

Functional Magnetic Resonance Imaging Examination of Two Modular Architectures for Switching Multiple Internal Models

Hiroshi Imamizu,¹ Tomoe Kuroda,² Toshinori Yoshioka,¹ and Mitsuo Kawato¹

¹Advanced Telecommunications Research Institute Computational Neuroscience Laboratories, and ²Japan Science and Technology Corporation, Exploratory Research for Advanced Technology Kawato Dynamic Brain Project, Keihanna Science City, Kyoto 619-0288, Japan

An internal model is a neural mechanism that can mimic the input–output properties of a controlled object such as a tool. Recent research interests have moved on to how multiple internal models are learned and switched under a given context of behavior. Two representative computational models for task switching propose distinct neural mechanisms, thus predicting different brain activity patterns in the switching of internal models. In one model, called the mixture-of-experts architecture, switching is commanded by a single executive called a “gating network,” which is different from the internal models. In the other model, called the MOSAIC (*MOD*ular *S*election *A*nd *I*dentification for *C*ontrol), the internal models themselves play crucial roles in switching. Consequently, the mixture-of-experts model predicts that neural activities related to switching and internal models can be temporally and spatially segregated, whereas the MOSAIC model predicts that they are closely intermingled. Here, we directly examined the two predictions by analyzing functional magnetic resonance imaging activities during the switching of one common tool (an ordinary computer mouse) and two novel tools: a rotated mouse, the cursor of which appears in a rotated position, and a velocity mouse, the cursor velocity of which is proportional to the mouse position. The switching and internal model activities temporally and spatially overlapped each other in the cerebellum and in the parietal cortex, whereas the overlap was very small in the frontal cortex. These results suggest that switching mechanisms in the frontal cortex can be explained by the mixture-of-experts architecture, whereas those in the cerebellum and the parietal cortex are explained by the MOSAIC model.

Key words: task switching; internal model; computational models; functional magnetic resonance imaging; tool use; cerebro-cerebellar communication loop

Introduction

Many studies have investigated the neural correlates of task switching using a variety of paradigms, such as switching a stimulus–response correspondence (Dove et al., 2000) or changing a game rule “cognitive set” (Konishi et al., 1998; Nakahara et al., 2002). These studies have revealed the involvement of the prefrontal (Konishi et al., 1998; Nakahara et al., 2002) and the parietal (Dove et al., 2000) cortices and the cerebellum (Bischoff-Grethe et al., 2002). However, little is known about the global mechanisms underlying the switching.

Two representative computational theories have most thoroughly been investigated for task switching: a mixture-of-experts model for centralized switching (Jacobs and Jordan, 1991; Gomi and Kawato, 1993; Graybiel et al., 1994; Jordan and Jacobs, 1994) and the MOSAIC (*MOD*ular *S*election *A*nd *I*dentification for *C*ontrol) model for parallel-distributed switching (Kawato and Wolpert, 1998; Wolpert and Kawato, 1998; Wolpert et al., 1998;

Wolpert and Ghahramani, 2000; Haruno et al., 2001; Doya et al., 2002; Wolpert et al., 2003). In the mixture-of-experts model (Fig. 1A), outputs from function approximators known as expert modules are linearly combined by a classifier known as a gating module. They are trained so as to split the input data into subparts in which particular experts are specialized. The gating module weights the contribution of the output of each expert module to the final output according to its estimation of the probability that each expert is appropriate for the current context. Thus, the selection is centralized in the gating module and segregated from the expert modules. A mixture-of-experts architecture predicts that switching activity is segregated from an activity related to expert modules.

In contrast, in the MOSAIC (Fig. 1B), internal models that are experts at mimicking the input–output properties of controlled objects (Kawato et al., 1987; Miall et al., 1993; Shadmehr and Mussa-Ivaldi, 1994; Wolpert et al., 1995; Flanagan and Wing, 1997; Kawato, 1999) crucially contribute to selection. Multiple pairs of forward internal models (predictors) and inverse internal models (controllers) are tightly coupled as functional units. For example, when we use a new pair of scissors, forward models (Blakemore et al., 1998; Iacoboni, 2001; Miall et al., 2001) of various types of scissors simultaneously predict sensory feedback

Received Aug. 31, 2003; revised Oct. 26, 2003; accepted Nov. 21, 2003.

This work was supported in part by the Telecommunications Advancement Organization of Japan.

Correspondence should be addressed to Dr. Hiroshi Imamizu, Keihanna Science City, 2-2-2 Hikaridai, Kyoto 619-0288, Japan. E-mail: imamizu@atr.co.jp.

DOI:10.1523/JNEUROSCI.4011-03.2004

Copyright © 2004 Society for Neuroscience 0270-6474/04/241173-09\$15.00/0

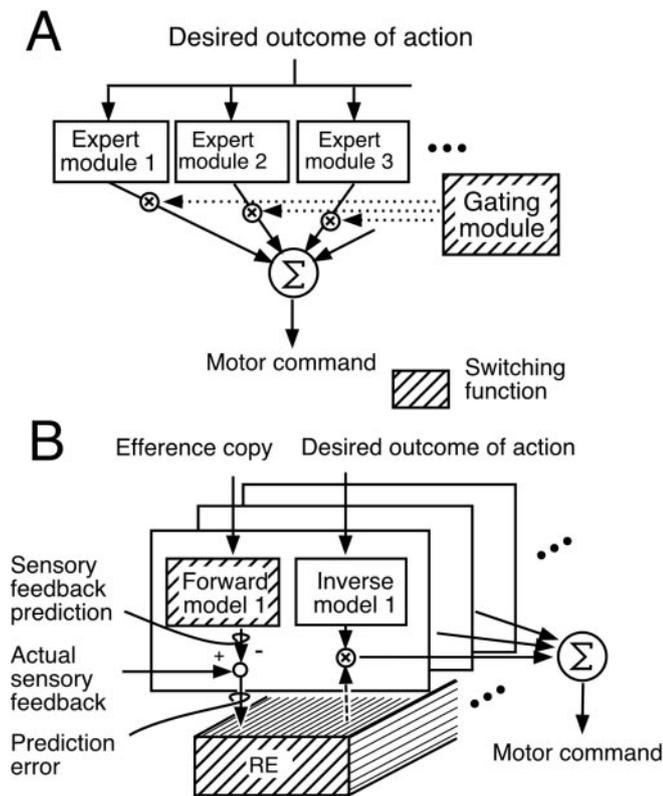


Figure 1. Two representative computational methods for switching internal models. *A*, Mixture-of-experts architecture: a gating module estimates the current context and decides the contribution of each expert (internal model) independently of the activity of the experts. Switching functions (hatched parts) are concentrated in the gating module. *B*, MOSAIC architecture: a forward model and an inverse model that are tightly coupled, and the RE decides the contribution of each pair of internal models according to the goodness of prediction made by the forward models. Internal models contribute to the switching functions.

from an efference copy of motor commands. The prediction of each forward model is then compared with actual feedback. The smaller the error, the more likely that the forward model must have been an appropriate predictor in the current context. The inverse model paired with the appropriate predictor is considered as an appropriate controller. Thus, the selection mechanism crucially depends on the internal models (i.e., forward models must be active when switching internal models). Consequently, the MOSAIC predicts that the switch activity spatially and temporally overlaps the internal model activity. Our objective is to examine the predictions of the two architectures investigating brain activity related to task switching for novel tools. Our previous study (Imamizu et al., 2003) found modular organization of the internal models in the cerebellum, but the cerebro-cerebellar neural mechanism for the switching of the internal models is unknown.

