Asymptotic limits of sensorimotor adaptation

Scott T. Albert[1], Jihoon Jang[1], Hannah Sheahan[2], Lonneke Teunissen[3],
Koenraad Vandevoorde[4], and Reza Shadmehr[1]

1. Department of Biomedical Engineering, Johns Hopkins School of Medicine, Baltimore MD
2. Dept. of Experimental Psychology, University of Oxford, Oxford UK
3. Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen Netherlands
4. Leuven Brain Institute, KU Leuven, Leuven Belgium.

Correspondence: Scott Albert, 416 Traylor Building, Johns Hopkins School of Medicine, 720 Rutland Ave., Baltimore, MD 21205, USA. Email: salbert8@jhmi.edu. Phone: 410-614-3424.

Acknowledgements: This work was supported by grants from the National Institutes of Health (R01NS078311, F32NS095706), the National Science Foundation (CNS-1714623), the Cambridge Trust, the Rutherford Foundation, and a travel grant from the Boehringer Ingelheim Fonds. Additionally, we would like to thank Hugo Fernandes and Konrad Kording for so graciously compiling and sharing their data with us. Finally, we would like to recognize the Summer School in Computational Sensory-Motor Neuroscience (CoSMo) and its organizers (Gunnar Blohm, Konrad Kording, and Paul Schrater), for giving us the opportunity to learn and develop the original idea for this work.

Abstract

In many forms of motor adaptation, performance approaches a limit at which point learning stops, despite the fact that errors remain. What causes this adaptation limit? Here we found that while reach adaptation exhibited an asymptotic limit, this limit was not fixed: when the variance of the perturbation decreased, the adaptation limit increased, and performance improved. Moreover, the limit could be altered in real-time by changing perturbation variance. The same was true at low reaction times, indicating that implicit processes influenced the adaptation limit. Using a mathematical model that describes adaptation as a balance between learning and forgetting, we found that participants altered their adaptation limit by changing their sensitivity to reaching errors. Changes in error sensitivity were linked to the consistency of past errors. These observations suggest that during adaptation, time-varying error sensitivities compete with constant forgetting, setting the boundary conditions that produce an apparent limit in total adaptation.

Introduction

During motor adaptation, humans and other animals experience perturbations that alter the sensory consequences of motor commands, yielding sensory prediction errors. The brain responds to each error by learning, thus adjusting its motor commands on the subsequent attempt. Over many trials, these adjustments to the motor commands accumulate, but surprisingly, adaptation often remains incomplete: even after an extended period of practice, residual errors persist. Residual errors are observed in various behaviors including reaching[1–4], saccades[5,6], and walking[7]. Thus, an unknown factor appears to limit the total amount of adaptation. Why does adaptation stop despite the fact that residual errors remain?

An important clue is the observation that the mean of the residual errors is not fixed, but varies with the variance of the past perturbations. For example, when people are exposed to perturbations that vary in magnitude from one trial to the next, an increase in the perturbation variance also increases the mean of the residual errors[8–10]. Thus, the asymptote of adaptation is not a hard limit, but a dynamic variable that depends on the second order statistics of the perturbation history.

To understand these observations, we begin with a well-established model in which the process of adaptation is driven by two competing forces: sensitivity to error which produces learning, and trial-by-trial forgetting which produces decay[11–14]. When errors are large, learning dominates, yielding changes in motor commands that improve performance. However, as errors become small, forces that encourage forgetting reach an equilibrium with forces that induce learning. At this equilibrium, performance stops changing, yielding persistent steady-state errors[15].

While this theoretical framework predicts that residual errors should vary with the *mean* of the perturbation, it fails to explain why the residual errors are altered by the *variance* of the perturbation. One possibility is that perturbation variance affects error sensitivity[12,14,16–21], while another possibility is that perturbation variance affects the forgetting rate. To compare these ideas, we trained volunteers to make reaching movements in the presence of variable visual or force field perturbations. In some experiments, we imposed a strict upper bound on reaction time, thus limiting the contributions of explicit strategy[22–25]. In all cases, we found that an increase in perturbation variance decreased asymptotic performance, thereby increasing residual errors. Critically, these changes in residual errors were caused by modulation of error sensitivity, not forgetting rates.

But why did increases in perturbation variance alter error sensitivity? To answer this question, we sorted pairs of movements according to error size, and found that sensitivity to each error size changed during adaptation: an error size that was consistently encountered was followed by an increase in sensitivity to that error[26]. This simple rule, increase error sensitivity when errors are consistent, correctly predicted not only changes to the rate of learning during adaptation, but also the magnitude of the residual errors that remained at the end of adaptation.

Together, the results suggest that the brain does not simply learn from error and then discard the error information. Rather, the effect of an error lasts beyond the trial in which it was experienced, forming a memory of past errors that serves to regulate error sensitivity. Changes in error sensitivity eventually reach an equilibrium with forgetting, resulting in an asymptotic limit on the total extent of adaptation.

Results

To quantify the relationship between mean residual error and perturbation variance, we started by re-analyzing data collected in an earlier study[8] in which the authors perturbed reaching movements in three groups of participants using visuomotor rotations (Fig. 1A, Rotation). All groups experienced perturbation sequences that had the same mean (30°), but different levels of trial-to-trial variability: one group experienced a constant perturbation of 30° (zero variance), while the other two groups experienced perturbations with low or high variance (Fig. 1B, top). At the end of training, there were residual errors in all groups. However, the residual errors increased with the perturbation variance (Fig. 1H, Fernandes, median residual error on last 10 trials; repeated measures ANOVA: $F(2,14)=17.8$, $p<0.001$, $\eta_p^2=0.54$).

Here we designed a set of experiments to answer two questions: (1) why does adaptation suffer from residual errors and (2) why does the size of the residual errors depend on perturbation variance?

Perturbation variance limits sensorimotor adaptation

We began by repeating the experiment performed by Fernandes and colleagues[8], but with an important adjustment. In the earlier work, all three perturbation conditions were experienced by the same set of subjects, raising the possibility that prior exposure to the visuomotor rotation could have altered subsequent learning in the other environments[26–28]. To avoid this possibility, we recruited different sets of participants for each perturbation condition.

In each of our experiments, participants held the handle of a robotic arm (Fig. 1A) and made reaching movements in a two-dimensional workspace. In Experiment 1, we introduced a visual perturbation and divided the participants into two groups: a zero-variance group (n=19) in which the perturbation magnitude remained invariant at 30° (Fig. 1C, black), and a high variance group (n=14) in which the perturbation was sampled on each trial from a normal distribution with a mean of 30° and standard deviation of 12° (Fig. 1C, red). Our results confirmed the earlier observation: after extended training, participants in the zero-variance group learned more than the high variance group (Fig. 1C, bottom; Fig. 1H, Exp. 1, mean error on last 10 epochs, two-sample t-test, $p=0.002$; Cohen's $d = 1.49$).

In Experiment 2, we tested the generality of this observation by measuring how participants responded to variability in force field perturbations (Fig. 1A, Force field). As before, we divided the participants into two groups, a zero-variance group (n=12) in which the perturbation magnitude remained constant at 14 N·sec/m (Fig. 1D, top, black), and a high variance group (n=13) in which the perturbation magnitude was sampled on each trial from a normal distribution with mean 14 N·sec/m and standard deviation of 6 N·sec/m (Fig. 1D, top, red). To track the learning process, we intermittently measured the forces that the subjects produced via channel trials [29] (Fig. 1A, channel). As in visuomotor adaptation, variance in the force field perturbation reduced the total amount of learning (Fig. 1D, bottom; Fig. 1H, Exp. 2, mean error on the last 5 epochs; two-sample t-test, $p=0.001$; Cohen's $d = 1.46$). Thus, perturbation variability consistently affected the adaptation limit across various modalities of reach adaptation.

Perturbation variance limits the total extent of adaptation

Closer examination of the late stage of training (Figs. 1B, 1C, and 1D, bottom) raised the possibility that adaptation had not completely saturated; perhaps with additional exposure, adaptation might have converged across variance conditions, even eliminating the residual errors. To examine this possibility,

we repeated Experiment 1, but this time more than doubled the training trials (Fig. 1E, top). Addition of these trials allowed performance to saturate, as evidenced by the slope of the reach angles (Fig. 1G, slope of the line fit to individual performance over the last 50 epochs was not different than zero; $p=0.71$ and $p=0.83$ for the low and high variance groups). Notably, despite extended training, errors persisted (Fig. 1H, Exp. 3, residual errors ± SD on last 50 epochs; zero-variance: 1.7 ± 0.9°; high variance: 8.7 ± 1.7°; t-test against zero; both groups, $p<0.001$). Furthermore, once again we found that increased perturbation variance coincided with an increase in residual error: at the end of adaptation, there was no overlap in the distributions of residual errors in the low and high variance groups (Fig. 1H, Exp. 3; two-sample t-test, $p<0.001$; Cohen's $d = 5.24$).

If perturbation variability causally modulated the asymptotic limit of adaptation, we reasoned that we could switch between two different asymptotic states by changing perturbation variance mid-experiment. To test this prediction, in Experiment 4 participants (n=14) first adapted to a zero-variance 30° visuomotor perturbation (Fig. 1F, black). With training, performance approached a plateau. We next increased the perturbation variance (while keeping the mean constant) by sampling from a normal distribution with a standard deviation of 12° (Fig. 1F, red). Coincident with the increase in perturbation variance there was a reduction in reaching angle (Fig. 1H, Exp. 4, mean residual error on last 10 epochs; two-sample t-test, $p=0.005$; Cohen's $d = 1.16$). Thus, despite having already learned to compensate for much of the perturbation, when the variance of the perturbation increased, the residual error in every subject also increased (Fig. 1H, Exp. 4).

Together, Experiments 1-4 demonstrated that even after extended practice, motor adaptation reached an asymptotic limit, resulting in small persistent errors. However, this asymptotic limit was dynamic, responding to the second order statistics of the perturbation.

Perturbation variance impairs the implicit component of learning

While reach adaptation can occur despite severe damage to the explicit, conscious learning system of the brain[30], under normal circumstances performance benefits from both implicit and explicit learning systems[31–34]. A plausible hypothesis is that when perturbation variance increases, there is impairment in the explicit system, resulting in increased residual error. One way to reduce or eliminate the influence of the explicit system during reach adaptation is to limit the preparation time that people have to initiate their movements[22–25]. If perturbation variance impairs explicit learning, we would predict that when reaction time is strongly limited, thus reducing or eliminating the influence of the explicit system, there will be increased residual errors. However, increased perturbation variance should have little or no additional effects on residual errors.

Alternatively, perturbation variance may affect predominately the implicit system. In that case, when reaction time is strongly limited, suppressing contributions of the explicit system, there will be increased residual errors. Critically, increased perturbation variance should produce further reductions in residual errors. In this scenario, the change in residual errors following increased perturbation variance should be the same regardless of whether reaction time is limited or not.

To suppress the contributions of explicit learning, we repeated Experiment 1, but forced the subjects to respond to the target with very low reaction times (less than half the time that they typically required to initiate a movement at the start of adaptation, Fig. 2A). As before, we divided participants into two groups: a zero-variance group (n=13) in which the perturbation magnitude remained invariant

at 30°, and a high variance group (n=12) in which the perturbation was sampled on each trial from a normal distribution with a mean of 30° and standard deviation of 12°.

Under normal conditions in which there was no constraint on reaction time (Fig. 2A, Experiment 1), introduction of the perturbation led to a dramatic increase in reaction time: participants nearly doubled their preparation time, potentially signaling the development and expression of explicit strategies. In contrast, in the constrained reaction time group (Fig. 2A, Experiment 5), subjects executed reaching movements at considerably lower latencies. In this group, the time required for movement preparation remained roughly constant throughout the experiment, even after the introduction of the perturbation.

As expected, limiting reaction time impaired adaptation rates. In the zero-variance (Fig. 2B) and high variance (Fig. 2C) conditions, performance at short reaction times was worse than that of long reaction times (two-sample t-test on last 10 epochs; p=0.041 and p=0.007 for zero and high variance; Cohen's d = 0.77 and 1.17 for zero and high variance). Thus, reducing reaction time impaired performance in both the zero and high variance perturbation groups.

