Human reaching control in dynamic environments

Hari T. Kalidindi[1,2], Frédéric Crevecoeur[1,2]

[1]*Institute of Information and Communication Technologies, Electronics and Applied Mathematics*

[2]*Institute of Neuroscience*

University of Louvain (UCLouvain), Belgium

Corresponding Author

F. Crevecoeur
4 avenue Georges Lemaitre
1348, Louvain-le-Neuve, Belgium

Tel.: +3210478040
Frederic.crevecoeur@uclouvain.be

Abstract

Closed-loop models of movement control have attracted growing interest in how the nervous system transforms sensory information into motor commands, and several brain structures have been identified as neural substrates for these computational operations. Recently, several studies have focused on how these models need to be updated when environmental parameters change. Current evidence suggests that when the task changes, rapid control updates enable flexible modifications of current actions and online decisions. At the same time, when movement dynamics change, humans use different strategies based on a combination of adaptation and modulation of controller sensitivity to exogenous perturbations (robust control). This review proposes a unified framework to capture these results based on online estimation of model parameters with dynamic updates in control. The reviewed studies also identify the time scales of associated behavioral mechanisms to guide future research on the neural basis of movement control.

Highlights

- Stochastic optimal control with flexible use of varying sensory inputs is a building block model for postural control and reaching.
- Current research has investigated the rapid updates required in a dynamic context.
- Task structure and target rewards update control in 150-300ms.
- Model-robust control strategies and online adaptive control explain feedback corrections and short-term adaptation.

Introduction

Current research into how the brain controls reaching movements builds on a large body of work exploiting the model introduced by Todorov and Jordan's seminal paper, in which stochastic optimal control (SOC) is used as a theory to explain how humans achieve consistent success despite motor variability [1]. The model features a controller that produces optimal motor commands based on an estimate of the limb's state, a function called control policy, derived to minimize a cost function (see Box 1). This theory has sparked great interest in the study of how humans react to a wide range of disturbances. Mechanical and visual disturbances elicit a sequence of responses characterized by rapid but automatic compensation for the disturbance, followed by a slower but more sophisticated response. By sophisticated, we mean that these responses vary according to a number of factors linked to the dynamics and instructions of the task, as predicted in theory [2].

Studies on feedback responses to mechanical loads have concluded that long-latency feedback pathways, which influence motor responses in the upper limb within 50-60ms of load onset, implement a flexible state feedback controller (see Box 1), that has the same properties as the controller derived in the context of SOC. Flexible responses to visual disturbances appeared at around 120ms [3,4]. Mechanical and visual perturbations also result in faster responses (20-30 ms and ~90 ms, respectively), but these components appear to be automatic and may not vary much with task requirements. Subsequent studies have characterized other sensory systems such as the vestibular system, which contributes to flexible hand control with a longer latency (>150ms) [5,6], and tactile feedback, which has a latency of around 100ms (Figure 1, colored arrows) [7–10]. Thus, the SOC framework with different sensory inputs and their respective latencies can be used as a building block to model postural control and reaching.

In general, modeling studies have considered static environments, meaning that the parameters of the corresponding SOC model were constant for a given movement. These parameters correspond to (1) movement costs, (2) internal models of the limb dynamics, and (3) hypotheses about exogenous perturbation signals. However, we often need to produce movements in dynamic environments, quickly selecting between different movement goals amidst numerous interactions. In a paradigmatic shift towards describing sensorimotor control in dynamic environments, current research has investigated situations in which some parameters changed rapidly across trials or even within movements: cost parameters are linked to movement planning or re-planning; the parameters linked to the internal models determine the adaptation of the controller; and the assumptions made about the exogenous disturbance signal determine whether a control policy must be more or less sensitive to the errors in the internal model of movement. We argue below that all of these components are not fixed, but modulate control during movement, which requires formulation of optimal control models in combination with online estimation of their parameters (see Box 2). Section

2 addresses dynamic changes in the task while Section 3 surveys dynamic changes in movement dynamics. We emphasize that multiple aspects of motor behavior including online decisions regarding the movement goal, re-planning, rapid motor adaptation, and model-robust changes in control policies can be modeled in this framework.

Changes in task parameters

Task instructions can be expressed as a parametric cost-function that the controller must minimize. In standard SOC, the cost-function takes the form of a quadratic penalty on deviations of the state from a set of target values (which can be a desired trajectory [11]), plus a quadratic penalty on motor commands which represents metabolic expenses. Such a cost-function can account for target redundancy by providing degrees of freedom in any subset of state variables. Indirectly, this formulation can encode how rewards affect a movement by varying the penalty associated with the target (it is costlier to miss a more rewarding target). As a consequence, one can model a wide range of experiments that probe the effects of task instructions, target structure, and rewards with dynamic updates to the cost-function.

