Thesis

Movement state noise and output noise relate to visuomotor adaptation rate in an optimal way

Author
Rick van der Vliet

Supervisors
Winfred Mugge
Alfred Schouten

June 2017

ABSTRACT

Human movement relies on noisy processes in neurons, muscle cells and sensory cells. Therefore, movements are variable and can never be exactly reproduced. The nervous system seems to exploit this movement noise for motor learning and specifically motor adaptation. However, a positive relation between movement noise and motor adaptation has not been consistently found in motor adaptation literature. Possibly, noise is comprised of distinct processes which contribute to motor adaptation in different ways. In Kalman filter theory, motor adaptation rate is calculated optimally from state noise and output noise, with state noise and adaptation rate positively correlated and output noise and adaptation rate negatively correlated. Therefore, if people learn (close) optimally from error, we would expect a similar relation. To investigate the relation between state noise, output noise and adaptation rate, we performed a visuomotor reaching adaptation experiment with a baseline and a perturbation block in 69 subjects. State noise, output noise and adaptation rate in the baseline and perturbation block were extracting using Bayesian fitting of a trial-to-trial state-space model. We found that adaptation rate in the perturbation block correlates positively with baseline state noise (r=0.27; 95%HDI=[0.05 0.50]) and negatively with baseline output noise (r=-0.41; 95%HDI=[-0.63 -0.16]). In addition, the steady-state Kalman gain calculated from baseline state and output noise correlated positively with adaptation rate in the perturbation block (r = 0.31; 95%HDI = [0.09 0.54]). Therefore, noise can be viewed both as a supporting factor for motor adaptation (state noise) and as a noise factor hampering optimal performance (output noise), and in order to understand the relationship of noise to learning, one must decompose noise into its constituent components.

KEYWORDS

Motor learning, noise, genetics, visuomotor adaptation, motor planning, human

INTRODUCTION

Humans have a unique ability to generate an almost infinitely diverse set of movements. This ability hinges on neuronal networks to orchestrate movements, on muscles cells to execute movement plans and on sensory cells to feedback the result [1]. Because all these processes are noisy (see [2] for a review), an important property of the motor system is that movements are variable and can never be exactly reproduced. Noise which affects movement in a dimension relevant for a particular task might therefore deteriorate performance [3]. Rather than movement noise being a purely negative trait though, the nervous system exploits movement noise for motor learning. In reward-based learning, movement noise is increased in novel or unrewarding situations to facilitate exploration of the task, and is attenuated once task details have been learned to maximize performance [4–7]. Similar mechanisms seem to be at play in error-based motor learning in humans and might explain differences in motor adaptation rate between subjects. During reaching movements in a novel force field, subjects with more movement noise during a baseline block were found to adapt more rapidly to a force field perturbation [8]. However, a recent meta-analysis of adaptation experiments could not confirm the positive relation between movement noise and motor adaptation rate as a general rule [9]. Perhaps only some components of movement noise increase adaptation rate whereas others might even decrease adaptation rate [9].

Trial-to-trial models of visuomotor adaptation have decomposed movement noise into state noise and output noise [9–11] (see Figure 1A). State noise is believed to arise from variability in processing of sensory information, computations underlying adaptation and maintenance of the states in time, which are predominantly neuronal processes [10]. Indeed, noise in preparatory areas of the brain can be linked to behavioral variability using electrophysiology in macaques [12–14] and fMRI in humans [15]. This noise source can be manipulated by dedicated neural networks. In Bengalese finches birds, a basal ganglia-premotor loop learns a melody from reward [16] and injects noise [4] to promote exploration [5] during training [6] and development [7]. Probably, dopamine acts to increase this neuronal noise source [17,18]. Indeed, explorative behavior and motor learning rate have been linked to a single nucleotide polymorphisms in BDNF [19,20], which affects synaptic plasticity [21,22], and to single nucleotide polymorphisms in dopamine receptors and enzymes [23,24]. Output noise is thought to originate from the sensorimotor pathway and encompasses sensory and muscle noise [10]. This noise source increases with muscle force because vigorous contraction recruits larger motor units which tend to fire at a lower frequency and therefore produce more unfused twitches [25,26]. Differences in output noise between individuals might be explained by the strength of the muscle [27] and the activation pattern [28], which is influenced by genetics and training, and the quality of the sensors, which could be decreased in for example polyneuropathy. How do these noise sources relate to movement adaptation rate?

If humans learn (close to) optimally from movement error, the relation between adaptation rate and the components of noise is governed by a steady-state Kalman filter [29]. Kalman filter design makes that adaptation rate should increase with state noise and decrease with output noise [30]. Practically, this model predicts that people with more state noise will adapt faster (see Figure 1B and Figure 1D) whereas people with more output noise will adapt slower (Figure 1C and Figure 1D). However, even though Kalman filters have been shown to capture properties of the motor system [29,31], no attempts have been made to relate the distinct components of movement noise to adaptation rate. We predict that state noise would be associated with more rapid motor adaptation rate whereas output noise would be associated with slower visuomotor adaptation. To test the association between state noise,

output noise and adaptation rate, we performed a visuomotor adaptation experiment in 69 healthy subjects. We fitted the state-space model of trial-to-trial behavior (equations (1)-(4)) using Bayesian statistics to extract state noise, output noise and adaptation rate for each subject, and confirmed the hypothesis.

MATERIALS AND METHODS

Subjects

We included right-handed subjects aged 18 – 35 years between October 2016 and December 2016, without any medical conditions that might interfere with motor performance. Prior to the experiment, right-handedness was confirmed in all subjects using the Edinburgh handedness inventory. Any score higher than 0 was accepted. Subjects were recruited from the Erasmus MC University Medical Centre and received a small financial compensation. The study was performed in accordance with the Declaration of Helsinki and approved by the medical ethics committee of the Erasmus MC University Medical Centre.

Visuomotor adaptation

Experimental procedure

Subjects performed a visuomotor adaptation task [32,33] using a custom-built robotic device (previously described in [34,35]). They were seated in front of a horizontal projection screen while holding a robotic handle in their dominant right hand. The projection screen displayed the location of the robotic handle ("the cursor"), start location of the movement ("the origin"), and target location of the movement ("the target") (see Figure 2A). Position of the origin on the screen was fixed throughout the experiment, approximately 40 cm in front of the subject at elbow height. To remove direct visual feedback of hand position, subjects wore an apron that was attached to the projection screen around their neck.

Subjects were instructed to make straight shooting movements from the origin towards the target and to decelerate only when they passed the target. A trial ended when the distance between the origin and cursor was at least 10 cm or when trial duration exceeded 2 seconds. At this point, movements were damped with a force cushion (damper constant 3.5 Ns/m, spring constant 35 N/m) and the cursor was displayed at its last position until the start of the next trial to provide position error feedback. Furthermore, velocity feedback was given to keep movement velocity in a tight range. The target dot turned blue if movement time on a particular trial was too long (>600 ms), red if movement time was too short (<200 ms) and remained white if movement time was in the correct time range (200-600 ms). During presentation of position and velocity feedback, the robot pushed the handle back to the starting position. Forces were turned off when the handle was within 0.5 cm from the origin. Concurrently, the cursor was projected at the position of the handle again and subjects had to keep the cursor within 1 cm from the origin for 1 second to start the next trial.

The entire experiment lasted 900 trials with all three target directions (angle of -45°, 0° or 45°) occurring 300 times in random order. The experiment included vision unperturbed, vision perturbed and no vision trials (see Figure 2B). The three different trial types were used to build a baseline and a perturbation block (see Figure 2C). The baseline block contained a high number of no vision trials and was specifically designed to quantify movement variability. The perturbation block incorporated incremental small-step perturbation and was designed to extract motor adaptation rate.

