Proceeding of the IEEE
International Conference on Robotics and Biomimetics
Dali, China, December 2019

A Decision-Making Model Based on Basal Ganglia Account of Action Prediction*

Yabin Liang[1], Zikai Yan[2], Qi Zhang[3], Hongyu Liang[4], Xiyu Ji[5], Yin Liu[6] and Rong Liu*

School of Biomedical Engineering, Dalian University of Technology, Dalian, Liaoning 116024, China
{ [1]ybliang & [2]yzk_zakko & [3]zhangqifd123 & [4]hyliang & [5]jxy & [6]0910ly }@mail.dlut.edu.cn
*rliu@dlut.edu.cn

***Abstract* - Reinforcement learning offers robotics a framework for the design of sophisticated behaviors. The theory of it is deeply rooted in psychological and neuroscientific perspectives. The basal ganglia, thalamus and cerebral cortex constitute an important neural decision circuit involving in the behavioral choice of reinforcement learning. However, the influence of the changes of various nucleuses in the basal ganglia on decision-making is still unclear. Considering the complexity and difficulty of brain experiments on both animals and humans, it is necessary to establish a neural computing model. This paper builds a decision-making model based on physiology and anatomy of basal ganglia. The feasibility of the model was verified by comparing the behavior data from experiments and simulated by the model as well as readiness potentials and discharges rate of nucleus. Moreover, the simulation data were then used to further analyze the effects of conflict, dopamine, and subthalamic nucleus changes on behavioral accuracy and response. The results show this model opens a window for exploring specific neural mechanisms associated with basal ganglia-related decisions.**

Index Terms* - *basal ganglia, reinforcement learning, behavioral data, discharge rate of nucleus*, *readiness potentials

I. INTRODUCTION

Artificial intelligence technology is now one of the research hotspots for robotics. The brain is one of the most important organ in the human body. The key point to improve "intelligence" is to understand the neural mechanism of information processing in the brain, hence to develop biomimic algorithm promoting the development of intelligent robots.

Multiple neural clusters or brain regions are interconnected into a complex network to process information through the interaction of various neural circuits. Therefore, the neural circuit is the basic unit of the brain in which a neuron is the basic unit. In many neural circuits, the loop between the basal ganglia (BG), cerebral cortex and thalamus is closely related to the decision-making. Therefore, different neural computational models of basal ganglia have been built to explore the correlation between this loop and decision parameters, and so as to provide theoretical support for clinical diagnosis and treatment. Rabinovich [1] proposed a winnerless competition (WLC) model composed of FitzHugh-Nagumo spiking neurons to efficiently transform input information to a spatiatemporal output for the olfactory nervous system. Pascual [2] used the hodgkin–huxley (HH) model to explore the treatment of Parkinson's disease (PD) with deep brain stimulation (DBS). However, their models only contain the basal ganglia. Frank [3] constructed a neural circuit consisting of cerebral cortex, basal-ganglia, and thalamus. With this model, they explored how factors such as dopamine (DA), subthalamic nucleus (STN) affect the decisions through probability selection, flipping, and reinforcement learning. The development of the decision-making neural computation model can not only verify the previous research results but also explore some research hypothesis to reduce the possibility of duplicate or meaningless experiments as well as to enrich the decision-making theory. However, the above models are limited to the prediction of behavioral data.

The electroencephalogram (EEG) is the sum of the potentials of a large number of neurons. Readiness potentials (RP) are components related to motion preparation of EEG. In fact, before action, the brain has completed the evidence accumulation, response selection and motion planning steps during the entire decision process, which can be reflected in EEG. Studies have shown that Parkinson's patients show abnormal EEG patterns associated with the basal ganglia and the premotor cortex (PMC), mainly in the range of 15-30 Hz (β) rhythm and 3-10 Hz (θ) rhythm [4]. PD's difficulty in initiating movement is related to decreased activity in the PMC in time domain, which is obviously reflected in the change of RP [5]. However, the specific relationship between the basal ganglia and the RP has not been elucidated. Therefore, combining the nucleus discharge rate in the basal ganglia with the actual RP, it is possible to explore which neural nucleus may play a specific modulation role in decision-making for motor preparation.

