Three timescales in prism adaptation

Masato Inoue, Motoaki Uchimura, Ayaka Karibe, Jacinta O’Shea, Yves Rossetti, and Shigeru Kitazawa


It has been proposed that motor adaptation depends on at least two learning systems, one that learns fast but with poor retention and another that learns slowly but with better retention (Smith MA, Ghazizadeh A, Shadmehr R. PLoS Biol 4: e179, 2006). This two-state model has been shown to account for a range of behavior in the force field adaptation task. In the present study, we examined whether such a two-state model could also account for behavior arising from adaptation to a prismatic displacement of the visual field. We first confirmed that an “adaptation rebound,” a critical prediction of the two-state model, occurred when visual feedback was deprived after an adaptation-extinction episode. We then examined the speed of decay of the prism aftereffect (without any visual feedback) after repetitions of 30, 150, and 500 trials of prism exposure. The speed of decay decreased with the number of exposure trials, a phenomenon that was best explained by assuming an “ultraslow” system, in addition to the fast and slow systems. Finally, we compared retention of aftereffects 24 h after 150 or 500 trials of adaptation. The speed of decay was significantly greater after 500 than 150 trials. This difference in retention could not be explained by the two-state model but was well explained by the three-state model as arising from the difference in the amount of adaptation of the “ultraslow” process. These results suggest that there are not only fast and slow systems but also an ultraslow learning system in prism adaptation that is activated by prolonged prism exposure of 150–500 trials.

prism adaptation; reaching; motor learning; two-state model; three-state model

Inoue M, Uchimura M, Karibe A, O’Shea J, Rossetti Y, Kitazawa S. Three timescales in prism adaptation. J Neurophysiol 113: 328–338, 2015. First published October 8, 2014; doi:10.1152/jn.00803.2013.—It has been proposed that motor adaptation depends on at least two learning systems, one that learns fast but with poor retention and another that learns slowly but with better retention (Smith MA, Ghazizadeh A, Shadmehr R. PLoS Biol 4: e179, 2006). This two-state model has been shown to account for a range of behavior in the force field adaptation task. In the present study, we examined whether such a two-state model could also account for behavior arising from adaptation to a prismatic displacement of the visual field. We first confirmed that an “adaptation rebound,” a critical prediction of the two-state model, occurred when visual feedback was deprived after an adaptation-extinction episode. We then examined the speed of decay of the prism aftereffect (without any visual feedback) after repetitions of 30, 150, and 500 trials of prism exposure. The speed of decay decreased with the number of exposure trials, a phenomenon that was best explained by assuming an “ultraslow” system, in addition to the fast and slow systems. Finally, we compared retention of aftereffects 24 h after 150 or 500 trials of adaptation. The speed of decay was significantly greater after 500 than 150 trials. This difference in retention could not be explained by the two-state model but was well explained by the three-state model as arising from the difference in the amount of adaptation of the “ultraslow” process. These results suggest that there are not only fast and slow systems but also an ultraslow learning system in prism adaptation that is activated by prolonged prism exposure of 150–500 trials.

prism adaptation; reaching; motor learning; two-state model; three-state model

IT HAS BEEN PROPOSED that motor adaptation depends on at least two learning systems with different speeds, one that learns fast but has poor retention and another that learns slowly but has better retention (Smith et al. 2006). The idea was proposed by Smith et al. (2006) to account for data obtained during adaptation of saccadic eye movements (Kojima et al. 2004) and was further tested in adaptation of reaching movements in response to a force field perturbation (Smith et al. 2006) and visuomotor rotations (Lee and Scheweighofer 2009; Shmuelof et al. 2012; Zarahn et al. 2008). The model was successful in explaining many phenomena, such as savings, adaptation rebound (Smith et al. 2006), spontaneous recovery (Ethier et al. 2008), memory decay (Shmuelof et al. 2012; Vaswani and Shadmehr 2013), and generalization (Tanaka et al. 2012).

Joiner and Smith (2008), using a force field adaptation paradigm, further suggested that the magnitude of retention at 24 h could be predicted by the amount by which the “slow” system had adapted: retention at 24 h reflected ~65% of the magnitude of adaptation of the slow process after adaptation with 11, 30, 103, and 160 trials. It was noteworthy that the amount of retention at 24 h approached a plateau as the number of adaptation trials increased to 103 and 160 trials (Fig. 3 in Joiner and Smith 2008). By extrapolating these results to a still larger number of trials, it is reasonable to predict that the magnitude of adaptation of the slow process (and the amount of retention at 24 h) in the force field task should converge to an asymptote after an even larger number of adaptation trials.

However, Yin and Kitazawa (2001) demonstrated in monkeys that 250 trials of prism adaptation (i.e., adaptation of reaching movements in response to an optical displacement of the visual field) did not result in significant retention of prism aftereffects at 24 h. By contrast, 500 trials of adaptation produced significant retention, which amounted to ~70% of the visual displacement (Fig. 2 in Yin and Kitazawa 2001). This difference in 24-h retention rates with prism adaptation contrasts with those reported above for force field adaptation. This observed difference appears to challenge the standard two-state model, which posits a single slow state to account for retention. We therefore set out to test the hypothesis that there is a third system for prism adaptation that learns much more slowly than the slow system and that this would be engaged and grow significantly over the period between 250 and 500 exposure trials.