Data Collection

The experiment was controlled by a C++ program developed in-house. Position and velocity of the robot handle were recorded continuously at a rate of 500 Hz. Velocity data was smoothened with an exponential moving average filter (smoothing factor=0.18s). Trials were analyzed from movement start (defined as the time point when movement velocity exceeds 0.03 m/s) to movement end (defined as the time point when the distance from the origin is equal to or larger than 9.5 cm). Aiming error was defined as the signed (+ or -) angle between the vector connecting origin and target and the vector connecting movement start and movement end including the visual perturbation. Clockwise errors were defined as positive, counter-clockwise errors as negative. Peak velocity was found by taking the maximum velocity in the trial interval. We calculated peak velocity to investigate its relation with state noise and output noise. Trials with (1) a maximal displacement below 9.5 cm, (2) an aiming error larger than 30° or (3) a duration longer than 2 seconds were removed from further analysis (2% of data).

State-space model estimation

We used a trial-to-trial learning model to capture state noise, output noise and adaptation rate [10]:

$$x[n+1] = Ax[n] - Be[n] + \eta \quad (1)$$

$$y[n] = x[n] + \epsilon \quad (2)$$

$$e[n] = y[n] + p[n] \quad (3)$$

$$\eta \sim N\left(0, \sigma_\eta^2\right), \epsilon \sim N(0, \sigma_\epsilon^2) \quad (4)$$

In this model, $x[n]$ is the internal aiming state, i.e. the motor command, and $y[n]$ the aiming output, i.e. the actual movement. Error e$[n]$ on a particular trial is the sum of $y[n]$ and the perturbation $p[n]$. The learning terms are A, which represents retention of the state over trials, and B, the fractional change from error $e[n]$. Aiming state and output are affected by noise processes η (state noise), modeled as a zero-mean Gaussian with variance σ_η^2, and ϵ (output noise), modeled as a zero-mean Gaussian with variance σ_ϵ^2. State noise is stored directly in the movement states and therefore leads to correlated fluctuations in movement behavior. In contrast, output noise ϵ is only stored indirectly in movement states and results in much faster, uncorrelated behavioral noise than state noise η.

We fitted the state-space model described in equations (1)-(4) to the data of individual subjects using Markov-chain Monte-Carlo sampling [36] implemented in OpenBugs (ver 3.2.3, OpenBugs Foundation available from: http://www.openbugs.net/w/Downloads) with three 50,000 samples chains and 20,000 burn-in samples. A single estimate per subject s was made for $A[s]$ and $B[s]$ using all trials. Separate estimates were made per subject in the baseline and perturbation block for $B_{Block}[s]$, $\sigma_{\eta,block}^2[s]$ and $\sigma_{\epsilon,block}^2[s]$ (similar to [8]). We used a logistic normal distribution as a prior for $A[s]$ and $B_{Block}[s]$, a normal distribution as a prior for $B[s]$ and an inverse gamma distribution as a prior for $\sigma_{\eta,block}^2[s]$ and $\sigma_{\epsilon,block}^2[s]$:

$$A[s] \sim \frac{1}{1 + \exp\left(N\left(\mu_A, \sigma_{A,subjects}^2\right)\right)}, \qquad B[s] \sim N\left(\mu_B, \sigma_{B,subjects}^2\right) \quad (5)$$

$$B_{Block}[s] \sim \frac{1}{1+\exp\left(N\left(B[s], \sigma^2_{B,blocks}\right)\right)} \tag{6}$$

$$\sigma^2_{\eta,block}[s] \sim 1/gamma(10^{-3}, 10^{-3}), \qquad \sigma^2_{\epsilon,block}[s] \sim 1/gamma(10^{-3}, 10^{-3}) \tag{7}$$

Priors for μ_A and μ_B were selected from a normal distribution and priors for $\sigma^2_{A,subjects}$, $\sigma^2_{B,subjects}$ and $\sigma^2_{B,blocks}$ from a gamma distribution:

$$\mu_A \sim N(logit(0.99), 10^{-3}), \qquad \mu_B \sim N(logit(0.04), 10^{-3}) \tag{8}$$

$$\sigma^2_{A,subjects} \sim \sigma^2_{B,subjects} \sim \sigma^2_{B,blocks} \sim 1/gamma(10^{-3}, 10^{-3}) \tag{9}$$

We used the mode of the samples per parameter and subject for further calculations. Similarly to Cheng and Sabes [10] we investigated the validity of the model estimates by correlating the estimates with the variance statistics of the data.

First, steady-state noise and lag-1 autocorrelation are closely linked to state noise σ_η and output noise σ_ϵ. Since learning is small in the baseline set, for this data we can neglect the effect of learning term B, in which case noise and lag-1 autocorrelation of the aiming direction can be expressed as:

$$\sigma_y = \sqrt{\left(\sigma^2_\epsilon + \sum_{k=0}^{\infty} A^{2k}\sigma^2_\eta\right)} \tag{10}$$

$$R(1) = \frac{\sum_{k=0}^{\infty}(A^{2k+1}\sigma^2_\eta)}{\sigma^2_\epsilon + \sum_{k=0}^{\infty} A^{2k}\sigma^2_\eta} \tag{11}$$

Therefore, aiming noise increases with both state noise σ_η (see figure 2D) and output noise σ_ϵ (see figure 2F) whereas aiming lag-1 autocorrelation increases with state noise σ_η (see figure 2E) but decreases with output noise σ_ϵ (see figure 2G). Figure 2H visualizes the effect of state and output noise on aiming direction, and we expect similar relations in the baseline block of our experiment.

Second, the covariance σ_{py} between the perturbation and aiming direction depends solely on the learning parameters A and B and is therefore useful to assess the validity of adaptation rate B in the perturbation block. The covariance σ_{py} becomes increasingly negative for higher adaptation rates (see Figure 2I). A visualization of slow and fast learners is given in Figure 2J. We expect a similar relation in the perturbation block of our experiment.

We investigated the relation between adaptation rate, state noise and output noise using the parameters $B_{Block}[s]$, $\sigma_{\eta,block}[s]$ and $\sigma_{\epsilon,block}[s]$. In addition, we calculated the steady-state Kalman gain for every subject $K_{Block}[s]$ in the baseline and perturbation blocks from $A[s]$, $\sigma_{\eta,block}[s]$ and $\sigma_{\epsilon,block}[s]$ by first solving the Riccati equation for the steady-state covariance $P_{\infty,block}[s]$:

$$A[s]^T P_{\infty,block}[s]A[s] - P_{\infty,block}[s] - A[s]^T P_{\infty,block}[s]\left(P_{\infty,block}[s] + \sigma_{\epsilon,block}[s]^2\right)^{-1} P_{\infty,block}[s]A[s] + \sigma_{\eta,block}[s]^2 = 0 \quad (12)$$

$$K_{Block}[s] = P_{\infty,block}[s]/(P_{\infty,block}[s] + \sigma_{\epsilon,block}[s]^2) \quad (13)$$

As control analyses for the fitting procedure, we generated two data sets for our experimental protocol using equations (1)-(4) and the estimated model parameters. In the first dataset, the relation between $B_{Block}[s]$, $\sigma_{\eta,block}[s]$ and $\sigma_{\epsilon,block}[s]$ was left unchanged (original dataset), whereas for the second dataset the noise parameters $\sigma_{\eta,block}[s]$ and $\sigma_{\epsilon,block}[s]$ were separately permuted in such a way that any regression coefficient between the noise parameters and the adaptation rate would be smaller than 0.05 (permuted dataset). For both datasets, we re-estimated the model parameters according to the procedure described above and expected high test-retest correlations. Second, we re-estimated the relation between movement noise and adaptation rate. We expected this relation to remain intact for the original dataset and to disappear for the permuted dataset.

