosing among 50 cereals at the grocery store, the problem becomes computationally unwieldy, both in evaluation and in comparison of options. One solution to this problem that works for some perceptual decisions, a spatial reference frame, seems difficult to apply to abstract decisions (decisions without an obvious associated action), like choosing a political candidate or a flavor of ice cream.

Some recent studies suggest an alternative solution to the Selectivity Problem: *attention*. The idea is that a neuron's firing rate refers to the relative value of the attended option (Lim et al., 2011, Krajbich et al., 2010, Strait et al., 2015, Blanchard et al., 2015). (Relative means that value is in proportion to the value of its alternatives.) This stratagem solves the binding problem because there is no doubt about what stimulus a neuron's firing rate refers to: it refers to the attended one at any moment. Likewise, the action associated with that option is the one prepared if that option is attended.

Direct evidence for this idea comes from a neuroimaging study showing that BOLD responses in vmPFC in a choice task are aligned to the value of the fixated option relative to the value of the unfixated one (Lim et al., 2011). Similar results are observed in single neurons when options are presented asynchronously. When the first option appears, neuronal responses in OFC, in ventromedial area 14 (vmPFC), and in ventral striatum all encode its value and then, when the alternative appears, they encode the value of the alternative, relative to the first (Strait et al., 2015, Blanchard et al., 2015).

Another recently published study of the OFC (Rich and Wallis, 2016) further strengthens the view that attention plays an important role in the decision process. In that study, monkeys chose between two options and the authors used linear discriminant analysis to decode representations of the offers. They found that ensemble responses in OFC oscillated between two discrete states corresponding to the two offers and gradually converged on the preferred option. The results are reminiscent of what one would
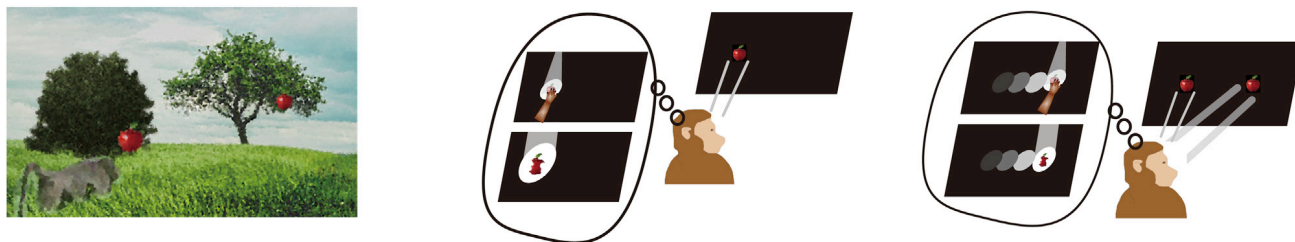
Figure 1. Possible Roles of Attentional Spotlights in Foraging and Binomial Choices
(Left) Foragers generally encounter a single option at a time and thus may be cognitively specialized for making accept-reject decisions.
(Middle) Similar to foraging situations in the natural environment, the subject in the report of McGinty and his colleagues encountered the reward predictive object one at a time. The object, its location, and the outcome (and the action) associated with it are all linked by attention (illuminated by a spotlight of attention). This possibility predicts a strong attentional modulation of the value signal.
(Right) Even animals forced by scientists to make binomial choices may instead make a series of (potentially interacting) accept-reject decisions, where the option under consideration is the attended one. The attentional focus shifts between options probably with some trace of the previously attended option.

imagine the spotlight of attention does, alternating between the two offers as the mind's eye considers each in sequence. Like the earlier studies, neurons' firing rates were not stably aligned to individual offers but instead were consistently aligned to the attended offer.

One interesting implication of these results is that we process value serially, not in parallel. In other words, we make a series of accept-reject decisions. Why would we do this? Presumably, value computation is somehow costly and is limited by the same type of bottleneck that limits other forms of attention. As a consequence, then, we may only consider one option at any given time. And normally, the option we are considering is the one we are looking at. So our gaze should tell us which option is attended, and which value is represented. A direct relationship between gaze and value signals would therefore provide important evidence in favor of the binding-by-attention hypothesis.

New Data Supporting the Relationship between Attention and Reward

A study in this issue of *Neuron* provides critical evidence in favor of the idea that attention is critically intertwined with value representation. McGinty and colleagues recorded activity of single neurons in the Area 13/OFC of monkeys performing a simple task. Monkeys first fixated at a cue that predicted the size of an upcoming reward, and, four seconds later, they got that reward. The key feature of the task is that monkeys were allowed to look wherever they wanted during that

four-second waiting period. The critical finding was that, as monkeys looked around the screen, firing rates were consistently modulated by the distance from the center of gaze to the cue.

