

## Journal Club

**Editor's Note:** These short, critical reviews of recent papers in the *Journal*, written exclusively by graduate students or postdoctoral fellows, are intended to summarize the important findings of the paper and provide additional insight and commentary. For more information on the format and purpose of the Journal Club, please see [http://www.jneurosci.org/misc/ifa\\_features.shtml](http://www.jneurosci.org/misc/ifa_features.shtml).

## Attaching Values to Actions: Action and Outcome Encoding in the Primate Caudate Nucleus

Christopher H. Donahue<sup>1</sup> and Hyojung Seo<sup>2</sup>

<sup>1</sup>Interdepartmental Neuroscience Program and <sup>2</sup>Department of Neurobiology, Yale University School of Medicine, New Haven, Connecticut 06520

Review of Lau and Glimcher (<http://www.jneurosci.org/cgi/content/full/27/52/14502>)

To make effective decisions while navigating uncertain environments, animals must develop the ability to accurately predict the consequences of their actions. Reinforcement learning has emerged as a key theoretical paradigm for understanding how animals accomplish this feat (Sutton and Barto, 1998). According to this framework, animals develop decision-making strategies through an iterative trial-and-error process. First, an action is selected based on a prediction of which choice will lead to the greatest payoff. After an action is completed, the prediction of future rewards from the same action, which is referred to as action value, is updated based on the outcomes of the action, enabling the animal to make a better decision the next time such a choice is encountered. Thus, decision-making processes become increasingly refined as the animal learns about its environment through experience, ultimately leading to more effective decisions.

In addition to successfully predicting the animal's choice behavior, the reinforcement learning model has been successfully used to elucidate the function of

the basal ganglia in goal-directed behavior. Dopaminergic neurons in the ventral tegmental area and the substantia nigra have been shown to encode a reward-prediction error, which is used to improve the outcomes of an animal's future choices (Schultz et al., 1997). Another study in monkeys engaged in a free-choice task showed that the activity of striatal neurons is correlated with action values, which were estimated by integrating the previous outcome history associated with each action (Samejima et al., 2005).

Although the basal ganglia play a key role in reinforcement learning, the specific relationship between striatal signals related to action values, choices, and outcomes is still poorly understood. Additionally, it is unknown how these signals are integrated within the larger cortico-basal ganglia circuitry to form a flexible and reliable decision-making network.

A recent study by Lau and Glimcher (2007) makes an important contribution to our understanding of how individual neurons in the basal ganglia encode action and outcome, and it provides valuable insights into the organization of the cortico-basal ganglia network. Lau and Glimcher recorded from phasically active neurons in the caudate nuclei of two monkeys that were engaged in a probabilistically rewarded delayed saccade task. Monkeys fixated on a central light-emitting diode (LED) for 400 ms before a peripheral LED was illuminated in one of eight target locations arranged symmetrically around

the fixation point. After a short delay, the fixation point was extinguished, signaling the monkey to make a saccade to the target. Rewards were delivered on 30–50% of correct trials, and the reward probability was held constant throughout the recording session [Lau and Glimcher (2007), their Fig. 1 (<http://www.jneurosci.org/cgi/content/full/27/52/14502/F1>)].

Interestingly, approximately one-half of neurons that were phasically active during the task displayed a peak response after the saccade had already been made, suggesting that they did not play a role in selecting movement [Lau and Glimcher (2007), their Fig. 3 (<http://www.jneurosci.org/cgi/content/full/27/52/14502/F3>)]. Lau and Glimcher next examined whether each neuron encoded reward outcome, direction (of action), or both action and reward. Approximately one-half (30 of 54) of the neurons showed statistically significant activity for only one category; they independently encoded either direction or reward history. Although the remaining neurons displayed a significant response to both factors, most were strongly biased toward only one of them: an analysis of the joint distribution of reward responsiveness and tuning sharpness showed that fewer than expected sharply tuned neurons had large differential reward responses [Lau and Glimcher (2007), their Fig. 8 (<http://www.jneurosci.org/cgi/content/full/27/52/14502/F8>)]. From this, Lau and Glimcher concluded

Received Feb. 26, 2008; revised March 24, 2008; accepted March 26, 2008.

C.H.D. was supported by National Institutes of Health (NIH) Training Grant 5 T32 NS 41228-07. H.S. was supported by NIH Grant MH073246. We thank Daeyeol Lee for his comments on this manuscript.

