

---

*This copy is for your personal, non-commercial use only.*

---

**If you wish to distribute this article to others**, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

**Permission to republish or repurpose articles or portions of articles** can be obtained by following the guidelines [here](#).

**The following resources related to this article are available online at [www.sciencemag.org](http://www.sciencemag.org) (this information is current as of October 3, 2014):**

**Updated information and services**, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/content/345/6202/1349.full.html>

**Supporting Online Material** can be found at:

<http://www.sciencemag.org/content/suppl/2014/08/13/science.1253138.DC1.html>

This article **cites 35 articles**, 17 of which can be accessed free:

<http://www.sciencemag.org/content/345/6202/1349.full.html#ref-list-1>

This article appears in the following **subject collections**:

Neuroscience

<http://www.sciencemag.org/cgi/collection/neuroscience>

(Fig. 3). The resulting map is a function of the taxonomic cutoff used. At less than 99%, the model does not show any diversity or biogeographic patterns. This is expected because over 1400 years the genomes evolve by only 0.5%, corresponding to a 99% identity. However, at this time there are still provinces that have not yet experienced a takeover or coalescence event (Fig. 1A), and those are expected to continue to diverge beyond the 99% threshold for longer times. This distribution can change temporarily as a result of takeover events. For example, at times the Central Pacific and North Pacific provinces were distinguishable at 99.5% (0.5% difference, see Fig. 2A).

We conclude that neutral evolution (neutral mutations and genetic drift) coupled with dispersal limitation can produce substantial biogeographic patterns in the global surface ocean microbe population. Microbes evolve faster than the ocean circulation can disperse them, a feature that can also be seen in molecular observations (10). The patterns are dynamic. Provinces gradually emerge as subpopulations diverge by neutral evolution and periodically collapse due to coalescence. Neutral processes, along with environmental selection, must be considered in future research on microbial biogeography, and our results provide a quantitative benchmark for their potential role. Our results conflict with the notion that “everything is everywhere” (11, 32) and may have important implications for how the oceans will respond to global change. Our model provides insights into the role of neutral evolution in shaping biogeographic patterns. The biology in the model is relatively simple, and future work may build on this by considering more spatial and temporal patterns (e.g., carrying capacity, as well as division and death rates based on ocean productivity) and more explicit representation of processes (e.g., recombination). Our modeling approach can theoretically be used to explore environmental selection as well. This will require relating genes to function, which is difficult but can be done for select genes or at the genome level (23, 33).

#### REFERENCES AND NOTES

1. A. Ramette, J. M. Tiedje, *Proc. Natl. Acad. Sci. U.S.A.* **104**, 2761–2766 (2007).
2. C. A. Hanson, J. A. Fuhrman, M. C. Horner-Devine, J. B. H. Martiny, *Nat. Rev. Microbiol.* **10**, 497–506 (2012).
3. R. T. Papke, D. M. Ward, *FEMS Microbiol. Ecol.* **48**, 293–303 (2004).
4. M. V. Brown, M. Ostrowski, J. J. Grzymalski, F. M. Lauro, *Mar. Genomics* **15**, 17–28 (2014).
5. R. H. MacArthur, E. O. Wilson, *The Theory of Island Biogeography* (Princeton Univ. Press, Princeton, NJ, 1967).
6. S. P. Hubbell, *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton Univ. Press, Princeton, NJ, 2001).
7. D. B. Rusch et al., *PLOS Biol.* **5**, e77 (2007).
8. B. K. Swan et al., *Proc. Natl. Acad. Sci. U.S.A.* **110**, 11463–11468 (2013).
9. J.-F. Ghiglione et al., *Proc. Natl. Acad. Sci. U.S.A.* **109**, 17633–17638 (2012).
10. A. C. Martiny, A. P. K. Tai, D. Veneziano, F. Primeau, S. W. Chisholm, *Environ. Microbiol.* **11**, 823–832 (2009).
11. S. M. Gibbons et al., *Proc. Natl. Acad. Sci. U.S.A.* **110**, 4651–4655 (2013).
12. J. B. H. Martiny et al., *Nat. Rev. Microbiol.* **4**, 102–112 (2006).

13. D. Wilkins, E. van Sebille, S. R. Rintoul, F. M. Lauro, R. Cavicchioli, *Nat. Commun.* **4**, 2457 (2013).
14. A. D. Barton, S. Dutkiewicz, G. Flierl, J. Bragg, M. J. Follows, *Science* **327**, 1509–1511 (2010).
15. R. Condit et al., *Science* **295**, 666–669 (2002).
16. W. J. Sul, T. A. Oliver, H. W. Ducklow, L. A. Amaral-Zettler, M. L. Sogin, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 2342–2347 (2013).
17. J. W. Drake, B. Charlesworth, D. Charlesworth, J. F. Crow, *Genetics* **148**, 1667–1686 (1998).
18. K. L. Vergin et al., *Environ. Microbiol.* **9**, 2430–2440 (2007).
19. M. Vos, X. Didelot, *ISME J.* **3**, 199–208 (2009).
20. J.-U. Kreft et al., *Proc. Natl. Acad. Sci. U.S.A.* **110**, 18027–18028 (2013).
21. F. L. Hellweger, V. Bucci, *Ecol. Modell.* **220**, 8–22 (2009).
22. P. G. Falkowski, C. D. Wirick, *Mar. Biol.* **65**, 69–75 (1981).
23. F. L. Hellweger, *Environ. Microbiol.* **11**, 1386–1394 (2009).
24. J. R. Clark, T. M. Lenton, H. T. Williams, S. J. Daines, *Limnol. Oceanogr.* **58**, 1008–1022 (2013).
25. W. B. Whitman, D. C. Coleman, W. J. Wiebe, *Proc. Natl. Acad. Sci. U.S.A.* **95**, 6578–6583 (1998).
26. M. Nei, W. H. Li, *Proc. Natl. Acad. Sci. U.S.A.* **76**, 5269–5273 (1979).
27. M. Slatkin, *Evolution* **47**, 264–279 (1993).
28. Materials and methods are available as supplementary materials on Science Online.
29. H. Ducklow, in *Microbial Ecology of the Oceans*, D. L. Kirchman, Ed. (Wiley-Liss, New York, 2000), chap. 4, pp. 85–120.

