Statistical determinants of visuomotor adaption along different dimensions during naturalistic 3D reaches

E. Ferrea[1]*, J. Franke[1,2], P. Morel[1,4†], A. Gail[1,2,3†]

1 German Primate Center, Sensorimotor Group, Goettingen, Germany.
2 University of Goettingen, Georg-Elias-Mueller Institute of Psychology, Goettingen, Germany,
3 Bernstein Center for Computational Neuroscience, Goettingen, Germany.
4 Univ. Lille, Univ. Artois, Univ. Littoral Côte d'Opale, ULR 7369 - URePSSS - Unité de Recherche Pluridisciplinaire Sport Santé Société, Calais, France.

† equal contribution

*Corresponding author: eferrea@dpz.eu

Abstract

Neurorehabilitation in patients suffering from motor deficits relies on relearning or re-adapting motor skills. Yet our understanding of motor learning is based mostly on results from one or two-dimensional experimental paradigms with highly confined movements. Since everyday movements are conducted in three-dimensional space, it is important to further our understanding about the effect that gravitational forces or perceptual anisotropy might or might not have on motor learning along all different dimensions relative to the body. Here we test how well existing concepts of motor learning generalize to movements in 3D. We ask how a subject's variability in movement planning and sensory perception influences motor adaptation along three different body axes. To extract variability and relate it to adaptation rate, we employed a novel hierarchical two-state space model using Bayesian modeling via Hamiltonian Monte Carlo procedures. Our results show that differences in adaptation rate occur between the coronal, sagittal and horizontal planes and can be explained by the Kalman gain, i.e., a statistically optimal solution integrating planning and sensory information weighted by the inverse of their variability. This indicates

that optimal integration theory for error correction holds for 3D movements and explains adaptation rate variation between movements in different planes.

Introduction

Studies of motor control and motor adaptation previously have been performed mostly in one or two-dimensional settings[1-7]. It is unclear, how well these findings generalize to movements in 3D, since 3D settings apply much less physical constraints on body pose and movements[8,9]. Unless movements are constrained in a horizontal plane by a table or supported by an exoskeleton, gravity induces a force anisotropy; unless movements are conducted in a frontoparallel plane, visual depth induces a perceptual anisotropy. It is unknown, if and how differences in the way subjects plan movements and in the way they perceive the environment and their own movement along different dimensions translate into motor adaptation anisotropy. Here we directly compared adaptation in different dimensions in the context of naturalistic movements in 3D.

Perceptual as well as other factors might contribute to adaptation anisotropies. Studies in 3D virtual reality (VR) found that adaptation along a vertical axis entails a reduced learning rate relative to the other two axes when a simultaneous triaxial perturbation was applied[10]. The authors proposed that their results are a consequence of the higher weight given to proprioception when adjusting movements along the vertical axis compared to sagittal or horizontal axis, considering/assuming that motor learning based on proprioceptive feedback entails a lower learning rate than based on visual feedback. We will refer the variability attributed to visual or proprioceptive feedback and affecting the estimate of the effector (hand) position as “measurement variability”[11]. However, not only measurement variability matters. According to optimal feedback control theory (OFC), the nervous system optimally (with respect to a specified goal) combines sensory feedback (e.g. visual and proprioceptive) information with a forward prediction of the

body's state to control movements and to correct errors[12-15]. During motor adaptation, this forward model is updated to reduce a sensory prediction error arising from a mismatch between its prediction and the sensory feedback. Any variability in this closed-loop process thus should affect the observed rate of adaptation. This includes variability in movement planning[11,16-21] which might result from the stochastic nature of neuronal processes in sensorimotor transformations, referred to as "planning variability", and which directly affects the reliability of the forward model prediction[22-26]. Here, we test the prediction of optimal integration theory that planning and measurement variability between different dimensions relative to the body determine the corresponding motor adaptation rates[16-19,27].

According to optimal integration theory, a statistically optimal solution for the update of the forward model during adaptation is the Kalman filter[28], which is updated from the error proportionally to the inverse of measurement variability and planning variability[11,29,30]. Higher variability in the forward model (high planning variability) would increase the Kalman gain and allow faster corrections of the experienced error, whereas a higher variability in the sensory feedback (high measurement variability) would decrease the gain and therefore lead to slower adaptation. Previous studies found that human behavior approaches that of an optimal learner[11,21,31] showing that adaptation rate positively correlates with the Kalman gain calculated from the planning and measurement variabilities. Also, individual differences between subjects in the rate of motor adaptation can be largely explained by individual differences in the Kalman gain[20], a prediction which we here test in the context of 3D movements. To study adaptation in naturalistic 3D movements, we leverage existing differences in planning and measurement variability in the different planes of space (coronal, sagittal, horizontal) instead of experimentally inducing variability. From previous work, we expect that the higher variability associated with depth perception reduces the rate of adaptation [14,32]. At the same time, planning variability might be elevated along the vertical axis given the stronger proprioceptive rather than visual component [33] due to the need of compensating gravitational forces [34-36].

Current experiments and models propose that sensorimotor adaptation is supported by two learning processes, which act on different timescales, one fast adapting and fast forgetting and one slow adapting and slow forgetting[7,37]. To estimate adaptation rates separately for each learning process and link them to Kalman gain, we will fit both fast- and slow-state processes during sensorimotor adaptation. Compared to one-state models[22,23,26], two-state models explain motor adaptation as a combination of explicit and implicit learning mechanisms, where explicit learning allows fast adaptation and implicit learning changes performance more slowly[4,6,38]. Two-state models also better explain phenomena like savings, i.e. improved learning due to repeated exposure[39-41], and anterograde interference, i.e. the negative influence of past learning on future learning[2,42], that was found to be partially correlated with explicit learning[5]. Moreover, neurophysiological evidence of two-state learning dynamics was found in humans and monkeys[43,44]. In a VR context, explicit learning was found to have a stronger influence on the final adaptation level than in non-VR adaptation paradigm[8]. Given that two-state dynamics were found to better model adaptive behaviours of visuomotor rotation paradigms, especially in VR settings, our method will be focused on a two-state space model but the results will be also compared with a one-state model.

By considering the stochastic nature of error-based learning, we can independently estimate the measurement and planning uncertainty from the empirical movement data during a 3D adaptation task using stochastic modelling. Stochastic modeling has been shown to better identify the time course of the hidden states[45]. Specifically, we use Bayesian inference via Hamiltonian Markov chain Monte Carlo sampling (referred to as Hamiltonian Monte Carlo, HMC) to implement a novel multilevel hierarchical version of the two-state space model fitting. Hierarchical modelling generally allows greater regularization of the subjects' parameters and simultaneous estimates of the parameter distributions ultimately resulting in less overfitting of the data[46]. We evaluated the hierarchical model and compared it to another state-of-the-art model performing Bayesian inference via an expectation-maximization (EM) algorithm[45] without a hierarchical structure. We show that HMC allows for a better extraction of data on the two

underlying motor learning states and the planning and measurement variabilities. Using hierarchical HMC models, we confirm that two-state models better explain our experimental three-dimensional data than single-state models. We, therefore, use the two-state hierarchical HMC model to test whether differences in adaptation dynamics can be found between movements and perturbations of different directions relative to the subject's own body and whether these differences comply with optimal integration theory.

Materials and Methods

Subjects

Data were collected from 26 healthy subjects (age range 20-32, 10 females, 16 males). The study was performed under institutional guidelines for experiments with humans, adhered to the principles of the Declaration of Helsinki, and was approved by the ethics committee of the Georg-Elias-Mueller-Institute for Psychology at the University of Goettingen. All the subjects were right-handed, had normal or corrected-to-normal vision and received financial compensation for their participation. Subjects gave written consent to participate and received detailed written instructions for the task. They were asked to paraphrase the instructions in their own words to make sure they understood the task. The experiment was carried out in one single appointment that lasted between 2.5 and 3 hours, depending on the subjects' freely chosen breaks among blocks (see later).

Virtual Reality Setup

Subjects sat in a darkened room at typical room temperature with no distracting noises. The 26 experiments were done individually at different times of the day and on different days. Subjects performed the visuomotor task in a three-dimensional (3D) virtual reality with Wheatstone stereoscopic view. Visual displays on two monitors were controlled by a customized C++ software similar to what was described in previous studies [47,48]. The images of the two monitors (one for each eye, BenQ XL2720T, 27-

129 inch diagonal, 1920x1080, 60 Hz refresh rate, BenQ, Taepei, Taiwan) were reflected by two semi-
130 transparent mirrors. (75 x 75 mm, 70/30 reflection/transmission plate beamsplitter, covered from the
131 back to block transmission; stock #46-643, Edmund Optics Inc., Barrington, New Jersey, United States)
132 that were angled 45° relative to the two monitors on each side of the subject (Fig. 1A).

133 The two screens render a 3D view of virtual objects on a dark background. To accommodate test subject
134 ergonomics, we tilted the plane of view (mirrors and monitors) downward from horizontal by 30° towards
135 the frontal direction (virtual horizontal). This ensured that arm movements in the physical space could be
136 carried out in an ergonomic posture in front of the body while the virtual representation of the hand (the
137 cursor) was aligned with the actual hand position. The virtual cube that defined the subjects' workspace
138 was centered on and aligned with the axis of gaze. For the rest of the paper, we describe orientation
139 relative to the axes of this cube, i.e. relative to the virtual, not the physical horizontal and vertical.

140 To optimize the 3D perception of the virtual workspace for each subject, the inter-pupillary distance was
141 measured (range: 51 to 68 mm, mean = 59.48 mm) and the 3D projection was modified in the custom-
142 made software accordingly. This has been shown to benefit depth perception, reduce image fusion
143 problems and reduce fatigue in participants[49].

144 Subjects performed controlled physical arm movements to control cursor movements in the 3D
145 environment. For this, they moved a robotic manipulator (delta.3, Force Dimension, Nyon, Switzerland)
146 with their dominant hand to place the 3D cursor position into visual target spheres. The manipulator
147 captured the hand position at a sampling frequency of 2 kHz with a 0.02 mm resolution. Subjects could
148 neither see their hand nor the robot handle. Instead, a cursor (yellow sphere, 6 mm diameter) indicated
149 the virtual representation of the robot handle and therefore the hand position (Fig. 1B).

150

151

Behavioral task

Subjects performed a 3D center-out reaching task, i.e. started their reaches from a central fixation position (indicated by sphere, semitransparent grey, 25 mm diameter) to one of eight potential target positions (grey sphere, 8 mm diameter) arranged at the corners of a 10 cm-sided cube centered on the fixation position (Fig. 1B-C). Each subject completed three blocks with breaks in between. One block consisted of 600 successful trials: 120 baseline trials, then 240 perturbed trials, and 240 washout trials again without a perturbation (Fig. 1D). Each block lasted between 20 to 25 minutes, depending on the performance of the subject. Breaks between blocks were adjusted according to the subject's needs and usually lasted between 10 to 15 minutes for compensating potential fatigue/savings/anterograde interference effects (see later). Before the actual experiment, a 10-minute training session was run to make subjects comfortable with the three-dimensional environment and to ensure a correct understanding and execution of the task.