Next, we compared the effects of perturbation variance on residual errors at low reaction times. We found that even when reaction times were constrained to suppress the explicit component of adaptation, increased perturbation variance produced a clear increase in the residual errors (Fig. 2D), reducing the total extent of learning by approximately 5° (Fig. 2E; difference in residual errors during the last 10 epochs; two-sample t-test, p<0.001; Cohen's d = 1.53). Therefore, despite suppression of explicit learning, the deficit in performance caused by high variance perturbations persisted: high variability led to a reduction in the extent of learning.

Critically, a 2-way ANOVA yielded a significant effect of both perturbation variance (F=33.08, p<0.001, η_p^2=0.38) and reaction time (F=13.76, p<0.001, η_p^2=0.20), but no interaction effect (F=1.07, p=0.31). The absence of an interaction effect indicated that the reduction in asymptotic performance induced by perturbation variability was no different when reaction time was high or low. If we assume that reduced reaction time impairs the ability of the explicit system to contribute to reach adaptation, then these results suggest that the reductions in performance observed in Experiments 1 and 5 were predominantly caused by impairments in implicit learning.

Perturbation variance reduces error sensitivity, but not forgetting rates

An examination of the data in Figs. 1 and 2 illustrates a frequently observed characteristic of motor adaptation: even at the late stages of training, performance continues to suffer from small, persistent errors[3,4,8,15,31,35–37]. Such persistent, steady-state errors are one of the fundamental predictions of mathematical models of adaptation[11–15] in which performance is driven by an interaction between two opposing forces, error-based learning, and trial-to-trial forgetting:

$$x^{(n+1)} = ax^{(n)} + be^{(n)} \tag{1}$$

Here the adapted state of the individual, *x*, changes from trial *n* to trial *n*+1 due to both forgetting, and learning. Forgetting is controlled by the retention factor *a*, which determines the fraction of memory retained from one trial to the next. Learning is controlled by sensitivity to error, denoted by *b*, which determines how much learning will occur from the experience of error *e*.

In this model, performance reaches a steady state (Fig. 3A) in which residual errors persist. This occurs because as training progresses, the errors which drive the learning process eventually become small enough that there is a balance between the forces that promote forgetting, and those that promote learning. At this stage learning appears to stop, despite the fact that errors remain. Thus, in principle, residual errors can increase because of a decrease in error sensitivity (Fig. 3C), or an increase in forgetting (Fig. 3B). Which factor changes due to increases in perturbation variance?

To answer this question, we estimated the forgetting rate (*a* in Eq. 1) of each participant by including an error-free movement period at the end of the experiments (gray region in Figs. 1C, 1D, 1F, and 2D). During these periods, behavior naturally decayed towards the baseline state (Fig. 3D), thus providing a means to isolate the rate of forgetting (i.e., the rate of decay of behavior). Interestingly, we found that in all experiments, the rate of forgetting was similar in the low and high variance groups (Fig. 3E, two-sample t-test; Exp. 1, p=0.72; Exp. 2, p=0.19; Exp. 5, p=0.79). Thus, perturbation variance did not alter the rate of forgetting.

Next, we empirically estimated error sensitivity (*b* in Eq. 1). To do this, we calculated the difference between the reach angle in each pair of consecutive trials (adjusting for forgetting) and divided this by the error experienced on the first of the two trials. By definition, this quotient represents one's sensitivity to error, i.e., the fraction of the error that is compensated for on the next trial. In sharp contrast to forgetting rates, we found consistent differences in sensitivity to error (Fig. 3F) among the various groups: in all experiments, participants who experienced a zero-variance perturbation exhibited an error sensitivity nearly twice that of individuals who were exposed to a high variance perturbation (two-sample t-test; Exp. 1, p=0.002, Cohen's d = 1.18; Exp. 2, p=0.039, Cohen's d = 0.87; Exp. 4, p=0.006, Cohen's d = 1.12; Exp. 5, p=0.016, Cohen's d = 1.05).

In summary, we asked whether perturbation variance affected forgetting rates, error sensitivity, or both. We found that in all experiments, the groups that had experienced high perturbation variance exhibited reduced error sensitivity, but no change in forgetting rates.

Perturbation variance reduced the ability to learn from small errors, not large errors

Our quantification of error sensitivity in Fig. 3F made the assumption that the brain is equally sensitive to errors of all sizes. However, it is well-documented that error sensitivity varies with magnitude of error: one tends to learn proportionally more from small errors[38–41]. In other words, error sensitivity is not constant, but declines as error size increases. How did perturbation variance alter the functional relationship between error magnitude and sensitivity to error?

To answer this question, we re-estimated error sensitivity, but this time controlled for the magnitude of error. We placed pairs of consecutive movements into bins according to the error experienced on the first trial, and then calculated error sensitivity within each bin. As expected, in both perturbation variance conditions, as error size increased, error sensitivity decreased (Fig. 4A; mixed-ANOVA, within-subjects effect of error size, F=24.69, p<0.001, η_p^2=0.30). This confirmed that indeed, people tended to learn less from larger errors. However, for a given error size, the high variance perturbation group exhibited lower error sensitivity than the zero-variance group (Fig. 4A; mixed-ANOVA, between-subjects effect of perturbation variance, F=10.9, p=0.002, η_p^2=0.16). Notably, increased perturbation variance reduced the ability to learn from small errors (<20°), but had no effect

on the larger errors (>20°) (Fig. 4A, post-hoc testing with t-test adjusted with Bonferroni correction, p=0.012 and Cohen's d = 0.75 for 5-10°, p=0.012 and Cohen's d = 0.78 for 10-15°, p=0.005 and Cohen's d = 0.83 for 15-20°, p=1 for other two bins). Why should increases in perturbation variance selectively affect learning from small errors, but not large errors?

The consistency of past errors modulated error sensitivity

A current model of sensorimotor adaptation[26] posits that the brain adjusts its sensitivity to error in response to the consistency of past errors. In this memory of errors model, when the error on trial *n* has the same sign as the error on trial *n*+1, it signals that the brain has undercompensated for error on trial *n*, and so should increase sensitivity to that error (Fig. 4B, left). Conversely, when the errors in two consecutive trials differ in sign, the brain has overcompensated for the first error, and so should decrease sensitivity to that error (Fig. 4B, right). Herzfeld and colleagues demonstrated that these changes in error sensitivity are local to specific errors, meaning that the brain can simultaneously increase sensitivity to one error size, while decreasing sensitivity to another[26]. Thus, in the context of a variable perturbation, the memory of errors model provides an interesting prediction: the perturbation affects the trial-to-trial consistency of errors, producing less consistency for some error sizes, particularly when perturbation variance increases.

We tested the predictions of this model by evaluating all consecutive trial pairs for each subject, and counting all instances of consistent and inconsistent errors. As expected, we found that in all five experiments, increased perturbation variance produced an increase in the probability of experiencing an inconsistent error (Figs. 4C; Exps. 1 & 3, p=0.017, Cohen's d = 0.68; Exp. 2, p<0.001, Cohen's d = 2.84; Exp. 4, p<0.001, Cohen's d = 2.22; Exp. 5, p=0.048, Cohen's d = 0.84). However, when we binned the data based on error size, the differences in the relative number of consistent and inconsistent errors exhibited a surprising pattern (Fig. 4D, mixed-ANOVA, between-subjects effect of perturbation variance, F=26.04, p<0.001, η_p^2=0.31; within-subjects effect of error size, F=61.02, p<0.001, η_p^2=0.51): for errors less than 20°, the zero-variance group had many more consistent error events and fewer inconsistent error events than the high variance group (Fig. 4D; post-hoc testing with t-test adjusted with Bonferroni correction, p<0.001 and Cohen's d = 1.99 for 5-10°, p<0.001 and Cohen's d = 1.47 for 10-15°, p=0.003 and Cohen's d = 0.78 for 15-20°). However, for errors larger than 20°, there was no difference in the relative number of consistent and inconsistent error pairs (Fig. 4D; post-hoc testing with t-test adjusted with Bonferroni correction, p=1 for 20-25° and p=0.161 for 25-30°).

In summary, the high variance perturbation led to a reduction in the trial-to-trial consistency of small errors, but not large errors (Fig. 4D). The memory of errors model (Fig. 4B) predicts that this pattern of error history will lead to a reduction in sensitivity to small errors for the high variance condition, but no difference in sensitivity to large errors. Indeed, both of these patterns were present in the measured behavior (Fig. 4A).

Error sensitivity changes throughout training according to the consistency of error

To determine if error sensitivity could truly have been altered by a memory of past errors, we next simulated a memory of errors model. This model expresses how error sensitivity varies both as a function of time, and error size. That is, on trial *n*, the sensitivity to error *e*, is expressed as $b^{(n)}(e)$.

Suppose that on trials *n*-1 and *n*, one experiences errors $e^{(n-1)}$ and $e^{(n)}$. The memory of errors model predicts that sensitivity to the first error will either increase or decrease, depending on the consistency of the two errors:

$$\Delta b^{(n+1)}\left(e^{(n-1)}\right) = \alpha \Delta b^{(n)}\left(e^{(n-1)}\right) + \beta \text{sign}\left(e^{(n)} e^{(n-1)}\right) \tag{2}$$

In the above equation, Δb represents the change in error sensitivity. Like Eq. 1, changes to error sensitivity are due to two forces: one that depends on consistency of consecutive errors (contained within the sign function), and one that imposes a decay on error sensitivity (parameter α). Thus, the model has only two unknown parameters, α and β.

To find these two parameters, we simulated the response of Eqs. (1) and (2) to the zero and high variance perturbation patterns we used in the various experiments (see Methods for a more complete model description; all other model parameters apart from α and β were obtained empirically from the data as in Figs. 3E and 3F). We varied our parameter set to identify a single pair of parameters (α and β) that best accounted for the measured reach angles in Experiments 3 and 4, in a least squares sense. The resulting parameter set yielded predictions that closely tracked subject behavior (Fig. 5A).

Having estimated the two unknown parameters of the model, we next applied Eq. 2 to predict how error sensitivity should have changed if the subjects stored a memory of past errors (Fig. 5B). To do this, we applied Eq. (2) to the actual sequence of errors experienced by each subject ($e^{(1)}$, $e^{(2)}$, etc.). Thus, the model predicted how error sensitivity should develop for each error size. To corroborate the model predictions, we measured error sensitivity on each trial (Fig. 5C).

We focused on Exp. 1 and 4, as these experiments included error-free periods, permitting empirical measurement of both error sensitivity and forgetting (note the difference in the ranges of the x-axes of Figs. 5B and 5C in the left and middle columns). In addition, we omitted measurements of error sensitivity over periods of the experiment where the number of observed data points was low or absent entirely (note gaps in time-courses in Fig. 5C left and right).

The model made the following predictions: First, the model predicted that with the zero-variance perturbation, error sensitivity should increase generally for all error sizes, but the extent of this increase should be greater for smaller errors (Fig. 5B, left). To test this prediction, for each subject we measured the change in error sensitivity across the first 60 epochs of Experiment 1. In the zero-variance group, error sensitivity indeed increased over the training period as predicted (Fig. 5D, left, zero variance, small errors. We could only estimate the change in sensitivity to small errors, because there were virtually no large errors at the end of the learning period).

Second, the model predicted that if perturbation variance was initially low, but then increased in the middle of adaptation (as in Exp. 4), error sensitivity should initially increase but then decline causing an increase in the residual errors (Fig. 5B, right). To test these predictions, for Exp. 4 we measured error sensitivity at the start of adaptation, at the end of the zero-variance period, and at the end of the high variance period. In accordance with model predictions, error sensitivity increased substantially over the initial zero-variance period (Fig. 5E, change from 1 to 80; paired t-test, p=0.003, Cohen's d = 0.96) and then dropped precipitously after increasing perturbation variance (Fig. 5E, data, change from 80 to 160; paired t-test, p=0.044, Cohen's d = 0.60).

