

---

# Benefit of Multiple Representations for Motor Sequence Control in The Basal Ganglia Loops

---

**Hiroyuki Nakahara**  
Lab. For Info. Synthesis  
Brain Science Institute, RIKEN  
2-1 Hirosawa Wako  
Saitama 351-0198, Japan  
hiro@brain.riken.go.jp

**Kenji Doya**  
Kawato Dynamic  
Brain Project, JST  
2-2 Hikaridai, Seika  
Soraku Kyoto Japan  
doya@erato.atr.co.jp

**Okihide Hikosaka**  
Dept. of Physiology  
Juntendo Univ.  
2-1-1 Hongo Bunkyo  
Tokyo Japan  
hikosaka@med.juntendo.ac.jp

## Abstract

Visual and motor coordinate systems have different computational advantages for motor sequence. This paper proposes that the multiple loops that link the basal ganglia (BG) and frontal cortical areas take their different advantages and specialize in acquisition and execution of visuomotor sequence. The central idea of the model is that a visuomotor sequence is easier to *learn* by visual coordinates but is easier to *control* by motor (joint-angle) coordinates [7]. Based on a hypothesis that the BG perform reinforcement learning using dopaminergic reinforcement signal, we constructed a network model that replicates both behavioral and physiological findings in recent experimental studies using "2x5 task" [3].

## 1 INTRODUCTION

We often feel as if our fingers move by themselves during playing piano with familiar scores, whereas we need pay attention to learn a new score. Thus, well-learned sequences seem to be acquired in form of body movements, whereas new ones don't. How the learning process progresses from new to well-learned sequences and whether they are acquired in the form of body movements or in a more abstract level have been subjects of much debate. Hikosaka and his colleagues recently demonstrated in a sequence learning task (2x5 task) in monkeys that there are indeed two components, short-term and long-term, in sequence learning [3] and found that the latter is specific to particular body usage [3, 9]. They further showed that different parts of the basal ganglia (BG) and the frontal cortex are involved differentially in the short-term and long-term component of sequence learning [5, 6].

Motivated by these findings, we propose a hypothesis that 1) humans (and monkeys) use

multiple representations of sequences, specifically those in visual and motor coordinates; 2) visual representation is dominant in early acquisition whereas motor representation is dominant in later execution; 3) among the BG loops, visual representation is used in the 'dorsolateral prefrontal (DLPF) loop' (called the visual loop in this paper) that links the DLPF area and the anterior striatum (St) in the BG while motor representation is used in the 'motor loop' that links the supplementary motor area (SMA) and the posterior St.

The BG have been considered to play a major role in sequential motor control using their multiple loop organization with the frontal cortex [1]. However, their specific role as well as how different loops interact with each other have been poorly understood. Based on our hypothesis and a theory that the BG perform temporal difference (TD) reinforcement learning using the reward signal carried by nigro-striatal dopamine neurons [4, 10], we construct a network model that replicates the behavioral/physiological findings of the 2x5 task experiment. The result supports our hypothesis of the use of different representations in the BG loops and also provides a new insight on the function of the presupplementary motor area (pre-SMA), which has been vaguely understood. Preliminary result was in [8].

## **2 VISUAL AND MOTOR REPRESENTATIONS FOR VISUOMOTOR SEQUENCES**

It should be noted that, in visuomotor sequence tasks like a serial reaching task, a correct sequence of movement can be encoded in at least two different ways. One is to encode the sequence in the visual space as if a target marker is moving around in the extrinsic, Cartesian coordinate and the visual target position is then transformed into a corresponding body posture on each step of the movement. Another is to encode the sequence in the somatic space as the sequence of target body postures using the intrinsic, joint-angle coordinate. Each representation has each strength and weakness for sequential motor control.

One of the most critical issues in sequential motor learning is the credit assignment problem whose difficulty relies on the representation used in the learning system. Visual representation is more advantageous in terms of credit assignment. The main reason is that the positions of candidate targets are given explicitly in the visual input space, whereas the candidates of target postures in the somatic space are not specified until the articulator (e.g. hand) reaches to one of them. Hence, search for a correct target in the visual space is just to select one among a few candidates whereas search in the somatic space involves exploration in a continuous, possibly high-dimensional state space. The illposedness of the visuomotor transformation further makes it difficult to consistently represent targets in the somatic space. Thus, it is much easier to learn a correct sequence in visual coordinates.