At the beginning of each trial, subjects moved the cursor (co-localized with the physical hand position) into the central fixation sphere and held the cursor there for 200 ms ("hold fixation"; Fig. 1B). After this holding period, one of the eight potential targets was briefly highlighted by a spatial cue (semitransparent grey sphere of 25 mm diameter, centered on the target). The corners of the cube, and hence also the target, remained visible after the cue stimulus disappeared to allow precise localization of the target in the 3D space. The size of the spatial cue corresponded to the eligible area in which the movement had to end to be considered successful. The spatial cue remained visible for 300 ms and then disappeared together with the fixation sphere. This disappearance served as a GO signal for the subjects. Subjects then had 500 ms time to leave the fixation sphere and begin their hand movement towards the target. For successful completion of the trial, it was sufficient that the cursor completely entered the target sphere, without the need to stop inside the eligible area. Success or failure was indicated by sound signals: high-pitch for success, low-pitch for failure. A trial was aborted if subjects had not left the starting sphere 500

ms after the GO signal, if the movement was stopped (speed below 3cm/s) at any point outside the target area or lasted longer than 10 seconds. After each trial, subjects could immediately start a new trial by moving the cursor back to the starting position.

Within each block, we consistently applied the same 2D visuomotor perturbation to the 3D position of the cursor during all perturbation trials. Each perturbation affected the cursor position either within the xy-, the yz- or the xz-plane, respectively. This means, for every perturbation, one out of the three dimensions was left unperturbed, corresponding to the axis of rotation for the perturbation. In the perturbed trials, the workspace was rotated by 30 degrees counterclockwise around either the x (rotation in the sagittal plane), y (rotation in the horizontal plane) or z (rotation in the coronal plane) axes in each of the three blocks (Fig. 1D). The rotation was applied in each block to one of the three planes. As proof of concept, we verified that the projections of the trajectories on the perturbed planes look similar (Fig. 2A). The order of the blocks (perturbation planes) were controlled over the experiment using stratified randomization between subjects, to prevent potential effects of fatigue, interference or savings between blocks.

In the three blocks, the subjects conducted the same movements in terms of starting and end positions. We chose the corners of a cube as targets since any movement from the center of the cube to a corner always is a diagonal movement that is composed of equally large x-, y- and z-components. Therefore, the same target movement can be subjected to perturbations in each of the three planes parallel to the surfaces of the cube. This allowed us to compare perturbations in different body-centered dimensions without changing movement start and end positions, thereby avoiding potentially confounding effects of posture.

197 Within each block of trials, the order of the targets was pseudo-randomized such that all eight targets
198 were presented within each set of eight successful trials. These 8-trial sets counted as epochs and the
199 model fitting analysis was carried out on epochs rather than on single trials.

200 **Data preprocessing**

201 Data were stored for offline analysis performed in MATLAB® R2018a (MathWorks, Inc., Natick,
202 Massachusetts, United States). Data plots were generated with gramm, a MATLAB plotting library[50].
203 Trajectories were aligned in time to the start of the movement, defined as a speed increase above 3 cm/s,
204 to when the hand reached the target, and resampled in 50 time steps, so that each trajectory was
205 normalized in duration.

206 To calculate the spatial deviation of the hand position for each perturbed plane independent of target
207 direction, the trajectory of each trial was projected on the axis orthogonal to the center-to-target direction
208 and laying in the plane of the applied perturbation (Fig. 2B, perturbed dimension). To remove biases that
209 would result from curved trajectories independent of any perturbation, we computed the average
210 trajectory during baseline trials for each subject and block and subtracted these mean trajectories from
211 each trajectory in the according block. Figure 2C (middle and bottom panel) shows, for one example
212 subject, the average of the projected trajectories at baseline, during the first eight perturbation trials
213 (early adaptation) and the last eight perturbation trials (late adaptation) for all three perturbed planes.
214 While baseline-corrected trajectories are straight during baseline (by construction), the early adaptation
215 phase is characterized by curved trajectories with online movement corrections to reduce the target error
216 induced by the perturbed visual feedback. Later during adaptation trajectories on average were straight
217 again, i.e. the subject aimed at a direction suited to compensate the effect of perturbed feedback from
218 beginning of the movement. To quantify changes of the initial aiming direction over the course of

adaptation, we measure the deviation of the trajectory along the perturbed dimension at the moment when half the distance along the center-to-target dimension is travelled (Fig. 2C) and quantify it as angle.

Fitting model and procedure

Existing Bayesian tools for fitting adaptation models[20,45] can only estimate parameters from single experimental runs (changes over repeated trials in one task condition and one subject) despite the often hierarchical nature of the experiment (repeated trials interleaved with multiple conditions, multiple subjects).

To get more reliable estimates of subject-level and population-level learning parameters, we developed a hierarchical-model version of single-state and two-state models of adaptation. Hierarchical models for Bayesian inference allow to regularize the single subject fitted values by introducing a hyper-parameter at the population level. For this, we used the probabilistic programming language Stan[51], through its MATLAB interface. Stan allowed us to perform Bayesian inference with its implementation of a Hamiltonian Markov chain Monte Carlo method using a NUTS sampler algorithm[52] for fitting the model parameters to the empirical data.

The underlying trial-by-trial adaptation model for the movement angle corrections is the following. For the trial n in subject j and perturbation plane p, the measured movement angle $y_{jp}(n)$ is:

$$y_{jp}(n) = x_{jp}(n) + \varepsilon_{u,jp}(n) \quad (1)$$

With $x_{jp}(n)$ being the hidden internal state (planning angle) and $\varepsilon_{u,jp}(n)$ the measurement noise of the movement due to limited sensory accuracy of the agent, described by a normal distribution with zero mean: $\varepsilon_{u,jp}(n) \sim N(0, \sigma_{u,jp})$. We refer to this term as measurement noise since it reflects the uncertainty of estimating the position of the hand based on the sensory feedback with zero mean and standard deviation $\sigma_{u,jp}$ from which we calculated the measurement variability by squaring the term $\sigma_{u,jp}$ ($\sigma^2_{u,jp}$). We use this term to differentiate it from the uncertainty arising at the central brain level (planning noise; see below). In our experimental paradigm/model, the measurement noise is also influenced and cannot be disentangled from the execution noise that is added to the movement during conduction at the "periphery", i.e., it includes any movement variability that is accessible to the central nervous representations only via sensory feedback.

The internal state $x_{jp}(n+1)$ in trial $n+1$ depends on the state in the previous trial n, with retention **A**, and the error e in the previous trial, with learning rate **B**. The error e is expressed as $e_{jp}(n) = p - y_{jp}(n)$, where p is equal to thirty degrees during perturbed trials and zero degrees otherwise. The state update is affected by a planning noise (or state update error) $\varepsilon_x(n)$, which gives us the following:

$$x_{jp}(n+1) = A_{jp}x_{jp}(n) + B_{jp}e_{jp}(n) + \varepsilon_{x,jp}(n) \quad (2)$$

In the single-state model, A_{jp} and B_{jp} are positive scalars, and $\varepsilon_{x,jp}(n) \sim N(0, \sigma_{x,jp})$. $\sigma_{x,jp}$ is the standard deviation of the internal state error and allows calculating the planning noise by squaring the term $\sigma_{x,jp}$ ($\sigma^2_{x,jp}$). For the two-state model:

$$A_{jp} = \begin{bmatrix} a_{s,jp} & 0 \\ 0 & a_{f,jp} \end{bmatrix} \text{ with constraints } a_{s,jp} > a_{f,jp} > 0$$

$$B_{jp} = \begin{bmatrix} b_{s,jp} \\ b_{f,jp} \end{bmatrix} \text{ with constraints } b_{f,jp} > b_{s,jp} > 0$$

$$\varepsilon_{x,jp}(n) \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{x,jp} & 0 \\ 0 & \sigma_{x,jp} \end{bmatrix}\right)$$

In our model, all subject and plane-level parameters: $\sigma_{u,jp}$, $a_{s,jp}$, $a_{f,jp}$, $b_{s,jp}$, $b_{f,jp}$, $\sigma_{x,jp}$ were drawn from population-level normal distributions, e.g. for $a_{s,jp}$ we define $a_{s,jp} \sim N(\mu_{a_s,p}, \sigma_{a_s})$. Here, $\mu_{a_s,p}$ and σ_{a_s} are population-level hyper-parameters, corresponding to the population (between-subject) average of the slow retention for perturbation plane p and the population standard deviation of the slow retention.

To keep the model simple, we did not include common per-subject offsets in the average hyper-parameters across planes; in other words, the model we used did not take into account the fact that the same subjects participated in experimental runs with different perturbed planes. A variation of the model which permitted per-subject offsets yielded equivalent results (data not shown).

Bayesian inference requires priors for the hyper-parameters, which we chose wide to not constrain the model. Retention average hyper-parameter ($\mu_{a_s,p}$, $\mu_{a_f,p}$ in the two-state model and $\mu_{A,p}$ in the single-state model) priors were $\sim N(1,9)$, learning rate average ($\mu_{b_s,p}$, $\mu_{b_f,p}$ and $\mu_{B,p}$) priors were $\sim N(0,9)$, and their population standard deviation hyper-parameter ($\sigma_{a_s}, \sigma_{a_f}, \sigma_{b_s}, \sigma_{b_f}$ and σ_A, σ_B) priors were half-Cauchy distributions (location = 0, scale = 5)[53]. Priors of the population average and standard deviation hyper-parameters of the measurement and planning noise ($\mu_{\sigma_x,p}, \mu_{\sigma_u,p}$ and $\sigma_{\sigma_x}, \sigma_{\sigma_u}$) were half-Cauchy

distributions (location = 0, scale = 15). The conclusions from the model fitting do not critically depend on the exact choice of any of these parameters.

Steady-state Kalman gain

From the variances of the noise processes, we calculated the steady-state Kalman gain, as derived by Burge and colleagues[11]. In short, by combining measurement and state-update equations of the Kalman filter and assuming that a steady-state is reached, we can introduce the term $\sigma^2_{X_t^+}$ which consist on the variance of the best estimates given the two noisy measurements (see Burge et al., 2008, for full formula derivation):

$$\sigma^2_{X_t^+} = \frac{-\sigma_x^2+\sqrt{(\sigma_x^2)^2+4\sigma_x^2\sigma_u^2}}{2} \quad (3)$$

The Kalman gain (K), weighting the contribution of planning and measurement variabilities in error correction, is calculated as follows and indicates the amount of correction that should be attributed to the error:

$$K = \frac{\sigma^2_{X_t^+} + \sigma_x^2}{\sigma^2_{X_t^+} + \sigma_u^2 + \sigma_x^2} \quad (4)$$

Validation of fitting procedure

To validate our hierarchical Bayesian fitting algorithm, we compared it to a non-hierarchical algorithm by applying both to surrogate data sets. For this, we generated N = 100 artificial experimental runs using an underlying two-state time course with the same number of baseline, perturbation, and washout trials as

in our experiment. For each simulation, each parameter was drawn from a random normal distribution approaching the values we observed in our data: as~N(0.93,0.03), af~N(0.55,0.2), bs~N(0.06,0.04), bf~N(0.18,0.08), planning variability~N(1.5,1), measurement variability~N(6,3). We then compared the accuracy of parameter estimations from different fitting methods, namely an EM algorithm [45] and a non-hierarchical and a hierarchical HMC algorithm.