In the present study, we tested the three-state hypothesis for prism adaptation by testing three interrelated predictions from that model, namely: 1) the process of prism adaptation can be approximated by the two-state model up to 250 trials of adaptation, but 2) the third component, an “ultraslow” system, becomes essential to explain data after 500 adaptation trials, and 3) long-term retention at 24 h reflects the amount of adaptation in the ultraslow system, not in the slow system.
To address the first issue, we first examined whether a so-called “adaptation rebound,” a critical prediction of the two-state model (Smith et al. 2006), could be observed after 250 trials of prism adaptation (experiment 1). In the work by Smith and colleagues, participants adapted to a force field perturbation (e.g., a counterclockwise force field), and errors induced by the perturbation decreased with learning on a trial-by-trial basis. After 200 trials of adapting to the first force field, a second field with the opposite perturbation direction was introduced (e.g., a clockwise force field) for 15 trials. When the second perturbation was removed aftereffects from the second force field disappeared quickly, whereas aftereffects from the first force field reappeared as an adaptation rebound. We designed a similar protocol with prism adaptation and first examined whether the phenomenon of adaptation rebound would occur in this task and whether the process can be approximated by the two-state model. Three- and four-state models were also fitted to the data for comparison.

To address the second issue, we designed experiment 2. Naive participants adapted their reaching movements while wearing prisms in a block of 30, 150, or 500 trials. As in experiment 1, the rate of memory decay was assayed by trials in which visual feedback was removed. We chose three adaptation doses (30, 150, or 500 trials), with the expectation that data involving 500 adaptation trials would be better explained by assuming the third ultraslow component.

To test the third prediction, we examined long-term retention by assessing the magnitude of prism aftereffects after a delay of 24 h for each of the 150- and 500-trial adaptation conditions. We critically examined whether the magnitude of the aftereffect correlated better with the amount of adaptation in the slow system or in the ultraslow system.

MATERIALS AND METHODS

Participants

Seven male volunteers (aged 27–47 yr) participated in experiment 1. Twenty-four volunteers (10 women and 14 men, aged 20–47 yr) participated in experiment 2. All participants were right-handed, had normal or corrected-to-normal visual acuity, and had no significant neurological history. The study received approval from the institutional ethics committee, and all participants gave informed written consent before the experiments.

Apparatus and Task Procedure

The participant was seated facing a tangent 17-in. CRT screen, 400 mm from the eyes, with the head restrained by a chin rest. The participant wore a pair of spectacles with liquid-crystal shutters (PLATO, Translucent Technologies) and viewed the CRT screen through opaque tubes that restricted the field of view. The shutters could change from opaque white to transparent by means of a control signal from a personal computer (Dell, Precision 370). A trial started when the participant pressed a button with the index finger. The button was positioned 300 mm below and 100 mm ahead of the eyes in the midsagittal plane. When the participant pressed the button for 1.5 s, a target (5-mm-diameter circle, superimposed on a cross, 15 mm × 15 mm) appeared in a random location within a 40-mm square in the center of the visual field. The participant was required to release the button within 300 ms after the appearance of the target, to touch the screen within 300 ms after the release, and to hold the final position of the hand for 1 s until a 1-s beep instructed the participant to return the hand to the starting position.

Vision of the hand and the arm was blocked at the release of the button by the liquid-crystal shutters to eliminate visual feedback during the movement. In trials with visual feedback (visual feedback trials), the shutters were opened again at the touch for 300 ms to allow participants to see the target and the final static position of the hand. In trials without visual feedback (no-visual feedback trials), the shutters were kept closed until the next trial, so that participants could not see the final static position of the hand and the target.

Whenever the visual field was displaced by wedge prisms placed within the opaque tubes, the CRT screen was also displaced (rotated) around the vertical axis that passed the middle point of the participant’s eye to compensate for the visual displacement of the screen caused by the prisms: it was displaced by the amount of refraction (e.g., 8.5°, 15 diopters). Because the distance between the screen and the eyes was 400 mm, the screen was displaced 60 mm along an arc of 400-mm radius. This procedure kept the visual image of the screen in the exposure period almost identical to that in the pretest period: the participant viewed the tangent CRT screen in the same straight-ahead direction as in the first period (Kitazawa et al. 1995).

Experimental Design

Experiment 1. The prism adaptation task consisted of four phases: 1) pretest (60 trials), 2) the first exposure (exposure 1, 200 trials), 3) the second exposure (exposure 2, 15 trials), and 4) posttest (50 trials). The pre- and posttest blocks were performed without visual displacement. During the first exposure block the visual field was displaced by 60 mm (8.5°, 15 diopters) to the left, and during the second exposure block it was displaced by 60 mm to the right.