30. Y. Masumoto et al., *J. Earth Simulator* **1**, 35–56 (2004).
31. J. M. Halley, Y. Iwasa, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 2316–2321 (2011).
32. P. Cermeño, P. G. Falkowski, *Science* **325**, 1539–1541 (2009).
33. A. M. Feist, M. J. Herrgård, I. Thiele, J. L. Reed, B. Ø. Palsson, *Nat. Rev. Microbiol.* **7**, 129–143 (2009).
34. S. J. Giovannoni et al., *Science* **309**, 1242–1245 (2005).

#### ACKNOWLEDGMENTS

We thank R. Stepanauskas for interesting discussion that inspired this research. Three anonymous reviewers provided constructive criticism. F.L.H. and N.D.F. are supported by grants from the NSF and National Oceanic and Atmospheric Administration Massachusetts Institute of Technology SeaGrant. E.V.S. is supported by ARC via grant DE130101336. The source code is available at [www.coe.neu.edu/~ferdi/Code/](http://www.coe.neu.edu/~ferdi/Code/).

#### SUPPLEMENTARY MATERIALS

[www.sciencemag.org/content/345/6202/1346/suppl/DC1](http://www.sciencemag.org/content/345/6202/1346/suppl/DC1)  
Materials and Methods  
Supplementary Text  
Figs. S1 to S9  
Table S1  
References (35–43)  
Movie S1

4 April 2014; accepted 1 August 2014  
10.1126/science.1254421

#### ERROR MEMORY

## A memory of errors in sensorimotor learning

David J. Herzfeld,<sup>1\*</sup> Pavan A. Vaswani,<sup>2</sup> Mollie K. Marko,<sup>1</sup> Reza Shadmehr<sup>1</sup>

The current view of motor learning suggests that when we revisit a task, the brain recalls the motor commands it previously learned. In this view, motor memory is a memory of motor commands, acquired through trial-and-error and reinforcement. Here we show that the brain controls how much it is willing to learn from the current error through a principled mechanism that depends on the history of past errors. This suggests that the brain stores a previously unknown form of memory, a memory of errors. A mathematical formulation of this idea provides insights into a host of puzzling experimental data, including savings and meta-learning, demonstrating that when we are better at a motor task, it is partly because the brain recognizes the errors it experienced before.

**H**ow does the brain alter behavior after experiencing an error? Classic theories assumed that the brain learns some fraction of the error regardless of its history or magnitude (1, 2). However, recent experiments (3–6) demonstrate that the brain learns relatively more from small errors than large errors, and can modulate its error sensitivity (7–9).

Understanding error sensitivity is important because it may provide insight into the phenomena of “savings” and “meta-learning.” Savings refers to the observation that when a person adapts to perturbation (A), and then the perturbation is removed (i.e., washout), they exhibit faster readaptation

to (A) (10). Remarkably, savings of (A) is present even when washout is followed by adaptation to (–A), a perturbation in the opposite direction (11, 12). Current error-dependent models of learning cannot account for these observations (13, 14), nor explain meta-learning, where prior exposure to a random perturbation produces savings (15, 16).

We begin with a standard model of motor learning (17–20) in which on trial  $n$ , a perturbation  $x$  is imposed on action  $u$  so that the sensory consequences are  $y^{(n)} = u^{(n)} + x^{(n)}$ . Based on their belief about the environment  $\hat{x}^{(n)}$ , the learner predicts the sensory consequences  $\hat{y}^{(n)} = u^{(n)} + \hat{x}^{(n)}$ , and then updates his or her belief from the prediction error  $e^{(n)} = y^{(n)} - \hat{y}^{(n)}$ . Such learning typically depends on a decay factor  $\alpha$ , and error sensitivity  $\eta$

$$\hat{x}^{(n+1)} = \alpha \hat{x}^{(n)} + \eta^{(n)} e^{(n)} \quad (1)$$

Consider an environment in which the perturbations persist from trial to trial, and another

<sup>1</sup>Department of Biomedical Engineering, Laboratory for Computational Motor Control, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. <sup>2</sup>Department of Neuroscience, Laboratory for Computational Motor Control, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA.

\*Corresponding author: E-mail: dherzfe1@jhmi.edu

environment in which the perturbations switch (Fig. 1A). In a slowly switching environment, the brain should learn from error because the perturbations are likely to persist (learning from error in one trial will improve performance on the subsequent trial). However, in a rapidly switching environment, the brain should suppress learning from error because any learning will be detrimental to performance on subsequent trials.

