RESEARCH ARTICLE

# Explicit contextual information selectively contributes to predictive switching of internal models

**Hiroshi Imamizu · Norikazu Sugimoto · Rieko Osu ·
Kiyoka Tsutsui · Kouichi Sugiyama · Yasuhiro Wada ·
Mitsuo Kawato**

**Abstract** Many evidences suggest that the central nervous system (CNS) acquires and switches internal models for adaptive control in various environments. However, little is known about the neural mechanisms responsible for the switching. A recent computational model for simultaneous learning and switching of internal models proposes two separate switching mechanisms: a predictive mechanism purely based on contextual information and a postdictive mechanism based on the difference between actual and predicted sensorimotor feedbacks. This model can switch internal models solely based on contextual information in a predictive fashion immediately after alteration of the environment. Here we show that when subjects simultaneously adapted to alternating blocks of opposing visuomotor rotations, explicit contextual information about the rotations improved the initial performance at block alternations and asymptotic levels of performance within each block but not readaptation speeds. Our simulations using separate switching mechanisms duplicated these effects of contextual information on subject performance and suggest that improvement of initial performance was caused by improved accuracy of the predictive switch while adaptation speed corresponds to a switch dependent on sensorimotor feedback. Simulations also suggested that a slow change in output signals from the switching mechanisms causes contamination of motor commands from an internal model used in the previous context (anterograde interference) and partial destruction of internal models (retrograde interference). Explicit contextual information prevents destruction and assists memory retention by improving the changes in output signals. Thus, the asymptotic levels of performance improved.

## Introduction

Internal models are neural mechanisms that mimic the input-output properties of controlled objects (Wolpert et al. 1995; Brashers-Krug et al. 1996; Kawato 1999; Imamizu et al. 2000). Empirically, two types of information are crucial for the switching of internal models: contextual information such as color or shape of the objects that can be perceived before movement execution, and information about the difference between actual and predicted sensorimotor

H. Imamizu (✉) · R. Osu · M. Kawato
Department of Cognitive Neuroscience,
ATR Computational Neuroscience Laboratories,
2-2-2, Hikaridai, Keihanna Science City,
Kyoto 6190288, Japan
e-mail: imamizu@atr.jp

N. Sugimoto
Department of Computational Neurobiology,
ATR Computational Neuroscience Laboratories,
2-2-2, Hikaridai, Keihanna Science City,
Kyoto 6190288, Japan

R. Osu
National Institute of Information and Communication
Technology, 2-2-2, Hikaridai, Keihanna Science City,
Kyoto 6190288, Japan

K. Tsutsui · K. Sugiyama · Y. Wada
Nagaoka University of Technology,
1603-1, Kamitomioka, Nagaoka, Niigata 9402188, Japan

feedbacks that can be calculated during or after execution. A computational model for switching of internal models, called MOdular Selection and Identification for Control (MOSAIC), has proposed two separate switching architectures for each type of information (Wolpert and Kawato 1998).

Multiple pairs of forward and inverse internal models exist in the MOSAIC model. Forward internal models predict sensory feedback from the efference copy of motor commands while inverse internal models transform desired sensory consequences into the motor commands that achieve them. The MOSAIC model has two architectures for switching the inverse models: a predictor and an estimator (Fig. 1). The predictor receives contextual information and calculates the degree to which each inverse model is responsible for the current environment. The estimator also calculates the degree based on prediction error, which is the difference between actual sensory feedback and sensory feedback predicted by a forward model. An inverse model paired with a forward model that has small prediction error is likely appropriate for the current environment. In Bayesian statistics, the degree from the predictor corresponds to prior probability regarding the environment before movement execution while the degree from the estimator corresponds to likelihood based on observed information. A responsibility signal is a product of prior probability and likelihood, and determines the final degree to which each inverse internal model is responsible for the current environment.

If contextual information is given, the predictor can calculate an appropriate responsibility signal independent of sensorimotor feedback, and the MOSAIC model can conduct an appropriate switch immediately after the alteration. If contextual information is not given, the estimator solely determines the responsibility signal. In this situation, switching ability depends on the appropriateness of the forward internal models gauged by prediction error. We hypothesize that the appropriateness of the internal models gradually changes in a similar fashion to performance changes in neural networks (Kawato et al. 1987). Thus, switching is not immediate or predictive but proceeds gradually in a trial-by-trial fashion. Because the predictor and estimator are separate architectures, the MOSAIC model predicts that the existence of contextual information selectively improves switching by the predictor immediately after the alteration while it does not affect the gradual switching by the estimator based on sensorimotor feedback.

Previous studies have investigated the effects of contextual information (e.g. color, auditory, or kinesthetic cues) on switching or learning in dual tasks, in which subjects were required to switch between opposing visuomotor rotations or force fields (e.g., Cunningham and Welch 1994; Gandolfo et al. 1996; Wada et al. 2003; Miall et al. 2004; Osu et al. 2004). However, little is separately known about the effects on the predictive switching immediately after the alteration and the gradual switching based on sensorimotor feedback. In the current study, we directly manipulated explicit contextual information produced by detailed instruction about the characteristics of opposing visuomotor rotations and investigated whether the information gives different effects to the predictive switch and switching based on sensorimotor feedback. We conducted a computer simulation of the MOSAIC model in the current behavioral task and examined whether the simulation results are consistent with behavioral results.
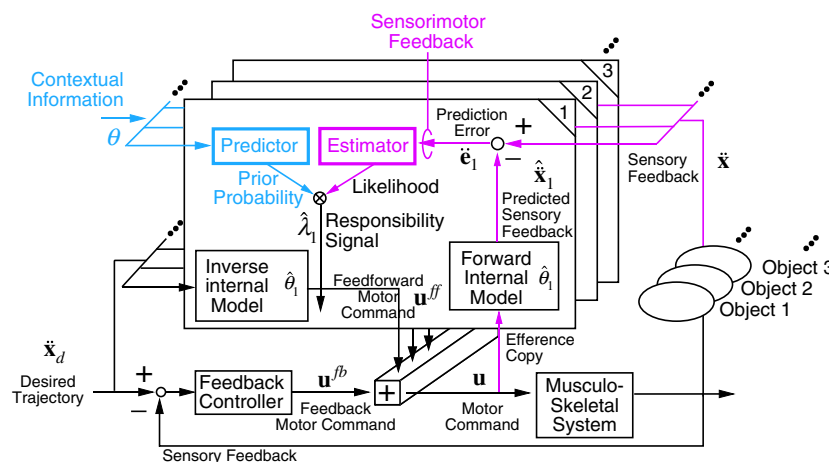


Fig. 1 Architecture of a MOSAIC model during manipulation of multiple objects. The model consists of feedback-controller parts (*bottom architecture*) and feedforward-controller parts (*upper architecture*). *Cyan* parts are related to switching mechanisms based on contextual information while *magenta* parts are related to those based on sensorimotor feedback. Nomenclature for symbols is described in the simulation parts of the METHODS section and Supplementary Material

Furthermore, to confirm significance of the separate architectures for the switching, we also conducted a simulation using a modular network that does not structurally distinguish the predictor and the estimator, and compared the results to those of the MOSAIC model.

## Methods

### Subjects

Ten male subjects (20–35 years of age) participated in this study after giving informed consent. All subjects were right-handed and naive to the purpose of this institutional review board-approved study.