Statistics

We calculated normalized Bayesian linear regression coefficients to investigate the relation between adaptation rate B_{Block} and noise $\sigma_{\eta,block}$ and $\sigma_{\epsilon,block}$ (Openbugs, three 50,000 samples chains and 20,000 burn-in samples). The dependent variable was modeled as a t-distribution with the regression model as the mean. As priors, we used uniform distributions (range -1 to +1) for the coefficients, normal distributions for the intercepts (zero mean, precision 10^{-6}), gamma distributions for the model error (shape and rate parameter 10^{-3}) and a shifted exponential prior (rate parameter 1/29) on the degrees of freedom [36]. This way, we evaluated (1) an intercept model, (2) an intercept with state noise $\sigma_{\eta,block}$ model, (3) an intercept with output noise $\sigma_{\epsilon,block}$ model and (4) an intercept with state noise $\sigma_{\eta,block}$ and output noise $\sigma_{\epsilon,block}$ model. The quality of a model was determined by calculating the difference in the deviance information criterion (DIC) between that model and the intercept model ($\Delta DIC = DIC_{Model} - DIC_{Intercept\ Model}$). The DIC assigns a score to a model by penalizing the complexity and rewarding the fit. Better models have lower DICs and better models therefore have a negative ΔDIC. In addition, we tested correlations between parameters with Bayesian Pearson correlation coefficients, using similar priors as for the linear regression.

Statistical results are reported as the mode of the effect size with 95% highest density intervals (HDIs). Model estimates are plotted as the mode with 68% HDIs, similar to the standard deviation interval.

RESULTS

Modelling learning and noise in reaching adaptation

The study sample consisted of 14 men and 55 women (age M=21 years, range 18 - 36 years). Mean handedness score was 79 (SD=19, range 45 – 100). Standard deviation of aiming direction (light brown) calculated across the 69 subjects illustrates the differences in movement behavior between people (Figure 3A). The state-space model (dotted green line) which captures retention of adaptation A,

learning from error B_{Block}, state noise $\sigma_{\eta,block}$ and output noise $\sigma_{\epsilon,block}$ shows good agreement with the average aiming direction (brown line).

Figures 3B and 3C show example subjects with low or high baseline state noise $\sigma_{\eta,baseline}$ (see Figure 3B) and low or high output noise $\sigma_{\epsilon,baseline}$ (see Figure 3C). Agreeing with our group level predictions (see Figures 2 D-G), we found a positive correlation between state noise $\sigma_{\eta,baseline}$ and aiming noise $\sigma_{y,baseline}$ (*r* = 0.30; 95%HDI = [0.08 0.54]), between state noise $\sigma_{\eta,baseline}$ and aiming lag-1 autocorrelation $R_{Baseline}(1)$ (*r* = 0.68; 95%HDI = [0.50 0.85]) and between $\sigma_{\epsilon,baseline}$ and aiming noise $\sigma_{y,baseline}$ (*r* = 1.00; 95%HDI = [0.96 1.00]) and a negligible correlation between $\sigma_{\epsilon,baseline}$ and aiming lag-1 autocorrelation $R_{Baseline}(1)$ (*r* = -0.06; 95%HDI = [-0.30 0.17]).

Example subjects with a low and high adaptation rate are shown in Figure 3H. Again, according to the model prediction (see Figure 2I), we found a negative correlation between adaptation rate $B_{Perturbation}$ and covariance σ_{py} on a group level (*r* = -0.83; 95%HDI = [-0.97 -0.69]).

Relation between state noise, output noise and adaptation rate

By comparing regression models with the DIC, we found that adaptation rate B_{Block} is best modeled with state noise $\sigma_{\eta,block}$ and output noise $\sigma_{\epsilon,block}$ (combined model) rather than with one of the noise processes (state model and output model) or only the intercept (intercept model) (see Table 1). This conclusion was valid for the within-baseline block regression, the between-baseline block and perturbation block regression and the within-perturbation block regression. In accordance with Kalman filter theory (see Figure 1B-1C), state noise $\sigma_{\eta,block}$ correlated positively with adaptation rate B_{Block} and output noise $\sigma_{\epsilon,block}$ correlated negatively with adaptation rate B_{Block} (see Table 1, and Figure 4A-B for the between-blocks results). This conclusion was also supported by positive correlations between steady-state Kalman gain $K_{Baseline}$, calculated from state noise $\sigma_{\eta,baseline}$ and output noise $\sigma_{\epsilon,baseline}$,, and adaptation rate $B_{Baseline}$ (*r* = 0.40; 95%HDI = [0.18 0.63] 188.0), between $K_{Baseline}$ and $B_{Perturbation}$ (*r* = 0.31; 95%HDI = [0.09 0.54]; see Figure 4C) and $K_{Perturbation}$, calculated from state noise $\sigma_{\eta,perturbation}$ and output noise $\sigma_{\epsilon,perturbation}$, and adaptation rate $B_{Baseline}$ (*r* = 0.39; 95%HDI = [0.17 0.62]).

As control measurements, we investigated for a generated ordered and permuted dataset (1) test-retest correlations of the model parameters and (2) the relations between $\sigma_{\eta,block}$, $\sigma_{\epsilon,block}$ and B_{Block}. High test-retest correlations were found for both the ordered and permuted dataset (see Table 2). Second, adaptation rate B_{Block} was better modeled for the ordered dataset but not the permuted dataset with state noise $\sigma_{\eta,block}$ and output noise $\sigma_{\epsilon,block}$ than with only an intercept (see Table 1). In addition, the correlations between state noise $\sigma_{\eta,block}$ and adaptation rate B_{Block} and output noise $\sigma_{\epsilon,block}$ and adaptation rate B_{Block} remained intact for the ordered dataset but disappeared for the permuted dataset. These results indicate that the relation between state noise $\sigma_{\eta,block}$, output noise $\sigma_{\epsilon,block}$ and adaptation rate B_{Block} originate from the data rather than the fitting procedure.

Finally, we investigated how state and output noise correlated with baseline peak velocity. A negligible correlation was found between baseline peak velocity and baseline state noise *r* = 0.03; 95%HDI=[-0.20 0.25]; whereas a small positive correlation was found between baseline peak velocity and baseline output noise *r* = 0.23; 95%HDI=[0.00 0.47].

DISCUSSION

We investigated the relation between components of movement noise and visuomotor adaptation rate. If humans learn (close to) optimally from movement error, it can be predicted from Kalman theory that state noise correlates positively with adaptation rate and output noise negatively [29]. To test this hypothesis, we performed a visuomotor reaching adaptation experiment in 69 subjects and extracted state noise, output noise and adaptation rate using a state-space model of trial-to-trial behavior. Indeed, we found that adaptation rate during the perturbation block correlates positively with baseline state noise (r=0.27; 95%HDI=[0.05 0.50]) and negatively with baseline output noise (r=-0.41; 95%HDI=[-0.63 -0.16]). In addition, the steady-state Kalman gain calculated from baseline state and output noise correlated positively with adaptation rate in the perturbation block (r = 0.31; 95%HDI = [0.09 0.54]). Therefore, noise can be viewed both as a supporting factor for motor adaptation (state noise) and as a negative factor hampering optimal performance (output noise). In order to understand the relationship of noise to learning, one must decompose noise into its constituent components.

State noise and output noise in error-based learning

Wu et al. showed that subjects with more noise in force production during unperturbed reaching movement in a baseline block had a higher adaptation rate when counteracting a novel velocity-dependent force field in a perturbation block [8]. Using similar state-space models to the one used in this paper [37,38] it would be possible to dissociate force noise in state and output noise as well. In light of our findings, we expect state noise to be the dominant process in their baseline block. In contrast, He et al. could not confirm any relation between movement noise and adaptation rate in a meta-analysis of adaptation experiments [9]. Possibly, this null result stems from lumping state and output noise, which we found to have opposing effects, together as a single noise source. However, it is also conceivable that the design of the experiments in the meta-analysis was not optimal for characterizing the subtle relation between noise and adaptation rate. For example, one of the five studies in the meta-analysis [39] used a large sudden 30° perturbation to characterize adaptation rate, which might activate explicit learning mechanisms [40,41]. We believe future studies into noise and learning would benefit from Bayesian state-space modeling of data and a system identification approach to the design of the experiment.