To quantify whether our two-state hierarchical algorithm introduces spurious correlations among the learning rates and the Kalman gain, we applied it to two versions of the surrogate datasets: One dataset was built with 200 completely uncorrelated samples, whereas in a second data set we imposed bs equal to the 0.3 times the Kalman gain calculated according to (4).

Results

The subjects performed the task (three conditions: baseline, perturbation and washout) with a high success rate in all perturbed planes (number of correct trials/number of total trials > 99%). We used a generalized linear mixed effect model (GLMM) with a binomial distribution to test the effect of the perturbation in the different planes (fixed effect) on the success rate, with the subjects as random effect. With the sagittal plane as intercept, there was no significant effect on the slope for the horizontal plane (tstat = 1.5477, p = 0.1217) and no significant effect for the slope of the coronal plane (tstat = 1.3976, p = 0.16223). This means that we did not observe differences in performance among the three perturbation planes.

The subjects showed stereotypical adaptation time courses (Figure 3A and supplementary figures 1-3) during and after introduction of the perturbation with typical error correction profiles that compensate for the applied rotation. To test differences among the three planes regarding the level of adaptation reached by the subjects (Fig. 3A), a one-way between-subjects ANOVA was conducted to compare the

321 effects of the perturbed plane over the final adaptation level. The final adaptation level was calculated as
322 the average of the adaptation time course over the last 4 epochs of the perturbation period (Fig. 3B). We
323 found that the subjects adapt differently to the perturbations depending on the plane of rotation ($F(2,69)$
324 $= 30.77$, $p = 2.8$ E-10). The final level of adaptation was significantly higher for perturbations in the coronal
325 plane compared to the sagittal plane (coronal-sagittal comparison, $p = 7.6$E-05, post-hoc multi-comparison
326 test corrected with Tukey's honestly significant difference) and horizontal plane (coronal-horizontal
327 comparison, $p = 1.0$ E-08) (Fig. 3B). Moreover, perturbations in the sagittal plane provoke stronger
328 adaptation than perturbations in the horizontal plane (sagittal-horizontal comparison, $p = 0.004$) (Fig. 3B).
329 To understand the underlying dynamic of the differences in adaptation, we modeled adaptation behavior
330 in the three different visuomotor perturbation planes with a two-state model, which separately identified
331 the slow and fast processes underlying adaptation. While our hierarchical Bayesian model fits data across
332 conditions and subjects (see Methods), Figure 3C shows the fitted curves for one example subject
333 performing three adaptation conditions.
334 Before comparing adaptation between the different planes, we evaluated whether in our three-
335 dimensional task, a two-state model fits the data better than a single-state model. The single-state model
336 shows larger root-mean-square residuals, i.e., larger difference between modeled and observed
337 adaptation trajectory (paired t-test, $p = 6.6$e-16). To take into account the number of free parameters when
338 comparing models, we additionally computed the widely applicable information criterion (WAIC), an
339 approximation of cross-validation for Bayesian procedures[54]. The two-state model shows a lower WAIC
340 compared to the single-state model (3.54E+04 and 3.58e+04, respectively), confirming the advantage on
341 using a two-stage model.
342 To validate our hierarchical fitting algorithm, we tested whether model parameters are better
343 reconstructed by our two-state Bayesian HMC hierarchical algorithm compared to a state-of-the-art EM
344 algorithm[45]. We also compared the HMC hierarchical model with the corresponding non-hierarchical HMC

model. For this, we generated multiple surrogate two-state adaptation datasets with randomly sampled parameters (see Methods). We then tested how well the three algorithms retrieve the parameters of the surrogate datasets. The hierarchical algorithm better reconstructs all parameters (error is smaller and correlation is higher) (Fig. 4). We therefore decided to use a two-state model and independently fit the parameters for each perturbation plane.

Optimal integration (OI) theory predicts that the rate of adaptation correlates with the weighing between internal estimation and external sensory information for assessing the state of the system. A learner optimally integrating these two types of information approximates a Kalman filter that combines them depending on their respective variabilities. If planning (internal estimate) variability is high and measurement (sensory information) variability low during adaptation, the experienced error will be given credibility, and therefore faster learning will happen. In the opposite case, the learner tends to disregard the experienced error and corrections would happen slowly.

We fitted all learning parameters of the two-state model and the measurement and planning variabilities from our empirical data (Fig. 5A). To test if the above-mentioned mechanism of OI can explain the observed differences between the perturbed planes in our experiment, we also correlated planning and measurement variabilities with the slow learning rate bs for each experimental run in our dataset (Fig. 5B). Correlations between planning variability and the learning rate as well as between measurement variability and learning rate were low (respective R2 : 0.13 and 0.11), but with a trend consistent to the OI predictions, i.e., low measurement noise and high planning noise correlate with higher values of the slow learning rate. When combining the contribution of the two noise parameters into an estimate of a Kalman gain for each subject, the slow learning rate was explained much better (R2 between K and bs: 0.48), with a high Kalman gain being associated with high learning rate. Notably, this correlation between Kalman gain (K) and slow learning rate (bs) is determined by the different levels for K and bs between the

368 perturbation planes. This means, the slow learning rate is mostly explained by the plane of the workspace
369 in which the perturbation is applied (R2 = 0.94), while it does not correlate with the Kalman gain variation
370 across the individual subjects within a perturbation condition. Last, the fast learning component shows a
371 correlation with the Kalman gain among perturbed planes (R2 between K and bf: 0.34).

372 A long washout plus additional waiting time between conditions prevented savings or anterograde effects
373 when switching between different planes of perturbation. To test for this, a forward stepwise linear
374 regression was used to identify possible predictors of the learning rates (bf or bs independently) out of
375 the order of execution (1st executed task, 2nd execution task or 3rd executed task) and perturbed axis.
376 At each step, variables were added based on p-values < 0.05. For both learning rates bs and bf, only the
377 perturbed axis was added to the model with p = 2.98E-44 and p = 8.22E-14 respectively. This means that
378 we did not observe a significant contribution of savings or anterograde effects on the learning rates.

379 Since the task uses multiple reach targets, one separate model could be fitted for each target unless one
380 assumes full generalization between sequentially visited targets. To avoid this pitfall and corresponding
381 increase in complexity, we fitted the model using epochs, motivated by the fact that in each epoch each
382 target was visited one time. When we also tested the model for the case where full generalization happens
383 between two consecutives targets, we found similar results as for the model trained on epochs (R2
384 between K and bs: 0.29, Supplementary Fig. 4).

385 We also validated the ability of our model to extract correlated learning parameters whenever present.
386 For this, we simulated an additional set of data where the slow learning rate bs was imposed to be 0.3
387 times the steady-state Kalman gain (see Methods). Figure 6A shows that we were able to reconstruct such
388 correlation with our model. Conversely, when no correlation between the slow learning rate and the
389 Kalman gain was introduced in the dataset, the model does not introduce spurious correlation. Figure 6
390 overall confirms the validity of our model for the study of correlations between model parameters.

Discussion

We tested if observations and model assumptions about optimal integration (OI) theory for motor learning, as previously derived from one- or two-dimensional highly confined reach tasks, also hold true for largely unconstrained movements in three-dimensional space. We show that the rate of visuomotor adaptation under 3D stereoscopic vision depends on the plane in which a 2D visuomotor rotation perturbation is applied. Subjects adapted best to perturbations in the coronal plane followed by sagittal and horizontal planes. The rates of slow and fast learning correlated well with the Kalman gain when considering the variation across perturbation conditions, but not across subjects within a condition. We could separately assess the within and across condition parameters with our novel hierarchical Bayesian HMC fitting of the data. These results show that important aspects of OI theories of motor control derived from 2D movements generalize to the largely unconstrained 3D movements studied here, at least at the average population level, but do not extend to the individual subject level for any of the applied perturbations[20].

A novel two-state hierarchical HMC algorithm for precise estimation of motor learning parameters

We developed a novel HMC Bayesian fitting procedure with a multilevel hierarchical structure to take into account effects from multiple subjects performing a motor adaptation task in three different learning conditions, i.e., when different planes relative to the body were perturbed. Naturally, such a hierarchical approach shows lower goodness-of-fit to single-subject adaptation curves than algorithms that focus on single-subject data since the hyper-parameter at the population level (i.e. all the samples coming from a same distribution) act as regularization term. But multi-subject experiments are typically designed to determine the effects of experimental variations on the motor learning behavior independent of test subject. Hence, it is most important to best estimate the underlying learning parameters.

For the current experimental design, our procedure is better able to reconstruct actual model parameters than a state-of-the-art EM fitting algorithm[45] or a non-hierarchical Bayesian HMC model, as revealed from the surrogate testing. This is due to the low identifiability of such two-state models for relatively short adaptation experiments comprising only a few dozen or even hundreds of trials per condition. The low identifiability is a result of random fluctuations in the shape of the adaptation curve and is easily associated with various learning parameters (see, for example, the outliers in Fig. 4). In the hierarchical model, each parameter is sampled from a population-level distribution, the parameters of which are also part of the model. This makes extreme values for the reconstructed parameter less likely.

Using our hierarchical model, we show that a two-state model fits the data significantly better than a single-state model. The two-state model is well established for subjects performing in a 2D environment[7,37] and better explains learning mechanisms such as savings and anterograde interference in 2D settings[38]. As previously hypothesized[4,6], the fast learning process can be partly associated to an explicit process and the slow learning process to an implicit process, both with anatomically distinct associated brain areas[6,19,43]. While it is plausible to assume that an interplay of explicit and implicit learning mechanisms is independent of the dimensionality of the task, to our knowledge, the two-state model to date has not been demonstrated for adaptation paradigms involving 3D movements. The fact that, in a 3D setting, we found a two-state model to better fit the adaptation data than a single-state model without introducing paradigms that explicitly enforce two-states dynamics (like re-learning after washout trials,[37]), might indicate that the explicit learning component in our three dimensional-paradigm is more pronounced than in classic two-dimensional paradigms. This finding is also in line with a recent study showing that VR visuomotor rotation tasks have a stronger explicit component than non-VR tasks [8]. The results therefore further strengthen the association between fast and slow states with explicit and implicit mechanisms.

438 **Differences in estimated planning and measurement noises explain differences in learning rate**

439 Motor learning studies are informative for motor rehabilitation protocols for patients that need to relearn
440 or re-adapt existing motor schemes. These patients will necessarily perform their movements in a three-
441 dimensional environment[55-58]. In our experiment, subjects adapted best to perturbations in the coronal
442 plane, followed by sagittal and horizontal planes. Differences in the slow and fast learning rates between
443 the planes were associated with different adaptation levels that were achieved during perturbation trials
444 in our data. In the following, we will discuss how these anisotropies in adaptation are related to
445 anisotropies in estimated planning and measurement noise, and, hence, could be explained by effects of
446 gravity and depth perception, respectively.

447 The slowest adaptation in the horizontal plane is likely explained by an anisotropy in planning variability
448 resulting from gravity. Previous studies suggested that movements along the vertical dimension have a
449 higher level of planning variability associated with gravity compensation[33-36]. Optimally combining
450 uncertainties on planning together with uncertainties on the estimation of the effector position here then
451 means that sensory measurements should be given more credibility in vertical movements due to the
452 higher planning variability. This leads to the prediction that adaptive corrections in the horizontal plane
453 arise slower than corrections in the sagittal or coronal plane, since the latter contain the vertical dimension
454 with the higher measurement credibility. Consistent with this view, perturbations in the horizontal plane
455 in our data were associated with the slowest adaptation rates and planning variability was lowest in the
456 horizontal plane, while measurement variability in the horizontal plane was close to the coronal plane
457 (Fig5B).

