

Chapter 9

Dopamine Prediction Errors and the Relativity of Value

Masamichi Sakagami and Shingo Tanaka

Abstract Value, or reward prediction, is thought to be generated by neural circuits among the basal ganglia and midbrain dopamine area. Midbrain dopamine neurons calculate the difference between estimated and actual rewards and send the reward prediction error to the basal ganglia, particularly striatum where the value is coded. The value reflects the quality and quantity of reward, the timing of reward delivery, etc. In many cases, however, actually experienced values depend on the context in which rewards are given and cannot be explained merely by physicochemical components of reward.

Keywords Dopamine neuron · Reward value · Reward prediction error · Midbrain · Random dot movement

We conducted two kinds of experiments on the reward prediction error in midbrain dopamine neurons of Japanese monkeys in order to investigate neuronal mechanisms of relative value coding. In Nomoto et al. [1], we recorded single unit activity from midbrain dopamine neurons, while monkeys performed the random dot discrimination task with an asymmetric reward schedule (one direction was associated with a large reward and the opposite direction with a small reward). We observed biphasic cue-evoked dopamine responses representing reward prediction errors at the times of stimulus detection (early component) and identification (late component). Both components reflected reward prediction errors. The early component, which was trial-type-independent, coded the prediction error of average value of all the cues in the task as the component could not discriminate the direction of random dot motion cue but could detect the onset of cue. The late component, which was trial-type specific, coded the prediction error of specific value of the cue. The quantity of late component corresponded to the difference between the value of specific cue and the average value coded by the early component. These results indicate that dopamine prediction error signals are computed by using

M. Sakagami (✉) · S. Tanaka
Brain Science Institute, Tamagawa University, Tokyo, Japan
e-mail: sakagami@lab.tamagawa.ac.jp

moment-to-moment reward prediction of cue stimuli and the amount of later one was modulated by that of preceding one. Therefore, the dopamine reward prediction error, particularly in later stages, can be relative, dependent on the context. Then we can postulate that the value itself in the basal ganglia may be relative, not absolute, because the value in the basal ganglia is dependent on the relative reward prediction error signal from dopamine neurons [2].

Value can be modulated also by the effort or the cost put into its achievement. Although recent reports indicate the cost information is transmitted to dopamine neurons, it is still unclear whether the paid cost modulates the activity of dopamine neurons. We examined whether the activity of the dopamine neurons in response to reward predictive cues was modulated by the cost preceding to these cues [3]. Two Japanese monkeys performed a saccade task. After fixation on a fixation point, subjects were required to make a saccade to a condition cue and then a target appeared. In the high-cost condition, long fixation to the target was required. In the low-cost condition, only a short fixation was required. After fixation on the target, the subjects made a saccade to the reward cue and obtained the reward. While the subjects performed the saccade task, the activities of dopamine neurons were recorded from the midbrain dopamine area (SNc). The dopamine neurons showed phasic responses to the condition cues and the reward cues. The neuronal response to the low-cost cue was larger than that to the high-cost cue. This difference in the activity of the dopamine neurons was independent of the subtype of dopamine neurons (value and salience types). In contrast, the responses to the reward cue after the high-cost condition were larger than those to the reward cue after the low-cost condition indicating that the paid cost increased the reward prediction error signal. We also showed that the paid cost enhanced the learning speed in an exploration task, in which the subjects have to learn the relationship between the reward cue and the reward trial-by-trial. Furthermore, the paid cost increased the response of the dopamine neurons to the reward itself. From these results, we suggest that information about the cost is integrated in the dopamine neurons and the paid cost amplifies the value of the reward.

According to recent studies on the value generation in the nigrostriatal circuit, the value learning can be explained by relatively simple reinforcement learning model [4–6], which means that the learnt value is basically dependent on the experienced probability of contingency between an event (stimulus or response) and outcome. Dopamine neurons calculate the prediction error to update the value more appropriately. As dopamine neurons calculate the prediction error moment-to-moment and the circumstances are usually not simple but structured, the moment-to-moment prediction error signal is not absolute but relative. As a result, we can postulate the value coded in the basal ganglia, which receives the prediction error signal from dopamine neurons, should be relative, which we have to clarify with future works.

References

1. Nomoto, K., Schultz, W., Watanabe, T., Sakagami, M.: Temporally extended dopamine responses to perceptually demanding reward-predictive stimuli. *J. Neurosci.* **30**(32), 10692–10702 (2010)
2. Schultz, W.: Multiple dopamine functions at different time courses. *Annu. Rev. Neurosci.* **30**, 259–288 (2008)
3. Tanaka, S., Sakagami, M.: The enhancement of the reward prediction error signal in the midbrain dopamine neuron by the cost paid for the reward. Abstract retrieved from Abstracts in Society for Neuroscience (Abstract No. 776.01/LLL19) (2013)
4. Daw, N.D., Niv, Y., Dayan, P.: Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* **8**(12), 1704–1711 (2005)
5. Rangel, A., Camerer, C., Montague, R.: A framework for studying the neurobiology of value-based decision-making. *Nat. Rev. Neurosci.* **9**(7), 545–556 (2008)
6. Samejima, K., Ueda, Y., Doya, K., Kimura, M.: Representation of action-specific reward values in the striatum. *Science.* **310**(5752), 1337–1340 (2005)