Materials and Methods

Experimental design. Our previous studies demonstrated that activity reflecting an internal model of a novel tool (a rotated mouse, the cursor of which appears in a rotated position) increases in the cerebellum after learning use of the tool (Imamizu et al., 2000) and that the internal models send output to the premotor (PM) regions (Tamada et al., 1999). We also investigated cerebellar activity after learning use of two novel tools (the rotated mouse and a velocity mouse, the cursor velocity of which is proportional to the mouse position). Activities of the two different tools were segregated spatially with a very small overlap, suggesting that the multiple internal models are acquired in a modular manner

(Imamizu et al., 2003). Based on these experimental paradigms, we investigated brain activity related to switching internal models for novel tools and examined the predictions of the two architectures.

Subjects manipulated the rotated, the velocity, or the normal mouse to track a randomly moving target on a screen with a cursor (a tracking task). The x and y components of the target path were each sums of sinusoids, the amplitude and frequency of which were determined randomly. After intensive training in using the novel mice, brain activity was scanned when the subjects conducted the tracking task using one of the three mice. The subjects manipulated the same type of mouse within a time block, the duration of which was determined randomly in a range of 26.5–42.4 sec. Four types of discontinuous transitions randomly split these blocks: (1) the mouse type changed, and cognitive cues were presented; (2) the mouse type did not change, and the cues were presented; (3) the mouse type changed, and the cues were not presented; and (4) the mouse type did not change, and the cues were not presented. The cognitive cues were cursor colors corresponding to the mouse type (red, violet, or green) and text presented on the screen (“rotate,” “velocity,” or “normal”). At every transition between blocks, the cursor was reset to the center of the screen while the target continued to move on a smooth path.

In our previous study (Imamizu et al., 2003), our purpose was to investigate regional difference in the cerebellar activity related to the internal models. Thus, the experimental design was simple (i.e., the type of mouse changed at a constant interval), and there was only one type of transition between the blocks (the mouse type changed, and cognitive cues were presented). However, in the current study, the experimental design was modified to investigate the switching mechanism of the internal models. The block was determined randomly so that the subjects could not anticipate when the mouse type would change. There were four types of transitions between blocks to dissociate various factors related to the switching (e.g., the mouse-type change, the cognitive cue, and the reset of the cursor). We scanned the whole brain to investigate the global switching mechanism.

Subjects. Ten neurologically normal subjects (21–39 years of age; two females and eight males) participated in the experiments. All participants were right-handed (Oldfield, 1971). Informed written consent was obtained from each subject. The protocol was approved by the ethics committee of Advanced Telecommunications Research Institute.

Tasks. A projector displayed the target and the cursor. A small white square target was presented on a dark background. The target moved within a square area, subtending horizontal and vertical visual angles of 23.36° and 19.71°, respectively. The subjects moved a small cross-hair cursor on the screen with the mouse.

The relationship between the cursor position and the mouse position for the rotated mouse was:

$$\begin{pmatrix} x_c \\ y_c \end{pmatrix} = \begin{pmatrix} \cos 120^\circ & \sin 120^\circ \\ -\sin 120^\circ & \cos 120^\circ \end{pmatrix} \begin{pmatrix} x_m \\ y_m \end{pmatrix}, \quad (1)$$

where (x_c, y_c) denotes the screen coordinates of the cursor (visual angle, °), and (x_m, y_m) denotes the mouse/hand coordinates (centimeters). In contrast, the velocity of the cursor (\dot{x}_c, \dot{y}_c) (degrees per second) was determined by the mouse position for the velocity mouse:

$$\begin{pmatrix} \dot{x}_c \\ \dot{y}_c \end{pmatrix} = 8.76 \begin{pmatrix} x_m \\ y_m \end{pmatrix}. \quad (2)$$

During the training sessions over the 5 d, brain activity was not scanned, but the subjects performed the task lying on a bed as they would in a magnetic resonance imaging (MRI) scanner. Two subjects did not participate in the training sessions because they had already received intensive training for using the novel mice in our other experiment (Imamizu et al., 2003). Each session lasted 5.63 min (days 1 and 2) or 14.22 min (days 3–5), and the mouse type did not change during the session. The subjects used the normal mouse in the first session of each morning and then were trained for only one of the two novel mice in the subsequent six sessions (days 1 and 2) or three sessions (days 3–5). Breaks between sessions lasted 5–10 min. After a 1–2 hr lunch break, the subjects used the normal mouse in the first session of each afternoon and then were trained

for the other novel mouse with the same number of sessions as in the morning.

All subjects underwent six functional MRI (fMRI) scanning sessions after the training sessions, with each session lasting 14.22 min. In each session, the rotated mouse, the velocity mouse, and the normal mouse changed in a pseudo-random order. A cognitive cue was presented in every other session. The length of the time block in which the subjects continuously used the same type of mouse was pseudo-randomly determined in a range of 26.5–42.4 sec (five to eight fMRI scanning intervals).

Analysis of behavioral data. The cursor (x_c, y_c), the target (x_t, y_t), and the mouse (x_m, y_m) positions were sampled at 1 kHz. The distance between the cursor and the target at each sampling point:

$$(\sqrt{(x_t - x_c)^2 + (y_t - y_c)^2}), \tag{3}$$

was accumulated over 5.3 sec (position tracking error). A velocity tracking error:

$$(\sqrt{(\dot{x}_t - \dot{x}_c)^2 + (\dot{y}_t - \dot{y}_c)^2}), \tag{4}$$

was also accumulated over 5.3 sec. Because results obtained from the velocity error were similar to those from the position error, we only present the position tracking error in this article.

MRI acquisition. A 1.5 T MRI scanner (Shimadzu-Marconi) was used to obtain blood oxygen level-dependent (BOLD) contrast functional images. Images weighted with the apparent transverse relaxation time were obtained with an echoplanar imaging sequence (repetition time, 5.3 sec; echo time, 65 msec; flip angle, 90°). One hundred sixty-three sequential whole-brain volumes ($64 \times 64 \times 44$ voxels at 3.5 mm isotropic resolution) were acquired in each session. High-resolution anatomical images of all subjects were also acquired with a T1-weighted sequence.

MRI analysis. We used SPM99 software (<http://www.fil.ion.ucl.ac.uk/spm/>) for image processing and analysis. The first two volumes of images were discarded to allow for T1 equilibration, whereas the remaining 161 image volumes were realigned to the first volume and normalized to the Montreal Neurological Institute (MNI; Montreal, Canada) reference brain. The data were smoothed spatially with a Gaussian kernel with a 7 mm full-width at half maximum (FWHM). The voxel time series was smoothed temporally with a Gaussian filter (FWHM of 4 sec).

A model for investigating activity related to switching of mouse type. We conducted a multiple regression analysis to find regions related to the switching of the mouse type. Three factors were included as main explanatory variables:

$$S_i^k = \alpha_i x^k + \beta_i y^k + \gamma_i z^k + e_i. \tag{5}$$

Here, S_i^k denotes the fMRI signal at the i -th voxel in the k -th scan, and x , y , and z are explanatory variables representing the switching of the mouse type, the presentation of the cognitive cues for switching, and the cursor reset, respectively. Each variable was a pulse function assigned 1 in the scan immediately after each event occurred and 0 otherwise and convolved with a canonical hemodynamic function as implemented in SPM99. The four transition types were represented by combinations of the explanatory variables (x, y, z): (1) the mouse type changed, and cognitive cues were presented: (1, 1, 1); (2) the mouse type did not change, and the cues were presented: (0, 1, 1); (3) the mouse type changed, and the cues were not presented: (1, 0, 1); and (4) the mouse type did not change, and the cues were not presented: (0, 0, 1). This analysis allows us to distinguish activity related to the switching of the mouse type from activities related to the cognitive cues and behavioral changes that were evoked by the cursor reset (e.g., transient increase of the tracking error, attention, effort levels, and eye/hand movements).

We performed a random effect analysis as follows. Images of parameter estimates for the contrast of interest (α, β , and γ) were created for each subject (first-level analysis) and were then entered into a second-level analysis to test whether each parameter was significantly larger than zero, using a one-sample t test across subjects ($t_{(9)} > 4.3$; $p < 0.001$ uncorrected; cluster size, > 50 voxels).