Third, and most surprisingly, the model predicted that even if participants were exposed to a high variance perturbation from the onset, error sensitivity should increase for all errors, but only by a

small amount (Fig. 5B, middle). This final prediction was interesting because it implied that the residual errors in the high variance perturbation were larger not because of a decrease in error sensitivity, but instead because trial-by-trial increases in error sensitivity were stunted relative to the zero-variance group (compare Figs. 5B left and middle). Indeed, we found that error sensitivity also increased for the subjects in the high variance group (Fig. 5D, high var.), but the magnitude of this increase was smaller relative to the zero-variance group (Fig. 5D, two-sample t-test, $p<0.001$ and Cohen's $d = 1.65$ for small-small error comparison, $p<0.001$ and Cohen's $d = 1.91$ for small-large error comparison). The same was true for Exp. 4, when comparing error sensitivity at the end of the high variance period, to the initial error sensitivity at the start of the zero-variance period (Fig. 5E, change from 1 to 160; paired t-test, $p=0.042$, Cohen's $d = 0.60$).

In summary, the memory of errors model made the surprising prediction that error sensitivity should generally increase in the early phase of training, more so for small than large errors, and then saturate or even decline as the training trials continued. It also predicted that introducing variance into the perturbation should not decrease error sensitivity, but rather stunt its growth. Our measurements confirmed these predictions.

Overall, variance in the perturbation altered the sequence of trial-to-trial errors. A memory of these errors modulated error sensitivity, which in turn, resulted in different asymptotic limits to adaptation.

Discussion

Adaptation exhibits a curious property: even after prolonged training, learning appears to stop, leaving behind small persistent residual errors[2,4,5,7]. Here we demonstrated that this asymptotic limit of adaptation is consistent with a learning system in which sensitivity to error is not constant, but rather changes as a function of the history of past errors. When the past errors are temporally consistent, which occurs when perturbation variance is low, error sensitivity rises quickly. However, as performance improves and errors become smaller, error sensitivity peaks and then declines. The history-dependent change in error sensitivity, which in turn is driven by the variance of the perturbation, appears to play a causal role in setting an upper bound on adaptation. Thus, errors are not simply used for learning and then discarded, but rather they appear to become a form of memory that describes how much the nervous system should learn from that error.

Controlling the total extent of adaptation through modulation of error sensitivity

Earlier studies that documented modulation of error sensitivity focused on understanding the rate of motor adaptation[16,42] in the context of savings[7,13,26,27,43–45], meta-learning[26,45], and anterograde interference[46]. Here we describe yet another feature of adaptation that appears to be controlled through error sensitivity: the total extent of motor adaptation. In Experiments 3 and 4 (Fig. 1) we demonstrated that differences in residual errors were coupled to differences in error sensitivity, but not the rate of forgetting. The balance between error sensitivity, which changed with perturbation variance, and forgetting, which remained constant, created different levels of asymptotic performance.

This observation raised an interesting question: why didn't the brain continue to increase its error sensitivity until residual errors were completely eliminated? Here we proposed a potential solution: increases in error sensitivity are bounded by decay. Our model (Eq. 2) accounts for saturation

in error sensitivity in the same way the state-space model (Eq. 1) accounts for saturation in adaptation: through forgetting. The counterbalancing of a memory of consistent errors (which increases error sensitivity) with decay (which decreases error sensitivity) leads to steady-state saturation in error sensitivity.

The addition of forgetting to the original memory of errors model[26] offers a potential explanation for at least three separate behavioral observations. First, when a participant is exposed to the same perturbation twice, they exhibit faster adaptation during the second exposure[11]. To account for this savings, the memory of errors model predicts that error sensitivity is up-regulated during the first exposure to the perturbation, leading to faster learning in the future. However, in some cases, long periods of washout or error-free trials appear to prevent the occurrence of savings[43,47]. This observation is consistent with Eq. 2: during sufficiently long periods of washout, error sensitivity would decay towards baseline due to forgetting, thereby preventing savings upon the next exposure to the perturbation.

Another peculiar feature of savings is the observation that the mean residual error is often indistinguishable during the first and second blocks of exposure to a perturbation, even though the initial rate of learning is faster during the second block[13,47,48]. Why doesn't the faster rate of learning during the second exposure also lead to a higher extent of learning? Our model provides an explanation: because error sensitivity tends to saturate during the initial exposure to the perturbation, behavior will reach the same asymptotic level during the second exposure despite presence of savings and higher initial rates of learning.

Finally, experiments that employ a constant error-clamp condition also demonstrate saturation in error sensitivity. In these experiments[6,39,40,49,50] subjects are exposed to the same error time and time again, irrespective of their motor output. Without decay, a memory of errors model would predict that error sensitivity, and thus the extent of learning, would increase without bound. However, in reality, learning reaches a steady-state in this constant error condition. In the context of visuomotor adaptation, an earlier report speculated that this saturation of learning implies a limitation in the amount of adaptation that can be stored within implicit motor learning systems[39]. However, more generally, this saturation in adaptation is also predicted by Eq. 2 due to the eventual balancing of learning and forgetting.

Plasticity in implicit learning systems

Motor adaptation is known to be supported by both implicit (subconscious) and explicit (cognitive) corrective systems[30–34]. Recent studies have interrogated the flexibility of these learning systems, with some[23,51–54] suggesting that implicit processes have a response to error that does not change with multiple exposures to a perturbation.

While our primary purpose was not to carefully tease apart implicit and explicit contributions to adaptation, we did explore this issue by limiting reaction time, which is thought to substantially reduce the contributions of explicit learning[22–25]. When reaction time was limited, the residual errors were not only present, but increased by the same amount if perturbation variance increased. This observation suggested that perturbation variance acted on the implicit learning system. These results argue against the idea that implicit learning systems generate inflexible and unchanging responses to sensorimotor errors.

This finding does not discount the possibility that explicit systems also contribute to residual errors. For example, recent work has demonstrated that declines in explicit learning associated with aging[55,56] also lead to differences in the total extent of adaptation[57]. It is tempting to suggest that these differences in asymptotic performance may also reflect a change in error sensitivity, like the implicit mechanism described herein. While this possibility remains untested, such an interpretation is consistent with recent work demonstrating that environment variability modulates explicit error sensitivity during adaptation to random walk perturbations[54].

In summary, our data support the more inclusive view that with experience, both implicit and explicit processes change the way they respond to error, and together determine the total extent of sensorimotor adaptation.

Alternate models

Perturbation variance can affect uncertainty of the learner. Over the past few decades, numerous studies[12,16,17] have used a Kalman filter[18] to study the relationship between uncertainty and learning rate. The Kalman filter describes the optimal way in which an observer should adjust their rate of learning in response to different sources of variability. This Bayesian framework has proved useful in understanding the slowing of adaptation in response to reductions in the reliability of sensory feedback[19–21], acceleration of adaptation in response to uncertainty in the state of the individual or environment[12,17,21], and even the optimal tuning of adaptation rates in individual subjects[14].

Could this Bayesian framework also account for our results? Indeed, it is possible that the inconsistent errors generated by the high variance perturbation are interpreted as an observation noise, thus decreasing the brain's confidence in its sensory feedback. In the Kalman framework, this would correctly predict that adaptation to the high variance perturbation would proceed more slowly than adaptation to the low variance perturbation (Figs. 1-2).

With that said, two of our behavioral observations do not easily fit into the Kalman framework. First, as expected, we observed that in both the zero-variance and high variance groups, error sensitivity declined as a function of error size[38–41]. However, across the variance conditions, differences in error sensitivity were specific to small errors, not large errors (Fig. 4A). The memory of errors model[26] provided a way to understand this pattern: differences in perturbation variability led to changes in the consistency of small errors, but not large errors (Fig. 4D) matching the pattern we observed in error sensitivity. It is unclear how to account for this phenomenon with a Kalman filter which responds to all error sizes with the same learning rate.

The second observation that does not easily comply with the Kalman view is the fact that sensitivity to error increased over its baseline level with repeated exposure to the high variance perturbation (Fig. 5D, high var.; Fig. 5E, change from 1 to 160). For a Kalman filter, the Kalman gain (i.e., rate of learning) would necessarily drop from its baseline level (where there is no variance) when the uncertainty in its observations increases (during the perturbation block where variance is added). This prediction of the Kalman framework is opposite to our data. In contrast, the memory of errors model (Figs. 5B-E) correctly predicts an increase in error sensitivity, because the experience of consistent errors remains more probable than inconsistent errors even in the high variance condition (Figs. 4C and 4D).

Neural basis of error sensitivity

Motor adaptation depends critically on the cerebellum[37,58–61], where Purkinje cells learn to associate efference copies of motor commands with sensory consequences[62]. This learning is guided by sensory prediction errors, which are transmitted to the Purkinje cells via the inferior olive, resulting in complex spikes. Notably, plasticity in Purkinje cells exhibits both sensitivity to error, and forgetting. Sensitivity to error is encoded via probability of complex spikes: in each Purkinje cell, the probability of complex spikes is greatest for a particular error vector[62,63], and this tuning modulates learning from error. Forgetting is present in the time-dependent retention of the plasticity caused by the complex spikes[64,65], resulting in decay of plasticity with passage of time. Thus, the presence of error sensitivity and forgetting in the plasticity of Purkinje cells provides one mechanism by which cerebellar-dependent adaptation can exhibit asymptotic performance with non-zero residual errors.

Our behavioral experiments demonstrated that during adaptation, error sensitivity increased with training, but this increase was suppressed if the perturbations were variable (Figs. 5D and 5E). Our model (Figs. 5B and 5D) suggested that these changes in error sensitivity may have arisen from the temporal consistency of errors such that the presence of two errors in the same direction would result in increased sensitivity. Thus, the theory makes the interesting prediction that in the framework of cerebellar learning, the temporal proximity of complex spikes might modulate error sensitivity. Specifically, when two consecutive errors of similar direction and magnitude occur, Purkinje cells that prefer that error are more likely to experience complex spikes in close temporal proximity. The model predicts that the temporal proximity of these complex spikes would result in up-regulation of error sensitivity in that Purkinje cell. This idea remains to be tested.

In this framework, Eq. 2 serves as a simple abstraction of a more complicated reality. That is, in contrast to the binary output of the sign function in Eq. 2, the consistency of error should vary along a spatial continuum of preferred and non-preferred errors, as well as a temporal continuum that separates error events. These predictions provide a platform for future experiments, both in the context of behavior and neurophysiology.

In summary, our work suggests that over the course of training, the brain relies on a memory of past errors to adjust error sensitivity. These variations in error sensitivity produce an asymptotic limit that prevents further improvements in performance, thus resulting in persistent residual errors.

Methods

Participants

A total of 117 volunteers participated in our experiments. All experiments were approved by the Institutional Review Board at the Johns Hopkins School of Medicine. In addition, we re-analyzed an earlier study with 16 participants[8].

Apparatus

In Experiments 1-5, participants held the handle of a planar robotic arm (Fig. 1A) and made reaching movements to different target locations in the horizontal plane. The forearm was obscured from view by an opaque screen. An overhead projector displayed a small white cursor (diameter = 3mm) on the screen that tracked the motion of the hand. Throughout testing we recorded the position of the robot handle using a differential encoder with submillimeter precision. We also recorded the forces produced on the handle by the subject using a 6-axis force transducer. Data were recorded at 200 Hz.

Visuomotor rotation

Experiments 1, 3, 4, and 5 followed a similar protocol. At the start of each trial, the participant brought their hand to a center starting position (circle with 1 cm diameter). After maintaining the hand within the start circle, a target circle (1 cm diameter) appeared in 1 of 4 positions (0°, 90°, 180°, and 270°) at a displacement of 8 cm from the starting circle. Participants then performed a "shooting" movement to move their hand briskly through the target. Each experiment consisted of epochs of 4 trials where each target was visited once in a pseudorandom order.

Participants were provided audiovisual feedback about their movement speed and accuracy. If a movement was too fast (duration < 75 ms) the target turned red. If a movement was too slow (duration > 325 ms) the target turned blue. If the movement was the correct speed, but the cursor missed the target, the target turned white. Successful movements (correct speed and placement) were rewarded with a point (total score displayed on-screen), an on-screen animation, and also a pleasing tone (1000 Hz). If the movement was unsuccessful, no point was awarded and a negative tone was played (200 Hz). Participants were instructed to obtain as many points as possible throughout the experimental session.

Once the hand reached the target, visual feedback of the cursor was removed, and a yellow marker was frozen on-screen to provide static feedback of the final hand position. At this point, participants were instructed to move their hand back to the starting position. The cursor continued to be hidden until the hand was moved within 2 cm of the starting circle. In most experiments, participants actively moved their hand back to the start position. However, in Experiments 3 and 5, the robot assisted the subject if their hand had not returned to the start position after 1 second.