458 The fastest adaptation rates in the coronal plane are likely explained by anisotropies in measurement
459 variability resulting from visual depth perception. The coronal plane in our data yielded the lowest
460 estimate of measurement noise and, correspondingly, subjects adapted faster to perturbations in this
461 plane (Fig. 5B). In our model, measurement noise represents the reliability of the subject measurements

of the cursor position, i.e., the visual feedback about the movement, since the cursor determines the motor goal. Subjects have to rely mainly on stereoscopy to estimate visual depth and visual localization is less precise in depth[14,32]. Since depth is not perturbed by a rotation in the coronal plane, measurement error is least in the coronal condition and adaptation should be fastest. On the other hand, visuomotor rotations in the sagittal and horizontal planes both resulted in perturbations of cursor depth compared to baseline. Corresponding to this, other studies also showed that subjects tend to adapt worse when the measurement variability (or sensory nose) was increased by practically blurring the error feedback[11,13,21].

Optimal integration theory in 3D

Previous studies[11,20,21,26,31], provided evidence for the OI theory during visuomotor adaptation. They showed that the planning component of motor noise correlates positively with the adaptation rate while the measurement component correlates negatively, at least when the respective noise component is experimentally kept constant. While the adaptation data from our 3D movements yielded similar trends, the correlation of slow learning rate with either measurement- or planning-variability was weak. For an optimal learner, the Kalman gain is indicative of the level of correction attributed to the experienced error [11,21,31]. This Kalman gain positively correlates with the slow learning rate in our data, supporting the idea that subjects optimally integrate the information relative to the internal state with experienced (visual) feedback. However, and as opposed to what was suggested in previous studies[20,31], we observe the correlation between Kalman gain and learning rates only across different perturbation conditions. Within each perturbation condition and across individual subjects, there was no relationship between Kalman gain and learning rates.

Conclusion

Here, we develop a novel HMC fitting procedure to correlate learning rates and noise levels estimated with the same model and without the need to experimentally manipulating these error variabilities. By applying this model to a sensorimotor adaptation task with naturalistic movements in 3D, we found that the rate of adaptation with which subjects counteract applied visuomotor rotations, depends on the statistics of the combined errors with which subjects plan and perceive movements. These errors differ between different axes relative to the body, likely due to gravity compensation during planning- and perception-induced anisotropies. This insight could be used in motor rehabilitation strategies by specifically targeting the orientation of the body relative to the target while performing reaching movements to accelerate learning.

Acknowledgments:

German Research Foundation (DFG, Germany, grant number FOR-1847-GA1475-B2), received by AG.

German Research Foundation (DFG, Germany, grant number SFB-889), received by AG.

Federal Ministry for Education and Research (BMBF, Germany, grant number 01GQ1005C), received by AG.

Federal Ministry for Education and Research (BMBF, Germany, grant numbers 01GQ0814), received by AG.

Figure Captions

Figure 1. Virtual reality setup with robotic manipulator for testing visuomotor rotation adaptation of naturalistic reaches in different planes relative to the body. A) Schematic representation of the 3D virtual reality setup. The two angled mirrors reflect the image from the two screens to the two eyes independently providing stereoscopic vision. The mirrors are semitransparent to allow calibration of the virtual reality space with the real hand position. During the experiments, transparency is blocked, so that the subjects do not see the robotic manipulator and their hand, but only the co-registered cursor. B) Spatial and temporal structure of the 3D reaching task. Following a brief holding period in the center of a virtual cube, the subject has to reach to one of eight targets arranged at the corners of the cube. For ergonomic reasons, the cube is orientated relatively to the slightly downward-tilted workspace and head

of the subject, not according to true horizontal. C) Example baseline trajectories of a single subject. D) Time course of the perturbation experiment. Following a baseline phase of 120 trials, a 30-degree visuomotor perturbation is applied for 240 trials, followed by a washout of 240 additional trials. In three separate blocks for each subject, the perturbation is applied in the sagittal, coronal and horizontal plane, respectively. Note that all reaches are conducted along a diagonal direction, i.e., each movement is affected by either of the perturbations.

Figure 2. Quantification of adaptation level. A) Example trajectories from one subject. Each row/color of the graph correspond to different planes of rotation. The trajectories are projected onto the plane which is perpendicular to the axis of visuomotor rotation for representation and analysis of the data in each of the three conditions (coronal, horizontal, sagittal). The trajectories show similarities across the three planes. B) Within each plane, the trajectories are quantified along the two task relevant dimensions, the center-to-target dimension and the dimension orthogonal to it in the same plane, which we refer to as "perturbed" dimension. To compare data across tasks, each trajectory was resampled and rotated into this reference frame. At half of the normalized trajectory, the adaptation angle was calculated and used for model fitting. C) Examples of transformed trajectories from one subject during baseline, early adaptation and late adaptation.

Figure 3. Visuomotor adaptation in three different planes relative to the body. A) Example of the time course of adaptation for one example subject with perturbation applied to the sagittal plane. B) Final adaptation level for each plane. C) Examples of data fitted with a one-state and a two-states learning model, respectively. Black points represent the subject's data (average movement angle at half-trajectory within 8-trial epochs; 105 epochs represent 360 concatenated baseline trials, 240 adaptation trials and 240 washout trials). The black dashed line shows the perturbation profile.

Figure 4. Comparison of hierarchical and non-hierarchical fitting algorithms. Parameter estimation quality for our HMC hierarchical fitting procedure (orange), an HMC non-hierarchical fitting procedure (red), and a EM fitting procedure (EM, blue; [45]). For each random parameter setting of the surrogate two-state datasets, we represent the error in parameter reconstruction (estimated minus set parameter) for each fitting procedure, including the distributions of the errors, average and 95% confidence intervals of the squared error; and the correlation (average and 95% CI) between the estimated and set parameters.

Figure 5. Comparison of two-state model parameters between the perturbation planes. A) Estimated fast and slow retention ($a_{f,jp}$, $a_{s,jp}$), learning rates ($b_{f,jp}$, $b_{s,jp}$), and measurement and planning variabilities ($\sigma_{x,jp}$, $\sigma_{u,jp}$) for the sagittal (S), horizontal (H) and coronal (C) perturbation planes. Each color represents one subjects. The error bars correspond to 95 % credible intervals (i.e., contain 95% of the posterior probability) for the population average hyper-parameters $\mu_{a_s,p}$, $\mu_{a_f,p}$, $\mu_{b_s,p}$, $\mu_{b_f,p}$, $\mu_{\sigma_x,p}$, $\mu_{\sigma_u,p}$. The steady-state Kalman gain K was computed from the posteriors of the measurement noise and planning noise. B) Relationships between the slow learning rate bs and variability related parameters: measurement noise variability, planning noise variability, and steady-state Kalman gain. Points correspond to individual subjects' parameters in the different rotation planes (coded by color). The

correspondingly colored lines represent linear fits based on the across-subject parameter values within each plane. The grey line corresponds to a linear fit based on subjects' parameters values across all planes. The horizontal and vertical error bars correspond to 95 % credible intervals for the corresponding parameter population average hyper-parameters.

Figure 6. Recovery of simulated correlations between model parameters by the HMC algorithm. A) A correlation among bs and K was introduced in the surrogate data by setting bs equal to 0.3 times K (dashed black line). B) When no correlation between bs and K was added to the surrogate data, the HMC fitting algorithm did not introduce artificial spurious correlations.

Supplementary Figures

Supplementary Figure 1. Results of fitting the hierarchical HMC model to the adaptation profile of each subject (n = 24) with perturbation applied to the Sagittal plane.

Supplementary Figure 2. Results of fitting the hierarchical HMC model to the adaptation profile of each subject (n = 24) with perturbation applied to the horizontal plane.

Supplementary Figure 3. Results of fitting the hierarchical HMC model to the adaptation profile of each subject (n = 24) with perturbation applied to the coronal plane.

Supplementary Figure 4. Similar to figure 5 but with the model applied on a trial base rather than epoch base (no average over 8 consecutive trials). A) Example of single subject time course of adaptation with 2-state model fits. B) Extrapolated model parameter for all the subjects (different colors). C) Slow learning rate (bs) vs Kalman gain relationship.

References

1 Krakauer, J. W. Motor learning and consolidation: the case of visuomotor rotation. *Adv Exp Med Biol* **629**, 405-421, doi:10.1007/978-0-387-77064-2_21 (2009).

2 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, doi:10.1152/jn.01055.2015 (2016).

3 Mazzoni, P. & Krakauer, J. W. An implicit plan overrides an explicit strategy during visuomotor adaptation. *J Neurosci* **26**, 3642-3645, doi:10.1523/JNEUROSCI.5317-05.2006 (2006).

4 McDougle, S. D., Bond, K. M. & Taylor, J. A. Explicit and Implicit Processes Constitute the Fast and Slow Processes of Sensorimotor Learning. *J Neurosci* **35**, 9568-9579, doi:10.1523/JNEUROSCI.5061-14.2015 (2015).

5 Morehead, J. R., Qasim, S. E., Crossley, M. J. & Ivry, R. Savings upon Re-Aiming in Visuomotor Adaptation. *J Neurosci* **35**, 14386-14396, doi:10.1523/JNEUROSCI.1046-15.2015 (2015).

6 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, doi:10.1523/JNEUROSCI.3619-13.2014 (2014).

7 Wolpert, D. M., Diedrichsen, J. & Flanagan, J. R. Principles of sensorimotor learning. *Nat Rev Neurosci* **12**, 739-751, doi:10.1038/nrn3112 (2011).

8 Anglin, J. M., Sugiyama, T. & Liew, S. L. Visuomotor adaptation in head-mounted virtual reality versus conventional training. *Sci Rep* **7**, 45469, doi:10.1038/srep45469 (2017).

9 Choi, W., Lee, J., Yanagihara, N., Li, L. & Kim, J. Development of a quantitative evaluation system for visuo-motor control in three-dimensional virtual reality space. *Sci Rep* **8**, 13439, doi:10.1038/s41598-018-31758-y (2018).

10 Lefrancois, C. & Messier, J. Adaptation and spatial generalization to a triaxial visuomotor perturbation in a virtual reality environment. *Exp Brain Res* **237**, 793-803, doi:10.1007/s00221-018-05462-2 (2019).

11 Burge, J., Ernst, M. O. & Banks, M. S. The statistical determinants of adaptation rate in human reaching. *J Vis* **8**, 20 21-19, doi:10.1167/8.4.20 (2008).

12 Ernst, M. O. & Banks, M. S. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* **415**, 429-433, doi:10.1038/415429a (2002).

13 Kording, K. P. & Wolpert, D. M. Bayesian integration in sensorimotor learning. *Nature* **427**, 244-247, doi:10.1038/nature02169 (2004).

14 van Beers, R. J., Wolpert, D. M. & Haggard, P. When feeling is more important than seeing in sensorimotor adaptation. *Curr Biol* **12**, 834-837, doi:10.1016/s0960-9822(02)00836-9 (2002).

