Development of a Visual-guided Probabilistic Selection Task for Rats

A. Trecker[1], T.W. Robbins[2], A.C. Mar[2]

[1]*Department of Comparative Psychology, Institute of Experimental Psychology, Heinrich-Heine-University Düsseldorf, Germany. andrea.trecker@email.de*

[2]*Department of Experimental Psychology, Behavioural and Clinical Neuroscience Institute, University of Cambridge*

Reinforcement learning can be defined as a trial and error learning of how to maximize rewards and/or minimize punishments through repeated sampling or interaction with the environment. This implicit, unsupervised form of learning is an essential component of daily adaptive behaviour and is observed to be disrupted in several neuropsychiatric disorders such as Parkinson's disease and schizophrenia [1, 2, 3]. The observation that phasic firing of midbrain dopamine neurons may encode reward prediction errors during reinforcement learning [4] has fueled considerable interest in the neural mechanisms of reinforcement.

While probabilistic reinforcement learning paradigms using visual objects are commonly-used in research with humans and non-human primates, to our knowledge, no fully-validated paradigm currently exists for the rodent. The aim of the present study was thus to develop a reinforcement learning paradigm (Probabilistic Selection Task) for rats with based on the experimental methods and computational framework of Frank and colleagues [1].

Experiments were conducted using 16 male lister hooded rats. Rats were tested in Med Associate Inc. operant chambers. One wall of a standard operant chamber was removed and replaced with an infrared touch sensitive flat-screen monitor at the front. House light and food magazine containing built-in photocells was located on the rear wall.

Behavioural training proceeded as follows:

In phase 1, rats were trained to press a white rectangle (3 x 15 cm) at the bottom centre of the screen in order to receive immediate delivery of a food pellet reward. After rats completed 180 trials within a session (max. 30 minutes), the size of the rectangle was reduced to its final size of 3 x 6 cm.

In phase 2, touching the rectangle resulted in trial initiation rather than a food reward. The rationale for using the initiation rectangle on the screen was to keep attention directed toward the screen and to prevent biases (sometimes observed in other touchscreen tasks) due to stereotyped turning responses from the food magazine at the rear of the chamber. Upon trial initiation, rats were presented with two distinct visual stimuli (e.g., spider and plane, each 8 x 8 cm) in two of three possible positions (y from bottom = 3 cm, x from left = 0, = 7,6, = 15,3 cm) on the touchscreen. The positions of the stimuli were randomized on every trial. Stimuli were also counterbalanced across the reward probability and presentation order of the reinforcement pairs. For designated AB pairs, touching stimulus A resulted in 90% chance of reward (10% chance of non-reward), while touching stimulus B only a 10% chance of reward (90% non-reward). Feedback for non-rewarded choices were followed by a 5s timeout with the house light illuminated. Similarly, designated CD pairs were associated with 80%:20% reward ratios and designated EF pairs with 70%:30% reward ratios. The trial initiation rectangle was presented immediately after reward collection or the timeout period to begin the next trial. Sessions lasted a maximum of 30 min or after completion of 180 trials.

In Phase 3, after achieving the learning criteria in Phase 2, subjects are then presented with the original stimuli in novel pairings to probe the contributions of positive and negative feedback to probabilistic learning by assessing whether rats displayed a bias for choosing frequently-reinforced stimuli, or for avoiding frequently-punished stimuli.

As expected we observed a significant effect of designated pair in phase 2, with the highest choice ratios (e.g., A/(A+B)) seen for AB (90:10) pairs and the lowest for EF (70:30) pairs. These differences in reinforcement-guided discrimination learning were apparent at the group level after only 7-8 sessions of training. Moreover, novel use of the initiation rectangle resulted in rapid choice latencies and a relative absence a priori stimulus or position biases. Trial by trial data were fit using standard reinforcement learning models to further characterise the differences in learning [5].

In sum, we have developed a visually-guided reinforcement learning task for rodents with parallels to the human Probabilistic Selection Task. This paradigm holds promise for the translation of human reinforcement learning in rodents, and may be of utility in future studies examining its neuropharmacological underpinnings.

References

1. Frank, M.J., Seeberger, L.C., O'Reilly, R.C. (2004). By Carrot or by Stick: Cognitive Reinforcement Learning in Parkinsonism. *Science* **306**, 1940-1093.

2. Waltz, J.A., Frank, M.J., Robinson, B.M., Gold, J.M. (2007). Selective reinforcement learning deficits in schizophrenia support predictions from computational models of striatal-cortical dysfunction. *Biol Psychiatry* **62**, 756-764.

3. Yilmaz, A., Simsek, F., Gonul, A.S. (2012). Reduced reward-related probability learning in schizophrenia patients. *Neuropsychiatr Dis Treat* **8**, 27-34.

4. Schultz, W., Dayan, P., Montague, P.R. (1997). A neural substrate of prediction and reward. *Science* **14**, 1593-1599.

5. Sutton, R.S., Barto, A.G. (1998). *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA.