

Title	Opposing history effect of preceding decision and action in the free choice of saccade direction.
Author(s)	Mochizuki, Kei; Funahashi, Shintaro
Citation	Journal of neurophysiology (2014), 112(4): 923-932
Issue Date	2014-08-15
URL	<a href="http://hdl.handle.net/2433/208398">http://hdl.handle.net/2433/208398</a>
Right	© 2014 the American Physiological Society; This is the accepted manuscript of the article is available at <a href="http://dx.doi.org/10.1152/jn.00846.2013">http://dx.doi.org/10.1152/jn.00846.2013</a> .
Type	Journal Article
Textversion	author

Opposing history effect of preceding decision and action in  
the free choice of saccade direction.

Kei Mochizuki<sup>1</sup> and Shintaro Funahashi<sup>1,2</sup>

1. Laboratory of Cognitive Brain Science,  
Department of Cognitive and Behavioral Sciences,  
Graduate School of Human and Environmental Studies,  
Kyoto University, Kyoto 606-8501, Japan.

2. Kokoro Research Center, Kyoto University,  
Kyoto 606-8501, Japan.

Correspondence:

Shintaro Funahashi  
Kokoro Research Center  
Kyoto University  
46 Yoshida-Shimoadachi-cho, Sakyo-ku,  
Kyoto 606-8501, Japan  
funahashi.shintaro.2z@kyoto-u.ac.jp

Running Head:

History effect in free choice of saccade direction

## **Abstract**

When we voluntarily act, we make a decision to do so prior to the actual execution. However, due to the strong tie between decision and action, it has been difficult to dissociate these two processes in an animal's free behavior. In the present study, we tried to characterize the differences in these processes based on their unique history effect. Using simple eye-movement tasks in which the direction of a saccade was either instructed by a computer or freely chosen by the subject, we found that the preceding decision and action had different effects on the animal's subsequent behavior. While choosing a direction (previous decision) produced a positive history effect that prompted the choice of the same saccade direction, making a saccadic response to a direction (previous action) produced a negative history effect that discouraged the monkey from choosing the same direction. This result suggests that the history effect in sequential behavior reported in previous studies was a mixture of these two different components. Future studies on decision-making need to consider the importance of the distinction between decision and action in animal behavior.

Keywords:

Free choice, History effect, Decision-making, Memory-guided saccade

## **Introduction**

An animal's behavior is composed of sequences of actions. To behave properly in a complex environment, an animal has to monitor the results of its own actions and be able to change the next behavior flexibly based on previous experiences. If the previous actions were valid and produced a positive result, the animal should maintain the same behavioral strategy and repeat the same action. However, when previous actions were unfruitful or had a negative result, the animal must change its behavior to prevent future losses.

In recent studies, how an animal's behavior is influenced by the past trial history has been extensively investigated using probability learning tasks. By assessing the animal's behavior during the exploration of multiple options with different reward probabilities, the experimenter can examine the animal's ability to flexibly change its behavior based on past experience. For instance, previous studies have reported a history effect of past trials in monkeys (Lau and Glimcher 2005) and rodents (Kim et al. 2009; Sul et al. 2010, 2011). In these studies, the animals tended to continue to choose previously rewarded options, along with a general tendency for alternation from their own previous response. This competition between repeating and changing from the previous response is often explained in the context of "exploration and exploitation", which is advantageous when foraging in a natural environment (Kumar and Varaiya 1986; Bertsekas 1987; Stephens and Krebs 1987). However, in the framework of probabilistic learning, animals are

highly trained to always beware of a possible changeover of the reward contingency and to search for a new response strategy (Cohen et al. 2007). Therefore, the animal's behavior and resultant positive and negative history effect observed in these studies might be caused by the specific task structure in a probabilistic learning paradigm with contingency reversal.

Moreover, previous studies have mainly focused on how an animal's choice history and reinforcement history could predict its response in the current trial (Kim et al. 2009; Sul et al. 2011). In these studies, "the decision to make a particular action" and "the physical execution of the chosen action" were not distinguished because they necessarily coincided in their paradigm. However, when animals make a voluntary action, there must have been a decision to do so before the action is executed. Although this close tie between decision and action is inevitable and the difference in their behavioral significance usually receives little attention, these two processes are not equal. While action execution requires precise planning and the actual transmission of motor commands to lower centers, decision-making involves the representation and evaluation of multiple options for motor output. The difference in the biological basis of these two processes may result in different effects on future cognitive functioning. Therefore, a preceding decision and action can have different effects on subsequent behavior, and they can be dissociated under an appropriate experimental setup.

In the present study, we investigated how an animal's behavior is modified by

the preceding trial history using a memory-guided saccade paradigm with macaque monkeys (Funahashi et al. 1989; Goldman-Rakic 1995). To dissociate the history effect of preceding action and decision, we used two types of tasks in which the monkey made a memory-guided saccade toward an instructed or personally chosen spatial location (Watanabe and Funahashi 2004; Watanabe et al. 2006). While the monkeys were required to choose the saccade direction by themselves in the latter task, there was no such requirement in the former task. Thus, while the action that the monkeys finally needed to make (a saccadic eye-movement toward one direction) was identical in the two tasks, only in the latter task did they have to decide where to make the saccade. By contrasting the behavior in these two tasks that did and did not require the animal's own decision-making with regard to the saccade direction, we could dissociate the influence of the preceding action and decision on the following behavior. In addition, the monkeys were always rewarded with the same amount of juice regardless of which task they performed or where they chose to make a saccade. Therefore, in the present study, there was no task-structural bias for switching the response among possible actions, and the animal's natural tendency to repeat and change from the previous behavior could be examined. As a result, the preceding decision history had a positive effect which promoted saccades toward the same direction in subsequent trials. On the other hand, the preceding action history had an opposite effect, preventing the animals from repeating saccades toward the same direction. This finding offers new insight into

studies of animal behavior and emphasizes the importance of distinguishing “the action itself” from “the decision to perform the action” in cognitive science.

## **Materials and Methods**

### **Animals**

We used two Japanese monkeys (*Macaca fuscata*; monkeys O and E). The monkeys were housed in individual stainless steel home cages. Water intake was restricted in the home cage but provided as a reward in the laboratory. Additional vegetables and fruits were provided to fulfill the required water intake if necessary. All experimental procedures were conducted in accordance with the guidelines provided by the Primate Research Institute of Kyoto University and were approved by the Animal Research Committee at the Graduate School of Human and Environmental Studies, Kyoto University.

### **Apparatus**

During experimental sessions, the monkey sat in a primate chair in a dark sound-attenuated room with its head movements restricted by a head-holding apparatus. We used TEMPO software (Reflective Computing, Olympia, WA, USA) for task control and data acquisition. Visual stimuli were presented on a 20-inch CRT monitor (Dell

UltraScan D2026T-HS, Dell, Round Rock, TX, USA) that was placed 40 cm from the subject's face. A scleral search coil system (Enzanshi Kogyo, Tokyo, Japan) was used to monitor the monkeys' eye movements (Robinson 1963; Judge et al. 1980).