In this paper, we build a decision-making model of basal ganglia for action prediction. The reliability of the model was verified by comparing the experimental behavior data with the simulated behavior data, and the RP with the discharge rate of nucleus to analyze the potential mechanism that how the changes in the basal ganglia affect decision-making.

II. MODELING

A. Model structure

BG not only participates in the movement choose and coordination, but also plays an important role in cognitive processes such as reward learning and working memory. BG mainly contains four nerve nuclei, striatum and substantia nigra (SN), globus pallidus (GP) and STN. The SN is divided into substantia nigra reticular (SNr) and substantia nigra pars

*This work was supported in part by the grants from the Natural Science Foundation of China (NSFC) (61573079,81741137), and the Fundamental Research Funds for the Central Universities to Dr. R. Liu.

978-1-7281-6321-5/19//$31.00 ©2019 IEEE

compacta (SNc). The GP is divided into globus pallidus interma (GPi) and globus pallidus external (GPe). Striatum and STN, as the main input nuclei of basal ganglia, receive information from stimuli of the sensory cortex and candidate behaviors from PMC, respectively. GPi and SNr, connecting to the thalamus which projects behavioral signals to the motor cortex of the brain for behavioral expression, are the main output nuclei of basal ganglia. They are only referred to GPi for simplicity because both of them have almost the same projected fibers. Based on the knowledge of anatomy and physiology, the structure diagram of cortex-basal ganglia-thalamus-cortex (CBGTC) loop is shown in Fig. 1.

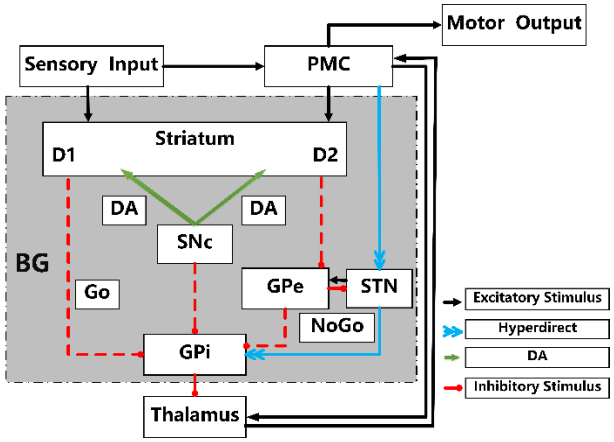

Fig.1 Structure diagram of cortex-basal ganglia-thalamus-cortex loop

There are three pathways in this loop including Go (cortex-striatum-GPi-thalamus-cortex), NoGo (cortex-striatum-GPe-GPi-thalamus-cortex) and Hyperdirect (cortex-STN-GPi-thalamus-cortex), which gating the information flow from and to cortex via thalamus. SNc neurons release DA into the striatum, which contains two receptors, D1 and D2. DA is directly projected into GPi through D1 receptor to form the Go pathway to inhibit GPi, causing the disinhibition of thalamus, thus promoting the response of PMC, and prohibiting the competitive response through the lateral inhibition connection in the meanwhile. DA is projected on GPe through D2 receptor and then GPi is inhibited, which forms NoGo pathway and producing inhibitory effect on thalamus. The interaction between the two pathways enables the neural model to illustrate the effects of different DA operations on the response and reinforcement learning of the entire neural network. In addition, STN is the only nucleus that releases excitatory neurotransmitters in the simplified model of CBGTC. It receives excitatory input from PMC to form a Hyperdirect pathway, and projects to GPi and GPe by sending the diffused excitatory input, GPe provides inhibitive feedback of STN and finally the function of inhibiting thalamus will be realized. In addition, the PMC selects a choice from the projection of sensory cortex directly, so the PMC integrates the information both of the sensory cortex and thalamus to make the final choice.

B. The modeling process

The CBGTC model is implemented with the emergent [6] software. The network framework (Fig. 2) is constructed according to the structure shown in Fig. 1. There are three module configurations in the modeling that is network framework, network input data and network program. The network framework module is divided into four sub-modules: network layer, network layer projection, network layer connection and network layer unit. The network layer realizes the stratification of different functional nerve nuclei, including the number of cells in each layer and the relative position of each layer. Then, the network input module was configured according to the experiment. Finally, the network program was configured. When the results were consistent with the expected results, the activity of the SNc layer was set as 1 to strengthen the connection with Go. However, when the results of the network are inconsistent with the expected results, the activity of the SNc layer is 0, which strengthens the connection with the NoGo. The function of the CBGTC model is that if one of the "Go" signals in the striatum is stronger than its corresponding "NoGo" signal, the other response is suppressed and then the CBGTC model will select a response. After each selection, the model will give positive or negative feedback by means of DA sharp increase or sharp decrease, which will enable Go and NoGo units to learn the previous selection and optimize the selection.