In the pretest block, visual feedback was given in the first and last 15 trials (trials 1–15 and 46–60) but was deprived for the intervening 30 trials (trials 16–45). In the first exposure block visual feedback was given during the first 30 trials and deprived during the next 10 trials, and this alternating cycle (30 visual feedback trials followed by 10 no-visual feedback trials, termed as an epoch hereafter) was repeated 5 times, yielding 200 trials in sum. This manipulation, interleaving blocks of trials with versus without visual feedback throughout the different phases of the experiment, was designed to track changes in the magnitude and stability of learning and prism aftereffects over time. By alternating between blocks of prism exposure versus aftereffect assessment, we expected to observe a "zig-zag" change in the sign and magnitude of reaching error, which would be beneficial for estimating adaptation and decay parameters for the computational model, formulated below in Eq. 1. The length of the first exposure block (a total of 150 trials with the initial visual displacement excluding 50 trials without visual feedback) was chosen to ensure a sufficient number of trials to trigger adaptation of the hypothesized slow system. During the second exposure block, the visual field was displaced by 60 mm in the direction opposite to that during the first exposure block and visual feedback was given in all 15 trials. This rather short block length was chosen following Smith et al. (2006) as sufficient to ensure that the fast system would adapt to the second displacement but the slow system would still retain adaptation to the initial displacement. In the posttest, participants performed 50 trials without visual feedback. Critically, we expected to observe adaptation rebound during the posttest, namely, a reemergence of the phenomenon of adaptation rebound.

Experiment 2. On day 1, participants performed blocks of pretest (30 trials), adaptation (30, 150, or 500 trials), and posttest (50, 100, or 100 trials). In the pretest visual feedback was present in the first 10 trials, deprived in the next 10 trials, and present again in the last 10 trials.

During the adaptation block, the visual field was displaced by 50 mm to the left (n = 12) or to the right (n = 12) by the prisms (7.1°, 12.5 diopters). The CRT screen was also displaced to compensate for the visual displacement of the screen, as described in experiment 1.
Adaptation was performed with visual feedback at the end of each movement, in a block of 30, 150, or 500 trials.

In the posttest block visual feedback was deprived in order to test the speed of prism aftereffect decay following a short delay. By contrast with experiment 1, in experiment 2 both the prismatic and CRT displacement were maintained throughout both the adaptation and the posttest block. Logically, during the posttest block the visual displacements imposed during the exposure period (i.e., by the prisms and the CRT) should be removed. However, as this removal did not change any visual experience, these displacements were maintained throughout the posttest block in experiment 2 to save time. That is, regardless of whether the prism and the CRT were in place or not, the initial visual input (before movement onset) was no different. Furthermore, the pointing measures taken during the posttest block were obtained under conditions of no visual feedback, so the presence or absence of the prism and CRT was irrelevant during the posttest period in which behavior was sampled. By skipping the process of removing the displacements in experiment 2, this afforded us greater temporal resolution, enabling us to test the speed of memory decay without a delay that would otherwise be unavoidable. The posttest consisted of 50 trials (in the 30-trial adaptation condition) and 100 trials (in the 150- and 500-trial adaptation conditions).

On day 2 retention of the prism aftereffect was tested, 24 h after the adaptation and posttest blocks of day 1. During the posttest on day 2, participants performed 30 trials of reaching without prisms and without visual feedback.

Each volunteer ($n = 24$) participated in three sessions, once for each of the 30-, 150-, and 500-trial adaptation conditions. The order of the three conditions was counterbalanced across participants. Intrasession intervals were 7 days on average (SE = 1 day). The direction of the visual displacement was fixed to the left or to the right for each participant.

**Data Analysis**

Two- , three-, and four-state models (experiments 1 and 2). We analyzed horizontal errors because the visual field was displaced in the horizontal direction. The error was defined as the discrepancy (distance) between the actual position of the target and the reach end point, irrespective of whether the participants were allowed or not allowed to see the errors. The mean horizontal error in trials with no visual feedback during the preexposure period was adjusted to 0. We then fitted the two-, three- and four-state models to the data from each participant and to the group mean of the horizontal errors averaged across the 7 (experiment 1) and 24 (experiment 2) participants. To further estimate confidence intervals (c.i.) for each parameter, we prepared 100 bootstrap data samples for each experiment by resampling 7 (experiment 1) and 24 (experiment 2) participants with replacement and applied the two-, three-, and four-state models to the mean of each bootstrap data sample (Burnham and Anderson 2002).

The two-state model was formulated after Smith et al. (2006) with a modification as follows:

$$e(n) = d(n) + x(n) - \text{bias}$$
$$x(n) = x_f(n) + x_s(n)$$
$$x_f(n) = A_x x_f(n - 1) - B_x e(n - 1)$$
$$x_s(n) = A_x x_s(n - 1) - B_x e(n - 1)$$

where $A_x$ and $B_x$ denote “retention factors,” which determine the speed of memory decay. The smaller the retention factor, the faster the decay. Thus the two inequalities reflect our assumption that the fast system learns faster and decays faster than the slow system.

$$x_f(n)$$, both of which were assumed to grow in proportion to the horizontal error on the previous trial, $e(n - 1)$, with constant “learning rates” of $B_f$ and $B_s$. The larger the learning rate, the faster the adaptation. $A_f$ and $A_s$, denote “retention factors,” which determine the speed of memory decay. The smaller the retention factor, the faster the decay. Thus the two inequalities reflect our assumption that the fast system learns faster and decays faster than the slow system.