The directions of correlations between state noise and adaptation rate and output noise and adaptation rate agree with predictions made using Kalman filter theory. This means that humans apply a (near-)optimal strategy for visuomotor adaptation. Indeed, previous experiments have shown that steady-state Kalman filters make appropriate predictions for (1) hand localization in the presence of supporting and resistive forces [29] and (2) visuomotor adaptation in the presence of varying error-signal reliability [31]. However, this prediction seems valid only for steady-state behavior: humans probably do not update their adaptation rate on every trial as is true for a Kalman filter [31]. Our study adds that the motor system also takes a Kalman-like approach to the calculation of the adaptation rate based on intrinsic levels of state noise and output noise.

State and output noise and optimal control

Wu et al. argue that their findings challenge the optimal control theory of movement [42,43] because the optimal control framework does not account for the motor system shaping and using motor noise to promote motor learning [8]. However, by dissociating motor noise into state noise and output noise these

two viewpoints might be reconciled. We hypothesize that the motor system uses strategies to minimize output noise arising from the sensorimotor system, agreeing with optimal control theory of movement, and regulates state noise in the central nervous system to promote explorative behavior, agreeing with the view from Wu et al. The motor system might control explorative behavior with state noise rather than output noise because state noise (1) is not signal-dependent [25], (2) can be fully regulated in the brain [4–7,12,16] , (3) can be easily stored in efference copies to create a movement history [44,45] and (4) correlates positively with learning rate in an optimal framework [30]. Accepting this perspective, optimal control models should incorporate an extra optimization term which regulates state noise to learn new motor strategies when performance is lower than expected.

Implications for rehabilitation

Motor adaptation is an important component of motor rehabilitation after brain damage. Skill acquisition (e.g mastering wheel chair skills or walking stairs with a hemiparesis) is a major part of rehabilitation programs. It involves acquiring new patterns of muscle activation over extended period ranging from days to months [46]. According to the optimal control model of movement, this process involves several steps relying on different areas of the brain: (1) acquiring an internal model that predicts sensory feedback for a given motor command (cerebellum), (2) combining these predictions with actual sensory information to form a belief about the states of the body (parietal cortex) and (3) setting feedback gains to optimally guide movement during execution (motor cortex) [1]. Brain injury changes the relation between a motor command and sensory feedback and therefore necessitates reacquiring (1) proper internal models, which is similar to movement adaptation and (2) optimal feedback gains though extensive practice. Helping patients relearn these internal models in an early stage after brain damage by providing error-based feedback of movement performance might therefore support skill learning and prevent adopting suboptimal movement in an early stage (compensation). Our results highlight the importance of analyzing the sources of movement variability to optimize such therapies as patients with high state variability are expected to learn fast and require few practice sessions whereas patients with high output variability are expected to learn slowly and require many practice sessions. In addition, our results suggest that interventions which can promote state variability, either behaviorally (reward-based feedback [47]) or pharmaceutically (transcranial random noise stimulation [48], dopamine agonists [18]) might help relearn movements after injury.

REFERENCES

1. Shadmehr, R. & Krakauer, J. W. A computational neuroanatomy for motor control. *Exp. brain Res.* **185,** 359–81 (2008).
2. Faisal, A. A., Selen, L. P. J. & Wolpert, D. M. Noise in the nervous system. *Nat. Rev. Neurosci.* **9,** 292–303 (2008).
3. Trommershäuser, J., Gepshtein, S., Maloney, L. T., Landy, M. S. & Banks, M. S. Optimal Compensation for Changes in Task-Relevant Movement Variability. *J. Neurosci.* **25,** (2005).
4. Kao, M. H., Doupe, A. J. & Brainard, M. S. Contributions of an avian basal ganglia-forebrain circuit to real-time modulation of song. *Nature* **433,** 638–43 (2005).
5. Tumer, E. C. & Brainard, M. S. Performance variability enables adaptive plasticity of 'crystallized'

adult birdsong. *Nature* **450,** 1240–4 (2007).

6. Stepanek, L. & Doupe, A. J. Activity in a cortical-basal ganglia circuit for song is required for social context-dependent vocal variability. *J. Neurophysiol.* **104,** 2474–86 (2010).
7. Olveczky, B. P., Andalman, A. S. & Fee, M. S. Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biol.* **3,** e153 (2005).
8. Wu, H. G., Miyamoto, Y. R., Gonzalez Castro, L. N., Ölveczky, B. P. & Smith, M. A. Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. *Nat. Neurosci.* **17,** 312–21 (2014).
9. He, K. *et al.* The Statistical Determinants of the Speed of Motor Learning. *PLOS Comput. Biol.* **12,** e1005023 (2016).
10. Cheng, S. & Sabes, P. N. Calibration of visually guided reaching is driven by error-corrective learning and internal dynamics. *J. Neurophysiol.* **97,** 3057–69 (2007).
11. Cheng, S. & Sabes, P. N. Modeling Sensorimotor Learning with Linear Dynamical Systems. *Neural Comput.* **18,** 760–793 (2006).
12. Chaisanguanthum, K. S., Shen, H. H. & Sabes, P. N. Motor variability arises from a slow random walk in neural state. *J. Neurosci.* **34,** 12071–80 (2014).
13. Churchland, M. M., Santhanam, G. & Shenoy, K. V. Preparatory activity in premotor and motor cortex reflects the speed of the upcoming reach. *J. Neurophysiol.* **96,** 3130–46 (2006).
14. Churchland, M. M., Afshar, A. & Shenoy, K. V. A Central Source of Movement Variability. *Neuron* **52,** 1085–1096 (2006).
15. Haar, S., Donchin, O. & Dinstein, I. Individual movement variability magnitudes are predicted by cortical neural variability. *bioRxiv* (2017). at <http://biorxiv.org/content/early/2017/01/03/097824>
16. Charlesworth, J. D., Warren, T. L. & Brainard, M. S. Covert skill learning in a cortical-basal ganglia circuit. *Nature* **486,** 251–5 (2012).
17. Costa, R. M. *et al.* Rapid Alterations in Corticostriatal Ensemble Coordination during Acute Dopamine-Dependent Motor Dysfunction. *Neuron* **52,** 359–369 (2006).
18. Galea, J. M., Ruge, D., Buijink, A., Bestmann, S. & Rothwell, J. C. Punishment-induced behavioral and neurophysiological variability reveals dopamine-dependent selection of kinematic movement parameters. *J. Neurosci.* **33,** 3981–8 (2013).
19. Fritsch, B. *et al.* Direct current stimulation promotes BDNF-dependent synaptic plasticity: potential implications for motor learning. *Neuron* **66,** 198–204 (2010).
20. McHughen, S. A. *et al.* BDNF val66met polymorphism influences motor system function in the human brain. *Cereb. Cortex* **20,** 1254–62 (2010).
21. Park, H. & Poo, M. Neurotrophin regulation of neural circuit development and function. *Nat. Rev. Neurosci.* **14,** 7–23 (2013).
22. Egan, M. F. *et al.* The BDNF val66met polymorphism affects activity-dependent secretion of BDNF and human memory and hippocampal function. *Cell* **112,** 257–269 (2003).
23. Frank, M. J., Doll, B. B., Oas-Terpstra, J. & Moreno, F. Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nat. Neurosci.* **12,** 1062–8 (2009).
24. Pearson-Fuhrhop, K. M., Minton, B., Acevedo, D., Shahbaba, B. & Cramer, S. C. Genetic variation

in the human brain dopamine system influences motor learning and its modulation by L-Dopa. *PLoS One* **8,** e61197 (2013).