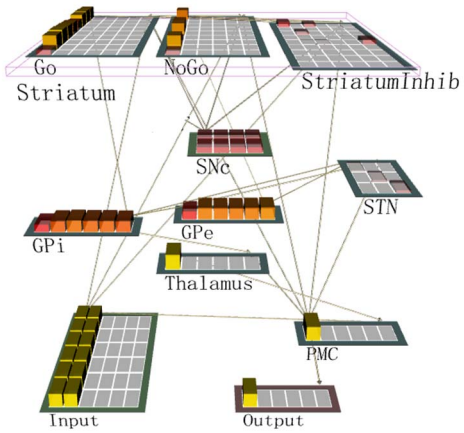

Fig.2 The CBGTC model implementing by the emergent software

III. METHODS

A. Subjects

There were 7 students from Dalian University of Technology participated in the experiment, including 4 male and 3 female, with an average age of 25.14±1.67 years old. All subjects are healthy with no cognitive and psychological problems. They all are right-handed. All the subjects have normal achromatic vision or have reached the standard after correction. Written informed consent was obtained from each subject.

B. Experimental Procedures

The experimental paradigm is similar to probabilistic two-choice reinforcement learning paradigm by Ratcliff and Frank[7]. To avoid the influence of reinforcement learning and

the continuous feedback in the test phase, we divided the original experimental test phase into two parts with no feedback. Moreover, a training stage was added between the two test stages to avoid the loss of learning information in the training stage due to long test time.

The experimental paradigm is shown in Fig. 3. After the experiment starts, a white symbol "+" appears in the center of the computer screen to remind the subject that the stimulus is to appear. Then, the white symbol "+" disappears, and the stimuli of a pair of letters is shown in the central position of the screen. The subject is asked to select one of the letters at will. If select the left letter, press the key "R" on the keyboard, and the right letter corresponds to the key "I". The period from the appearance of the letter to the keystroke of the subject is defined as reaction time (RT).