In an event-related analysis of activity related to the switching of the mouse type (as shown in Figs. 5 and 6), we investigated time courses of

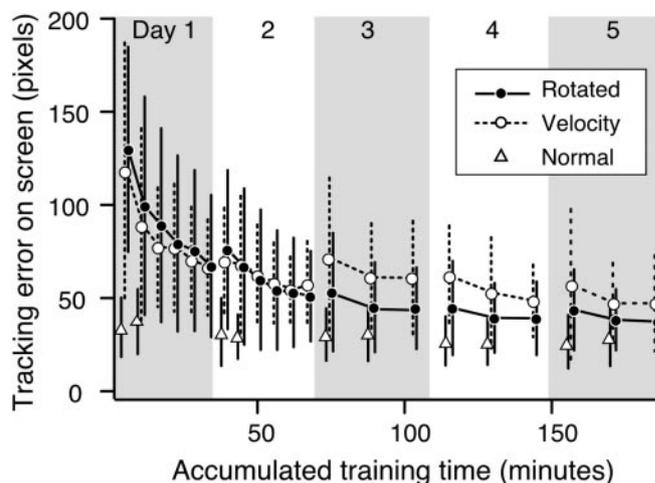


Figure 2. Tracking errors (across-subjects; mean \pm SD) when subjects manipulate the rotated mouse (solid lines with filled circles), velocity mouse (broken lines with open circles), and normal mouse (triangles) in training sessions.

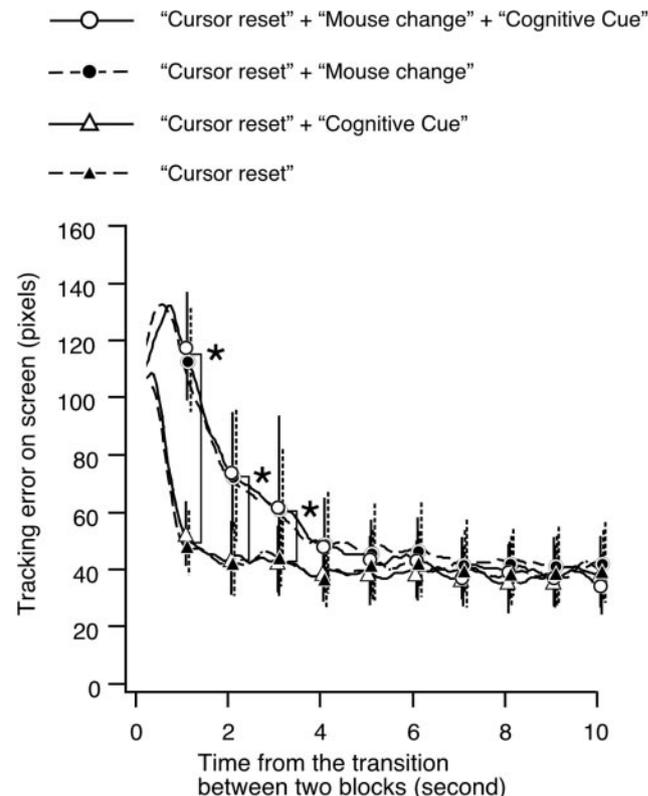


Figure 3. Time courses of errors during scanning sessions for four types of transition between blocks. The time courses were aligned on the transition. The markers (circles and triangles) and bars indicate across-subjects mean \pm SD at every second. The asterisk indicates a significant difference in the errors between when the mouse type was changed and when the type was not changed ($p < 0.05$; multiple comparisons using Student's t test).

the relative BOLD signal that is expressed as a percentage of signal increase at each voxel from the averaged signal across the whole brain and a session.

A model for investigating spatial overlap between activity related to behavioral switch and activity related to internal model. We conducted the second regression analysis to find regions in which behavioral switch

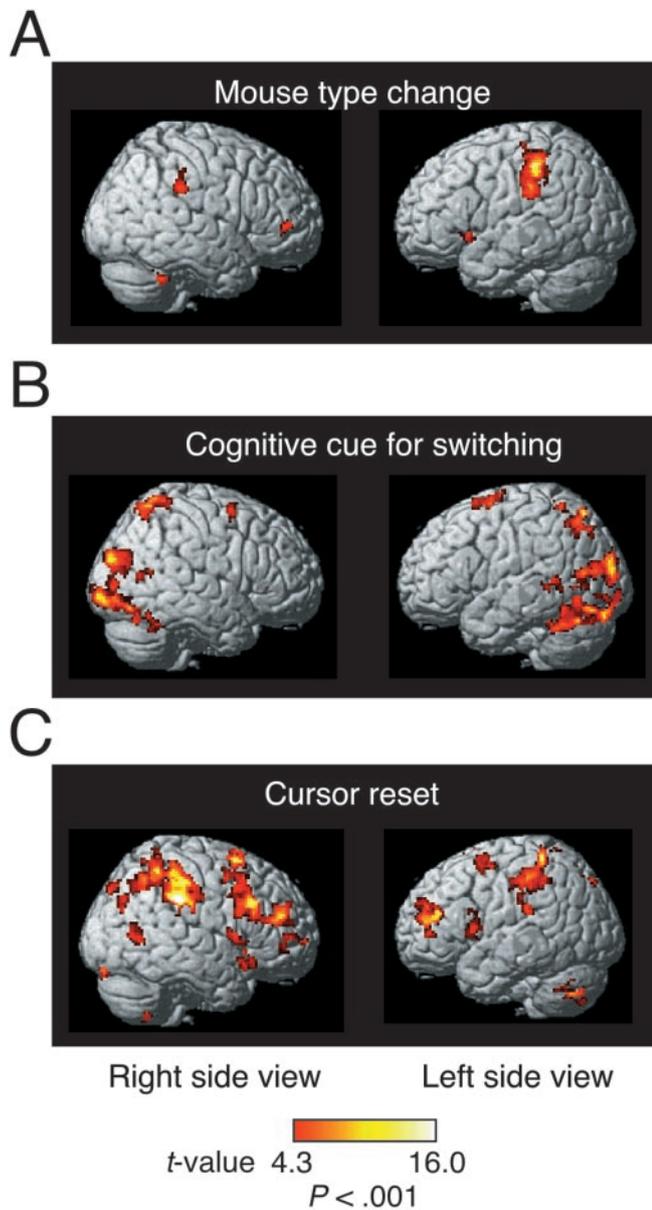


Figure 4. Regions related to the switching of the mouse type (A), cognitive cues (B), and cursor reset (C) (random effect model; $t_{(9)} > 4.3$; $p < 0.001$ uncorrected; cluster size, > 50 voxels).

activity spatially overlaps the internal model activity. In this regression analysis, the linear model was:

$$S_i^k = \delta_i s^k + \varepsilon_i v^k + \zeta_i n^k + \eta_i n^k + e_i, \quad (6)$$

where s is an explanatory variable representing the behavioral switch. It was assigned 1 in the scans immediately after the switch of the mouse type, the cognitive cue presentation, or the cursor reset or 0 in the other scans. In the previous analysis, we investigated activity evoked by mouse-type switching, whereas, in the current analysis, we investigated activity related to the behavioral switch that was evoked by the discontinuous transitions between blocks (the switch of the mouse type, the cognitive cue presentation, or the cursor reset). We also included three explanatory variables (r , v , and n) representing the mouse type to investigate internal model activity. They correspond to the rotated mouse, the velocity mouse, and the normal mouse and were assigned 1 if the scan corresponded to its mouse type and 0 otherwise (assigned 0 in the first scan after the mouse type changed). Contrasts of interest in a first-level anal-