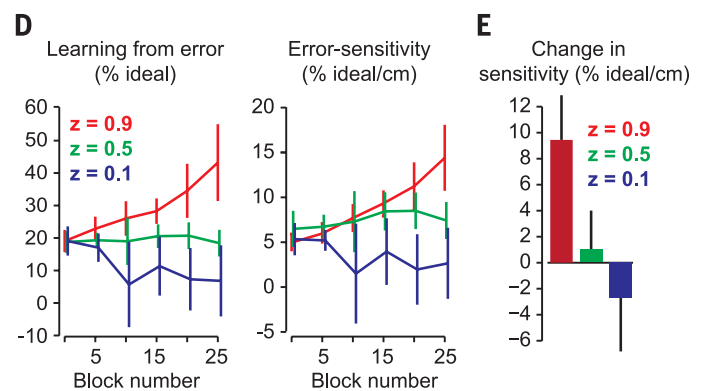
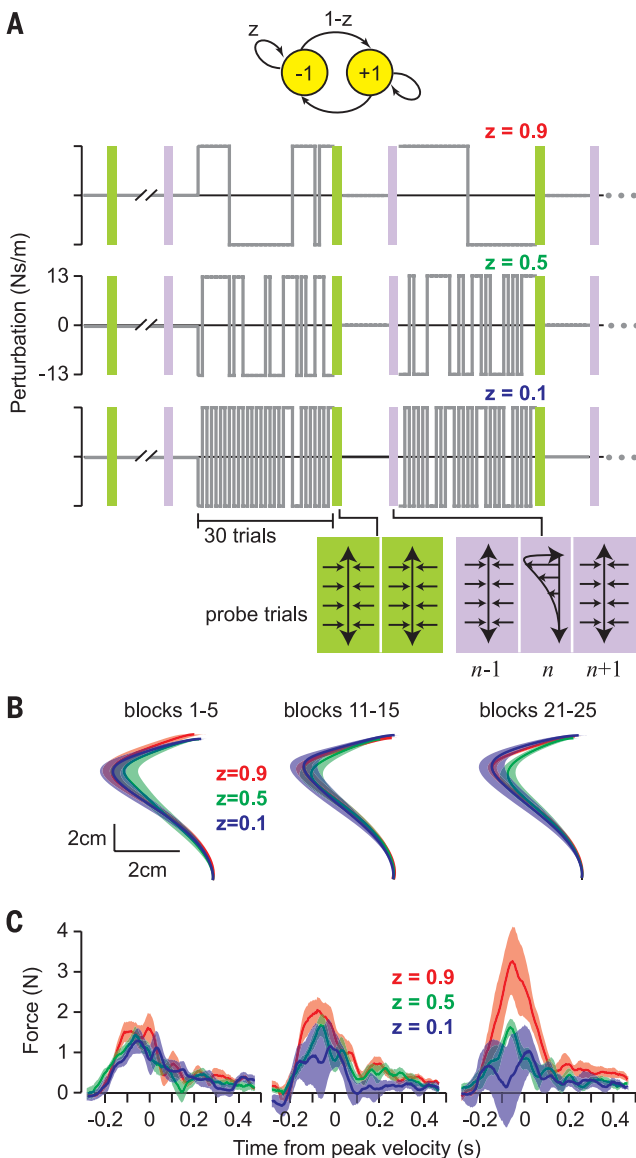
Three groups of volunteers ( $n = 9$  per group) made reaching movements while experiencing force perturbations from either a slow, medium, or rapidly switching environment (Fig. 1A). The mean of the perturbations was zero for all blocks (consisting of 30 trials). We measured error in a given trial and then computed the amount that was learned from that error (probe trials, purple bars, Fig. 1A). To quantify learning from error on trial  $n$ , we measured the change in force from the trial before to the trial after the perturbation,  $f^{(n+1)}(t) - f^{(n-1)}(t)$  (Fig. 1C). In block 1, learning from error was similar in the three groups ( $P > 0.99$ ),

and in all probe trials the perturbation produced similar errors [Fig. 1B; repeated measure-analysis of variance (RM-ANOVA), effect of group  $P > 0.8$ , interaction,  $P > 0.7$ ]. However, individuals who experienced the slowly switching environment increased their learning from error (Fig. 1C), whereas those who experienced the rapidly switching environment suppressed this learning.

We measured the force produced on a given trial and computed a coefficient representing percent ideal (Fig. 1D). RM-ANOVA indicated a significant block by group interaction ( $P < 0.05$ ), suggesting that the history of perturbations altered the amount of learning from error. Posthoc tests showed that in the slowly switching environment, participants learned more from error than in the rapidly switching environment ( $P < 0.03$ ). This change in error sensitivity developed gradually with training (Fig. 1D). The slowly switching environment induced an increase in error sensitivity (Fig. 1E; changes in sensitivity from the first half to second half of the experiment, ANOVA,  $P < 0.05$ ).

Is control of error sensitivity local to the experienced errors? In experiment 2, participants performed rapid out-and-back movements for which no visual feedback was available during the outward part of the reach, with the aim of hitting a target at the turn-around point of their movement. An occasional perturbation altered the feedback regarding hand position at the turn-around point (Fig. 2A). We measured the relation between error  $e^{(n)}$  and learning from error (change in reach extent).