## **Tasks**

We used two types of memory-guided saccade tasks (Fig. 1a): the Instructed Choice Task (ICT) and Free Choice Task (FCT). In both tasks, a trial started with the presentation of a fixation point (white cross,  $0.5^\circ$  in visual angle) at the center of the monitor. After the monkey maintained fixation on the fixation point for 1.0 sec (fixation period), eight peripheral targets (white cross,  $0.75^\circ$  in visual angle) were presented at an eccentricity of  $13^\circ$  in visual angle ( $0^\circ$ – $315^\circ$ , separated by  $45^\circ$ ). The monkey had to neglect these targets and keep watching the fixation point for another 1.0 sec (pre-cue period). Next, one or two visual cues (filled white circle,  $2.5^\circ$  in visual angle) were briefly blinked over the peripheral targets for 0.5 sec (cue period). In the ICT, one cue was presented at one of the eight target locations. In the FCT, two identical cues were simultaneously presented at two peripheral locations. After the cues disappeared, the monkey had to maintain fixation for the following 1.5–3.0 sec (delay period). At the end of the delay period, the fixation point was turned off, and the monkey was required to make a memory-guided saccade toward the cued location. A saccade to either of the two locations was regarded as correct in FCT trials. Every correct response was rewarded



by a drop of juice, and there was no difference in the amount of reward regardless of the monkey's choice in the FCT or the type of the task.

ICT and FCT trials were not arranged in a blockwise fashion but were intermingled in a random order from trial to trial. The location of the cue in ICT trials was randomly determined as one of eight peripheral target locations. In FCT trials, possible cue locations were limited to four vertical and horizontal locations (0°, 90°, 180° and 270°) to reduce the number of combinations of cue locations. In an FCT trial, cues were randomly presented at two of these four possible locations. Accordingly, trials consisted of eight cue conditions in the ICT and six pair conditions in the FCT. The proportion of FCT trials was approximately three times greater than that of the ICT to promote data collection regarding choice behavior.

## **Surgery and Training Procedure**

We implanted a stainless steel head-holding device and a scleral search coil in the monkeys. A scleral search coil was implanted onto the right eye globe by dissecting the conjunctiva (Judge et al. 1980). The monkeys were first anesthetized by an intramuscular injection of ketamine hydrochloride (10 mg/kg) and then an intravenous injection of pentobarbital sodium (10–15 mg/kg). Heart rate and respiration were monitored during the surgery. Stainless steel screws were put into the skull to ensure firm adhesion of the head-holding device. The connector for the search coil and the

head-holding device were fixed to the skull with dental acrylic. All of the surgical procedures were performed under aseptic conditions.

After the monkeys recovered from surgery, we started the training of the tasks. We first trained the monkeys with the ICT. When the monkeys learned to perform the ICT (about 85% correct on more than five consecutive experimental sessions), we started to intermingle FCT trials with ICT trials.

## **Data Analysis**

All statistical analyses and data plotting were performed using the statistical software R 3.0.2 (R Core Team 2013). We used Bonferroni's correction method in testing statistical significance in multiple comparisons unless otherwise noted. The proportion correct was separately calculated for the ICT and the FCT by dividing the number of trials with correct target capture by the number of trials that the animal reached the response period. In order to examine the individual differences of directional preference between the monkeys, we applied chi-squared tests to the number of choices in the FCT. We first pooled the sessions within individuals and then performed chi-squared tests to compare the number of choices in each pair condition between the subjects (pair-based preference). We also compared the number of the choice of a given direction against the number of trials in which that direction was available between the subjects (direction-based preference).

## Logistic regression

To quantify how behavioral history in past trials would influence an animal's own choice of saccade direction in future trials, we performed a logistic regression analysis. We investigated how the monkey's choice in FCT trials could be explained based on the behavioral history in the 10 preceding trials. In this analysis, the effectiveness of preceding ICT and FCT trials for predicting the monkey's choice was separately estimated. Additionally, we divided preceding FCT trials into two types: those in which the monkey was presented with the same cue pair (FCTs trial) and those with different cue pairs (FCTd trial) compared to the current trial. Based on this categorization, the 10 preceding trials were classified into three trial types (ICT, FCTs and FCTd) with regard to the current trial. By applying a logistic regression analysis, we measured how the monkey was likely to repeat choosing the same direction for a saccade that was instructed or freely chosen in preceding trials. For each monkey and for an arbitrary pair condition with directions A and B, we first collected the subject's choice in FCT trials with that pair condition and the prior behavioral sequence up to 10 trials before each of those FCT trials from all the sessions available for that subject. Then we estimated the history effect of preceding trials on the monkey's choice in the current trial using the following logistic regression model:

$$\log\left(\frac{p_A}{p_B}\right) = \sum_{j=-1}^{-10} \iota_j \cdot i_j + \sum_{j=-1}^{-10} \sigma_j \cdot s_j + \sum_{j=-1}^{-10} \delta_j \cdot d_j + \sum_{j=-1}^{-10} \nu_j \cdot u_j + \gamma$$

where  $p_A$  and  $p_B$  are the probabilities of choosing directions A and B. The variables  $i_j$ ,  $s_j$ ,  $d_j$  correspond to the effect of the previously executed behavior and are 1 if the monkey responded to direction A, -1 if the monkey responded to direction B, and 0 if the monkey responded to neither direction A nor direction B in the  $j$ -th preceding ICT, FCTs and FCTd trials, respectively. The variables  $u_j$  correspond to the effect of the previously withdrawn behavior and are 1 if the monkey gave up direction A, -1 if the monkey gave up direction B, and otherwise 0 in the  $j$ -th preceding FCTd trials. The coefficients  $\iota_j$  (ICT),  $\sigma_j$  (FCTs),  $\delta_j$  (FCTd chosen) and  $\nu_j$  (FCTd unchosen) denote the four types of the history effect explained above, and  $\gamma$  is a bias term. Note that there is no direction given up in the ICT, because only one cue was presented. Also, for the preceding FCTs trials, the effect of not choosing a direction is not dissociable from the effect of choosing, because not choosing direction B is the same to the choice of direction A. Therefore, the effect of not choosing could be a subject for the investigation only in the preceding FCTd trials. The lags for the trials ( $j$ ) were computed based on the number of trials from the current trial counting all the types of previous trials (ICT, FCTs and FCTd) altogether. For example,  $\iota_{-3}$  is the effect of the previous ICT trial located exactly three trials before the current trial, and not the effect of the third latest ICT trial by selectively counting the previous ICT trials (Fig. 1b). Estimated coefficients were then averaged across six pair conditions, resulting in 40 coefficients for each monkey (excluding the bias term).

To examine the statistical significance of the estimated history effect, we used a shuffling method. For each session's data, we shuffled the indices for FCT trials respectively among each of the six pair conditions. Thus, the monkey's choice in every FCT trial was exchanged with that of the other FCT trial with the same cue pair in that session. By this manipulation, we uncoupled the result of the monkey's choice from the trial history prior to each FCT trial. The sequence of task and pair conditions as well as the overall number of choices in each FCT pair condition were kept unchanged. Since the potential influence of the particular order of trial conditions possible in the data was preserved in this shuffled sample, we could dissociate the genuine effect of the behavioral history by comparing the results of regression for real data to those for shuffled data. We performed resampling 2000 times for each subject and calculated the confidence intervals of every behavioral measure in the logistic regression analysis.