Studies investigating the influence of goal structure and target rewards have concluded that changes in task demands update control very rapidly, and even online. Česonis and Franklin [12] varied the task from trial to trial, by instructing participants whether they had to stop at a target or hit it. Implicitly, these instructions induce different penalties on the terminal velocity, with a higher cost of non-zero velocity at the target for the stopping task. This led to flexible feedback responses, as shown previously [13], with participants making more complete responses to target jumps when they were not required to stop. It was suggested that partial corrections in the stopping task resulted from the cost of terminal velocity, which outweighed the cost of terminal error, making participants to deliberately undershoot the target. Česonis and Franklin further emphasized that the switching of the controllers could be performed between trials on a randomised schedule.

A similar question was investigated with geometric changes in the structure of the goal presented as a narrow square or a wide rectangle by de Comité et al. [14,15]. The purpose of using a wide target was to provide a degree of freedom in the movement goal, which evokes selective control of lateral velocity and force, without constraining the lateral position. Here, the target was changed during movement and lateral forces were used to determine whether the nervous system took the new structure into account. The results showed that when switching the target from a square to a rectangle or vice versa, the feedback responses were decreased or increased, respectively (Figure 2, a). Mainly, the study showed through muscle activity measurements that the new target structure influenced control with an average latency of 150ms after the target switch. In the follow up study, de Comité et al. highlighted that the modulation of control policies scaled with the rate of change of the target width [14], suggesting that dynamic variables were taken into account in this process. In all, when we start a movement, the control policy is not fixed from the end of the planning phase until the end of the movement, but is updated in flight as circumstances change.

Besides the effect of task instructions, it was recently shown that rewards modulate feedback gains [16], showing that they are an input to the building block of Figure 1. Cos and Cisek's group demonstrated that non-uniform distributions of visual rewards over a wide target, together with the actual limb position, determined the initial direction, changes of mind, and rerouting to new locations following changes in reward distribution (Figure 2, b) [17,18]. The authors suggested that the mechanism underlying target selection allows for online updates of control parameters at latencies of ~300ms as measured from hand kinematics. In an experiment involving multiple targets with different rewards, de Comité and colleagues observed that larger rewards for the alternative options, and larger perturbation magnitudes both elicited more frequent switches to alternative goals [19]. These results supported the theory proposed by Cisek about a distributed consensus across perceptual and motor systems [20], in which decisions are made by combining top-down expectations and bottom-up biomechanical factors to select an action. This idea can be extended to online control as it turns out that the distributed consensus was at play during movement, where both the state of the limb and the rewards of alternative targets influence the decision to select a new goal. This led to the hypothesis that the cost-to-go, that is the total expected cost from the current time until movement end, is processed continuously in the brain to update the controller. This hypothesis has recently been used to explain why the timing of the decisions to change goal during a movement varied as a function of the rewards and ongoing trajectories [21].

Thus, the commitment to a control policy after the planning stage was not fixed until the end of the movement, instead changes in task requirements could modify the controller online. From a computational perspective, these observations can be added to an optimal control model with feedback on task parameters and modification of the control online (Figure 2, c). Adding feedback on task parameters captures the fact that the neural mechanism that determines a control policy based on task structure operates both before and during movement on a timescale of less than few hundreds of milliseconds.

Changes in movement dynamics

Classic adaptation paradigms [22] have shown that the nervous system acquires internal models of the body and environmental dynamics, a mechanism that can be used to produce anticipatory compensation for sustained changes in the environment. But how does the nervous system respond to less predictable disturbances in movement dynamics? A first response is simply to correct the limb trajectory based on the control architecture in Figure 1. However, both the controller and the state-estimator use knowledge of limb dynamics (Box 1), which if erroneous, may produce motor commands that are inoperant. There are two alternative approaches to address this problem that go beyond the capabilities of SOC: (1) robust control that limits the impact of disturbances without making assumptions about them, and (2) rapid adaptation, which consists, if possible, in acquiring knowledge of the dynamics and modifying the control during movement. Mapping these hypotheses onto behavioural and neurophysiological data has revealed that both solutions may be used in the brain.

Experimental observations have shown that humans respond to disturbances by co-contracting muscles. This response has been observed across several types of perturbations, including visual and mechanical loads [23–26]. Early responses to novel perturbations are not specific to the type of perturbation, and it has been hypothesized that their goal is to increase the mechanical impedance of the joints. However, standard measurements of impedance also depend on neural feedback via reflexes. An alternative view suggested that co-contraction reflects a robust control strategy [27]. Participants reached a visual target while force fields were applied randomly. They responded by increasing the strength of feedback corrections, but importantly they also increased hand velocities in the direction in which the external forces were absent – indicating a general increase in control gains. This pattern of responses was similar to simulated movements obtained with robust control. It impacted control strategies immediately after a single force field trial, and tended to decay in ~10 trials, showing that this process contributes to reducing motor errors in the short-term. Compatible with this idea, Coltman and colleagues highlighted modulation of long-latency stretch responses co-occurring with the fast time scale of adaptation [28].