The three- and four-state models were formulated as follows:

$$e(n) = d(n) + x(n) - \text{bias}$$
$$x(n) = x_f(n) + x_s(n) + x_{us}(n)$$
$$x_f(n) = A_x x_f(n - 1) - B_x e(n - 1)$$
$$x_s(n) = A_x x_s(n - 1) - B_x e(n - 1)$$
$$x_{us}(n) = A_{us} x_{us}(n - 1) - B_{us} e(n - 1)$$

where $A_{us}$ and $B_{us}$ denote “ultraslow factors,” which determine the speed of memory decay. The smaller the retention factor, the faster the decay. Thus the two inequalities reflect our assumption that the ultrafast system learns faster and decays faster than the ultrafast system.

and

$$e(n) = d(n) + x(n) - \text{bias}$$
$$x(n) = x_f(n) + x_s(n) + x_{us}(n) + x_{hs}(n)$$
$$x_f(n) = A_x x_f(n - 1) - B_x e(n - 1)$$
$$x_s(n) = A_x x_s(n - 1) - B_x e(n - 1)$$
$$x_{us}(n) = A_{us} x_{us}(n - 1) - B_{us} e(n - 1)$$
$$x_{hs}(n) = A_{hs} x_{hs}(n - 1) - B_{hs} e(n - 1)$$

where $A_{hs}$ and $B_{hs}$ denote “hyperslow factors,” which determine the speed of memory decay. The smaller the retention factor, the faster the decay. Thus the two inequalities reflect our assumption that the hyperslow system learns faster and decays faster than the ultrafast system.

The models were fitted to the data by applying the least squares method with MATLAB (version R2013a; optimization and global optimization toolboxes). The optimization under the condition of inequalities was achieved by using the fmincon function implemented in the optimization toolbox and repeated 100 times with different initial conditions by using the multistart function in the global optimization toolbox. We further restricted the range of search for the retention factor of the fast system between 0.7 and 0.95 (0.7 $< A_f < 0.95$) and the range of search for the learning rate between 0 and 0.3 (0 $< B_f < 0.3$). These ranges cover the reported values for the fast system ($A_f = A_s = 0.9$, $B_f < 0.11$) in previous studies (Joiner and Smith 2008; Smith et al. 2006). We set upper bounds for the learning rates ($B_f = 0.10$, $B_s = 0.03$, $B_{us} = 0.033$, $B_{hs} < 0.011$), so that the upper bounds decreased by a factor of 1/3.

To compare the two-, three-, and four-state models for their relative goodness of fit, the determination coefficient (d.c.) and Akaike’s information criterion (AIC) were calculated for each model. The d.c. and AIC were calculated with formulas as follows:

$$d.c. = 1 - \text{var(residual error)}/\text{var(horizontal error)}$$

and

$$AIC = n \ln[\text{var(residual error)}] + 2k$$

where $n$ denotes the number of data points ($n = 50 + 100 + 100$) and $k$ denotes the degrees of freedom in each model ($k = 5$ for the 2-state, 7 for the 3-state, and 9 for the 4-state model). The d.c. represents the proportion of the variance of the data explained by the model; it takes the maximum value of 1 when there are no residual errors. However, the d.c. is not suitable for choosing the best of the three models, because the d.c. increases with the number of parameters. AIC takes not only the residual error but also the degrees of freedom ($k$) into account: the degrees of freedom are added as a penalty to the natural
logarithm of the variance of the residual error. Thus the model that yields the smallest AIC can be judged as the best model. When we applied the two-, three-, and four-state models to each data set, the AIC difference was calculated by subtracting the smallest value of the three: the model that yielded 0 was judged to be the best model (Burnham and Anderson 2002). To test whether the AIC difference was significant or not, we applied the sign test (nonparametric paired-samples sign test) to the results obtained for the 7 (experiment 1) and 24 (experiment 2) participants. We further applied the sign test to the results obtained for the 100 bootstrap samples for each experiment. The nonparametric sign test was used because the AIC difference was bounded by 0 and did not follow the normal distribution.

An exponential model for approximating the immediate decay (experiment 2). To assess the speed of aftereffect decay after a short delay, we analyzed horizontal errors during the day 1 posttest in experiment 2. As the error tended to increase across trials (reflecting decay) to reach a plateau, the mean horizontal error was fitted to the exponential function with a plateau, formulated as follows:

\[ e(n) = [e(1) - \text{plateau}] \exp[-k(n-1)] + \text{plateau} \quad (6) \]

where \( e(n) \) is the horizontal error in the \( n \)th trial and \( k \) is the constant of decay. We further tested whether the mean horizontal error at the end of the posttest period (i.e., on trial 50) was the same across the 30-, 150-, and 500-adaptation conditions. Student t-tests were used after correcting the level of significance to 0.05/3 (Bonferroni correction).