Table 1. Increment of activation

Area	Coordinates	T value
Switching of mouse type		
L inferior parietal gyrus	(−58, −30, 50)	14.90
L insula	(−32, 18, −4)	8.16
R supramarginal gyrus	(64, −34, 32)	7.51
R cerebellum (Crus 1)	(38, −48, −36)	7.43
R middle frontal gyrus	(44, 48, 4)	6.45
Cognitive cues		
R precuneus	(10, −52, 72)	11.65
R vermis 6	(10, −72, −14)	11.60
L cerebellum (Crus 1)	(−26, −86, −22)	10.21
L superior frontal gyrus	(−18, 0, 70)	10.02
R calcarine fissure and surrounding cortex	(2, −96, 0)	9.56
R superior occipital gyrus	(30, −70, 26)	9.08
L middle occipital gyrus	(−22, −90, 14)	8.44
L superior parietal gyrus	(−28, −66, 48)	7.88
L middle temporal gyrus	(−52, −50, 8)	7.24
L middle temporal gyrus	(−42, −68, 12)	6.55
R middle temporal gyrus	(46, −58, 14)	6.53
L superior parietal gyrus	(−30, −52, 64)	6.09
Cursor reset		
R supramarginal gyrus	(55, −32, 34)	14.26
L supramarginal gyrus	(−62, −32, 22)	11.66
R precuneus	(8, −44, 54)	10.95
L middle frontal gyrus	(−30, 36, 20)	10.49
R inferior frontal gyrus, triangular part	(42, 24, 24)	10.37
R lingual gyrus	(22, −88, −20)	9.16
R superior frontal gyrus	(20, 14, 64)	9.09
R calcarine fissure and surrounding cortex	(16, −78, 8)	7.97
L supramarginal gyrus	(−14, −72, 44)	7.89
R middle temporal gyrus	(52, −68, 12)	7.77
L middle frontal gyrus	(−28, 4, 54)	7.76
R middle frontal gyrus	(34, 62, 2)	7.71
R middle occipital gyrus	(42, −74, 32)	7.61
R precentral gyrus	(30, −8, 50)	7.57
R supplementary motor area	(2, 4, 58)	7.32
L inferior frontal gyrus, opercular part	(−56, 10, 10)	7.26
L cerebellum 7b	(−26, −66, −46)	6.74
R rolandic operculum	(56, 6, 6)	6.43
R putamen	(28, 16, −8)	6.31
R cerebellum 8	(30, −56, −52)	5.92
R superior temporal gyrus	(50, 16, −18)	4.89

Activation was thresholded at $t_{(9)} > 4.3$ ($p < 0.001$ uncorrected) and cluster size at > 50 voxels. A random effect model was used. L, Left; R, right.

ysis were δ , $(\varepsilon - \eta)$ and $(\zeta - \eta)$ for each subject. In a second-level analysis, we tested whether an estimated parameter for the behavioral switch was significantly larger than zero ($\delta > 0$) and whether a parameter for the rotated mouse or the velocity mouse was larger than a parameter for the normal mouse ($\varepsilon - \eta > 0$ or $\zeta - \eta > 0$), using a one-sample t test across subjects ($t_{(9)} > 4.3$; $p < 0.001$ uncorrected; cluster size, > 50 voxels). Thus, regions related to the rotated mouse (or the velocity mouse) are those significantly more activated during use of the rotated mouse (or the velocity mouse) than the normal mouse.

We searched for regions in which the behavioral switch activity overlaps with activity related to the rotated mouse or the velocity mouse and defined the anatomical volume(s) of interest (VOI), including the overlapping regions. Regarding the cerebral cortex, the VOI was defined according to Tzourio-Mazoyer et al. (2002). The cerebellum was divided into six VOI: medial volume, including the vermis and the intermediate parts ($-20 < x < 20$ in MNI coordinates), and the remaining two lateral volumes were each divided into the superior and the inferior volumes by a horizontal plane ($z = 40$) that roughly corresponds to the horizontal fissure. We calculated the percentage of overlapping volume to the behavioral switch activity within each VOI.

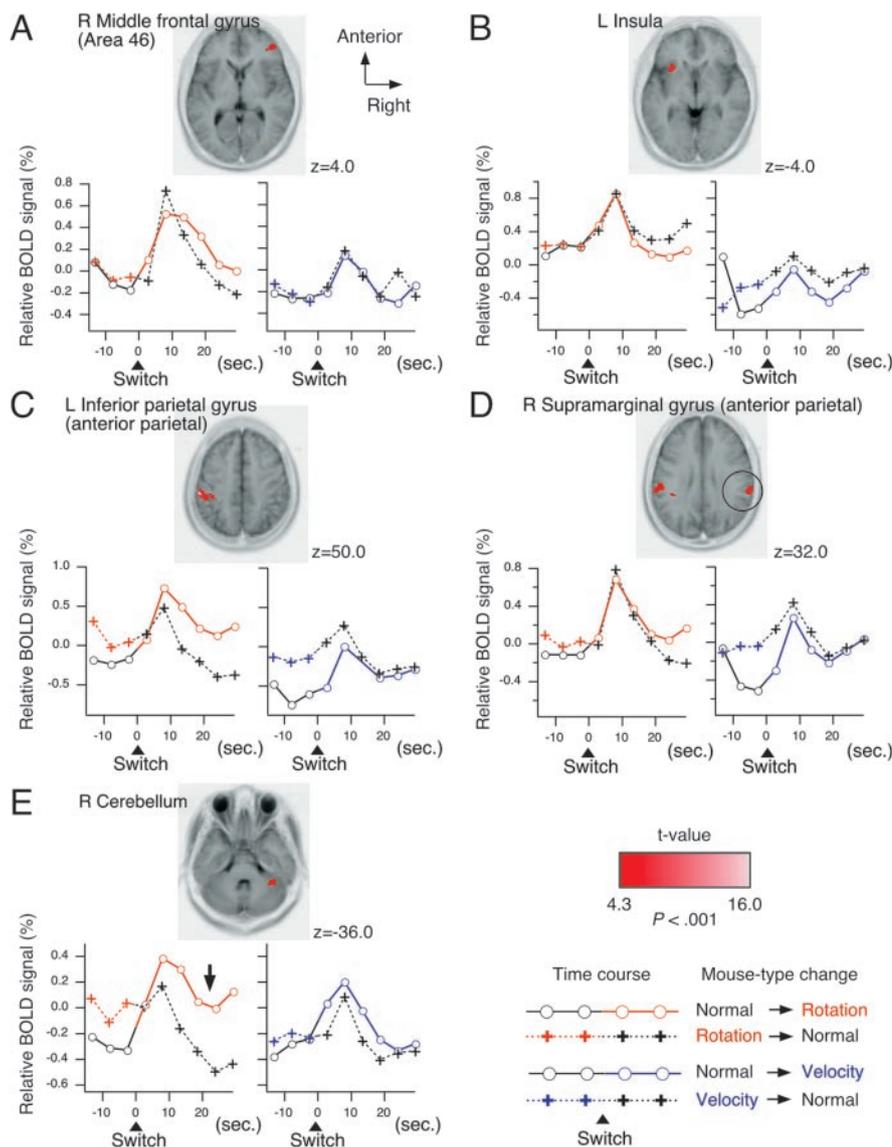


Figure 5. Time courses of activation (relative BOLD signal; see Materials and Methods) in regions related to the switching of the mouse type. The activated regions are the same as those in Figure 4A (random effect model; $t_{(9)} > 4.3$; $p < 0.001$ uncorrected; cluster size, > 50 voxels) but shown in transverse sections. The time courses were averaged across repetitions, subjects, and voxels in each region. The abscissa represents time from the switching of mouse type in seconds. The color of lines and markers indicates the mouse type being used at each time point (orange, rotation; blue, velocity; black, normal). As illustrated in the bottom left corner, the solid lines with open circles indicate time courses when the mouse type changed from the normal mouse to the rotated mouse (the left figure in each section) or velocity mouse (the right figure). The broken lines with crosses indicate those when the mouse type changed from the rotated mouse (the left figure) or velocity mouse (the right figure) to the normal mouse. The time course marked with an arrow (E) will be used as an example in Figure 6.