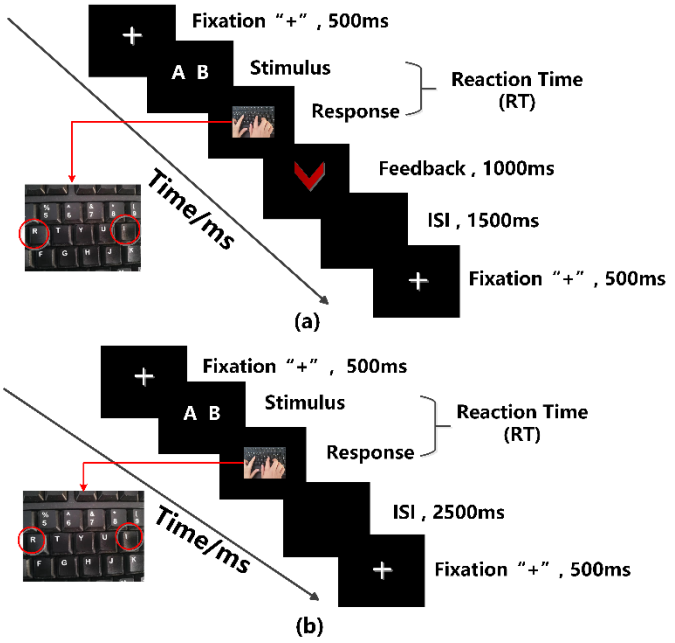

Fig. 3 The experimental paradigm. (a) Training phase; (b) Test phase

In order to learn the law of stimuli appearance and test the learning results, the experimental paradigm will be adjusted in different experimental phases. In the training phase (Fig 3 (a)), when the subject makes a choice, the letter will disappear from the screen, and the feedback symbol "√" or "x" will appear to show the choice is good or not. In the next trial, the subject will adjust choice by choosing letters with higher positive feedback probability or avoid selecting letters with higher probability with negative feedback. In addition, to prevent pressing unintentional button, a response of "reaction too fast" will occur when the RT is less than 100ms. To make the subject concentrate on the experiment, the letter will disappear when the button is not pressed for more than 1500 ms, and then "no feedback" displayed on the screen. The feedback will retain 1000 ms, followed by a black screen that lasts 1500 ms until the next trial begins. In the period test phase (Fig 3 (b)), when the subject makes a choice, the letter on the screen will disappear with no feedback. Then a black screen that lasts 2500 ms will continue to ensure the consistency of the overall experimental paradigm.

As shown in Fig. 4, in the training phase, for the AB pair, 80% of the time when A is chosen providing correct feedback, and 20% of the time with error feedback. For the CD and EF pairs, C and E were reinforced in 70% and 60% of time, respectively, The probability of reinforcement for other alternative stimuli (B, D, F) is the supplementary value of (A, C, E) described previously, i.e. (1-p). For the test phase, 10 pairs of letter pairs are added. According to the training phase, the conflicts generated by the training phase are roughly divided into three categories, i.e. high, medium and low conflict. In win-win option of high conflict, both letters will be positively strengthened (AC, AE, CE). In lose-lose option of high conflict, both letters will be negative strengthened (BD, BF, DF). In the test pair of low conflict, one will get positively strengthened and the other will be negative (AD, BC, BE, AF).

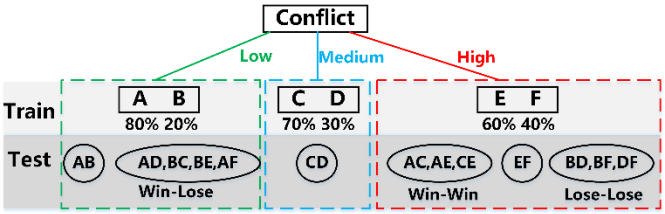

Fig. 4 Experimental paradigm of probabilistic reinforcement learning

C. *Data Acquisition*

Both EEG and behavioral data were collected in the experiment. The EEG data were collected by ANT eego™mylab 32-channel EEG system. The reference electrode is the left mastoid M1. The impedance of all the electrodes were adjusted to below 10kΩ, and the sampling rate was 500Hz. A 0.1-50 Hz bandpass filtering was applied to remove noises. The recording of behavioral data was based on e-prime software.

D. *Data Processing*

The RP was extracted on the basis of EEG with the signal amplitude was at the micro-amplitude level and very susceptible to interference. The processing process is shown in Fig. 5.

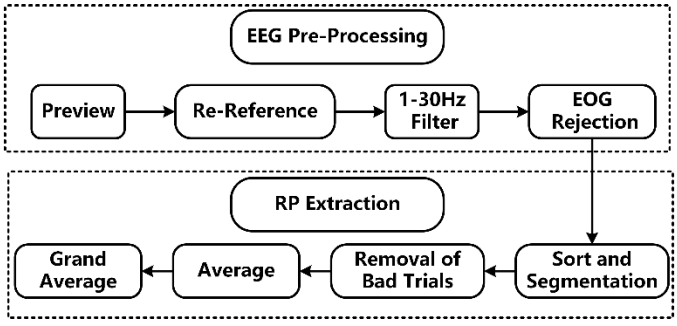

Fig. 5 The analysis process of RP

The process includes EEG pre-processing and RP extraction. The former mainly includes preview, re-reference,

digital filterring and electro-oculogram rejection, while the steps in the RP extraction consists of sort and segmentation, removal of bad trials, average and grand average.

IV. Results

A. Model function verification

The most basic function of the CBGTC model is to satisfy the real biological interaction relations between the various layers of neural nuclei. Therefore, we simulated the binary choice experiment (set two columns randomly from the input layer of the model, and choose one column from the output layer) to test whether the model satisfies these relations. Fig. 6 shows the activity of STN layer, Go group, NoGo group and thalamus layer of the model in the selection task.