Assessments of long-term retention (experiment 2). We examined long-term retention by assessing the magnitude of prism aftereffects after a delay of 24 h. The aftereffect, defined as the mean error on the initial five trials on day 2 (with no visual feedback), was calculated for each participant for each of the 150- and 500-adaptation conditions, and conditions were compared with a paired-samples t-test. We further examined whether the magnitude of the aftereffect correlated better with the amount of adaptation in the slow system or in the ultraslow system.

RESULTS

Experiment 1

Figure 1A shows the average data from seven participants. During the pretest, the horizontal error was distributed around 0. During the first exposure phase (exposure 1), participants initially made leftward errors in the direction of the prismatic displacement, with maximal errors of \( \sim 60 \text{ mm} \) (trial 61), i.e., the magnitude of the displacement, but they learned to gradually reduce their errors, with the initial error decreasing to \( < 15 \text{ mm} \) within 30 trials. When visual feedback was deprived, during the subsequent 10 trials (trials 91–100) the error increased from 14 to 26 mm on average, reflecting decay of what was learned. Further learning occurred in the next set of visual feedback trials, reflected in a reduction of horizontal errors, but these increased (i.e., decayed) again in the next phase when visual feedback was deprived, although the magnitude of error increase (decay) was smaller in this block than in the previous one. In the fifth epoch of no-visual feedback trials, the error was small but still increased from 4.8 mm on the first trial (trial 251) to 8.3 mm on the 10th trial (trial 260) (Fig. 1B). During the second exposure phase (exposure 2), in which the visual field was displaced 60 mm to the right, participants initially made a large leftward error (\( \sim 100 \text{ mm} \)), but then the error decreased rapidly. During the first few trials (trials 276–279) of the posttest, participants made errors as small as 5 mm to the left (Fig. 1, A and C). This may indicate that the adaptation to the initial visual displacement was completely washed out and an aftereffect to the second visual displacement was observed. However, the errors did not simply converge to 0, but rather crossed the zero-line and had increased up to 15 mm rightward within \( \sim 10 \) trials (trials 286–295; Fig. 1, A and C), indicating that clear “adaptation rebound” had occurred. Error bars show SE.

We assessed goodness of fit of the models to individual participant data as well as to the group mean data. Data from a representative individual are shown in Fig. 2A, showing changes in horizontal error and a clear adaptation rebound as for the group mean data, although the noise level is of course much larger. The data agreed quantitatively with the two-state model.

Fig. 1. Adaptation rebound in prism adaptation (experiment 1). A: group mean horizontal errors plotted against trial sequence (\( n = 7 \)). Thick solid line indicates the size of the prismatic displacement of the visual field. Open circles indicate horizontal errors in the visual feedback trials; filled circles indicate horizontal errors in the no-visual feedback trials. Note the adaptation rebound during the posttest period (trials 276–325). B: mean horizontal errors in trials 1 and 10 during the 1st and 5th epochs of no-visual feedback trials during the 1st exposure period. C: mean horizontal error during the posttest period. Note that the mean error during trials 11–20 is inverted and larger than the mean error during trials 1–3, indicating that clear “adaptation rebound” had occurred. Error bars show SE.
Fig. 2. Application of the multistate models to the data showing adaptation rebound. A and B: 2 (A) and 3 (B)-state models fitted to horizontal errors from a typical participant (subject 1). C and D: 2 (C) and 3 (D)-state models fitted to horizontal errors from the group mean across 7 participants (average). Data in C and D are the same as those in Fig. 1. Green, blue, and magenta traces show development of the fast (green), slow (blue), and ultraslow (magenta) systems, and red traces show the motor output predicted from the 2 (A and C)- and 3 (B and D)-state models. E and H: comparison of Akaike’s information criterion (AIC) among the 3 models applied to data from individual participants (E, n = 7) and the mean of bootstrap data sample (n = 100). Note that the smallest AIC (ΔAIC = 0) was yielded by the 2-state model in 6 of 7 participants (E) and 91 of 100 bootstrap data samples. Brackets show pairs of comparison with significant differences (* P < 0.05/3, Bonferroni correction). F, G, I, and J: distribution of the retention factors (F and I) and learning rates (G and J) estimated by the 2-state model for each participant (F and G) and for each of 100 bootstrap data samples. Notches and horizontal lines in the box plots show 2.5th, 25th, 50th, 75th, and 97.5th percentiles.
panioned by an increase in AIC of $\sim 2$ ($\Delta$AIC = 1.7). No further improvement was observed with the four-state model in terms of d.c. As a result, the AIC increased by 4 in the four-state model compared with the three-state model (AIC = 5.8). Therefore, the two-state model was judged to be the best in terms of AIC for this particular participant.

Likewise, the two-state model performed best in six of seven participants and the four-state model was the worst in all (Fig. 2E). The paired sign tests showed that the AIC was significantly larger with the four-state model compared with the two- and three-state models (Fig. 2E; $P = 0.016 < 0.05/3$; Bonferroni correction), but the difference between the two- and three-state models did not reach significance.