Theoretical developments have further extended the application of robust control principles to sensorimotor learning, showing that efficient controllers can be constructed despite internal model errors [29,30]. However, the general increase in feedback gains observed experimentally is costly in terms of metabolic effort, so it is inefficient to apply it in situations where the nature of the perturbation can be learned and recognized. In this case, adaptation can help.

Adaptation can be defined as an internal update in the representation of movement dynamics. Recently it has been studied from the perspective of feedback control. There has been a traditional separation between *feedforward* (a pre-defined control sequence) and feedback control pathways, but adaptation in one or the other pathway is mirrored in the other [31–35]; furthermore, correlations between feedback responses and learning have been documented at a single-trial scale [36]. In addition, Joiner and Smith [37] showed that the adaptive responses to a single perturbation trial were temporally tuned to its profile as a function on position or velocity. Therefore, the learning rule in the brain may use a common state-space as SOC and robust control models.

To describe the computational mechanisms behind rapid adaptation, it has been shown that the feedback controller can adapt within an ongoing movement (as early as 250ms after the movement onset) [38,39]. The authors used randomly interleaved force field perturbations while participants made reaching movements. The feedback responses improved, without any anticipation, and became temporally tuned to the perturbation profile, similar to adaptive responses shown in consistent environments (Figure 3a). These adaptive responses could be explained in the framework of *adaptive optimal control* with online updates of the parameters describing limb dynamics. Similar updates were observed earlier in a visuomotor rotation task [40]. The learning rule in adaptive optimal control uses a gradient that computes the sensitivity of the sensory prediction error with respect to the system parameters, and then updates the estimation and control policy based on the novel representation of dynamics.

Interestingly, a different task appears to be governed by a similar gradient-based learning rule. Hadjiosif and colleagues asked participants to perform reaching movements under mirror reversal of hand feedback – a perturbation that changes the sign of the impact of motor output on sensory errors – and observed that performance deteriorated (Figure 3b) [41]. This result showed that policy updates govern short-term adaptation by assuming a relationship between errors and corrections that is invalidated by the mirror reversal. From the perspective of internal model theory [42], this result sets constraints on the nature of the internal model used for learning, or equivalently, on the class of disturbances to which the sensorimotor system can adapt in the short-term. Interestingly, the result is expected in the context of adaptive control [38] (Figure 3c), since the learning rule calculates the gradient of the estimated state as a function of the model parameter [43], which, without knowledge of the sign change induced by the mirror reversal, can produce the error expansion shown in Figure 3b.

Conclusion and Possible Directions

We proposed that SOC with dynamic parameter updates can explain the human ability to quickly incorporate contextual information to adjust the ongoing action. The formalism of model predictive control has been earlier suggested to capture updates to the control gains [44], which is extended here to a wide range of sensorimotor functions, including motor decisions and adaptation in a general framework. The framework in Box 2 allows for the updating of parameters, and the results reviewed above have identified the associated timescales. This framework also leaves a series of open questions that provide avenues and challenges for future work.

An obvious challenge to validate the algorithm proposed in Box 2 is to describe how parameter changes can be identified in the nervous system. While some parameters may be explicitly signaled, such as the target structure, others may be implicitly signaled by differences between actual and expected sensory information. For example, one can investigate which variables or error sizes may or may not produce dynamic updates to the controller. Concerning the implicit error signals, the question of parameter estimation remains open: for adaptation, the algorithm must be equipped with a learning or selection rule, which can feature parametric estimation and heuristic selection of "good" controllers according to hierarchical criteria [45]. Finally, this model invites us to revisit the standard assumptions about the existence of a feedforward controller in the brain. The theory proposed here can produce anticipatory actions as an expression of priors, as well as short-term adaptation without requiring any feedforward pathway [46], which, without questioning its existence, suggests at least a critical reexamination of what are the unequivocal behavioral proxies of this putative control scheme.

The formalism is useful for bridging sensorimotor control and the role of contextual factors. Recent models posit that probabilistic inference about the context determines updates across movements [47], leaving aside the formulation of a continuous-time control problem. As a result, there is a gap between control models that focus on time-varying state variables and control policies, and discrete-time models that use a movement as a unit of time and characterize updates across trials in more abstract terms. As parameter estimation must be included in dynamic control models, we believe that these two approaches are ready to meet.