Group 1 ( $n = 10$ ) experienced a perturbation schedule that transitioned from slow, medium, to rapid switching (Fig. 2B), whereas group 2 ( $n = 10$ ) experienced the reverse. In group 1, error sensitivity decreased, whereas in group 2, error sensitivity increased (Fig. 2C). We measured the mean error sensitivity in each environment, resulting in three measurements for each subject across the experiment. RM-ANOVA showed a significant main effect of group ( $P < 0.005$ ) and block ( $P < 0.001$ ) and group-by-block interaction



**Fig. 1. History of error alters error sensitivity.** (A) Reaching paradigm with force-field perturbations. The yellow circles note a perturbations state, and  $z$  indicates probability of remaining in that state. The slow, medium, and rapidly switching environments are shown. One group of volunteers was trained in each environment. We measured error sensitivity through probe trials in which participants experienced a constant perturbation, sandwiched between two error-clamp trials. (B) Movement trajectories in the perturbation trial of the probe trials. Trajectories were averaged over five successive presentations of the probe. The errors in probe trials did not differ between groups. (C) Learning from error in the probe trials, measured as the change in force from the trial prior to the trial after the perturbation. (D) Learning from error in the probe trials, plotted as a percentage of the ideal force (left). Error sensitivity  $\eta$  was measured as the trial-to-trial change in the percentage of ideal force divided by error (right). (E) Change in error sensitivity between the baseline block and the last five error-clamp triplets. Data are mean  $\pm$  SEM.

( $P < 0.001$ ). As the statistics of the perturbation changed, so did the error sensitivity.

We measured learning from error as a function of error in each environment (Fig. 2D). A given error produced greater learning when that error was experienced in a slowly switching environment (Fig. 2D, red line) (RM-ANOVA main effects of error size  $P < 10^{-4}$ , and environment  $P < 0.001$ , posthoc between slow versus medium or fast,  $P < 0.001$ ). We quantified error sensitivity at each error size (Fig. 2E) and found that error sensitivity had not changed globally, but had changed predominantly for smaller error sizes. RM-ANOVA of the absolute sensitivities between 0.25 and 2 cm showed a significant main effect of environment ( $P < 10^{-4}$ ), as well as a significant environment by error size interaction ( $P < 0.05$ ). We found a significant difference in error sensitivity across environments for an error size of 0.25 cm ( $P < 0.05$ ), but no significant difference for an error size of 2 cm ( $P > 0.1$ ). Interestingly, the small error sizes for which the participants had shown the largest change in error sensitivity were also the most frequent errors (Fig. 2F). This hinted that control of error sensitivity was a function of error.

Current models of sensorimotor learning assume that error sensitivity  $\eta^{(n)}$  is independent of error  $e^{(n)}$ . This is true for state-space models of learning (18, 21–24), as well as Kalman filter models of learning (5, 25–27). However, suppose that sensory prediction errors are encoded in the nervous system with a set of basis elements, where each basis element  $g_i$  has a preferred error  $\tilde{e}_i$ . Further, suppose that error sensitivity is determined by a population coding

$$\begin{aligned}\eta(e^{(n)}) &= \sum_i w_i g_i(e^{(n)}) \\ g_i(e^{(n)}) &= \exp\left(-\frac{(e^{(n)} - \tilde{e}_i)^2}{2\sigma^2}\right)\end{aligned}\quad (2)$$

On trial  $n - 1$ , the motor command  $u^{(n-1)}$  produces an error  $e^{(n-1)}$ , as illustrated in the top part of Fig. 3A. The nervous system learns from this error and produces motor command  $u^{(n)}$  on the subsequent trial, resulting in  $e^{(n)}$ . In a slowly switching environment (top part of Fig. 3A),  $e^{(n)}$  has the same sign as  $e^{(n-1)}$ . In this case, error sensitivity should increase around  $e^{(n-1)}$  (Fig. 3B, red line). By contrast, in a rapidly switching environment (Fig. 3A, bottom),  $e^{(n)}$  has a different sign than  $e^{(n-1)}$ . In this case, error sensitivity should decrease

$$\mathbf{w}^{(n+1)} = \mathbf{w}^{(n)} + \beta \text{sign}(e^{(n-1)}e^{(n)}) \frac{\mathbf{g}(e^{(n-1)})}{\mathbf{g}^T(e^{(n-1)})\mathbf{g}(e^{(n-1)})}\quad (3)$$

In Eq. (3),  $\mathbf{w} = [w_1 \ w_2 \ \dots \ w_N]^T$ ,  $\mathbf{g} = [g_1 \ g_2 \ \dots \ g_N]^T$ , and superscript  $T$  is the transpose operator. This rule is similar to the RPROP algorithm, a heuristic for adjusting the learning rate of machines (28), but has the unique feature of assuming that error sensitivity is via population coding of the error space.

