A sub-movement time scale of human motor adaptation

Abbreviated Title: Adaptive Control of Reaching

F. Crevecoeur[1,2], J.-L. Thonnard[1,2,3], P. Lefèvre[1,2]

[1]*Institute of Communication and Information Technologies, Electronics and Applied Mathematics (ICTM), UCLouvain, Belgium*

[2]*Institute of Neuroscience, UCLouvain, Belgium*

[3]*Cilniques Universitaires Saint-Luc, UCLouvain, Belgium*

Corresponding Author:

F. Crevecoeur

4 avenue Georges Lemaitre

1348 Louvain-la-Neuve, Belgium

email: frederic.crevecoeur@uclouvain.be

Abstract

Humans and animals adapt motor commands to predictable disturbances within tens of trials in laboratory conditions. However in natural conditions, adaptation likely occurs continuously and it remains unknown how the brain handles novel disturbances across time. Because motor commands and sensory feedback together carry continuous information about limb dynamics, we hypothesized that the nervous system could adapt to unexpected disturbances online. We tested this hypothesis in two reaching experiments during which opposing velocity-dependent force fields were randomly applied. We found that within-movement feedback corrections gradually improved, despite the fact that the perturbations were unexpected. Moreover, when participants were instructed to stop at a via-point, the application of a force field prior to the via-point induced mirror-image after-effects after the via-point, indicative of within-trial adaptation to the unexpected dynamics. These findings highlight a fast time-scale of motor learning, which supplements feedback control and supports adaptation of an ongoing movement.

Introduction

Neural plasticity in the sensorimotor system enables adaptive internal representations of movement dynamics, which supports motor adaptation and acquisition of motor skills with practice [1]. In the context of reaching movements, studies have documented that healthy humans and animals can learn to anticipate the impact of a force field applied to the limb, and recover straight reach paths within tens to hundreds of trials [2-7]. Importantly, studies on motor learning have consistently highlighted the presence of an after-effect, which mirrors the initial movement deviation and indicates that adaptation was supported by a novel internal model of movement dynamics [8-10].

To date, motor learning and adaptation have been exclusively studied on a trial-by-trial basis, highlighting context-dependent learning rates and memory dynamics across movements [11-15]. Although this approach has revealed several fundamental properties of the motor system, it has left largely unexplored the problem of online control during early exposure to the force field. More precisely, it remains unknown how the nervous system controls reaching movements in the presence of unexpected dynamics, or errors in the internal models, as occurs when a force field is introduced experimentally. To address this question, we sought to investigate whether within-movement error corrections could be mediated by rapid adaptation, allowing partial compensation for model errors in real-time.

In fact, it is often assumed that unexpected disturbances during movements are automatically countered by the limb's mechanical impedance and by reflexes [3,16,17]. However, recent work suggests that muscles viscoelastic properties, as well as the gain of the spinal reflex (latency ~20ms for upper limb muscles) are low at spontaneous levels of muscle activation [18]. Furthermore, long-latency feedback (latency ≥50ms) is also based on internal models of the limb and environmental dynamics [19-21], even when correcting for very small disturbances [22]. Thus, the presence of model errors is equally challenging for rapid feedback control as it is for movement planning, and yet healthy humans seem to handle the presence of an unexpected force field relatively well.

In theory, an approximate internal model can be deduced during movement because motor commands and sensory feedback together carry information about the underlying dynamics. This problem in engineering were studied in the framework of *adaptive control*[23-25]. The idea of adaptive control is to adapt the model continuously based on sensory feedback about the unexpected movement error. During reaching experiments, the unexpected introduction of a force field produces a model error, which generates a hand path deviation that can be used to estimate the underlying dynamics in addition to the state of the limb. It is important to distinguish adaptive control from standard (extended) stochastic optimal control. Indeed, stochastic optimal control is based on the assumption that there is no fixed model error, and that the disturbances in the process follow a known distribution [26,27]. The goal of adaptive control is to correct for the possible presence of model errors continuously.

Considering the very rich repertoire of rapid sensorimotor loops [28], we explored the possibility that mechanical disturbances to the limb can update the internal models online following the principles of adaptive control. Online estimation of a task-related parameter was reported in the context of visuomotor rotations [29], where participants had to estimate the mapping between hand and cursor motion. Here we address this problem in the context of reaching in a force field to probe specifically the internal representations of limb dynamics. We found that movements displayed online adjustments consistent with real-time motor adaptation. We further illustrate in a simple model that these effects, as well as standard learning curves, can be simulated in the framework of adaptive control. We suggest that adaptive control likely constitutes an important component of motor learning, as it bridges online motor corrections with learning across few trials. We also highlight the associated computational challenges for the nervous system that have yet to be fully characterized.

Results

Participants grabbed the handle of a robotic device and performed forward reaching movements towards a visual target (Fig. 1a). In the first experiment, clockwise and counterclockwise orthogonal force fields (FF) were randomly applied as catch trials, such that neither the occurrence nor the direction of the force field could be anticipated. Reach paths were of course strongly impacted by the presence of the force field, however the online correction became smoother with practice (compare trials #1 and 20 in Fig. 1b). To quantify this, we extracted the maximum hand deviation along the direction of the force field, and the target overshoot near the end of the movement (Fig. 1b: Max. and Osh., respectively). We found that the maximum deviation did not evolve significantly across force field trials (Fig. 1c), whereas the overshoot displayed a significant exponential decay (Fig. 1d, $R^2 = 0.17$ and 0.56 for clockwise and counterclockwise force fields, respectively). We also observed that exposure to a single force field trials induced significant after-effect on the next trial, as previously reported in a random adaptation paradigm [30] (Supplementary figure S1). The absence of a clear change in the maximum deviation confirms that there was no anticipation, or default strategy that could counter the unexpected disturbance, yet the online corrections improved across trials.

(Figure 1 about here)

The reduction in target overshoot resulted from a smoother correction for the lateral displacement. Indeed, there was no significant change in the peak forward hand velocity across trials for both clockwise (rmANOVA, $F_{(23,299)} = 0.56$, $P = 0.9$) and counterclockwise force fields ($F_{(22,286)} = 0.87$, $P = 0.62$). Thus the perturbation applied to the limb remained statistically similar. However, the absolute peak hand force near the end of the movement decreased significantly. This is illustrated in Fig. 2a for average continuous traces, and in Fig. 2b showing the significant exponential decay of the terminal peak hand force (R^2 =0.17 and 0.11 for clockwise and counterclockwise perturbations). As

a consequence, the correlation between the measured lateral force and the force field perturbation, corresponding to the forward velocity multiplied by the scaling coefficient of the force field, significantly increased across trials (Fig. 2c). In the case of perfect compensation, the measured force should equal the action of the force field, and a correlation of 1 is expected. Thus, the increase in correlation between the measured force and the force field within trials indicated that the online correction became progressively tuned to the properties of the force field.