Movements were performed in one of three conditions: null trials, rotation trials, and no feedback trials. On null trials, veridical feedback of hand position was provided. On rotation trials, once the target appeared on screen (the cue for the reaching movement), the on-screen cursor was rotated relative to the start position (Fig. 1A). Each rotation experiment terminated with a period of no feedback trials. On these trials, the subject cursor was hidden during the entire trial. In addition, the subject was provided no feedback regarding movement endpoint, accuracy, or timing.

As a measure of adaptation, we analyzed the reach angle on each trial. The reaching angle was measured as the angle between the line segment connecting the start and target positions, and the line

segment connecting the start and final hand position. The final hand position was taken to be the location of the hand right after the hand exceeded 95% of the target displacement. The sign of the reaching angle was determined by whether the final hand position was rotated CW or CCW relative to the line connecting the start and target positions. For analysis of reaching errors, we computed the same quantity, but for the final cursor position rather than the final hand position. On perturbation trials, these two quantities are related by the rotation that occurred on that trial.

Force field adaptation

In Experiment 2, participants were perturbed by a velocity-dependent force field (Fig. 1A), as opposed to a visuomotor rotation. At trial onset, a circular target (diameter = 1 cm) appeared in the workspace, coincident with a tone that cued subject movement. Participants then reached from the starting position to the target. The trial ended when the hand stopped within the target location. After stopping the hand within the target, movement timing feedback was provided. If the preceding reach was too slow, the target turned blue and a low tone was played. If the reach was too fast, the target turned red and a low tone was played. If the reach fell within the desired movement interval (450-550 ms), the subject was rewarded with a point (total score displayed on-screen), an on-screen animation, and also a pleasing tone (1000 Hz). Participants were instructed to obtain as many points as possible throughout the experimental session. After completing each outward reaching movement, participants were instructed to then bring their hand back to the starting position. This return movement was not rewarded and was always guided by a "channel" (see description below in this section).

As in the rotation experiments, the target appeared in 1 of 4 positions (0°, 90°, 180°, and 270°) at a displacement of 10 cm from the starting circle. Each experiment consisted of epochs of 4 trials where each target was visited once in a pseudorandom order. The experiment began with a set of null field trials (no perturbations from the robot). After this period, participants were exposed to a force field. The force field was a velocity-dependent curl field (Fig. 1A) in which the robot generated forces proportional and perpendicular to the velocity of the hand according to:

$$\begin{bmatrix} f_x \\ f_y \end{bmatrix} = b \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} v_x \\ v_y \end{bmatrix} \tag{3}$$

where v_x and v_y represent the x and y velocity of the hand, f_x and f_y represent the x and y force generated by the robot on the handle, and b represents the magnitude of the force field. Note that the sign of b determines the orientation of the field (the field if CW is $b > 0$ and CCW if $b < 0$).

Subject reaching forces were measured on designated "channel" trials[29] where the motion of the handle was restricted to a linear path connecting the start and target locations (Fig. 1A). To restrict hand motion to the straight-line channel trajectory, the robot applied perpendicular stiff spring-like forces with damping (stiffness = 6000 N/m, viscosity = 250 N-s/m). Reaching forces were measured on every 5th epoch of movements with a cycle of 4 channel trials, one for each target. Additionally, the experiment terminated with a block of channel trials to measure how well the adapted state of each participant was retained over time.

Offline we isolated forces produced on each trial against the channel wall, perpendicular to the direction of the primary movement. To do this, we calculated the average force during baseline channel trials. We then subtracted this baseline force timeseries from all of the force timeseries recorded during

channel trials throughout the experiment. After correcting for this baseline force, we then computed an adaptation index. The adaptation index represents the scaling factor relating the force produced on a given trial and the ideal force the subject would produce if they were fully adapted to the perturbation[11]. To calculate this scaling factor, we linearly regressed the ideal force timecourse onto the actual force timecourse. The ideal force was calculated by multiplying the velocity timecourse of the movement by the perturbation magnitude. The adaptation index served as our behavioral measurement for model fitting and parameter estimation.

In addition to analyzing the forces produced on channel trials, we also analyzed the trajectory of the hand on perturbation trials. From each trajectory we isolated a signed movement error, which we used to calculate the probability that an error switched sign from one trial to the next (Fig. 4C, Exp. 2). To calculate the movement error, we isolated the portion of each reaching movement between 20% and 90% of target displacement. Within this region we detected the maximum absolute error and treated this as the error magnitude. We signed this error according to whether the hand was to the left or right (or top or bottom) of the line connecting the start position and target position. To prevent minor overcompensations for the force field magnitude from being treated as movement errors, deviations that fell within 3 mm of the line connecting the start and target locations were not treated as errors. Using smaller thresholds of 1 or 2 mm did not qualitatively affect our results.

Statistics

In this work, we employed several types of statistical tests: repeated measures ANOVA, two-way ANOVA, and mixed-ANOVA. These tests were carried out in IBM SPSS 25. In all cases we report the p-value, F-value, and η_p^2 for each test. For post-hoc testing we employed t-tests with Bonferroni corrections. For these tests, we report the p-value and Cohen's d as a measure of effect size. Our mixed-ANOVA contained a between-subjects factor and a within-subjects repeated measure. For the within-subjects repeated measure, data are binned within small windows defined by differences in error size. In the event that a participant is missing data within a bin (data are missing in approximately 15% of all bins), we replaced the missing data point with the mean of the appropriate distribution.

Experiment 1

We tested how variance in the perturbation affected the total extent of visuomotor adaptation. The experiment started with 10 epochs (40 trials) of no perturbation. After this a perturbation period began that consisted of 60 rotation epochs (240 trials total). At the end of the perturbation period, retention of the visuomotor memory was tested in a series of 15 epochs (60 trials) of no feedback.

To test the effect of perturbation variance on behavior, participants were divided into 1 of 2 groups. In the zero-variance group, participants (n=19) were exposed to a constant visuomotor rotation of 30°. In the high variance group, participants (n=14) were exposed to a visuomotor rotation that changed on each trial. The rotation was sampled from a normal distribution with a mean of 30° and a standard deviation of 12°.

Experiment 2

We found that perturbation variance reduced the total amount of adaptation in Experiment 1. To test if this impairment was a general property of sensorimotor adaptation, we tested another group of

subjects with a force field. The experiment started with 10 epochs (40 trials) of no perturbation (2 of these epochs were channel trials). After this a perturbation period began that consisted of 75 epochs (300 trials, 20% were channel trials) of force field perturbations. At the end of the perturbation period, retention of the adapted state was tested in a series of 10 epochs (40 trials) of channel trial movements.

To test the effect of perturbation variance on behavior, participants were divided into 1 of 2 groups. In the zero-variance group, participants (n=12) were exposed to a constant force field magnitude of 14 N-s/m. In the high variance group, participants (n=13) were exposed to a force field magnitude that changed on each trial. The force field magnitude was sampled from a normal distribution with a mean of 14 N-s/m and a standard deviation of 6 N-s/m.

Experiment 3

Inspection of the learning curves in Experiment 1 indicated that performance may not have completely saturated by the end of the perturbation period. Therefore, to confirm that perturbation variance induces different performance saturation levels, we repeated Experiment 1, but this time more than doubled the number of perturbation trials. The experiment started with 5 epochs (20 trials) of no perturbation. The following perturbation period consisted of 160 rotation epochs (640 trials). As in Experiment 1, participants were divided into a zero-variance group (n=10) and a high variance group (n=10). Perturbation statistics remained identical to Experiment 1.

Experiment 4

To determine if perturbation variance causally altered the total extent of adaptation, we designed a control experiment. In this experiment, participants started with a visuomotor rotation in the zero-variance condition, and then after reaching asymptotic performance, were exposed to the high variance condition. If variance causally determined the total amount of learning, we expected that asymptotic performance would decrease after the addition of variability to the perturbation.

Participants (n=14) began the experiment with 5 epochs (20 trials) of null trials. After this, the zero-variance period started. Participants were exposed to either a CW or CCW visuomotor rotation of 30° for a total of 80 epochs (320 trials). At the end of this period, participants switched to a high variance condition where the rotation was sampled on each trial from a normal distribution with a mean of 30° and a standard deviation of 12°. This period lasted for an additional 80 epochs (320 trials). Finally, the experiment concluded with 15 epochs (60 trials) of no feedback.

Experiment 5

Sensorimotor adaptation is supported by both explicit strategy and implicit learning[31]. To determine which of these types of learning were impaired by perturbation variance we performed an experiment where we limited the time participants had to prepare their movements. Limiting reaction time is known to suppress explicit strategy[22].

To limit reaction time, we instructed participants to begin their reaching movement as soon as possible, after the target location was revealed. To enforce this, we limited the amount of time available for the participants to start their movement after the target location was shown. This upper bound on reaction time was set to either 225, 235, or 245 ms (taking into account screen delay). To enforce the desired preparation time, if the reaction time of the participant exceeded the desired upper bound, the

644 participant was punished with a screen timeout after providing feedback of the movement endpoint. In
645 addition, a low unpleasant tone (200 Hz) was played, and a message was provided on screen that read
646 "React faster". As in Experiment 1, participants were divided into a zero-variance perturbation group
647 (n=13) and a high variance group (n=12). All other details were identical to Experiment 1.
648
649 *Re-analysis of prior work*
650 In Fig. 1B, we reference earlier work from a study by Fernandes and colleagues[8]. The experiment
651 methodology is fully described in their original manuscript. Briefly, participants (n=16) made a center-
652 out reaching movement to a target. After the movement ended, participants were shown the endpoint
653 location of an otherwise hidden cursor that tracked the position of the right index finger.
654 Participants performed three experimental blocks. Each block had the same general structure.
655 At the start of the block, participants made 40 reaching movements to 8 different targets (5 for each
656 target) with continuous visual feedback of the cursor. Next, participants made an additional 80 reaching
657 movements to 8 different targets (10 for each target) using only endpoint feedback of the cursor
658 position. After this baseline period, a single target position was selected, and 240 reaching movements
659 were performed under the influence of a visuomotor rotation. The visuomotor rotation was sampled on
660 each trial from a normal distribution with a mean of 30° and a standard deviation of either 0°, 4°, or 12°.
661 The block ended in a set of 160 generalization trials that are not relevant to the current study.
662 The experiment had a within-subject design. Each participant was exposed to all three
663 perturbation variances, but in a random order. The orientation of the rotation (CW or CCW) was
664 randomly chosen on each block. In addition, the target selected during the adaptation period was
665 randomly chosen from 1 of the 4 diagonal targets on each block.
666
667 *State-space model of learning*
668 After the experience of a movement error, humans and other animals change their behavior on future
669 trials. In the absence of error, adapted behavior decays over time. Here we used a state-space model [66]
670 to capture this process of error-based learning. Here, the internal state of an individual *x*, changes from
671 trials *n* to *n*+1 due to learning and forgetting.
672 $$x^{(n+1)} = ax^{(n)} + b^{(n)}e^{(n)} + \varepsilon_x^{(n)} \quad (4)$$
673 Forgetting is controlled by the retention factor *a*. The rate of learning is controlled by the error
674 sensitivity *b*. We describe modulation of error sensitivity in a later section. Learning and forgetting are
675 stochastic processes affected by internal state noise ε_x: a normal random variable with zero-mean and
676 standard deviation of σ_x.
677 While we cannot directly measure the internal state of an individual, we can measure their
678 movements. The internal state *x* leads to a movement *y* according to:
679 $$y^{(n)} = x^{(n)} + \varepsilon_y^{(n)} \quad (5)$$
680 The desired movement is affected by execution noise, represented by ε_y: a normal random variable
681 with zero-mean and standard deviation of σ_y.

To complete the state-space model in Eqs. 3 and 4, we must operationalize the value of an error, *e*. In sensorimotor adaptation, movement errors are determined both by motor output of the participant (*y*) and the size of the external perturbation (*r*):

$$e^{(n)} = r^{(n)} - y^{(n)} \tag{6}$$

In our studies, the perturbation took the form of either a visuomotor rotation or a velocity-dependent force field. Eq. 6 clearly demonstrates why perturbation variance influences error variance.

