NeuroImage 157 (2017) 1–12

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Neural substrates of updating the prediction through prediction error during decision making

Ying Wang[a], Ning Ma[a], Xiaosong He[a], Nan Li[a], Zhengde Wei[a], Lizhuang Yang[a], Rujing Zha[a], Long Han[a], Xiaoming Li[a,b], Daren Zhang[a], Ying Liu[c,**], Xiaochu Zhang[a,d,e,*]

[a] *Key Laboratory of Brain Function and Disease, School of Life Sciences, University of Science and Technology of China, Chinese Academy of Sciences, Hefei, Anhui 230026, China*
[b] *Department of Medical Psychology, Anhui Medical University, Hefei, Anhui 230032, China*
[c] *Provincial Hospital Affiliated to Anhui Medical University, Hefei, Anhui 230001, China*
[d] *School of Humanities & Social Science, University of Science and Technology of China Hefei, Anhui 230026, China*
[e] *Centers for Biomedical Engineering, University of Science and Technology of China Hefei, Anhui 230027, China*

ARTICLE INFO

Keywords:
Decision making
Updating the prediction
Prediction error
fMRI
HD-tDCS

ABSTRACT

Learning of prediction error (PE), including reward PE and risk PE, is crucial for updating the prediction in reinforcement learning (RL). Neurobiological and computational models of RL have reported extensive brain activations related to PE. However, the occurrence of PE does not necessarily predict updating the prediction, e.g., in a probability-known event. Therefore, the brain regions specifically engaged in updating the prediction remain unknown. Here, we conducted two functional magnetic resonance imaging (fMRI) experiments, the probability-unknown Iowa Gambling Task (IGT) and the probability-known risk decision task (RDT). Behavioral analyses confirmed that PEs occurred in both tasks but were only used for updating the prediction in the IGT. By comparing PE-related brain activations between the two tasks, we found that the rostral anterior cingulate cortex/ventral medial prefrontal cortex (rACC/vmPFC) and the posterior cingulate cortex (PCC) activated only during the IGT and were related to both reward and risk PE. Moreover, the responses in the rACC/vmPFC and the PCC were modulated by uncertainty and were associated with reward prediction-related brain regions. Electric brain stimulation over these regions lowered the performance in the IGT but not in the RDT. Our findings of a distributed neural circuit of PE processing suggest that the rACC/vmPFC and the PCC play a key role in updating the prediction through PE processing during decision making.

Introduction

During decision making, prediction error (PE) arises when there is a difference between expected and actual outcomes (Lak et al., 2014). The utilization of PE depends on the decision environment: in probability-known cases, because we never know the final result of this probabilistic event, there will be a PE after each event even if predictions are correct on average. However, the occurrence of PE does not indicate PE learning (i.e., using PEs to update the prediction) because we know that the probability never changes (Bach and Dolan, 2012). In an instructive task, there is no learning from PE for updating the prediction either (Li et al., 2011). However, in some other cases, for example, when the association between two consecutive events is unknown, such association can be learned via learning of PE. That is, the occurrence of PE would further be used to reshape the prediction and optimize strategies, which is the main idea of reinforcement learning (RL) (Davidow et al., 2016). Hence, it is evident from human behaviors that PE is used at different levels under different conditions.

Based on recent studies on learning, RL and Bayesian learning are the two major current frameworks (Mathys et al., 2011). When the environment's stochastic properties change over time (i.e., the environment is volatile) (Silvetti et al., 2013), "learning" refers to plasticity or switch, and Bayesian learning can beautifully capture it. However, in a stable environment with an unknown association between two consecutive events, "learning" refers to updating the prediction through PE. Under this condition, RL has been widely used in the analysis (d'Acremont et al., 2009; Glascher et al., 2010; Davidow et al., 2016).

Learning from PE for updating the prediction in RL has been a

---

[*] Corresponding author at: Key Laboratory of Brain Function and Disease, School of Life Sciences, University of Science and Technology of China, Chinese Academy of Sciences, Hefei, Anhui, 230026, China.
[**] Corresponding author.
E-mail addresses: felice828@126.com (Y. Liu), zxcustc@ustc.edu.cn (X. Zhang).

http://dx.doi.org/10.1016/j.neuroimage.2017.05.041
Received 1 September 2016; Accepted 17 May 2017
Available online 20 May 2017

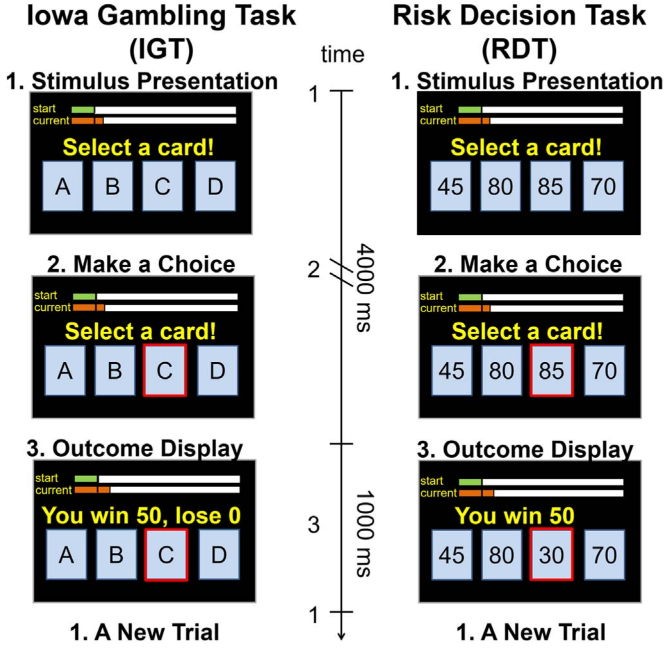

Fig. 1. An illustration of the experimental tasks. On the IGT and the RDT, each trial was divided into the following events: (1) during the choice phase, four seconds were provided for choice pondering and selection, and if no choice was made within 4 s, the computer would make a random choice; (2) during the outcome evaluation phase, the outcome was presented on the screen for 1 s. After each trial, the next trial began immediately without inter-trial intervals (ITIs). The green bar continually shows the base endowment throughout the task, and the orange bar shows the real-time accumulation of points.