### **Reinforcement learning models**

To directly estimate the history effect of preceding decision and action, we used a reinforcement learning model to fit the monkeys' performance (Sutton and Barto 1998). For each session, the values of each saccade direction were maintained separately and updated based on the monkey's behavior. We assumed that choosing, not choosing, and responding to a direction would independently influence the values of the chosen, unchosen and responded directions. In contrast to the probabilistic learning literature (Lau and Glimcher 2005; Kim et al. 2009; Sul et al. 2010, 2011), the monkeys in the

present study could obtain the same amount of reward for every correct trial, regardless of the task or response direction. Therefore, the value of a direction was updated by simply adding a different increment based on how the response was made. We used three increment parameters for values of directions, that were responded ( $\Delta_R$ ), chosen ( $\Delta_C$ ) or unchosen ( $\Delta_U$ ) by the animal. The value for direction A in the  $i$ -th trial ( $Q_{A, i}$ ) was updated by:

$$Q_{A, i+1} = \alpha \cdot Q_{A, i} + \Delta$$

where  $\Delta$  is  $\Delta_R$  for the response direction in an ICT trial,  $\Delta_R + \Delta_C$  for the response direction in an FCT trial,  $\Delta_U$  for the direction presented but unchosen in an FCT trial, and 0 for directions that were not presented in that particular trial. Coefficient  $\alpha$  is a discount factor. For an arbitrary pair condition with directions A and B, the probability to choose direction A in trial  $i$  ( $p_{A, i}$ ) was calculated according to a softmax function as follows:

$$p_{A, i} = \frac{\exp(Q_{A, i})}{\sum_{a \in \{A, B\}} \exp(Q_{a, i})}$$

We compared three reinforcement learning models by embedding different combinations of the increment parameters to examine which model best illustrates the animal's behavior. In the first model, we implemented only  $\Delta_R$  parameter for the value update and other two parameters ( $\Delta_C$  and  $\Delta_U$ ) were set to 0 (R model). Therefore, the value of the response direction was equally updated in the ICT and FCT. Although there is no variable to denote the presence or absence of the reward because of the

guaranteed reward delivery in the present study, this model corresponds to a standard Q-learning model in which the values are updated only by the experienced outcome (Watkins and Dayan 1992; Sutton and Barto 1998). In the second model, we used  $\Delta_R$  and  $\Delta_C$  parameters but  $\Delta_U$  was set to 0 (R+C model). Therefore, the value of the response direction in the FCT was updated by  $\Delta_C$  in addition to  $\Delta_R$ . In the third model, we used all the  $\Delta_R$ ,  $\Delta_C$  and  $\Delta_U$  parameters (R+C+U model). Therefore, the value of the unchosen direction presented in the FCT was differently updated. This model corresponds to the models with value updates for unchosen alternatives used in previous studies (Lee et al. 2005; Abe and Lee 2011; Hayden et al. 2009). For each session's data, we applied all the three models to the animal's behavior and calculated the parameters in the models using the maximum-likelihood estimation method with the optim function in R (Nelder and Mead 1965; Nash 1980). We compared the goodness of the model by calculating Akaike's information criteria (AIC) to investigate whether choosing and not choosing a direction in the animal's own decision exhibited unique effect on the values of the spatial locations. We also compared each model's prediction performance of the animal behavior by a leave-one-out cross-validation. For each model, we averaged the parameters of the model across all the sessions except for one test session, and then applied the same model with the averaged parameters to the test session. The prediction performance was calculated as a proportion of FCT trials in which the value of the chosen direction estimated under the model was greater than that of the unchosen

direction.

## Results

### General Performance

Both monkeys learned to perform well in the tasks. Monkey O performed 98 sessions. Figure 2a shows the example of saccadic responses in the ICT and FCT. Each session consisted of about 600 trials and the average proportion correct was 96.1% for the ICT and 98.0% for the FCT. Monkey E performed 98 sessions. Each session consisted of about 500 trials and the average proportion correct was 98.1% for the ICT and 97.9% for the FCT. The paired *t*-test revealed a slightly higher task performance for the FCT in monkey O ( $p < .001$  for monkey O,  $p = .40$  for monkey E). This might be related to the higher chance level to capture a correct location in the FCT if the monkey makes a saccade to one of the eight target locations by guessing. (There were two correct locations in the FCT, rather than just one.) However, in general, the monkeys seldom made erroneous responses in either task. In total, data for 59454 and 47301 correct trials for each monkey were acquired and used in the further analysis.

The monkeys exhibited directional preferences, but chose both directions in each FCT pair condition (Fig. 2b). Overall, the choice proportion was most biased in the upper vs lower condition (0.26:0.74) for monkey O and right vs lower (0.64:0.36) for



monkey E. To compare the preference for saccade directions between the monkeys, we applied chi-squared tests to the proportion of choices in each pair conditions. The proportion of choosing each of the two directions was different between the monkeys in all pair conditions except for one condition (chi-squared test, corrected  $p = .74$  for right vs lower pair condition,  $p < .001$  for other pair conditions). The proportions of trials in which the monkey chose right, upper, left and lower direction out of the total number of trials in which that direction was presented in the FCT were 65.1%, 31.3%, 49.5% and 54.7% for monkey O and 53.6%, 53.8%, 52.5% and 40.0% for monkey E, respectively. This direction-based rather than the pair-based preference was different between the monkeys for all the four directions (chi-squared test, corrected  $p < .01$  for all directions). The inconsistency of directional preference between the monkeys suggests that the preference for a particular direction was not a principal source of the characteristic history effect reported in the subsequent analysis.

### **History Effect based on the Logistic Regression Analysis**

To quantify how preceding trials influence the free choice of saccade direction in the current trial, we performed a logistic regression analysis. In this model, the history effect of 10 preceding trials was used as regressors to explain the choice in the current FCT trial. Regression was performed separately for each FCT pair condition. The effects of preceding ICT and FCT trials were independently estimated as different

predictors. The effects of preceding FCT trials were further differentiated into two conditions: i.e., the effects of preceding trials with the same (FCTs) and different (FCTd) cue pairs compared to the current trial (see Materials and Methods). Also, the effect of choosing and not choosing a direction was separately examined for preceding FCTd trials. Therefore, there were three conditions of preceding trials (ICT, FCTs and FCTd) that constituted four types of history effect (preceding ICT, FCTs, FCTd chosen/unchosen), and a set of 41 regression coefficients (10 preceding trials for 4 conditions, and a bias term) were estimated for each FCT pair condition.

Figure 3a,c shows the averaged coefficients of logistic regression for the six FCT pair conditions. In both monkeys, the coefficients for preceding ICT drastically decreased 3–4 trials prior to the current trial. This indicates that the monkeys were less likely to choose the direction that they had been forced to respond to immediately before that trial. The coefficients of preceding FCT with different cue pairs (FCTd) were similar to those in the ICT except for a slight positive shift in the chosen direction and negative shift in the unchosen direction. This suggests that the monkeys had a general tendency to avoid choosing the directions available in the previous trials, regardless of whether they were chosen or not.

On the other hand, the coefficients of preceding FCT with a cue pair identical to that in the current trial (FCTs) were largely higher than those of the ICT in both monkeys. In monkey E, the coefficients for the FCTs condition were positive throughout the

preceding trials, indicating that the monkey was likely to choose the same direction when an identical cue pair was repeatedly presented in adjacent FCT trials. There was a slight increase in the effect in the latest and second latest preceding FCTs trials. In monkey O, the coefficients for the FCTs condition were also positive for most of the preceding trials, but decreased to nearly zero in one trial before the current trial. This indicates that the monkey's tendency to choose the same direction in the FCTs condition disappeared when the same FCT condition was repeatedly presented in two consecutive trials. Thus, there was an individual difference in history effect when the preceding trials were divided into four conditions.

In the logistic regression analysis based on the shuffled data, the results were identical between two monkeys (Fig. 3b,d). Because the choice in the current FCT trial is uncoupled from its original history of trial sequence, the coefficients remained constant across preceding trials in each condition. In both monkeys, the regression coefficients for preceding ICT trials was virtually zero. For the FCT, the coefficients for preceding FCTs trials and choosing in preceding FCTd trials were positive. The coefficients for not choosing in preceding FCTd trials was negative. These history effects in FCT trials can be attributed to the animal's overall directional preference in each session, which lead to correlated choices by chance in FCT trials within the same session.