25. Harris, C. M. & Wolpert, D. M. Signal-dependent noise determines motor planning. *Nature* **394,** 780–4 (1998).
26. Jones, K. E., Hamilton, A. F. & Wolpert, D. M. Sources of signal-dependent noise during isometric force production. *J. Neurophysiol.* **88,** 1533–44 (2002).
27. de C. Hamilton, A. F., Jones, K. E. & Wolpert, D. M. The scaling of motor noise with muscle strength and motor unit number in humans. *Exp. Brain Res.* **157,** 417–430 (2004).
28. Selen, L. P. J., Beek, P. J. & Dieën, J. H. van. Can co-activation reduce kinematic variability? A simulation study. *Biol. Cybern.* **93,** 373–381 (2005).
29. Wolpert, D. M., Ghahramani, Z. & Jordan, M. I. An internal model for sensorimotor integration. *Science* **269,** 1880–2 (1995).
30. Kalman, R. E. A New Approach to Linear Filtering and Prediction Problems. *J. Basic Eng.* **82,** 35 (1960).
31. van Beers, R. J., Sabes, P., Gon, J. D. van der, Ghez, C. & Brenner, E. How Does Our Motor System Determine Its Learning Rate? *PLoS One* **7,** e49373 (2012).
32. Hadipour-Niktarash, A., Lee, C. K., Desmond, J. E. & Shadmehr, R. Impairment of retention but not acquisition of a visuomotor skill through time-dependent disruption of primary motor cortex. *J. Neurosci.* **27,** 13413–9 (2007).
33. Tseng, Y.-W. 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).
34. 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,** (2009).
35. Donchin, O. *et al.* Cerebellar regions involved in adaptation to force field and visuomotor perturbation. *J. Neurophysiol.* **107,** 134–47 (2012).
36. Kruschke, J. K. . *Doing Bayesian data analysis : a tutorial with R, JAGS, and Stan*. (Academic Press, 2010).
37. Thoroughman, K. A. & Shadmehr, R. Learning of action through adaptive combination of motor primitives. *Nature* **407,** 742–7 (2000).
38. Donchin, O., Francis, J. T. & Shadmehr, R. Quantifying generalization from trial-by-trial behavior of adaptive systems that learn with basis functions: theory and experiments in human motor control. *J. Neurosci.* **23,** 9032–45 (2003).
39. Fernandes, H. L., Stevenson, I. H., Kording, K. P., Movshon, J. & Nawrot, M. Generalization of Stochastic Visuomotor Rotations. *PLoS One* **7,** e43016 (2012).
40. Taylor, J. A., Ivry, R. B., Bizzi, E., Bastian, A. & Tenenbaum, J. Flexible Cognitive Strategies during Motor Learning. *PLoS Comput. Biol.* **7,** e1001096 (2011).
41. Taylor, J. A., Krakauer, J. W. & Ivry, R. B. Explicit and implicit contributions to learning in a sensorimotor adaptation task. *J. Neurosci.* **34,** 3023–32 (2014).
42. Todorov, E. & Jordan, M. I. Optimal feedback control as a theory of motor coordination. *Nat. Neurosci.* **5,** 1226–35 (2002).
43. Todorov, E. Optimality principles in sensorimotor control. *Nat. Neurosci.* **7,** 907–15 (2004).

44. Voss, M., Ingram, J. N., Haggard, P. & Wolpert, D. M. Sensorimotor attenuation by central motor command signals in the absence of movement. *Nat. Neurosci.* **9,** 26–7 (2006).

45. Wolpert, D. M. & Ghahramani, Z. Computational principles of movement neuroscience. *Nat. Neurosci.* **3 Suppl,** 1212–7 (2000).

46. Krakauer, J. W. Motor learning: its relevance to stroke recovery and neurorehabilitation. *Curr. Opin. Neurol.* **19,** 84–90 (2006).

47. Galea, J. M., Mallia, E., Rothwell, J. & Diedrichsen, J. The dissociable effects of punishment and reward on motor learning. *Nat. Neurosci.* **18,** 597–602 (2015).

48. Prichard, G., Weiller, C., Fritsch, B. & Reis, J. Effects of different electrical brain stimulation protocols on subcomponents of motor skill learning. *Brain Stimul.* **7,** 532–40 (2014).

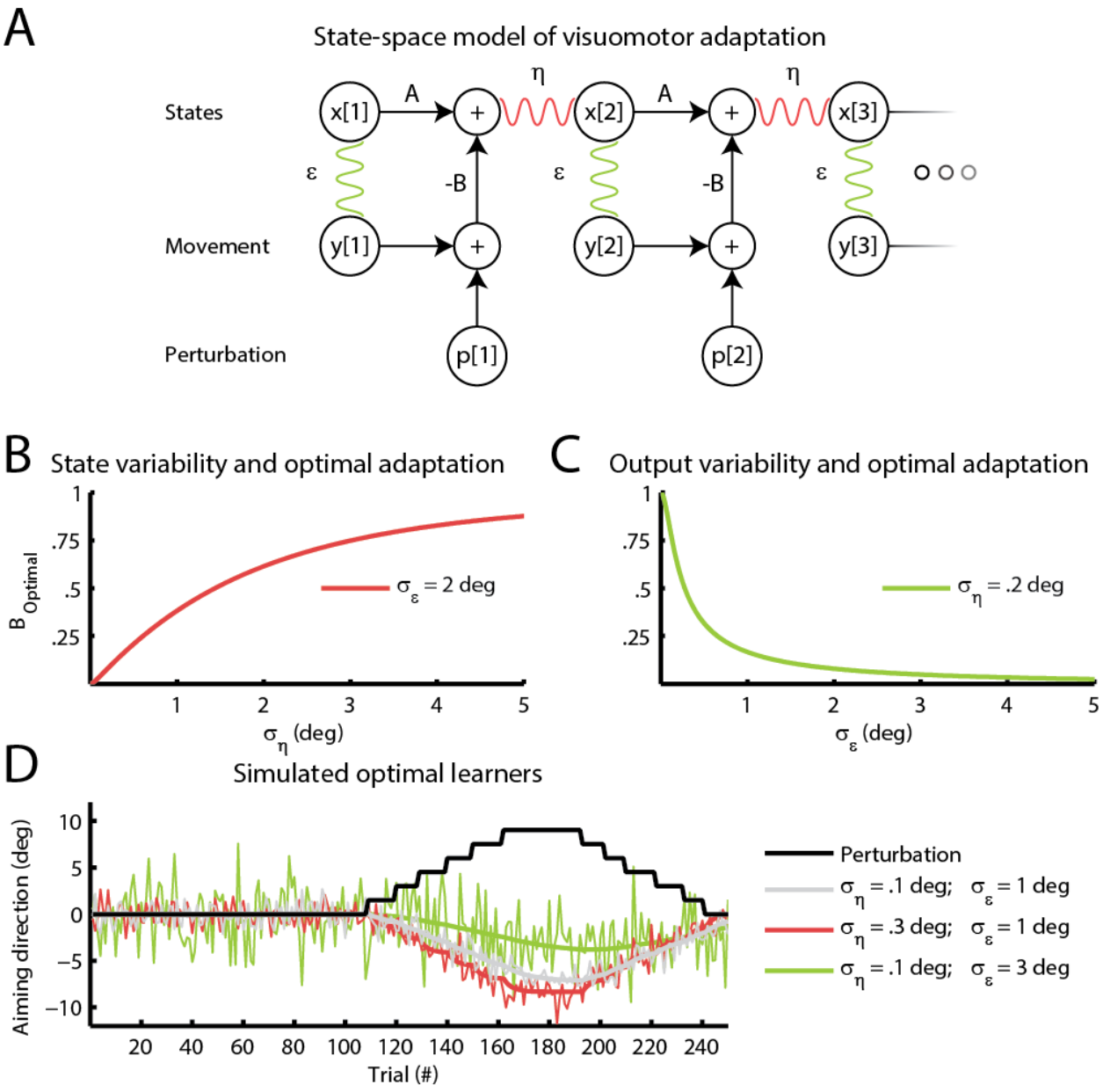
A
State-space model of visuomotor adaptation
States
Movement
Perturbation
x[1]
x[2]
x[3]
y[1]
y[2]
y[3]
p[1]
p[2]
A
η
ε
-B
+
B
State variability and optimal adaptation
B Optimal
σε = 2 deg
ση (deg)
C
Output variability and optimal adaptation
ση = .2 deg
σε (deg)
D
Simulated optimal learners
Aiming direction (deg)
Trial (#)
Perturbation
ση = .1 deg; σε = 1 deg
ση = .3 deg; σε = 1 deg
ση = .1 deg; σε = 3 deg

Figure 1 State and output noise have opposing effects on visuomotor adaptation.