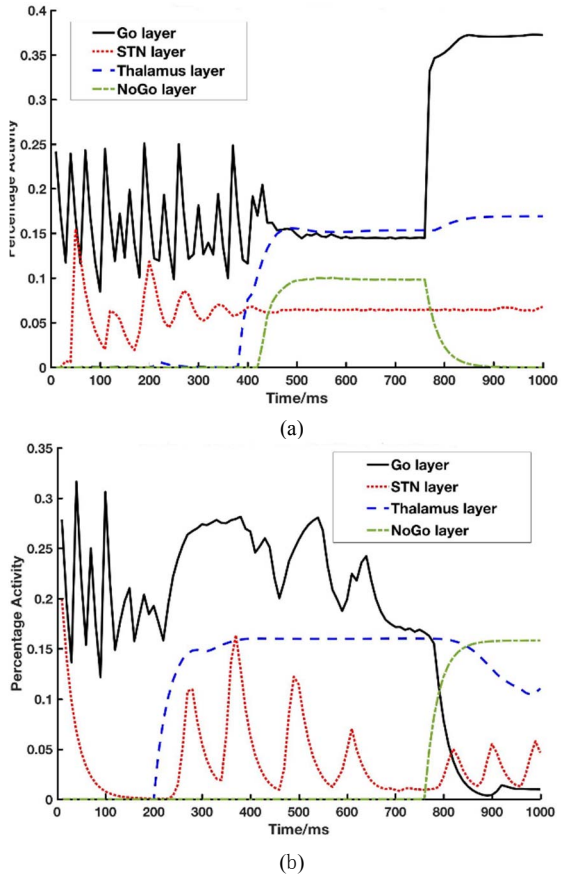

Fig 6 Percentage activity of STN, Go, NoGo and the corresponding thalamus layer, in a binary task. (a)The situation when the model chooses the right result and gets rewarded; (b) The situation when the model chooses the wrong result and gets punished

After accepting the stimulus from the sensory input (the feeling of the current state) and the PMC (the current candidate response), the group of Go and NoGo start to accumulate information respectively and compete with each other. At the same time, the STN accepts the input of the PMC and is activated, thus impose a global NoGo signal to network. Go group promotes thalamic selection, while NoGo group and STN layer inhibit thalamic selection. There are two possible cases between Go group and NoGo group, one is that active Go group and active NoGo group promotes and inhibits different choices respectively (400 ms of Fig. 6 (a)), the other is that the Go group active enough to promote any choice (200 ms of Fig. 6 (b)). In both cases, if the STN activity is reduced, the activity of thalamus will increase quickly, allowing the PMC to make a choice. After the selection, the model will judge whether the selection is correct or not. If it is correct, reward will be given, that is, increase the activity of Go group and decrease the activity of NoGo group (750ms of Fig. 6 (a)). If it is wrong, punishment will be given, that is, decrease the activity of Go group and enhance the activity of NoGo group (750ms of Fig. 6 (b)). In conclusion, the activity of each layer of BG conforms to the physiological and anatomical knowledge described in section II, which verifies that the CBGTC model constructed meets the main biological constraints and functions

B Behavioral data

After verifying the basic function of the CBGTC model, the probability reinforcement learning experiment simulation was carried out. The network training input and test input module were configured according to Fig. 4. The six columns of input represent the stimulus letters A, B, C, D, E, F respectively. The two combinations constituted 6 stimulus conditions (AB, CD, EF, AC\AE\CE, BD\BF\DF, AD\BC\BE\AF). During the analysis of the behavioral data, the 10 cases in the test phase were classified into three categories, that is, (AC, AE, CE) is classified as Win-Win, (BD, BF, DF) is classified as Lose-Lose and (AD, BC, BE, AF) is classified as Win-Lose, and together with the three cases learned in the training (AB,CD,EF).

As shown in Table I, the average accuracy and the distribution of RT on 0.1, 0.5 and 0.9 quantile of experimental data and BG model fitting data were analyzed. The results show that with the increase of experimental conflict (AB<CD<EF or Win-Lose<Win-Win<Lose-Lose), the average accuracy of both data shows a declining trend and the RT of both data shows an increasing trend. Therefore, it is of great value to prove that even if the currently constructed CBGTC model is not suitable for quantitative biometric analysis, but can be used to simulate experimental data to qualitatively predict the trend change of the RT and the accuracy.