Acknowledgements

The authors are supported by grants from F.R.S.-FNRS (grant numbers HK: FC43127 and ; FC: 1. C.033.18), and by a grant from UCLouvain (Concerted Research Action, ARC “coAction”).

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used “DeepL” and “Google Translate” in order to minimize grammatical errors. After using these tools, the author reviewed and edited the content as needed and take full responsibility for the content of the publication.

BOX 1: Stochastic Optimal Control

Standard models of sensorimotor control are based on *stochastic optimal control* (SOC). Here, the state vector is denoted by x_t, the control is u_t, the measurement of the state is y_t, and the relationship between these quantities is described by a linear dynamical system of the form:

$$x_{t+1} = Ax_t + Bu_t + Dv_t, \qquad (1)$$

$$y_t = Hx_t + Gw_t. \qquad (2)$$

It is assumed that the system is simulated over a fixed time horizon ($t = 1, 2, \ldots, N$), and that the initial condition is known, or up to an error of known distribution. The variable v_t is the exogenous perturbation signal, and w_t is the measurement noise. The noise vectors v_t and w_t are multivariate, zero-mean Gaussian random disturbances with unit variance. The covariance GG^T must be positive definite. This control system is coupled with cost-function:

$$J(x,u) = x_N^T Q_N x_N + \sum_{t=1}^{N-1} x_t^T Q_t x_t + u_t^T R u_t, \qquad (3)$$

with $Q_{t,t=1,\ldots,N}$ symmetric and non-negative, and R symmetric positive definite. The problem is to ***find the control sequence that minimizes the expected value of*** $\boldsymbol{J(x,u)}$**.** The optimal control action is a ***linear state-feedback controller,*** that is a linear mapping of state into a control vector as follows:

$$u_t = L_t x_t, \qquad (4)$$

where L_t are time varying ***feedback gains*** dependent on A, B, and on the cost parameters (Q_t and R). Equation 4 defines the ***control policy,*** which is rule applied to the current state of the system. When the state of the system is not available, an optimal linear observer can be defined as follows (Kalman filter):

$$\hat{x}_{t+1} = A\hat{x}_t + Bu_t + K_t(y_t - H\hat{x}_t), \qquad (5)$$

and the matrices K_t are calculated as a function of the system dynamics, observation matrix H, and the covariance matrices to minimise the estimation error.

Box 2: Optimal Control in Changing Environments: Model Predictive Control using continuous parameter estimation

Calling $\Theta \coloneqq \{A, B, Q_{t:N}, R\}$ the set of parameters known to the controller at each time step, we can rewrite:

$$u_t = L_t(\Theta)\hat{x}_t, \qquad (5)$$

and formulate the problem of control in dynamic environments by considering that the set of parameters varies over time. Defining $\Theta^{(i)} := \{A, B, Q_{t:N}^{(i)}, R\}$ the set of parameters corresponding to task i, the response to a change in task (Section 2) corresponds to changing the controller from $L_t(\Theta^{(1)})$ to $L_t(\Theta^{(2)})$. Note that movement time (N) is in Θ, and it can be selected dependent on the cost of time and its relationship with efforts [48,49].

Motor adaptation can be expressed as dynamic changes in the control policy in response to novel dynamics (Section 3). For example, a velocity dependent force field changes the matrix A. Assume the force field is $\vec{f} = C\dot{\vec{p}}$, with p representing the effector coordinate (C is a matrix), the dependency of the matrix A on C is denoted by $A(C)$, where a submatrix of A contains C. Adaptation can be defined as follows: let $\hat{A}$ be the estimate of $A(C)$, adapting the controller consists in updating $\hat{A}$ according to a learning rule that has yet to be defined. Note that it is possible to find models $\hat{A} \neq A(C)$ producing efficient behaviour.

A full control algorithm capture responses to changes in task and adaptation can be described as follows [50]:

1- Estimate Θ_t, where the time index was added to highlight time varying parameters,
2- Calculate $L_{t:N}(\Theta_t)$, i.e. the feedback gains with the current cost-function ($Q_{t:N}$ of the current task, and R), internal models ($\hat{A}$ and $\hat{B}$), and time horizon (N); use robust control if perturbations are unpredictable,
3- Apply $u_t = L_1(\Theta_t)\hat{x}_t$, $u_{t+1} = L_2(\Theta_t)\hat{x}_{t+1}$, etc., if there is no change in the parameters, otherwise Step 1.