Results

Behavioral data

As shown in Figure 2, the tracking errors for the novel mice decreased as the number of training sessions increased. A repeated-measures ANOVA on the errors indicated the significant effect of sessions for the rotated mouse ($F_{(20,140)} = 13.19$; $p < 0.0001$) and for the velocity mouse ($F_{(20,140)} = 12.07$; $p < 0.0001$), suggesting that learning had occurred.

Figure 3 shows time courses of the tracking errors separately for the four types of transitions between blocks. The time courses were aligned on the transition point. There was no significant difference ($F_{(1,9)} = 0.81$) in the averaged tracking errors between the cognitive cue conditions (solid curves) and the no-cognitive

cue conditions (broken curves), and, thus, the effects of the cue presentation could not be identified, at least in the behavioral data. In contrast, the averaged error when the mouse type changed (curves with circles) was significantly larger ($F_{(1,9)} = 30.76$; $p < 0.0001$) than the error when the mouse type did not change (curves with triangles). We examined the difference at every second (1.0, 2.0, . . . and 10.0 sec) and found a significant difference between the error time courses when the mouse type changed and those when the mouse type did not change within 3 sec after the transition ($p < 0.05$; multiple comparisons using Student’s *t* test). Although a statistically significant difference was identified only within 3 sec, there remained a marked difference between the solid and the broken curves until 6 sec. Even if a delay between the switching of internal models and the change in performance is taken into account, the difference in the tracking errors suggests that the switching of internal models is a time-consuming process requiring more than a few seconds.

Activity related to switching of mouse type

In the first analysis of brain activity, we conducted a multiple regression analysis to find regions related to the switching of the mouse type, cognitive cues for switching, and cursor reset (see Materials and Methods). To specify the activation related to the three factors, we searched regions in which the estimated parameters for explanatory variables were significantly larger than zero according to *t* statistics (random effect model; $t_{(9)} > 4.3$; $p < 0.001$ uncorrected; cluster size, > 50 voxels). Figure 4 shows the regions, and Table 1 lists the coordinates and *t* value of an activation peak in each region. Activity related to the switching of the mouse type was observed in the anterior parietal (AP) regions (left inferior parietal gyrus and right supramarginal gyrus), the left insula, the right cerebellum (anterior part of the lateral cerebellum), and the right prefrontal region (Brodmann area 46 in the middle frontal gyrus). Activity related to the cognitive cues was observed in regions associated with visual information processing (i.e., the occipital, the temporal, and the superior parietal regions). Activity related to the cursor reset was observed in various regions associated with relatively higher-order sensory–motor functions (i.e., the parietal, the PM, and the prefrontal regions).

Figure 5 shows time courses of activation (relative BOLD signal; see Materials and Methods) in the regions related to the switching of the mouse type (see also Fig. 4A). The time courses were aligned with the switching and averaged across voxels in each region and subject. The color of the lines and markers indicates the mouse type being used at each time point (normal,

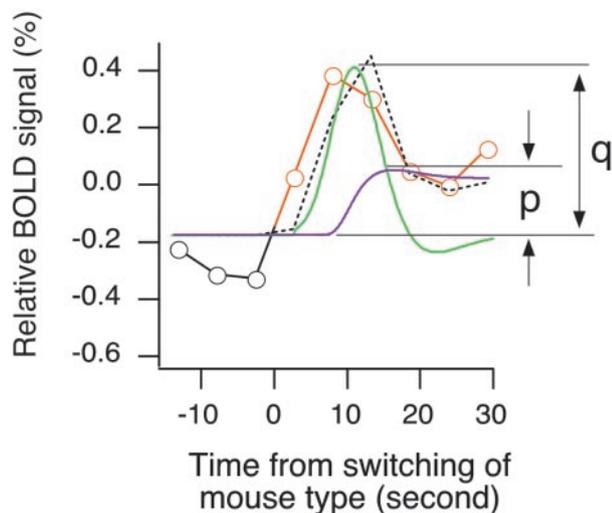


Figure 6. An example for quantitative analysis of the activation time courses to separate and compare the sustained component and transient component. A solid line with open circles indicates one of the time courses observed in the right cerebellum (marked with an arrow in Fig. 5E). The time course was fitted using a general linear model by a weighted sum of three components: a step function modeling the sustained activity (magenta curve); a pulse function modeling the transient activity (green curve); and a constant component. The height of the curve peaks from the baseline (p and q) indicates the estimated weights of the sustained component and the transient component. A broken curve represents the summation of the three estimated components (a fitted time course).

black; orange, rotation; blue, velocity). The bottom left figure in each panel indicates the time courses at transition between the normal and the rotation mouse, whereas the bottom right figure indicates those at transition between the normal and the velocity mouse. As shown at the bottom right corner of Figure 5, the solid lines with open circles indicate the time courses when the mouse type changed from the normal (black) to the rotation (orange) or velocity (blue). The broken lines with crosses indicate those when the mouse type changed from the rotated or velocity mouse to the normal.

The time courses in area 46 (Fig. 5A) and the insula (Fig. 5B) indicate that the activation transiently increased immediately after the switching, whereas the levels of sustained background activity >10 – 15 sec after the switching were almost the same as those before the switching. In contrast, in the cerebellum (Fig. 5E) and occasionally in the parietal regions (Fig. 5C), we could observe not only a transient increase of activity but also a change in the sustained activity level. For example, in the bottom left panel of Figure 5E, the level of activation indicated by the solid line and open circles was low before the switching when the subjects used the normal mouse (black). It transiently increased immediately after the switching and then remained high as long as the subjects used the rotated mouse (orange). A time course indicated by a broken line and crosses presented the inverse temporal pattern: the high-level-sustained activity became low after the switching. Similar patterns were observed in the bottom right panel of Figure 5E. The colored parts of the lines (orange or blue) indicate activity when the subjects used the rotated or velocity mouse. Only in the cerebellum were the colored parts always positioned above the black parts, suggesting that activity specific to the novel mice was the most prominent in the cerebellum among the regions related to the switching of the mouse type. In light of our previous studies (Tamada et al., 1999; Imamizu et al., 2000, 2003), this activity most probably reflects internal models for the novel mice.

Table 2. Overlap between regions related to behavioral switch and those related to the rotated mouse and the velocity mouse

Area	Voxels	Coordinates	t value
R cerebellum	161	(46, -58 , -34)	9.23
L cerebellum	17	(-34 , -50 , -40)	5.34
R dorsal premotor region (superior frontal gyrus)	40	(14, -10 , 66)	6.08
R parietal region (supramarginal gyrus and inferior parietal gyrus)	96	(60, -30 , 32)	6.07

Activation was thresholded at $t_{(p)} > 4.3$ ($p < 0.001$ uncorrected) and cluster size at >50 voxels. A random effect model was used. R, Right; L, left.

Temporal overlap between activity related to switching of mouse type and activity related to internal models

We conducted quantitative analysis of the event-related time courses to separate and compare the transient component related to the switching and the sustained component related to the novel mice. As illustrated in Figure 6, each time course was fitted by a weighted sum of three components: (1) a pulse function modeling the transient activity that was assigned a value of 1 only in the scan immediately after the switching of the mouse type and 0 otherwise (a green curve indicating the weighted one); (2) a step function modeling the sustained activity, the value of which changed from 0 to 1 (from the normal mouse to the novel mouse) or from 1 to 0 (from the novel mouse to the normal mouse) after the switching (a magenta curve); and (3) a constant component. The onset–offset of the step function was not set at the middle of the first scan ($5.3/2 = 2.65$ sec after the switching) but at the second scan ($5.3 + 5.3/2 = 7.95$ sec), because the behavioral data suggested that the switching of the internal models takes >3 sec.