Motor (somatic) representation is more advantageous for control of movement in real time. Once a sequence is encoded by a sequence of body postures, step-by-step coordinate visual-to-motor transformation is not necessary. Optimization of the sequence of body postures further enables fast and robust reach to the targets. Hence, it is more robust, more accurate and quicker to control a sequential movement by use of optimized motor representations.

## **3 2x5 TASK AND ITS FINDING**

Figure 1 shows an example of the sequence of events in a single trial of the 2x5 task (refer to [3] for detail). When the animal pressed the home key at the start of a trial, two of the 16 LEDs (a 'set' of stimulus) were turned on. The animal had to press the illuminated LEDs in a predetermined order, which should be found by trial-and-error. If successful, another pair of LEDs (the second set) was illuminated which the monkey had to press again in a predetermined order. A fixed sequence of 5 sets (a 'hyperset') was presented in a trial. After each successful set, the animal was given a reward. The completion of a hyperset gave the

animal the maximum reward. When the animal pressed a wrong button, the animal had to start from the 1st set. The same hyperset was used through a 'block' until completing a certain number of successful trials (criterion). On each day in training period, the monkey performed several blocks with different hypersets. Some hypersets were used everyday ('learned' hypersets), while others were randomly generated and used only once ('new' hypersets). Order of presentation of learned and new hypersets are randomized everyday.

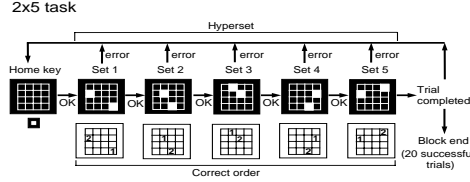


Figure 1: Procedure of 2x5 task with an example of a hyperset .

The implication of some findings in the 2x5 task are only briefly mentioned here.

- **Different learning levels** were observed in training period, namely *short-term* and *long-term* levels (improved performance within a block and across days respectively) [3]. Note that the memory of learned hypersets was sequence-dependent and was incompletely transferred between different hands (the reverse hyperset and opposite hand experiments) [9].
- **Functional differentiation over different areas** was observed in blockade experiment by muscimol injection. Comparing the blockade of the anterior St and posterior St (putamen), the anterior St is more critical for new hypersets, whereas the posterior St is more critical for learned ones [5]. Comparing the SMA and pre-SMA blockades, the pre-SMA is more critical for new hypersets, whereas the SMA is more critical for learned ones [6].

#### 4 HYPOTHESIS OF MULTIPLE REPRESENTATIONS IN THE BG LOOPS AND MODEL FOR THE 2X5 TASK SIMULATION

Considering the experimental evidence of the BG loops and the 2x5 task findings, we propose that the visual loop learns sequences in visual representation, while the motor loop learns in motor representation. Because of different nature of these representations as discussed, we hypothesize that a visuomotor sequence is first learned in the visual loop and then learned in the motor loop, even though both loops may learn the sequence concurrently. The pre-SMA is a pivotal area to coordinate the visual and motor loops, since it connects the DLPF and the SMA and projects to both anterior/posterior St.

The simulation study of the 2x5 task was employed to test our hypothesis. Figure 2 shows an overall structure of the network model (top) and corresponding brain areas (bottom). The network model consisted of the visual network (the visual loop), the motor network (the motor loop), the critic network (dopamine system), and inverse kinematics module.

The visual network is a function of the immediate visual input  $\mathbf{v}^I$  and the visual context  $\mathbf{v}^C$  to predict the spatial position of the LED button to be pressed (the visual target prediction  $\mathbf{v}^P$ ). The motor network is a function of the visually-dependent motor input  $\mathbf{m}^I$  and the motor context  $\mathbf{m}^C$  to predict the arm posture necessary for pressing the correct button (the motor target prediction  $\mathbf{m}^P$ ). The critic network learns to predict an expected reward given the visual input  $\mathbf{v}^I$  and also helps learning of the visual and motor networks. The final motor command  $\mathbf{m}^O$  is selected stochastically based on the visual prediction  $\mathbf{v}^P$  and the motor prediction  $\mathbf{m}^P$ . All the literature cannot be mentioned that supports the anatomical correspondence for the computational component (Fig. 2) but we only mention that the visual and the motor predictions are made at the level of the basal ganglia output which

selectively disinhibits reciprocally connected thalamo-cortical units in the DLPF and SMA.

