Systems and Computers in Japan, Vol. 33, No. 11, 2002
Translated from Denshi Joho Tsushin Gakkai Ronbunshi, Vol. J83-D-II, No. 2, February 2000, pp. 814–826

Composition and Decomposition Learning of Reaching Movements Under Altered Environments: An Examination of the Multiplicity of Internal Models

Eri Nakano,[1*] John R. Flanagan,[2] Hiroshi Imamizu,[3] Rieko Osu,[3] Toshinori Yoshioka,[3†] and Mitsuo Kawato[3,4†]

[1]Advanced Telecommunications Research Institute International, Kyoto, 619-0299 Japan

[2]Department of Psychology, Queen's University, Kingston, Canada

[3]ERATO, Kawato Dynamic Brain Project, Kyoto, 619-0299 Japan

[4]ATR Human Information Processing Research Laboratories, Kyoto, 619-0299 Japan

SUMMARY

We have studied the learning processes of reaching movements under novel environments whose kinematic and dynamic properties are altered. In the experiments, we have used, as the kinematic transformation, a rotational transformation which is displayed by rotating a cursor indicating hand position in the orthogonal coordinate system on a CRT; a viscous transformation using viscous field as the dynamic transformation; and a combined transformation of these two transformations. It is observed that the hand trajectory approaches a straight line along with learning and accurately reaches the target. When the combined transformation is learned after the rotational transformation and viscous transformation are learned first, respectively, the final error becomes smaller and the path length also becomes shorter than the case when the combined transformation is learned first. Moreover, the final error and path length of the movement under rotational transformation and viscous transformation when the combined transformation is learned first also become smaller than the case when the rotational and viscous transformations are learned first. These results suggest that the central nervous system has learned separately the multiple internal models which compensate the respective transformations, and has composed or decomposed the respective internal models in accordance with the environmental changes. It may be considered that such multiplicity of internal models makes it possible for the living body to flexibly cope with the environments or tools having various dynamic and kinematic properties. © 2002 Wiley Periodicals, Inc. Syst Comp Jpn, 33(11): 80–94, 2002; Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/scj.1166

Key words: reaching movement; internal model; modular learning; kinematic transformation; dynamic transformation.

[*]Now affiliated with Digital Human Laboratory, National Institute of Advanced Industrial Science and Technology.
[†]Now affiliated with Human Information Science Laboratories.

Contract grant sponsors: A portion of this research was performed as part of the "Research on communication of primates including humans," based on the 1999-FY Adjustment Fund for Science and Technology Promotion of the Science and Technology Agency, and moreover, it has received the assistance of HFSP.

© 2002 Wiley Periodicals, Inc.

1. Introduction

Humans acquire adaptive behaviors for various changes of tools and environments, and can quickly cope with the changeover of the environments and tools after acquisition. In such adaptive process, the central nervous system (CNS) has learned to make smaller the error between the movement targeted and the movement realized. At this time, to accurately achieve the given target in external space, it is thought that the CNS has learned to generate appropriate series of movement commands in the physical space by using the internal models for the properties of the controlled system and environment [1–3].

The movement commands depend not only on the kinematic quantities [4] such as the length and position of the physical structure system and the posture in the polar coordinate system of the shoulder or the dynamic quantities [5] related to forces such as mass or moment of inertia and viscosity/elasticity, but also on the kinematics of tools and environments, for example, the tool shapes or the relationship between positions of limbs in the visual coordinate system and working coordinate system or the dynamics, for example, the mass or moment of inertia of tools and the external force in the water or environment without gravity. Therefore, when these properties change, the movement commands required for carrying out the objective movement will also change. To achieve the objective movement, the CNS must be capable of generating the series of movement commands, respectively, in accordance with the various kinematic properties or dynamic properties which the tools and environments possess.

If the internal model only generates an appropriate movement command, it must be learned anew every time the property of the tool and environment is altered. Therefore, even if there is an experience that the altered property has been learned before, the initial error or the time used in learning always becomes close whenever it corresponds to a novel property. However, if the internal models have learned separately different properties by multiple modules, it is not necessary to learn the already-learned property again. This kind of learning method is called the modular scheme. The CNS has the following advantages in adapting this scheme. The initial error becomes smaller and the time used in readaptation also becomes shorter by using the low-level module which has learned the property in accordance with the alternation of the property. Moreover, by various combinations of multiple low-level modules, considerable numbers of kinds of tools and environments can be dealt with. Recently, the proposed multiple internal models consider the mechanism of control and learning based on such modular scheme [6–8].

In the conventional researches, the kinematic properties [9–14] or dynamic properties [15–18] of the environments are altered experimentally, and the adaptive processes to the altered environments have been investigated through movement learning tasks. It is thought that the adaptive mechanism reported in these researches suggests that in addition to the internal models of the arms which have already been acquired in the developmental process, the internal models of the altered environments have been prepared separately. Moreover, it has been reported that the complicated visual movement task is decomposed into simpler elements in the brain and they are respectively learned by multiple modules [14]. Furthermore, Imamizu and colleagues have shown that the learnings of the kinematic properties which were altered by different rules are coped by different cerebellum parts. However, there have been no studies on whether or not the internal models, respectively corresponding to the kinematic properties and dynamic properties, are prepared separately. If they were prepared separately, would the internal models for the respective properties be combined, or would a separate internal model learn the alternations of both properties together?

We have regarded the environment in which the kinematics only is altered and the environment in which the dynamics only is altered as the environments having different internal models. We have set a rotational transformation as the alternation of the former and a viscous transformation using a viscous field as the alternation of the latter. These two transformations will be called the "single transformation." With respect to this, the transformation using both rotational transformation and viscous transformation simultaneously will be called the "combined transformation." The details of the rules of these transformations will be described later. If the CNS has learned by the multiple internal models the movements under the transformations, the following can be predicted. If the rotational transformation and viscous transformation are learned by separate internal models as the elements of the combined transformation, the learning of the combined transformation may probably be promoted by learning the rotational transformation and viscous transformation beforehand. Conversely, after the combined transformation is learned, the tasks under rotational transformation and viscous transformation may probably be executed relatively easily.

To verify these predictions, the following two experiments are performed on separate days using the same subject. In one experiment, the combined transformation task is performed after the single transformation is experienced (Fig. 1A). In this case, the CNS may have generated the movement commands for executing the combined transformation task by composing the internal models corresponding to the respective single transformations. Therefore, this experiment is referred to as the composition experiment. In another experiment, the single transformation tasks are performed after the combined transformation