The fitting was also conducted by a linear regression analysis. The linear model was:

$$w_t = ap_t + bq_t + c. \quad (7)$$

Time t ranged from -13.5 to 29.5 sec at intervals of 5.3 sec. The mouse type changed at $t = 0$. Here, w is the fMRI signal averaged across repetitions, subjects, and voxels in each region. If a transition between blocks occurred <20 sec before, data were excluded to avoid the effects of the previous transition. p is the step function, and q is the pulse function. Both functions were convolved with the canonical hemodynamic function. c is a constant term.

We calculated the ratio of the magnitude of the sustained component (p) to that of the transient component (q) for each region. The magnitude was defined as the difference between the peak and the baseline of the weighted waveform, as shown in Figure 6. The ratios (p/q) were obtained from time courses at four types of mouse transitions (rotated–normal, normal–rotated, velocity–normal, and normal–velocity), and they were averaged. The averaged ratio was 0.11 in the right prefrontal region (area 46), 0.21 in the left insula, 0.55 in the left inferior parietal gyrus, 0.22 in the right supramarginal gyrus, and 0.63 in the cerebellum. Thus, the sustained component reflecting the activity of internal models was most prominent in the cerebellum among the regions related to the switching.

Spatial overlap between activity related to behavioral switch and activity related to internal model

We conducted the second regression analysis to find regions in which the behavioral switch activity spatially overlaps the internal model activity (see Materials and Methods). As the switching activity, we investigated activity related to the behavioral switch that was evoked by the discontinuous transitions between blocks

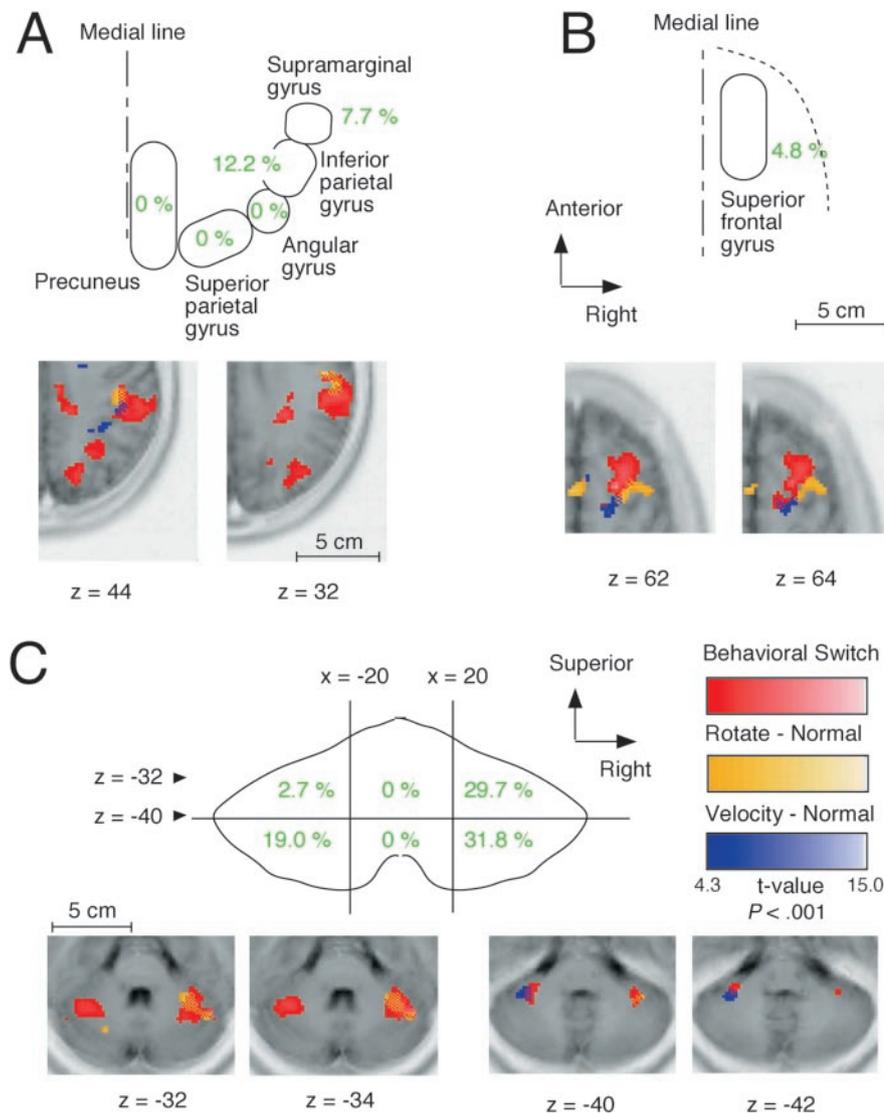


Figure 7. Activations related to the behavioral switch (red), rotated mouse (orange), and velocity mouse (blue) in the parietal regions (A), the PM regions (B), and the cerebellum (C) (random effect model; $t_{(9)} > 4.3$; $p < 0.001$ uncorrected; cluster size, > 50 voxels). The illustrations above the activation maps indicate the VOI and the percentage of overlap between the switching-related volume and the volume related to the rotated or velocity mouse (mosaic of red and orange, or red and blue) to the switching-related volume (red) within each VOI. The illustrations in A and B show the horizontal sections, whereas the illustration in C shows the coronal section. All activation maps show transverse sections.

(the switch of the mouse type, the cognitive cue presentation, or the cursor reset). The overlap was found in the anterior parts of the lateral cerebellum, the right AP regions, and the right dorsal PM cortex in the superior frontal gyrus (random effect model; $t_{(9)} > 4.3$; $p < 0.001$ uncorrected; cluster size, > 50 voxels). Here, we emphasize that no spatial overlap was found at this threshold in the right prefrontal region, the left insula, or the left inferior parietal gyrus, which was activated by the switch of the mouse type. Table 2 lists the volume of the overlapping regions, activation peaks, and peak t values in the regions. Figure 7 shows activity related to the behavioral switch (red), the rotated mouse (orange), and the velocity mouse (blue) in the parietal regions (Fig. 7A), the PM regions (Fig. 7B), and the cerebellum (Fig. 7C).

We defined the anatomical VOI, including the overlapping regions, and investigated the percentage of the overlapping volume (mosaic of red and orange or red and blue) to the switching-related volume (red) for each VOI, as described in Materials and

Methods. As indicated in Figure 7, the percentages in the cerebral regions ranged between 4.8 and 12.2%, whereas the maximum percentage was 31.8% in the cerebellum. Thus, the switching activity more markedly overlapped the internal model activity in the cerebellum than in the cerebral cortex. The averaged volume of the VOI was 1.94 cm³ in the cerebral regions, whereas it was 3.15 cm³ in the cerebellum.

We conducted a volumetric analysis on the activation map of each subject (fixed effect model; $p < 0.05$ corrected) and obtained consistent results; that is, the overlap was large in the PM region (left, 15.6%; right, 16.7%) and the cerebellum (maximum percentage, 16.1% in the middle superior cerebellum) and was small in area 46 (left, 0.7%; right, 1.9%) and the insula (left, 2.8%; right, 3.0%). A spatial filter was not applied to preserve the resolution of the raw images in the individual analysis (see Discussion).

Discussion

Activity associated with the mouse-type change was found in area 46, the insula, the AP regions, and the anterior part of the lateral cerebellum. According to the event-related analysis, the ratio of the sustained internal model activity to the transient switching-related activity was large in the cerebellum and the AP region, whereas it was small in area 46 and the insula. In accordance with this temporal overlap analysis, spatial overlap between the behavioral switch activity and the internal model activity was not observed in area 46 or the insula but observed in the PM, the AP regions and the cerebellum. Furthermore, the largest overlapping volume was found in the cerebellum. Based on the predictions of the mixture-of-expert and MOSAIC architectures, the current results can be interpreted within the two different schemata as follows.