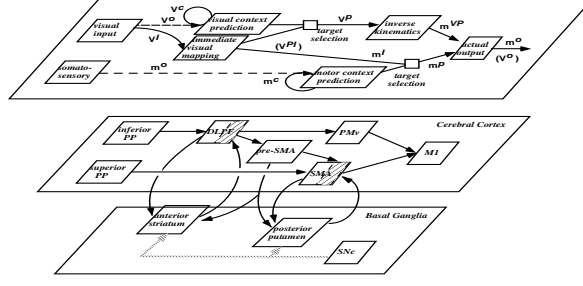


Figure 2: The proposed model of functions of the basal ganglia loops: the computational components (top) and the corresponding parts of the BG loops (bottom).

#### 4.1 Visual and motor representations

The state of the visual network was encoded by a 16 dimensional vector, each component of which stood for a spatial position of the LEDs on the panel. The immediate visual input  $\mathbf{v}^I$  was a binary vector encoding the illuminated LEDs as 1 and the not-illuminated as 0.

We modeled an arm by two joint angles in 2D dimension,  $(\theta^1, \theta^2)$ , which denote shoulder and elbow angles respectively, for simplicity. Then, the state of the motor network was encoded by a population vector of joint angles in motor coordinates, for which we used a vector of normalized Gaussian activation functions (allocated by  $8 \times 8$  grid in this study,

hence,  $j = 1, \dots, 64$ ):  $\mathbf{m}_j^I = \exp[-\frac{1}{2} \sum_{k=1}^2 \left( \frac{\theta^k - \hat{\theta}_j^k}{\delta^k} \right)^2] / \sum_i \exp[-\frac{1}{2} \sum_{k=1}^2 \left( \frac{\theta^k - \hat{\theta}_i^k}{\delta^k} \right)^2]$ , where  $\hat{\theta}_j^k$  denotes the preferred joint angle for the  $j$ -th unit.

This study assumed the inverse kinematic transformation to be learned prior to the experiment so that an analytic calculation was used to transform the visual vector  $\mathbf{v}$  into a corresponding motor vector  $\mathbf{m}$ . Hence, provided the joint angles  $(\theta_j^1, \theta_j^2)$  corresponding to each of the LED buttons  $v^i = (x_i, y_i)$  was calculated, the motor vector  $\mathbf{m}^i$  corresponding to each  $i$ -th button position was then obtained. For a given visual vector  $\mathbf{v} = \{v^i\}_{i=1}^{16}$ , the corresponding motor vector was given by  $\mathbf{m} = \sum_{i=1}^{16} v^i \mathbf{m}^i$ , which is denoted by  $\mathbf{m} = K^{-1}(\mathbf{v})$ .

#### 4.2 Prediction of target output

Both visual and motor networks had the same structure. The output vectors  $\mathbf{v}^P$  and  $\mathbf{m}^P$  represent the probability of the target position in visual and motor coordinates, respectively. The output of the visual and motor networks were given in the same manner by

$$\mathbf{v}^P(t) = S(\mathbf{W}^{VI} \mathbf{v}^I(t) + \mathbf{W}^{VC} \mathbf{v}^C(t)),$$

$$\mathbf{m}^P(t) = S(\mathbf{W}^{MI} \mathbf{m}^I(t) + \mathbf{W}^{MC} \mathbf{m}^C(t)),$$

where  $\mathbf{W}^{VI}$  and  $\mathbf{W}^{VC}$  are the weights for the immediate and contextual visual inputs and  $\mathbf{W}^{MI}$  and  $\mathbf{W}^{MC}$  are the weights for the visually-dependent motor and contextual motor inputs.  $S$  denotes a so-called softmax function,  $S(\mathbf{u})_i = e^{\zeta u_i} / \sum_k e^{\zeta u_k}$ .