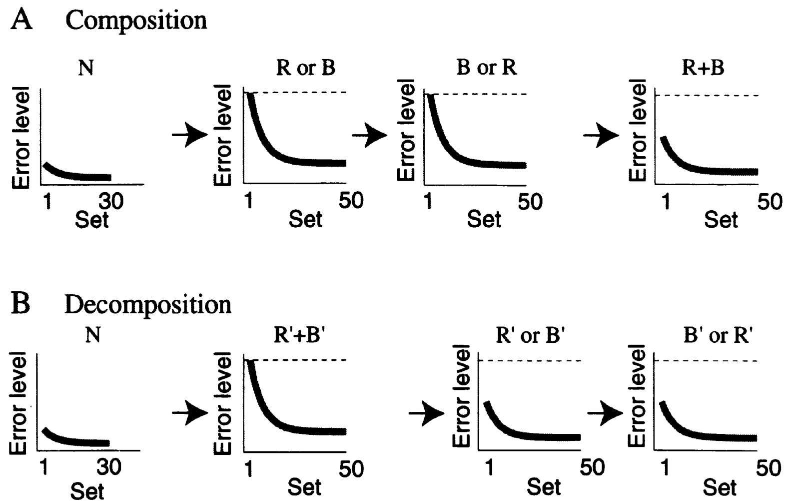


Fig. 1. Order of transformation tasks in the (A) composition and (B) decomposition experiments. N, R and R′, B and B′, and R+B and R′+B′ denote normal, rotational, viscous, and combined transformations. The thick lines denote error levels in reaching movements as a function of the number of trial sets. Ten trials were included in a set. The dashed lines indicate error levels in the early stage of learning.

is experienced (Fig. 1B). In this case, the internal models corresponding to the rotational transformation and viscous transformation prepared simultaneously in the combined transformation may probably be divided and used separately in the single transformations. Therefore, this experiment is called the decomposition experiment. In order to look into the effect of the preceding learning on the later learning, it is necessary that the internal model about the transformation which experiences first in each experiment is not acquired. In these experiments, between the transformations whose positive and negative signs of rotational angles or whose positive and negative signs of viscous terms are mutually different, it is assumed that the learning of transformations will not transfer. Accordingly, the transformations whose signs are mutually different are adopted between the composition experiment and the decomposition experiment. In this paper, the rotational transformation, viscous transformation, and combined transformation are denoted by R, B, and R+B, respectively. R′, B′, R′+B′ show the transformation having symbols different from R, B, R+B (Fig. 1). Based on the earlier assumption, the learnings of R and R′, B and B′, as well as R+B and R′+B′ are regarded as mutually equivalent in the degree of difficulty and the time constant of learning curve.

We have performed the following comparisons using the data obtained in the composition and decomposition experiments. The data measured under combined transformation are compared before (R′+B′) and after (R+B) learning the single transformation; moreover, the data under the single transformation are compared before (R or B) and after (R′ or B′) learning the combined transformation (see Fig. 1). At the time of learning the novel transformation, the error is large initially but the learning curve is probably observed, in which the errors decrease exponentially as the trials advance (for example, the second and third ones from the left in Fig. 1A). After the transformation has already been learned, the error of the reaching movement becomes smaller than that before learning, and the rate of decrease of the errors accompanying the trials may also be small (for example, the right end of Fig. 1A).

In the design of the experiments, the transformation task, in which it is predicted that the error of reaching movement will became smaller, is always performed last in the experiment. Therefore, even if a result which agrees with the prediction has been obtained, there is a possibility that not only the promotional effect of learning by the same kind of transformation is learned beforehand but also the simple effect of order which is not related to the type of transformation are included. Accordingly, we have investigated whether or not the difference in the errors of reaching movements depends only on the order of tasks. If the error in the case of the order which is always behind is significantly small, it is possible that the effect of order has influenced the learning effect.

Moreover, in order to investigate the effect of experimental days, namely, whether or not the transformation learning performed on the first day will transfer to the learning of the second day, the experimenter himself learns preliminarily the single transformation on the first day and then the single transformation task is performed again by changing the sign of the transformation one week later. If the error of the reaching movement of the second day is always significantly smaller than that of the first day, there is a possibility that the experience of the first-day experiment has promoted the learning of the transformation task of the second day. This phenomenon is called positive transfer. Conversely, if the error of the second day is larger than the first day, there is a possibility that the experience of the first-day experiment has hindered the learning of the task of the second day. This phenomenon is called negative transfer.

2. Experiments

2.1. Method

2.1.1. Experimental equipment and procedures

Eight subjects (21 to 35 years old; six males and two females) sit on chairs; with the right arm supported by a strap suspended from the ceiling, the right hand grasps the tip of a two-link manipulandum (PFM : Parallel-link air-magnet Floating direct-drive Manipulandum) and the visual reaching movement task is executed on a horizontal plane.* The wrist is fixed by a brace. The position of the hand is measured by the PFM, and its present position (a cursor of diameter 0.4 cm) and the target (a circle of diameter 1 cm) are displayed on a CRT screen installed in front of the subject. The scales of the CRT coordinate system and the working coordinate system of the hand are the same. The distance between CRT and subject is about 1.6 m. The sampling frequency is 500 Hz. Since the arms of the subject are hidden by a shielding plate, they perform the movements by looking at the screen only.

To get used to the experimental environment and task, all subjects have participated in training tasks to make the cursor reach the target in the space which has no transformations beforehand. This task is called the no-transformation task (normal). Before the experiment starts, the types of transformation are briefly explained to the subjects, and it is determined in such a way that the cursor expressing the hand will be put into the circle which is the target as accurately as possible with a short distance, within the duration (600 ms) of target presentation. The task is performed within a circle of radius 14 cm shown on the screen. For the purpose of making the movement distance constant, the target coordinates are selected randomly from the point on a circle of radius 10 cm with the final position of the previous trial as center. The times from the presentation of the previous target to the presentation of the next target are 4000 ms for subjects RB, AO, NY, HM and 1400 ms for subjects CY, SN, TT, KH. Since the target is presented randomly with a short-time interval, it is difficult during the execution of the task under transformation for the subject to perform the movement by recognized transformation without learning the transformation.

Ten trials are regarded as one set for the task, and 30 or 50 sets are performed for one session. At the first trial of each set, the cursor is always returned to the same position (center of the circle of radius 14 cm). The transformation is always imposed during the session of the transformation tasks. A short recess is inserted between sessions.

In the composition and decomposition experiments (Fig. 1), 10 sets of no-transformation tasks are inserted after the end of the session of the respective transformation tasks. All subjects perform these two kinds of experiments with an opening of an interval of over 1 week. Following the procedures of performing the experiments, the eight subjects are divided into two groups: C-D or D-C. Subjects RB, AO, SN, CY belonging to the C-D group perform the composition experiment on the first day and the decomposition experiment on the second day. Subjects NY, HM, TT, KH belonging to the D-C group perform the experiments in reversed order. The experimental conditions for the respective subjects are summarized in Table 1.

2.1.2. Rules of transformations

In the rotational transformation, the position (p, q) of the hand in the orthogonal coordinate system of the subject is rotated on the CRT screen and displayed on screen as a

Table 1. Experimental condition of each subject

Subject	Order of transformations: Day 1	Day 2
C-D group	Composition	Decomposition
RB	[N] [R^+] [B^-] [R^++B^-]	[N] [R^-+B^+] [R^-] [B^+]
AO	[N] [B^+] [R^-] [R^-+B^+]	[N] [R^++B^-] [B^-] [R^+]
CY	[N] [R^-] [B^-] [R^-+B^-]	[N] [R^++B^+] [R^+] [B^+]
SN	[N] [B^+] [R^+] [R^++B^+]	[N] [R^-+B^-] [B^-] [R^-]
D-C group	Decomposition	Composition
NY	[N] [R^++B^-] [B^-] [R^+]	[N] [B^+] [R^-] [R^-+B^+]
HM	[N] [R^-+B^+] [R^-] [B^+]	[N] [R^+] [B^-] [R^++B^-]
TT	[N] [R^++B^-] [B^-] [R^+]	[N] [B^+] [R^-] [R^-+B^+]
KH	[N] [R^-+B^-] [R^-] [B^-]	[N] [R^+] [B^+] [R^++B^+]

Transformation tasks were performed in the left to right order. [N], [R], [B], and [R+B] denote normal, rotational, viscous, and combined transformations, respectively. Positive or negative signs were added to each transformation on different days for each subject.