(Figure 2 about here)

These results are expected in the framework of adaptive control. However it remains possible that participants learned to produce smooth corrective movements without learning about the force field specifically. To address this, we sought a more direct link between the online correction and the force field by using a via-point located at two thirds of the reach path. In this experiment, participants were given scores for good trials, and bonuses when they stopped at the via-point. The bonus was awarded if the hand velocity inside the via-point dropped below 3cm/s. Force field trials were randomly interleaved, with and without via-points. In a random subset of trials, the force field was turned off at the via-point. We reasoned that if the feedback correction during force field trials reflected an online update of the internal models, then these corrections should evoke almost instantaneous after-effects for the remainder of the movement from the via point to the goal target.

Our results confirmed this prediction. First we reproduced the observations made in the first experiment for the trials without via-point (see supplemental figure S2). Second, we found a systematic after-effect between the via-point and the goal target that mirrored the impact of the force field experienced between the start target and the via-point (Fig. 3a for average hand paths and 3b). The stopping at the via-point was necessary for two reasons: first it allowed us to turn off the force field unnoticeably because the velocity was close to zero. Second, it also ensured that the

135 after-effects were not due to momentum following the online correction, because the velocity was
136 again very close to zero at the via-point. Thus the after-effects between the via-point and the goal
137 target were due to a re-acceleration in the direction opposite to the perturbation induced by the
138 force field. We extracted the minimum lateral hand velocity in the via-point and the lateral hand
139 velocity at the second peak (Fig. 3c: squares and open discs on exemplar traces). As expected, there
140 was no difference across clockwise, baseline, and counterclockwise trials for the minimum velocity
141 (Fig. 3d). In contrast, the lateral hand velocity measured at the second maximum displayed a very
142 strong modulation consistent with an after-effect (Fig. 3d, rmANOVA on average lateral hand speed
143 across force field: $F_{(2,14)} = 17.23$, $P < 0.001$).
144
145 (Figure 3 about here)
146
147 Finally we were interested to see how early the modulation of the lateral hand velocity
148 following the via-point occurred. In addition to the main effect of the force field, we found a
149 significant interaction between the force field and the block number ($F_{(10,70)} = 2.34$, $P < 0.05$).
150 Strikingly, the effect reported in Fig. 3 was already present in the first block (Fig. 4b, paired t-test on
151 data from block 1: $t_8 = 3.6$, $P = 0.007$, effect size: 1.2). This difference increased across blocks,
152 indicating an increase in the rate of online adjustments of control (data from block 6: $t_8 = 5.8$, $P <$
153 0.001, effect size: 1.95).
154
155 (Figure 4 about here)
156
157 The presence of a significant modulation in the first block indicated that online adaptation
158 occurred during the first few trials (there were 5 trials with via-point and switching-off of the force
159 field per direction and per block). This striking result warranted further investigation: can these
160 adjustments occur within the very first trial? Our dataset cannot provide a definitive answer because,

due to randomization, some via-point trials were preceded by force field trials without via-point for some participants. Nevertheless, we used a statistical model of the modulation of lateral velocity as a function of trial number to obtain the beginnings of an answer.

We first changed the sign of velocity during CW trials such that an after effect would correspond to a positive velocity for all trials. Then, we accounted for idiosyncrasy by including participants as a random factor in a mixed linear model. We found a significant positive correlation between the lateral hand velocity and the trial index. Most importantly, the regression included an offset that was significantly distinct from 0. The intercept value was $0.014ms^{-1}$, and the *P*-values for the slope and intercept were 0.014 and 0.006, respectively. The same regression performed on the baseline trials, where no force field was applied prior to the vial point, revealed an offset of -0.007, which was not significant ($P = 0.06$). Thus, we can reject the hypothesis that the lateral hand velocity was zero in the very first trial. In other words, our dataset supports the hypothesis that online adjustments occurred even during the very first trial. Lastly, we compared the lateral hand velocity from the very first successful trial in each direction and from each participant, and also found a significant modulation of hand velocity following the via-point based on a single trial in each direction. Indeed the lateral velocity following CCW perturbations was significantly higher than following CW perturbations (one-tail paired t-test, $t_{(8)} = 1.91$, $P = 0.044$).

These results indicate that the internal representation dynamics was updated during movements, and could modify the previously planned movement from the via-point to the goal target in a way that was consistent with a near-instantaneous after-effect. Furthermore, statistical modeling suggested that such adjustments occurred during the very first trial. It is possible that participant used three internal models, one for null-field trials and one per force-field direction, and switched between them dependent on the ongoing movement. Although plausible, this view does not easily explain how these models are learned in the first place nor how they are refined across trials to account for the larger amounts of online adaptation observed across the blocks (Fig. 4). A

more compelling view in our opinion is the hypothesis that the nervous system uses adaptive control, describing how sensory feedback and motor commands are used to track model parameters online.

To illustrate this, we now present the theoretical model and simulations based on adaptive control. First it is important to highlight that adaptive changes in the internal models differ from the sensorimotor loops that correct for perturbations altering the state of the limb (Fig. 5a: inner loop, solid). In the inner loop, sensory feedback is used to estimate the state of the system through an observer (e.g. Kalman filter), and the controller generates motor commands that correct for motor errors. This loop is responsible for adjusting motor commands following mechanical disturbances to the limb or following target jumps, two situations requiring an update of the ongoing command to comply with the goal of the task.

(Figure 5 about here)

In the framework of adaptive control, there is a second loop (Fig. 5a: outer loop, dashed), which estimate the model based on current sensory data and motor commands [23-25]. This estimation is indirect because there is no direct measurement of the model parameters, thus the motor command and sensory feedback must be used to deduce the structure of the underlying dynamics. The simulations performed in this study are based on iterative least square identification (RLS) [25], but other techniques such as expectation-maximization may be considered [31]. However it is important to note that the increase in after-effect observed across the blocks of Experiment 2 required the presence of a parameter to change the online learning rate, which is why we privileged the RLS formalism as it includes such parameter by design (see Methods, γ). The outer loop updates the model used in the estimator and in the controller of the inner loop (Fig. 5a, oblique arrows). The second experiment was designed to highlight the presence of this loop: at the via-point with near-zero velocity, there is no deviation or error in the state of the limb that requires correction by the

inner loop. Yet, participants did not simply reach straight ahead because of change in the internal models performed by the outer loop during the first portion of the reaching movement.