Figures

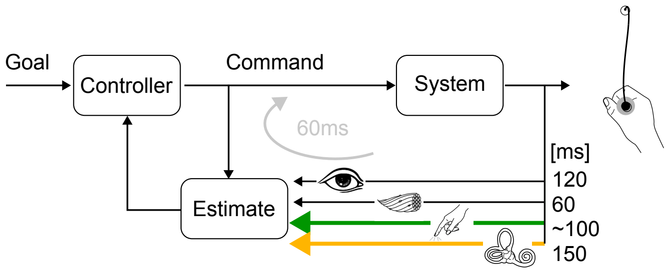

Figure 1. *Schematic representation of the components of an optimal feedback controller: the controller transforms the current estimated state in to a motor command following a feedback control policy that minimizes the expected value of movement cost. Flexible control is mediated by different sensory modalities, and their respective latencies is reported on the figure. The shortest latency (~50-60ms for muscle afferent feedback) must be considered in instances of SOC used as building block. Colored arrows highlight the contribution of sensory modalities to flexible reaching control that were not previously emphasized.*

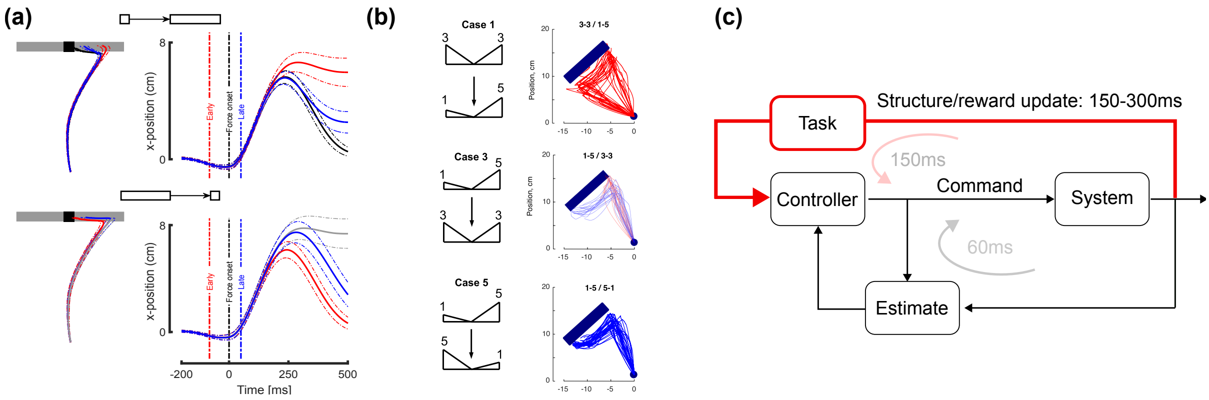

***Figure 2. (a)** Online changes in control following target switches. The goal target switched from a rectangle to a square or vice versa. The switches occurred early (red) or late (blue) during movement and a mechanical perturbation was applied to evaluate whether the controller considered changes in the task redundancy occurring online. Hand traces (left), temporal evolution of lateral coordinate (right). Black and gray traces on top and bottom traces are trajectories without switches. Adapted with permission from de Comite et al., 2021. **(b)** Rapid changes in movement goal following updates in reward distribution. Reward followed a triangular function with minimum at the center and linear increases such that the extreme location was more rewarding. The distribution of visual reward was presented throughout movement and changed as indicated in the left column. Hand traces started from the initial location (black dot), then reached the location of the bar target revealing participant's intended rewards. The color code corresponds to the final location. The hand traces show that when the reward distribution changed, participants rerouted their movements to the more rewarding end. Adapted from Marti-Marca et al., 2020. **(c)** Computational representation of dynamic updates in movement goal, featuring a feedback loop that monitor changes in task parameters online with latency ≥150 ms. The basic building block consisting in a SOC control model is in black, and the addition to this model is the red arrow that updates task parameters online.*