15 Wolpert, D. M. & Ghahramani, Z. Computational principles of movement neuroscience. *Nat Neurosci* **3 Suppl**, 1212-1217, doi:10.1038/81497 (2000).

16 Friston, K. What is optimal about motor control? *Neuron* **72**, 488-498, doi:10.1016/j.neuron.2011.10.018 (2011).

17 Scott, S. H. Optimal feedback control and the neural basis of volitional motor control. *Nat Rev Neurosci* **5**, 532-546, doi:10.1038/nrn1427 (2004).

18 Shadmehr, R. & Krakauer, J. W. A computational neuroanatomy for motor control. *Exp Brain Res* **185**, 359-381, doi:10.1007/s00221-008-1280-5 (2008).

19 Shadmehr, R., Smith, M. A. & Krakauer, J. W. Error correction, sensory prediction, and adaptation in motor control. *Annu Rev Neurosci* **33**, 89-108, doi:10.1146/annurev-neuro-060909-153135 (2010).

20 van der Vliet, R. *et al.* Individual Differences in Motor Noise and Adaptation Rate Are Optimally Related. *eNeuro* **5**, doi:10.1523/ENEURO.0170-18.2018 (2018).

21 Wei, K. & Kording, K. Uncertainty of feedback and state estimation determines the speed of motor adaptation. *Front Comput Neurosci* **4**, 11, doi:10.3389/fncom.2010.00011 (2010).

22 Cheng, S. & Sabes, P. N. Modeling sensorimotor learning with linear dynamical systems. *Neural Comput* **18**, 760-793, doi:10.1162/089976606775774651 (2006).

23 Cheng, S. & Sabes, P. N. Calibration of visually guided reaching is driven by error-corrective learning and internal dynamics. *J Neurophysiol* **97**, 3057-3069, doi:10.1152/jn.00897.2006 (2007).

24 Korenberg, A. & Ghahramani, Z. A Bayesian view of motor adaptation. *Cahiers de Psychologie Cognitive* **21** (2002).

25 Morel, P. D. & Baraduc, P. in *Cinquième conférence plénière française de Neurosciences Computationnelles, "Neurocomp'10"* (Lyon, France, 2010).

26 van Beers, R. J. Motor learning is optimally tuned to the properties of motor noise. *Neuron* **63**, 406-417, doi:10.1016/j.neuron.2009.06.025 (2009).

27 Diedrichsen, J., Shadmehr, R. & Ivry, R. B. The coordination of movement: optimal feedback control and beyond. *Trends Cogn Sci* **14**, 31-39, doi:10.1016/j.tics.2009.11.004 (2010).
28 Kalman, R. E. A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering* **82**, 35-45, doi:10.1115/1.3662552 (1960).
29 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).
30 Ghahramani, Z., Wolptrt, D. M. & Jordan, M. I. in *Advances in Psychology* Vol. 119 (eds Pietro Morasso & Vittorio Sanguineti) 117-147 (North-Holland, 1997).
31 He, K. *et al.* The Statistical Determinants of the Speed of Motor Learning. *PLoS Comput Biol* **12**, e1005023, doi:10.1371/journal.pcbi.1005023 (2016).
32 Gepshtein, S. & Banks, M. S. Viewing geometry determines how vision and haptics combine in size perception. *Curr Biol* **13**, 483-488, doi:10.1016/s0960-9822(03)00133-7 (2003).
33 Apker, G. A., Karimi, C. P. & Buneo, C. A. Contributions of vision and proprioception to arm movement planning in the vertical plane. *Neurosci Lett* **503**, 186-190, doi:10.1016/j.neulet.2011.08.032 (2011).
34 Le Seac'h, A. B. & McIntyre, J. Multimodal reference frame for the planning of vertical arms movements. *Neuroscience letters* **423**, 211-215 (2007).
35 Papaxanthis, C., Pozzo, T. & McIntyre, J. Kinematic and dynamic processes for the control of pointing movements in humans revealed by short-term exposure to microgravity. *Neuroscience* **135**, 371-383 (2005).
36 Sciutti, A. *et al.* Visual gravity influences arm movement planning. *Journal of neurophysiology* **107**, 3433-3445 (2012).
37 Smith, M. A., Ghazizadeh, A. & Shadmehr, R. Interacting adaptive processes with different timescales underlie short-term motor learning. *PLoS Biol* **4**, e179, doi:10.1371/journal.pbio.0040179 (2006).
38 Huberdeau, D. M., Krakauer, J. W. & Haith, A. M. Dual-process decomposition in human sensorimotor adaptation. *Curr Opin Neurobiol* **33**, 71-77, doi:10.1016/j.conb.2015.03.003 (2015).
39 Avraham, G., Morehead, J. R., Kim, H. E. & Ivry, R. B. Re-exposure to a sensorimotor perturbation produces opposite effects on explicit and implicit learning processes. *bioRxiv* (2020).
40 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, doi:10.1016/j.neuron.2011.04.012 (2011).
41 Krakauer, J. W., Ghez, C. & Ghilardi, M. F. Adaptation to visuomotor transformations: consolidation, interference, and forgetting. *J Neurosci* **25**, 473-478, doi:10.1523/JNEUROSCI.4218-04.2005 (2005).
42 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**, doi:10.1371/journal.pcbi.1000893 (2010).
43 Della-Maggiore, V., Malfait, N., Ostry, D. J. & Paus, T. Stimulation of the posterior parietal cortex interferes with arm trajectory adjustments during the learning of new dynamics. *J Neurosci* **24**, 9971-9976, doi:10.1523/JNEUROSCI.2833-04.2004 (2004).
44 Inoue, M. & Kitazawa, S. Motor Error in Parietal Area 5 and Target Error in Area 7 Drive Distinctive Adaptation in Reaching. *Curr Biol* **28**, 2250-2262 e2253, doi:10.1016/j.cub.2018.05.056 (2018).
45 Albert, S. T. & Shadmehr, R. Estimating properties of the fast and slow adaptive processes during sensorimotor adaptation. *J Neurophysiol* **119**, 1367-1393, doi:10.1152/jn.00197.2017 (2018).

46 Browne, W. J. & Draper, D. A comparison of Bayesian and likelihood-based methods for fitting multilevel models. *Bayesian analysis* **1**, 473-514 (2006).

47 Morel, P. *et al.* Long-term decoding of movement force and direction with a wireless myoelectric implant. *J Neural Eng* **13**, 016002, doi:10.1088/1741-2560/13/1/016002 (2016).

48 Morel, P., Ulbrich, P. & Gail, A. What makes a reach movement effortful? Physical effort discounting supports common minimization principles in decision making and motor control. *PLoS Biol* **15**, e2001323, doi:10.1371/journal.pbio.2001323 (2017).

49 Rosenberg, L. *The effect of interocular distance upon operator performance using stereoscopic displays to perform virtual depth tasks*. (1993).

50 Morel, P. Gramm: grammar of graphics plotting in Matlab. *The Journal of Open Source Software* **3**, doi:10.21105/joss.00568 (2018).

51 Gelman, A., Lee, D. & Guo, J. Stan: A probabilistic programming language for Bayesian inference and optimization. *Journal of Educational and Behavioral Statistics* **40**, 530-543 (2015).

52 Hoffman, M. D. & Gelman, A. The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.* **15**, 1593-1623 (2014).

53 Gelman, A. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian analysis* **1**, 515-534 (2006).

54 Vehtari, A., Gelman, A. & Gabry, J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and computing* **27**, 1413-1432 (2017).

55 Holden, M. K. Virtual environments for motor rehabilitation: review. *Cyberpsychol Behav* **8**, 187-211; discussion 212-189, doi:10.1089/cpb.2005.8.187 (2005).

56 Turolla, A. *et al.* Virtual reality for the rehabilitation of the upper limb motor function after stroke: a prospective controlled trial. *J Neuroeng Rehabil* **10**, 85, doi:10.1186/1743-0003-10-85 (2013).

57 Deutsch, J. E., Merians, A. S., Adamovich, S., Poizner, H. & Burdea, G. C. Development and application of virtual reality technology to improve hand use and gait of individuals post-stroke. *Restor Neurol Neurosci* **22**, 371-386 (2004).

58 Pisella, L., Rode, G., Farne, A., Tilikete, C. & Rossetti, Y. Prism adaptation in the rehabilitation of patients with visuo-spatial cognitive disorders. *Curr Opin Neurol* **19**, 534-542, doi:10.1097/WCO.0b013e328010924b (2006).

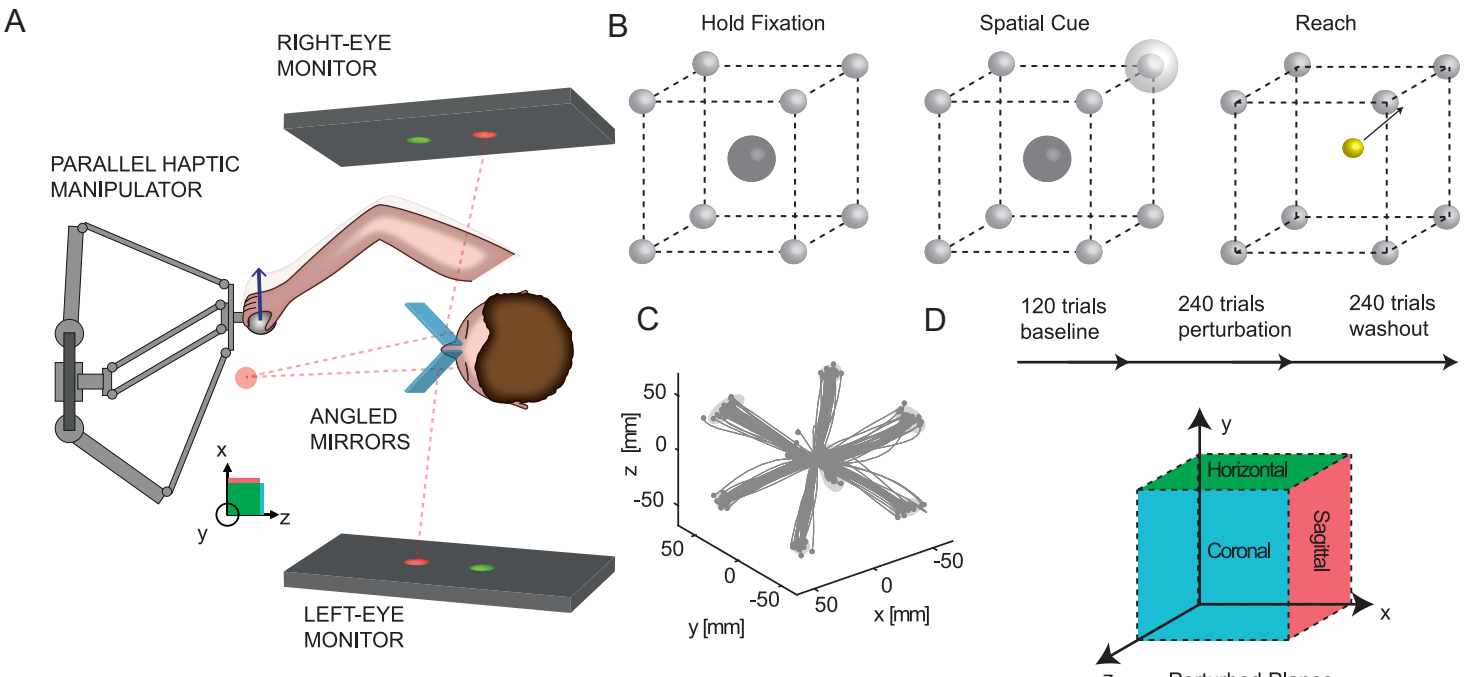

A
RIGHT-EYE MONITOR
PARALLEL HAPTIC MANIPULATOR
ANGLED MIRRORS
LEFT-EYE MONITOR
x
y
z
B
Hold Fixation
Spatial Cue
Reach
C
50
0
-50
z [mm]
50
0
-50
y [mm]
50
0
-50
x [mm]
D
120 trials baseline
240 trials perturbation
240 trials washout
y
Horizontal
Coronal
Sagittal
x
z
Perturbed Planes

