

Effects of reward history on current certainty reflected by wagering time during reversal-learning in monkeys

Chenxi Zhai^a, Zhiyong Jin^a, Yudian Cai^a, Xuanlong Zhu^a, Zhaoxin Wang^a, Zheng Wang^d, Sze Chai Kwok^{a,b,c}

a. Shanghai Key Laboratory of Brain Functional Genomics, Key Laboratory of Brain Functional Genomics Ministry of Education, School of Psychology and Cognitive Science, East China Normal University, Shanghai, 200001, China

b. Shanghai Key Laboratory of Magnetic Resonance, East China Normal University, Shanghai, 200001, China

c. Division of Natural and Applied Sciences, Duke Kunshan University, Kunshan, Jiangsu, 215300, China

d. School of Psychological and Cognitive Sciences, Peking University, Beijing, 100000, China

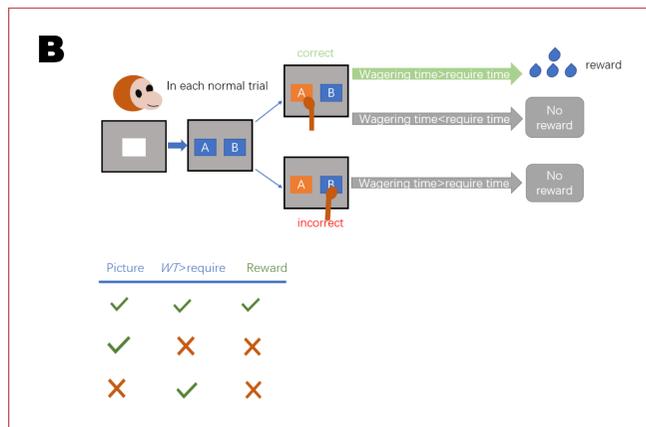
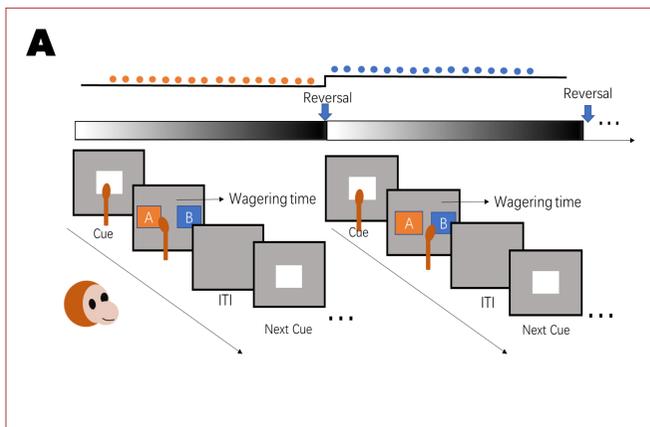


Abstract:

Humans and animals face ever changing environment and make decisions by learning the contingencies between stimulus and their outcome. The current choice is not only affected by the accumulated perception evidence but also influenced by past decisions. Confidence is considered as a process related to self-awareness. Some study report that several species besides human are capable of confidence judgments, like macaques and rats. Also, many studies report that confidence can bias subsequent decision. In this study, we ask monkey to wait for a certain amount of time on the selected stimuli as a wager for its own decision. We divided trials into reached and unreached trials. We found reached trials have a significantly higher accuracy rate than unreached trials ($p < 0.01$), which indicates that monkeys are more willing to wait when their responses were correct. We constructed a logistic general linear model to estimate probability of the chosen stimuli. The probability on the current trial could be predicted by factors including reward, chosen side, picture identity, interactions of reward \times side and reward \times picture of the preceding one trial. The reward and the reward \times side interaction factors are significant predictor of the decision in next trial. Furthermore, the absolute differential probability of chosen and unchosen stimuli was positively proportional to the waiting time of macaques on the same trial (two animals correlation coefficient are 0.32 and 0.44, all p -values < 0.001). These results indicate that (1) rhesus monkeys' waiting time can reflect the certainty of their own choices, and (2) as affected by whether the previous trial was rewarded, the macaque will adjust its ongoing choices and the associated invested time.