Using this framework, we reproduced standard learning curves in computer simulations, with exponential decay of the lateral hand displacement, mirror-image catch trials, and unlearning following the catch trials as in typical human learning experiments (Fig. 5b) [6]. This model also reproduced the reduction in peak end force occurring near the end of movement as observed in Experiment 1 by assuming that the online learning rate increased across force field trials (Fig. 5c, see also Methods), that is the online adjustments were smaller in the first trials than in the last trials. This idea is consistent with the presence of savings, characterizing a faster re-learning upon exposure to a novel environment experienced before [11,13]. Here we simply hypothesized a similar faster relearning when exposed to a previously experienced perturbation based on the continuous inflow of sensory information and an increased online learning rate (γ parameter). Finally, we also reproduced the presence of after-effects after switching off the force field at the via-point and extracting the lateral velocity exactly as in Experiment 2 (Fig. 3c and 5d). Again, the modulation of online adaptation documented across trials and blocks could be reproduced by increasing the online learning rate similar to the faster trial-by-trial re-adaptation observed in previous learning experiments [13].

Discussion

Controlling movements in the presence of model errors is a challenging task for robotics as well as for the nervous system. However healthy humans can handle model errors like those arising when a force field is introduced experimentally during reaching. To understand this, we explored the possibility that the nervous system reduces the impact of model errors online following the principles of adaptive control. We found that participants' feedback corrections to unexpected force fields became gradually better (Experiment 1), and evoked within-trial after-effects for an ongoing sequence of movements (Experiment 2). We showed that these behaviors, as well as standard learning curves, were captured in a simple adaptive control model in which motor commands, state estimates, and sensory feedback are used to update the internal models of dynamics during movement. The strengths of the theoretical model were that it reproduced some aspects of our experimental findings, along with the typical exponential decay observed in human motor adaptation, and based on online continuous processing of sensory feedback. Thus it constitutes a powerful candidate model for linking online control with motor learning.

We should emphasize that there is no disagreement between adaptive control and stochastic optimal control (LQG, [26]), but instead there is complementarity between these models. Stochastic optimal control is based on the assumption that the system dynamics are known exactly, and that the process disturbance is known in distribution [32]. The unexpected introduction of a force field introduces a fixed error in the model parameters, which is not a stochastic (random) disturbance. This violates the assumption of stochastic optimal control and warrants the use of control design that explicitly considers the possibility of an error in the model parameters. Adaptive control aims at solving this problem.

These results provide evidence that separate motor memories of opposite perturbations can be learned and retrieved within movement. Previous work showed that distinct planning conditions were critical to learn opposite perturbations, such as explicit cues [33], distinct representations of movements [34], or even distinct prior or follow-through movements associated with the same

movement [35,36]. In our study, there was no cue given about the upcoming perturbation, and movements' internal representations were identical across baseline and force field trials. Thus, it was the physical state of the limb during movement execution that evoked changes in neural representations of reach dynamics.

The retrieval of distinct motor memories was highlighted by the increase in online adaptation rate responsible for the gradual increase in lateral modulation of hand velocity (Fig. 4). This result highlights that participants were able to recognize a previously experienced perturbation, or assumed that the ongoing perturbation was the same as the previously experienced ones, and increased the within-trial adaptation rate. However, it may not be necessary to assume that previously learned internal models are retrieved. Indeed, our data indicated that adjustments could happen within the very first trial, and this is captured in the model in which adaptation can occur in real-time through parameter tracking without assuming a pre-stored alternative model. We suggest that online tracking in real-time combined with a change in online learning rate upon re-exposure to a previously experienced perturbation is a parsimonious account of our data.

It remains unclear how quickly these adjustments occurred within movements, and this represents an important but challenging question for future research. Anatomically, long-latency corrections for perturbation to the limb (~60ms) are supported by a distributed network through cerebellum [37] primary somatosensory area, motor and pre-motor cortices, and parietal cortex [28,38,39]. This network clearly overlaps with the main regions associated with short-term plasticity [1]. Behaviorally, long-latency feedback exhibits exquisite sensitivity to even very small disturbances [22]. Thus, the network of brain regions engaged in trial-to-trial learning is also likely recruited very early during unexpected force field trials, and it is conceivable that rapid adjustments occurred within this rapid pathway.

Computationally, however, the correction of internal models involves a less direct mapping between sensory feedback and the internal representations of movement dynamics, in addition to a change in the neural controller. This computational load may require longer processing times

associated with model updates in comparison with feedback responses following perturbations to the limb or to the movement goal. As an absolute upper bound, we may consider that adjustments in Experiment 2 were performed between reach onset and the moment when participants exited the via-point, which was typically less than 0.5s. We expect that future work aim at measuring the latency of this outer loop more accurately.

Another clear challenge for future work is to unravel the underlying neural mechanism. The key point is to understand how the brain generates a near instantaneous after effect like those observed in Experiment 2. Indeed, at the via-point with near-zero velocity across all trials, there was no motor error to correct; yet the subsequent motor output reflected previously encountered disturbances. One possibility to account for this result is through synaptic plasticity [1]: that is the gain of the same sensorimotor loop changes due to re-afference of sensory prediction errors inducing changes in synaptic strength. Another possibility is that the unexpected force field generated changes in the state of the sensorimotor system along a dimension which does not influence the motor output [40,41], allowing similar biomechanical state (i.e. stopping at the via-point) followed by a distinct motor output across force field trials when leaving the via-point. Such mechanism does not necessarily engage changes in synaptic weights, which is computationally advantageous [23]. Unraveling the latency of the model updates will likely set physiological constraints on the candidate neural underpinnings.

Finally, our study also opens questions from a theoretical standpoint. Indeed, we must recall that our model was excessively simple: we considered the translation of a point-mass in the plane with full state information (see Methods and Supplementary Material). We thus expect that future modeling work unveil the computational challenges that arise when considering more realistic models of the neuro-musculo-skeletal system, including nonlinear dynamics, sensorimotor noise, and temporal delays.

305 # Methods

306 **Experimental Procedures**

307 Healthy volunteers (n=24) provided informed written consent and the experimental
308 procedures were approved by the ethics committee at the host institution (UCLouvain). Participants
309 grabbed the handle of a robotic device (KINARM, BKIN Technologies, Kingston, Canada) and were
310 instructed to perform reaching movements towards a visual target. The movement onset was cued
311 by filling in the goal target (Fig. 1a), and feedback was provided about movement timing (good trials
312 were between 600ms and 800ms following the go signal). In Experiment 1 (n = 14), baseline trials
313 were randomly interleaved with orthogonal force field trials mapping the forward hand velocity onto
314 a lateral force ($F_x = \pm L\dot{y}, L = 13\ Nsm^{-1}$). Movements consisted of 15cm reaches in the forward
315 direction. To make the task less predictable, other trial types were also randomly interleaved
316 including trials with a via-point situated on the side of the straight reach path, and trials with a
317 constant background load applied to the hand ($F_x = \pm 4N$). Participants performed six blocks of 80
318 trials including 4 force field trials per direction per block, summing to a total of 24 force field trials
319 per direction and participant.