### Task and apparatus

In each trial, subjects made an out-and-back pointing movement from the center start zone to one of eight radial targets (Fig. 2a). The target and a cursor corresponding to the fingertip position were displayed on a computer screen. A
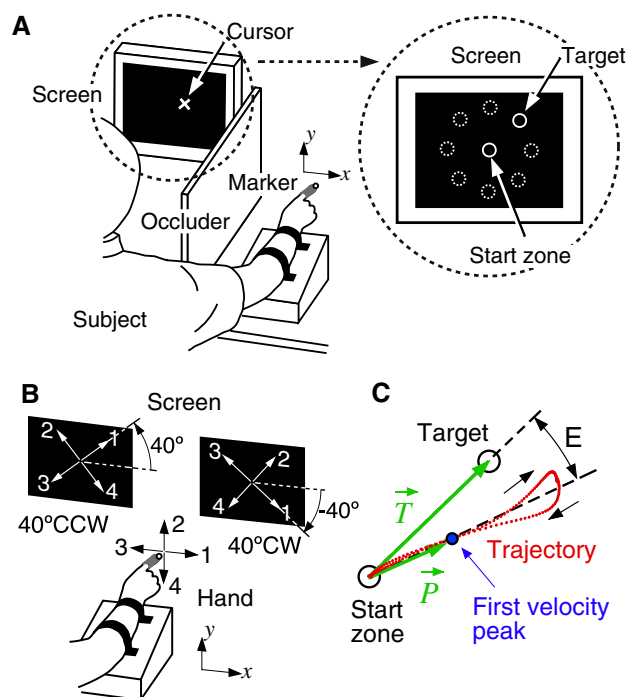


**Fig. 2** **a** Experimental apparatus (*left*) and positions of the start zone and targets on the screen (*right inset*). **b** Relationship between direction of finger movements (*black arrows*) and cursor movements (*white arrows*) under 40° CCW and 40° CW rotations. *Numbers* indicate correspondence between *black* and *white arrows*. **C** Angular error (*E*) in each trial was measured by angular difference between the vector from center of start zone to target ($\vec{T}$) and the vector from center of start zone to the cursor position at the moment of maximal outward velocity ($\vec{P}$)

position recording system (OPTOTRAK, NorthernDigital, Canada) was attached to the right index fingertip, and its position was projected along the coronal plane ($x$–$y$ plane). The cursor (a small ''$x$'') position on the screen was determined by the projected position. The cursor was visible at all times, but subjects could not see their arm or hand. The forearm of the subjects was fixed on a platform, and movements of the wrist and the metacarpophalangeal joint of the right index finger were allowed. Surgical tape fixed the distal and proximal interphalangeal joints of the finger. Arm and hand positions were carefully adjusted to remain comfortable to prevent fatigue. Although the pointing movement may be unstable and inaccurate in comparison to arm movement using the elbow and shoulder, which is frequently used in studies on sensorimotor control and learning, we adopted it considering our future study in the limited space in an fMRI scanner.

The subjects were asked to start movements immediately after the target appears and to move in a fast and smooth motion without trajectory correction. The eight targets were located radially in 45° increments (inset in Fig. 2a). Only one of these targets appeared in each trial. The distance between the center of the start zone and the target was 8.0 cm in hand space. The distance was slightly magnified on the screen (9.0 cm, a visual angle of 2.29°) so that subjects could easily see the cursor trajectory. The average interval was 0.52 s (SD 0.13) between the target onset and the movement onset and 0.56 s (SD 0.28) between the movement onset and its termination.

### Procedure

Several days before the main experiment, subjects conducted a baseline block consisting of 120 trials without visuomotor rotations to become habituated to the pointing task. During the main experiment, the cursor position was rotated 40° clockwise (CW) or counterclockwise (CCW) around the center of the screen (Fig. 2b). The ten subjects were assigned to either an instructed group ($n = 5$) or a non-instructed group ($n = 5$). While Fig. 2b was shown to the instructed subjects, they were informed of the rotation as follows. ''The cursor position will be rotated 40° clockwise or counter clockwise around the center of the screen. Thus, when you move your finger straight right, the cursor will move in the lower right direction under the clockwise condition but the cursor will move in the upper right direction under the counter clockwise condition. Similarly, when you move the finger straight up, the cursor will move in the upper right direction under the clockwise condition but the cursor will move in the upper left in the counter clockwise rotation, etc. To hit the target, you need to rotate the finger trajectory in the opposite direction to the imposed rotation.'' The non-instructed subjects were not

informed of the rotations, and their verbal reports indicated that none used any conscious or systematic strategy.

The eight targets were presented randomly in each cycle of eight trials. One block consisted of 15 cycles (120 trials) and lasted 9 min. Each subject underwent ten blocks with a 3-min break between the blocks. The rotation type (CW or CCW) alternately changed at every block, and the order of rotations was counterbalanced between subjects. The instructed subjects were informed of the rotation type by an experimenter at the beginning of each block.

Analysis

To gauge performance accuracy, we measured the angular error between the target direction and the direction of the cursor movement at the moment of first maximal velocity (Fig. 2c). Outliers within each cycle (eight trials) were excluded using Grubb's test (Grubbs 1969). That is, the maximum or minimum value was expunged from the dataset if $(\overline{Y} - Y_{\max})/\mathrm{SD}$ or $(\overline{Y} - Y_{\min})/\mathrm{SD}$ was larger than

$$\frac{(N-1)}{\sqrt{N}} \sqrt{\frac{t^2_{(\alpha/(2N),N-2)}}{N-2+t^2_{(\alpha/(2N),N-2)}}}.$$

Here, $N, \overline{Y}$, and SD denote the sample size, mean, and standard deviation, respectively. $t^2_{(\alpha/(2N),N-2)}$ denotes the critical value of $t$-distribution with a significance level of $\alpha/(2N)$. $\alpha$ was set at 0.05. The test was iterated until no outliers were detected. The maximum number of outliers in each cycle was two. The errors in the first cycle of each block were included in our analyses, whereas they have sometimes been excluded in previous studies (e.g., Krakauer et al. 1999). We confirmed that similar results were obtained in our study if they were excluded. When we separately computed the mean error for each subject and rotation type (CW or CCW) and performed a $t$-test, there was no significant difference between rotation types ($t(18) = 1.2$). Thus, we omitted the rotation type effects in the statistical analyses described below.

To examine whether behavioral measures change across blocks, we used a nonparametric test for trend (Lehmann 1975). A statistical measure $D$ is defined as

$$D = (T_1 - 1)^2 + (T_2 - 2)^2 + \cdots + (T_n - n)^2.$$

Here, $T_i$ is the rank (from the smallest to the largest) of the $i$th value ($x_i$) in a given time series of data ($x_1, x_2, x_3, ..., x_n$). The $Z$ statistic is calculated as

$$Z = \frac{D - E}{\sqrt{V}}.$$

Here, $E = \frac{n^3 - n}{6}$ and $V = \frac{n^2(n+1)^2(n-1)}{36}$. Because the distribution of $Z$ is approximately a normal distribution, we can reject a null hypothesis that there is not a decreasing trend if $Z$ is larger than the critical value of a normal distribution with a significance level of $\alpha$. This test can also examine an increasing trend. We computed the mean values of the behavioral measure for each block and applied the test to the values.

