Probabilistic Discounting for Modeling Behaviors in Iowa Gambling Task

Takaki Makino [1) 3)], Taiki Takahashi [2)], Hirofumi Nishinaka[3)], Hiroki Fukui [3)]
1) Division of Project Coordination, University of Tokyo Kashiwa-no-ha 5-1-5, Kashiwa 277-8568 Japan mak@scint.dpc.u-tokyo.ac.jp
2) Dept. Behavioral Science, Hokkaido University 3) Dept. Forensic Psychiatry, National Center of Neurology and Psychiatry

1. Introduction

Iowa Gambling Task (IGT) [1,2] is a behavioral experiment task that investigates action selection under an environment with probabilistic outcome. It is known that patients with lesions on ventromedial prefrontal cortex (VMPFC) and amygdala tend to show impaired action selection, that is, failure in learning to select beneficial actions and continue choosing non-beneficial actions. These brain areas are said to be related with Somatic Marker hypothesis [5] (the function to learn outcome of actions through emotional reactions), and the impairment of action selection has been attributed to the damage of Somatic Marker functions. Thus, the valence model was widely used to analyze the result of IGT tasks, which is parameterized by the weight *w* between positive and negative stimulus. However, the impairment of action selection by VMPFC-lesioned patients may come from the myopic behavior of the patients, making larger discount on temporally distant results [6]. Since IGT provides a probabilistic but static environment, we cannot directly test myopicity from its results. However, if we use a behavioral model based on Rachlin's probabilistic discounting theory [3], which describes correlation between probabilistic discounting and temporal discounting, it is possible to make an indirect analysis of myopicity on the result of IGT [7].

In this study, we compared the estimated parameters by the existing valence model and proposed probability discounting model on the set of IGT behavioral data, which include both the original IGT and variant IGT (with inversed penalties and rewards) for healthy subjects and VMPFC-lesioned patients.

2. Iowa Gambling Task

In Iowa Gambling Task (IGT), the subject repeatedly chooses one card deck out of four. The subject gains some virtual money, whose amount is fixed for each deck. After that, the subject may lose some money, whose amount and probability also depend on the chosen deck. The goal of the task is to increase the virtual money, starting with $2,000, as much as possible through the game.

Two of the four decks (A and B) have assigned a large fixed gain $100 with large loss so that the mean benefit from the decks is negative, while the other two (C and D) have assigned a small fixed gain $50 and smaller loss so that the mean benefit is positive.

It is known that healthy subjects are tend to choose good decks (C and D) in the long run, unlike the patients with lesions on ventromedial prefrontal cortex (VMPFC) or amygdala, who tend to choose bad decks (A and B). It is said that the lesions on these brain regions causes deficit in somatic marker, which plays an important role in learning against large penalty.

IGT Original Task

	A	B	C	D
a. Fixed gain	$100	$100	$50	$50
b. Loss probability	50%	10%	50%	10%
c. Average loss value	- $250	- $1250	- $50	- $250
d. Mean Benefit (a+bc)	- $25	- $25	+ $25	+ $25

IGT Variant Task

	E	F	G	H
a. Fixed loss	- $100	- $50	- $100	- $50
b. Gain probability	10%	50%	50%	10%
c. Average gain value	$1250	$50	$250	$250
d. Mean benefit (a+bc)	+ $25	- $25	+ $25	- $25

Figure 1. Iowa Gambling Task

In this study, we performed experiments with the variant version of IGT as well as original IGT. The variant IGT has the reversed setting, that is, fixed loss and probabilistic gain.

3. Data Analysis with RL Models

In this study, we analyzed the behavioral data with reinforcement learning-based models to extract the characteristics of the subjects' choice. In the theory of reinforcement learning, the choice of action is considered as a decision process based on the (subjective) values of actions, and the values are decided from the past experience of the choice of action and its result. We can model this process using a pair of a value judgment model and an action selection model.

In this study, we compared two value judgment models in terms of most likely parameters. The parameters of models for each subject are calculated so that the probability distribution of action selection predicted by the model give the maximum likelihoo]d for the actual action selections taken by the subject. As for action selection model, we adopt Soft-max model, which gives the probability $P_t(a)$ of selecting action a at time t, using the estimated value $V_t(a)$ of choosing deck a at time t and reverse temperature β (corresponding to the consistency of action selection):

$$P_t(a) = \frac{e^{\beta V_t(a)}}{\sum_{a'} e^{\beta V_t(a')}}$$

4. Data and Models

In this study, we compared parameter distributions for healthy and VMPFC-lesioned subjects between existing valence model and proposed probability-discounting model. To exclude inaccurate parameter values caused by

bad fitting, we used the parameter values only when the Akaike Information Criterion of the model for a subject is above the baseline (uniform distribution model).