- **A.** State-space model of visuomotor adaptation. Aiming directions are planned on trial $x[2]$ as a linear combination of the state on the previous trial $x[1]$ multiplied by a retentive factor A minus the error $e[1]$ on the previous trial multiplied with learning factor B. In addition, the movement state is distorted by the random process η. The actual aiming direction $y[2]$is the planned movement distorted by the random process ϵ. The error $e[1]$ is the sum of the aiming direction relative to the target $y[1]$ and external perturbation $p[1]$.
- **B.** State noise and optimal adaptation rate $B_{Optimal}$ (defined as the Kalman gain). The optimal adaptation rate increases with state noise σ_η. In this figure, σ_ϵ was kept constant at $2°$.
- **C.** Output noise and optimal adaptation rate $B_{Optimal}$ (defined as the Kalman gain). The optimal adaptation rate decreases with output noise σ_ϵ. In this figure, σ_η was kept constant at $0.2°$.
- **D.** Simulated optimal learners. At trial 110, a perturbation (black line) is introduced that requires the optimal learner to adapt their movement. The gray learner has low state noise $\sigma_\eta = 0.1°$ and output noise $\sigma_\epsilon = 1°$. The red learner has a higher state noise $\sigma_\eta = 0.3°$ than the gray learner $\sigma_\eta = 0.1°$. This causes the red learner to adapt faster. The green learner has a higher output noise than the gray learner $\sigma_\epsilon = 3°$. This causes the green learner to adapt more slowly. For all learners, the thick line shows the average, thin line a single realization.

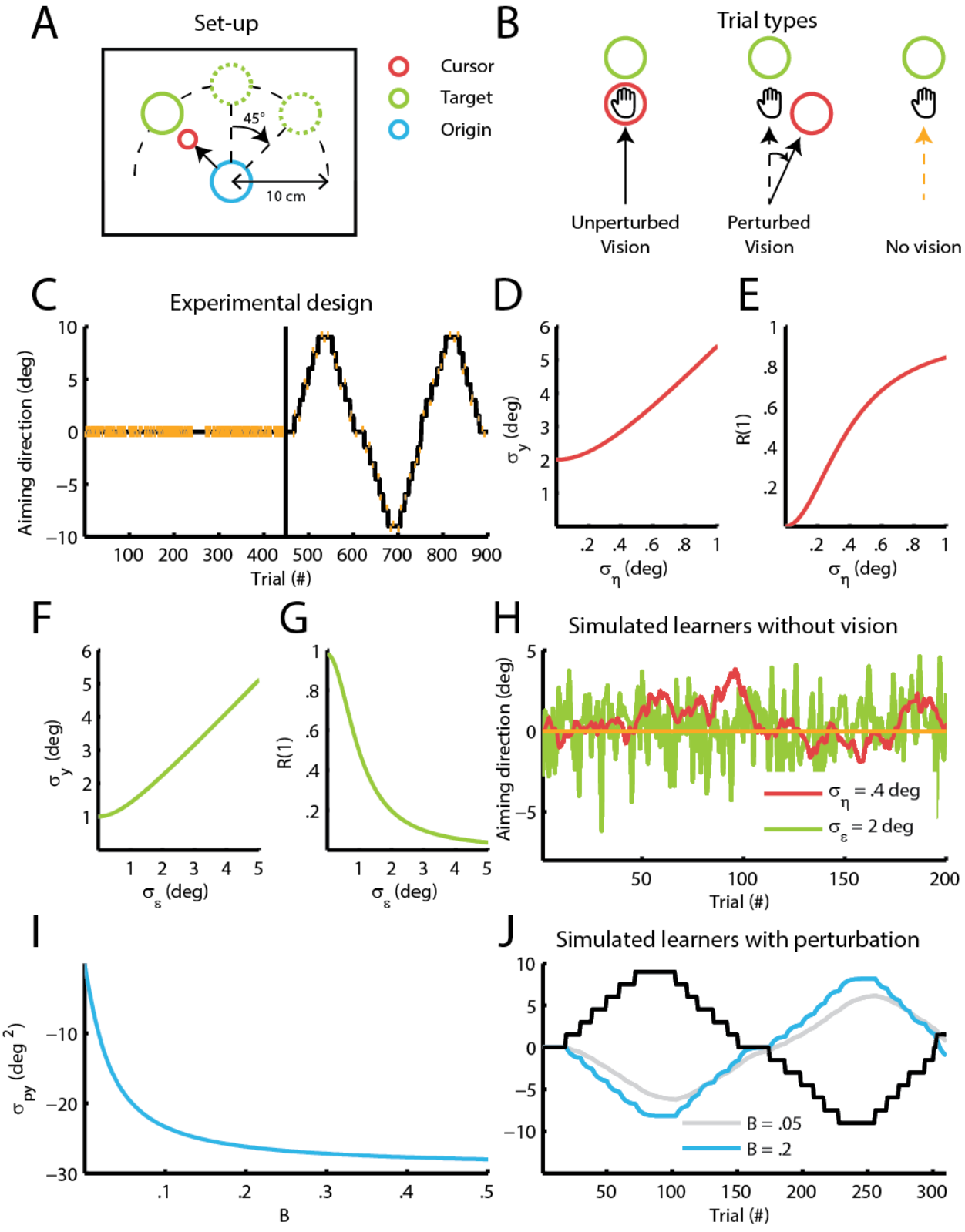
A
Set-up
45°
10 cm
Cursor
Target
Origin
B
Trial types
Unperturbed Vision
Perturbed Vision
No vision
C
Experimental design
Aiming direction (deg)
Trial (#)
D
σy (deg)
ση (deg)
E
R(1)
ση (deg)
F
σy (deg)
σε (deg)
G
R(1)
σε (deg)
H
Simulated learners without vision
Aiming direction (deg)
ση = .4 deg
σε = 2 deg
Trial (#)
I
σpy (deg 2)
B
J
Simulated learners with perturbation
B = .05
B = .2
Trial (#)

Figure 2 Measurements of state and output noise and adaptation rate in a visuomotor adaptation experiment.

- **A.** Set-up. The projection screen displayed the location of the robotic handle ("the cursor"; yellow circle 5 mm radius), start location of the movement ("the origin", white circle 5 mm radius), and target location of the movement ("the target", white circle 5 mm radius) on a black background The position of the origin on the screen was fixed throughout the experiment, while the target was placed 10 cm from the origin at an angle of -45°, 0° or 45°.
- **B.** Trial types. The experiment included vision unperturbed and perturbed trials and no vision trials. In vision unperturbed trials, the cursor was shown at the position of the handle during the movement. The cursor was also visible in vision perturbed trials but at a predefined angle from the vector connecting the origin and the handle. In no vision trials, the cursor was turned off when movement onset was detected and therefore only visible at the start of movement to help subjects keep the cursor at the origin.
- **C.** Experimental design. The baseline block consisted of 225 vision unperturbed trials and 225 no vision trials. Order of the vision unperturbed trials and no vision trials was randomized except for trials 181-210 (no vision trials) and trials 241-270 (vision unperturbed trials). We incorporated these non-random parts in order to allow a direct comparison of trials in which state noise played a larger and smaller role. The perturbation block had 50 no vision trials and 400 vision trials, with every block of nine trials containing one no vision trial. Every eight to twelve trials, the cursor was perturbed with an incremental 1.5° step. These steps started in the positive direction until reaching 9° and then switched sign to continue in the opposite direction until reaching -9°. This way, a perturbation signal was constructed with three "saw waves" lasting 150 trials each. The experiment was briefly paused every 150 trials.
- **D.** Simulation of noise and σ_η. Aiming noise σ_y increases with σ_η (calculated for $A = 0.98$, $B = 0$, $\sigma_\epsilon = 2°$).
- **E.** Simulation of autocorrelation and σ_η. Lag-1 autocorrelation $R(1)$ increases with σ_η (calculated for $A = 0.98$, $B = 0$, $\sigma_\epsilon = 2°$).
- **F.** Simulation of noise and σ_ϵ. Aiming noise σ_y increases with σ_ϵ (calculated for $A = 0.98$, $B = 0$, $\sigma_\eta = 0.2°$).
- **G.** Simulation of autocorrelation and σ_η. Lag-1 autocorrelation $R(1)$ decreases with σ_ϵ (calculated for $A = 0.98$, $B = 0$, $\sigma_\eta = 0.2°$).
- **H.** Simulated learners without vision. The green and red traces show a single realization of two learners with either high state noise (red learner $\sigma_\eta = 0.4°$ and $\sigma_\epsilon = 0°$) or high output noise (green learner $\sigma_\eta = 0°$ and $\sigma_\epsilon = 2°$). Both sources increase the aiming noise, but state noise leads to correlated noise whereas output noise leads to uncorrelated noise. This property can be seen from the relation between sequential trials. For the red learner sequential trials are often in the same (positive or negative) direction. For the green learner sequential trials are in random directions. This is captured by the lag-1 autocorrelation.
- **I.** Simulation of covariance and B. The covariance σ_{py} between the perturbation p and aiming direction y gets stronger for increasing B (simulated with $A = 0.98$).
- **J.** Simulated learners with perturbation. The gray and blue lines show a simulated slow ($A = 0.98$,