Correspondence should be addressed to Christopher H. Donahue, Interdepartmental Neuroscience Program, Yale University School of Medicine, New Haven, CT 06520. E-mail: christopher.donahue@yale.edu.

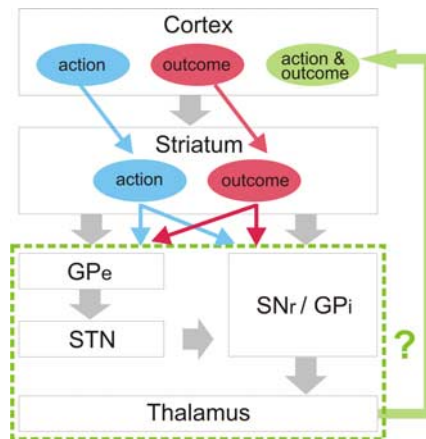
DOI:10.1523/JNEUROSCI.0858-08.2008

Copyright © 2008 Society for Neuroscience 0270-6474/08/284579-02\$15.00/0

that action and outcome were encoded in largely separate channels in the caudate.

These separately encoding populations could be used together to update the predicted value of actions. Lau and Glimcher suggest that the signals corresponding to retrospective movement direction could serve as what is called an “eligibility trace” in the reinforcement learning literature. Eligibility traces are signals that can act as a short-term memory of the animal’s own behavior, so that rewards can be properly associated with previous actions (Sutton and Barto, 1998). Neural activity encoding previous choices has also been found in the dorsolateral prefrontal cortex (DLPFC) (Seo et al., 2007), suggesting that the signals related to previous actions could be used to update action values in the corticostriatal pathway.

Further results from the prefrontal cortex highlight the importance of considering Lau and Glimcher’s findings within the context of a broader corticobasal ganglia decision-making network. In a study using a task similar to Lau and Glimcher’s, Tsujimoto and Sawaguchi (2005) found that both reward information and directional preference are jointly encoded in individual neurons of the DLPFC. In that study, monkeys were trained on both a memory-guided and a visually guided saccade task. Tsujimoto and Sawaguchi (2005) concluded that each neuron’s postmovement activity was significantly modulated by the directional preference, the reward outcome, and the specific task category. Neurons in the sup-



**Figure 1.** A schematic diagram of the corticobasal ganglia loop involved in encoding actions and outcomes. It is not yet known where in this pathway these signals are combined. GPe, Globus pallidus pars externa; GPi, globus pallidus pars interna; SNr, substantia nigra pars reticulata; STN, subthalamic nucleus.

plementary eye fields also conjunctively encode action and outcome (Uchida et al., 2007). Altogether, these results suggest an important contrast between how the prefrontal cortex and striatum encode information related to actions and outcomes. Because neurons in the caudate nucleus receive dense projections from the DLPFC, the data suggest that the neurons projecting to the caudate originate from separately encoding populations. These separate channels could combine somewhere in the corticobasal ganglia loop downstream of the caudate before reaching a distinct area of cortex containing

neurons with overlapping representations (Fig. 1).

Future research should focus on recording areas of the corticobasal ganglia loop downstream of the caudate to identify where the signals related to action and outcome are combined. Additionally, it would be informative to use tasks that require an animal to use reward information to select later actions. Such tasks could further elucidate how separate signals are used to update action values and could lead to a better understanding of the organization of the corticobasal ganglia network.

## References

- Lau B, Glimcher PW (2007) Action and outcome encoding in the primate caudate nucleus. *J Neurosci* 27:14502–14514.
- Samejima K, Ueda Y, Doya K, Kimura M (2005) Representation of action-specific reward values in the striatum. *Science* 310:1337–1340.
- Schultz W, Dayan P, Montague R (1997) A neural substrate of prediction and reward. *Science* 275:1593–1599.
- Seo H, Barraclough DJ, Lee D (2007) Dynamic signals related to choices and outcomes in the dorsolateral prefrontal cortex. *Cereb Cortex* 17:110–117.
- Sutton RS, Barto AG (1998) Reinforcement learning: an introduction. Cambridge, MA: MIT.
- Tsujimoto S, Sawaguchi T (2005) Context-dependent representation of response-outcome in monkey prefrontal neurons. *Cereb Cortex* 15:888–898.
- Uchida Y, Lu X, Ohmae S, Takahashi T, Kitazawa S (2007) Neuronal activity related to reward size and rewarded target position in primate supplementary eye field. *J Neurosci* 27:13750–13755.