Simulations

We defined a simplified environment of the pointing task under the visuomotor rotations. When $\mathbf{x}(t) = [x(t) \quad y(t)]^{\mathrm{T}}$ and $\mathbf{u}(t) = [u_x(t) \quad u_y(t)]^{\mathrm{T}}$ represent the cursor position and the motor command at time $t$ respectively, the equation of motion can be described as

$$\begin{bmatrix} \ddot{x}(t) \\ \ddot{y}(t) \end{bmatrix} = R(\theta) \begin{bmatrix} u_x(t) \\ u_y(t) \end{bmatrix}, \tag{1}$$

where $R(\theta)$ is the rotation matrix, and the mass of the finger tip is 1 for simplicity. $T$ represents the transpose of a matrix, and $\ddot{x}$ indicates the second time derivative (acceleration) of $x$. Thus, the cursor will be accelerated in the direction deviating by $\theta$ [radian] from that of the intended motor command. Although subjects made an out-and-back pointing movement in the behavioral experiment, we simulated only the out part of the movement. As mentioned above, the average interval was 0.56[s] between the movement onset and its termination in out-and-back movements. Thus, we set the interval in the simulation at half of 0.56[s]; that is, the out movement began at $t = 0$[s] and terminated at $t = 0.28$[s]. The initial position of the cursor was $[0 \quad 0]^{\mathrm{T}}$ at the beginning of each trial. The path of the desired trajectory ($\mathbf{x}_{\mathrm{d}}(t) = [x_{\mathrm{d}}(t) \quad y_{\mathrm{d}}(t)]^{\mathrm{T}}$) was a line connecting the initial position with one of the eight targets. The velocity-profile of the desired trajectory was bell-shaped. The time interval ($\Delta t$) in the simulation was 0.01[s].

The MOSAIC model observes the position ($\mathbf{x}(t)$) and velocity ($\ddot{\mathbf{x}}(t)$) of the cursor at each time point, and calculates a motor command ($\mathbf{u}(t)$). Each forward internal model in the MOSAIC model estimates the angle of the rotation ($\theta$). The estimated parameter was represented by $\hat{\theta}_i$ ($i = 1, 2, ..., M$). Because there were two types of the rotations in the current simulation, $M$ was 2 (see Supplementary Material for simulations when $M$ was four). An initial value of $\hat{\theta}_i$ was set at 5 before learning. The forward model predicts the acceleration vector of the cursor ($\hat{\ddot{\mathbf{x}}}_i$) from the motor command ($\mathbf{u}$) based on the estimated angle:

$$\hat{\ddot{\mathbf{x}}}_i(t) = R(\hat{\theta}_i)\mathbf{u}(t). \tag{2}$$

An inverse model calculates feedforward motor command ($\mathbf{u}_i^{\mathrm{ff}}$) from the acceleration of the desired trajectory ($\ddot{\mathbf{x}}_{\mathrm{d}}$) :

$$\mathbf{u}_i^{\mathrm{ff}}(t) = R(\hat{\theta})^{-1}\ddot{\mathbf{x}}_{\mathrm{d}}(t). \tag{3}$$

The total output from the MOSAIC is a summation of feedback motor command ($\mathbf{u}^{\mathrm{fb}}$, see the lower parts of Fig. 1) and the feedforward motor commands weighted by the responsibility signal.

The simulation was conducted in instructed and non-instructed conditions. The explicit contextual information, i.e. the current angle of visuomotor rotation ($\theta$), was fed into the predictor in the instructed condition and contributed to calculation of the responsibility signal but it was not fed into the predictor in the non-instructed condition. Because we investigated the effect of explicit contextual information produced by detailed instruction about characteristics of visuomotor rotations, we provided the predictor in simulations with the exact current angle of visuomotor rotation. Details of simulations are described in Supplementary Material.

We also conducted a simulation using a modular architecture that has a single switching mechanism (see Fig. S1 in Supplementary Material). This architecture was constructed in a simulation study of object manipulation (Gomi and Kawato 1993) based on a mixture-of-experts model (Jacobs et al. 1991). The mixture-of-experts model involves expert modules, which are equivalent to internal inverse models, and a gating module. The gating module is a single switching mechanism that does not structurally distinguish contextual information and sensorimotor feedbacks.

## Results

### Behavioral results

Figure 3 shows angular errors as a function of trial number from subjects in the instructed and non-instructed groups. Regarding the baseline block (leftmost gray lines), we computed the angular error averaged across trials for each subject and performed a $t$-test to check differences in baseline pointing accuracy between the two groups. Mean error ($\pm$SD) across subjects was 6.89° ($\pm$2.06°) for the instructed group and 6.73° ($\pm$1.83°) for the non-instructed group. There was no significant difference between groups ($t(8) = 0.11$). To check the effect of target direction on angular error in the baseline block, we computed the error averaged across trials for each subject and target direction and analyzed the averaged error with a one-way ANOVA. The directional effect was not significant ($F(7,72) = 1.75$). Therefore, we could not identify any difference in pointing accuracy between the groups or among the target directions in the natural condition.
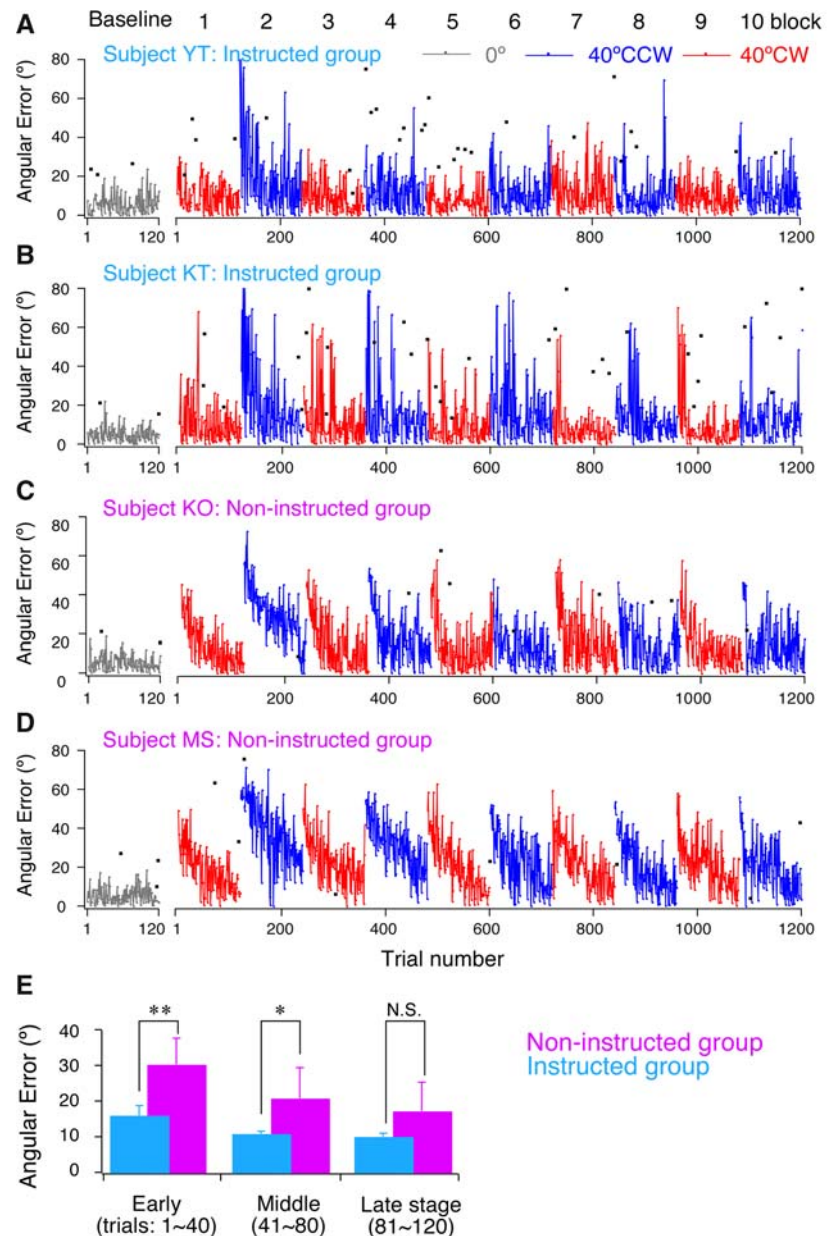
Regarding angular errors as a function of trial number in the main experiment (blue or red lines), the increase of errors at the beginning of each block replicated performance interference between opposing rotations (Krakauer et al. 1999; Bock et al. 2001; Tong et al. 2002; Miall et al. 2004). The decrease in the errors within each block has been characterized as short-term readaptation (Welch et al. 1993). Although trial-by-trial error fluctuates in both the instructed and non-instructed subjects, the error of instructed subjects sometimes increased more abruptly than non-instructed subjects. Correspondingly, the number of detected outliers (black squares in Fig. 3; see ''Methods'') from a total of 1,200 trials was 28.6 (2.38% of the total trials, SD ±6.69) averaged across the instructed subjects while it was 8.4 (0.70%, SD ±0.89) averaged across the non-instructed subjects.