The visual and motor context vectors were updated in the same manner, of which the former  $\mathbf{v}^C$  is only shown here by  $\mathbf{v}^C(t+1) = \mathbf{v}^C(t) + (1/\tau_V)(\mathbf{v}^O(t) - \mathbf{v}^C(t))$  where  $\tau_V$  was the time constant. The initial context vector  $\mathbf{v}^C(0)$  was set as  $\mathbf{v}^I(0)$  for simplicity.

Two important assumptions were made: (1) that the immediate visual mapping ( $\mathbf{W}^{VI}$ ) was reset in the beginning of each block. This is computationally advantageous to prevent interference of learning between different hypersets. Note that such a property as temporally storing information and canceling it rapidly is a central feature of working memory in the DLPF [2]; (2) considering the pre-SMA neural responses, that the visually-dependent motor input was given by the immediate visual prediction  $\mathbf{v}^{PI}$  as  $\mathbf{m}^I(t) = K^{-1}(\mathbf{v}^{PI}(t))$ , where  $\mathbf{v}^{PI} = S(\mathbf{W}^{VI}\mathbf{v}^I(t))$ . In addition,  $\mathbf{W}^{VC}$  was kept as identity for simplicity.

### 4.3 Selection of final motor command

To select the final motor command  $\mathbf{m}^O$ , the visual prediction  $\mathbf{v}^P$  was transformed by  $\mathbf{m}^{VP} = K^{-1}(\mathbf{v}^P)$  as in motor coordinates. It was then combined with the motor prediction  $\mathbf{m}^P$  using a component-wise product

$$p_j(t) = \mathbf{m}_j^{VP}(t)\mathbf{m}_j^P(t) / \sum_k \mathbf{m}_k^{VP}(t)\mathbf{m}_k^P(t) \quad (j = 1, \dots, 64)$$

This is sensitive to both values of visual and motor networks componentwise. Finally one of the outputs was stochastically set as  $\mathbf{m}_j^O(t) = 1$  with  $\text{Prob}(\mathbf{m}_j^O(t) = 1) = p_j(t)$ .

### 4.4 Reinforcement learning algorithm in the critic network

We used TD learning in the actor-critic scheme [4] based on aforementioned hypothesis of the BG. The critic was defined by  $P(t) = \mathbf{w}^R \cdot \mathbf{v}^I(t) + b^R$ , where  $\mathbf{w}^R$  and  $b^R$  are the weight vector and bias. The TD error  $\hat{r}(t)$  was then defined by  $\hat{r}(t) = r(t) + \gamma P(t+1) - P(t)$ , where  $r(t)$  is a given reward and  $\gamma$  is a discount factor. The critic weight matrix  $\mathbf{w}^R$  and bias  $b^R$  were updated using TD error by  $\Delta \mathbf{w}^R(t) \propto \hat{r}(t)\mathbf{v}^I(t)$  and  $\Delta b^R(t) \propto \hat{r}(t)$ . The weight matrices  $\mathbf{W}^{VI}$ ,  $\mathbf{W}^{VC}$ , and  $\mathbf{W}^{MC}$  of the visual and motor networks were updated in the same manner, of which the update of  $\mathbf{W}^{VC}$  is:

$$\Delta \mathbf{W}_i^{VC}(t) \propto \hat{r}(t)\mathbf{v}_i^O(t)\mathcal{G}(\mathbf{v}_i^O(t), \mathbf{v}_i^P(t))\mathbf{v}_i^C(t)$$

where  $\mathcal{G}(p, q)$  is the gain term to realize the stochastic gradient descent for a squared error  $(p - q)^2$  in terms of the softmax function and is defined by  $\mathcal{G}(p, q) = (p - q)q(1 - q)$ .

In the next section, we report the results of the training, the reverse hyperset, the opposite hand, and blockade simulations. Each simulation method is then mentioned.