TABLE I Accuracy and Quantile RT of data from the experiment and the CBGTC model

			Quantile RT (ms)		
	Simulation	Accuracy	0.1Q	0.5Q	0.9Q
Experiment	AB	0.953	379	458	614
	CD	0.854	378	497	663
	EF	0.819	401	514	734
	Win-Win	0.673	371	456	692
	Lose-Lose	0.328	411	599	845
	Win-Lose	0.762	366	451	682
Simulation	AB	0.957	239	441	664
	CD	0.813	247	520	745
	EF	0.727	274	560	857
	Win-Win	0.529	272	677	868
	Lose-Lose	0.444	351	743	1202
	Win-Lose	0.692	257	581	839

Previous studies have shown that the physiology and pathology of PD is cognitive and motor impairment caused by

the decrease of DA, and the surgical target is mainly STN. In our work, in order to observe the impact of changes with three factors on decision-making, we simulated different levels of DA and different levels of the weight from STN to GPi by adjusting the parameters of BG model. In addition, the different levels of conflict are simulated by experimental paradigm. Finally, we statistically analyzed the accuracy and quantile RT of the behavioral data obtained from the simulation.

The first is to explore the effect of three factors on RT. Fig. 7 shows quantile RT distributions at different DA levels, different conflict levels and different STN levels.

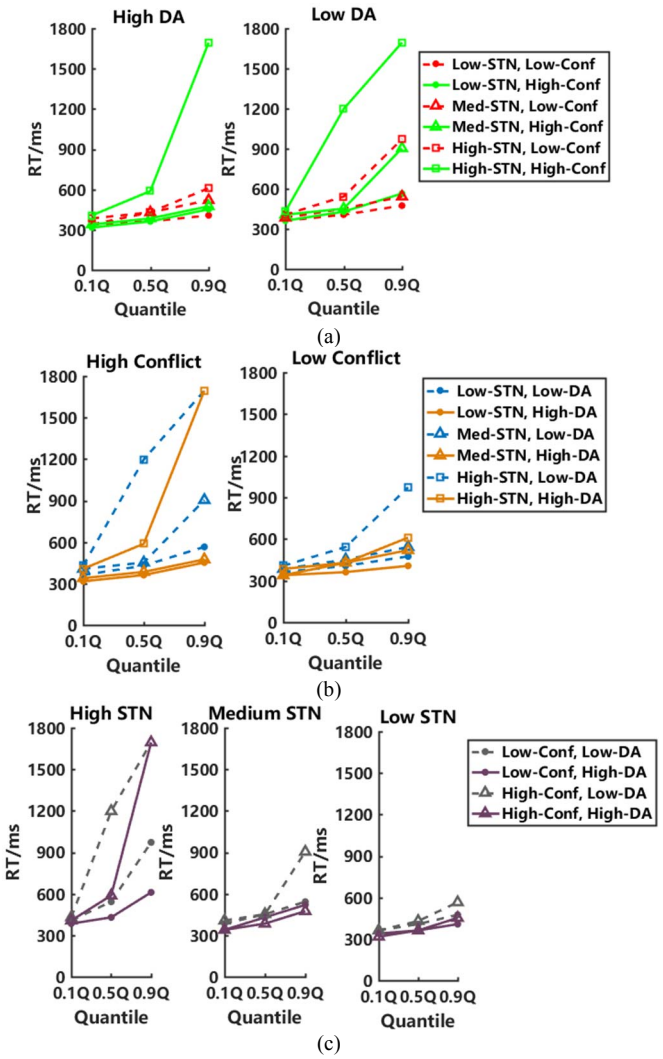

Fig. 7 The Quantile RT distributions producing by different factors. (a)Different DA levels; (b)Different conflict levels; (c)Different STN levels

It was observed that the rise slope of 0.9Q to 0.5Q in each figure was larger than that of 0.5Q to 0.1Q, indicating that the RT distribution was right-sided, which was consistent with the typical RT distribution. Lowering the level of DA, we can see that except for the trend line of high STN and high conflict combination, each line is increased, and it has greater influence on the RT of the second half of the bit line. The low level DA will increase the RT and reduce the reaction rate, which can explain the slow reaction of PD caused by the decrease of DA content. As for the difference of high STN and high conflict combination, it is speculated that the influence of the two factors on the RT drowns out the contribution of DA, and the higher the level of the two factors, the greater the RT. The effect of increasing conflict is to scatter the RT distribution from a dynamic range of about 300 ms-900 ms to 300 ms-1600 ms. The increase of STN intensity is similar to that of conflict. These findings are basically consistent with Frank who describes that removing STN completely from the model under high conflict conditions will lead to premature reaction. This also explains why the STN is the target of PD surgery: STN can remove the inhibition for response.