*However, among the eight subjects, the arms of four subjects (CY, SN, TT, KH) are supported by supporting rods installed at the manipulandum.

cursor (x, y) (Fig. 2). R is a constant matrix expressing the rotational movement.

$$\begin{pmatrix} x \\ y \end{pmatrix} = R \begin{pmatrix} p \\ q \end{pmatrix}$$

If the hand is moved in the same way as the case of no transformation without learning this rotational transformation, the distance between target and cursor will became large on the screen. The rotational angles used in the transformation are +60° (Fig. 2A) and –60° (Fig. 2B). The transformations based on these rotational angles are expressed by R^+ and R^-, respectively.

In the viscous transformation, the hand of the subject is generated with the following forces (f_x, f_y) by PFM during task execution (upper sides in A and B of Fig. 3).* If the hand is to be made reaching the target directly as in the case of no transformation, the path of the hand will bend due to the influence of this force field, and so the length will become longer than the straight-line path:

$$\begin{pmatrix} f_x \\ f_y \end{pmatrix} = B \begin{pmatrix} \dot{p} \\ \dot{q} \end{pmatrix}$$

This force is calculated proportional to the velocity of the hand. B is a second-order square matrix expressing the viscous coefficients. The viscous field used in this experiment is the same kind of rotational force field used by Shadmehr and Mussa-Ivaldi [15]. We have used two kinds of force fields—B^+ (lower side in A of Fig. 3) and B^- (lower side in B of Fig. 3)—in which the positive and negative signs of the diagonal components of B are mutually different. All of the subjects are inexperienced for this kind of transformation task.

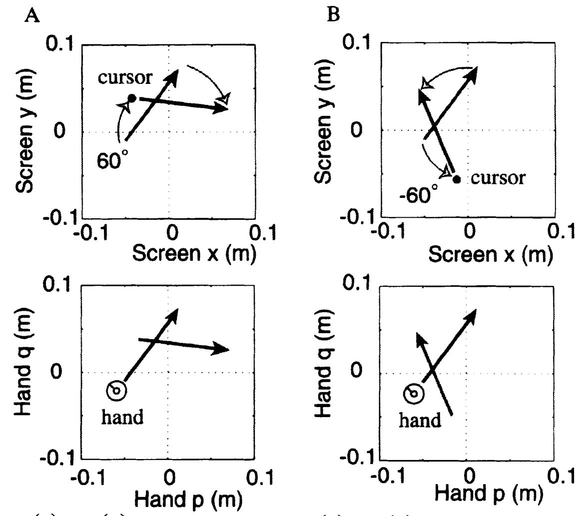
2.2. Analyses

2.2.1. Filtering

The position data are smoothed by a fourth-order Butterworth filter with a cutoff frequency 20 Hz. The differential values of the position data are calculated analytically by estimating a second-order equation passing through three points. The starting time point and ending time point of the respective movements are determined by using the curvature [13]. The threshold value of the curvature is 3 mm^{-1}. Among the 10 trials contained in each set, the first trial is excluded from the analysis.

*In the viscous transformation and combined transformation, subjects RB, AO, NY, HM have experienced the following viscous field which is somewhat smaller than the value of the diagonal components of the other subjects.

$$B^+ = \begin{pmatrix} +12 & -13 \\ -13 & -12 \end{pmatrix} \quad B^- = \begin{pmatrix} -12 & -13 \\ -13 & +12 \end{pmatrix}$$



$$\begin{pmatrix} x \\ y \end{pmatrix} = R^+ \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} \cos 60° & -\sin 60° \\ \sin 60° & \cos 60° \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix}$$

$$\begin{pmatrix} x \\ y \end{pmatrix} = R^- \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} \cos(-60°) & -\sin(-60°) \\ \sin(-60°) & \cos(-60°) \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix}$$

Fig. 2. Rule for rotational transformations. A cursor indicating the hand position was projected onto a CRT screen with constant rotation.

The errors of the reaching movements are calculated by using the measured position data of the hand. In the case of rotational transformation, the data of rotational transformation of the measured positions are used in order to realize the positions displayed on the screen during experiments.

First, the distance between the hand position and the target position at the ending point of the movement decided by the curvature is determined. This distance will be called the target error. When the rules of the rotational transformation are learned, the hand will accurately reach the target and the target error will become smaller [12]. Moreover, we have calculated the length of the path along which the hand has moved during one trial. This length will be called the path length. The viscous force field will influence the hand path as a way of bending it. However, the hand will compensate this influence and it will reach the target directly accompanying the learning, and the path length will become shorter [15]. Since the values of the target error and path length will decrease accompanying the learning, they can be regarded as the indices of learning (Fig. 4). In this paper, to express the progress of learning from now on, the fact that these values are small or large will be represented as: the error level is low or high. Namely, if the error level is high, it is regarded as not learned, and if the error level is low, it is regarded as learned.

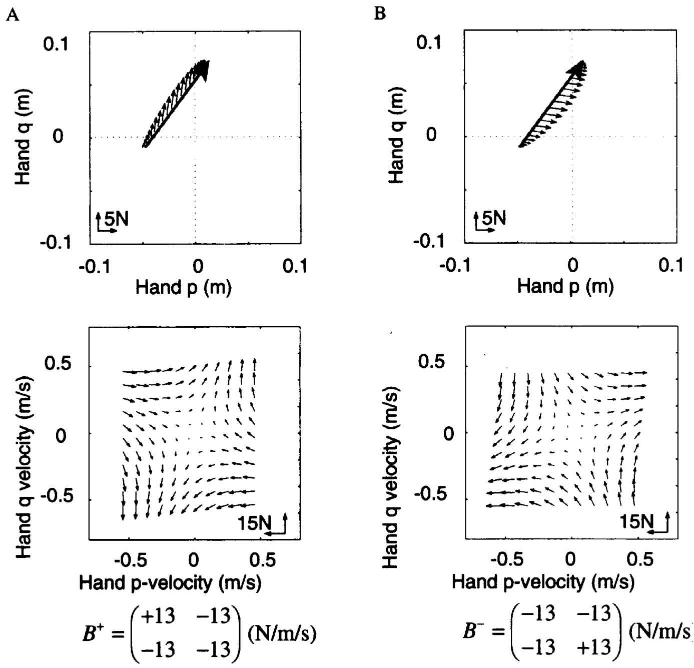


Fig. 3. Rule for viscous transformations. External forces perturbed the hand (p,q) in proportion to the hand velocity. The upper sides in A and B are examples of forces that acted on the hand during reaching movements. These forces were predicted using minimum jerk trajectories (movement distance: 10 cm, movement duration: 500 ms) according to Flash and Hogan [20]. The lower sides in A and B show force fields. These figures are based on force data predicted by each viscosity, B^{+} or B^{-}.