320 In Experiment 2, movements consisted of 18cm reaches in the forward direction, and the via-
321 point (radius: 1cm) was located at 12cm on the straight line joining the start and goal target.
322 Participants performed six blocks of 80 trials composed of 45 baseline trials, 5 force field trials, 5
323 baseline trials with a via-point, 5 force field trials with a via-point, and 5 force field trials with a via-
324 point and with a switching off of the force field. All trials and force field directions were randomly
325 interleaved. The experimental setup monitored the hand velocity in real time allowing to turn off the
326 force field if (1) the hand cursor was in the via-point and (2) the hand velocity dropped below 3cm/s
327 while in the via-point (determined based on preliminary testing). Feedback about a successful
328 stopping at the via-point was given online. This task was difficult and we recorded an average of 21,
329 25 and 17 trials for counterclockwise, baseline and clockwise via-point trials meeting these two
330 conditions, respectively (range across subjects: [5, 30], [18, 29], [2, 30]).

Analysis and Statistical Design

Position and force signals were sampled at 1kHz. Velocity signals were derived from numerical differentiation (2nd order finite difference algorithm). The presence of significant effects was assessed with repeated-measures ANOVAs (rmANOVA). In experiment 1, one participant missed a force-field trial such that the analyses were performed on 23 trials for the corresponding force field direction. Post-hoc tests were performed based on paired t-test with Bonferroni correction for multiple comparisons. The exponentials were fitted in the least-square sense and the significance of the fit was determined based on whether the 95% confidence interval of the exponent responsible for the curvature of the fit included or not the value of zero. An R^2 statistics was derived as follows:

$$R^2 = 1 - \frac{V_{RES}}{V_{TOT}}, \quad (1)$$

where V_{RES} denotes the variance of the residuals and V_{TOT} is the variance of the data. We adopted Cohen's d to quantify the effect size for paired comparison as the mean difference between two populations divided by the standard deviation of the paired difference between the populations [42]. The mixed linear model of lateral hand velocity as a function of trial number was based on standard techniques [43]. We changed the sign of the lateral velocity for clockwise force field trials such that for all trials, a positive modulation reflected the presence of an after effect. The dependent variable was the lateral hand velocity measured at the moment of the second velocity peak, the fixed factor was the trial index, and the random effect was the participant.

Model

The model is based on the translation of a point-mass with unity mass in the horizontal plane. The state-space representation is:

$$\ddot{x} = -k\dot{x} + F_x + F_{ENV}, \quad (2)$$

$$\ddot{y} = -k\dot{y} + F_y, \quad (3)$$

$$\tau\dot{F}_x = u_x - F_x, \quad (4)$$

$$\tau \dot{F}_y = u_y - F_y, \quad (5)$$

where k is a dissipative constant (0.1 Nsm), Equations 4 and 5 capture the first order muscle dynamics with time constant set to 0.1s, and F_{ENV} represents the unknown environmental perturbation: $F_{ENV} = 0$ for the null field, or $F_{ENV} = \pm L\dot{y}$ for force field trials.

This system of differential equations can be transformed in a discrete-time state space representation allowing us to add stochastic disturbances. Let x^* and y^* denote the coordinates of the goal target relative to the start, $\boldsymbol{z}_t := [x_t,\ y_t,\ \dot{x}_t,\ \dot{y}_t,\ F_{x,t},\ F_{y,t},\ x^*,\ y^*]^T$ represent the state vector at the time step t, and $\boldsymbol{u}_t := [u_{x,t}, u_{y,t}]^T$ represent the two-dimensional control vector, then the system can be written as follows:

$$\boldsymbol{z}_{t+1} = \left(\hat{A}_t + \Delta A\right)\boldsymbol{z}_t + B\boldsymbol{u}_t + \boldsymbol{\varepsilon}_t, \quad (6)$$

where $\hat{A}_t$ is the current (time varying) expected dynamics, ΔA is the unknown model error containing the unmodeled environmental disturbance (F_{ENV}), and $\boldsymbol{\varepsilon}_t$ is a Gaussian disturbance with zero-mean and covariance $\Sigma_\varepsilon \coloneqq BB^T$. We denote the true unforced dynamics by the matrix $A \coloneqq \hat{A}_t + \Delta A$. Details about these matrices can be found in the Supplementary Material.

We assumed for simplicity that the controller has perfect state measurement, meaning that $\boldsymbol{z}_t$ is known (fully observable case). The problem is that the controller does not know ΔA. This problem was approached in the context of sensorimotor control assuming that ΔA followed a known (Gaussian) distribution [27], which induces state-dependent noise. Observe that this assumption implies that the force field intensity varies randomly over time, which is not consistent with our experiment for which the force field varied randomly across trials, but remained fixed within trials. In other words, stochastic optimal control does not apply in the context of this paper because ΔA depends on F_{ENV}, which is not a stochastic variable but instead fixed bias or error in the model. Thus the initial controller is derived assuming that the system dynamics corresponds to A. To further simplify the problem, we assume that the controller knows which parameter is unknown (L). Let $\hat{A}_t$ denote the expected dynamics at time t dependent on the current estimate of the force field

intensity noted $\hat{L}_t$, a Linear-Quadratic-Gaussian controller can be derived giving a time series of control signals based on linear feedback gains that depend on $\hat{A}$:

$$\boldsymbol{u}_t = -C_t(\hat{A}_t)\boldsymbol{z}_t. \quad (7)$$

Now, under this assumption, we may predict the next state vector, which we designate by $\hat{\boldsymbol{z}}_t$. When the true dynamics is distinct from the expected dynamics ($\Delta A = A - \hat{A}_t \neq 0$), there is a prediction error, $\boldsymbol{e}_t \coloneqq \boldsymbol{z}_t - \hat{\boldsymbol{z}}_t$, which can be used to correct the estimate of the model. We follow standard Recursive Least Square identification techniques (RLS), and use the following adaptive rule to update the estimate of L [25]:

$$\hat{L}_{t+1} = \hat{L}_t + \gamma \frac{\partial \hat{\boldsymbol{z}}_t}{\partial L} \boldsymbol{e}_t, \quad (8)$$

where γ is the online learning rate. This new value of estimated force field intensity is then used to update in $\hat{A}_{t+1}$, and a new series of feedback gains is computed and applied to the next state as control law: $\boldsymbol{u}_{t+1} = -C_{t+1}(\hat{A}_{t+1})\boldsymbol{z}_{t+1}$. In all, the closed loop controller consists in iteratively deriving optimal control gains based on an adaptive representation of the unknown parameters recursively estimated using Eqn 8. It can be shown in theory that under reasonable assumptions, this adaptive optimal control scheme robustly stabilizes the closed loop system [25].