$B = 0.05$) and fast learner ($A = 0.98$, $B = 0.2$). The fast learner tracks the perturbation signal more closely than the slow learner. This property is captured by the covariance between the perturbation and the aiming direction.

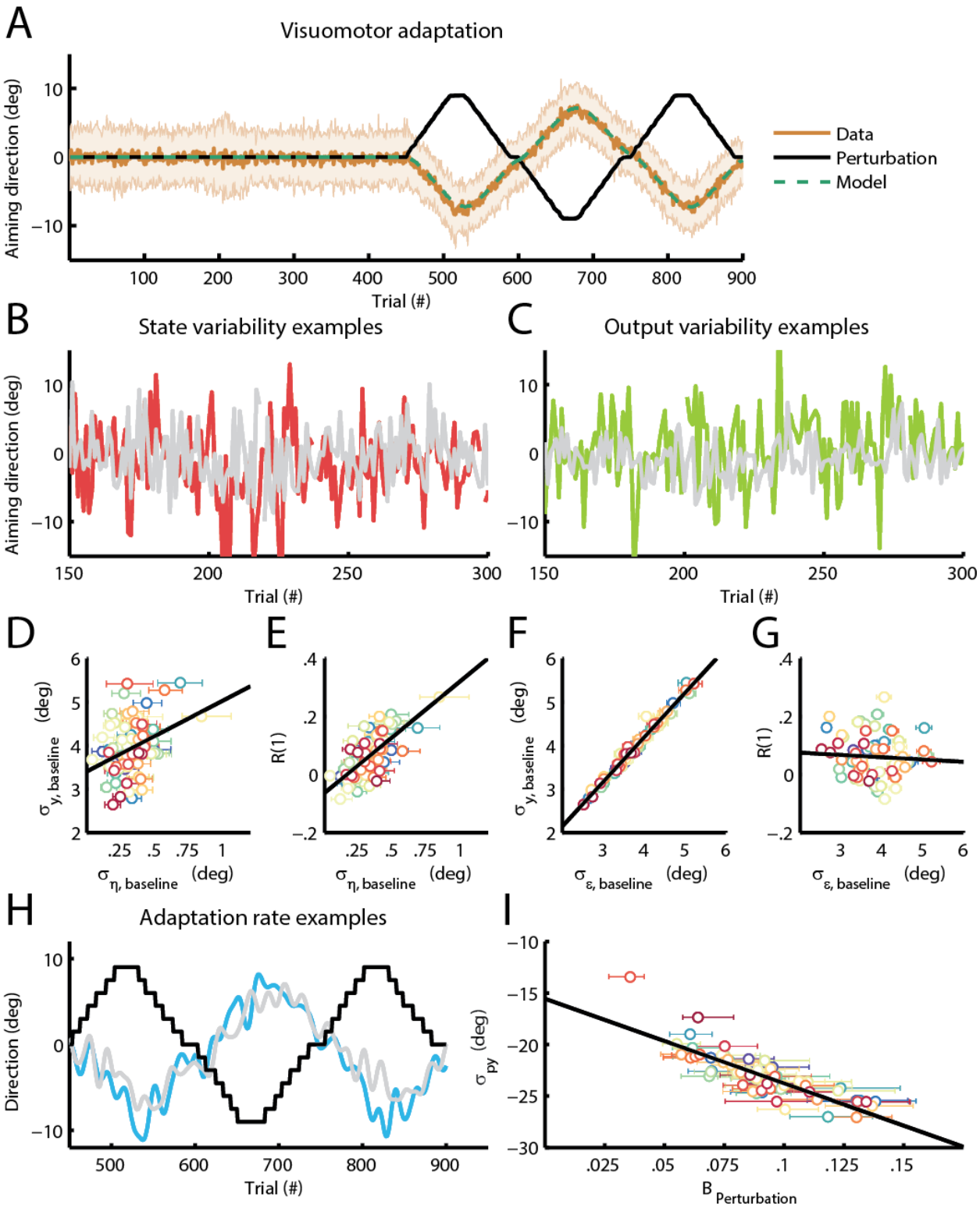
A
Visuomotor adaptation
Aiming direction (deg)
Trial (#)
Data
Perturbation
Model
B
State variability examples
C
Output variability examples
D
E
F
G
H
Adaptation rate examples
Direction (deg)
I
R(1)

Figure 3 State-space model of visuomotor adaptation.

- **A.** Visuomotor adaptation. Average aiming traces of the 69 subjects with standard deviations are shown in brown tone colors. The black indicates the average perturbation signal, the green line the model average.
- **B.** State noise examples. The gray line shows a subject with low state noise ($\sigma_{\eta,baseline} = 0.11°$ $\sigma_{\epsilon,baseline} = 4.0°$), the red line a subject with high state noise ($\sigma_{\eta,baseline} = 0.69°$ $\sigma_{\epsilon,baseline} = 5.0°$).
- **C.** Output noise examples. The gray line shows a subject with low output noise ($\sigma_{\eta,baseline} = 0.33°$ $\sigma_{\epsilon,baseline} = 2.7°$), the green line a subject with high output noise ($\sigma_{\eta,baseline} = 0.27°$ $\sigma_{\epsilon} = 5.1°$).

  D-G Relation between model estimates and baseline parameters. Models estimates and 68% confidence intervals are shown for every subject as a dot with error bars. The black line is a linear regression between the model estimates and baseline parameters. Panel **D** shows the relation between model estimate $\sigma_{\eta,baseline}$ and baseline parameter $\sigma_{y,baseline}$,, panel **E** the relation between model parameter $\sigma_{\eta,baseline}$ and baseline parameter lag-1 autocorrelation $R_{Baseline}(1)$, panel **F** the relation between model estimate $\sigma_{\epsilon,baseline}$ and baseline parameter noise $\sigma_{y,baseline}$ and panel **G** the relation between model estimate $\sigma_{\epsilon,baseline}$ and the baseline parameter lag-1 autocorrelation $R_{Baseline}(1)$.
- **H.** Adaptation rate examples. The thick lines show a slow (gray, $B = 0.055$) and fast subject (blue, $B = 0.14$) smoothened with a 6th order Butterworth filter. The black shows the perturbation signal for the fast subject.
- **I.** Input-output covariance. Relation between the model parameter $B_{Perturbation}$ and the input output covariance σ_{py} in the perturbation block. Models estimates and 68% confidence intervals are shown for every subject as a dot with error bars. The black line is a linear regression between the model estimates and baseline parameters.