As for the group mean data, >90% of the variance was explained by the two-state model (Fig. 2C; d.c. = 0.934). However, the d.c. was not improved by adding the ultraslow component because the learning constant converged to 0 (Fig. 2D). Accordingly, the two-state model yielded the smallest AIC. By applying the three models to the mean of 100 bootstrap data samples, we found that the two-state model was selected 91 times (model selection frequency $= 0.91$), the three-state model was selected 9 times (0.09), but the four-state model never was selected (0). Accordingly, the sign tests showed that the two-state model was significantly better than the three- and four-state models in terms of the AIC difference (Fig. 2H; $P < 0.001 < 0.05/3$, in all combinations).

The results show that the process of adaptation rebound after ~200 trials of the first exposure and 15 trials of the second exposure was best explained by the two-state model, as we initially expected. The medians of the estimated parameters were 0.834 ($A_1$), 0.994 ($A_2$), 0.182 ($B_1$), 0.0379 ($B_2$), and 9.4 mm (bias), with 95% c.i. of [0.741, 0.890] ($A_1$), [0.988, 1] ($A_2$), [0.128, 0.234] ($B_1$), [0.0198 0.0640] ($B_2$), and [7.1 11.6] (bias) (Fig. 2, I and J).

Experiment 2

Figure 3A shows the average data from 24 participants. During the pretest period, the horizontal error was distributed around 0. During the exposure period, the initial error was approximately the magnitude of the visual displacement (~50 mm) but decreased across trials to ~10 mm within 30 trials and reached a plateau of approximately ~5 mm thereafter.

During the posttest period, the horizontal error increased, that is, the aftereffect decayed in an exponential manner across trials in all conditions (Fig. 3, A and B). The rate of decay was the largest in the 30-trial adaptation condition ($k = 0.11$), was nearly halved in the 150-adaptation condition (0.058), and was further near-halved in the 500-adaptation condition (0.026). The mean error at trial 50 was significantly different in any combination of the three conditions ($P < 0.0077 < 0.05/3$, paired t-tests). This result indicates that the rate of decay was significantly different across the three conditions.

Next, we examined which of the two-, three-, and four-state models could best explain these data. In a typical example (participant 19), the d.c. improved from 0.471 (Fig. 4A, 2-state) to 0.475 (Fig. 4B, 3-state), and the AIC was also improved by 4.4, when the ultraslow component was added to the two-state model. However, there was no further improvement when the hyperslow component was added to the three-state model. As for the averaged data, the d.c. was improved from 0.888 (2-state) to 0.8928 (3-state) and the AIC was improved by 37 (Fig. 4, C and D) by addition of the ultraslow component to the two-state model. The d.c. was further improved by addition of the hyperslow component (d.c. = 0.8931), but the AIC increased by 1.7, showing that the three-state model was the best. This was further confirmed by applying the models to the 100 bootstrap data samples (Fig. 4H): the model selection frequency was largest for the three-state model (0.41), followed by the four-state model (0.32), and least for the two-state model (0.27). The three AIC comparisons were all significant (paired sign tests, $P < 0.0005 < 0.05/3$). The results clearly show that the three-state model best explained the different decays after 30-, 150-, and 500-adaptation trials.

The medians of the estimated parameters were 0.914 ($A_1$), 0.995 ($A_2$), 0.9996 ($A_3$), 0.202 ($B_1$), 0.0612 ($B_2$), 0.0330 ($B_3$), and −3.0 mm (bias), with 95% c.i. of [0.795, 0.95] ($A_1$), [0.987, 0.999] ($A_2$), [0.997, 1] ($A_3$), [0.119 0.296] ($B_1$), [0.0307 0.10] ($B_2$), [0.0082 0.0330] ($B_3$), and [−6.9 0.47] (bias) (Fig. 4, I and J).

The long-term aftereffect at 24 h was tested in the 150- and 500-adaptation conditions. We defined the magnitude of long-term retention as the mean of the horizontal errors in the first five trials in the posttest period on day 2. The mean prism aftereffect magnitude after 500 trials of adaptation (18 ± 12 mm, mean ± SD) was significantly larger than that after 150 trials of adaptation (8.7 ± 11 mm) (Fig. 5B; $P = 0.0002$, paired t-tests).

Previous work (Joiner and Smith 2008) suggested that retention at 24 h depended on the amount of adaptation of the slow process of the two-state model. Hence, we calculated the magnitude of adaptation of the slow process at the end of day 1 for each participant, based on the two-state model, and compared that value with the magnitude of prism aftereffect retained at 24 h (Fig. 5C). There was no significant correlation between the magnitude of adaptation of the slow process and the size of the aftereffect at 24 h ($r = 0.10, P = 0.50$). In contrast, the amount of adaptation of the ultraslow process in the three-state model did correlate with the amount of prism aftereffect retained at 24 h (Fig. 5D; $r = 0.52, P = 0.00016$). The slope of an orthogonal regression line was close to 1 (1.0; 95% c.i. = [0.55 1.84]; bootstrap method, $n = 10,000$) with an y-intercept of −6.1 mm (c.i. = [−18 3.1]), suggesting that the adaptation of the ultraslow process decayed little after an interval as long as 24 h. (Alternatively, the adaptation of the ultraslow process may have deadapted during the 24 h and copied to another unknown system for long-term retention.)