Asymptotic properties of learning

State-space models of learning predict that performance can saturate despite presence of residual errors. This saturation is caused by a steady state condition where the amount of learning from error is exactly counterbalanced by the amount of forgetting (Fig. 3A). The steady state can be derived from Eqs. 4-6:

$$y_{ss} = \frac{b\bar{r}}{1-a+b} \tag{7}$$

The formula for steady-state adaptation (y_{ss}) shows that one's learning extent depends on 3 factors: (1) error sensitivity *b*, (2) retention factor *a*, and (3) the mean of the perturbation $\bar{r}$. If there is no forgetting (a = 1), an individual will adapt completely to the mean of the perturbation. However, if retention is incomplete (a < 1), the steady state behavior (y_{ss}) will always fall short of the mean of the perturbation, resulting in residual errors.

Eq. 7 is important for three reasons. (1) It demonstrates why the total extent of learning varies with a change in forgetting rate (Fig. 3B). (2) It demonstrates why the total extent of learning varies with a change in error sensitivity (Fig. 3C). (3) It demonstrates that the total amount of learning does not directly depend on variability in the perturbation, only the mean of the perturbation (Fig. 3A).

Calculation of the retention factor

To determine if differences in learning extent were caused by a change in the rate of forgetting, we estimated the retention factor (*a*) of each participant. To do this, we quantified how behavior decayed during the error-free periods that terminated Experiments 1, 2, 4, and 5 (Figs. 1-3). During these error-free periods, trial errors were either hidden (no feedback condition in visuomotor rotation experiments) or fixed to zero (channel trials in the force field adaptation experiment). In the absence of error (e=0), our state-space model simplifies to exponential decay (omitting noise terms):

$$y^{(n)} = a^{n-m} y^{(m)} \tag{8}$$

Eq. 8 relates the motor output (*y*) on trial *n* of the error-free period to the initial motor behavior measured at the end of the adaptation period, $y^{(m)}$. The term $n-m$ represents the number of trials that elapsed from the start of the error-free period until the current trial *n*.

For visuomotor rotation experiments, we estimated the retention factor separately for each target by fitting Eq. 8 to subject behavior in the least-squares sense. We report the mean retention factor in Fig. 3E. For force field adaptation, we estimated a single retention factor, by first averaging the adaptation index across the 4 targets in each epoch, and then fitting Eq. 8 to the epoch-by-epoch behavior in the least-squares sense. In Fig. 3E, we converted this epoch-based retention factor to a trial-

based retention factor by raising the epoch-based retention factor to the power of 1/4 (an epoch of 4 trials has 4 trial-by-trial decay events).

Calculation of error sensitivity

Using Eq. 8, we found that changes in learning saturation were not caused by modulation of forgetting rates. Next, we determined how variability impacted error sensitivity (*b*), using its empirical definition:

$$b^{(n_1)} = \frac{y^{(n_2)} - a^{n_2 - n_1} y^{(n_1)}}{e^{(n_1)}} \tag{9}$$

Eq. 9 determines the sensitivity to an error experienced on trial n_1 when the participant visited a particular target T. This error sensitivity is equal to the change in behavior between two consecutive visits to target T, on trials n_1 and n_2 (i.e., there are no intervening trials where target T was visited) divided by the error that had been experienced on trial n_1. In the numerator, we account for decay in the behavior by multiplying the behavior on trial n_1 by a decay factor that accounted for the number of intervening trials between trials n_1 and n_2. For each target, we used the specific retention factor estimated for that target with Eq. 8.

We used Eq. 9 to calculate error sensitivity for all of our visuomotor rotation experiments. When reporting error sensitivity, we averaged across the four targets (Figs. 3F, 4A, 5C, 5D, and 5E). In some cases (Fig. 3F) we collapsed trial-by-trial measurements of error sensitivity across all trials and all errors. In other cases, we calculated the change in error sensitivity over different periods of training. For Fig. 5D, we measured the change in sensitivity from the beginning (epochs 1-15) to the end (epochs 40-59) of the perturbation block in Exp. 1. For this, we calculated two error sensitivities, one for errors less than 20° and the other for errors greater than 20°. To remove outliers, we identified error sensitivity estimates that deviated from the population median by over two median absolute deviations. We used a similar process for our analysis of Exp. 4 in Fig. 5E. Here we had three periods of interest, the start of adaptation (epochs 1-3), the end of the zero-variance period (epochs 78-80), and the end of the high variance period (epochs 157-159). Our second to last method for analyzing error sensitivity is shown in Fig. 4A. Here, we calculated sensitivity to errors of specific sizes. For this, we separated trial pairs into bins that depended on error size: 5°-10°, 10°-15°, 15°-20°, 20°-25°, and 25°-30°. For each subject-error bin pair, we required there to be at least 10 measurements. We did not consider errors smaller than 5° because the empirical estimator in Eq. 9 becomes unstable for small error sizes. Our final method of analyzing error sensitivity is shown in Fig. 5C. Here we calculated error sensitivity both as a function of error size and also trial number. These measurements are particular noisy because on any given trial, only a subset of participants experienced errors of a given size. Therefore, to reduce this noise, we calculated error sensitivity within bins of consecutive trials. For the zero-variance perturbation we generally included 5 epochs (20 trials) in each bin. For the high variance perturbation (Fig. 5C middle) we included 10 epochs in each bin.

For force field adaptation, we could not empirically estimate error sensitivity, as this approach requires the measurement of forces directly before and after the experience of an error. However, in reality, forces are measured only on infrequent channel trials, making such an empirical calculation impossible. For this reason, we used a model-based approach to measure error sensitivity (Fig. 3F, Exp. 2). We fit our state-space model Eqs. 4-6 to single subject data in the least-squares sense, over the last 5

channel trial epochs of the adaptation period. To do this, we needed to describe four states of learning (one for each target). We describe multitarget state-space models in more detail in an earlier work [66]. As a brief summary, we modeled our multitarget experiment by applying Eqs. 4-6 separately for each target. On any given trial, the state corresponding to the relevant target learned from the error on that trial. The other three states exhibited only decay on that trial. We described the perturbation *r* in terms of the force field magnitude on that trial (14 N-s/m was considered a perturbation of unit 1 in the model). Using this framework, we found the error sensitivity that minimized the squared difference between our model simulation and participant behavior.

Memory of errors model

Using Eq. 9, we found that the capacity for learning changed in the different perturbation environments due to a modulation in error sensitivity (Fig. 3F). Curiously, error sensitivity differed for errors of certain sizes, not for all errors (Fig. 4A). To account for these findings, we used a memory of errors model that was recently proposed by Herzfeld and colleagues[26].

This model uses a simple normative framework. When the errors on trial *n* and trial *n*+1 have the same sign (a consistent error), this signals that the brain under-corrected for the first error (Fig. 4B). Therefore, the brain should increase its sensitivity to the initial error. On the other hand, when the errors on trials *n* and *n*+1 have opposite signs (an inconsistent error), this signals that the brain over-corrected for the first error. Therefore, the brain should decrease its sensitivity to the initial error. These rules are encapsulated by Eq. 2. Note that unlike the original memory of errors model, here, we allow for decay in error sensitivity over time through the decay factor α.

In our simulations, we applied Eq. 2 to errors of different sizes. We divided up the error space into 5° error bins. We assumed that all errors in that bin shared the same error sensitivity (Fig. 5B), but sensitivity could differ across bins.

Simulation of the memory of errors model

In Fig. 5, we asked if our memory of errors model (Eq. 2) would accurately predict the behavior of our subject population. To answer this question, we first fit the free parameters in Eq. 2 to our data (Fig. 5A), and then used the fully specified model to simulate error sensitivity patterns (Fig. 5B). These free parameters included two terms: α and β in Eq. 2.

We fit these two model parameters to the mean behavior in Experiments 3 and 4, in the least-squares sense. We focused on Experiments 3 and 4, as these experiments included the longest training periods, and therefore produced reliable steady-state adaptation limits. We fit the model to the combined behavior of the zero-variance group in Experiment 3 (Fig. 5A, left), the high variance group in Experiment 3 (Fig. 5A, middle), and the zero-to-high variance group in Experiment 4 (Fig. 5A, right). In this way, we identified a single parameter set that minimized the sum of squared errors across all three groups. The error we considered was the difference between mean subject reaching angles and the mean reaching angle predicted through simulation of our model. By model, we are referring to the reaching angle predicted by our state-space model (Eqs. 4 and 5) that used an error sensitivity that started at 10% (chosen to match the initial error sensitivity empirically measured in subject behavior; see curves in Fig. 5C) and then varied according to a memory of errors (Eq. 2). Our model assumed that error sensitivity was common to errors within bins of 5° but could vary across bins (as in Fig. 4A). We

spaced these bins from -45 to 45°. For our simulations, we used a retention factor of a = 0.9736 (the mean retention factor measured in the error-free period in Experiment 1).

Due to variability in the process of learning, moving, and the high variance perturbation, the output of our model was stochastic. Therefore, to obtain the model prediction for a given set of free parameters, we calculated the mean output of the model over 10,000 simulations. For each simulation, we resampled the trial-to-trial variations in the high variance perturbation. We also resampled trial-to-trial variations in the state of the learner (ε_x in Eq. 4) and the motor output (ε_y in Eq. 5). For these simulations, we set σ_x (Eq. 4) and σ_y (Eq. 5) both equal to 2°. We chose this variability level by calculating the standard deviation of the last 100 reaching angles in the zero-variance group of Experiment 3 (median standard deviation was 4.1°, and we divided this up evenly for state noise and motor noise to arrive at our 2° estimate).

To identify the optimal parameter set we first attempted to use *fmincon* in MATLAB R2019a. We found however, that *fmincon* yielded different parameter estimates with each change to the algorithm's initial conditions. Therefore, to confirm that we identified a global, rather than a local minimum, in the squared-error cost function, we performed a secondary grid search in the proximity of the parameter sets identified by *fmincon*. For this grid search, we tested all 2,601 pairwise combinations of α and β, where α was varied from 0.95 to 1 in increments of 0.001, and β was varied from 0.03 to 0.08 in increments of 0.001. This grid search identified an optimal parameter set of α = 0.987 and β = 0.042.

We used this parameter set for our simulations in Figs. 5B-E. In Fig. 5A, we simulated our model a total of 100,000. In Figs. 5C-5E we combined our model parameters with the actual error sequences experienced by individual participants. In Fig. 5B, we tracked subject errors in each 5° bin, and used Eq. 2 to predict if error sensitivity should increase or decrease from one trial to the next. For the left and middle insets, we combined subjects across Experiments 1 and 3. However, Experiment 1 ended at epoch 60, so after this point, only subjects in Experiment 3 are represented (hence the change in the size of the error bars after the vertical black line in the left and middle insets in Fig. 5B). To transition between epoch 60 and 61, and maintain continuity in our predictions, we used a bootstrapping approach. For each bootstrap, we sampled the initial error sensitivity on epoch 61 in each error bin from a normal distribution with a mean and variance that were determined from the model predictions on epoch 60. With that particular set of error sensitivities, we simulated the response from epochs 61 to 160 for subjects in Experiment 3. We repeated this procedure a total of 1,000 times, each time resampling the initial error sensitivity on epoch 61.

Finally, to obtain the model predictions in Figs. 5D and 5E, we used the error sensitivity timecourses depicted in Fig. 5B. In Fig. 5D, we focused on participants in Experiment 1, where we included an error-free period, thus permitting robust measurement of error sensitivity and forgetting rates. In all cases, we measured the differences in error sensitivity for small (between 5° and 20°) and large (between 20° and 30°) errors from the start (epoch 1) to the last epoch (epoch 60), predicted by the memory of errors model (Fig. 5C). We used a similar method in Fig. 5E, only for Experiment 4. Here we measured the change in error sensitivity for all errors (5° to 30°) over three separate periods. From the beginning of the zero-variance period to the end of the zero-variance period (Fig. 5E, change from 1 to 80), from the end of the zero-variance period to the end of the high variance period (Fig. 5E, change from 80 to 160), and from the beginning to the end of the experiment (Fig. 5E, change from 1 to 160). In

both Figs. 5D and 5E, we obtained single estimates of error sensitivity, by collapsing across the appropriate 5° error sensitivity bins. When collapsing across bins, we weighted each bin by the number of times errors in that bin were experienced by the subject population, over the corresponding time window.