To examine the effects of the explicit contextual information on the short-term readaptation process, we separately calculated mean angular error for the early (trial numbers 1–40), middle (41–80), and late (81–120) stages within each block and investigated the instruction effects on the mean error. The errors were separately averaged across blocks for the stages within each subject and analyzed with a two-way (group × stage) ANOVA. Figure 3E shows the errors averaged across blocks and subjects. The effect of group was significant ($F(1,8) = 8.46$, $P < 0.05$), and error averaged across stages, blocks and subjects in the instructed group (mean ± SD 12.12 ± 3.24) was smaller than in the non-instructed group (22.60 ± 9.51). This indicates a positive instruction effect on overall performance. An interaction effect between the group and the stage was significant ($F(2, 16) = 7.87$, $P < 0.01$), suggesting a different instruction effect that depended on the stages. When we compared error averaged within each stage between groups, error in the instruction group was most significantly smaller than in the non-instructed group in the early stage (Tukey's HSD post hoc test at $P < 0.01$ level), but there was no significant difference in the late stage. This result suggests that the effect of the contextual information was prominent immediately after the alteration of the environment.

To examine a long-term change of subject performances across blocks, we averaged angular error within each block across trials and subjects and plotted the averaged error as a function of block number in the main experiment (Fig. 4a, b). We considered a two-way (instructed/non-instructed group × block) ANOVA as a simple analysis, but we also examined the difference between odd (filled circles) and even (open circles) blocks because the rotation type was different between them. Specifically, blocks were divided into the odd or even groups, and then given an order within each group, e.g., the first block is the first block in the odd group, the second block is the first block in the even group, the third block is the second block in the odd group, the fourth block is the second block in the even group, and so

**Fig. 3 a–d** Angular errors as a
function of trials of four
representative subjects from
instructed and non-instructed
groups. *Black squares* represent
outliers detected by Grubb's test
(see ''Methods''). Values larger
than 80 are plotted at 80. **e**
Angular errors averaged across
trials blocks and subjects (±SD)
separately for the early, middle,
and late stages of the
readaptation process. *$P < 0.05$,
**$P < 0.01$, *NS* not significant
according to Tukey' HSD post
hoc test



on. Then, we applied a three-way (instructed/non-in-
structed group × even/odd block group × block order
within the block group) ANOVA to error averaged within
each block. Error averaged across blocks and subjects in
the instructed group was smaller than in the non-instructed
group ($F(1,8) = 8.49$, $P < 0.05$), indicating a positive
instruction effect on overall performance, as in previous
analysis. The errors in the odd blocks were smaller than
those in the even blocks ($F(1,8) = 7.1$, $P < 0.05$), indicat-
ing that the errors remained small for the rotation type
presented in the first block. We identified a significant
effect of block order ($F(4,32) = 11.16$, $P < 0.01$), sug-
gesting that error averaged within each block changed
across blocks.

Time courses along the abscissa in Fig. 4a and b indi-
cates how the error averaged within each block changed
across blocks. Error averaged within the first block was
small in both the instructed and non-instructed groups, and
reflects initial performance before subjects experienced the
opposing rotations. We could not identify a significant
difference in the error averaged within the first block
between the instructed and non-instructed groups (Tukey's
HSD post hoc test at $P < 0.05$ level). Error averaged
within each block increased in the second and third blocks in
which subjects experienced opposing rotations and gradu-
ally decreased as block number increased. To examine if a
long-term change of performance across blocks occurred
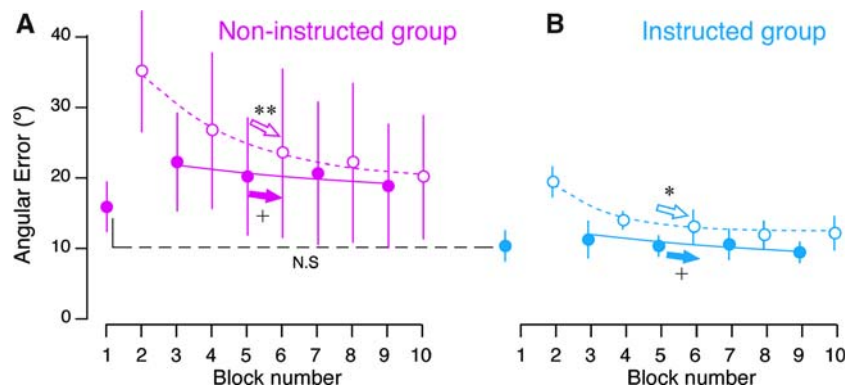under interference between opposing rotations, we

**Fig. 4 a, b** Angular error averaged within each block as a function of block number. Error within each block is averaged across trials and subjects (±SD). *Filled* and *open circles* correspond to odd and even blocks, respectively. *Broken curves* indicate exponential functions fitted to the averaged errors in even blocks. *Solid curves* indicate functions to the averaged errors in odd blocks (from the third to the ninth blocks) + $P < 0.10$ (marginally significant), *$P < 0.05$, **$P < 0.01$, *NS* not significant. *Arrows* indicate results of tests for trends. *Dashed line* across panels indicate results of Tukey' HSD post hoc test

separately applied a test for trend (see ''Methods'') to the error averaged within each block for the odd and even blocks. We did not include the error averaged within the first block in the test because performance in the first block is free from interference. A decreasing trend was significant for even blocks in both the instructed ($P < 0.05$) and non-instructed groups ($P < 0.01$), and marginally significant for odd blocks in both groups ($P < 0.10$). These results suggest a positive long-term change of performance, especially in the even blocks. However, we could not determine whether the decrease in error was caused by the improvement of performance immediately after block alterations, the speed of short-term readaptation, the asymptotic performance after readaptation within each block, or combinations of these components because error averaged within each block can be affected by all of these components.