## 5 Results

**Different learning levels in training period:** we let the model learn 2 learned and 1 new hypersets per a simulated day for 10 days, approximately keeping their ratio in experiment. The criterion for the successful trials in a block was set as 10. Both short-term and long-term learning levels were also observed in our simulation. The model makes more errors during an early stage *within a block*, particularly for the 1st and 3rd days (short-term learning; Fig. 3, left), while the model improved its performance for the same hypersets *across days* (long-term learning; Fig. 3, left & middle). For comparison, the performance across days is shown for the case only the motor network is used, i.e.  $p_j(t) = \mathbf{m}_j^P(t)$  (Fig. 3, right). It is clear that learning in the motor network was very slow in simulation, which should be expected from our hypothesis. When only the visual network is used, it was observed that, while the performance for learned hypersets in a first few days was almost as good as that of the normal simulation, its performance in a last few days was slightly but significantly worse than that of the normal simulation, which was also expected.

**Reverse hyperset simulation:** the performance for reversed hypersets, where all sets were the same as the learned ones but the sequence of the sets were reversed, was tested. The

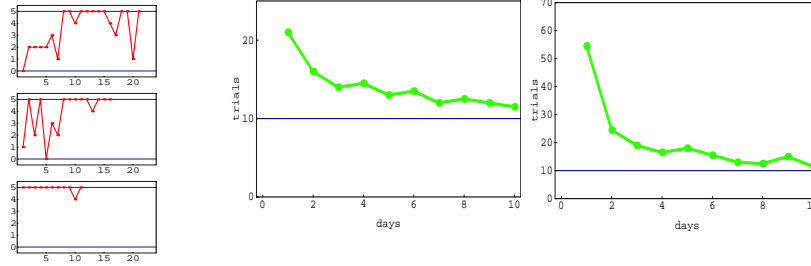


Figure 3: The number of completed sets (ordinate) across trials (abscissa) is shown for the same hyperset for the 1st, 3rd, 10th days from top to bottom (left). The mean number of trials for learned hypersets across days in training simulation (middle) and in training simulation in case only the motor network is used (right).

performance for reversed ones was significantly worse than that for learned ones, but not different from that for new ones (Fig. 4, left), similar to the experiment result [9].

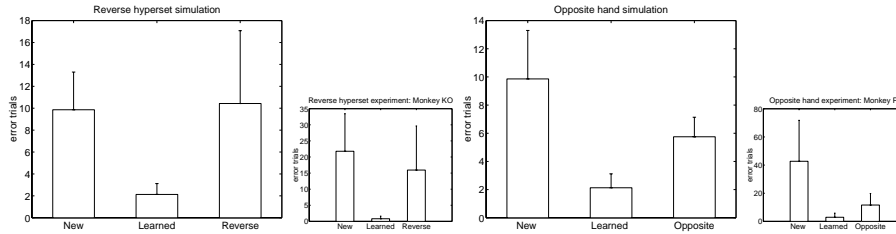


Figure 4: (left) reversed hyperset simulation: the comparison of the number of error trials for new, learned, and reversed hypersets. (right) opposite hand simulation: new hypersets (left) and learned hypersets using the trained hand (center) and the opposite hand (right). Error bars stand for SD. Experimental results [9] are also shown next to each simulation.

**Opposite hand simulation:** the opposite hand simulation for learned hypersets was simulated by initializing the motor context prediction. The performance for learned hypersets by the opposite hand was significantly worse than that for learned ones by the trained hand, but significantly better than that for new ones (Fig. 4, right), similar to the experiment [9].

**Blockade of visual and motor networks:** three blockade conditions was tested; (1) the visual network blockade, in which we substituted the visual input  $\mathbf{v}^I$  for the immediate visual prediction  $\mathbf{v}^{PI}$  and the visual target prediction  $\mathbf{v}^P$ ; (2) the pre-SMA blockade, in which the visual input  $\mathbf{v}^I$  is substituted for the immediate visual prediction  $\mathbf{v}^{PI}$ ; (3) the motor network blockade, in which we let  $\mathbf{m}^P(t) = S(\mathbf{m}^I(t))$  where  $\mathbf{m}^I(t) = K^{-1}(\mathbf{v}^I(t))$ .