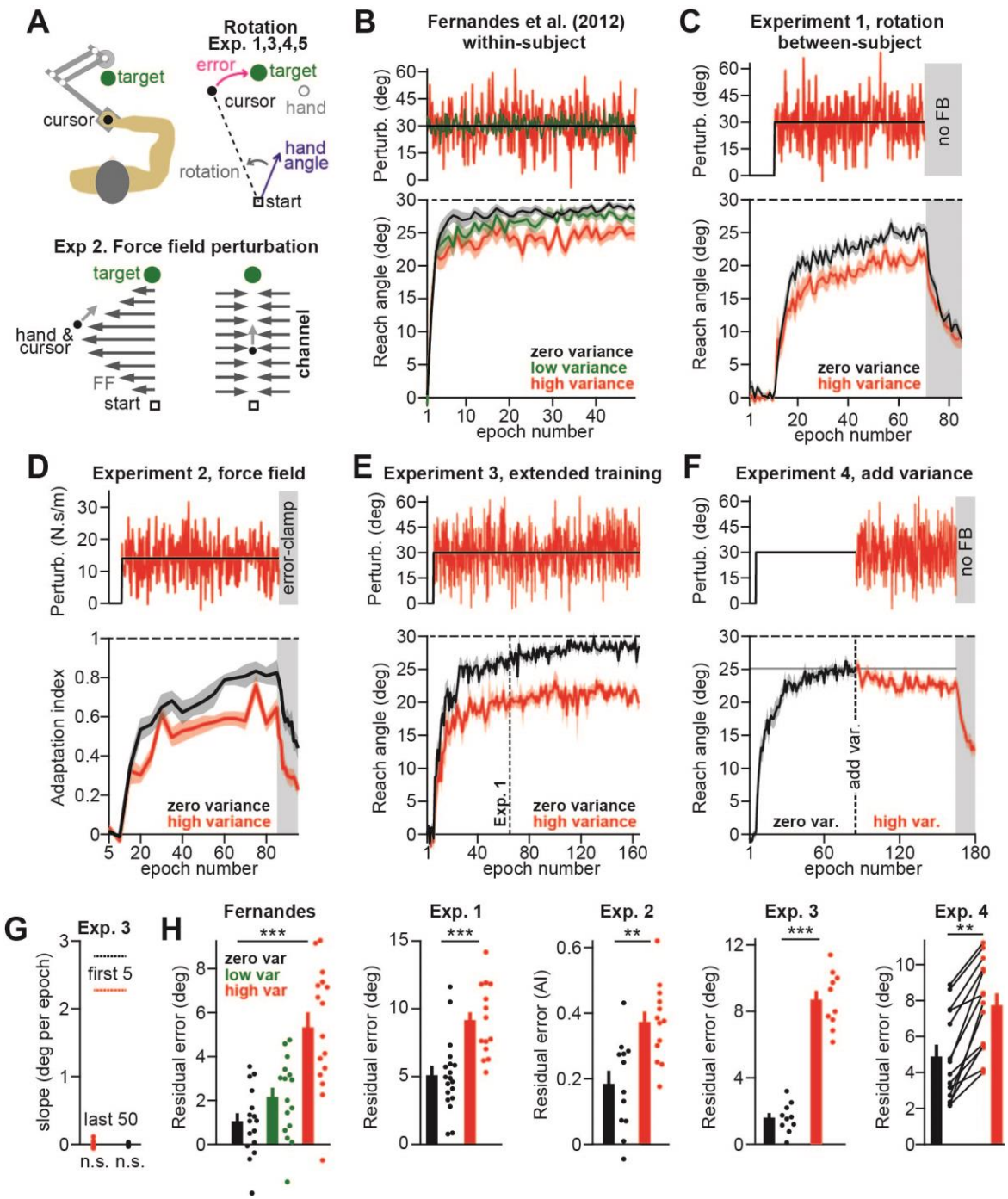
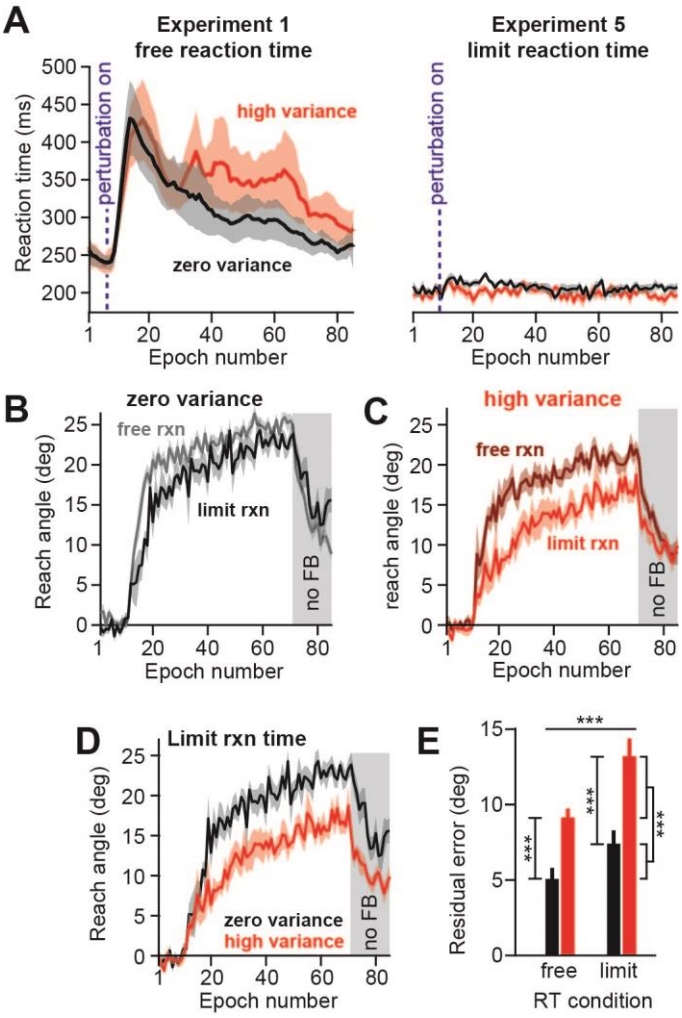


Figure 1. Perturbation variance impairs sensorimotor adaptation. **A.** Schematic of our experiment setup. **B.** Fernandes and colleagues [8] measured the reach angle of participants (bottom, n=16) during adaptation to variable visuomotor rotations (top: SD = 0, 4, and 12° for zero, low, and high variance; mean is 30° for all). Participants demonstrated differing residual errors (reported in inset **H**, **Fernandes**; median error on the last 48 trials). **C.** In Experiment 1, we repeated the experiment of Fernandes et al. (2012) with a between-subjects design. Participants adapted to a zero (n=19) or high (n=14) variance perturbation (SD = 0 and 12° for zero and high variance; mean is 30° for both). The residual error is shown in **H**, **Exp. 1** (median of the last 48 trials). **D.** In Experiment 2, we tested force field adaptation. Occasionally, we measured reaching forces on channel trials that restricted motion of the

hand to a straight path. Participants experienced a zero (n=12) or high (n=13) variance perturbation (top: SD = 0 and 6 N-s/m for zero and high variance; mean = 14 N-s/m for both). We computed an adaptation index on each channel trial (bottom). Residual error (inset **H**, **Exp. 2**) is one minus mean adaptation index on last 5 error clamp trials. **E.** In Experiment 3, we exposed participants to an extended period of visuomotor rotations (160 epochs = 640 trials). The vertical dashed line indicates the total number of rotation trials in Experiment 1. Participants adapted to a zero (n=19) or high (n=14) variance perturbation (top: SD = 0 and 12° for zero and high variance; mean is 30° for both). Mean residual error (inset **H**, **Exp. 3**) was computed over the last 50 epochs. To confirm that performance had reached a plateau, we measured the slope of a line fit to the same period (inset **G**). For comparison, horizontal dashed lines show the mean slope over the first 5 epochs of the perturbation. **F.** In Experiment 4, we adapted participants (n=14) to a zero-variance perturbation, and then abruptly switched to a high variance perturbation. Residual errors (inset **H**, **Exp. 4**) were computed over the last 10 epochs of each period. Error bars are mean ± SEM. Statistics denote the result of a repeated-measured ANOVA (**H**, **Fernandes**) or two-sample t-tests (**H**, all other insets). Statistics: **$p<0.01$ and ***$p<0.001$.



870

871 **Figure 2.** Perturbation variance decreases the total amount of implicit adaptation. **A.** We measured reaction time
872 (time to start the reach in response to presentation of a visual target) during adaptation to a visuomotor rotation.
873 In Experiment 1 (left) we did not limit reaction time. Reaction time greatly increased after the introduction of the
874 perturbation (blue dashed line). In Experiment 5 (right) we repeated the paradigm used in Experiment 1, but
875 limited reaction time. **B.** Limiting reaction time reduced adaptation rates in the zero-variance perturbation group.
876 **C.** Limiting reaction time reduced adaptation rates in the high variance perturbation group. **D.** We compared
877 adaptation to the zero-variance and high variance perturbations when reaction time was limited. Even in the limit
878 rxn condition, perturbation variance impaired adaptation. **E.** To measure differences in adaptation extent as a
879 function of reaction time and perturbation variance, we used a two-way ANOVA. For this test, we measured the
880 total amount of adaptation over the last 10 epochs of Experiments 1 and 5. The top-most and right-most statistical
881 bars denote the significance of the main effects of reaction time and perturbation variance, respectively. The other
882 two vertical statistical bars display the results of two-sample t-tests. Error bars are mean ± SEM. Statistics:
883 ***$p<0.001$.

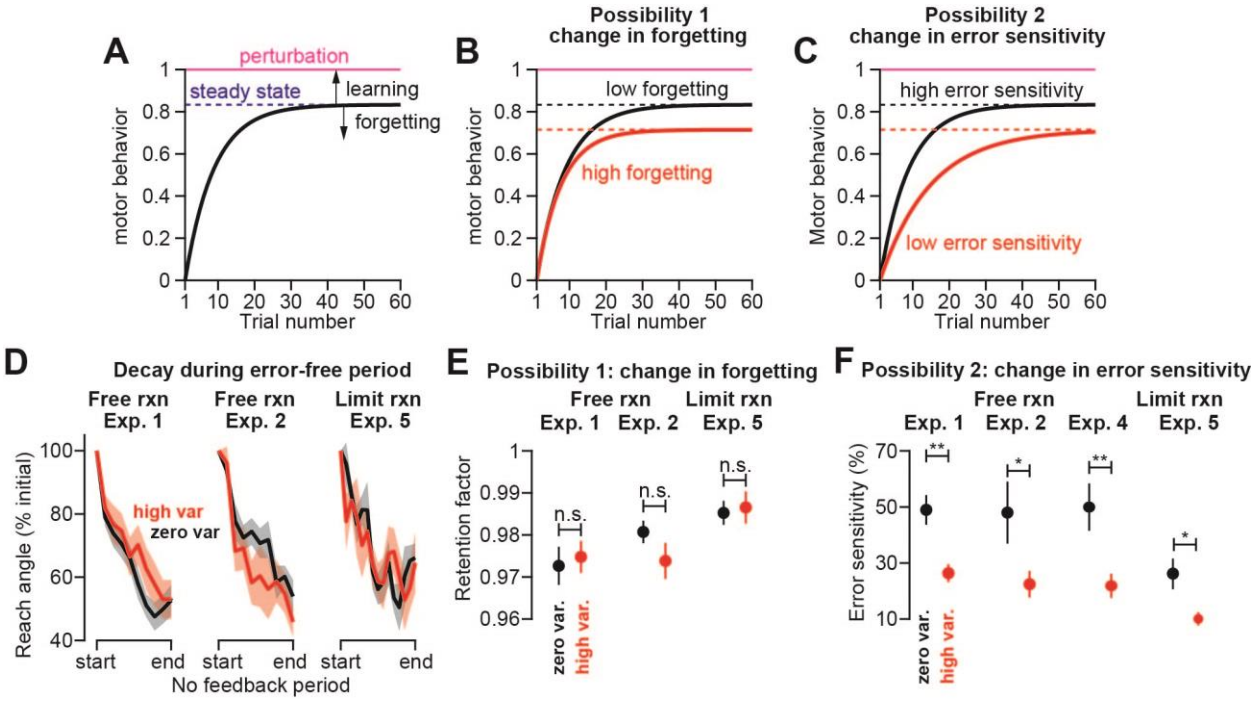


Figure 3. Perturbation variance decreases error sensitivity, not decay rates. **A.** State-space model of adaptation predicts that learning will reach an asymptote when the amount of learning from an error exactly counterbalances the amount of forgetting that occurs between trials. The plot demonstrates the behavior of such a model during adaptation to a perturbation of unit 1. According to the model, changes in asymptotic levels of performance can occur because of changes in forgetting (**B**, Possibility 1 schematic; a = 0.98 for low forgetting and 0.96 for high forgetting), or changes in error sensitivity (**C**, Possibility 2 schematic; b = 0.05 for low error sensitivity and 0.1 for high error sensitivity). **D.** To test Possibility 1, we measured the retention during error-free periods at the end of Experiments 1 (Exp. 1), 2 (Exp. 2), and 4 (Exp. 5). We normalized reach angle to the first trial in the no-feedback period. Decay rates did not differ between the low and high variance groups. Each point on the x-axis is a cycle of 4 trials. **E.** We measured the retention factor during error-free periods in each experiment. We found no difference in retention for the zero-variance and high variance groups. **F.** To test Possibility 2, we measured sensitivity to error in each experiment. Error sensitivity was greater for the zero-variance perturbation in every experiment. Error bars are mean ± SEM. Statistics: *p<0.05, **p<0.01, and n.s. indicates no statistical significance.