The free parameters of the model are the cost-function used to derive the LQG controller, and the online learning rate γ. We fixed the cost-function to generate smooth movements with bell-shape velocity profiles comparable with human movements qualitatively and allowed γ to vary in order to capture the greater online adjustments observed across trials in Experiment 1. This parameter was fixed for the simulations of the learning curve, and for the simulation of the via-point experiment. The simulation of a learning paradigm (Fig. 5b) was performed by letting $\hat{A}_t$ evolve during movements and storing the value obtained at the end of the trial for the beginning of the next trial ($\gamma = 0.05$). The reduction in peak end-force observed in Experiment 1 was simulated by using low or high values of γ during movement ($\gamma = 0.01$ and 0.1, respectively, Fig. 5c). Finally Experiment

3 was simulated as a series of null field trials interleaved with force-field trials with switching-off at the via-point ($\gamma = 0.25$, Fig. 5c). Fig. 5c reports the force field trials, and the baseline trials preceded by at least 3 baseline trials to washout the impact of each force field trial.

411 ## Figure Legends

412 **Figure 1: Experiment 1, Behavior.** a: Illustration of a successful trial in Experiment 1. Participants
413 waited for the go signal, and were instructed to reach for the goal target within 0.6 s to 0.8 s of the
414 go signal. Feedback about movement timing was provided to encourage participants to adjust their
415 movement speed, but all trials were included in the dataset. b: Hand paths from all participants for
416 the first force field trial (top) and the force field trial number 20 (bottom), were selected to illustrate
417 the change in online control. Counter-clockwise (CCW) and clock-wise (CW) FF trials are depicted in
418 blue and red, respectively. The black traces (baseline) are randomly selected trials from each
419 participant. The baseline trials were performed in the null field, following at least 3 null-field trials.
420 The maximum hand path deviation in the direction of the force field (Max.) and the maximum target
421 overshoot in the opposite direction (Osh.) were extracted. c: Maximum hand path deviation for CCW
422 and CW force fields across trials. The dashed traces illustrate the non-significant exponential fits. d:
423 Maximum target overshoot across FF trials. Solid traces indicate significant exponential decay.
424 Shaded blue and red areas around the means indicate the SEM across participants (n=14).

425

426 **Figure 2. Improvement of Online Corrections.** a: lateral force during the first (dashed/light) and last
427 (solid/dark) trials. Shaded areas represent one SEM across participants (n=14) with the same color
428 code as in Fig.1. The reduction in peak end force is highlighted with the black arrows. b: Peak end
429 force across force field trials (mean±SEM). The solid traces show the significant exponential decay. c:
430 Correlation between the lateral hand velocity and the force generated by the force field during the
431 time interval corresponding to the gray rectangle in panel a (from 100ms to 700ms). The correlations
432 were averaged across participants; the shaded area represents one SEM. The significant linear
433 regression of the correlations across blocks is illustrated.

434

435

Figure 3: Experiment 2 Behavior. a: Average hand traces from Experiment 2 for trials with a via-point with the same color code as in Experiment 1 (black: null-field, red: CW, blue: CCW). For all trials, the second portion of the reach, from the via-point to the goal target, was performed without force field (FF Off). b: Average x-coordinate as a function of time across participants (n=10). Shaded areas represent one SEM across participants. The gray rectangle illustrates the portion of the movement spent in the via-point approximately. Observe the displacement following the via-point in the opposite direction as the perturbation-related movement. c: Illustration of the trial-by-trial extraction of parameters on five randomly selected trials from one participant. The force field was turned off when the hand speed dropped below 3 cm/s (dashed trace). The parameters extracted were the lateral hand velocity at the moment of the hand speed minimum (squares), and the lateral hand velocity at the moment of hand speed maximum (open discs). d: Lateral hand velocity at the moment of hand speed minimum or at the moment of the second hand speed maximum averaged across trials for each participant (gray dots), and averaged across participants (blue, black and red, mean±SEM). Individual means across the three trial types were subtracted from the data for illustration purpose. The two stars represent significant pair-wise differences at the level $P<0.01$ with Bonferroni correction for multiple comparisons.

Figure 4: Learning across Blocks. a: Lateral hand velocity measured at the second peak hand speed following CCW and CW force-fields (red and blue, respectively), and baseline trials (black) across blocks. There were blocks in which participants did not succeed any force-filed trial, thus displays are the mean±SEM across participants in each block with $8 \leq n \leq 10$. The insets highlight the first and last block. b: Individual average across trials for each participant (gray dots), and group mean±SEM in the first and 6th block (n = 9 in both cases). The two stars highlight significant difference based on paired t-test ($P < 0.01$).

Figure 5. Model Simulations. a: Schematic illustration of an adaptive controller: sensory feedback is used to estimate both the state of the system and the model, which in turn updates the controller and estimator online. The solid arrows represent the inner loop correcting for state deviations and the dashed arrows represent the outer loop correcting for model errors. b: Reproduction of a standard learning experiment (hand paths and maximum lateral deviation across trials), with mirror-image catch trials (ct), and unlearning following the catch trials (difference between ct-1 and ct+1). The trial-by-trial changes were reproduced with online computations exclusively (inspired by Thoroughman and Shadmehr [6]). c: Simulation of the results from Experiment 1: the adaptive controller reduces the peak end force by increasing the online learning rate, which improves the tuning of online corrections to the properties of the force field. d: Simulations of behavior from Experiment 2, the force field was turned off at the via-point, and the second portion of the reach displays an after effect (see Methods).

Acknowledgements

FC is supported by a grant from the Belgian Fund for Scientific Research (F.R.S.-FNRS, Chercheur Qualifié, grant number 1.C.033.18F). This work was supported by grants from PRODEX and by the Interuniversity Attraction Pole of the Belgian Science Policy (BELSPO).

References

1 Dayan, E. & Cohen, L. G. Neuroplasticity subserving motor skill learning. *Neuron* **72**, 443-454, doi:10.1016/j.neuron.2011.10.008 (2011).