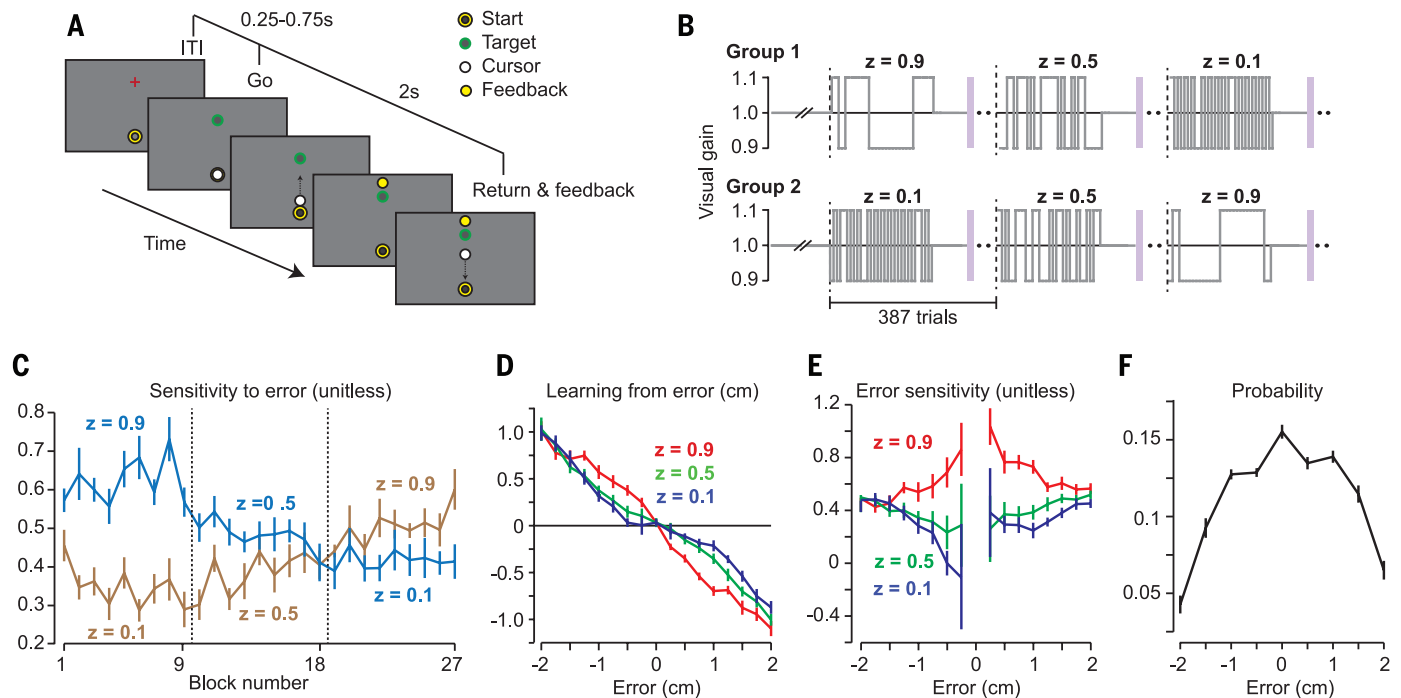
Equations 1 to 3 represent a learner that stores two kinds of memory: a memory of the state of environment ( $\hat{x}$ , Eq. 1) and a memory of errors ( $\mathbf{w}$ , Eq. 3). We simulated the model (Fig. 3C, gray line) and found that in the slowly switching environment, error sensitivity increased in the neighborhood of the experienced errors, whereas in the

rapidly changing environment, error sensitivity decreased (Fig. 3D).

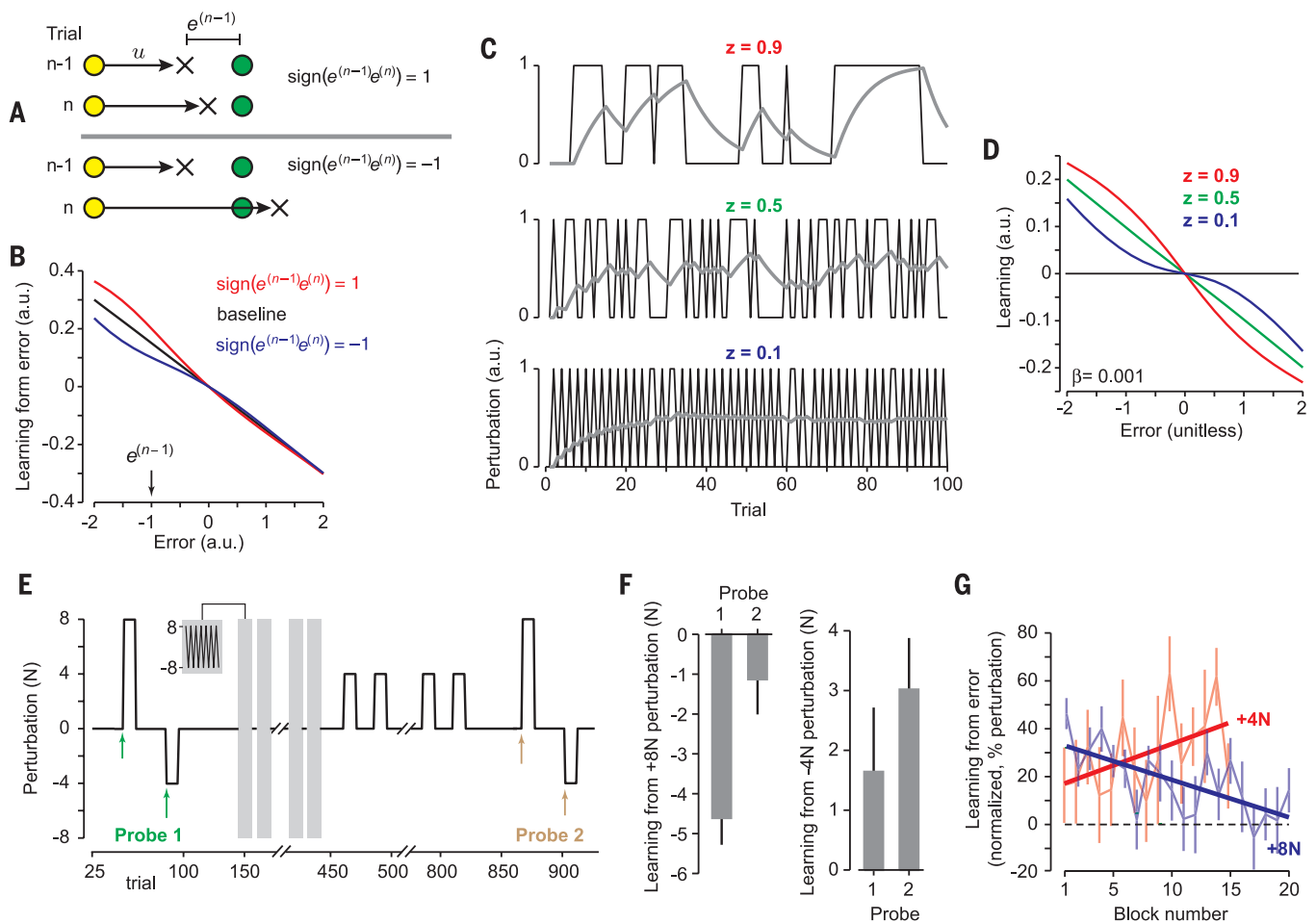
Our model made a critical prediction: If the brain controlled error sensitivity via memory of errors, then it should be possible to simultaneously increase sensitivity for one error, while decreasing it for another. In experiment 3, we considered an isometric task in which participants ( $n = 16$ ) produced a force to match a target (16 N) in the face of a perturbation. The perturbations were designed so that, according to our model, individuals would increase their sensitivity to  $-4$  N errors, while simultaneously decreasing their sensitivity to  $+8$  N errors.