2.2.2. Dispersion analysis

We have investigated the effect of the experiment performed on the first day on the experimental result of the second day; the effect of the previous transformation learning on the later transformation learning; and the learning effect due to the repeated practice learnings. Then, by using the values of the target errors and path lengths, we have performed the dispersion analyses of the three factors of experimental days (the first day and second day), transformations (no transformation, rotational transformation, viscous transformation, and combined transformation), and number of sets (1 to 30). Here, in accordance with the number of trials of no-transformation tasks, the data from the 1st set to the 30th are used.

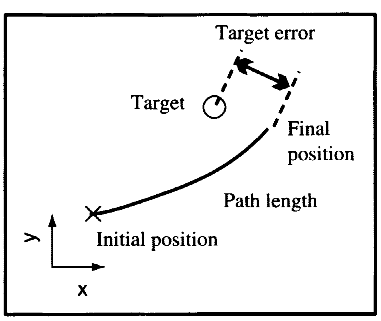


Fig. 4. Indices of learning. The target error was calculated as the distance between the final hand position after movement and the target position. The path length was the length from the initial to final hand position.

Next, as to the effect of order, the dispersion analysis of one factor is performed by using the data of all subjects. In the composition experiment, order 1 contains the data of no transformation only, orders 2 and 3 contain the data of rotational transformation and viscous transformation, and order 4 contains the data of combined transformation only. In the decomposition experiment, order 1 contains the data of no transformation only, order 2 contains the data of combined transformation only, and orders 3 and 4 contain the data of rotational transformation and viscous transformation (see Table 1). The types and numbers of the transformations contained in orders 2 and 3 in the composition experiment are the same; and those contained in orders 3 and 4 in the decomposition experiment are the same. Therefore, here, orders 2 and 3 or orders 3 and 4 are compared in pair by t examination, respectively.

As to the preliminary experiment investigating the effect of the experimental days, the respective single transformations of the first day and the corresponding single transformations of the second day are compared in pair by t examination. The Bonferroni–Dun method is used in all examinations of significant errors.

2.2.3. Modeling by exponential function

The data of target errors or path lengths are averaged for every set in each session. Based on the prediction derived from the hypothesis of multiple internal models described in the Introduction of this paper, the following analyses are performed. Since the combined transformation task of the composition experiment (Fig. 5A right, Composition R+B) is performed after 50 sets of the same transformation have been learned by the single transformation task (Fig. 5A left, Composition R or B), it can be considered to start from the 51st set. On the other hand, the combined transformation of the decomposition experiment (Fig. 5A right, Decomposition R′+B′) has been learned as the first 50 sets and the single transformation (Fig. 5A left, Decomposition R′ or B′) performed next may be considered to start from the 51st set. As to the relationship between the number of trials and the change of error level, the data (1 to 50 sets) obtained in the single transformation of the composition experiment and the data (regarded as 51 to 100 sets) of the decomposition experiment are joined and regarded as the successive data (Fig. 5A left). In reality, the data of the first-half 50 sets and the data of the second-half 50 sets originate from the tasks under the transformations of different signs, respectively. Since the second-half data are the data measured after learning the transformations with the same signs, their initial values and rates of decrease must be lower than those of the first-half data which have not been learned beforehand. Therefore, the error level of the successive data may drop exponentially relatively smoothly. The same can be said for the combined transformation of the composition experiment and that of decomposition experiment (Fig. 5A right). However, when the data themselves of the single transformations of the decomposition experiment and composition experiment (Fig. 5B left) or the combined transformations of the composition experiment and decomposition experiment (Fig. 5B right) are joined, they may not became successive. These data will be called the inverted data. As to the no transformation, the data of joining the first day and second day are regarded as the successive data and the data of joining the second day and first day are regarded as the inverted data.

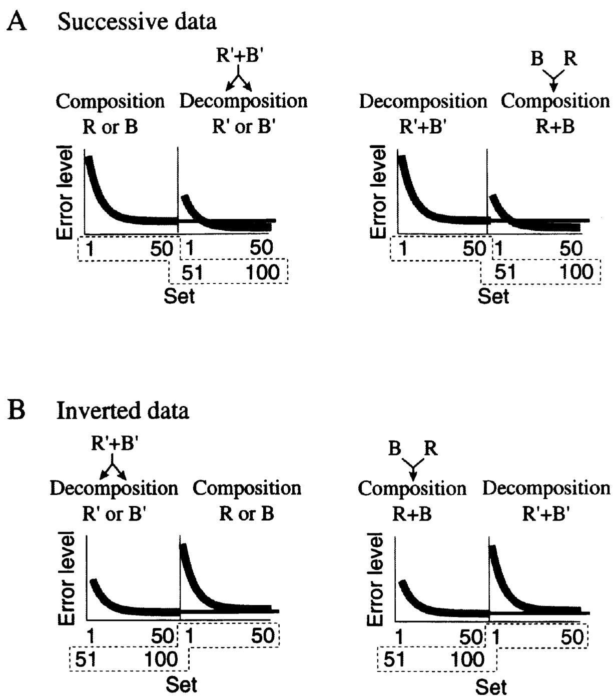


Fig. 5. "Modeling by exponential" of data for mean target errors and mean path lengths. R, B, and R+B denote rotational, viscous, and combined transformations, respectively. R′, B′, and R′+B′ show that the positive and negative signs of transformations in the decomposition experiment are different from those in the composition experiment. The successive (A) or inverted data (B) consist of the sequences of trial sets enclosed by a dashed line. The first half of the joined data is modeled exponentially. The black solid lines show the relationship between the joined data and the model prediction.

First, we have modeled the first-half 30 or 50 sets of the successive data and inverted data by

$$V_{model} = k_0 + k_1 \exp(-k_2 n) \tag{1}$$

n is each set and k_i is a coefficient. V_{model} expresses the value of target error or path length predicted by Eq. (1). For the respective transformation tasks, in order to investigate how much the second half of the data joined by the exponential model obtained here can be explained, the predicted value of the model V_{model} and the absolute error of the successive data or inverted data V_{data} are calculated for the second half of the joined data:

$$| V_{model} - V_{data} | \tag{2}$$

By using t examination, the significant difference between the mean absolute errors obtained from the successive data and inverted data are investigated.

3. Results

For the composition experiment and decomposition experiment, Fig. 6 shows the examples of the hand paths of the early stage and later stage of learning of the respective transformations (from the data of subject RB). In the no-transformation tasks, the hand paths are straight lines and the target errors are also smaller for both. In the early stage of learning of the composition experiment (Fig. 6A), the hand has moved to a direction completely different from the target in the rotational transformation task (R^+) , and the hand has shown a trend of deviating from the target near the target in the viscous transformation task (B^-). However, in the combined transformation task ($R^+ + B^-$) in which the single transformation has been learned, the deviation from the target of the hand is relatively smaller compared to the above-mentioned two transformation tasks. In the early stage of learning of the decomposition experiment (Fig. 6B), the hand has moved to a direction completely different from the target in the combined transformation task ($R^- + B^+$); however, in the rotational transformation task (R^-) and viscous transformation task (B^+) after learning the combined transformation, the deviation from the target of the hand is small. In either case, the hand has reached the target almost directly in the later stage of learning similar to the no-transformation task.