## Dissociation of Decision and Action History

The logistic regression model described above was based solely on the categorization of the trial type, and does not consider the cognitive processes in each condition. Therefore, we further examined the effect of preceding decision and action history based on the coefficients obtained from the logistic regression. In ICT trials, the monkeys were only required to make a forced response toward the instructed direction and there was no room for a choice. Therefore, the regression coefficients obtained from preceding ICT trials could be directly regarded as the effect of preceding actions. On the other hand, in FCT trials, the monkeys needed to both choose the saccade direction by themselves and make a response toward it. In this regard, preceding FCT trials have both decision and action aspects, which could independently influence the current trial. Therefore, the estimated coefficients for preceding FCT trials would be a mixture of decision and action history effects.

To separate the effect of decision history from that of action history and solely examine the former, we subtracted the regression coefficients for the ICT from those for the FCT. The history effect of decision in the  $j$ -th preceding trial was calculated as  $\sigma_j - \iota_j$  for FCTs trials and  $\delta_j - \iota_j$  for FCTd trials. These differences of the regression coefficients between FCT and ICT trials could represent how the monkey's tendency to make a particular choice overcame the effect of the past saccadic response toward that direction.

Figure 4 shows the result of the dissociation of decision history from action history. The effect of decision history in preceding FCTs trials was positive for both monkeys throughout the 10 trials of the analytic period and exceeded the confidence intervals calculated from the shuffled data (Fig. 4b) during the 3–4 trials before the current trial. On the other hand, the effect of decision history in preceding FCTd trials were weak for both monkeys and did not exceed the confidence intervals from the shuffled samples. These results indicate that the monkeys tended to repeat the same decision in adjacent trials, but this effect of decision history was specific to the pair conditions. The individual difference of the regression coefficients for preceding FCTs trials between the monkeys (Fig. 3) could be the result of a strong negative action history effect in monkey O, which could have canceled the positive effect of decision history and led to a deterioration of choice predictability in the FCTs condition in one trial before the current trial.

### **History Effect based on the Reinforcement Learning Models**

The results of the logistic regression analysis suggested that the preceding decision and action have different effects on the choice in subsequent trials. Therefore, we tried to directly characterize these effects by applying reinforcement learning models in which the effects of decision and action on the value for each saccade direction were separately implemented (see Materials and Methods). In the models, each direction retained its

own value, and the choice in FCT trials was determined based on a comparison of the values of the two presented directions. We used three models. In all the models, the value of the responded direction was updated by  $\Delta_R$  in both ICT and FCT trials. In two of the models, we tested the effect of the animal's own decision of saccade direction by further updating the value of the responded direction by  $\Delta_C$  in FCT trials. In one of these models, we also tested the effect of not choosing a direction by updating the value of the unchosen direction by  $\Delta_U$  in FCT trials.

We fit the models to the animal's behavior separately for each session. Table 1 shows the averaged parameters and the goodness of each model based on the AIC and the prediction performance. The models with  $\Delta_C$  parameter (R+C model and R+C+U model) performed better than the model with  $\Delta_R$  parameter alone (R model) in most of the sessions (96 of 98 sessions for monkey O, 95 of 98 sessions for monkey E). Between the former models, the model with  $\Delta_U$  parameter was more frequently selected, but the prediction performance was similar between the two models. In both R+C and R+C+U models,  $\Delta_R$  was significantly smaller than zero, indicating that making a saccadic response to a direction decreased its value (Wilcoxon signed rank test, corrected  $p < .01$  and  $p < .001$  for R+C and R+C+U models in monkey O,  $p < .01$  and  $p < .05$  in monkey E). On the other hand,  $\Delta_C$  was significantly larger than zero, indicating that choosing a direction increased its value (Wilcoxon signed rank test, corrected  $p < .001$  and  $p < .05$  for R+C and R+C+U models in monkey O,  $p < .001$  for both models in monkey E).

In R+C+U model,  $\Delta_U$  was significantly smaller than zero (Wilcoxon signed rank test, corrected  $p < .001$  in both monkeys). These results indicate that “deciding to make an action” and “making an action itself” have opposite effects on the value of the executed behavior, which leads to the differential effects of decision and action history observed in the logistic regression analysis.

## **Discussion**

In the present study, we found that preceding decision and action had different effects on the subsequent behavior of the animal. These two behavioral components are generally considered to be indissociable in studies on decision-making, since the result of a decision always needs to be carried out as a physical action to be observed by the experimenter. In other words, the strong tie between decision and action makes it difficult to differentiate their effects on animal behavior. However, in the present study, we could successfully separate these two components by comparing the history effects in the ICT and the FCT. Choosing a direction produces a positive history effect that promotes the choice of the same saccade direction when the identical choice situation is presented repeatedly. On the other hand, making a saccadic response to a direction produces a negative history effect and discourages the monkey from choosing the same direction. Both of these effects were strongly manifested during 3–4 trials. Consistent results were obtained from the analysis using the reinforcement learning model.

## **Competing Tendencies for Repetitive Behavior and Repetition**

### **Avoidance**

It is known that human subjects generally tend to avoid repeating the same behavior (Brugger et al. 1996). For example, subjects make more changes in their response than mathematically expected when asked to generate a random sequence of letters or numbers (Rabinowitz et al. 1989; Brugger 1997; Falk and Konold 1997). On the other hand, it is also known that humans sometimes show a robust tendency to repeat the same behavior (Sandson and Albert 1984). Persistence in repetitive behavior becomes particularly apparent in patients with brain damage and disorders, such as perseveration in motor control (Annoni et al. 1998; Na et al. 1999), verbal expression (Cohen and Dehaene 1998) and non-verbal expression (Goldberg and Tucker 1979; Cosentino et al. 2004). Perseveration is evident even in higher-order cognitive functions, as seen in the Wisconsin card sorting test (Milner 1963; Nelson 1976). Thus, humans seem to possess opposing behavioral tendencies that support and discourage repetitions.

However, in animal studies on decision-making, most previous experiments have been performed with tasks that provide behavioral options with different reinforcement probabilities. For example, in studies using a concurrent reward schedule, animals are provided with two response options (e.g. left or right key) and allowed to choose one of them (Herrnstein and Loveland 1974; Shapiro and Allison 1978). These two options differ with respect to their reward ratio or interval and the animals need to find which



one is more advantageous by trial and error. In addition, after the animals learn the highly rewarding option, the reward contingency with two options changes without any announcement. Under such a task contingency, animals should continue to choose the same option once they identify which is advantageous. Nevertheless, they also have to keep monitoring the results of past trials and sometimes seek other choice options to deal with a sudden change in the reward contingency. Accordingly, animals often exhibit a mixed strategy of choice repetition and occasional alternation, which is reminiscent of the competing trends for repetition and repetition avoidance. However, this behavior may be an artifact of reinforcement schedules which include contingency reversal that are characteristic in animal studies. Recent studies using reinforcement learning models for probabilistic learning (Sutton and Barto 1998; Samejima et al. 2005) may also have the same problem. Therefore, a detailed investigation of an animal's choice behavior under a task structure without contingency reversal was needed to investigate the conflict between repetition and repetition avoidance that has been suggested in human studies.

In the present study, there was no task-structural constraint for the monkeys to alternate choices because the same amount of reward was always delivered regardless of the saccade direction chosen in the FCT. With this experimental setup, we could avoid the undesirable bias for response alternation present in the probabilistic learning paradigm, and examine the animal's innate tendency for repetition and repetition avoidance. As shown in Results, both monkeys exhibited opposing history effect which

promoted and restrained response repetition. This finding suggests that the competition between repetition and repetition avoidance is a common behavioral principle in humans and non-human animals which organizes the animal's choice behavior. Our work links human and animal studies regarding the choice between equally valuable options, and provides insights for future investigations of the nature of internally driven free choice behavior.