2 Lackner, J. R. & DiZio, P. Rapid adaptation to Coriolis-force perturbations of arm trajectory. *Journal of Neurophysiology* **72**, 299-313 (1994).

3 Shadmehr, R. & Mussa-Ivaldi, F. A. Adaptive representation of dynamics during learning of a motor task. *Journal of Neuroscience* **14**, 3208-3224 (1994).

4 Krakauer, J. W., Ghilardi, M. F. & Ghez, C. Independent learning of internal models for kinematic and dynamic control of reaching. *Nature Neuroscience* **2**, 1026-1031 (1999).

5 Wagner, M. J. & Smith, M. A. Shared Internal Models for Feedforward and Feedback Control. *Journal of Neuroscience* **28**, 10663-10673, doi:10.1523/jneurosci.5479-07.2008 (2008).

6 Thoroughman, K. A. & Shadmehr, R. Learning of action through adaptive combination of motor primitives. *Nature* **407**, 742-747 (2000).

7 Singh, K. & Scott, S. H. A motor learning strategy reflects neural circuitry for limb control. *Nature Neuroscience* **6**, 399-403, doi:10.1038/nn1026 (2003).

8 Lackner, J. R. & DiZio, P. Motor control and learning in altered dynamic environments. *Current Opinion in Neurobiology* **15**, 653-659, doi:10.1016/j.conb.2005.10.012|ISSN 0959-4388 (2005).

9 Shadmehr, R., Smith, M. A. & Krakauer, J. W. Error Correction, Sensory Prediction, and Adaptation in Motor Control. *Annual Review of Neuroscience, Vol 33* **33**, 89-108 doi:10.1146/annurev-neuro-060909-153135 (2010).

10 Wolpert, D. M., Diedrichsen, J. & Flanagan, J. R. Principles of sensorimotor learning. *Nature Reviews Neuroscience* **12**, 739-751, doi:10.1038/nrn3112 (2011).

11 Smith, M. A., Ghazizadeh, A. & Shadmehr, R. Interacting adaptive processes with different timescales underlie short-term motor learning. *Plos Biology* **4**, 1035-1043, doi:10.1371/journal.pbio.0040179 (2006).

12 Kording, K. P., Tenenbaum, J. B. & Shadmehr, R. The dynamics of memory as a consequence of optimal adaptation to a changing body. *Nat Neurosci* **10**, 779-786, doi:10.1038/nn1901 (2007).

13 Gonzalez Castro, L. N., Hadjiosif, A. M., Hemphill, M. A. & Smith, M. A. Environmental consistency determines the rate of motor adaptation. *Curr Biol* **24**, 1050-1061, doi:10.1016/j.cub.2014.03.049 (2014).

14 Wei, K. & Koerding, K. Uncertainty of feedback and state estimation determines the speed of motor adaptation. *Frontiers in Computational Neuroscience* **4**, doi:10.3389/fncom.2010.00011 (2010).

15 Diedrichsen, J., White, O., Newman, D. & Lally, N. Use-dependent and error-based learning of motor behaviors. *J Neurosci* **30**, 5159-5166, doi:10.1523/JNEUROSCI.5406-09.2010 (2010).

16 Franklin, D. W. & Wolpert, D. M. Computational Mechanisms of Sensorimotor Control. *Neuron* **72**, 425-442, doi:10.1016/j.neuron.2011.10.006 (2011).

17 Burdet, E. *et al.* A method for measuring endpoint stiffness during multi-joint arm movements. *Journal of Biomechanics* **33**, 1705-1709, doi:10.1016/s0021-9290(00)00142-1 (2000).

18 Crevecoeur, F. & Scott, S. H. Beyond Muscles Stiffness: Importance of State Estimation to Account for Very Fast Motor Corrections. *PLoS Computational Biology* **10**, e1003869 (2014).

19 Kurtzer, I. L., Pruszynski, J. A. & Scott, S. H. Long-latency reflexes of the human arm reflect an internal model of limb dynamics. *Current Biology* **18**, 449-453, doi:10.1016/j.cub.2008.02.053 (2008).

20 Cluff, Y. & Scott, S. H. Rapid feedback responses correlate with reach adaptation and properties of novel upper limb loads. *Journal of Neuroscience* **33**, 15903-15914 (2013).

21 Crevecoeur, F. & Scott, S. H. Priors Engaged in Long-Latency Responses to Mechanical Perturbations Suggest a Rapid Update in State Estimation. *Plos Computational Biology* **9**, doi:10.1371/journal.pcbi.1003177 (2013).
22 Crevecoeur, F., Kurtzer, I. & Scott, S. H. Fast corrective responses are evoked by perturbations approaching the natural variability of posture and movement tasks. *Journal of Neurophysiology* **107**, 2821-2832, doi:10.1152/jn.00849.2011 (2012).
23 Fortney, K. & Tweed, D. B. Computational Advantages of Reverberating Loops for Sensorimotor Learning. *Neural Computation* **24**, 611-634 (2012).
24 Ioannou, P. & J, S. *Robust Adaptive Control*. (Prentice Hall Inc., 1996).
25 Bitmead, R. R., Gevers, M. & Wertz, V. *Adaptive optimal control : the thinking man's GPC*. (Prentice Hall, 1990).
26 Todorov, E. Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system. *Neural Computation* **17**, 1084-1108 (2005).
27 Izawa, J., Rane, T., Donchin, O. & Shadmehr, R. Motor adaptation as a process of reoptimization. *Journal of Neuroscience* **28**, 2883-2891, doi:10.1523/jneurosci.5359-07.2008 (2008).
28 Scott, S. H. A Functional Taxonomy of Bottom-Up Sensory Feedback Processing for Motor Actions. *Trends Neurosci* **39**, 512-526, doi:10.1016/j.tins.2016.06.001 (2016).
29 Braun, D. A., Aertsen, A., Wolpert, D. M. & Mehring, C. Learning Optimal Adaptation Strategies in Unpredictable Motor Tasks. *Journal of Neuroscience* **29**, 6472-6478, doi:10.1523/jneurosci.3075-08.2009 (2009).
30 Scheidt, R. A., Dingwell, J. B. & Mussa-Ivaldi, F. A. Learning to move amid uncertainty. *Journal of Neurophysiology* **86**, 971-985 (2001).
31 Ghahramani, Z. & Hinton, G. E. Parameter estimation for linear dynamical systems. *Technical Report CRG-TR-96-2, University of Toronto, Dept. of Computer Sceince* (1996).
32 Astrom, K. J. *Introduction to stochastic control theory*. (Academic Press, 1970).
33 Schaal, S., Sternad, D., Osu, R. & Kawato, M. Rhythmic arm movement is not discrete. *Nat Neurosci* **7**, 1136-1143 (2004).
34 Hirashima, M. & Nozaki, D. Distinct Motor Plans Form and Retrieve Distinct Motor Memories for Physically Identical Movements. *Current Biology* **22**, 432-436, doi:10.1016/j.cub.2012.01.042 (2012).
35 Howard, I. S., Ingram, J. N., Franklin, D. W. & Wolpert, D. M. Gone in 0.6 seconds: the encoding of motor memories depends on recent sensorimotor states. *J Neurosci* **32**, 12756-12768, doi:10.1523/JNEUROSCI.5909-11.2012 (2012).
36 Sheahan, H. R., Franklin, D. W. & Wolpert, D. M. Motor Planning, Not Execution, Separates Motor Memories. *Neuron* **92**, 773-779, doi:10.1016/j.neuron.2016.10.017 (2016).
37 Hore, J. & Flament, D. Changes in motor cortex neural discharge associated with the development of cerebellar limb ataxia. *J Neurophysiol* **60**, 1285-1302 (1988).
38 Pruszynski, J. A. *et al.* Primary motor cortex underlies multi-joint integration for fast feedback control. *Nature* **478**, 387-390, doi:10.1038/nature10436 (2011).
39 Omrani, M., Murnaghan, C. D., Pruszynski, J. A. & Scott, S. H. Distributed task-specific processing of somatosensory feedback for voluntary motor control. *Elife* **5**, doi:10.7554/eLife.13141 (2016).
40 Kaufman, M. T., Churchland, M. M., Ryu, S. I. & Shenoy, K. V. Cortical activity in the null space: permitting preparation without movement. *Nat Neurosci* **17**, 440-448, doi:10.1038/nn.3643 (2014).
41 Stavisky, S. D., Kao, J. C., Ryu, S. I. & Shenoy, K. V. Motor Cortical Visuomotor Feedback Activity Is Initially Isolated from Downstream Targets in Output-Null Neural State Space Dimensions. *Neuron* **95**, 195-208 e199, doi:10.1016/j.neuron.2017.05.023 (2017).
42 Lakens, D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front Psychol* **4**, 863, doi:10.3389/fpsyg.2013.00863 (2013).