To separately investigate the effect of explicit contextual information for these components, we fitted the following curve to a time course of trial-by-trial error in each block, as shown by the green curves in Fig. 5a and b:
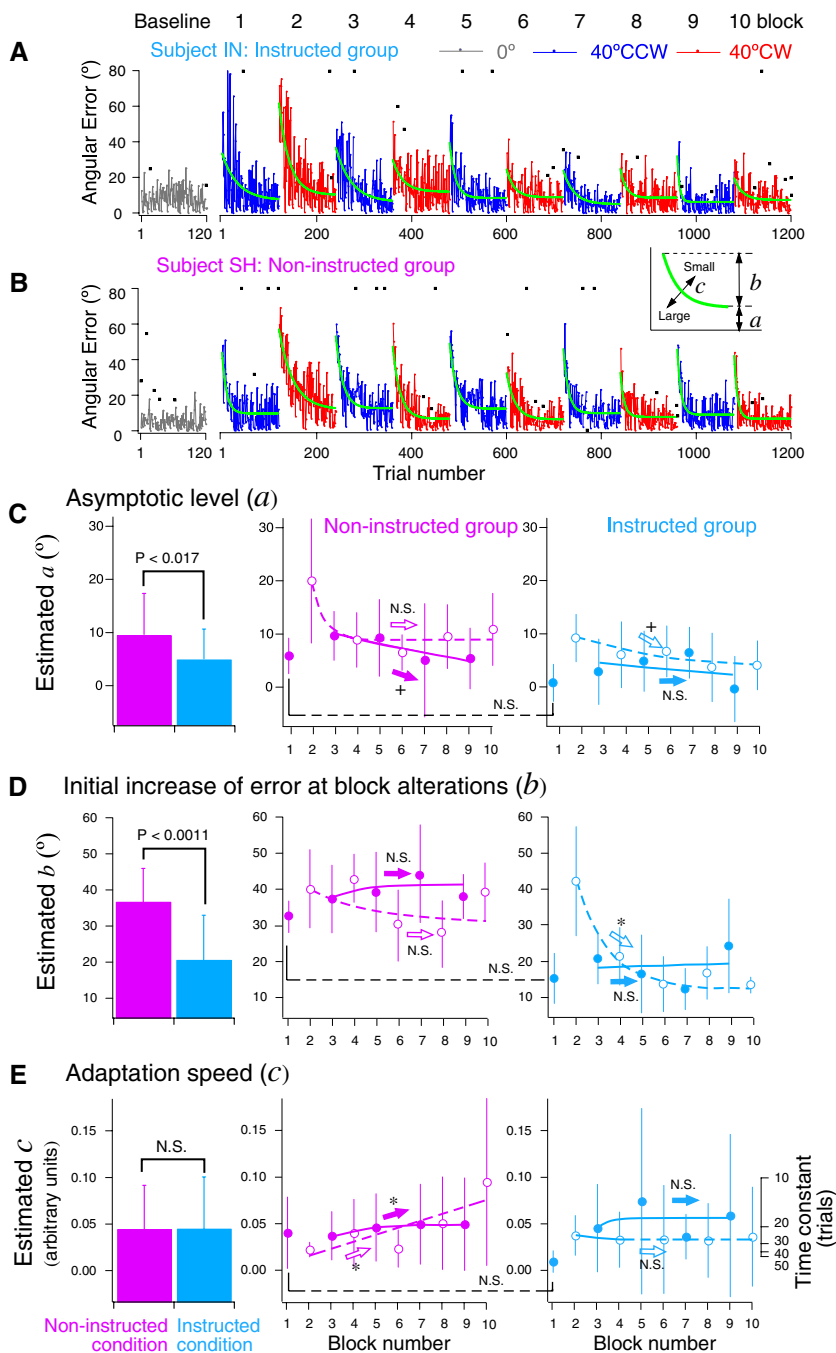
$$y = a + b \exp(-c \cdot x).$$

Here, $x$ and $y$ represent trial number and error, respectively. As illustrated in the inset of Fig. 5b, parameter $a$ corresponds to the asymptotic level within each block when short-term adaptation is completed. Parameter $b$ corresponds to the initial increase of the error at block alterations from the asymptotic level. Parameter $c$ reflects adaptation speed within in each block. Each estimated parameter was separately analyzed with a two-way (instructed/non-instructed group × block) ANOVA. The left panels in Fig. 5c, d and e shows the parameters averaged across blocks and subjects. A significant effect of instruc-

tion was identified in asymptotic level ($a$) ($F(1,8) = 9.00$, $P < 0.017$) and initial increase of error ($b$) ($F(1,8) = 24.9$, $P < 0.0011$). Specifically, the parameter values of $a$ and $b$ in the instructed group were smaller than the values in the non-instructed group. The small value of $b$ in the instructed group corresponds to the small angular error in the early stages within each block in comparison to the error in the non-instructed group (Fig. 3e). However, we could not identify a significant effect on adaptation speed ($c$) ($F(1,8) = 0.01$).

The middle and right panels in Fig. 5c, d and e shows the parameters averaged across subjects as a function of block number. We investigated long-term changes of parameter values across blocks (from the second to the tenth blocks) in the same fashion as Fig. 4. Regarding parameter $a$, time courses of parameter value were similar to those in Fig. 4. We could not identify a significant decreasing trend of the value across the even blocks in the non-instructed group (dashed curve in middle panel of Fig. 5c), but the value markedly decreased from the second to the remaining even blocks. Error averaged within each block also markedly decreased in the even blocks in the non-instructed group (dashed curve in Fig. 4a). A marginally significant decreasing trend ($P < 0.10$) was found for the odd blocks in the non-instructed group (solid curve in middle panel of Fig. 5c). The same trend was also found for error averaged within each block in the odd blocks in the non-instructed group (solid curve in Fig. 4a). A decreasing trend of $a$-value was marginally significant ($P < 0.10$) in the even blocks in the instructed group (dashed curve in right panel of Fig. 5c), and the trend of error averaged within each block in the even blocks in the instructed group was significant ($P < 0.05$, dashed curve in Fig. 4b). Although a decreasing trend was marginally significant in error

averaged within each block in the odd blocks in the
instructed condition (solid curve in Fig. 4b), a significant
trend could not be identified in the odd blocks in the
instructed condition (solid curve in right panel of
Fig. 5c). Although statistical significances weakened in
Fig. 5c, we found similar patterns of time courses in
Figs. 4 and 5c, suggesting a similarity to a certain degree
between error averaged within each block and asymptotic
level.

Regarding the initial increase of error at block altera-
tions (Fig. 5d), the *b*-value in the non-instructed group did
not decrease across blocks and remained around 40° for
both the odd and even blocks (middle panel). We identified
a significant decreasing trend ($P < 0.05$) in the instructed
group in the even blocks but not in the odd blocks (right
panel). The *b*-value in the odd blocks was small across
blocks in the instructed group. Regarding adaptation speed
(Fig. 5e), *c*-values significantly increased ($P < 0.05$) in the

non-instructed group both in the odd and even blocks (middle panel). However, we could not identify a significant trend in the instructed group (right panel).

To examine the validity of exponential fitting, we computed the sum of squares due to error (SSE) associated with the exponential model ($y = a + b \exp(-c \cdot x)$):

$$\text{SSE}(\text{Exponential Model}) = \sum_{i=1}^{n} (y_i - \hat{y}_i),$$

where $n$, $y_i$ and $\hat{y}_i$ denote the number of trials in the main experiment for each subject, observed angular errors and values predicted by the model, respectively. The outliers above mentioned were not included in this analysis. As a reduced model for comparison, we fitted a linear model ($y = a + b x$) to the angular errors, computed SSE associated with the linear model, and calculated the ratio

$$F = \frac{[\text{SSE}(\text{Linear Model}) - \text{SSE}(\text{Exponential Model})]/(p-k)}{\text{SSE}(\text{Exponential Model})/(n-p)}.$$

Here, $p$ and $k$ denote, the numbers of parameters in the exponential model (i.e., 3) and in the linear model (i.e., 2), respectively. The ratio has $F$ distribution with $(p - k)$ and $(n - p)$ degrees of freedom, and the large ratio indicates that the exponential model is more effective than the reduced linear model (Chatterjee and Price 1991). We conducted this analysis for each subject and identified the significant effectiveness of exponential fitting in all subjects. That is, even the smallest ratio among subjects ($F(1, 1188) = 4.64$ in subject MS) was larger than the $F$ value at the $P < 0.05$ level. This result suggests the significance of parameter $c$.

Simulation results

Figure 6 shows results of the simulation using the MOSAIC model. Figure 6a and c indicates time courses of the responsibility signals for the CW and the CCW internal models. A significant difference between the instructed and the non-instructed conditions was observed immediately after alteration of the environment. The responsibility signal rapidly changed in the instructed condition (cyan curves) while the signal changed slowly in the non-instructed condition (magenta curves).