Figure 1

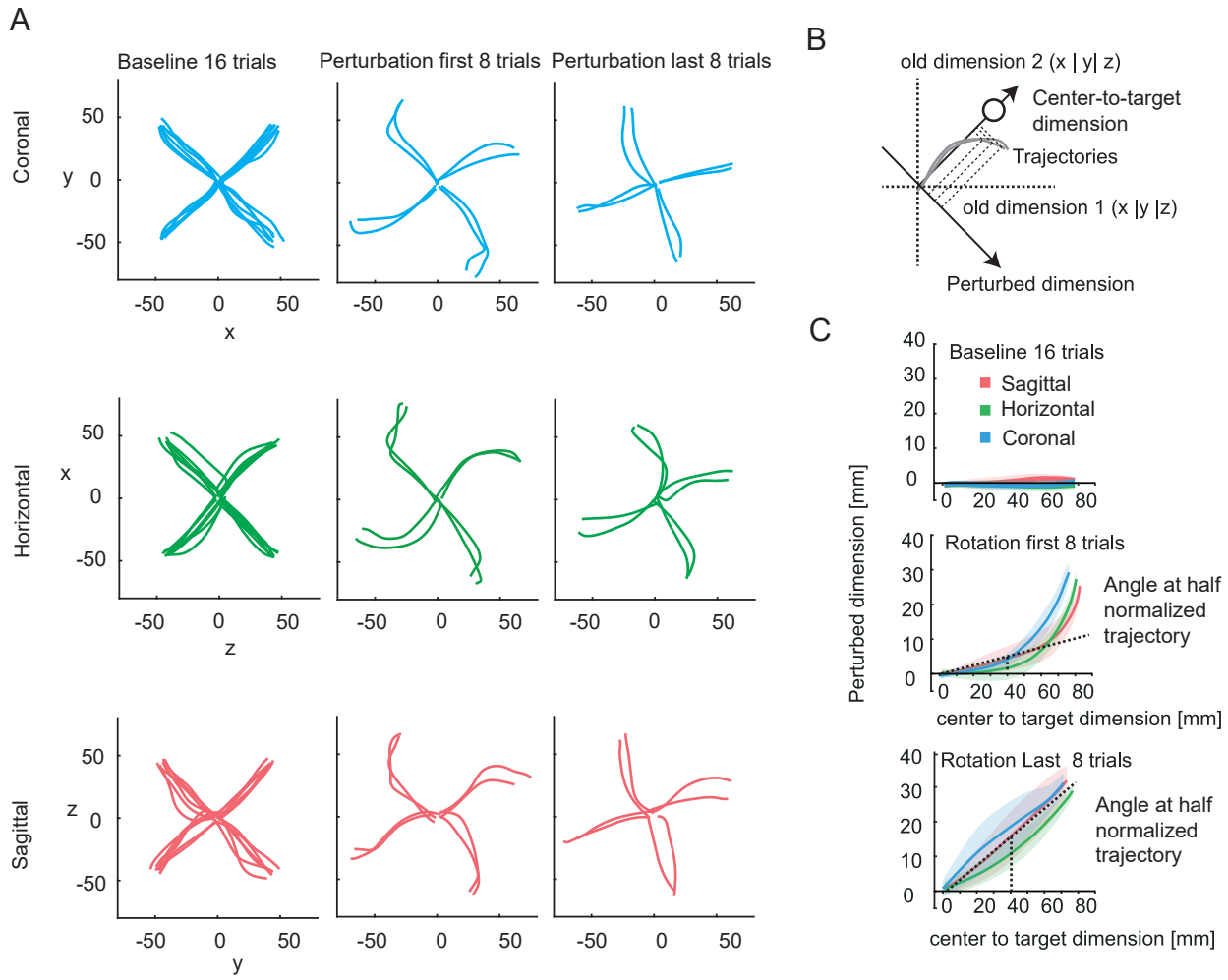

A
Baseline 16 trials
Perturbation first 8 trials
Perturbation last 8 trials
Coronal
Horizontal
Sagittal
x
y
z
B
old dimension 2 (x | y| z)
Center-to-target dimension
Trajectories
old dimension 1 (x |y |z)
Perturbed dimension
C
Baseline 16 trials
Sagittal
Horizontal
Coronal
Rotation first 8 trials
Rotation Last 8 trials
Angle at half normalized trajectory
Perturbed dimension [mm]
center to target dimension [mm]

Figure 2

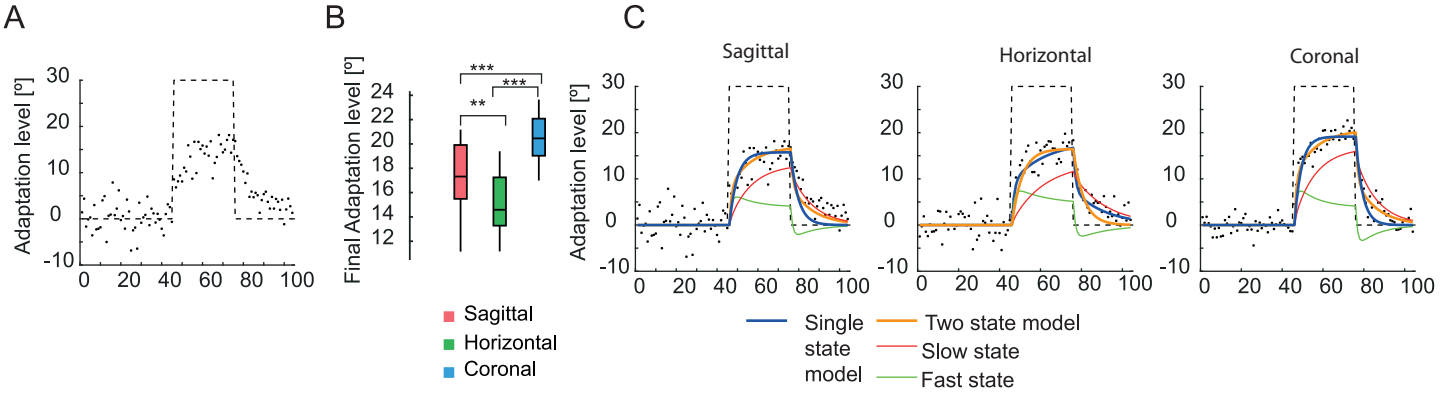

A
B
C
Adaptation level [º]
Final Adaptation level [º]
Sagittal
Horizontal
Coronal
Single state model
Two state model
Slow state
Fast state

Figure 3

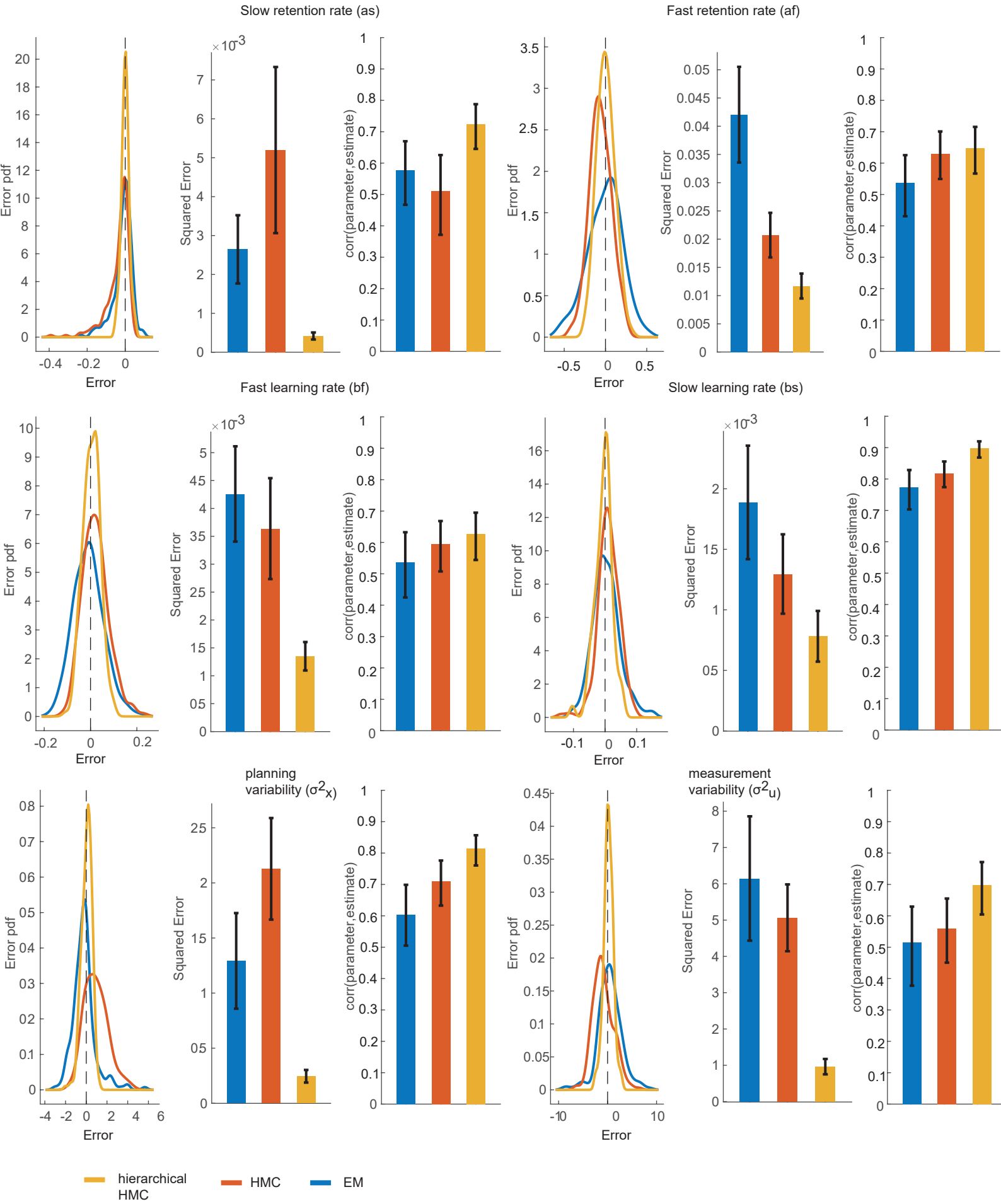

Slow retention rate (as)
Fast retention rate (af)
Fast learning rate (bf)
Slow learning rate (bs)
planning variability (σ^2_x)
measurement variability (σ^2_u)
Error pdf
Squared Error
corr(parameter,estimate)
Error
hierarchical HMC
HMC
EM

