Exploring the Cognitive Process of Learning the Latent Structures in a Probabilistic Reversal Learning Task

Akira Masumi *and Takashi Sato*

Department of Media Information Engineering, National Institute of Technology, Okinawa College,905 Henoko, Nago, Okinawa, Japan, masumi.a@okinawa-ct.ac.jp.

In our daily life, we frequently face a situation in which we need to make decision flexibly. Driving a car is one of the examples; a "right speed" depends on traffic lights, degree of congestion on the road, pedestrians around there and so on, furthermore those conditions may change dynamically. In such a situation, we have to pay attention to the surrounding situations and what we should do. This kind of flexibility in decision making, known as cognitive flexibility, is results of complex and higher level cognition, and its neural basis and computational accounts have been studied in many experimental and theoretical studies [1, 2].

As one of the experimental paradigms, a probabilistic reversal learning task has been used to characterize cognitive flexibility of a subject [2]. In this task, a subject first learns stimulus-response contingencies with probabilistic errors, which of two options is a rewarding option or not. After sufficient number of trials, these contingencies are reversed during the task, then the subject have to adjust their behavior from the previously (before reversal) learned one to the reversed one. With this procedure, we can assess the cognitive flexibility of a subject by measuring the extent to which the subject can appropriately adjust their behavior. In recent studies, it is suggested that a subject may utilize a "cognitive map" representing latent structures of the task in which the task has two kinds of a contingency and the reversal occurs during the task, and so on [3]. However, it is not fully understood yet whether subjects actually utilize the "cognitive map" of the task, and in that case, how do participants acquire it through the trials?

We conducted the experiments with the task to address the questions. In the present study, there are two conditions with different difficulties of the task in which an occurrence probability of errors differ. If the occurrence of errors is rare, it is easy to discriminate a rewarding option from a punishing one, and vice versa. We carried out the experiments under easy and difficult condition for each 30 participants (27 males and 3 females, mean age 19.27 years, SD=1.14).

From the results of the experiments, some behavioral indicators suggested that subjects make a decision based on their own estimation about the latent structures of the task. Model-based analysis by using model-free and model-based reinforcement learning model also supported same result. In addition, we found that a reaction time from displaying the stimulus to key pressing is shorter in a later phase of the task than the earlier one in the easy condition, and it was suggested that a shape of distributions of the reaction time qualitatively change. This results implies a mode of decision making is different between conditions and changes across the different phases of the task.

Acknowledgements (optional)

This study was supported by JSPS KAKENHI Grant Number JP16H06397.

References

[1] W.A. Scott, *Sociometry*, pp. 405–414, 1962.

[2] A. Izquierdo, et al, *Neuroscience*, vol. 345, no. 2, pp. 12–26, 2017.

[3] R. C. Wilson, et al, *Neuron*, vol. 81, no. 2, pp. 267–279, 2014.