Figure 6b and d indicates the angle of the visuomotor rotation estimated by each internal model ($\hat{\theta}_i$ see ''Methods'' section). These parameters are closely related to the degree of learning acquisition of internal models. Their values were initially set at 5° and approached 40° as learning proceeds. The gray regions correspond to blocks in which the other internal model was appropriate for the environment. Time courses of the parameter values slightly

dropped in the gray regions. The drop of the time course in the instructed condition was smaller than in the non-instructed condition. Consequently, time courses in the instructed condition were superior to those in the non-instructed condition.
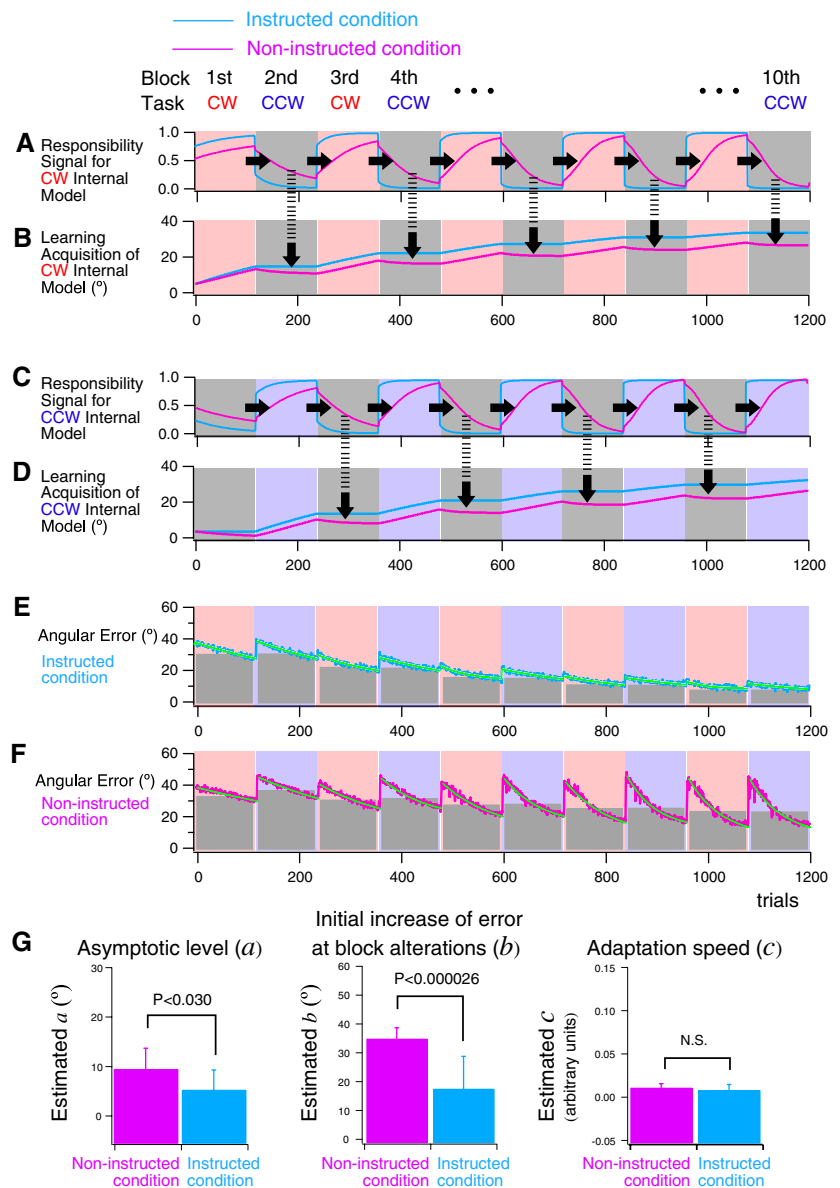
The increase in angular error immediately after the alteration was small in the instructed condition (Fig. 6e) corresponding to rapid change in the responsibility signals. In contrast, the error greatly increased after the alteration in the non-instructed condition (Fig. 6f). Gray bars in Fig. 6e and f indicate error averaged within each block across trials. As the block number increased, the averaged error in the instructed condition decreased more rapidly than in the non-instructed condition, suggesting a positive instruction effect on performance. We fitted exponential curves to the time courses of trial-by-trial error in each block in simulations (green curves) and compared the parameters averaged across blocks (from the second to the tenth blocks) in the instructed condition to those in the non-instructed condition, as shown in Fig. 6g. A significant difference was identified in both parameters $a$ ($t(16) = 2.39$, $P < 0.030$) and $b$ ($t(16) = 5.81$, $P < 0.000026$), but not in parameter $c$ ($t(16) = 0.28$), which is consistent with the behavioral results (left panels in Fig. 5c–e).

Figure 7 shows results of the simulation using the modular architecture based on a single switching mechanism (i.e., the mixture-of-expert model). Figure 7a and c indicates time courses of the responsibility signals. The amplitude of the correct responsibility signal tends to be larger in the instructed condition (cyan curves) than that in the non-instructed condition (magenta curves). However, there was not much difference in the values of the responsibility signals for the two conditions immediately after the alteration of the environment. Figure 7b and d indicates the angle of the visuomotor rotation estimated by each internal model ($\hat{\theta}_i$). We could not find any significant effect of the instruction on the time courses of the learning. Figure 7e indicates angular error as a function of trial numbers. There was little difference in error immediately after the alteration between the instructed and the non-instructed conditions in parallel with the time courses of the responsibility signals. The learning in the modular architecture was very slow (Fig. 7b, d, e: note that horizontal axes are partially magnified). We tried a large parameter value for the learning rate ($\eta^{\text{expert}}$, see Supplementary Material), but the switching occurred much too frequently within each block and the performance became unstable.

## Discussion

We confirmed performance interference (Krakauer et al. 1999; Bock et al. 2001; Tong et al. 2002; Miall et al. 2004)

**Fig. 6** Results of simulation using MOSAIC. Change in responsibility signals (**a**, **c**) and learning acquisition of internal models (**b**, **d**) and angular errors (**e**, **f**) when the MOSAIC learns the CW task in the first block. *Green curves* indicate the fitted exponential curves. **g** Estimated parameters for exponential curves averaged across blocks
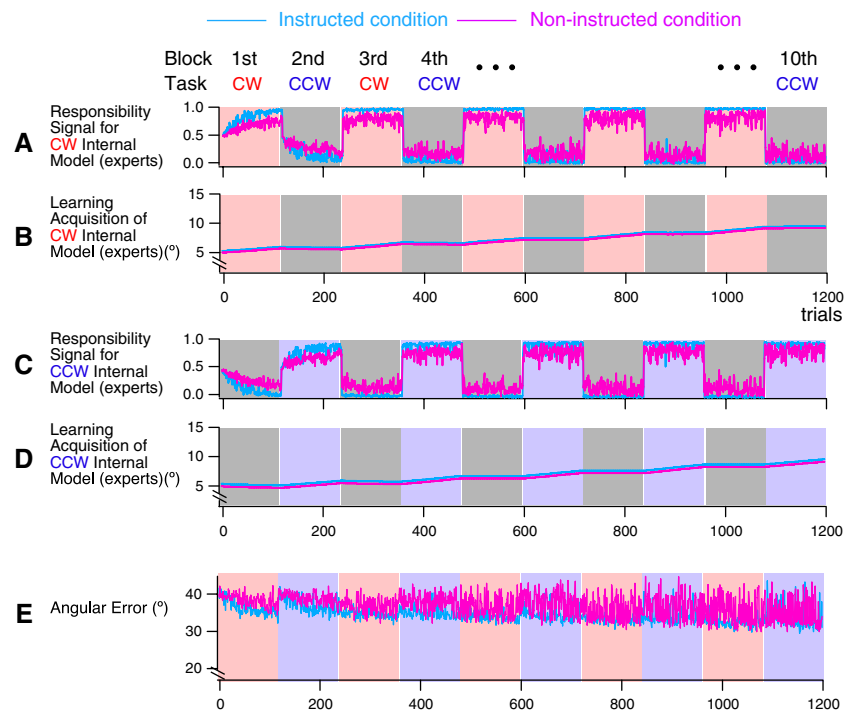
and short-term readaptation (Welch et al. 1993) when subjects simultaneously adapted to alternating blocks of opposing rotations (Fig. 3a–d). Regarding short-term re-adaptation, the effect of explicit contextual information was prominent immediately after alteration of the environment (Fig. 3e). To examine the long-term change of subject performance under interference, we investigated how angular error averaged within each block changed across blocks (from the second to the last blocks). Especially in the even blocks, error averaged within each block decreased both in the instructed and non-instructed groups as block number increased (Fig. 4), suggesting a positive long-term performance change under interference.