4.1. Existing Method: Valence model

In Valence Model, the estimated value is updated based on the following formula.

$$V_t(a) = V_{t-1}(a) + \gamma[r_t(a) - V_{t-1}(a)]$$
$$r_t(a) = w \cdot r_t^+(a) + (1-w) \cdot r_t^-(a)$$

where γ ($0 \leqq \gamma \leqq 1$) is the parameter for update speed (larger γ makes the estimation closer to the outcome of the last action r_t and ignores older outcomes), and w ($0 \leqq w \leqq 1$) is the weighting parameter between positive outcome r_t^+ and negative outcome r_t^- (larger w overweights positive outcome.)

Table 1. Summary of Subjects.

	# subjects	Fit for probability-discount model	Fit for valence model
Healthy	38	25	29
Original	20	18	18
Variant	18	9	13
VMPFC	12	8	9
Original	10	7	8
Variant	2	1	1

4.2. Proposed Method: Probability-Discounting model

In probability discounting model [3], an event with probability p is cognitively processed through the average waiting time, that is, the expected number of trials until the next occurrence of the event ($1/p - 1$). Such a model can deal with "myopic" subjects, who tend to discount events with large average waiting time. Such a discounting can be modeled by hyperbolic delay discounting using parameter h:

$$d(p) = \frac{1/p}{1 + h(1/p - 1)}$$

When $h > 1$, the model tend to neglect probabilistic events; when $h < 1$, the model gets sensitive for probabilistic events.

Our model independently updates the value V^S of fixed outcomes and value V^P of probabilistic outcomes, and calculates the final estimated value V using the estimated probability $p_t(a)$ of probabilistic outcomes.

$$V_t^S(a) = V_{t-1}^S(a) + \gamma[r_t^S(a) - V_{t-1}^S(a)]$$
$$V_t^P(a) = V_{t-1}^P(a) + \gamma[r_t^P(a) - V_{t-1}^P(a)]$$
$$V_t(a) = V_t^S(a) + d(p_t(a)) \cdot V_t^P(a)$$

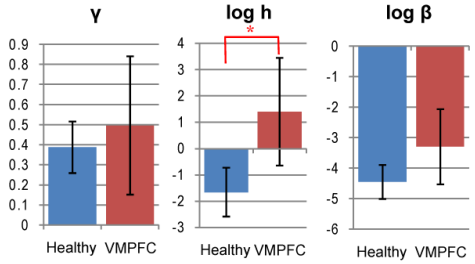

Figure 2. Estimated Parameters by Valence Model

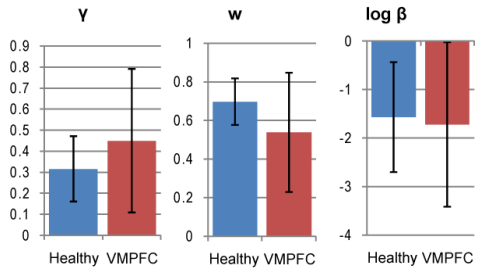

Figure 3. Estimated Parameters by Probabilistic Discounting Model

5. Results and Discussions

Figures 2 and 3 show the distribution of parameters for the model-fit cases. In existing valence model, the estimated parameter w showed no significant difference between healthy subject group and VMPFC-lesioned patient group. On the other hand, using the proposed probability discounting model, we found significant different (p<.05) of discounting parameter h between healthy and VMPFC-lesioned subjects. This indicates that the impaired action selection by patients can be described by them myopicity but not by function of Somatic Marker. Since the distribution of parameter w was significantly different between original tasks and variant tasks by healthy subjects (Figure 4), it is suggested that valence model are affected by the task difference, and the parameter of probability discounting model corresponds to some neuronal mechanism.

6. Summary

The simulation results suggested the advantage of probability-discounting model as a model of human decision-making in probabilistic environments. However, we could not confirm that the probability discounting parameter h represents the characteristic of the value estimation of the subject because the data for original IGT task is provided from the subjects different from subjects for the data for variant IGT. One future work is to test the correlation of parameters for original and variant IGT resuls from the same subject, and to find corresponding brain areas through fMRI experiments.

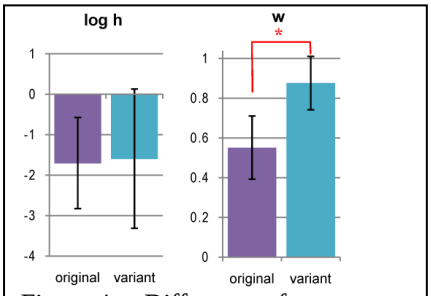

Figure 4. Difference of parameters by the tasks (Healthy subjects)

Reference

[1] Bechara A et al. (1994) *Cognition* **50**:7-15.
[2] Fukui H et al. (2005) *Neuroimage* **24**(1):253-259.
[3] Rachlin H et al. (1991) *J Exp Anal Behav* **55**:233–244.
[4] Yechiam E et al. (2005) *Psychol Sci* **16**(12):973-978.
[5] Damasio H et al. (1994) *Science* **264**:1102–1105.
[6] Kalidindi K, Bowman H (2007) *Neural Networks* **20**(6):676-689.
[7] Takahashi T, Makino T et al. (2009) *Intl J Psychology Research* (To appear).