### **Generalization to Other Decision-Making Tasks**

In the present study, the monkey could obtain a reward every trial regardless of the task or response direction. Application of reinforcement learning models to the monkey's behavior revealed that making a response decreased the value function of the response direction (Table 1). However, in standard reinforcement learning models, the value of a given action increases if the animal receives a reward after that response (Sutton and Barto 1998; Samejima et al. 2005). In fact, the increment of the value function after a reward is one of the most fundamental features of the reinforcement learning model, which enables the animal to learn an appropriate behavioral strategy. In this regard, the result reported in the present study may appear to be strange because it claims a decrease in the value function every time the monkey made an action to a direction and got rewarded. This discrepancy could be attributed to the task design used in the present study. In contrast to the decision situation in probabilistic learning, the animals in the

present study were able to receive a reward every trial, and didn't need to search for an advantageous behavioral strategy. Thus, the animals' behavior in the present study was not determined in terms of reward maximization. Therefore, the reinforcement learning models and estimated value updates in the present study cannot be directly compared to the other existing models for probabilistic learning.

The difference in task structures between the present and previous experiments could raise a limitation in generalizing the present results. The decision-making tasks used in cognitive science are composed of various complex features such as reward expectations and risks in value-based decision-making (Platt and Glimcher 1999; Barraclough et al. 2004; Sugrue et al. 2004; Dorris and Glimcher 2004), discrimination about physical properties of the stimuli in perceptual judgement (Newsome et al. 1989; Shadlen and Newsome 1996; DeAngelis et al. 1998; de Lafuente and Romo 2005) and the relationship and interaction between actors in social decision-making (Yoshida et al. 2011; Chang et al. 2013). It is not clear whether the animal's tendency to avoid and prefer repetitive behaviors observed under a free choice condition with none of these features similarly operates in other decision situations. However, because the free choice paradigm we used did not depend on any of these specific task components, the present results could represent an innate behavioral tendency of animals that may take place in any decision-making situations. According to the task constraints which strongly determine the optimal behavioral strategy in probabilistic learning and other

types of tasks, whether the animal's natural tendency can be actually observed or not depends on each experimental paradigm. Nevertheless, the knowledge about the animal's fundamental traits may help to understand the behavioral patterns under various task constraints which take place over the pre-existing natural tendency, and possibly resolves the inconsistency in observed behavior among different studies and tasks. Further studies are needed to investigate how the animal's own tendency in behavioral repetition is taken over by the requirements of the tasks.

### **Context Specificity of the Decision History Effect**

In the logistic regression analysis, we differentiated the preceding FCT trials with the same (FCTs) and different (FCTd) cue pair compared to the current trial, and examined their effects separately. Although these two conditions were identical in respect to “the effect of the past freely-chosen response toward one of the two directions available in the current trial”, they had quite different effects on the monkey's behavior (Fig. 3). The choice in preceding FCTd trials exhibited a similar effect to the preceding ICT trials, and strongly prevented a repeated response to the same direction in consecutive trials. In contrast, the choice in preceding FCTs trials showed consistently greater coefficients than those of FCTd and ICT trials, and mostly facilitated the repetitive response to the same direction throughout the 10 trials of the preceding trial period, except for the -1 trial for monkey O. Even in this exceptional case, the coefficients for the FCTs condition

was neutral (about zero) and much greater than the coefficients for the ICT and FCTd conditions. In other words, according to the effect of preceding ICT and FCTd trials, monkey O had a robust tendency to avoid repeating the same response, but this tendency was counteracted for preceding FCT trials with a cue pair identical to the current trial. This result clearly indicates that making a decision promotes the same decision in subsequent trials, but this effect is highly context-dependent and only manifests in trials with identical choice options. By cancelling out the effect of preceding actions from the effect of preceding decisions (Fig. 4), a consistent result was observed. While the decision history in preceding FCTs trials positively influenced the choice in the current trial, the decision history in preceding FCTd trials exhibited only a small effect that did not exceed the chance level taken by a shuffling method.

These results suggest the importance of the decisional situation in sequential behavior and the history effect it exerts. Not only which action was chosen by the animal, but also among which of the alternatives it was chosen may be a key to the history effect it exerts on subsequent trials. However, the effect of choosing in preceding trials was examined independently from the effect of not choosing only in FCTd trials. The dissociation of the effects of choosing and not choosing was possible because the cues presented in preceding FCTd trial were different from those in the current trial. In FCTs trials, two cues available in the preceding and current trials were identical and therefore not choosing a direction necessarily means the choice of the

other direction. Because of this coupling, we could not dissociate the effect of choosing and not choosing in preceding FCTs trials. Therefore, it is not clear whether the effect of the preceding FCTs trials was derived from a unique effect of choosing in the same decisional context, or can be fully explained in terms of the negative history effect of not choosing that might have taken place in these trials. Because the strong coupling of choosing and not choosing in the same set of alternatives will cause a multicollinearity of the explanatory variables, the dissociation of these features is difficult. Future studies using experimental tasks with three or more action alternatives in a trial to sufficiently control the risk of multicollinearity will reveal the respective effects of choosing and not choosing under the same and different decisional context in animal's free choice behavior.

### **Dependence on the Effector**

In the present study, we used memory-guided saccade tasks in which the animals had to express their decision by saccadic eye movement. However, with respect to information-seeking, making a saccade may not be a mere designation of a location but rather a completion of scanning that visual field. Therefore, it is possible that the monkeys' choice behavior in the present study was affected by their innate behavior in visual foraging, and thus limited to the saccadic motor domain. Since the monkeys could freely see everywhere during the inter-trial interval, it is not likely that the urge to

watch a certain location of the display made a large contribution to the choice behavior in the present study. Still, it is not clear our results regarding decision and action history effects can be generalized to other effectors.

Previously, Lee and Schieber (2006) reported a monkey's tendency to switch its response direction in a hand-reaching task. When monkeys were allowed to choose from left and right targets by themselves, they frequently changed their response from the previous trial. This switching bias may seem to contradict the positive history effect of preceding decision observed in the present study. However, they also reported that the monkeys showed a robust tendency to change the hand used in target-reaching in consecutive trials, as well as a strong preference to choose a target that was ipsilateral to the hand in use. Based on this finding, Lee and Schieber (2006) suggested that part of the monkeys' tendency to change the target choice could have been driven by the bias to switch the hand in use. Also, the effects of preceding decision and action were not explicitly separated in their analysis. Therefore, further studies will be needed to determine whether the difference between our result and those in the previous study was caused by a difference in the effector or some other experimental factors.

## **Grants**

This research was supported by Grants-in-Aid for Scientific Research (21240024 and 25240021) from the Ministry of Education, Culture, Sports, Science and Technology

(MEXT) of Japan to S. F., and a Grant-in-Aid for JSPS Fellows (23-7155) from the Japan Society for the Promotion of Science to K. M. The animals were provided by the National BioResource Project “Japanese Monkeys” supported by the MEXT, Japan.