Next, we need to analyze the impact of three factors on the accuracy. As shown in Figure 8.

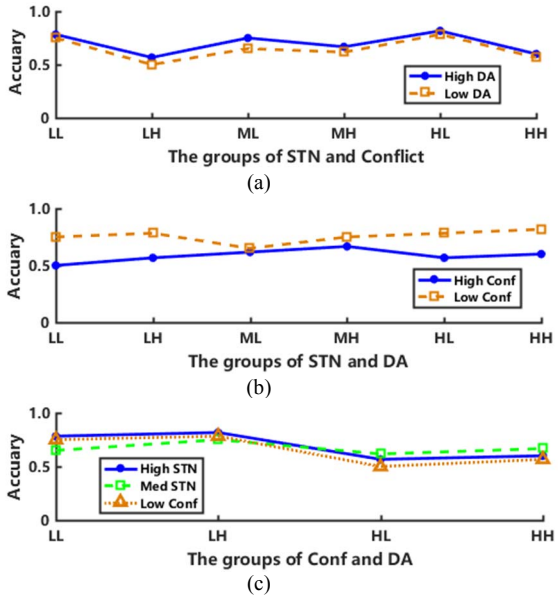

Fig. 8 The accuracy producing by different factors. (a)Different DA levels; (b)Different conflict levels; (c)Different STN levels

The accuracy of high DA is higher than that of low DA. Through variance analysis, we get $F(1,10)=76.26>F\text{-crit}=4.96$, $P=0.4029>0.05$. This shows that DA has an impact on the accuracy, but not significant. It is speculated that in the process of reinforcement learning regulated by DA, the decrease of DA leads to the decrease of the intensity of reward, but the intensity of punishment does not change, so the accuracy rate is rarely reduced. Compared with the low conflict level, the accuracy of high conflict decreases. $F(1,10)=26.56>F\text{-crit}=4.96$ and $P=0.0004<0.05$ are obtained by variance analysis, which shows that conflict has significant influence on the accuracy. It proves that conflict represents the difficulty of experiment, so the accuracy of difficult decision-making will decrease accordingly. Compared with the low STN level, the accuracy of high STN increases. Through variance analysis, we can get $F(1,6)=0.19>F\text{-crit}=5.98$, $P=0.6711>0.05$, which shows that STN has little influence on accuracy. Considering the inconsistent trend line of STN, we can infer that the influence of STN on the accuracy is almost nothing, so that when STN is in the medium level, the accuracy rate is covered by two other factors, showing that the accuracy of low conflict and high DA

is the highest. Previous descriptions of the impact of DA and conflict on accuracy are consistent, which proves that the explanation is reasonable.

C. Neural discharge rate and RP

Fig. 9 is the RP of the prefrontal electrode Fpz. From it, we can see that the curves of low conflict and high conflict are different between 200ms and 600ms. This result is consistent with previous study conclusion that the prefrontal cortex can monitor conflicts [8].

The amplitude of low conflict is higher and reach to the peak faster. Similar phenomenon was observed in the PMC layer of the model. Fig. 10 shows the neural discharge rate (percentage activity) curve of the PMC layer under different conflicts in the model, which illustrates that the neural discharge rate of the model can be equivalent to the RP under the same experimental paradigm.