The amount of adaptation of the “hyperslow process” in the ultraslow system of the three-state model, with a retention contrast, the amount of adaptation of the ultraslow process in the three-state model did correlate with the amount of prism aftereffect retained at 24 h (Fig. 5E; $r = 0.37, P = 0.0103$). It is also worth noting that there was no significant correlation between the magnitude of adaptation, quantified as the last five trials of the adaptation block on day 1, and the size of the aftereffect at 24 h ($r = 0.08, P = 0.60$). The results clearly indicate that what was carried over to the next day was proportional to the adaptation in the ultraslow system of the three-state model, with a retention ratio close to 1.
DISCUSSION

In experiment 1, we confirmed that an adaptation rebound, a critical prediction of the two-state model, occurred in the posttest block (in which visual feedback was deprived) following a block of adaptation and deadaptation. To our knowledge, this is the first demonstration of adaptation rebound during visuomotor adaptation that occurs immediately after an adaptation-extinction episode (however, see Hatada et al. 2006 for adaptation rebound that developed over days). The fit of the two-state model to the data supports the hypothesis that prism adaptation also depends on two systems, one that learns fast but forgets fast and another that learns slowly but forgets less. We further fitted three-state and four-state models to the data in experiment 1, which contained 150 adaptation trials. We predicted that adding the slower systems would not improve the fit of the model with this number of adaptation trials. As expected, the two-state model was shown to be the best in terms of the AIC.

In experiment 2, which included 500 adaptation trials in one condition, we predicted that the three-state model would better represent the data as a whole than the two-state model. Our prediction was confirmed again: the three-state model was shown to be better than the two- or four-state models in terms of the AIC. In addition, the amount of retention on day 2 significantly correlated with the amount of adaptation in the ultraslow system of the three-state model at the end of day 1 but not with the amount of adaptation in the slow system (2-state model). These results consistently support our hypothesis that there is a third slower component in prism adaptation in addition to the conventional fast and slow systems.

Is the Ultraslow System Unique to Prism Adaptation?

Why were three systems required to account for these data in prism adaptation, whereas only two systems were sufficient to explain short-term as well as long-term retention at 24 h in
Fig. 4. Application of multistate models to the data in experiment 2. A and B: 2 (A)- and 3 (B)-state models fitted to horizontal errors from a typical participant (subject 19). C and D: 2 (C)- and 3 (D)-state models fitted to the group mean across 24 participants (average). Data in C and D are the same as those in Fig. 3A. Green, blue, and magenta traces show development of the fast (green), slow (blue), and ultraslow (magenta) systems, and red traces show the motor output predicted from the 2 (A and C)- and 3 (B and D)-state models. E and H: comparison of AIC among the 3 models applied to data from individual participants (E, n = 24) and the mean of bootstrap data sample (n = 100). Brackets show pairs of comparison with significant differences (*P < 0.05/3, Bonferroni correction). Note in H that the AIC yielded by the 3-state model was significantly smaller than those yielded by the 2-state (P = 0.0000068) and 4-state (P = 0.00047) models. F, G, I, and J: distribution of the retention factors (F and I) and learning rates (G and J) estimated by the 3-state model for each participant (F and G) and for each of 100 bootstrap data samples. Other conventions are the same as in Fig. 2.
force field adaptation (Joiner and Smith 2008; Smith et al. 2006)? Recently, Vaswani and Shadmehr (2013) examined decay of aftereffects after adaptation to a force field perturbation for 300 trials. In the posttest period, errors were clamped to 0 (error-clamp trials). They found that the aftereffect decayed but reached a plateau, similar to our present results. They examined whether the fast-slow model could explain their data. With the best-fit parameters for the period of adaptation ($A_t = 0.876, B_t = 0.315, A_s = 0.995, B_s = 0.056$), ~70% of the variance during a long postexposure period was explained by the two-state model, but the model failed to reproduce the plateau: the two-state model predicted that the aftereffects would approach 0. This was true even when we fitted the two-state model to the overall data during and after adaptation (Fig. 6A; $A_f = 0.886, B_f = 0.266, A_s = 0.995, B_s = 0.055$, d.e. = 0.864). This discrepancy between the model predictions and the observed data can be resolved by positing an additional ultraslow system (Fig. 6B; $A_f = 0.690, B_f = 0.189, A_s = 0.977, B_s = 0.147, A_us = 0.9991, B_us = 0.013$). With the three-state model, the end point of the decay (~0.2) was also reproduced (Fig. 6B). This result suggests that the ultraslow system may not be unique to prism adaptation but might be at work even during motor learning in response to force field perturbation.

Neural Mechanisms

The present study has demonstrated that prism adaptation behavior can be explained by a simple state-space model that posits the existence of three adaptive processes operating in parallel with different learning rates and retention capacities. One question is where in the brain these proposed functional mechanisms are implemented.