Keywords: reversal learning, confidence, serial dependence, macaque

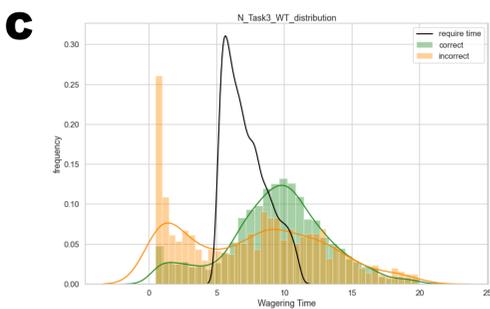
1. Behavioral experiment:



-A. Paradigm of reversal learning. Rule of stimuli-reward reversed after monkeys reached a required correct rate in last 20 trials. The required correct rate was randomly chosen in 75%, 80%, 85%.

-B. Details and reward conditions of time wagering. Only when the monkey chooses the correct picture and waits enough (longer than trial-specific required time) will it be rewarded. The waiting time starts from when one picture is selected until their hands leave the screen.

2. Trial definition:



	Unreached trial	Reached trial
Description	Trial waiting time < required waiting time	Trial waiting time \geq required waiting time
Indication	Low confidence	High confidence
Reward	No reward	If correct, 2 ml water

-C. Example of the distribution of required waiting time from monkey N. Figure shows the distributions of required time (black) and actual waiting time (green for correct and yellow for incorrect trials) of all sessions. For each monkey, the trial-specific required waiting time was randomly chosen from the pre-determined distribution.

-D. Trials were divided in reached and unreached trials, depended on whether the actual waiting time reached required waiting time for each trial.

3. Model

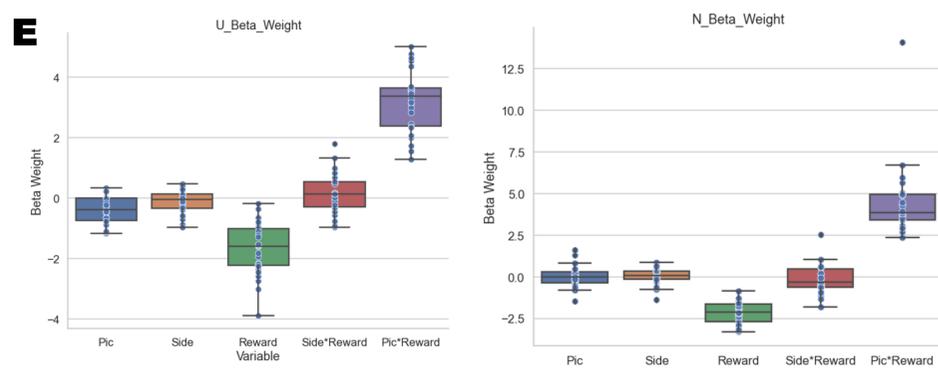
3. Model

- We run general linear model (GLM) analysis (Equation 1).
 - We estimated the effect of last trial's chosen image (Pic_{t-1}), reward outcome (R_{t-1}), chosen side (S_{t-1}), and interactions ($R_{t-1} \times S_{t-1}$, $R_{t-1} \times Pic_{t-1}$) to predict which image will be chosen in this trial and probabilities of choosing them ($P(A_t)$, $P(B_t)$).

$$\log\left(\frac{P(A_t)}{P(B_t)}\right) = \beta_0 + \beta_1 Pic_{t-1} + \beta_2 R_{t-1} + \beta_3 S_{t-1} + \beta_4 R_{t-1} \times S_{t-1} + \beta_5 R_{t-1} \times Pic_{t-1} + \varepsilon \quad (1)$$