In the baseline block, we probed sensitivity to  $+8$  N and  $-4$  N perturbations (probe 1, Fig. 3E). The resulting learning from error is plotted in Fig. 3F (probe 1). At baseline, participants responded to the  $+8$  N and  $-4$  N perturbations by learning a fraction of each error (Fig. 3F). We next produced 20 repetitions of a rapidly switching environment in which the perturbations were  $\pm 8$  N (Fig. 3E, inset). After a period of washout, we then produced 15 repetitions of a slowly switching environment in which the perturbations were 0 N or  $+4$  N. The critical aspect of our design was that the participants were never exposed to a  $-4$  N perturbation. They nevertheless experienced  $-4$  N errors (because removal of a learned  $+4$  N perturbation results in a  $-4$  N error).

The 8 N environment induced a decrease in sensitivity to a  $+8$  N error, and subsequent exposure to the  $+4$  N environment resulted in an increase in sensitivity to a  $-4$  N error [Fig. 3G; RM-ANOVA showed a significant main effect of perturbation ( $P < 0.03$ ) as well as a perturbation by block



**Fig. 2. Error sensitivity is a local function of experienced errors.** (A) Paradigm with visuomotor gain perturbations. (B) Perturbation schedule. Dashed lines indicate changes in the statistics of the environment. (C) Error sensitivity averaged over all error sizes measured over each environment block. (D) Learning from error measured at various error sizes. (E) Error sensitivity as a function of error magnitude. (F) Probability of error.



**Fig. 3. Theoretical model and experiment 3.** (A) On trial  $n - 1$ , the motor command  $u^{(n-1)}$  is generated, resulting in error  $e^{(n-1)} = -1$ . If the error in trial  $n$  is of the same sign as  $e^{(n-1)}$ , then error sensitivity should increase (top). However, if the error experienced in trial  $n$  has a different sign than  $e^{(n-1)}$ , then error sensitivity should decrease (bottom). (B) Learning from error after experience of two consecutive errors from (A). Error sensitivity around  $e^{(n-1)}$  increases if  $\text{sign}(e^{(n-1)}e^{(n)}) = 1$  and decreases otherwise. (C) Model performance for slow, medium, and rapidly switching environments (gray line represents  $\hat{x}^{(n)}$ ). Learning from error (D) is increased in the slowly

switching environment and decreased in the rapidly switching environment. (E) Experiment 3 perturbation protocol. (F) Single-trial learning from a +8 N perturbation and a -4 N perturbation in probes 1 and 2. Learning is increased for the -4 N perturbation, while simultaneously decreased for a +8 N perturbation. (G) Learning from error normalized by the perturbation magnitude (4 or 8 N) in the first trial of each repetition of the rapidly (blue) and slowly switching (red) environments. Learning increased in the slowly switching (4 N) environment but decreased when the perturbation was rapidly switching (8 N). Error bars are SEM.

interaction,  $P < 0.01$ ]. The critical question, however, was whether both of these changes in sensitivity were simultaneously present. After the slowly switching block of perturbations, we again probed sensitivity to +8 N and -4 N errors (probe 2, Fig. 3E). Compared to the baseline block (probe 1), learning from a +8 N error had decreased ( $P < 0.005$ ), while simultaneously, learning from a -4 N error had increased ( $P < 0.05$ ) (Fig. 3F). When we ran our model on the same sequence of errors that participants had experienced, the change in error sensitivity predicted by the model was highly correlated with the change observed in our participants ( $R^2 = 0.65$ ;  $P < 10^{-8}$ , fig. S1), suggesting that history of error induced changes in error sensitivity in the region of the experienced errors.