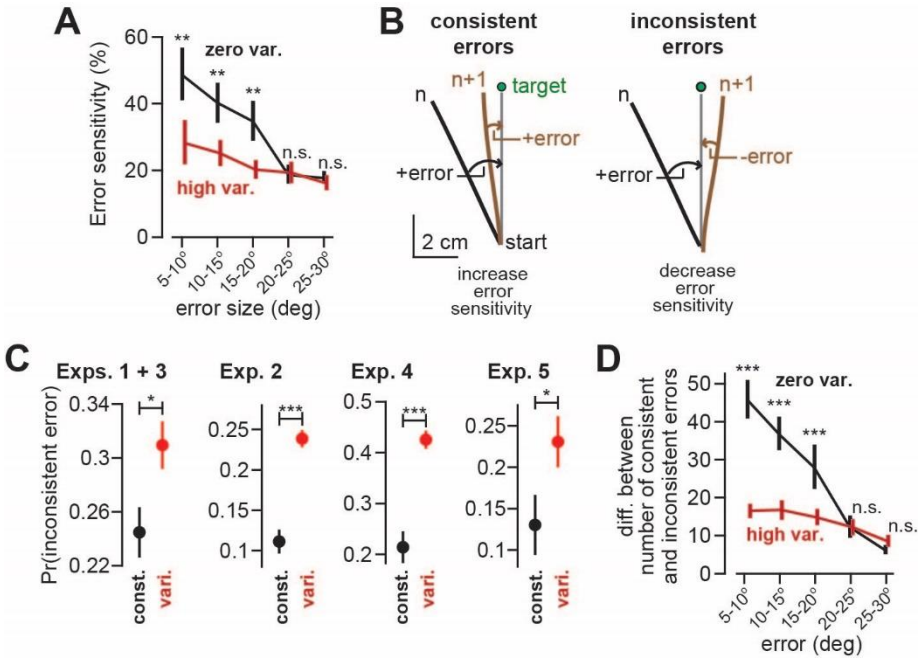


Figure 4. The consistency of trial-to-trial errors coincides with changes in error sensitivity. **A.** To determine how error sensitivity varied as a function of error size, we sorted pairs of movements into different bins (5° width) according to the size of the error on the first movement. Next, we computed the mean error sensitivity across all trials within each error size bin. In the zero-variance environment, subjects exhibited greater error sensitivity to error sizes that were less than 20°. **B.** We considered the possibility that the trial-to-trial consistency of errors caused changes in error sensitivity. Consistent errors (left) are consecutive pairs of trials where the errors have the same sign. Inconsistent errors (right) are consecutive pairs of trials where the errors have opposite signs. The black and brown traces show example reach trajectories from a single participant. **C.** We measured the total fraction of inconsistent error trials. The high variance perturbation caused a higher probability of inconsistent errors in every experiment. **D.** The difference between the number of consistent and inconsistent errors during adaptation to the visuomotor rotation. We found that high variance caused a change in consistency of small errors (less than 20°) but not large errors. Error bars are mean ± SEM. For **A** and **D**, we used a mixed-ANOVA followed by post-hoc two-sample t-tests with Bonferroni corrections. In **C**, two-sample t-tests were used for statistical testing. Statistics: *p<0.05, **p<0.01, ***p<0.001 and n.s. indicates no statistical significance.

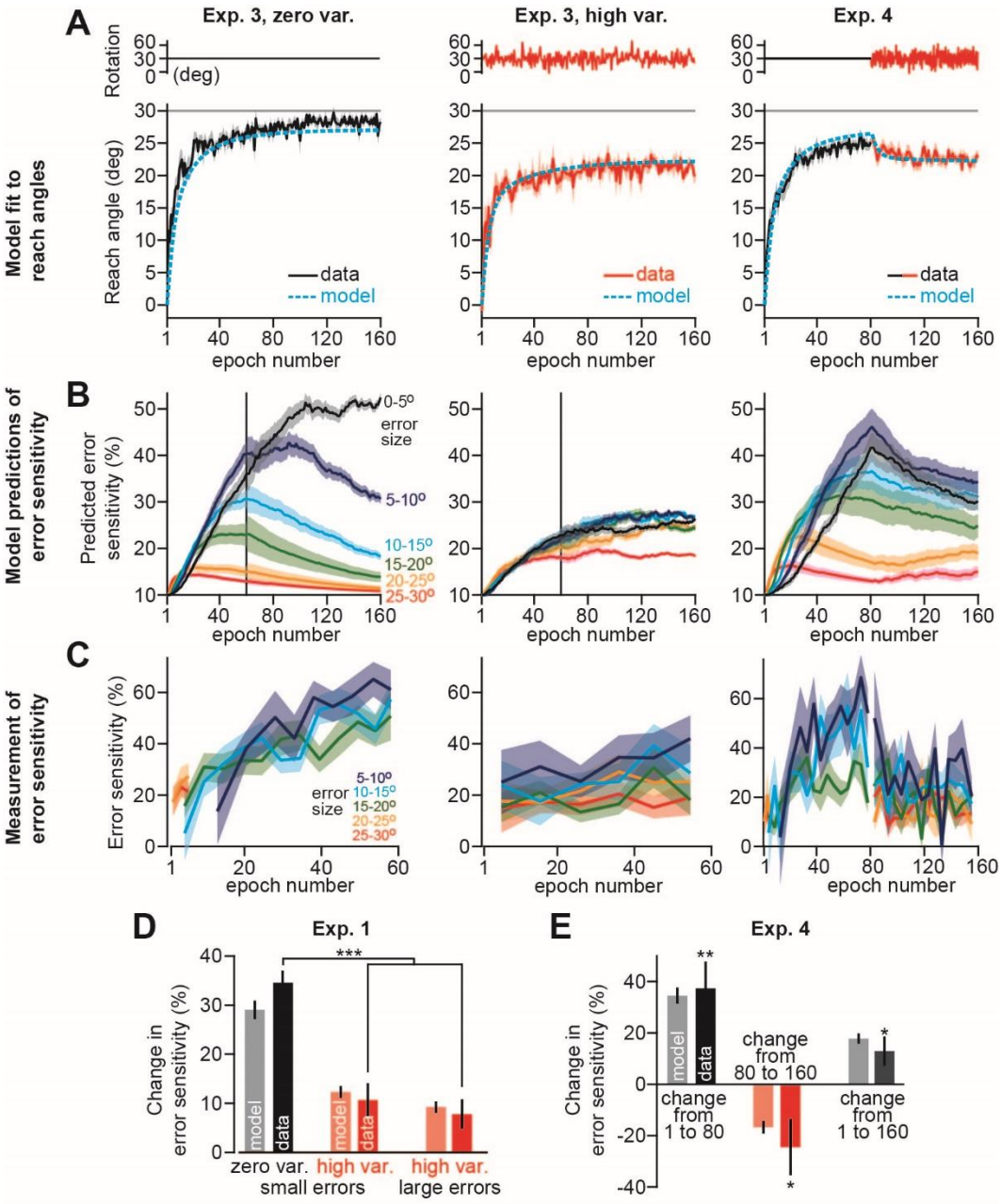


Figure 5. The memory of errors model predicts that error sensitivity changes over the course of training. We simulated a memory of errors model that increases/decreases error sensitivity from one trial to the next when a consistent/inconsistent error occurs. **A.** We simulated the response of this model to the zero-variance perturbation in Exp. 3 (left, Exp. 3, zero var.), the high variance perturbation in Exp. 3 (middle, Exp. 3, high var.), and the zero-to-high variance perturbation in Exp. 4 (right, Exp. 4). The same parameter set was used for each simulation. The model predictions represent the mean model prediction over 100,000 simulations. **B.** Next, we took the specific errors experienced by each participant and used the memory of errors model to predict trial-by-trial changes in error sensitivity. At left, the plot shows the predictions for the zero-variance group in Experiments 1 and 3 (epochs 1-60 represent both experiments, but from 61 onwards only Experiment 3). The same is true for the middle plot, but for the high variance group predictions. At right, the plot shows predictions for the zero-to-high variance group in Experiment 4. **C.** To corroborate the model predictions, we measured error sensitivity on each trial as a function of error size. Then, we binned data across sequences of consecutive trials. Missing data points in each curve

represent periods over which errors were scarce or entirely absent. **D.** The memory of errors model predicted that error sensitivity should increase for both the zero-variance and high variance perturbation, but less so for the high variance condition. Light gray bars show the predicted change in error sensitivity. Dark gray bars show the actual change in error sensitivity measured in the subject population (difference in error sensitivity from epochs 1-15 to epochs 40-59). Predictions are split for small (less than 20°) and large (greater than 20°) errors. **E.** In Experiment 4, the model predicts error sensitivity should first increase in the zero-variance perturbation, and then decrease after adding variance to the perturbation. Light gray bars show the predicted change in error sensitivity. Dark gray bars show the actual change in error sensitivity measured in the subject population over three separate periods (change from 1 to 80 shows the change from the beginning to the end of the zero-variance period; change from 80 to 160 shows the effect of introducing variance to the perturbation; change from 1 to 160 shows the change in error sensitivity from the very start of the experiment, to the very end of the experiment). All bars include errors in the range of 5° to 30°. Error bars are mean ± SEM. Statistics: *$p<0.05$, **$p<0.01$, and ***$p<0.001$.