43 Laird, N. M. & Ware, J. H. Random-effects models for longitudinal data. *Biometrics* **38**, 963-974 (1982).

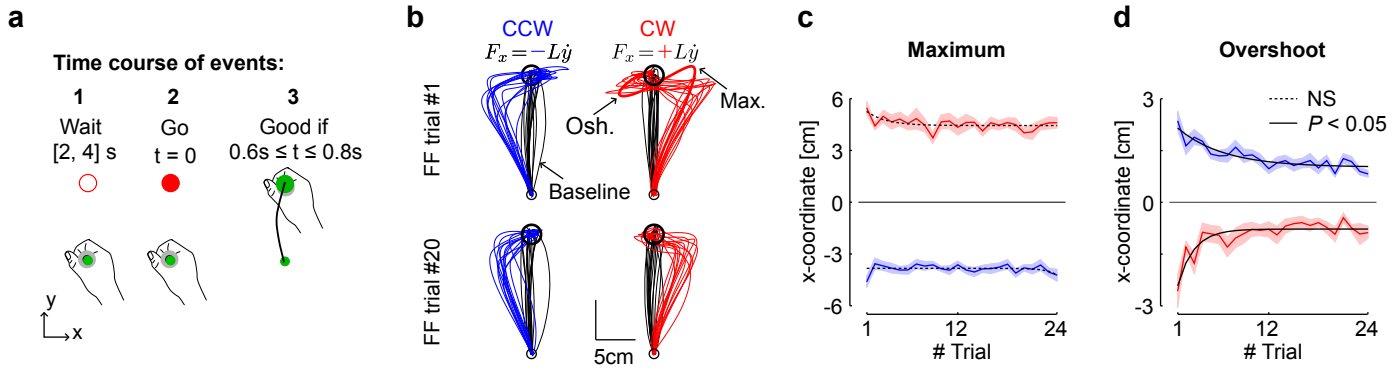
a
Time course of events:
1
Wait
[2, 4] s
2
Go
t = 0
3
Good if
0.6s ≤ t ≤ 0.8s
y
x
b
CCW
$F_x = -L\dot{y}$
CW
$F_x = +L\dot{y}$
FF trial #1
FF trial #20
Osh.
Max.
Baseline
5cm
c
Maximum
x-coordinate [cm]
6
3
0
-3
-6
1
12
24
Trial
d
Overshoot
NS
$P < 0.05$
x-coordinate [cm]
3
0
-3
1
12
24
Trial

Figure 1

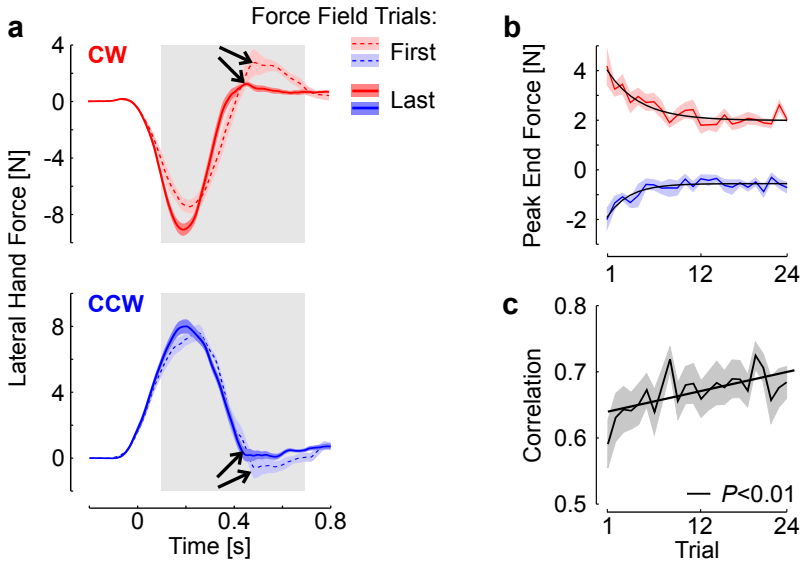
a
Force Field Trials:
First
Last
CW
CCW
Lateral Hand Force [N]
4
0
-4
-8
8
4
0
0
0.4
0.8
Time [s]
b
Peak End Force [N]
4
2
0
-2
1
12
24
c
Correlation
0.8
0.7
0.6
0.5
— P<0.01
1
12
24
Trial