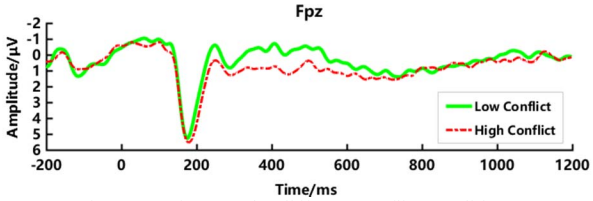

Fig. 9 RP of Fpz under different conflict conditions

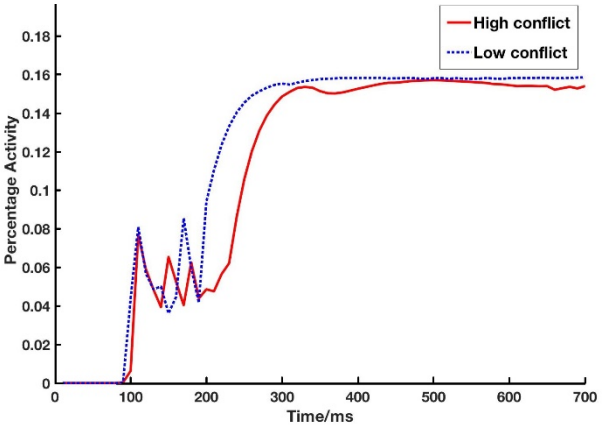

Fig. 10 The neural discharge rate (percentage activity) curve of PMC under different conflict conditions

V · Conclusion and Discussion

A biomimic neural computing decision-making model based on BG was constructed. With the probabilistic reinforcement learning paradigm, the accuracy and RT from the model simulation were compared with the experimental behavior data. In addition, the neural discharge rate curve of the PMC layer in the model compared with the RP of the prefrontal cortical electrode Fpz. The results show that the model can predict the behavior choice. Changing the levels of conflict, DA and STN pathways in the model, it is confirmed that these three factors may influence the RT and accuracy, although differ in degree.

Associated with neurophysiological and pharmacological data, the proposed CBGTC model supports the decision-making process by simulating the nonlinear dynamics of multiple neural node populations. Although this model cannot accurately fit the individual's behavioral data quantitatively, it can qualitatively predict the behavioral data. If the model becomes more complex, it is difficult to understand the principle of its core function. As a result, this model is not designed to focus on the entire perception-behavior output process, but only to explain some phenomena by parameter fitting. This model provides an open platform for different experimental situations, which greatly reduces the complexity of the experiment. The BG model in this paper simplifies some neural pathways, and future work will put on optimizing the model to obtain more accurate simulation data.

In future work, we will continue to explore the influence of other factors in the CBGTC model on decision-making. At the same time, the variable factors of the experiment should be enriched, so as to establish more relations between the discharge rate curve of the model and the RP.

Acknowledgment

This work was supported in part by the grants from the Natural Science Foundation of China (NSFC) (61573079,81741137), and the Fundamental Research Funds for the Central Universities to Dr. R. Liu. We appreciate all the participants and interns involved in this study.

References

[1] M. Rabinovich, A. Volkovskii, P. Lecanda, et al, "Dynamical encoding by networks of competing neuron groups: winnerless competition," Physical Review Letters[J], vol. 87, no. 6, p. 68102, 2001.

[2] A. Pascual, J. Modolo, A. Beute, "Is a computational model useful to understand the effect of deep brain stimulation in Parkinson's disease?" Journal of Integrative Neuroscience[J], vol. 5, no. 04, pp. 541-559, 2006.

[3] M. J. Frank, "Dynamic Dopamine Modulation of Striato-Cortical Circuits in Cognition: Converging Neuropsychological, Psychopharmacological and Computational Studies," University of Colorado at Boulder, Boulder, CO, 2004.

[4] C. Seer, F. Lange, D. Georgiev, et al, "Event-Related Potentials and Cognition in Parkinson's Disease: An Integrative Review," Neurosci Biobehav Rev[J], vol. 71, pp. 691-714, 2016.

[5] J. P. Dick, J. C. Rothwell, B. L. Day, et al,"The Bereitschaftspotential is abnormal in Parkinson's disease." Brain[J], vol. 112, no. Pt 1, pp. 233-244,.1989.

[6] B. Aisa, B. Mingus, R. O Reilly, "The Emergent neural modeling system," Neural Networks[J], vol. 21, no. 8, pp. 1146-1152, 2008.

[7] R. Ratcliff, M. J Frank, "Reinforcement-based decision making in corticostriatal circuits: mutual constraints by neurocomputational and diffusion models," Neural Comput[J], vol. 24, no. 5, pp. 1186-1229, 2012.

[8] Y. Si, X. Wu, F. Li, et al, "Different Decision-Making Responses Occupy Different Brain Networks for Information Processing: A Study Based on EEG and TMS," Cerebral Cortex[J], 2018.