3.1. Effect of previous transformation learning on later learning

As a result of dispersion analyses, for the target error, except the main effect of day and the interaction between day and transformation of subject CY as well as the main effects of day of subjects HM and TT, all main effects of day, transformation, and number of sets as well as the interaction between day and transformation are obtained (*p*

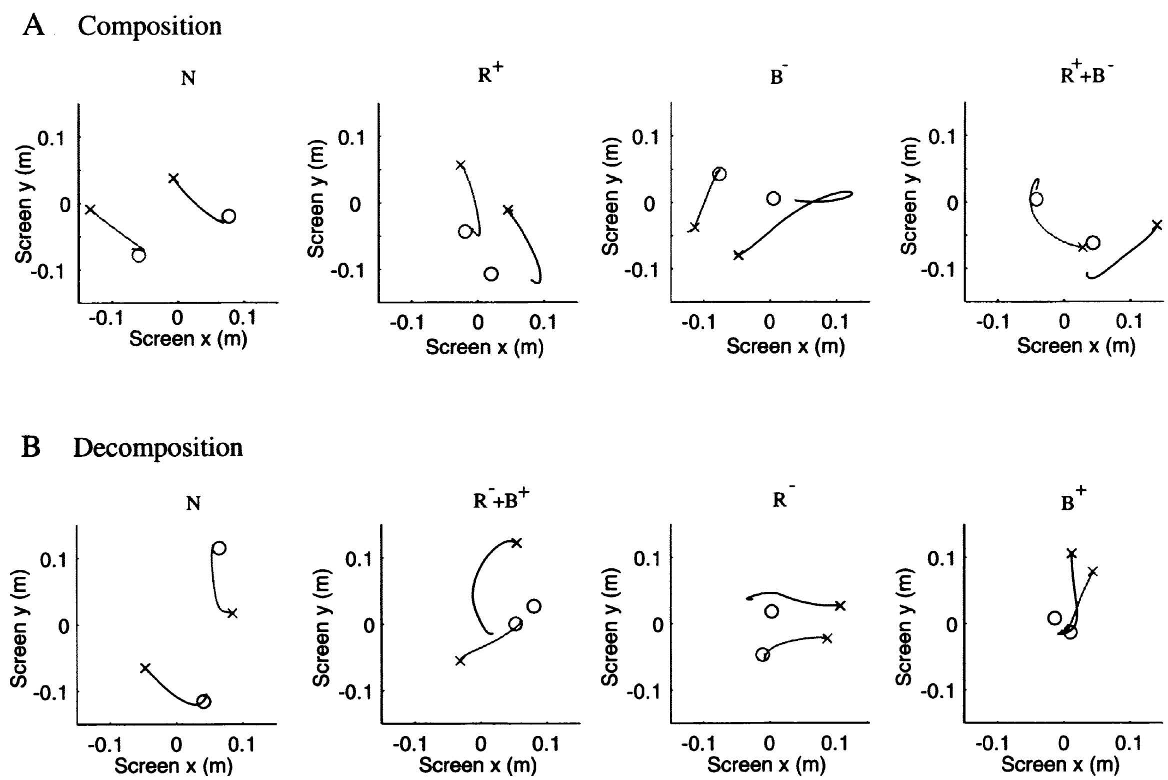


Fig. 6. Hand paths measured under each transformation. This figure shows hand paths in the early stage of learning (1st set) and in the late stage of learning (30th set) of the composition experiment (A) as well as in the early stage of learning (1st set) and in the late stage of learning (30th set) of the decomposition experiment (B). Each path from left to right is derived from data obtained under a normal, rotational, viscous, or combined transformation denoted by N, R^+ and R^-, B^- and B^+, or $R + B^+$ and $R^+ + B^-$, respectively. Black and gray lines indicate paths in the early and late stage of learning, respectively. X and O denote the initial position of the hand and the target position. The origin of this figure is the center of a circle with a radius of 14 cm on a CRT screen where the tasks were performed. The data are from subject RB.

< 0.05). For the path length, except the main effect of day of subjects AO and KH and the interaction between day and transformation of subject KH, all main effects of day, transformation, and number of sets as well as the interaction between day and transformation are obtained ($p < 0.05$). For the two indices of target error and path length, the combined transformations themselves or the single transformation themselves of composition experiment and decomposition experiment are compared in pair. In the pair comparison, the data of seven subjects excluding CY in the target error and seven subjects excluding KH in the path length, whose interaction between day and transformation has been seen, are used. As a result, for the comparison of 33 pieces (21 with significant difference) among 42 pieces (7 subjects × 3 types of transformations × 2 indices), the error level of decomposition experiment is lower for single transformation and the error level of composition experiment is lower for combined transformation (left in upper row of Table 2). Conversely, for the comparison of 9(5) pieces, the error level of composition experiment is lower for single transformation and the error level of decomposition experiment is lower for combined transformation (left in lower raw of Table 2).

From the data of subject RB, the relationships of the number of trials with the successive data and inverted data are shown in Fig. 7. In the case of this subject, almost all successive data decrease exponentially relatively smoothly accompanying the number of trials, and the absolute error with the exponential model is also small. From the early stage to the later stage of learning (1 to 50 or 51 to 100 sets) in the respective sessions of transformation tasks, the mean target error for every set has decreased (Fig. 7A upper) and the mean path length has became shorter (Fig. 7A lower). The error level of no-transformation task is low throughout all trials and is almost the same as the error level of the later stage of learning of the respective transformation tasks. On the other hand, particularly for the rotational and combined transformations, the inverted data are not smooth and the absolute error with the exponential model is large (Fig. 7B). Table 3 summarizes the results of the *t* examinations of the significant differences of the mean absolute errors between the successive data and the prediction by model and between the inverted data and the prediction by model. For the comparison of 33 pieces (16 with significant difference) among 48 pieces (8 subjects × 3 types of transformations × 2 indices), the mean absolute error between the successive data and the prediction by model is smaller (upper row, right of Table 2); and for the comparison of 15(6) pieces, the mean absolute error between the inverted data and the prediction by model is smaller (lower row, right of Table 2).

3.2. Effect of order

As a result of dispersion analyses, the main effects of the order have been obtained [target error: $F(df)$ = 231.87(7), $p < 0.0001$; path length: $F(df)$ = 185.74(7), $p < 0.0001$]. In the decomposition experiment, the target error of order 4 is significantly larger than that of order 3 ($p < 0.0001$). In the composition experiment, the target error of order 3 is significantly larger than that of order 2 ($p < 0.0001$). In both experiments, there is no significant difference between orders for the path length.

3.3. Effect of experimental day

As to the target error, there is no significant difference between the first and second day for all transformations. As to the path length, the second day is significantly longer for all transformations ($p < 0.0001$).