## References

- Abe H, Lee D. Distributed coding of actual and hypothetical outcomes in the orbital and dorsolateral prefrontal cortex. *Neuron* 70: 731–741, 2011.
- Annoni G, Pegna A, Michel C, Estade M, Landis T. Motor perseverations: a function of the side and the site of a cerebral lesion. *Eur Neurol* 40: 84–90, 1998.
- Barracough DJ, Conroy ML, Lee D. Prefrontal cortex and decision making in a mixed-strategy game. *Nat Neurosci* 7: 404–410, 2004.
- Bertsekas DP. *Dynamic Programming: Deterministic and Stochastic Models*. Englewood Cliffs, NJ: Prentice Hall, 1987.
- Brugger P. Variables that influence the generation of random sequences: an update. *Percept Mot Skills* 84: 627–661, 1997.
- Brugger P, Monsch AU, Johnson SA. Repetitive behavior and repetition avoidance: the role of the right hemisphere. *J Psychiatry Neurosci* 21: 53–56, 1996.
- Chang SWC, Garipis JF, Platt ML. Neuronal reference frames for social decisions in primate frontal cortex. *Nat Neurosci* 16: 243–250, 2013.
- Cohen JD, McClure SM, Yu AJ. Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philos Trans R Soc Lond B Biol Sci* 362: 933–942, 2007.
- Cohen L, Dehaene S. Competition between past and present. assessment and interpretation of verbal perseverations. *Brain* 121 (Pt 9): 1641–1659, 1998.
- Cosentino S, Jefferson A, Chute DL, Kaplan E, Libon DJ. Clock drawing errors in dementia: neuropsychological and neuroanatomical considerations. *Cogn Behav Neurol* 17: 74–84, 2004.
- de Lafuente V, Romo R. Neuronal correlates of subjective sensory experience. *Nat Neurosci* 8: 1698–1703, 2005.
- DeAngelis GC, Cumming BG, Newsome WT. Cortical area mt and the perception of stereoscopic depth. *Nature* 394: 677–680, 1998.
- Dorris MC, Glimcher PW. Activity in posterior parietal cortex is correlated with the relative subjective desirability of action. *Neuron* 44: 365–378, 2004.
- Falk R, Konold C. Making sense of randomness: Implicit encoding as a basis for judgment. *Psychol Rev* 104: 301–318, 1997.

- Funahashi S, Bruce CJ, Goldman-Rakic PS. Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J Neurophysiol* 61: 331–349, 1989.
- Goldberg E, Tucker D. Motor perseveration and long-term memory for visual forms. *J Clin Neuropsychol* 1: 273–288, 1979.
- Goldman-Rakic PS. Cellular basis of working memory. *Neuron* 14: 477–485, 1995.
- Hayden BY, Pearson JM, Platt ML. Fictive reward signals in the anterior cingulate cortex. *Science* 324: 948–950, 2009.
- Herrnstein RJ, Loveland DH. Hunger and contrast in a multiple schedule. *J Exp Anal Behav* 21: 511–517, 1974.
- Judge SJ, Richmond BJ, Chu FC. Implantation of magnetic search coils for measurement of eye position: an improved method. *Vision Res* 20: 535–538, 1980.
- Kim H, Sul JH, Huh N, Lee D, Jung MW. Role of striatum in updating values of chosen actions. *J Neurosci* 29: 14701–14712, 2009.
- Kumar P, Varaiya P. *Stochastic Systems: Estimation, Identification and Adaptive Control*. Englewood Cliffs, NJ: Prentice Hall, 1986.
- Lau B, Glimcher PW. Dynamic response-by-response models of matching behavior in rhesus monkeys. *J Exp Anal Behav* 84: 555–579, 2005.
- Lee D, McGreevy BP, Barraclough DJ. Learning and decision making in monkeys during a rock-paper-scissors game. *Brain Res Cogn Brain Res* 25: 416–430, 2005.
- Lee D, Schieber MH. Serial correlation in lateralized choices of hand and target. *Exp Brain Res* 174: 499–509, 2006.
- Milner B. Effects of different brain lesions on card sorting: The role of the frontal lobes. *Arch Neurol* 9: 90–100, 1963.
- Na DL, Adair JC, Kang Y, Chung CS, Lee KH, Heilman KM. Motor perseverative behavior on a line cancellation task. *Neurology* 52: 1569–1576, 1999.
- Nash JC. *Compact Numerical Methods for Computers: Linear Algebra and Function Minimisation*. Adam Hilger, Bristol, UK: Adam Hilger, 1980.
- Nelder JA, Mead R. A simplex method for function minimization. *Comput J* 7: 308–313, 1965.
- Nelson HE. A modified card sorting test sensitive to frontal lobe defects. *Cortex* 12: 313–324, 1976.

- Newsome WT, Britten KH, Movshon JA. Neuronal correlates of a perceptual decision. *Nature* 341: 52–54, 1989.
- Platt ML, Glimcher PW. Neural correlates of decision variables in parietal cortex. *Nature* 400: 233–238, 1999.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2013.
- Rabinowitz F, Dunlap WP, Grant MJ, Campione JC. The rules used by children and adults in attempting to generate random numbers. *J Math Psychol* 33: 227–287, 1989.
- Robinson DA. A method of measuring eye movement using a scleral search coil in a magnetic field. *IEEE Trans Biomed Eng* 10: 137–145, 1963.
- Samejima K, Ueda Y, Doya K, Kimura M. Representation of action-specific reward values in the striatum. *Science* 310: 1337–1340, 2005.
- Sandson J, Albert ML. Varieties of perseveration. *Neuropsychologia* 22: 715–732, 1984.
- Shadlen MN, Newsome WT. Motion perception: seeing and deciding. *Proc Natl Acad Sci U S A* 93: 628–633, 1996.
- Shapiro N, Allison J. Conservation, choice, and the concurrent fixed-ratio schedule. *J Exp Anal Behav* 29: 211–223, 1978.
- Stephens DW, Krebs JR. *Foraging Theory*. Princeton, NJ: Princeton University Press, 1987, 1st edition.
- Sugrue LP, Corrado GS, Newsome WT. Matching behavior and the representation of value in the parietal cortex. *Science* 304: 1782–1787, 2004.
- Sul JH, Jo S, Lee D, Jung MW. Role of rodent secondary motor cortex in value-based action selection. *Nat Neurosci* 14: 1202–1208, 2011.
- Sul JH, Kim H, Huh N, Lee D, Jung MW. Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. *Neuron* 66: 449–460, 2010.
- Sutton RS, Barto AG. *Reinforcement Learning: An Introduction*. Cambridge, MA: The MIT Press, 1998.
- Watanabe K, Igaki S, Funahashi S. Contributions of prefrontal cue-, delay-, and response-period activity to the decision process of saccade direction in a free-choice ODR task. *Neural Netw* 19: 1203–1222, 2006.

Watanabe Y, Funahashi S. Neuronal activity throughout the primate mediodorsal nucleus of the thalamus during oculomotor delayed-responses. I. cue-, delay-, and response-period activity. *J Neurophysiol* 92: 1738–1755, 2004.

Watkins C, Dayan P. Q-learning. *Mach Learn* 8: 279–292, 1992.

Yoshida K, Saito N, Iriki A, Isoda M. Representation of others' action by neurons in monkey medial frontal cortex. *Curr Biol* 21: 249–253, 2011.

## Figure Captions

Fig. 1

(a) Schematic illustration of the two tasks. In the ICT, the monkey was required to make a memory-guided saccade toward the cued location. In the FCT, the monkey also needed to choose one of two cued locations before making a saccade. (b) Example of variables in the logistic regression analysis. The values of  $i_j$  (ICT),  $s_j$  (FCTs),  $d_j$  (FCTd chosen) and  $u_j$  (FCTd unchosen) in relation to direction A (right) and B (upper) that are available in the current trial are shown in a matrix. Circles and arrows indicate the location of the cue(s) and the monkey's response in each trial, respectively.

Fig. 2

(a) Example of saccadic responses in a session for monkey E. Saccade trajectories for all of the correct trials in a session are plotted in different colors and symbols (ICT: light crosses, FCT: dark circles). Eight black circles indicate the location and the size of the visual cues. Axes are scaled in visual angles. (b) Choice proportion for each pair condition in the FCT. For each box, the location of the median line (overall session average) and the small dots (each session) indicate the proportion of choosing the direction labeled at the top from the bottom (98 sessions each for both monkeys O and E).