long-lasting question in neuroscience. Many studies have found extensive of PE-related brain activations (e.g., the dopamine system including the striatum, the prefrontal cortex, etc.) (Schultz et al., 1997; Preuschoff et al., 2008; d'Acremont et al., 2009; Glascher et al., 2010; Bach and Dolan, 2012; Robinson et al., 2013). However, considering different levels of PE utilization under different conditions, we do not think that all these PE-related brain activations directly contribute to updating the prediction. However, most previous studies have only focused on general neural coding of PE and its alignment with behavior but did not compare the different kinds of PE processing directly (Schultz et al., 1997; Preuschoff et al., 2008; d'Acremont et al., 2009; Glascher et al., 2010; Bach and Dolan, 2012; Robinson et al., 2013). Therefore, the neural mechanism that specifically underlies learning from PE for updating the prediction during the decision making "still awaits formalization" (Bach and Dolan, 2012). Clarifying the specific roles of PE-related brain regions in RL may help us understand the function of the dopamine system.

In this study, we adopted the probability-unknown Iowa Gambling Task (IGT) (Bechara et al., 1994) and the probability-known Risk Decision Task (RDT) (Preuschoff et al., 2006) (Fig. 1). The invariable probability to win or lose on the IGT is unknown to the participants before the experiment and requires the participants to learn, while the probability to win or lose on the RDT is explicitly known to the participants. The only difference of PE processing between these two tasks is updating the prediction (Oya et al., 2005; Worthy et al., 2013). PEs are only represented on the RDT but are represented and used for updating the prediction on the IGT. Furthermore, we hypothesize a set of common PE-related brain regions between the IGT and the RDT that may be related to PE representation. We also hypothesize an additional set of PE-related brain regions preferentially activated on the IGT (i.e., IGT > RDT) that may be specifically associated with updating the prediction.

Most studies considered the reward PE in the RL (Glimcher, 2011; Schultz, 2016), but the risk PE should also be considered for RL as shown in this and a previous study (d'Acremont et al., 2009). In addition, d'Acremout et al. (2009) reported that reward and risk PE were processed by distinct neural circuits during RL. In this study, we tested whether this distinction is universal for both PE representation and updating the prediction through learning from PE. We hypothesize that reward PE-related brain regions are also related to risk PE during updating the prediciton because updating the prediction through learning from reward PE is assumed to be risk-sensitive (Tobler et al., 2005; Preuschoff and Bossaerts, 2007; Bossaerts, 2010). Therefore, brain regions specifically related to reward or risk PE and brain regions responding to both PEs need to be separated as clear as possible. For this purpose, we adopted a Multi-Voxel Pattern Analysis (MVPA)-like pattern analysis to improve the sensitivity of the response characteristics of a voxel (Rivolta et al., 2014; Dubois et al., 2015).

Through comparing the IGT and the RDT, we identified brain regions associated with PE representation and updating the prediction. Furthermore, two more ROI-based analyses were conducted to confirm our hypothesis. We explored the modulatory effect of uncertainty on the activation in these brain regions (Fan, 2014), and we tested functional connectivity between these ROIs and prediction-associated brain regions via psychophysical interaction (PPI) analysis.

Finally, because most functional magnetic resonance imaging (fMRI) data can provide correlation but cannot readily demonstrate necessity (Wang et al., 2014), we used high-definition transcranial direct current stimulation (HD-tDCS) to alter the activation in updating the prediction-related brain regions. We hypothesize that electric stimulation over these regions influences the performance on the IGT but not the RDT.

Materials and methods

Participants

Forty-nine healthy participants performed the IGT, and an independent sample of forty-one healthy participants performed the RDT. Data from eight participants on the IGT and one participant on the RDT were discarded owing to head movement in the MRI scanner (more than 2 mm) or to outlier properties (more than five missed choices or worse performance than the random data). Therefore, data from forty-one participants on the IGT (5 female; mean age ± SD, 22.42 ± 1.81 years; mean education ± SD, 16.02 ± 1.77 years) and forty participants on the RDT (5 females; mean age ± SD, 23.55 ± 2.14 years; mean education ± SD, 16.88 ± 1.87 years) were included in subsequent behavioral and fMRI analysis.

All participants were graduate or undergraduate students from the University of Science and Technology of China and reported no prior knowledge of the task. All participants were right-handed, and participants were excluded if they had any major medical illnesses, major psychiatric disorders, or neurological illnesses, or displayed gross structural abnormalities based on their T1-weighted images. No participants had a history of dependence (current or past) on any drug. Informed consent was obtained from all participants, and the study was approved by the Research Ethics Committee of the University of Science and Technology of China and conformed to the tenets of the Declaration of Helsinki.

Procedure

On both the IGT and the RDT, participants were first instructed to make repeated choices from four decks to gain or lose certain points and were informed that their payment would depend on their final accumulated points. Further, on the IGT, participants were told that there were "good decks" and "bad decks" that would result in a net gain or a net loss, respectively, in the long term, thus requiring them to learn; on the RDT, participants were instructed in accordance with the indicated task description (Fig. 1). Subsequently, participants com-