Gaze was not just a weak modulator of a value signal; instead, the distance from the reward-predicting cue was represented as strongly as value was. Nor was gaze encoding the property of a specialized set of neurons; it was strongest in the same neurons that encoded value. These findings are particularly surprising given that OFC is often seen as a pure reward area.

This result is important because gaze is an excellent measure of the locus of attention. These results therefore confirm, for the first time, a direct link between value coding in OFC and the status of gaze/attention in a freely viewing animal.

The type of coding the authors observed was very different from the types of gaze encoding observed in other brain areas like the parietal lobe. It was not spatially selective, meaning it did not depend on whether the animal's gaze was to the left or the right of the cue. It was not lateralized or contralaterally enhanced. It did not depend on angle of gaze or any other ocular metric. OFC is not just a noisy version of LIP. Instead, the most parsimonious theory is that its firing rate reflects the value of the option modulated by the amount of attention assigned to it.

Are Binomial Choices Actually Accept-Reject Problems?

What does it mean if we evaluate options one at a time? One implication is that we

may not so much choose between options as consider each option by itself, decide whether to accept or reject it, and then move on (and often, back to the first option). This serial view of choice is foreign in economics but is quite natural in foraging theory (Stephens and Krebs, 1986). Behavioral ecologists have long noted that foraging animals seldom encounter two options at a time, but instead almost always encounter one at a time and then decide, on their own, whether to pursue or ignore it (Figure 1, left panel). It may even be the case that the psychological—and underlying neural—mechanisms of choice are designed for single-offer accept-reject type decisions. In other words, accept-reject decisions are fundamentally decisions about whether to engage—to choose to act—or to refrain from action—to not choose that option. While we can handle simultaneous encounters—we are flexible, after all—we do it in a way that reflects our basic cognitive adaptations for sequential encounters.

This possibility accounts for a remarkable feature of McGinty et al. (2016)'s study (Figure 1, middle panel). They found classic OFC reward-related responses in the absence of any economic choice task. Indeed, monkeys in their study were highly trained, meaning they knew there was no possibility they might have to eventually choose. It seems that an overt choice context is not critical for OFC to serve its value-computing role.

This idea also raises another interesting possibility. If the relative value of an option is high enough, a decision maker should accept the offer, and, if it is not, the

decision maker should reject it. If value representations are formatted to the attended option, then the ensemble firing rate of OFC can be said to embody the evidence in favor of selecting the attended option (Rich and Wallis 2016; Strait et al., 2015). And if that's the case, then economic choices are fundamentally a series of stopping problems. Or, more precisely, a series of *accept-reject problems* (Figure 1, right panel). This possibility has the potential to unify models of economic choice with models of other forms of choice to form more a general understanding of decision-making in the brain.

REFERENCES

Blanchard, T.C., Hayden, B.Y., and Bromberg-Martin, E.S. (2015). Neuron 85, 602–614.

Treisman, A.M., and Gelade, G. (1980). Cognit. Psychol. 12, 97–136.

Hunt, L.T., Behrens, T.E., Hosokawa, T., Wallis, J.D., and Kennerley, S.W. (2015). eLife 4, e11945.

Krajbich, I., Armel, C., and Rangel, A. (2010). Nat. Neurosci. 13, 1292–1298, http://dx.doi.org/10.1038/nn.2635.

Lim, S.L., O'Doherty, J.P., and Rangel, A. (2011). J. Neurosci. 31, 13214–13223.

McGinty, V., Rangel, A., and Newsome, W. (2016). Neuron 90, this issue, 1299–1311.

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

Rich, E., and Wallis, J.D. (2016). Nat. Neurosci., in press.

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

Stephens, D.W., and Krebs, J.R. (1986). Foraging theory (Princeton University Press).

Strait, C.E., Sleezer, B.J., Blanchard, T.C., Azab, H., Castagno, M.D., and Hayden, B.Y. (2016). J. Neurophysiol. 115, 1098–1111.

Strait, C.E., Sleezer, B.J., and Hayden, B.Y. (2015). PLoS Biol. 13, e1002173.

Wallis, J.D. (2007). Annu. Rev. Neurosci. 30, 31–56.