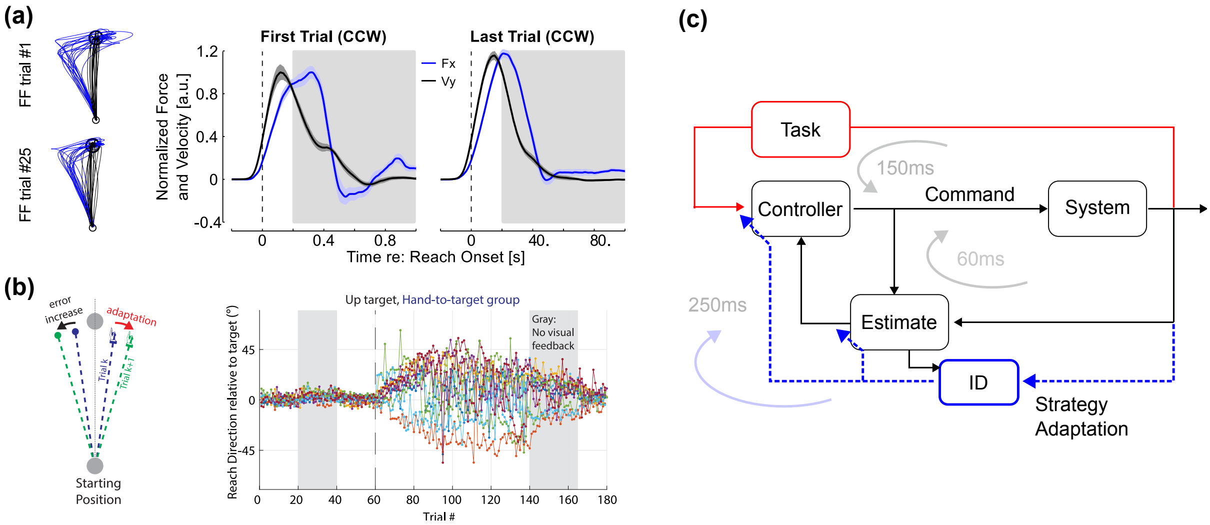

Figure 3. (a) *Changes in feedback responses to unexpected force fields. Individual traces are shown from each participant. Black traces are baseline trials, blue traces are force field trials, thus the first force field trial (top) and force field trial #25 are shown for each participant. The correction to the 25th force field disturbance displayed characteristics of adapted movement shown in the right panels: the correlation between commanded force proportional to forward velocity (black, V_y) and the applied forces (blue, F_x) increased. Adapted from Crevecoeur et al., 2020b.* ***(b)*** *Direct policy updates in a mirror reversal task (left). A linear mapping of the error into an update in movement trajectory in the case of a mirror reversal would produce expansion of the error. Right: experiment data showing the expansion of the error across movements. Displays are the reach direction relative to the target as a function of trial number. Adapted from Hadjiosif et al., (2021)* ***(c)*** *Addition to the model consist in estimating the parameters characterizing movement dynamics, and updating the control strategy with robust and adaptive mechanisms. Feedback adaptation and the expansion phenomenon set constraint on the learning rules used in short-term adaptation. The ID box illustrates the identification of model parameters used to update the controller. The delay of 250ms refers to the moment when correction becomes tuned to the force field (from {Crevecoeur, 2020b}). Again, the basic building block is in black, and the updates to the model are in colored arrows. This section emphasizes updates in model parameters and control strategies highlighted in blue.*

References

1. Todorov E, Jordan MI: **Optimal feedback control as a theory of motor coordination**. *Nature Neuroscience* 2002, **5**:1226–1235.

2. Scott SH: **A Functional Taxonomy of Bottom-Up Sensory Feedback Processing for Motor Actions**. *Trends Neurosci* 2016, **39**:512–526.

3. Franklin DW, Wolpert DM: **Specificity of reflex adaptation for task-relevant variability**. *J Neurosci* 2008, **28**:14165–14175.

4. Cross KP, Cluff T, Takei T, Scott SH: **Visual Feedback Processing of the Limb Involves Two Distinct Phases**. *J Neurosci* 2019, **39**:6751–6765.

5. Moreau-Debord I, Martin CZ, Landry M, Green AM: **Evidence for a reference frame transformation of vestibular signal contributions to voluntary reaching**. *Journal of Neurophysiology* 2014, **111**:1903–1919.

6. Keyser J, Medendorp WP, Selen LPJ: **Task-dependent vestibular feedback responses in reaching**. *J Neurophysiol* 2017, **118**:84–92.

7. Pruszynski JA, Johansson RS, Flanagan JR: **A Rapid Tactile-Motor Reflex Automatically Guides Reaching toward Handheld Objects**. *Curr Biol* 2016, **26**:788–792.

8. Crevecoeur F, Barrea A, Libouton X, Thonnard J-L, Lefèvre P: **Multisensory components of rapid motor responses to fingertip loading**. *Journal of Neurophysiology* 2017, **118**:331–343.

9. Reschechtko S, Pruszynski JA: **Voluntary modification of rapid tactile-motor responses during reaching differs from its visuomotor counterpart**. *Journal of Neurophysiology* 2020, **124**:284–294.

10. Forgaard CJ, Reschechtko S, Gribble PL, Pruszynski JA: **Skin and muscle receptors shape coordinated fast feedback responses in the upper limb**. *Current Opinion in Physiology* 2021, **20**:198–205.

11. Cluff T, Scott SH: **Apparent and Actual Trajectory Control Depend on the Behavioral Context in Upper Limb Motor Tasks**. *J Neurosci* 2015, **35**:12465–12476.