Figure 2

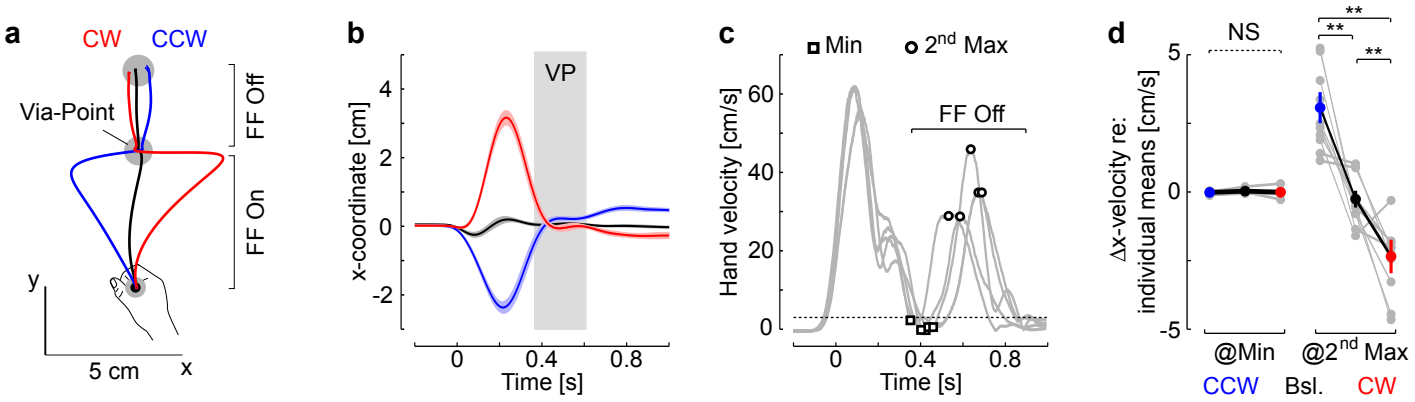
a
CW
CCW
Via-Point
FF Off
FF On
y
5 cm
x
b
VP
x-coordinate [cm]
4
2
0
-2
0
0.4
0.8
Time [s]
c
Min
2nd Max
FF Off
Hand velocity [cm/s]
60
40
20
0
0
0.4
0.8
Time [s]
d
NS
**
**
**
Δx-velocity re: individual means [cm/s]
5
0
-5
@Min
@2nd Max
CCW
Bsl.
CW

Figure 3

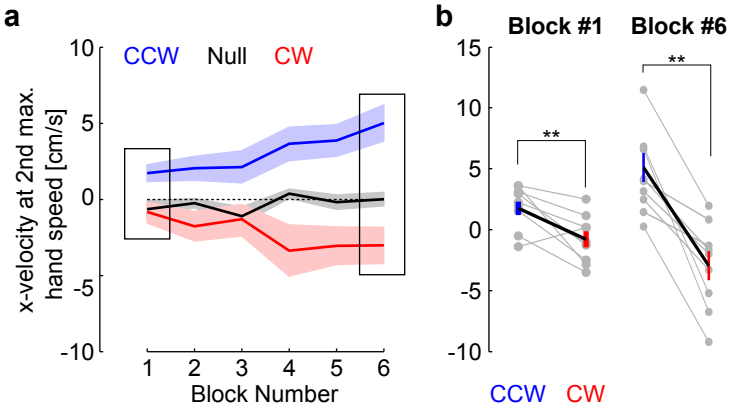
a
CCW
Null
CW
x-velocity at 2nd max. hand speed [cm/s]
10
5
0
-5
-10
1
2
3
4
5
6
Block Number
b
Block #1
Block #6
15
10
5
0
-5
-10
**
**
CCW
CW

Figure 4

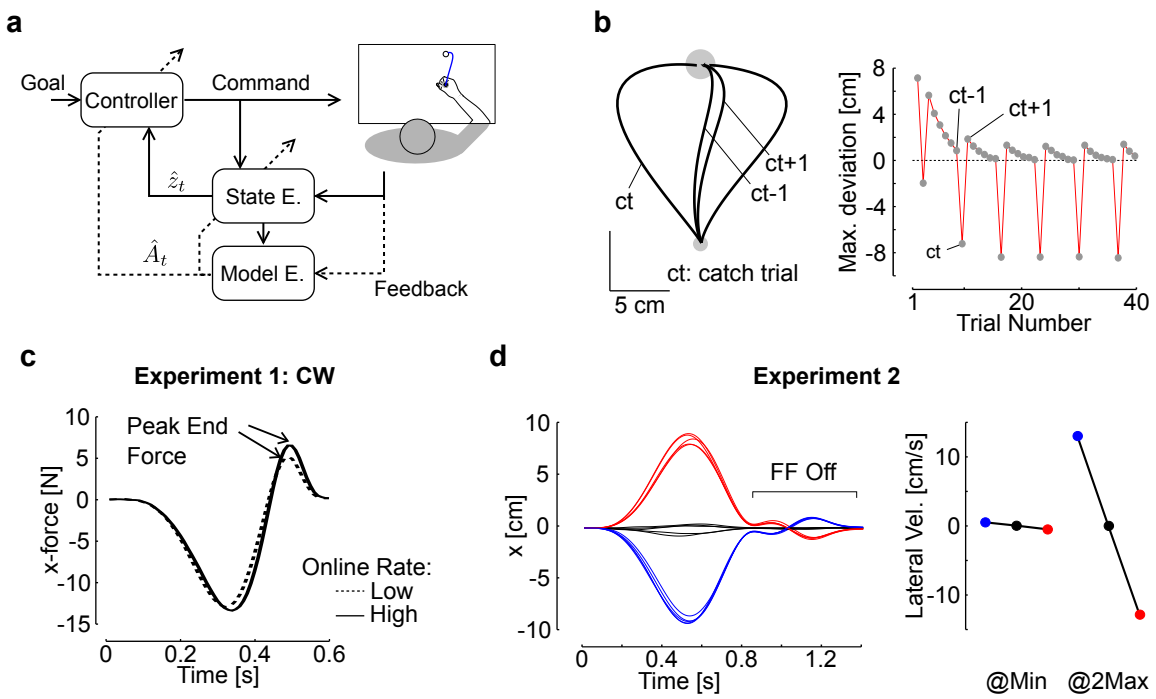
a
Goal
Controller
Command
State E.
Model E.
Feedback
b
ct
ct+1
ct-1
ct: catch trial
5 cm
Max. deviation [cm]
Trial Number
c
Experiment 1: CW
Peak End Force
x-force [N]
Time [s]
Online Rate:
Low
High
d
Experiment 2
FF Off
x [cm]
Time [s]
Lateral Vel. [cm/s]
@Min
@2Max

Figure 5