Table 2. The number of subjects whose results support or do not support the multiple internal model hypothesis for an ANOVA and exponential modeling

	Supported							
	ANOVA				Modeling			
	R	B	R+B	total	R	B	R+B	total
Target error	6 (2) / 7	4 (3) / 7	7 (5) / 7	18 (12) / 21	6 (1) / 8	4 (2) / 8	7 (3) / 8	17 (6) / 24
Path length	7 (4) / 7	4 (3) / 7	5 (4) / 7	16 (11) / 21	7 (6) / 8	3 (2) / 8	6 (2) / 8	16 (10) / 24
Total	13 (6) / 14	8 (6) / 14	12 (9) / 14	33 (21) / 42	13 (7) / 16	7 (4) / 16	13 (5) / 16	33 (16) / 48
	Not supported							
	ANOVA				Modeling			
	R	B	R+B	total	R	B	R+B	total
Target error	1 (0) / 7	3 (3) / 7	0 / 7	4 (3) / 21	2 (0) / 8	4 (2) / 8	1 (0) / 8	7 (2) / 24
Path length	0 / 7	3 (1) / 7	2 (1) / 7	5 (2) / 21	1 (0) / 8	5 (3) / 8	2 (1) / 8	8 (4) / 24
Total	1 (0) / 14	6 (4) / 14	2 (1) / 14	9 (5) / 42	3 (0) / 16	9 (5) / 16	3 (1) / 16	15 (6) / 48

The number supporting the hypothesis / the total number of subjects. Significant numbers are shown in parentheses.
R, B, and R+B denote rotational, viscous, and combined transformations, respectively.

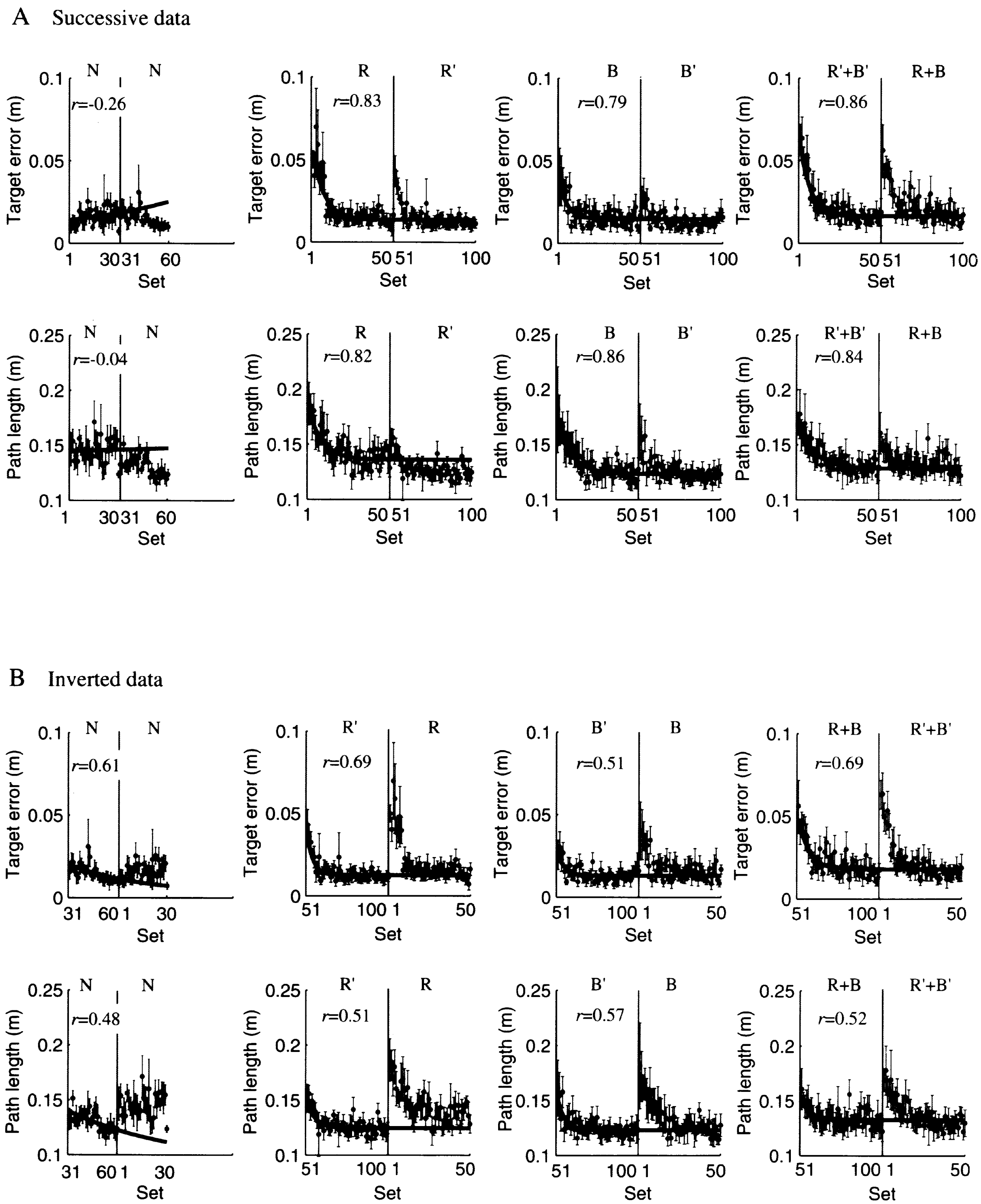


Fig. 7. Relationship between the number of trial sets and (A) successive data or (B) inverted data for mean target errors (A and B, upper) and mean path lengths (A and B, lower). N, R and R′, B and B′, and R+B and R′+B′ denote normal, rotational, viscous, and combined transformations, respectively. R, B, and R+B represent transformations in the composition experiment, and R′, B′, and R′+B′ show those in the decomposition experiment (cf. Fig. 5). The error bars and thick lines indicate the standard errors and the exponential model predictions. The correlation coefficients r between the first half of the successive or inverted data and the model predictions are shown in the upper left of each panel. The data are from subject RB.

4. Discussion

In this research, the environment in which the CNS has altered the kinematic property and the environment in which it has altered the dynamic property are learned by separate internal models, respectively, and their outputs are combined to generate optimal movement commands. The experiments for verifying the hypothesis of multiple internal models are performed. It has been reported by noninvasive measurement using fMRI [19] that after learning different kinematic transformations, nervous activities have been observed in different parts of the cerebellum corresponding to the respective transformations. This report suggests that different internal models have been used for the different kinematic transformations. Therefore, for the kinematic transformation and dynamic transformation used in the experiments, namely, for the rotational transformation and viscous transformation, it may be considered that different internal models have also learned the respective transformations. If the model for a certain transformation has already been acquired, the adaptation to the situation of performing the movement under the same transformation may be faster.

4.1. Result of supporting hypothesis

If the rotational transformation, viscous transformation, and combined transformation are all learned by different internal models or by the same one model, their error levels must be similar regardless of the order of learning the transformation. However, over 80% of the results of dispersion analyses have shown that after learning the single (or combined) transformation first, the error level of target error or path length in the task under combined (or single) transformation is generally low throughout the session. Moreover, less than 70% of the mean absolute error between the successive data and the exponential model is higher than that of the inverted data. Namely, the exponential model for the first half of the successive data has predicted the second half of the data relatively well. Taking the successive data of the rotational transformation task in Fig. 7A as an example, the error level in the early stage of learning drops by learning the combined transformation beforehand and for that reason, the data are joined, relatively smoothly, with the data of the later stage of learning of the task performed without experiencing the combined transformation. Moreover, within each experiment, since the result that the error level is lower for the task whose order is behind, it is good enough to consider that it is not the reason that the error level of transformation learning performed later by the effect of order only has dropped. Namely, the possibility of the simple effect of order which is not related to the type of transformation is rejected.