See Figure 5. By the pre-SMA blockade, the performance for new hypersets was significantly deteriorated but not for learned ones. By the visual network blockade, the performance for new ones was severely deteriorated, whereas the performance for learned ones was mildly but significantly deteriorated. In contrast, by the motor network blockade, the performance for learned ones was significantly deteriorated, while the performance for new ones was also mildly deteriorated. Hence, it is clear that the motor network is more critical for learned hypersets, whereas the visual network is more critical for new ones. It is also clear that the function of the pre-SMA is critical for new hypersets.

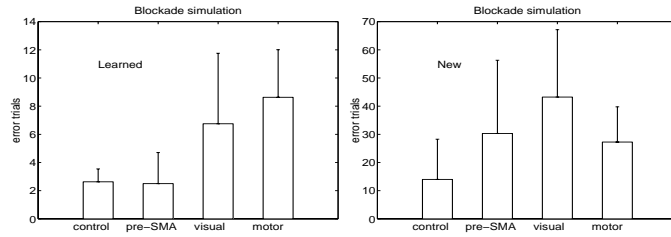


Figure 5: The comparison of the number of error trials for learned and new hypersets (left and right), in each of which the control, the pre-SMA blockade, the visual network blockade, the motor network blockade conditions are shown from left to right.

## 6 DISCUSSION

We hypothesized that the BG contribute to sequential motor control based on RL with multiple representations. The visual loop contributes particularly to early acquisition of sequences in visual coordinates while the motor loop contributes to late execution, once acquired. Simulation results resembled those of the 2x5 experiment: different learning levels (training simulation), sequential-dependency of well-learned sequences (reverse hyperset simulation), incomplete transfer of the motor memory between different hands (opposite hand simulation), and functional differentiations of different loops (blockade simulation).

**Acknowledgments:** supported by Uehara Memorial Foundation and JSPS Research for the Future Program to OH. HN acknowledges insights from colleagues in Hikosaka lab.

## References

- [1] G. E. Alexander, M. D. Crutcher, and M. R. DeLong. Basal ganglia-thalamocortical circuits: Parallel substrates for motor, oculomotor, “prefrontal” and “limbic” functions. In H. B. M. Uylings et al., editor, *Progress in Brain Res.*, volume 85, pages 119–146. Elsevier, 1990.
- [2] P. S. Goldman-Rakic. Circuitry of primate prefrontal cortex and regulation of behavior by representational memory. In F. Plum and V. Mountcastle, editors, *Handbook of Physiology - The Nervous System V*, volume 5, chapter 9, pages 373–417. 1987.
- [3] O. Hikosaka, M. Kato, S. Miyachi, and K. Miyashita. Learning of sequential movements in the monkey: Process of learning and retention of memory. *J.of Neurophys.*, 74(4):1652–1661, 1995.
- [4] J. C. Houk, J. L. Adams, and A. G. Barto. A model of how the basal ganglia generate and use neural signals that predict reinforcement. In *Models of Information Processing in the Basal Ganglia*, pages 249–270. MIT Press, Cambridge, Massachusetts, 1995.
- [5] S. Miyachi, O. Hikosaka, K. Miyashita, Z. Kari, and M. Rand. Differential roles of monkey striatum in learning of sequential hand movement. *Exp. Brain Res.*, 115:1–5, 1997.
- [6] K. Miyashita and O. Hikosaka. Activity of pre-SMA neuron for the performance of sequential movement in monkeys. *Proc. of 27th Annual Meeting, Soc. for Neuroscience.*, 23:778, 1997.
- [7] H. Nakahara. *Sequential Decision Making in Biological Sys.* PhD thesis, Tokyo Univ., 1997.
- [8] H. Nakahara, K. Doya, O. Hikosaka, and S. Nagano. Multiple representations in the basal ganglia loops for acquisition and execution of sequential motor control. *Proc. of 27th Annual Meeting, Soc. for Neuroscience.*, 23:778, 1997.
- [9] M. Rand, O. Hikosaka, S. Miyachi, X. Lu, and K. Miyashita. Characteristics of a long-term procedural skill in the monkey. *Exp. Brain Res.*, (118):293–297, 1998.
- [10] W. Schultz, P. Dayan, and R. Montague. A neural substrate of prediction and reward. *Science*, 275:1593–1599, 1997.