Figure 4

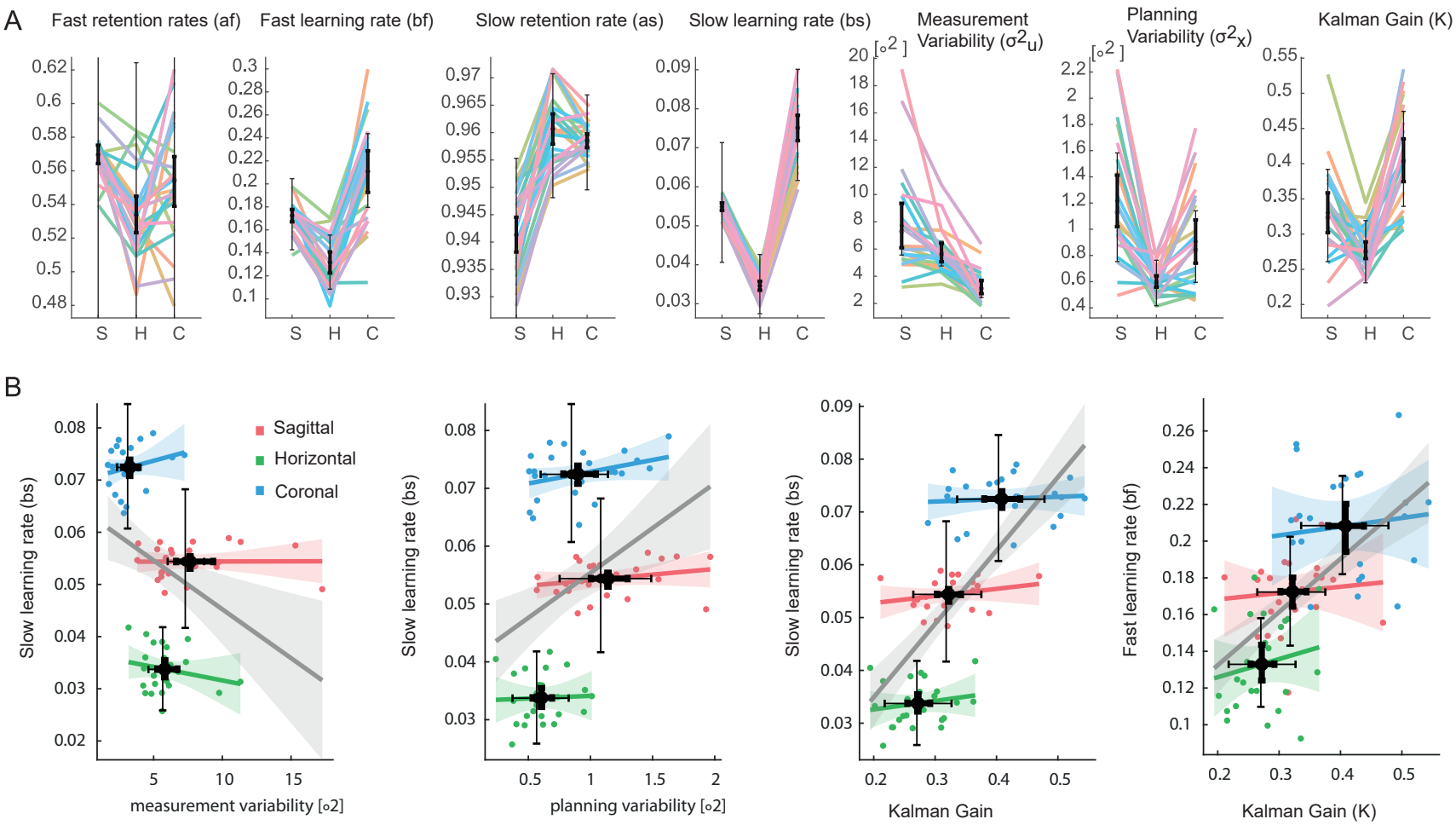

A
Fast retention rates (af)
Fast learning rate (bf)
Slow retention rate (as)
Slow learning rate (bs)
Measurement Variability (σ^2_u)
Planning Variability (σ^2_x)
Kalman Gain (K)
S H C
B
Sagittal
Horizontal
Coronal
Slow learning rate (bs)
measurement variability [∘2]
planning variability [∘2]
Kalman Gain
Fast learning rate (bf)
Kalman Gain (K)

Figure 5

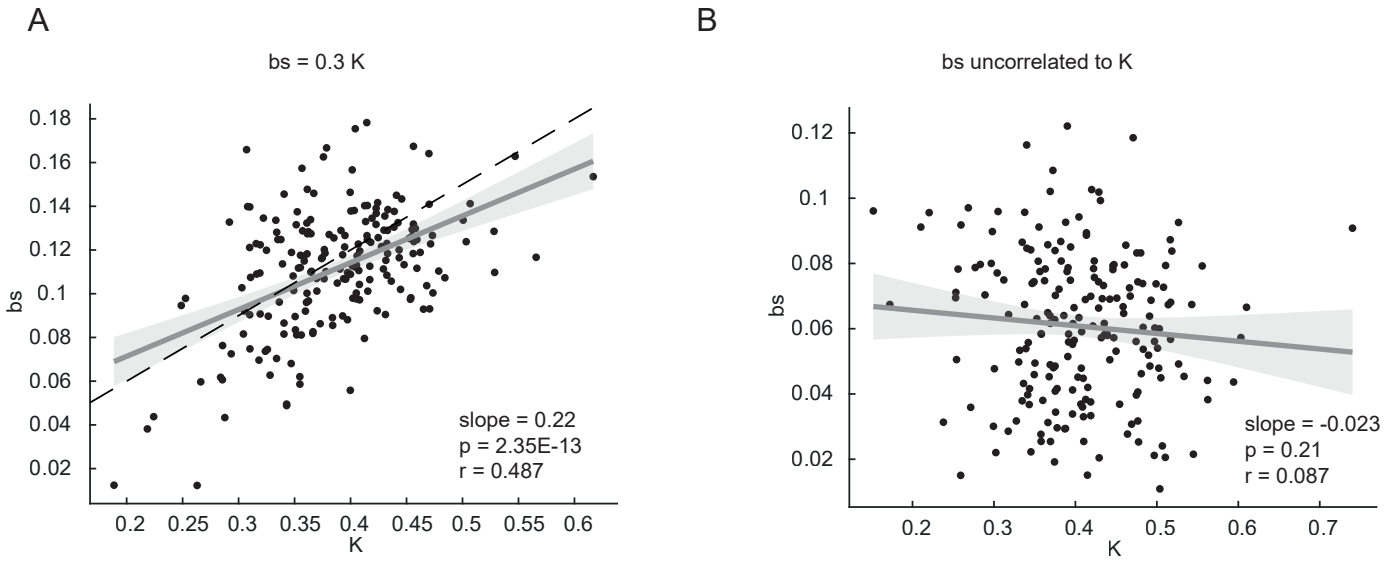
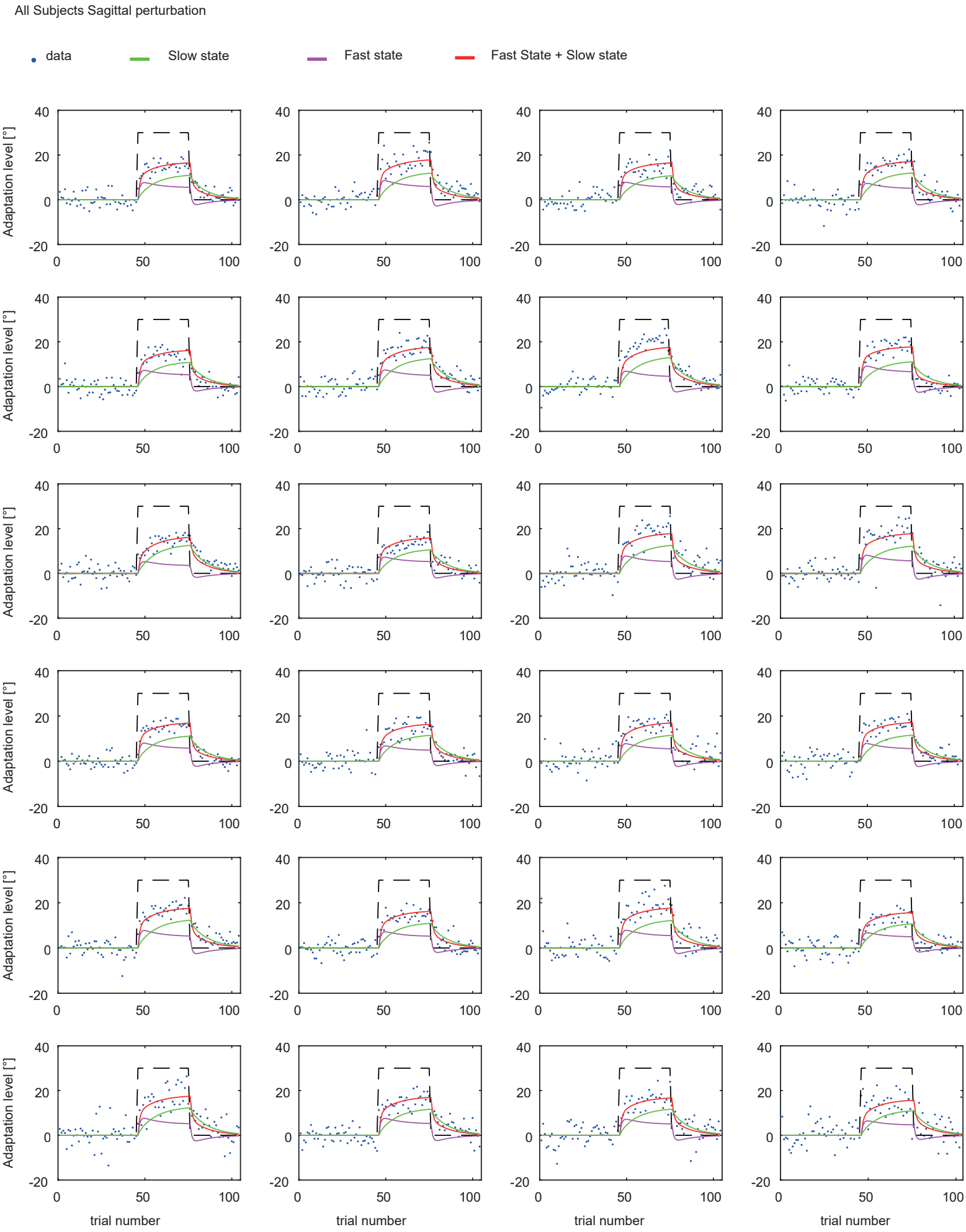
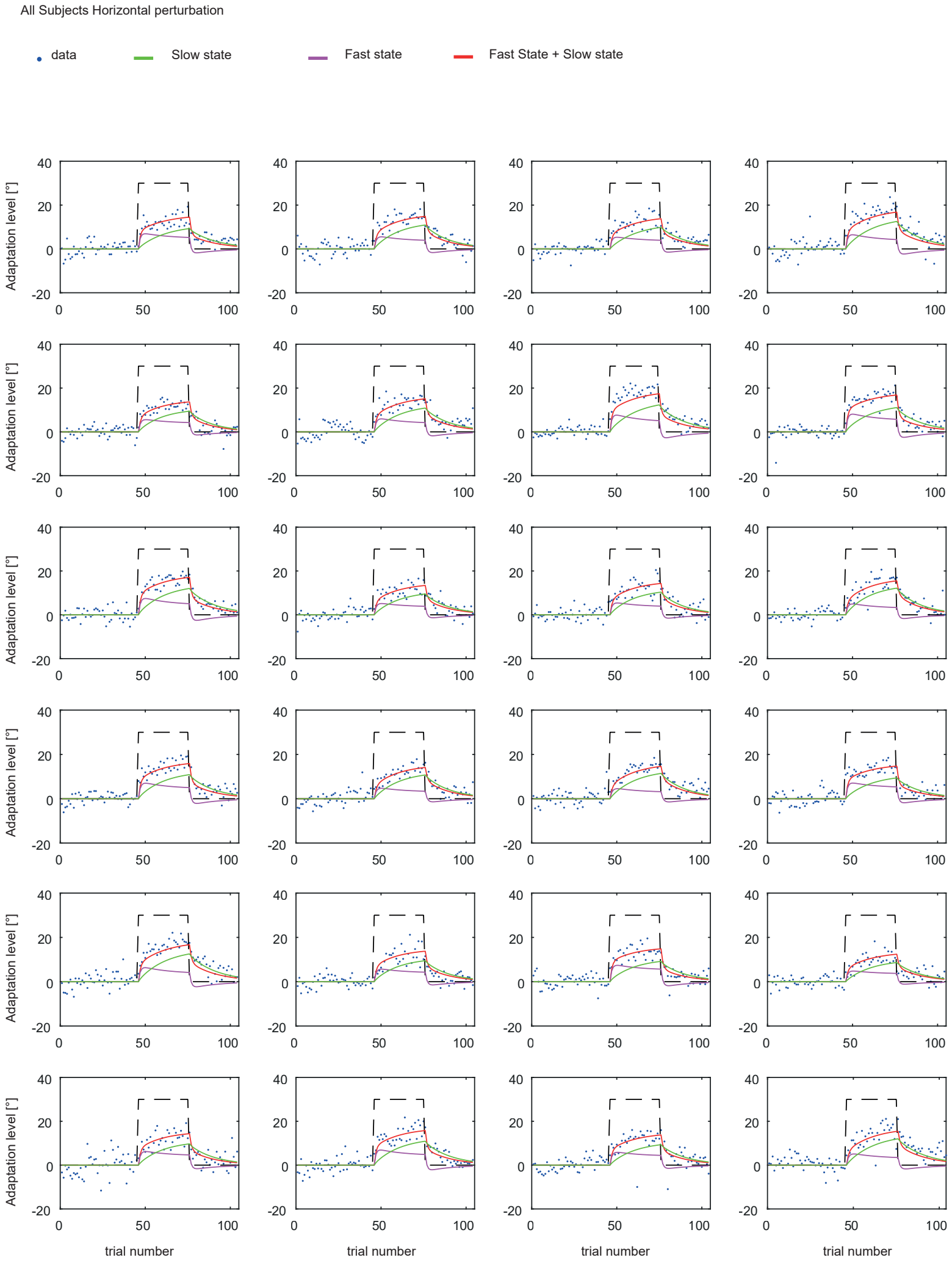

