Fig.1. $log(t)$-$log(M)$ plot of our psychological experiments. The solid line computed using least squares program. The slopes are inverse D in the article. We suppose that data in our experiment include errors in the ratio of 10 percent. The error bar is calculated based on the assumption. (a) is a data of a female, (b) is of a male.

Fig.2. Distribution of learning time. The BP network learns 8 patterns with criterion 0.01. Each session starts from a random initial condition. Total number of data is 10000.

Fig.3. $log(M)$-$log(t)$ plot of our simulation. These are obtained from different size of network and different criterion value. (a) is of the mini-structre(8-3-8) and criterion is 0.10. (b) is of the size 16-10-16 with criterion 0.01. (c) is of the size 20-10-20 with criterion 0.05. 'Sumples' in the graph means the number of BPNs. We can see the tendency in this plot with small vibrating for all case. The learning time is the average of each distribution, not the time we proposed in the article. But it is almost no difference between the average case and the 20 percent one.