References

1. Donchin, O. *et al.* Cerebellar regions involved in adaptation to force field and visuomotor perturbation. *J. Neurophysiol.* **107**, 134–147 (2012).
2. Brashers-Krug, T., Shadmehr, R. & Bizzi, E. Consolidation in human motor memory. *Nature* **382**, 252–255 (1996).
3. Tseng, Y.-W., Diedrichsen, J., Krakauer, J. W., Shadmehr, R. & Bastian, A. J. Sensory prediction errors drive cerebellum-dependent adaptation of reaching. *J. Neurophysiol.* **98**, 54–62 (2007).
4. Krakauer, J. W., Pine, Z. M., Ghilardi, M. F. & Ghez, C. Learning of visuomotor transformations for vectorial planning of reaching trajectories. *J. Neurosci.* **20**, 8916–8924 (2000).
5. Ethier, V., Zee, D. S. & Shadmehr, R. Spontaneous Recovery of Motor Memory During Saccade Adaptation. *J. Neurophysiol.* **99**, 2577–2583 (2008).
6. Robinson, F. R., Noto, C. T. & Bevans, S. E. Effect of Visual Error Size on Saccade Adaptation in Monkey. *J. Neurophysiol.* **90**, 1235–1244 (2003).
7. Malone, L. A., Vasudevan, E. V. L. & Bastian, A. J. Motor Adaptation Training for Faster Relearning. *J. Neurosci.* **31**, 15136–15143 (2011).
8. Fernandes, H. L., Stevenson, I. H. & Kording, K. P. Generalization of stochastic visuomotor rotations. *PLoS One* **7**, e43016 (2012).
9. Therrien, A. S., Wolpert, D. M. & Bastian, A. J. Increasing Motor Noise Impairs Reinforcement Learning in Healthy Individuals. *eNeuro* **5**, (2018).
10. Havermann, K. & Lappe, M. The Influence of the Consistency of Postsaccadic Visual Errors on Saccadic Adaptation. *J. Neurophysiol.* **103**, 3302–3310 (2010).
11. Smith, M. A., Ghazizadeh, A. & Shadmehr, R. Interacting adaptive processes with different timescales underlie short-term motor learning. *PLoS Biol.* **4**, e179 (2006).
12. Kording, K. P., Tenenbaum, J. B. & Shadmehr, R. The dynamics of memory as a consequence of optimal adaptation to a changing body. *Nat. Neurosci.* **10**, 779–786 (2007).
13. Coltman, S. K., Cashaback, J. G. A. & Gribble, P. L. Both fast and slow learning processes contribute to savings following sensorimotor adaptation. *J. Neurophysiol.* **121**, 1575–1583 (2019).
14. van der Vliet, R. *et al.* Individual Differences in Motor Noise and Adaptation Rate Are Optimally Related. *eNeuro* **5**, (2018).
15. Vaswani, P. A. *et al.* Persistent Residual Errors in Motor Adaptation Tasks: Reversion to Baseline and Exploratory Escape. *J. Neurosci.* **35**, 6969–6977 (2015).
16. Gonzalez Castro, L. N., Hadjiosif, A. M., Hemphill, M. A. & Smith, M. A. Environmental Consistency Determines the Rate of Motor Adaptation. *Curr. Biol.* **24**, 1050–1061 (2014).
17. Baddeley, R. J., Ingram, H. A. & Miall, R. C. System identification applied to a visuomotor task: near-optimal human performance in a noisy changing task. *J. Neurosci.* **23**, 3066–3075 (2003).
18. Kalman, R. A New Approach to Linear Filtering and Prediction Problems. *ASME J. Basic Eng.* 34–45 (1960).
19. Burge, J., Ernst, M. O. & Banks, M. S. The statistical determinants of adaptation rate in human reaching. *J. Vis.* **8**, 1–19 (2008).
20. van Beers, R. J. How does our motor system determine its learning rate? *PLoS One* **7**, e49373–e49373 (2012).
21. Wei, K. & Körding, K. Uncertainty of feedback and state estimation determines the speed of motor adaptation. *Front. Comput. Neurosci.* **4**, 11 (2010).
22. Fernandez-Ruiz, J., Wong, W., Armstrong, I. T. & Flanagan, J. R. Relation between reaction time and reach errors during visuomotor adaptation. *Behav. Brain Res.* **219**, 8–14 (2011).
23. Haith, A. M., Huberdeau, D. M. & Krakauer, J. W. The Influence of Movement Preparation Time on the Expression of Visuomotor Learning and Savings. *J. Neurosci.* **35**, 5109–5117 (2015).
24. McDougle, S. D. & Taylor, J. A. Dissociable cognitive strategies for sensorimotor learning. *Nat.*

Commun. **10**, 40 (2019).
25. Leow, L.-A., Gunn, R., Marinovic, W. & Carroll, T. J. Estimating the implicit component of visuomotor rotation learning by constraining movement preparation time. *J. Neurophysiol.* **118**, 666–676 (2017).
26. Herzfeld, D. J., Vaswani, P. A., Marko, M. K. & Shadmehr, R. A memory of errors in sensorimotor learning. *Science (80-.).* **345**, 1349–1353 (2014).
27. Leow, L.-A., de Rugy, A., Marinovic, W., Riek, S. & Carroll, T. J. Savings for visuomotor adaptation require prior history of error, not prior repetition of successful actions. *J. Neurophysiol.* **116**, 1603–1614 (2016).
28. Sing, G. C. & Smith, M. A. Reduction in learning rates associated with anterograde interference results from interactions between different timescales in motor adaptation. *PLoS Comput. Biol.* **6**, e1000893 (2010).
29. Scheidt, R. A., Reinkensmeyer, D. J., Conditt, M., Rymer, W. Z. & Mussa-ivaldi, F. A. Persistence of motor adaptation during constrained, multi-joint, arm movements. *J. Neurophysiol.* **84**, 853–862 (2000).
30. Shadmehr, R., Brandt, J. & Corkin, S. Time-dependent motor memory processes in amnesic subjects. *J. Neurophysiol.* **80**, 1590–1597 (1998).
31. Taylor, J. A., Krakauer, J. W. & Ivry, R. B. Explicit and Implicit Contributions to Learning in a Sensorimotor Adaptation Task. *J. Neurosci.* **34**, 3023–3032 (2014).
32. Taylor, J. A. & Ivry, R. B. Flexible cognitive strategies during motor learning. *PLoS Comput. Biol.* **7**, e1001096 (2011).
33. Mazzoni, P. & Krakauer, J. W. An implicit plan overrides an explicit strategy during visuomotor adaptation. *J. Neurosci.* **26**, 3642–3645 (2006).
34. Hwang, E. J., Smith, M. A. & Shadmehr, R. Dissociable effects of the implicit and explicit memory systems on learning control of reaching. *Exp. brain Res.* **173**, 425–437 (2006).
35. Krakauer, J. W., Mazzoni, P., Ghazizadeh, A., Ravindran, R. & Shadmehr, R. Generalization of motor learning depends on the history of prior action. *PLoS Biol.* **4**, e316 (2006).
36. Rabe, K. *et al.* Adaptation to Visuomotor Rotation and Force Field Perturbation Is Correlated to Different Brain Areas in Patients With Cerebellar Degeneration. *J. Neurophysiol.* **101**, 1961–1971 (2009).
37. Galea, J. M., Vazquez, A., Pasricha, N., Orban De Xivry, J. J. & Celnik, P. Dissociating the roles of the cerebellum and motor cortex during adaptive learning: The motor cortex retains what the cerebellum learns. *Cereb. Cortex* **21**, 1761–1770 (2011).
38. Marko, M. K., Haith, A. M., Harran, M. D. & Shadmehr, R. Sensitivity to prediction error in reach adaptation. *J Neurophysiol* **108**, 1752–1763 (2012).
39. Kim, H. E., Morehead, J. R., Parvin, D. E., Moazzezi, R. & Ivry, R. B. Invariant errors reveal limitations in motor correction rather than constraints on error sensitivity. *Commun. Biol.* **1**, 19 (2018).
40. Morehead, J. R., Taylor, J. A., Parvin, D. E. & Ivry, R. B. Characteristics of Implicit Sensorimotor Adaptation Revealed by Task-irrelevant Clamped Feedback. *J. Cogn. Neurosci.* **29**, 1061–1074 (2017).
41. Wei, K. & Kording, K. Relevance of error: what drives motor adaptation? *J. Neurophysiol.* **101**, 655–664 (2009).
42. Smith, M. A. & Shadmehr, R. Modulation of the rate of error-dependent learning by statistical properties of the task. in *Advances in Computational Motor Control* (2004).
43. Kojima, Y. Memory of Learning Facilitates Saccadic Adaptation in the Monkey. *J. Neurosci.* **24**, 7531–7539 (2004).
44. Mawase, F., Shmuelof, L., Bar-Haim, S. & Karniel, A. Savings in locomotor adaptation explained by

changes in learning parameters following initial adaptation. *J. Neurophysiol.* **111**, 1444–1454 (2014).
45. Zarahn, E., Weston, G. D., Liang, J., Mazzoni, P. & Krakauer, J. W. Explaining Savings for Visuomotor Adaptation: Linear Time-Invariant State-Space Models Are Not Sufficient. *J. Neurophysiol.* **100**, 2537–2548 (2008).
46. Lerner, G. *et al.* The origins of anterograde interference in visuomotor adaptation. *bioRxiv* 593996 (2019). doi:10.1101/593996
47. Kitago, T., Ryan, S. L., Mazzoni, P., Krakauer, J. W. & Haith, A. M. Unlearning versus savings in visuomotor adaptation: comparing effects of washout, passage of time, and removal of errors on motor memory. *Front. Hum. Neurosci.* **7**, 307 (2013).
48. Huang, V. S., Haith, A., Mazzoni, P. & Krakauer, J. W. Rethinking motor learning and savings in adaptation paradigms: model-free memory for successful actions combines with internal models. *Neuron* **70**, 787–801 (2011).
49. Kojima, Y. & Soetedjo, R. Change in sensitivity to visual error in superior colliculus during saccade adaptation. *Sci. Rep.* **7**, 9566 (2017).
50. Kojima, Y. & Soetedjo, R. Elimination of the error signal in the superior colliculus impairs saccade motor learning. *Proc. Natl. Acad. Sci.* **115**, E8987--E8995 (2018).
51. Wilterson, S. A. & Taylor, J. A. Implicit visuomotor adaptation remains limited after several days of training. *bioRxiv* (2019). doi:10.1101/711598
52. Morehead, J. R., Qasim, S. E., Crossley, M. J. & Ivry, R. Savings upon Re-Aiming in Visuomotor Adaptation. *J. Neurosci.* **35**, 14386–14396 (2015).
53. Bond, K. M. & Taylor, J. A. Flexible explicit but rigid implicit learning in a visuomotor adaptation task. *J. Neurophysiol.* **113**, 3836–3849 (2015).
54. Avraham, G., Keizman, M. & Shmuelof, L. Environmental Consistency Modulation of Error Sensitivity During Motor Adaptation is Explicitly Controlled. *J. Neurophysiol.* (2019). doi:10.1152/jn.00080.2019
55. Hegele, M. & Heuer, H. The impact of augmented information on visuo-motor adaptation in younger and older adults. *PLoS One* **5**, e12071–e12071 (2010).
56. Heuer, H. & Hegele, M. Adaptation to visuomotor rotations in younger and older adults. *Psychol. Aging* **23**, 190–202 (2008).
57. Vandevoorde, K. & Orban de Xivry, J.-J. Internal model recalibration does not deteriorate with age while motor adaptation does. *Neurobiol. Aging* **80**, 138–153 (2019).
58. Xu-Wilson, M., Chen-Harris, H., Zee, D. S. & Shadmehr, R. Cerebellar Contributions to Adaptive Control of Saccades in Humans. *J. Neurosci.* **29**, 12930–12939 (2009).
59. Herzfeld, D. J. *et al.* Contributions of the cerebellum and the motor cortex to acquisition and retention of motor memories. *Neuroimage* **98**, 147–158 (2014).
60. Hanajima, R. *et al.* Modulation of error-sensitivity during a prism adaptation task in people with cerebellar degeneration. *J. Neurophysiol.* **114**, 2460–2471 (2015).
61. Kim, S., Ogawa, K., Lv, J., Schweighofer, N. & Imamizu, H. Neural Substrates Related to Motor Memory with Multiple Timescales in Sensorimotor Adaptation. *PLoS Biol.* **13**, e1002312 (2015).
62. Herzfeld, D. J., Kojima, Y., Soetedjo, R. & Shadmehr, R. Encoding of action by the Purkinje cells of the cerebellum. *Nature* **526**, 439–442 (2015).
63. Soetedjo, R., Kojima, Y. & Fuchs, A. F. Complex spike activity in the oculomotor vermis of the cerebellum: a vectorial error signal for saccade motor learning? *J. Neurophysiol.* **100**, 1949–1966 (2008).
64. Yang, Y. & Lisberger, S. G. Role of plasticity at different sites across the time course of cerebellar motor learning. *J. Neurosci.* **34**, 7077–90 (2014).
65. Herzfeld, D. J., Kojima, Y., Soetedjo, R. & Shadmehr, R. Encoding of error and learning to correct

that error by the Purkinje cells of the cerebellum. *Nat. Neurosci.* **21**, 736–743 (2018).
66. Albert, S. T. & Shadmehr, R. Estimating properties of the fast and slow adaptive processes during sensorimotor adaptation. *J. Neurophysiol.* **119**, 1367–1393 (2018).