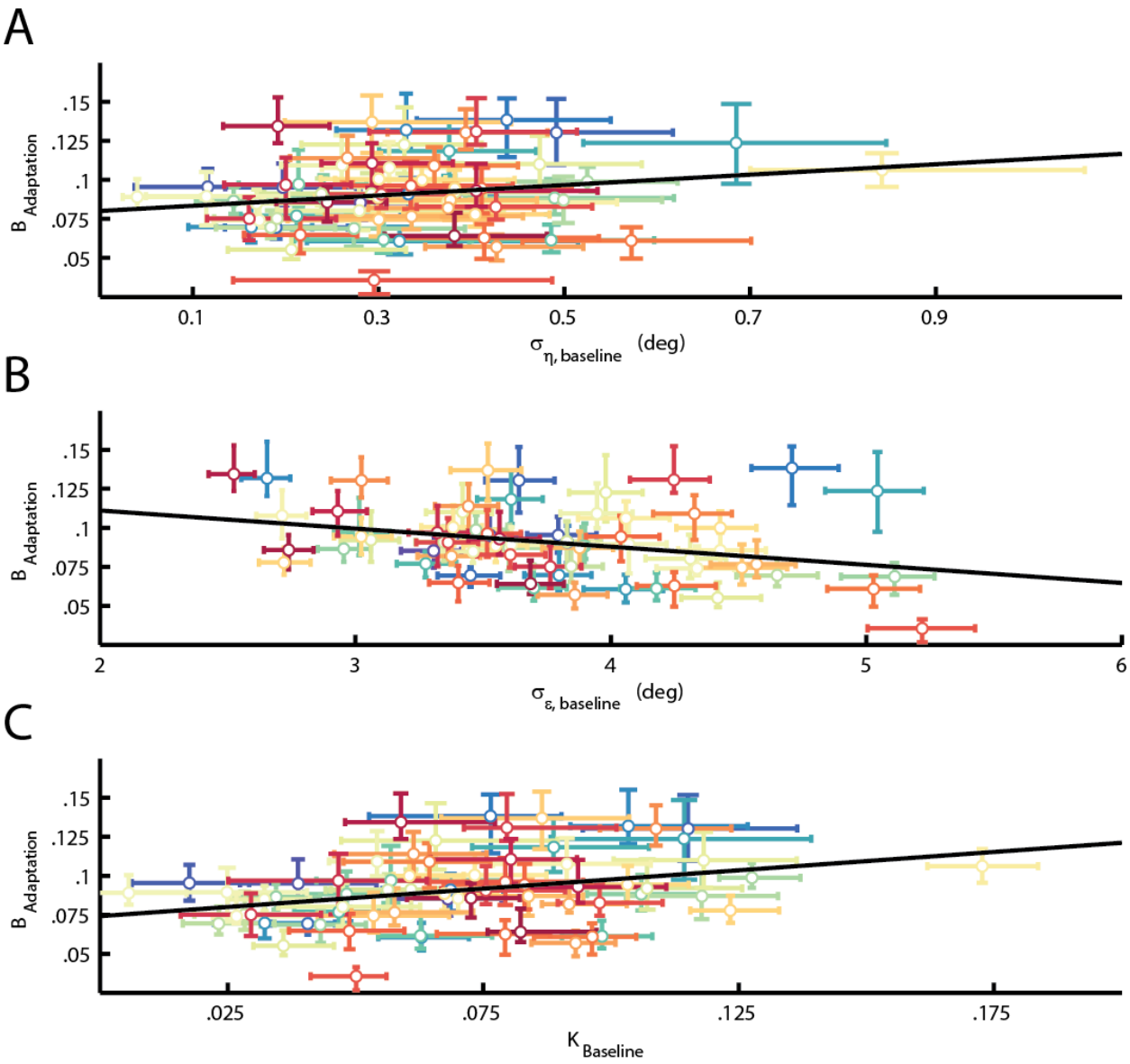
A
B Adaptation
.15
.125
.1
.075
.05
0.1
0.3
0.5
0.7
0.9
σ η, baseline (deg)
B
B Adaptation
.15
.125
.1
.075
.05
2
3
4
5
6
σ ε, baseline (deg)
C
B Adaptation
.15
.125
.1
.075
.05
.025
.075
.125
.175
K Baseline

Figure 4 Relation between noise and visuomotor adaptation.

A-C. Relation between model estimates. Models estimates and 68% confidence intervals are shown for every subject as a dot with error bars. The black line is a linear regression between the model estimates. Panel **A** shows the relation between the model estimates $\sigma_{\eta,baseline}$ and $B_{Perturbation}$, panel **B** the relation between the model estimates $\sigma_{\epsilon,baseline}$ and $B_{Perturbation}$ and panel **C** the relation between $K_{Baseline}$ and $B_{Perturbation}$.

TABLES

			Within-baseline block regression		Between blocks regression		Within-perturbation block regression	
	Model	Var.	Coefficient	ΔDIC	Coefficient	ΔDIC	Coefficient	ΔDIC
Original dataset	State	σ_η	0.32 [0.06 0.53]	-4.4	0.20 -0.04 0.44	-1.0	0.24 [0.00 0.48]	-1.8
	Output	σ_ϵ	-0.30 [-0.54 -0.06]	-3.9	-0.35 [-0.59 -0.11]	-6.2	-0.53 [-0.74 -0.31]	-18.7
	Combined	σ_η	0.38 [0.15 0.61]	-12.4	0.27 [0.05 0.50]	-9.9	0.17 [-0.05 0.38]	-18.8
		σ_ϵ	-0.36 [-0.60 -0.14]		-0.41 [-0.63 -0.16]		-0.49 [-0.70 -0.27]	
Generated dataset ordered	Combined	σ_η	0.25 [0.00 0.47]	-8.3	0.23 [0.01 0.48]	-8.1	0.26 [0.04 0.47]	-14.9
		σ_ϵ	-0.37 [-0.59 -0.13]		-0.37 [-0.61 -0.15]		-0.39 [-0.61 -0.17]	
Generated dataset permuted	Combined	σ_η	-0.09 [-0.33 0.16]	3.8	-0.05 [-0.31 0.19]	4.0	0.06 [-0.18 0.30]	2.5
		σ_ϵ	0.02 [-0.23 0.25]		-0.01 [-0.26 0.23]		0.14 [-0.09 0.39]	

Table 1. Regressions analysis for the adaptation rate B_{Block} with independent parameters state noise $\sigma_{\eta,block}$, and output noise $\sigma_{\epsilon,block}$. Adaptation rate B_{Block} was determined in the baseline block (first data column) or the perturbation block (second and third data column). The noise parameters $\sigma_{\eta,block}$ and $\sigma_{\epsilon,block}$ were determined in the baseline block (first and second data column) or the perturbation block (third data column). Linear regressions were calculated on the original dataset and two datasets generated from the estimated model parameters. For the permuted generated dataset both the noise parameters were permuted over the different subjects. We compared (1) an intercept with state noise $\sigma_{\eta,block}$ model, (2) an intercept with output noise $\sigma_{\epsilon,block}$ model and (3) an intercept with state noise $\sigma_{\eta,block}$ and output noise $\sigma_{\epsilon,block}$ model with an intercept model using the difference in the deviance information criterion. Model coefficients are reported as the mode of the posterior distribution with 95% highest density intervals. ΔDIC is the difference in the deviance information criterion of a model compared to the intercept model.

Parameter	Generated dataset ordered	Generated dataset permuted
$B_{Baseline}$	0.79 [0.65 0.93]	0.75 [0.61 0.91]
$B_{Adaptation}$	0.89 [0.79 0.99]	0.89 [0.78 0.99]
$\sigma_{\eta,baseline}$	0.79 [0.64 0.94]	0.87 [0.77 0.99]
$\sigma_{\eta,adaptation}$	0.70 [0.53 0.87]	0.73 [0.55 0.91]
$\sigma_{\epsilon,baseline}$	0.97 [0.93 1.00]	0.98 [0.93 1.00]
$\sigma_{\epsilon,adaptation}$	0.98 [0.93 1.00]	0.98 [0.94 1.00]

Table 2. Test-retest correlations for an ordered and permuted dataset generated from the model parameters. Correlation coefficients are reported as the mode of the posterior distribution with 95% highest density intervals.