A
bs = 0.3 K
bs
K
slope = 0.22
p = 2.35E-13
r = 0.487
B
bs uncorrelated to K
bs
K
slope = -0.023
p = 0.21
r = 0.087

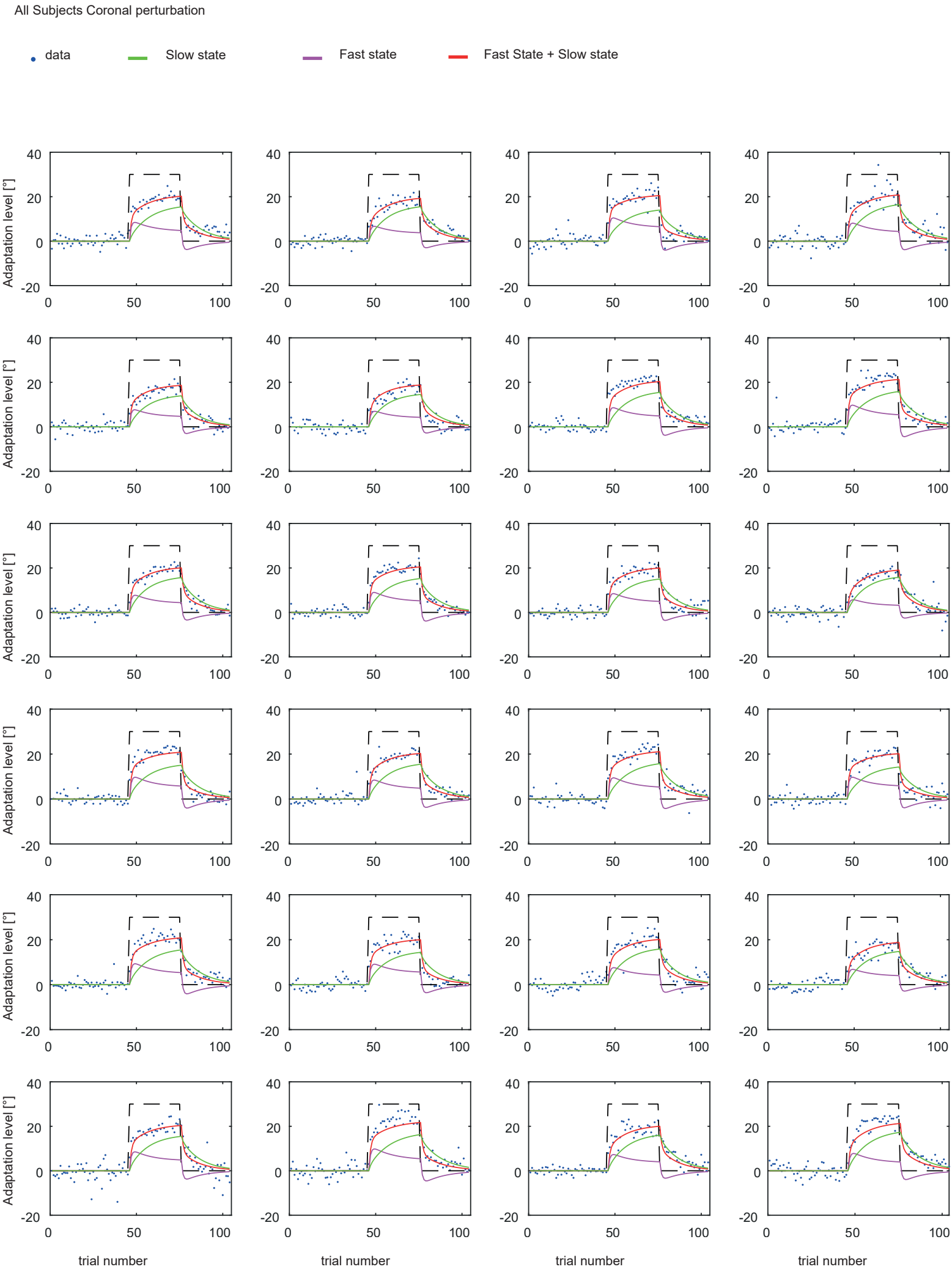
Figure 6



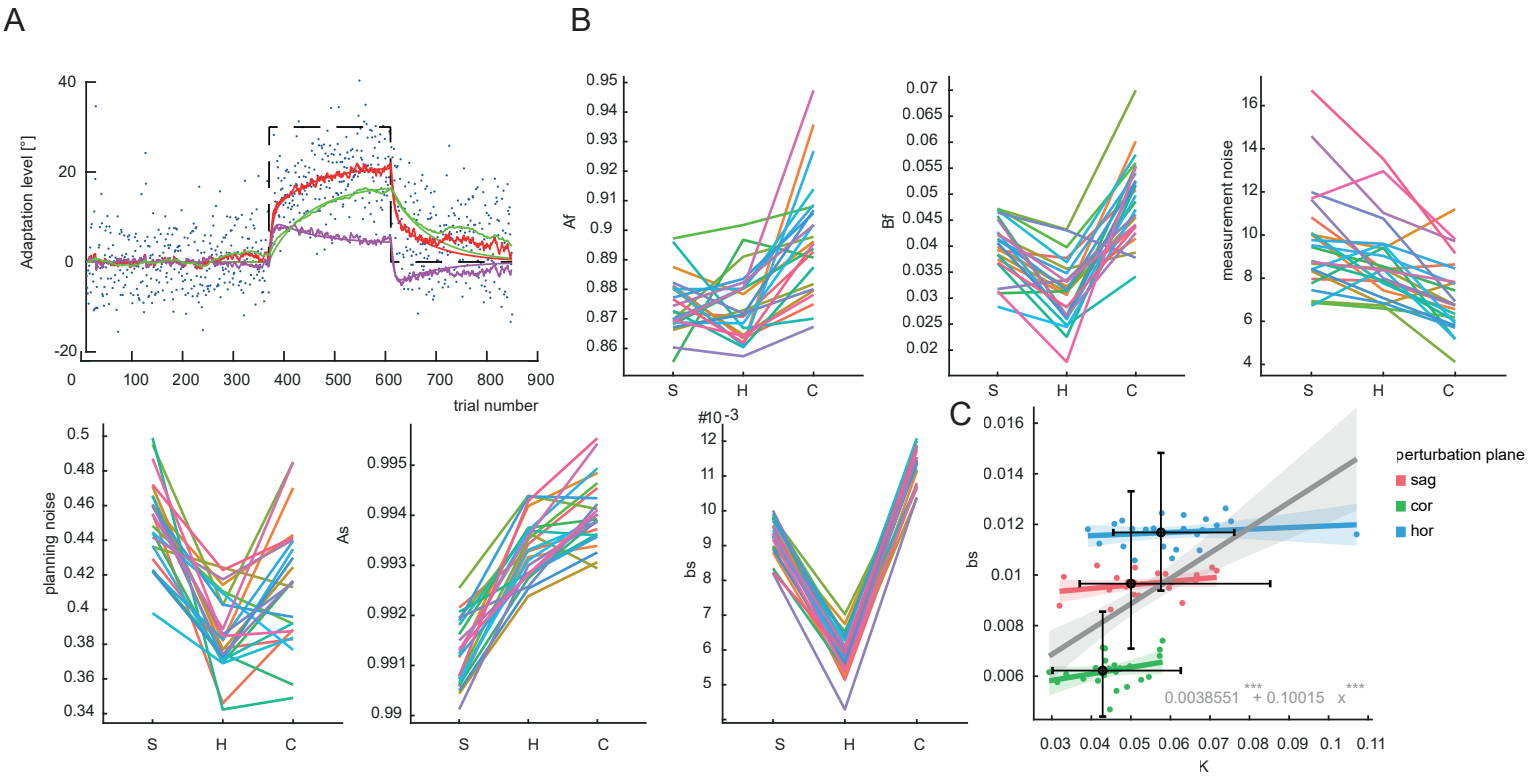
Supplementary Figure 1



Supplementary Figure 2



Supplementary Figure 3



Supplementary Figure 4