One of schema assumes, as shown in Figure 8A, that the internal models are located in the cerebellum, whereas the central executives of the MOSAIC are located among the cerebral regions. The responsibility estimator (RE) that decides the responsibility of each module for the current context may consist of several components. The components closely related to internal models (the blue and red part of the RE) are in the premotor (PM) and the AP regions, where the sustained activity was observed, whereas the components conducting the final computation of the responsibility (“genuine” switch functions; the red part) are in the insula and area 46, where the transient activity was dominant. This model provides one concrete computational mechanism for prefronto–parieto–cerebellar connectivity revealed by electrophysiological studies (Sasaki et al., 1977) and transneuronal–tracer studies (Middleton and Strick, 1994; Clower et al., 2001; Middleton and Strick, 2001; Dum and Strick,

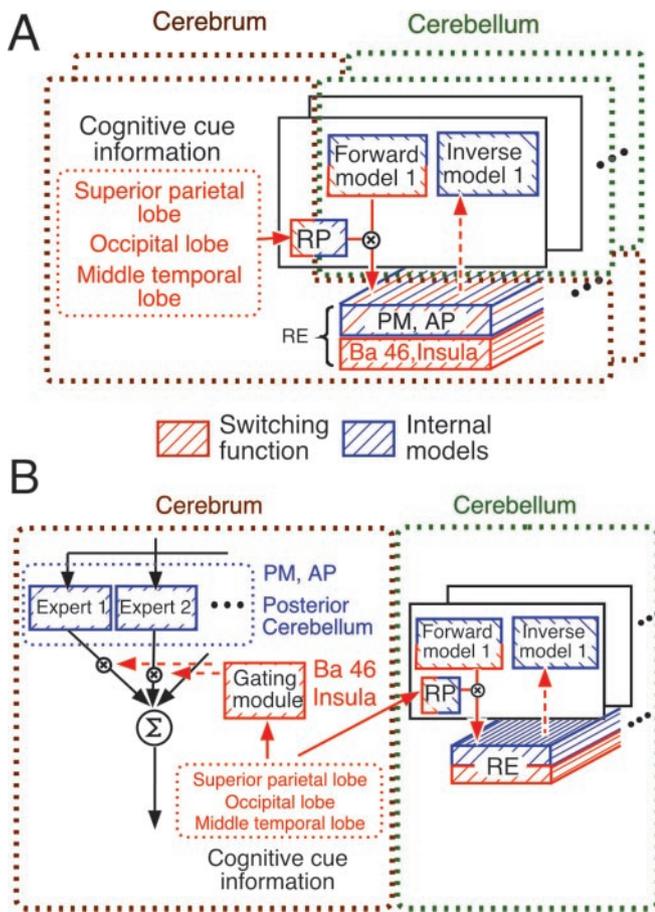


Figure 8. Two possible schemata regarding the switch mechanisms of internal models implicated by the current results. *A* assumes that the switching is conducted only by the MOSAIC architecture. *B* assumes that it is conducted by the mixture-of-experts and the MOSAIC depending on task requirements (see Discussion). The blue parts are related to the internal models, whereas the red parts are related to the switching functions. RP, Responsibility predictor.

2003; Kelly and Strick, 2003). In particular, a recent study (Kelly and Strick, 2003) identified putative, minute reciprocal one-to-one connections between the prefrontal (area 46) and the cerebellum. Internal models in the cerebellum compete to estimate the current context, and their estimation errors are sent to the prefrontal RE. Based on the error information, the prefrontal estimator decides the responsibility of each internal model and gives the cerebellum feedback about the degree of the contribution of the internal model to the total output. Therefore, reciprocal one-to-one connections between each cerebellar internal model and the prefrontal counterpart are essential for this neural implementation of the MOSAIC model.

Another component of the MOSAIC architecture is a responsibility predictor (RP) that predicts the responsibility of the module from sensory and cognitive information. For example, before lifting an object, we can roughly estimate its weight from visual information (e.g., object size and surface texture). The responsibility predictor can foresee the responsibility before movement onset. The activity in the occipital, temporal, and superior parietal regions increased immediately after the presentation of the cognitive cues for switching. Visual information processed in the occipital and temporal regions converges at the superior parietal regions. The responsibility predictor may also consist of several components distributed between the cerebrum (red parts) and the cerebellum (blue parts), and components closely related to

sensory processing may exist in the superior parietal regions. This is also consistent with anatomical connectivity between the posterior parietal cortex and the cerebellum (Clower et al., 2001).

The other schema, illustrated in Figure 8*B*, assumes that the mixture-of-experts and MOSAIC architectures exist in distinct brain regions. Area 46 and the insula, where the transient switch activity was dominant and no spatial overlap with internal model activity was detected, probably assume the role of the gating module. The PM, the AP, and the cerebellar regions, where sustained activity specific to each novel mouse was found (Fig. 7), probably correspond to expert modules. Why are the two types of architecture necessary in the brain? Sensory feedback as a consequence of our own behavior is essential for the MOSAIC because the responsibility of the modules is determined according to the error between the prediction of the forward model and the actual sensory feedback. However, direct and continuous sensory feedback is sometimes unavailable for switching in several artificial circumstances with which modern people need to deal. In such cases, the mixture-of-experts or other mechanisms in the cerebral cortex are expected to play an important role. Many imaging studies have used tasks in which direct sensory feedback was unavailable. For example, Dove et al. (2000) used a task in which a simple relationship between a stimulus and a response changed repeatedly (switching of S-R mappings). They found that the supplementary motor area, the lateral prefrontal cortex, the parietal regions, the insula, and the cuneus/precuneus were related to the switching but the cerebellar activity was not reported. Bischoff-Grethe et al. (2002) suggested that the cerebellum also contributes to reassignment of the S-R mappings. The cerebellar activity (in about $y = -66$) was located posterior to that in our experiment ($y = -48$). This is consistent with the assumption that the more abstract cognitive functions such as arbitrary S-R mapping are located in posterior regions in comparison with more basic sensory-motor integration (for review, see Desmond and Fiez, 1998).

In the current study, transient activity related to the switching of the mouse type was observed in the dorsolateral prefrontal cortex (area 46). According to a previous study (D'Esposito et al., 1995), this region includes a central executive system that selects information stored in short-term memory buffers and plays an important role in dual-task performance (e.g., concurrent accomplishment of semantic-judgment tasks and spatial-judgment tasks). Other studies used a spatio-temporal working memory task and reported that area 46 was active when the subjects selected items from working memory rather than when they maintained the items in memory (Rowe and Passingham, 2001; Sakai et al., 2002). These studies suggest that area 46 is involved in the control of information flow according to contexts (i.e., what the gating module is supposed to do in the mixture-of-experts architecture). The Wisconsin Card Sorting Test (WCST) has been used to probe behavioral flexibility [i.e., the ability to shift from one response tendency (cognitive set) to another that is suitable for the current context]. Previous studies (Konishi et al., 1998; Nakahara et al., 2002) indicated that the posterior part of the inferior frontal sulcus (area 44/45) plays an essential role in the WCST. This region is located next to area 46 and probably has similar functions.

We examined the possibilities that the spatial and temporal overlaps found in the cerebellum were consequences of method limitations or artifacts of data processing. If the gating module and the expert module are located so closely that we cannot distinguish between them within the limit of fMRI spatial resolution, the transient activity related to the switching and the sus-

tained activity specific to the novel mice may overlap each other even if they are anatomically distinct. However, this is improbable, first because we successfully separated the transient activity from the sustained activity in the cerebral cortex for the same subjects while using the same methods as in the cerebellum. Second, we conducted individual–subject analysis on the data, to which a spatial filter was not applied to preserve the resolution of the raw images, and obtained consistent results with Figure 7 (see the last paragraph in Results). Furthermore, we observed a marked temporal overlap of the transient and sustained activities for a single voxel of a single subject in the event-related analysis for several brain regions, including the cerebellum.