12. Česonis J, Franklin DW: **Contextual cues are not unique for motor learning: Task-dependant switching of feedback controllers**. *PLOS Computational Biology* 2022, **18**:e1010192.

* This study documents context dependent changes in reaching control by varying the instructions from trial to trial and showing that participants switched controllers accordingly.

13. Liu D, Todorov E: **Evidence for the flexible sensorimotor strategies predicted by optimal feedback control**. *J Neurosci* 2007, **27**:9354–9368.

14. De Comite A, Crevecoeur F, Lefèvre P: **Continuous Tracking of Task Parameters Tunes Reaching Control Online**. *eNeuro* 2022, **9**:ENEURO.0055-22.2022.

15. De Comite A, Crevecoeur F, Lefèvre P: **Online modification of goal-directed control in human reaching movements**. *Journal of Neurophysiology* 2021, **125**:1883–1898.

16. Codol O, Kashefi M, Forgaard CJ, Galea JM, Pruszynski JA, Gribble PL: **Sensorimotor feedback loops are selectively sensitive to reward**. *eLife* 2023, **12**:e81325.

** The authors provide a thorough investigation of the effect of rewards in seven specific and well defined sensorimotor loops. Only the earliest responses were insensitive to reward, and the subsequent ones exhibited an increase in feedback gains.

17. Cos I, Pezzulo G, Cisek P: **Changes of Mind after movement onset depend on the state of the motor system**. *Eneuro* 2021, **8**.

18. Marti-Marca A, Deco G, Cos I: **Visual-reward driven changes of movement during action execution**. *Scientific reports* 2020, **10**:1–12.

* This study demonstrated that changes in reward distribution along a redundant axis evoked online rerouting of an ongoing movement towards more rewarding state. The latency of the update was ~300ms as measured from hand kinematics.

19. De Comite A, Crevecoeur F, Lefèvre P: **Reward-Dependent Selection of Feedback Gains Impacts Rapid Motor Decisions**. *eNeuro* 2022, **9**:ENEURO.0439-21.2022.

20. Cisek P: **Making decisions through a distributed consensus**. *Current opinion in neurobiology* 2012, **22**:927–936.

21. Comite AD, Lefèvre P, Crevecoeur F: **Continuous evaluation of cost-to-go for flexible reaching control and online decisions**. *PLOS Computational Biology* 2023, **19**:e1011493.

22. Krakauer JW, Hadjiosif AM, Xu J, Wong AL, Haith AM: **Motor Learning**. *Compr Physiol* 2019, **9**:613–663.

23. Babadi S, Vahdat S, Milner TE: **Neural Substrates of Muscle Co-contraction during Dynamic Motor Adaptation**. *J Neurosci* 2021, doi:10.1523/JNEUROSCI.2924-19.2021.

24. Franklin DW, Burdet E, Peng Tee K, Osu R, Chew C-M, Milner TE, Kawato M: **CNS Learns Stable, Accurate, and Efficient Movements Using a Simple Algorithm**. *J Neurosci* 2008, **28**:11165–11173.

25. Franklin S, Wolpert DM, Franklin DW: **Visuomotor feedback gains upregulate during the learning of novel dynamics**. *J Neurophysiol* 2012, **108**:467–478.

26. Dimitriou M: **Task-dependent modulation of spinal and transcortical stretch reflexes linked to motor learning rate**. *Behavioral Neuroscience* 2018, **132**:194–209.

27. Crevecoeur F, Scott SH, Cluff T: **Robust Control in Human Reaching Movements: A Model-Free Strategy to Compensate for Unpredictable Disturbances**. *J Neurosci* 2019, **39**:8135–8148.

28. Coltman SK, Gribble PL: **Time course of changes in the long latency feedback response parallels the fast process of short term motor adaptation**. *Journal of Neurophysiology* 2020, doi:10.1152/jn.00286.2020.

29. Bian T, Wolpert D, Jiang Z-P: **Model-Free Robust Optimal Feedback Mechanisms of Biological Motor Control**. *Neural Computation* 2020, **32**:1–34.

* This study provides an algorithm to synthesize optimal feedback controllers with inexact models of the limb dynamics based on sensory data. It further formalizes the principles of robust control and develops the correspondence with reinforcement learning (see also reference 31).

30. Pang B, Cui L, Jiang Z-P: **Human motor learning is robust to control-dependent noise**. *Biol Cybern* 2022, **116**:307–325.

** The study documents based on mathematical developments and simulations of bio-inspired reaching movements that it is possible to construct efficient controllers that are not necessarily based on an exact model of limb dynamics. The conditions are documented to ensure that the algorithm iteratively produces controllers that are close to the optimal set of gains, and converge to the optimal one asymptotically.