Fig. 3

(a, c) The coefficients of logistic regression calculated from the real data. For both monkeys, the preceding ICT trials (light gray) had negative regression coefficients especially just before the current trial. The coefficients for preceding FCTd trials was slightly positive for the chosen direction (solid dark gray) and negative for the unchosen direction (dotted dark gray) compared to that of the ICT, but they were generally parallel among these three conditions. On the other hand, the coefficients for preceding FCTs trials (black) was highly elevated from those in the other conditions, and the difference drastically increased as a function of the number of preceding trials. Error bars indicate the standard error based on the six pair conditions. Filled symbols indicate the statistical significance based on the shuffled samples (upright and inverted triangles show a significant increase and decrease, respectively). (b, d) The confidence intervals for the coefficients based on the shuffled data (95% confidence intervals with correction for multiple comparisons).

Fig. 4

(a) The history effect of preceding decision and action calculated from the coefficients of a logistic regression analysis. Data for monkey O (solid) and monkey E (dotted) are plotted together. The history effect of preceding action is a direct reproduction of the regression coefficients for the ICT in Fig. 3. Making a response to a direction

decreased the probability of choosing that direction in subsequent trials (light gray). On the other hand, choosing a direction increased the probability of the same choice in subsequent trials when the same cue pair was presented (black), but this effect of past decision did not take place when the cue pair was different (dark gray). Error bars indicate the standard error based on the six pair conditions. Filled symbols indicate the statistical significance based on the shuffled samples (upright and inverted triangles show a significant increase and decrease, respectively). (b) The confidence intervals for the history effect of preceding decision and action based on the shuffled data (95% confidence intervals with correction for multiple comparisons).

## Tables

Table 1 Average parameter estimates in the reinforcement learning models.

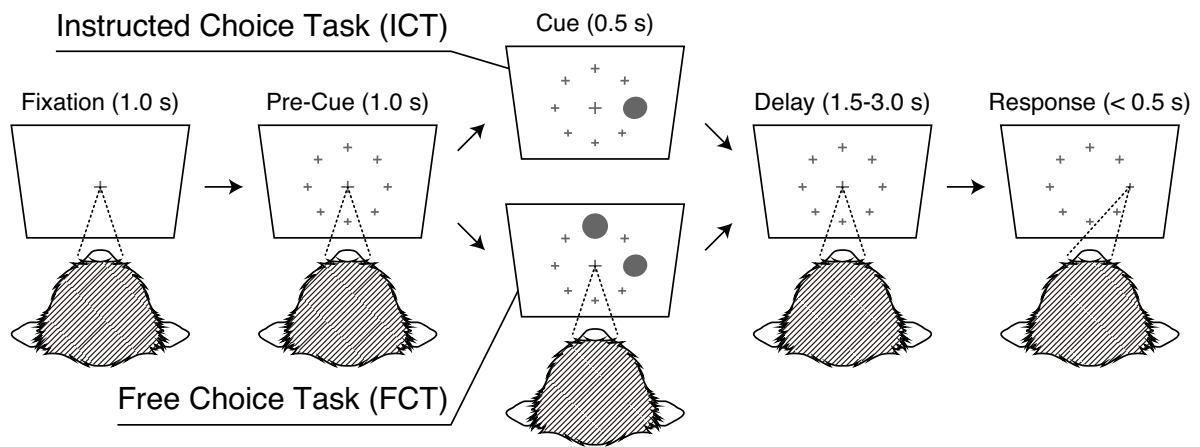
	$\alpha$	$\Delta_R$	$\Delta_C$	$\Delta_U$	AIC	Prediction (%)
Monkey O (98)						
R (2)	$0.96 \pm 0.17$	$-0.16 \pm 0.35$	—	—	$1.34 \pm 0.05$	$52.8 \pm 4.5$
R+C (12)	$0.98 \pm 0.08$	$-0.08 \pm 0.28$	$0.13 \pm 0.30$	—	$1.26 \pm 0.11$	$63.0 \pm 6.9$
R+C+U (84)	$0.90 \pm 0.22$	$-0.16 \pm 0.35$	$0.06 \pm 0.28$	$-0.21 \pm 0.29$	$1.23 \pm 0.11$	$62.6 \pm 4.2$
Monkey E (98)						
R (3)	$0.98 \pm 0.06$	$-0.09 \pm 0.27$	—	—	$1.35 \pm 0.06$	$52.0 \pm 6.6$
R+C (15)	$0.98 \pm 0.03$	$-0.06 \pm 0.18$	$0.16 \pm 0.23$	—	$1.26 \pm 0.11$	$64.0 \pm 7.7$
R+C+U (80)	$0.92 \pm 0.12$	$-0.13 \pm 0.28$	$0.11 \pm 0.25$	$-0.28 \pm 0.27$	$1.23 \pm 0.11$	$66.2 \pm 5.7$

AIC was standardized by dividing the AIC in each session by the number of FCT trials in the same session that were used in the maximum-likelihood estimation. Data are mean  $\pm$  standard deviation. Values in parentheses next to the names of the animals and the models are the number of total sessions for each animal and the number of sessions in which the model was selected to be optimal based on the AIC, respectively. The models with  $\Delta_C$  parameter (R+C model and R+C+U model) were frequently selected than the model without  $\Delta_C$  parameter. In both of the R+C and R+C+U models,  $\Delta_R$  was significantly smaller than zero and  $\Delta_C$  was significantly larger than zero.



Fig. 1

(a) Behavioral tasks

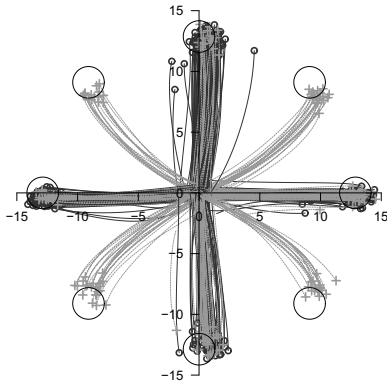


(b) Example of variables in the logistic regression analysis

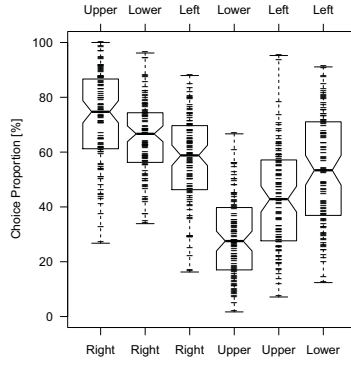
previous trials						current trial
-6	-5	-4	-3	-2	-1	$j$
0	0	0	0	0	1	$i_j$
0	0	-1	0	0	0	$s_j$
0	0	0	0	1	0	$d_j$
1	-1	0	0	0	0	$u_j$