Moreover, if the results of supporting the hypothesis are obtained in the single transformation of C-D group in which the composition experiment is performed first and in the combined transformation only of D-C group in the decomposition experiment is performed first, there is a possibility that the experience of the first-day experiment has made the execution of the second-day experiment easier. Conversely, if the results of supporting the hypothesis are obtained in the combined transformation of C-D group and the single transformation only of D-C group, there is a possibility that the experience of the first-day experiment has hindered the learning in the second-day experiment. However, in both transformation learnings in both groups, the results that over half have supported the hypothesis are obtained (see Table 2). Furthermore, in the preliminary experiment in which the effect of experimental days is studied, a negative transfer has been seen only for the path length between the first- and the second-day experiment; however, the positive transfer is not seen. Therefore, it cannot be said that the effect of experimental days has necessarily influenced the results.

From these results, the possibility that the three types of transformations adopted in the experiments will be learned by three types of internal models, respectively, as well as the possibility that the single internal model will learn on all such occasions are rejected. Therefore, in the composition experiment, the internal models corresponding to the respective viscous transformation and rotational transformation are prepared, and both of them may be used in the combined transformation task. Moreover, in the decomposition experiment, the internal model corresponding to the rotational transformation and the internal model corresponding to the viscous transformation are acquired simultaneously under the combined transformation and the internal model corresponding to the same transformation may be used separately under the single transformation task.

4.2. Result of not supporting hypothesis

On the other hand, a little under 20% of the result in the dispersion analyses and a little over 30% of the result in the modeling by exponential function have not supported the hypothesis (see Table 2). As to the reason for this result, the fact that the scheme adopted by the subject differs depending on the day, the interference by the first-day experience (negative transfer), and the hesitation about the inexperienced environment may be considered.

For the rotational transformation, the subject cannot accurately reach the target if he does not learn the rules of transformation. Subject SN has reported that he felt that the execution is more difficult in the initial stage of the task because the second-day learning is interfered with by the first-day learning. For subject CY, the target error after combined transformation learning is smaller than that before learning, although the absolute error of the successive data is larger than that of the inverted data. The result for the exponential model is one among four which have not supported the hypothesis (see Table 2). Since the subject has known the type of transformation by teaching, there is a possibility that he has the attitude of ready to use the internal model acquired in the first day. In that case, since the sign of the rotation is the opposite, its internal model has an error of 120° appearing between the predicted position and the hand position present on CRT. Therefore, the target error becomes larger than the first day temporarily. However, according to the multiple internal models [6], the gain of the learning for the model which presented a large error becomes smaller in such a way that the representation of the internal model which has already been acquired will not be altered, and the different internal models start to learn. If this idea is correct, the results of these subjects, which are inconsistent at a glance, can be interpreted as follows. Although the opposite-sign transformation is learned anew in the second-day combined transformation, since the first-day internal model is activated again in the next rotational transformation and the initial value of the error becomes larger, it has immediately changed over to the internal model acquired immediately before to reduce the error. Moreover, the negative transfer has been observed not only in the data of subject CY but also in the preliminary experiment on the effect of experimental days. If the aforementioned interpretation on the negative transfer is correct, it may show that the transformations whose signs are different may be learned by the different internal models. However, this hypothesis does not explain the mechanism itself that the internal model acquired the last time is activated by mistake and the negative transfer occurs.

For the viscous transformation, the subject does not necessarily learn the rules of learning and can achieve the task by commonly exciting the arm muscles to generate, using the viscous field, the force for resisting the disturbance. In the viscous transformation, subjects AO and SN belonging to C-D group showing the results of 8 pieces among 15 pieces (see Table 3) which do not support the hypothesis have experienced the viscous transformation on the beginning of the first day. Since they are unfamiliar with both the transformation and experimental condition, they are strained and achieved the task with arm stiffened. As a result, there is a possibility that the error level has dropped generally. Therefore, even if the error level has dropped due to the quickening of the learning effect of the combined transformation in the decomposition experiment of the second day, it is possible that the difference between both days cannot be obtained statistically. There is a possibility that the four subjects (seven pieces) belonging to D-C group in which the error level is low on both days may have also

Table 3. Mean absolute errors between joined data and exponential model

	Target error								
	R			B			R+B		
Subject	t	df		t	df		t	df	
RB	2.11*	67	s	2.13*	69	s	0.04	93	i
AO	1.24	65	s	3.02**	73	i	2.49**	58	s
SN	0.73	97	s	3.01**	67	i	5.60**	50	s
CY	0.86	95	i	0.39	97	s	0.03	93	s
NY	0.62	71	s	0.67	97	i	0.47	85	s
HM	1.07	73	s	6.90**	65	s	1.23	64	s
TT	0.26	93	s	2.31*	64	i	4.08**	77	s
KH	0.26	68	i	0.31	97	s	1.58	53	s

	Path length								
	R			B			R+B		
Subject	t	df		t	df		t	df	
RB	4.21**	60	s	2.75**	80	s	0.82	90	s
AO	1.72*	78	s	1.49	97	i	7.80**	90	s
SN	2.13*	88	s	3.03**	86	i	4.26**	81	s
CY	1.93*	63	s	0.96	71	s	0.95	95	s
NY	0.32	94	s	1.99*	88	i	4.63**	61	s
HM	1.61	74	s	0.90	97	i	0.05	97	i
TT	4.82**	78	s	6.63**	70	s	8.56**	97	i
KH	0.01	97	i	0.21	93	i	1.08	69	s

t: t values, df: degrees of freedom, s: succesive data, i: inverted data.
R, B, and R+B denote rotational, viscous, and combined transformations, respectively. The type of joined data having a smaller number of mean absolute errors is shown to the right of the degree of freedom.
$*p < .05$ $**p < .01$

taken the scheme of common excitation; however, it may not be due to the stiffness because they have experienced the combined transformation earlier than the viscous transformation.

For the combined transformation, the results which do not support the hypothesis are five pieces. Among them, four pieces shown by subjects HM and TT belonging to D-C group are all the results for the path length. All of them have experienced the combined transformation on the beginning of the first day. Since the most difficult transformation task has been performed first, the scheme is adopted that the hand will not be moved too much because of stiffness, and so it is possible that the results of large target error and short path length are obtained.

5. Conclusions

In the preceding research, it has been reported that after a certain transformation task is learned, the next learning is interfered with depending on the time at which the next different transformation task is performed [16]. In the experiments, since the interval time between the respective transformation tasks is too short, there is also a possibility that the learning has been interfered with depending on the subjects. It may be necessary that the time between tasks must be controlled in the future.

However, we are able to explain more than 90% of the results of the experiments by the multiple internal models or the interpretation which does not confront them. Namely, the following two possibilities have been suggested. The internal model for kinematic transformation and the internal model for dynamic transformation have been learned separately as the elements of the combined transformation. In the single (or combined) transformation performed after learning the combined (or single) transformation, it is not the case that separate internal models are prepared anew but rather the internal model acquired in the previous transformation learning is utilized. Therefore, the CNS has learned separately the multiple internal models which compensate the respective transformations, and composed or decomposed the respective' internal models in accordance with the change of the environment.