Based on the results of exponential fitting, both the asymptotic level after short-term readaptation (*a*) and the

initial increase of error at block alterations (*b*) became small if explicit contextual information existed (left panels in Fig. 5c, d). We investigated long-term changes of these parameters from the second to the last blocks. Regarding the even blocks in the instructed group, a significant decreasing trend was identified in both the *a* and *b*-values (dashed curves in right panels of Fig. 5c and d), although it was only marginally significant in the *a*-value. This suggests that both the asymptotic level and the initial performances at block alterations improved as block number increased. We could not identify a significant decreasing trend in the odd blocks in either the *a*- or *b*-values (solid lines in right panels of Fig. 5c and d). This is probably because the values were small already in the early blocks, e.g., the third block, and difficult to decrease further in the

**Fig. 7** Results of simulation using a modular architecture based on the mixture of experts model. Change in responsibility signals (**a**, **c**), learning acquisition of internal models (**b**, **d**) and angular errors (**e**) when the architecture learns the CW task in the first block. Note that the horizontal axes in **b**, **d**, and, **e** are partially magnified



late blocks ("floor effect"). These results indicate that a significant long-term decrease of error averaged within each block in the even blocks of the instructed group (Fig. 4b) was caused by improvement of both the asymptotic level and the initial performance at block alterations.

Regarding the non-instructed group, we could not identify a significant decreasing trend of the *a*-value in the even blocks (dashed curve in middle panel of Fig. 5c) but the value markedly decreased from the second to the remaining even blocks. The *a*-value in the odd blocks (solid line) decreased as block number increased, although the trend was marginally significant. The *b*-values did not change from around 40° in either the even or odd blocks (middle panel of Fig. 5d). The *c*-value, which corresponds to adaptation speed, gradually and significantly increased across blocks in both the odd and even blocks in the non-instructed group (middle panel of Fig. 5e). Therefore, a long-term decrease of error averaged within each block in the non-instructed group (Fig. 4a) was caused by improvement of both the asymptotic level and short-term adaptation speed.

In the simulation results using the MOSAIC model, we found a difference in the changes of responsibility signals (Fig. 6a, c), the learning acquisition of internal models (Fig. 6b, d), and the error time courses (Fig. 6e–g) between the instructed and non-instructed conditions. However, in our simulation using modular architecture based on a single switching mechanism (mixture-of-expert model), we could not find a difference in the change of responsibility signals at block alterations (Fig. 7a, c), the learning acquisition of

internal models (Fig. 7b, d), or the error time courses (Fig. 7e) between the two conditions. These results suggest that the MOSAIC model can better explain differences in the behavioral results between the instructed and non-instructed groups than modular architecture based on a single switching mechanism.

In the MOSAIC model, the changes of responsibility signals affect the short-term changes of errors and the long-term acquisition of internal models as follows. Below, (CW, CCW) denotes a pair of responsibility signal values for the CW and CCW internal models. In the instructed condition (cyan curves in Fig. 6a, c), the predictor can set the values at (0.75, 0.25) with a priority on the CW rotation based on instructions at the beginning of the first block. In the non-instructed condition (magenta curves), however, the predictor cannot contribute to the calculation of the values, and thus the values are (0.5, 0.5). At the beginning of the second block in the instructed condition, the predictor contributes to the calculation and the responsibility-signal values rapidly shift to nearly (0, 1). However, the change in the responsibility signal is slow in the non-instructed condition. The intermediate values of the responsibility signals, e.g., (0.7, 0.3), during the transition evokes two types of negative effects on switching and learning: (1) output signals from the CW internal model consist of 70% of the motor command, despite the fact that the subject should perform the CCW task in the second block. Thus error increases more than when a subject performs the CCW task in the first block (Fig. 6f). (2) Memory in the CW internal models is disrupted (the downward black

arrows) in proportion to the responsibility signal values. The degrees of the negative effects in the non-instructed condition are larger than those in the instructed condition because of the slow change in the responsibility signals.

In the above framework, the initial increase of error at block alterations (b) is caused by contamination of motor commands from an internal model used in the previous block (first type of negative effect). If explicit contextual information is given, the predictor can calculate appropriate responsibility signals immediately after block alteration. Thus, the initial increase of error at block alterations is small. Therefore, the b-value reflects the accuracy of the predictive switch based on explicit contextual information. Asymptotic level (a) within each block corresponds to the error when responsibility signals completely change to appropriate values for the current context and switching is completed. Therefore, the a-value reflects the performance of an internal model for each rotation. Adaptation speed (c) corresponds to the speed of the gradual change of responsibility signals based on sensorimotor feedback.

As mentioned in the introduction, the MOSAIC model predicts that the existence of explicit contextual information selectively improves the predictive switch immediately after alteration without affecting the gradual switch based on sensorimotor feedback. This prediction is consistent with our behavioral results indicating that the initial increase of error at block alterations (b) became small if explicit contextual information existed, but that adaptation speed (c) did not change (left panels of Fig. 5c–e).

In MOSAIC simulations, the time courses of the learning acquisition of internal models in the instructed condition (cyan curves in Fig. 6b–d) were always superior to the non-instructed condition (magenta curves), suggesting better performance of internal models in the instructed than in non-instructed conditions. This is consistent with our behavioral results indicating that asymptotic level (a) averaged across blocks was smaller in the instructed group than in the non-instructed group (right panel of Fig. 5c).

Regarding long-term performance changes, MOSAIC simulations indicated that performance of the internal models gradually improved across blocks, especially in the instructed condition (cyan curves in Fig. 6b and d), and that initial error at block alterations decreased across blocks in the instructed condition (Fig. 6e). Our behavioral data indicated decreases of both the asymptotic level and the initial increase of error at block alterations in the even blocks of the instructed group (dashed curves in right panels of Fig. 5c and d). These results suggest that long-term improvements of the internal models and the predictive switch occur if explicit contextual information is given. In simulations, performance of internal models also gradually improved in the non-instructed condition across blocks (magenta curves in Fig. 6b and d), but error at block

alterations did not decrease (Fig. 6f). These results correspond to behavioral results indicating that asymptotic level decreased to some degree (middle panel of Fig. 5c), but that initial increase of error at block alteration did not change from around 40° (middle panel of Fig. 5d).

Anterograde and retrograde interferences have been suggested in adaptation to opposing kinematics and/or dynamics. Anterograde interference is the negative influence of the preceding task on the subsequent task while retrograde interference is a consequence of the subsequent task disrupting memory in the preceding task (Robertson et al. 2004). We found similar interferences in the MOSAIC simulations. That is, the contamination of motor commands from the internal model used in the previous context (rightward arrows in Fig. 6) corresponds to anterograde interference. The partial destruction of internal models (degradation of $\hat{\theta}$) when they become modifiable states in inappropriate contexts (downward arrows) corresponds to retrograde interference. These interferences occur when the responsibility signal does not rapidly change in response to context changes.