Fig. 2

(a) Example of saccade



(b) Choice proportion  
Monkey O



Monkey E

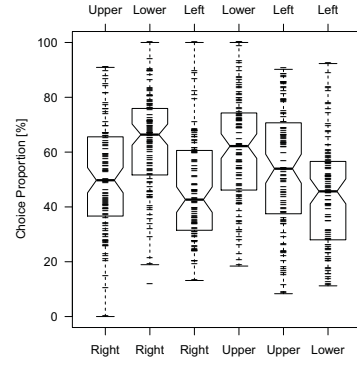
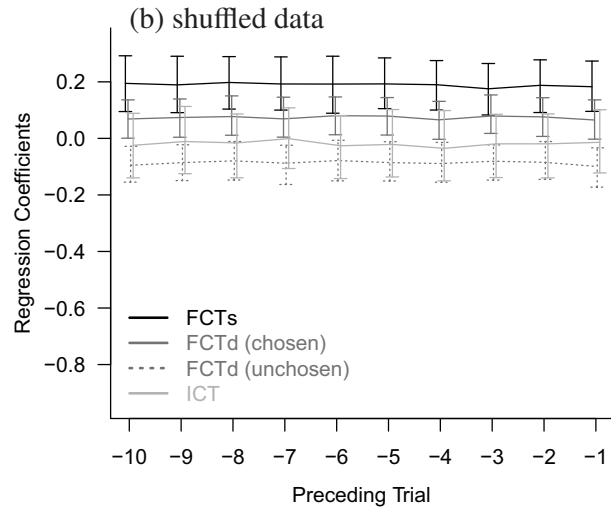
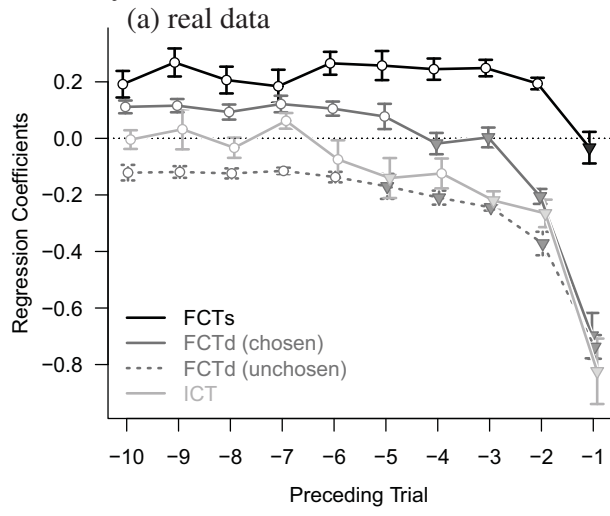


Fig. 3

### Monkey O



### Monkey E

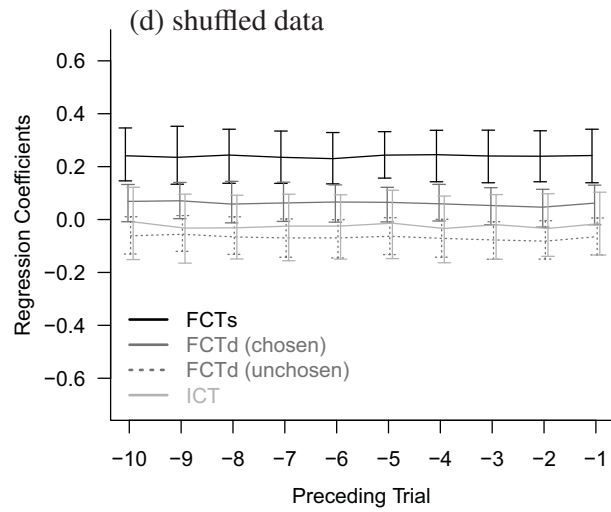
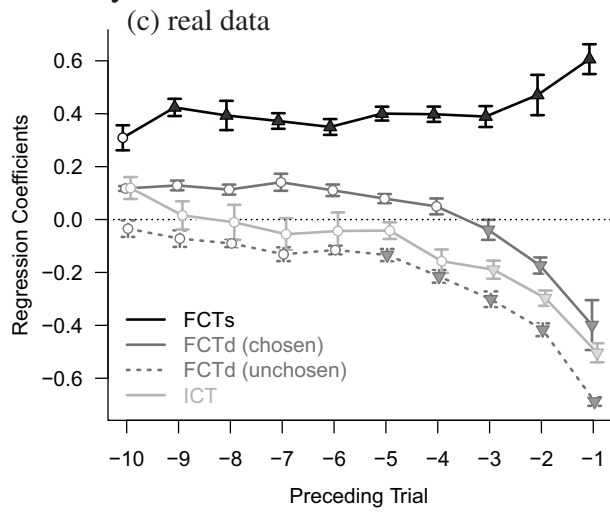


Fig. 4

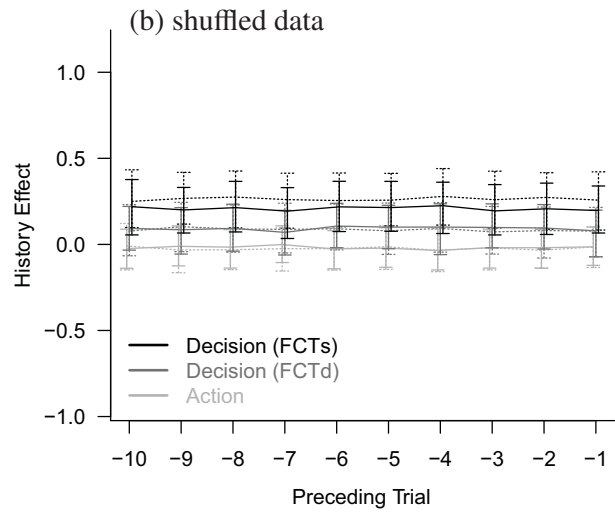
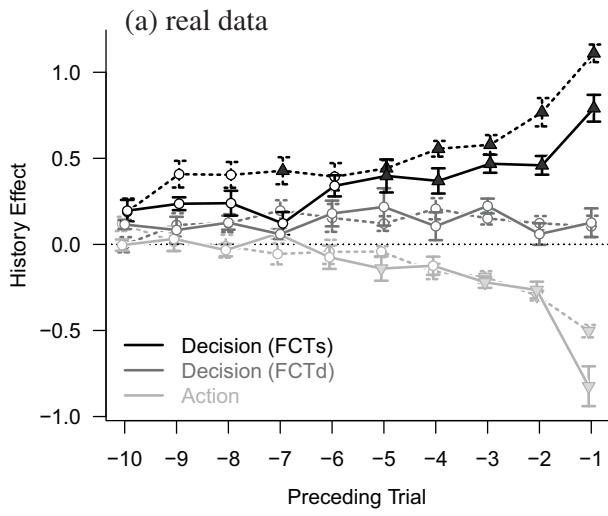


Table 1 Average parameter estimates in the reinforcement learning models.

	$\alpha$	$\Delta_R$	$\Delta_C$	$\Delta_U$	AIC	Prediction (%)
Monkey O (98)						
R (2)	$0.96 \pm 0.17$	$-0.16 \pm 0.35$	—	—	$1.34 \pm 0.05$	$52.8 \pm 4.5$
R+C (12)	$0.98 \pm 0.08$	$-0.08 \pm 0.28$	$0.13 \pm 0.30$	—	$1.26 \pm 0.11$	$63.0 \pm 6.9$
R+C+U (84)	$0.90 \pm 0.22$	$-0.16 \pm 0.35$	$0.06 \pm 0.28$	$-0.21 \pm 0.29$	$1.23 \pm 0.11$	$62.6 \pm 4.2$
Monkey E (98)						
R (3)	$0.98 \pm 0.06$	$-0.09 \pm 0.27$	—	—	$1.35 \pm 0.06$	$52.0 \pm 6.6$
R+C (15)	$0.98 \pm 0.03$	$-0.06 \pm 0.18$	$0.16 \pm 0.23$	—	$1.26 \pm 0.11$	$64.0 \pm 7.7$
R+C+U (80)	$0.92 \pm 0.12$	$-0.13 \pm 0.28$	$0.11 \pm 0.25$	$-0.28 \pm 0.27$	$1.23 \pm 0.11$	$66.2 \pm 5.7$

AIC was standardized by dividing the AIC in each session by the number of FCT trials in the same session that were used in the maximum-likelihood estimation. Data are mean  $\pm$  standard deviation. Values in parentheses next to the names of the animals and the models are the number of total sessions for each animal and the number of sessions in which the model was selected to be optimal based on the AIC, respectively. The models with  $\Delta_C$  parameter (R+C model and R+C+U model) were frequently selected than the model without  $\Delta_C$  parameter. In both of the R+C and R+C+U models,  $\Delta_R$  was significantly smaller than zero and  $\Delta_C$  was significantly larger than zero.