Acknowledgment. A portion of this research was performed as part of the "Research on communication of primates including humans," based on the 1999-FY Adjustment Fund for Science and Technology Promotion of the Science and Technology Agency, and moreover, it has received the assistance of HFSP.

REFERENCES

1. Kawato M, Furukawa K, Suzuki R. A hierarchical neural network model for the control and learning of voluntary movements. Biol Cybern 1987;56:1–17.
2. Kawato M. Computational theory of brain. Sangyotosho; 1996.
3. Kawato M. Internal model of cerebellum exterior—computational engine of human intelligence: Imagination, communication, language, thinking and consciousness. Separate Volume, Mathematical Science, p 194–208, 1997.
4. Lacquaniti F, Guigon E, Bianchi L, Ferraina S, Caminiti R. Representing spatial information for limb movement: Role of area 5 in the monkey. Cerebral Cortex 1995;5:391–409.
5. Kalaska JF, Cohen AD, Hyde ML, Prud'homme M. A comparison of movement direction-related versus load direction-related activity in primate motor cortex, using a two-dimensional reaching task, J Neurosc 1989;9:2080–2102.
6. Kawato M, Wolpert DM. Internal models for motor control. In Glickstein M (editor). Sensory guidance of movement. John Wiley & Sons; 1998. p 291–307.
7. Wolpert DM, Kawato M. Multiple paired forward and inverse models for motor control. Neural Networks 1998;11:1317–1329.
8. Wolpert DM, Miall RC, Kawato M. Internal models in the cerebellum. Trends Cogn Sci 1998;2:338–347.
9. MacGonigle BO, Flook JP. Long-term retention of single and multistate prismatic adaptation by humans. Nature 1978;272: 364–366.
10. Welch RB. Adaptation to space perception. In Boff KR, Kaufman L, Thomas JP (editors). Handbook of perception and human performance, Vol. 24. Wiley; 1986. p 24–25.
11. Welch RB, Bridgeman B, Anand S, Browman K. Alternating prism exposure causes dual adaptation and generalization to a novel displacement. Percept Psychophys 1993;54:195–204.
12. Imamizu H, Shimojo S. The locus of visual-motor learning at the task or manipulator level: Implications from intermanual transfer. J Exp Psych Hum Percept Perform 1995;21:719–733.
13. Imamizu H, Uno Y, Kawato M. Internal representations of the motor apparatus: Implications from generalization in visuomotor learning. J Exp Psychol Hum Percept Perform 1995;21:1174–1198.
14. Ghahramani Z, Wolpert DM. Modular decomposition in visuomotor learning. Nature 1997;386:392–395.
15. Shadmehr R, Mussa-Ivaldi FA. Adaptive representation of dynamics during learning of a motor task. J Neurosc 1994;14:3208–3224.
16. Brashers-Krug T, Shadmehr R, Bizzi E. Consolidation in human motor memory. Nature 1996;382:252–255.
17. Conditt MA, Gandolfo F, Mussa-Ivaldi FA. The motor system does not learn the dynamics of the arm by rote memorization of past experience. J Neurophysiol 1997;78:554–560.
18. Sheidt RA, Conditt MA, Reinkensmeyer DJ, Mussa-Ivaldi FA. Motor adaptation persists in the absence of kinematic errors. Soc Neurosci Abstr 85.4, 1997.
19. Imamizu H, Miyauchi S, Tamada T, Sasaki Y, Takino R, Yoshioka T, Putz B, Kawato M. Multiple representations for visuomotor learning in the cerebellum: A functional MRI study. Neuro Image Vol. 7, No. S819, 1998.
20. Flash T, Hogan N. The coordination of arm movements: An experimentally confirmed mathematical model. J Neurosci 1985;5:1688–1703.

AUTHORS (from left to right)





Eri Nakano completed her doctoral course at Kobe University in 1999 and became a visiting researcher at Advanced Telecommunications Research Institute International. She holds a Ph.D. degree. She is engaged in psychological and computational theoretic research on human movement control and learning mechanism. She is a member of the Japanese Psychological Association, the Japan Society of Neuroscience, and the North American Society of Neuroscience. Since 2001, she has been with the Digital Human Laboratory, National Institute of Advanced Industrial Science and Technology.

John R. Flanagan obtained a Ph.D. (psychology) degree from McGill University in 1992. He was an MRC postdoctoral fellow at the University of Cambridge in 1992. He was a research associate at Columbia University in 1994, and has been one at Queen's University since 1995. He is engaged in psychological and computational theoretic research on human movement control and learning mechanism. He is a member of the North American Society of Neuroscience and the Society for the Neural Control of Movement.

Hiroshi Imamizu obtained his doctoral course credits and left from the Cultural Science Research Division of the University of Tokyo in 1987. He was a visiting researcher at ATR Human Information Processing Research Laboratories in 1992. He has been a Group Leader in the Kawato Dynamic Brain Project of the Japan Science and Technology Corporation. He is engaged in psychological and computational theoretic research on human movement and learning mechanism. He holds a doctoral (psychology) degree. He is a member of the Japanese Psychological Association, the Japan Society of Fundamental Psychology, the Japan Society of Developmental Psychology, the North American Society of Vision and Ophthalmology, the North American Society of Neuroscience, and the Japan Society of Neuroscience.

Rieko Osu obtained the recognition of research guidance in the second half of her doctoral course and left the Literature Research Division of Kyoto University. She became a Training Researcher at ATR Human Information Processing Research Laboratories in 1996. She holds a Ph.D. degree. She has been a researcher at the Kawato Dynamic Brain Project of the Japan Science and Technology Corporation since 1996, engaged in research on movement function of the living body. She is a member of the Japanese Psychological Association, the Japan Society of Neuroscience, and the North American Society of Neuroscience.

Toshinori Yoshioka graduated from the Department of Physics of Kyoto Sangyo University in 1988 and joined CSK Co., Ltd. He moved to the ATR Vision and Hearing Mechanism Research Laboratory in 1992, and beginning in 1993 he was engaged in the development of research applications at ATR Human Information Processing Research Laboratories. In 1996, he joined the Kawato Dynamic Brain Project of the Japan Science and Technology Corporation. He has been with Human Information Science Laboratories since 2001.

Mitsuo Kawato (member) completed his doctoral course at Osaka University in 1981. After serving as a research associate and lecturer there, he moved to the ATR Vision and Hearing Mechanism Research Laboratory in 1988. In 1992, he became Section Head of the Third Research Section of ATR Human Information Processing Research Laboratories. He has been with Human Information Science Laboratories since 2001. He was a visiting professor at the Research Institute for Electronic Science of Hokkaido University from 1992 to 1995, at Genoa University (Italy) and Kanazawa Institute of Technology since 1994, and has concurrently been a General-Responsible-Person of the Kawato Dynamic Brain Project of the Japan Science and Technology Corporation since 1996. He is engaged in research on computational theoretic neuroscience of the brain. He has received a Yonezawa Award, Osaka Science Award, Director-General Award of the Science and Technology Agency, and Tsukahara Award. He is Editor-in-Chief of *Neural Networks* as well as Editorial Director and a member of the Award Selection Committee of the Japan Society of Neural Networks.