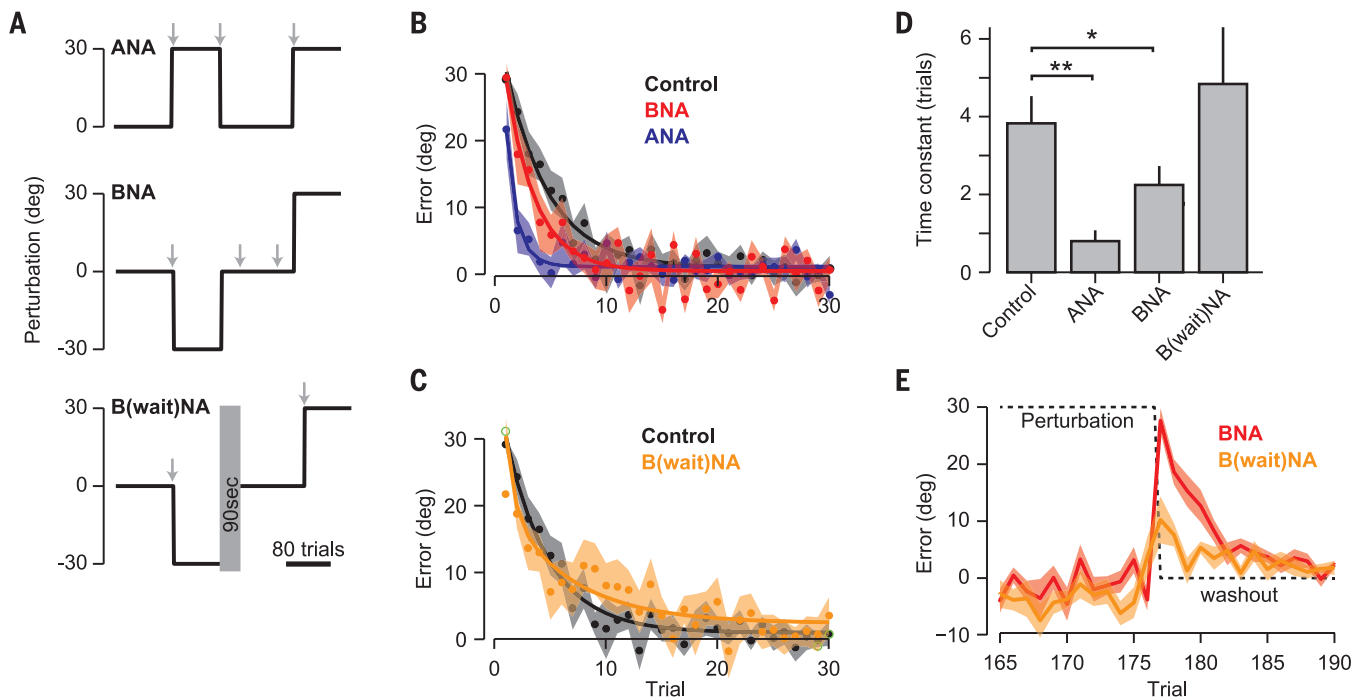
This new model of learning provided insights on a wide range of puzzling experiments, including the phenomena of savings and meta-

learning (fig. S2). It predicted that when one is better at a task than before, it is not because the brain recalled the motor commands, but because it recognized the errors—the errors for which error sensitivity had been altered. In addition, the model predicted that savings and meta-learning could be blocked by controlling the errors that are experienced during learning.

In experiment 4, volunteers participated in a visuomotor rotation experiment (Fig. 4A and fig. S4,  $n = 10$  per group). The control group (ANA) experienced a +30° perturbation followed by extended washout and then relearning of +30°, a protocol that should produce savings (13). According to our model, savings occurs because during the initial exposure to (A), the stable sequence of perturbations increase error sensitivity, and these errors are revisited in the subsequent test of (A). If so, we should be able to block savings by presenting (A) gradually (GNA group,

fig. S4), preventing prior exposure to errors that are visited at the onset of (A).

Furthermore, we should be able to produce savings in a very different way: Expose participants to perturbation (B) and then present sudden washout (Fig. 4A, BNA). During washout, they are exposed to a sequence of stable errors, which increase error sensitivity for those errors. Notably, the washout-induced aftereffects are errors that are also experienced during subsequent test of (A). If the meta-learning in BNA is due to errors that are experienced during washout of (B), we should be able to eliminate it by reducing the washout-induced errors. In BwaitNA, a wait period was inserted between -30° training and washout, reducing the size of aftereffects associated with the transition from -30° to washout (Fig. 4E). We also tested this idea in a different way: gradual washout (B) (group BGNA, fig. S4). In summary, the model predicted



**Fig. 4. Saving occurs only when previously experienced errors are revisited. (A)** A visuomotor perturbation experiment. Gray arrows indicate 1- to 2-min set breaks. **(B)** Performance in the final +30° perturbation. ANA and BNA groups show savings, i.e., faster learning of the perturbation compared to control (naive). Exponential fits are shown for the group data. **(C)** The  $B_{\text{wait}}$ NA group does not exhibit savings. **(D)** Exponential time constants are compared to controls ( $*P < 0.05$ ,  $**P < 0.01$ ). A smaller time constant indicates faster learning. **(E)** Comparison of the errors (i.e., aftereffects) experienced by the BNA and  $B_{\text{wait}}$ NA groups. The  $B_{\text{wait}}$ NA group experienced smaller errors owing to the presence of the set break.

savings in ANA but not GNA, and meta-learning in BNA but not  $B_{\text{wait}}$ NA and BGNA. Our experimental results confirmed these predictions (Fig. 4, C and D, and fig. S4).

We found that during learning, the brain controlled error sensitivity in a principled way: learning more from error when perturbations were likely to persist, and less when perturbations were likely to change. Error-sensitivity modulation was specific to the experienced errors, suggesting that training produced a memory of errors. This idea accounted for a host of puzzling observations, including saturation of error sensitivity (5, 6, 29), the phenomenon of meta-learning (16), examples of savings (10–12), and reinforced repetition (15).