Our simulation using the MOSAIC model could not duplicate the following behavioral results. First, the increase of errors at the beginning of each block in the instructed condition was small in the simulation (Fig. 6e) in comparison to the behavioral result (e.g., Fig. 5a). This suggests that the predictor in the simulation is more complete than that in the behavioral data, and the responsibility signal rapidly changed. Second, errors averaged within each block tended to be larger in the even blocks than in the odd blocks (Fig. 6e, f). However, the difference was very small in comparison to the behavioral data (Fig. 4a, b), and the simulation could not duplicate the difference in slopes of error reduction between the conditions (curves in Fig. 4a, b).

On questioning after the experiment, most of the non-instructed subjects did not realize the rotations, and attributed reasons of the error increase after the alterations to loss of skill or knack during the rest periods (see Supplementary Material). According to a previous study (Vetter and Wolpert 2000), when a virtual reality system shifted a perceived finger position gradually and alternately upward or downward from the actual position, subjects could adaptively track a moving target without awareness of the shift. These suggest that switching internal models in the non-instructed condition is based on sensorimotor feedback that can be processed at the subconscious level.

Although a systematic questioning was not done for the instructed subjects, some subjects reported that they tried to use the cognitive strategy to counteract the imposed rotation (see ''Methods'' section) but that the strategy often failed because they were required to start movements

immediately after a target appears. Consistently with this report, the trial-by-trial error of the instructed subjects sometimes abruptly increased and became noisy, and the number of detected outliers in the instructed group was larger than in the non-instructed group (e.g., Fig. 3a, b). We calculated correlation coefficients between the measured error in each trial and the error predicted by the exponential model in 100 cases (10 sessions × 10 subjects) and compared the coefficients between the instructed and non-instructed groups. The coefficients in the instructed group were significantly smaller than those in the non-instructed group ($t(98) = 6.97$, $P < 0.0001$), suggesting that the model for trial-by-trial readaptation based on sensorimotor feedback fits the error time courses of the instructed subjects less than those of the non-instructed subjects. The failure of the cognitive strategy and noisy performance in the instructed group is consistent with a study suggesting a conflict between the cognitive strategy and implicit learning (Mazzoni and Krakauer 2006).

The reduction of the error during the early stage of each block in the instructed condition was unlikely caused by intentional cancellation of the rotation for the following reasons. First, we could not find significant difference in parameter $b$ of the second block between the instructed and the non-instructed groups (Tukey's HSD post hoc test at $P < 0.05$ level). If the subject could reduce the initial error by using conscious cancellation, it would also be possible in the second block but this was not possible, suggesting that the interference could not be reduced solely by such a strategy. Second, Osu et al. (2004) also instructed subjects how to cancel the force fields by presenting figures illustrating effects of the fields. However, significant interference could be identified between adaptations to the opposing force field, suggesting the interference could not be reduced by only the conscious compensation. They also could not identify a significant progress of the conscious compensation. Thus, the main reason for the reduction in the error is probably the rapid change of the responsibility signal as our simulation with the MOSAIC model suggests.

Previous functional imaging studies (Imamizu et al. 2003, 2004) also supported the MOSAIC model and suggested that a prefrontal region (Brodmann area 46) contributes to the predictor while loops between the parietal regions (Pisella et al. 2000) and the cerebellum contribute to the estimator. The current behavioral study in the context of previous computational (Wolpert et al. 1995; Wolpert and Kawato 1998; Kawato 1999) and imaging studies (Imamizu et al. 2004) suggests that, to achieve adaptation to multiple environments, a predictive mechanism, possibly located in the frontal cortex, needs to switch multiple internal models residing in the cerebellum.

## References

Bock O, Schneider S, Bloomberg J (2001) Conditions for interference versus facilitation during sequential sensorimotor adaptation. Exp Brain Res 138:359–365

Brashers-Krug T, Shadmehr R, Bizzi E (1996) Consolidation in human motor memory. Nature 382:252–255

Chatterjee S, Price B (1991) Regression analysis by example. Wiley, New York

Cunningham HA, Welch RB (1994) Multiple concurrent visual-motor mappings: implications for models of adaptation. J Exp Psychol Hum Percept Perform 20:987–999

Gandolfo F, Mussa-Ivaldi FA, Bizzi E (1996) Motor learning by field approximation. Proc Natl Acad Sci USA 93:3843–3846

Gomi H, Kawato M (1993) Recognition of manipulated objects by motor learning with modular architecture networks. Neural Netw 6:485–497

Grubbs F (1969) Procedure for detecting outlying obserbations in samples. Technometrics 11:1–21

Imamizu H, Miyauchi S, Tamada T, Sasaki Y, Takino R, Putz B, Yoshioka T, Kawato M (2000) Human cerebellar activity reflecting an acquired internal model of a new tool. Nature 403:192–195

Imamizu H, Kuroda T, Miyauchi S, Yoshioka T, Kawato M (2003) Modular organization of internal models of tools in the human cerebellum. Proc Natl Acad Sci USA 100:5461–5466

Imamizu H, Kuroda T, Yoshioka T, Kawato M (2004) Functional magnetic resonance imaging examination of two modular architectures for switching multiple internal models. J Neurosci 24:1173–1181

Jacobs RA, Jordan MI, Nowlan SJ, Hinton GE (1991) Adaptive mixture of local experts. Neural Comput 3:79–87

Kawato M (1999) Internal models for motor control and trajectory planning. Curr Opin Neurobiol 9:718–727

Kawato M, Furukawa K, Suzuki R (1987) A hierarchical neural-network model for control and learning of voluntary movement. Biol Cybern 57:169–185

Krakauer JW, Ghilardi MF, Ghez C (1999) Independent learning of internal models for kinematic and dynamic control of reaching. Nat Neurosci 2:1026–1031

Lehmann EL (1975) Nonparametrics statistical methods based on ranks. McGraw-Hill, New York

Mazzoni P, Krakauer JW (2006) An implicit plan overrides an explicit strategy during visuomotor adaptation. J Neurosci 26:3642–3645

Miall RC, Jenkinson N, Kulkarni K (2004) Adaptation to rotated visual feedback: a re-examination of motor interference. Exp Brain Res 154:201–210

Osu R, Hirai S, Yoshioka T, Kawato M (2004) Random presentation enables subjects to adapt to two opposing forces on the hand. Nat Neurosci 7:111–112

Pisella L, Grea H, Tilikete C, Vighetto A, Desmurget M, Rode G, Boisson D, Rossetti Y (2000) An 'automatic pilot' for the hand in human posterior parietal cortex: toward reinterpreting optic ataxia. Nat Neurosci 3:729–736

Robertson EM, Pascual-Leone A, Miall RC (2004) Current concepts in procedural consolidation. Nat Rev Neurosci 5:576–582

Tong C, Wolpert DM, Flanagan JR (2002) Kinematics and dynamics are not represented independently in motor working memory: evidence from an interference study. J Neurosci 22:1108–1113

Vetter P, Wolpert DM (2000) The CNS updates its context estimate in the absence of feedback. Neuroreport 11:3783–3786

Wada Y, Kawabata Y, Kotosaka S, Yamamoto K, Kitazawa S, Kawato M (2003) Acquisition and contextual switching of

multiple internal models for different viscous force fields. Neurosci Res 46:319–331

Welch RB, Bridgeman B, Anand S, Browman KE (1993) Alternating prism exposure causes dual adaptation and generalization to a novel displacement. Percept Psychophys 54:195–204

Wolpert D, Kawato M (1998) Multiple paired forward and inverse models for motor control. Neural Netw 11:1317–1329

Wolpert DM, Ghahramani Z, Jordan MI (1995) An internal model for sensorimotor integration. Science 269:1880–1882