31. Wagner MJ, Smith MA: **Shared internal models for feedforward and feedback control**. *J Neurosci* 2008, **28**:10663–10673.

32. Cluff T, Scott SH: **Rapid feedback responses correlate with reach adaptation and properties of novel upper limb loads**. *J Neurosci* 2013, **33**:15903–15914.

33. Ahmadi-Pajouh MA, Towhidkhah F, Shadmehr R: **Preparing to reach: selecting an adaptive long-latency feedback controller**. *J Neurosci* 2012, **32**:9537–9545.

34. Maeda RS, Cluff T, Gribble PL, Pruszynski JA: **Feedforward and Feedback Control Share an Internal Model of the Arm's Dynamics**. *J Neurosci* 2018, **38**:10505–10514.

35. Maeda RS, Gribble PL, Pruszynski JA: **Learning New Feedforward Motor Commands Based on Feedback Responses**. *Current Biology* 2020, **30**:1941-1948.e3.

36. Albert ST, Shadmehr R: **The Neural Feedback Response to Error As a Teaching Signal for the Motor Learning System**. *J Neurosci* 2016, **36**:4832–4845.

37. Joiner WM, Sing GC, Smith MA: **Temporal specificity of the initial adaptive response in motor adaptation**. *PLoS Comput Biol* 2017, **13**:e1005438.

38. Crevecoeur F, Thonnard J-L, Lefèvre P: **A Very Fast Time Scale of Human Motor Adaptation: Within Movement Adjustments of Internal Representations during Reaching**. *eNeuro* 2020, **7**.

** This study highlights that online feedback responses to unexpected force field disturbances evoke adaptation in the feedback pathway that can be explained in the context of adaptive control. A second experiment shows that unexpected disturbances prior to a via-point evokes after effects after the via-point, indicating that adaptation occurred during the ongoing sequence of actions.

39. Crevecoeur F, Mathew J, Bastin M, Lefèvre P: **Feedback Adaptation to Unpredictable Force Fields in 250 ms**. *eNeuro* 2020, **7**.

40. Braun DA, Aertsen A, Wolpert DM, Mehring C: **Learning optimal adaptation strategies in unpredictable motor tasks**. *J Neurosci* 2009, **29**:6472–6478.

41. Hadjiosif AM, Krakauer JW, Haith AM: **Did We Get Sensorimotor Adaptation Wrong? Implicit Adaptation as Direct Policy Updating Rather than Forward-Model-Based Learning**. *J Neurosci* 2021, **41**:2747–2761.

** The study shows that short term adaptation to a visual perturbation consisting in mirror reversal of the cursor motion relative to the axis aligned with movement direction is governed by direct policy update, which produces an expansion of the error across trials

42. Bin M, Huang J, Isidori A, Marconi L, Mischiati M, Sontag E: **Internal Models in Control, Bioengineering, and Neuroscience**. *Annu Rev Control Robot Auton Syst* 2022, **5**:55–79.

43. Bitmead RR, Gevers M, Wertz V: **Adaptive Optimal Control The Thinking Man's GPC**. 1990,

44. Dimitriou M, Wolpert DM, Franklin DW: **The Temporal Evolution of Feedback Gains Rapidly Update to Task Demands**. *J Neurosci* 2013, **33**:10898–10909.

45. Ikegami T, Ganesh G, Gibo TL, Yoshioka T, Osu R, Kawato M: **Hierarchical motor adaptations negotiate failures during force field learning**. *PLOS Computational Biology* 2021, **17**:e1008481.

** Towards better characterization of the learning or adaptation rules used in the brain, it is shown that different criteria linked to task success or error in dynamics organize adaptation in an hierarchical manner.

46. Crevecoeur F, Mathew J, Lefèvre P: **Separability of Human Motor Memories during reaching adaptation with force cues**. *PLOS Computational Biology* 2022, **18**:e1009966.

* This study highlights trial-by-trial and online adaptation in a dual-adaptation task based on adaptive control.

47. Heald JB, Lengyel M, Wolpert DM: **Contextual inference underlies the learning of sensorimotor repertoires**. *Nature* 2021, **600**:489–493.

48. Berret B, Baud-Bovy G: **Evidence for a cost of time in the invigoration of isometric reaching movements**. *Journal of Neurophysiology* 2022, **127**:689–701.

49. Shadmehr R, Huang HJ, Ahmed AA: **A Representation of Effort in Decision-Making and Motor Control**. *Current Biology* 2016, **26**:1929–1934.

50. De Nicolao G, Magni L, Scattolini R: **Stability and Robustness of Nonlinear Receding Horizon Control**. In *Nonlinear Model Predictive Control*. Edited by Allgöwer F, Zheng A. Birkhäuser; 2000:3–22.