There are several human neuroimaging studies that have used target-reaching with visual displacement induced by wedge prisms (Clower et al. 1996; Danckert et al. 2008; Luauté et al. 2009). However, Luauté et al. (2009) was the only study that examined blood oxygenation level-dependent (BOLD) signal changes during exposure to a visual displacement over an extended period, sufficient for the fast and the slow systems to develop (i.e., >20 trials; see traces in Fig. 4). Extrapolating from Fig. 4D, and assuming that fMRI signals in the fast system grow over the first few trials, peak at around trial 10, and gradually subside thereafter, BOLD signal changes in the left parieto-occipital sulcus (POS) in the Luauté et al. study would appear to match the profile of the fast system. On the other hand, slower signal development was observed in lobules IV and V of the cerebellar cortex, in the hemisphere ipsilateral to the hand that was used for reaching: signals in the cerebellar cortex reached a peak at trial 16 and subsided thereafter (Luauté et al. 2009). These results raise the possibility that the fast system is implemented in the cerebral cortex (POS), whereas the slow system is implemented in lobules IV and V in the cerebellar cortex. In monkeys, complex spikes recorded from Purkinje cells in lobules IV–VI in the cerebellar hemisphere ipsilateral to the hand appear to encode end-point errors.
in reaching (Kitazawa et al. 1998). We may speculate that long-term depression, triggered by complex spikes in the cerebellar cortex (Ito 1989, 2001), is a part of the slow learning process. A causal role for the cerebellum in prism adaptation is indicated by impaired performance following cerebellar lesions in humans (review in Jacquin-Courtois et al. 2013; Martin et al. 1996; Pisella et al. 2005; Thach et al. 1992; Weiner et al. 1983) as well as in monkeys (Baizer et al. 1999).

Studies have recently suggested that the memory trace of adaptation in ocular reflexes is initially acquired in the cerebellar cortex (floculus) and undergoes short-term storage and is then "transferred" to the vestibular nuclei, whereby the memory lasts over days (Anzai et al. 2010; Kassardjian et al. 2005; Shutoh et al. 2006). Extending these speculations regarding the ocular reflex to prism adaptation, we may suggest that plastic changes in the deep cerebellar nuclei are a likely candidate for the ultraslow process that would follow fast-process changes in the cerebellar cortex (Nagao et al. 2013).

The above-mentioned possibility that the fast, slow, and ultraslow processes implicated in prism adaptation are implemented in the left POS (fast), the cerebellar cortex (slow), and the deep cerebellar nuclei (ultraslow) may seem to contradict a report by Galea et al. (2011) in which the cerebellar cortex was implicated in a fast-learning/forgetting process and the primary motor cortex (M1) was implicated in a slow-learning/forgetting process during adaptation of reaching movements to visuomotor rotation. However, this seeming contradiction may be resolved by taking the following points into account. First, adaptation to visuomotor rotation generally occurs much more slowly than prism adaptation. From the results of Galea et al., it took at least 20 trials of learning for the amount of initial error to be halved (from 30° to 15°), whereas, by contrast, it took only a few trials (<5 trials) for the initial error to be halved in the present and other studies on prism adaptation (Kitazawa et al. 1997; Luauté et al. 2009; Weiner et al. 1983). Hence, the fast process for prism adaptation may be distinct from the fast process implicated in adaptation to visuomotor rotation, with the latter possibly corresponding to the slow process in prism adaptation. Assuming this, both studies agree in that each posits that the cerebellar cortex is involved in learning with a similar time constant, irrespective of whether that process is termed "fast" as in Galea et al. (2011) or "slow" as in the present study. Second, involvement of M1 in retention does not, of course, preclude additional involvement of the deep cerebellar nuclei, which receive inputs from cerebellar lobules IV–VI, particularly since this part of the dentate nucleus is closely connected with M1 via the thalamus and the pontine nucleus (Dum et al. 2002). It would not be surprising if two brain structures, both located within the same circuit interconnecting M1 and the dentate nuclei, were involved in long-term retention of a motor memory. We therefore speculate that the ultraslow process may be implemented in a more or less distributed fashion across a cerebro-cerebellar loop interconnecting M1, pontine nucleus, dentate nucleus, thalamus, and back again to M1.

**Future Directions**

In the present study, we assumed that the same error drives all three components simultaneously, as has been hypothesized by the original study of Smith and colleagues (Joiner and Smith 2008; Smith et al. 2006). However, it is possible that the brain is solving a credit assignment problem so that the error is efficiently assigned to one of the many components by using a Bayesian solution (Kording et al. 2007). It is also possible that reinforcement learning partially contributes to the acquisition and stabilization of the ultraslow component (Shmuelof et al. 2012).

Future work combining computational modeling (Berniker and Kording 2011; Kording et al. 2007; Lee and Schweighofer 2009) with detailed characterization of adaptation behavior (O’Shea et al. 2014), together with brain imaging techniques, should help to further elucidate the temporal dynamics of the learning systems mediating prism adaptation and help to determine the underlying neural substrates.

**GRANTS**

This research was supported by a research grant from the Human Frontiers Science Program Organization and by a Grant-in-Aid for Scientific Research on Innovative Areas (no. 25119002). J. O’Shea is supported by the Royal Society (United Kingdom).

**DISCLOSURES**

No conflicts of interest, financial or otherwise, are declared by the author(s).

**AUTHOR CONTRIBUTIONS**

manuscript; M.I., M.U., A.K., J.O., Y.R., and S.K. approved final version of manuscript.

REFERENCES