The model predicted that meta-learning vanishes when a small delay or gradual washout alters the history of errors (Fig. 4A), demonstrating that savings depends crucially on the memory of errors that is accumulated during training. This memory of errors likely exists in parallel with the two traditional forms of motor memory, memory of perturbations (23) and memory of actions (30).

In our model, we chose to describe the learner as a process with a single time scale. However, data suggest that learning from error depends on a fast and a slow process with different error sensitivities (23, 25, 31). We speculate that the memory of errors exerts its influence through the error sensitivity of the fast process, and its manipulation through history of errors may be a useful strategy to speed recovery during rehabilitation (32).

#### REFERENCES AND NOTES

- M. I. Jordan, D. E. Rumelhart, *Cogn. Sci.* **16**, 307–354 (1992).
- M. Kawato, K. Furukawa, R. Suzuki, *Biol. Cybern.* **57**, 169–185 (1987).
- F. R. Robinson, C. T. Noto, S. E. Bevens, *J. Neurophysiol.* **90**, 1235–1244 (2003).
- R. Soetedjo, A. F. Fuchs, Y. Kojima, *J. Neurosci.* **29**, 15213–15222 (2009).
- K. Wei, K. Körding, *J. Neurophysiol.* **101**, 655–664 (2009).
- M. K. Marko, A. M. Haith, M. D. Harran, R. Shadmehr, *J. Neurophysiol.* **108**, 1752–1763 (2012).
- M. A. Smith, R. Shadmehr, *Advances in Computational Motor Control* **3**, (2004).
- L. N. Gonzalez Castro, A. M. Hadjiosif, M. A. Hemphill, M. A. Smith, *Curr. Biol.* **24**, 1050 (2014).
- M. C. Trent, A. A. Ahmed, *Front. Comput. Neurosci.* **7**, 118 (2013).
- Y. Kojima, Y. Iwamoto, K. Yoshida, *J. Neurosci.* **24**, 7531–7539 (2004).
- L. A. Malone, E. V. Vasudevan, A. J. Bastian, *J. Neurosci.* **31**, 15136–15143 (2011).
- A. M. Sarwary, L. P. Selen, W. P. Medendorp, *J. Neurophysiol.* **110**, 1269–1277 (2013).
- E. Zarahn, G. D. Weston, J. Liang, P. Mazzoni, J. W. Krakauer, *J. Neurophysiol.* **100**, 2537–2548 (2008).
- F. Mawase, L. Shmuelof, S. Bar-Haim, A. Karniel, *J. Neurophysiol.* **111**, 1444–1454 (2014).
- D. A. Braun, A. Aertsen, D. M. Wolpert, C. Mehring, *Curr. Biol.* **19**, 352–357 (2009).
- E. J. Turnham, D. A. Braun, D. M. Wolpert, *J. Neurophysiol.* **107**, 1111–1122 (2012).
- N. J. Mackintosh, *Psychol. Rev.* **82**, 276–298 (1975).
- K. A. Thoroughman, R. Shadmehr, *Nature* **407**, 742–747 (2000).
- R. A. Rescorla, A. R. Wagner, in *Classical Conditioning*, A. H. Black, W. F. Prokasy, Eds. (Appleton-Century-Crofts, New York, 1972), vol. 2.
- J. M. Pearce, G. Hall, *Psychol. Rev.* **87**, 532–552 (1980).
- R. A. Scheidt, J. B. Dingwell, F. A. Mussa-Ivaldi, *J. Neurophysiol.* **86**, 971–985 (2001).
- O. Donchin, J. T. Francis, R. Shadmehr, *J. Neurosci.* **23**, 9032–9045 (2003).
- M. A. Smith, A. Ghazizadeh, R. Shadmehr, *PLOS Biol.* **4**, e199 (2006).
- S. Cheng, P. N. Sabes, *Neural Comput.* **18**, 760–793 (2006).
- K. P. Körding, J. B. Tenenbaum, R. Shadmehr, *Nat. Neurosci.* **10**, 779–786 (2007).
- R. J. van Beers, *Neuron* **63**, 406–417 (2009).
- R. J. van Beers, *PLOS ONE* **7**, e49373 (2012).
- M. Riedmiller, H. Braun, in *IEEE International Conference on Neural Networks* (1993), vol. 1, pp. 586–591.
- M. S. Fine, K. A. Thoroughman, *J. Neurophysiol.* **96**, 710–720 (2006).
- V. S. Huang, A. Haith, P. Mazzoni, J. W. Krakauer, *Neuron* **70**, 787–801 (2011).
- D. J. Herzfeld et al., *Neuroimage* **98**, 147–158 (2014).
- J. L. Patton, Y. J. Wei, P. Bajaj, R. A. Scheidt, *PLOS ONE* **8**, e46466 (2013).

#### ACKNOWLEDGMENTS

This work was supported by NIH grants T32EB003383, T32GM007057, R01NS078311, and 1F31NS079121. We thank J. Krakauer, J.-J. Orban de Xivry, and A. A. Ahmed for comments. D.J.H., P.A.V., M.K.M., and R.S. conceived and designed the experiments. D.J.H. performed the experiments, analyzed the data, and performed the simulations. R.S. and D.J.H. wrote the paper. The data reported in this paper and in the supplementary materials are available from the corresponding author.

#### SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6202/1349/suppl/DC1  
Supplementary Text  
Figs. S1 to S4  
References (33–40)

10 March 2014; accepted 24 July 2014  
Published online 14 August 2014;  
10.1126/science.1253138