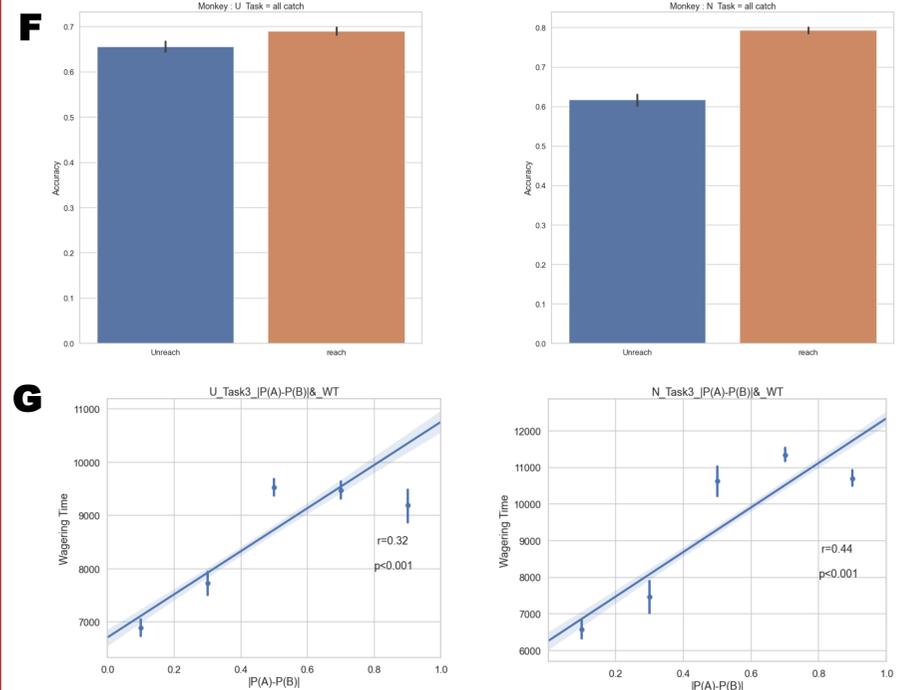
- We run this model per animal per sessions. In this poster we show data of two monkeys and each monkey has 25 sessions.

3. Model fitting



-Significant effects of preceding *Reward* and *Pic*Reward* on the current choice of pictures ($ps < 0.05$). The box plot shows the quartiles of beta weights. Each dot represents corresponding GLM coefficient fitted on one session. Models run on each monkey separately, monkey U (left) and monkey N (right).

3. Result



-F. Accuracy differs between two types of trial. Reach trials (orange) have higher performance than unreach trials (blue), on monkey U (left, $p < 0.05$) and monkey N (right, $p < 0.001$).

-G. Certainty about current choice positively correlated with monkeys' waiting time. Certainty is measured by the absolute value of difference between predicted probability of two choice. Results are consistent on the two monkey (left, monkey U, Spearman correlation $r = 0.32$, $p < 0.001$; right, monkey N, Spearman correlation $r = 0.44$, $p < 0.001$).

Reference and Acknowledgment

Lak, A., Hueske, E., Hirokawa, J., Masset, P., Ott, T., Urai, A. E., ... Kepecs, A. (2020). Reinforcement biases subsequent perceptual decisions when confidence is low, a widespread behavioral phenomenon. *Elife*, 9. doi:10.7554/eLife.49834
 Izquierdo, A., Brigan, J. L., Radke, A. K., Rudebeck, P. H., & Holmes, A. (2017). The neural basis of reversal learning: An updated perspective. *Neuroscience*, 345, 12-26. doi:10.1016/j.neuroscience.2016.03.021
 Lak, A., Costa, G. M., Romberg, E., Koulakov, A. A., Mainen, Z. F., & Kepecs, A. (2014). Orbitofrontal Cortex Is Required for Optimal Waiting Based on Decision Confidence. *Neuron*, 84(1), 190-201. doi:10.1016/j.neuron.2014.08.039



kwoklab.org
Scan to get more information