Our previous studies indicated that the internal models for the rotated mouse and the velocity mouse are acquired in the neighboring regions near the posterior superior fissure (Imamizu et al., 1998, 2000, 2003; Tamada et al., 1999). We have already confirmed that these sustained activations were not caused by behavioral factors such as tracking error, attention, effort, or eye/hand movements (Imamizu et al., 1998, 2000, 2003). The first regression analysis of the current study additionally suggests that the transient activity related to the mouse-type change was also free from the above confounding behavioral factors. This is because the mouse-type change activated only a few brain regions (Table 1), whereas the cursor reset that increased the above behavioral factors activated different various sensory–motor regions, as listed in Table 2.

In conclusion, the switch activity was segregated from the internal model activity in the prefrontal regions (area 46 and the insula), whereas it largely overlapped the internal model activity in the cerebellum and the parietal cortex. Therefore, switching mechanisms modeled by the gating module and the RE can be achieved in the prefrontal regions and that internal models of the MOSAIC can be achieved in the cerebellum and the parietal cortex. Our results suggest globally distributed switching mechanisms over the cerebro-cerebellar communication loop and distinct computational roles assigned to different parts of this loop in selecting appropriate internal models for given contexts.

References

- Bischoff-Grethe A, Ivry RB, Grafton ST (2002) Cerebellar involvement in response reassignment rather than attention. *J Neurosci* 22:546–553.
- Blakemore SJ, Wolpert DM, Frith CD (1998) Central cancellation of self-produced tickle sensation. *Nat Neurosci* 1:635–640.
- Clower DM, West RA, Lynch JC, Strick PL (2001) The inferior parietal lobe is the target of output from the superior colliculus, hippocampus, and cerebellum. *J Neurosci* 21:6283–6291.
- D'Esposito M, Detre JA, Alsop DC, Shin RK, Atlas S, Grossman M (1995) The neural basis of the central executive system of working memory. *Nature* 378:279–281.
- Desmond JE, Fiez JA (1998) Neuroimaging studies of the cerebellum: language, learning and memory. *Trends Cogn Sci* 2:355–362.
- Dove A, Pollmann S, Schubert T, Wiggings CJ, Cramon DY (2000) Prefrontal cortex activation in task switching: an event-related fMRI study. *Cognit Brain Res* 9:103–109.
- Doya K, Samejima K, Katagiri K, Kawato M (2002) Multiple model-based reinforcement learning. *Neural Comput* 14:1347–1369.
- Dum RP, Strick PL (2003) An unfolded map of the cerebellar dentate nucleus and its projections to the cerebral cortex. *J Neurophysiol* 89:634–639.
- Flanagan JR, Wing AM (1997) The role of internal models in motion planning and control: evidence from grip force adjustments during movements of hand-held loads. *J Neurosci* 17:1519–1528.
- Gomi H, Kawato M (1993) Recognition of manipulated objects by motor learning with modular architecture networks. *Neural Netw* 6:485–497.
- Graybiel AM, Aosaki T, Flaherty AW, Kimura M (1994) The basal ganglia and adaptive motor control. *Science* 265:1826–1831.
- Haruno M, Wolpert DM, Kawato M (2001) Multiple paired forward-inverse models for sensorimotor learning and control. *Neural Comput* 13:2201–2220.
- Iacoboni M (2001) Playing tennis with the cerebellum. *Nat Neurosci* 4:555–556.
- Imamizu H, Miyauchi S, Tamada T, Sasaki Y, Takino R, Putz B, Yoshioka T, Kawato M (1998) Modular organization of multiple internal models for visuomotor learning: a functional MRI study. *Soc Neurosci Abstr* 24:166.
- 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.
- Jacobs RA, Jordan MI (1991) A competitive modular connectionist architecture. In: *Advances in neural information processing systems 3* (Lippmann RP, ed), pp 767–773. San Mateo, CA: Morgan Kaufmann.
- Jordan MI, Jacobs RA (1994) Hierarchical mixture of experts and the EM algorithm. *Neural Comput* 6:181–214.
- Kawato M (1999) Internal models for motor control and trajectory planning. *Curr Opin Neurobiol* 9:718–727.
- Kawato M, Wolpert DM (1998) Internal models for motor control. In: *Sensory guidance of movement* (Glickstein M, ed), pp 291–307. Sussex: Wiley.
- Kawato M, Furukawa K, Suzuki R (1987) A hierarchical neural-network model for control and learning of voluntary movement. *Biol Cybern* 57:169–185.
- Kelly RM, Strick PL (2003) Cerebellar loops with motor cortex and prefrontal cortex of a nonhuman primate. *J Neurosci* 23:8432–8444.
- Konishi S, Nakajima K, Uchida I, Kameyama M, Nakahara K, Sekihara K, Miyashita Y (1998) Transient activation of inferior prefrontal cortex during cognitive set shifting. *Nat Neurosci* 1:80–84.
- Miall RC, Weir DJ, Wolpert DM, Stein JF (1993) Is the cerebellum a Smith predictor? *J Motor Behav* 25:203–216.
- Miall RC, Reckess GZ, Imamizu H (2001) The cerebellum coordinates eye and hand tracking movements. *Nat Neurosci* 4:638–644.
- Middleton FA, Strick PL (1994) Anatomical evidence for cerebellar and basal ganglia involvement in higher cognitive function. *Science* 266:458–461.
- Middleton FA, Strick PL (2001) Cerebellar projections to the prefrontal cortex of the primate. *J Neurosci* 21:700–712.
- Nakahara K, Hayashi T, Konishi S, Miyashita Y (2002) Functional MRI of macaque monkeys performing a cognitive set-shifting task. *Science* 295:1532–1536.
- Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9:97–113.
- Rowe JB, Passingham RE (2001) Working memory for location and time: activity in prefrontal area 46 relates to selection rather than maintenance in memory. *NeuroImage* 14:77–86.
- Sakai K, Rowe JB, Passingham RE (2002) Active maintenance in prefrontal area 46 creates distractor-resistant memory. *Nat Neurosci* 5:479–484.
- Sasaki K, Oka H, Kawaguchi S, Jinnai K, Yasuda T (1977) Mossy fibre and climbing fibre responses produced in the cerebellar cortex by stimulation of the cerebral cortex in monkeys. *Exp Brain Res* 29:419–428.
- Shadmehr R, Mussa-Ivaldi FA (1994) Adaptive representation of dynamics during learning of a motor task. *J Neurosci* 14:3208–3224.
- Tamada T, Miyauchi S, Imamizu H, Yoshioka T, Kawato M (1999) Cerebro-cerebellar functional connectivity revealed by the laterality index in tool-use learning. *NeuroReport* 10:325–331.
- Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M (2002) Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage* 15:273–289.
- Wolpert D, Kawato M (1998) Multiple paired forward and inverse models for motor control. *Neural Netw* 11:1317–1329.
- Wolpert DM, Ghahramani Z (2000) Computational principles of movement neuroscience. *Nat Neurosci [Suppl]* 3:1212–1217.
- Wolpert DM, Ghahramani Z, Jordan MI (1995) An internal model for sensorimotor integration. *Science* 269:1880–1882.
- Wolpert DM, Miall RC, Kawato M (1998) Internal models in the cerebellum. *Trends Cogn Sci* 2:338–347.
- Wolpert DM, Doya K, Kawato M (2003) A unifying computational framework for motor control and social interaction. *Philos Trans R Soc Lond B Biol Sci* 358:593–602.