

Evidence for the Flexible Sensorimotor Strategies Predicted by Optimal Feedback Control

Dan Liu and Emanuel Todorov

Department of Cognitive Science, University of California, San Diego, La Jolla, California 92093-0515

Everyday movements pursue diverse and often conflicting mixtures of task goals, requiring sensorimotor strategies customized for the task at hand. Such customization is mostly ignored by traditional theories emphasizing movement geometry and servo control. In contrast, the relationship between the task and the strategy most suitable for accomplishing it lies at the core of our optimal feedback control theory of coordination. Here, we show that the predicted sensitivity to task goals affords natural explanations to a number of novel psychophysical findings. Our point of departure is the little-known fact that corrections for target perturbations introduced late in a reaching movement are incomplete. We show that this is not simply attributable to lack of time, in contradiction with alternative models and, somewhat paradoxically, in agreement with our model. Analysis of optimal feedback gains reveals that the effect is partly attributable to a previously unknown trade-off between stability and accuracy. This yields a testable prediction: if stability requirements are decreased, then accuracy should increase. We confirm the prediction experimentally in three-dimensional obstacle avoidance and interception tasks in which subjects hit a robotic target with programmable impedance. In additional agreement with the theory, we find that subjects do not rely on rigid control strategies but instead exploit every opportunity for increased performance. The modeling methodology needed to capture this extra flexibility is more general than the linear-quadratic methods we used previously. The results suggest that the remarkable flexibility of motor behavior arises from sensorimotor control laws optimized for composite cost functions.

Key words: arm movement; optimal feedback control; composite cost; obstacle avoidance; stability–accuracy trade-off; behavioral flexibility

Introduction

Humans interact with a diverse and uncertain environment requiring flexible motor behavior. Here, we show that the optimal feedback control theory that we (Todorov and Jordan, 2002a,b; Todorov, 2004, 2005; Todorov et al., 2005) and others (Meyer et al., 1988; Loeb et al., 1990; Hoff, 1992; Kuo, 1995; Scott, 2004) have pursued affords the flexibility apparent in behavioral data, in contrast with more traditional theories (Flash and Hogan, 1985; Uno et al., 1989; Bizzi et al., 1992; Feldman and Levin, 1995). We distinguish two forms of flexibility. From the perspective of motor planning or preparation, flexibility entails taking into account multiple task requirements and properties of the environment known before movement, and preparing a sensorimotor strategy with both open-loop and closed-loop components customized for the present task and circumstances. From the perspective of motor execution, a strategy is flexible if its closed-loop component makes on-line adjustments that exploit the multiple ways in which a redundant musculoskeletal plant can achieve the same behavioral goal. Both forms of flexibility are

obvious desiderata for a well designed estimation-and-control system such as the sensorimotor system.

Previous work has emphasized the evidence for flexibility during execution, in particular the structure of motor variability (which is larger in task-irrelevant dimensions) and the goal-directed nature of on-line corrections (Bernstein, 1967; Cole and Abbs, 1987; Scholz and Schoner, 1999; Domkin et al., 2002; Todorov and Jordan, 2002b). We explained such phenomena with the minimal intervention principle, which states that task-irrelevant deviations from the average behavior should be left uncorrected to maximize performance (Todorov and Jordan, 2002a,b; Todorov, 2004). This argument is further advanced here by showing that (1) target displacement can cause correction before the hand has cleared an intermediate obstacle, ruling out the imaginary via points postulated by alternative models, and (2) end-point variability matches the shape of an elongated target and feedback corrections in the redundant dimension are suppressed, as the minimal intervention principle predicts. Apart from these findings, however, our emphasis here is on flexibility in motor planning/preparation.

With the exception of speed–accuracy trade-offs (Jeannerod, 1988), the systematic relationship between sensorimotor strategies and mixtures of task goals (as well as properties of the environment) has received surprisingly little attention. Optimal control models, which dominate the thinking on trajectory planning, have traditionally optimized a homogeneous cost and treated all other goals as hard constraints; the latter are supposed to be specified externally, outside the scope of such models. The ho-

Received March 15, 2006; revised June 4, 2007; accepted July 9, 2007.

This work was supported by National Institutes of Health Grant NS-045915 and National Science Foundation Grants ECS-0524761 and SBE-0542013. We thank Javier Movellan and Howard Poizner for comments on this manuscript.

Correspondence should be addressed to Emanuel Todorov at the above address. E-mail: todorov@cogsci.ucsd.edu.

DOI:10.1523/JNEUROSCI.1110-06.2007

Copyright © 2007 Society for Neuroscience 0270-6474/07/279354-15\$15.00/0

mogeneous cost could be energy consumption (Nelson, 1983; Anderson and Pandy, 2001), derivative of hand acceleration (Flash and Hogan, 1985), derivative of joint torque (Uno et al., 1989), endpoint variance (Harris and Wolpert, 1998). The constraints include endpoint position, final velocity and acceleration (typically zero), movement time, intermediate points along the trajectory. However, these hypothetical constraints are rarely explicit in real-world tasks, raising two questions: (1) how are their values being chosen; (2) are their values “chosen” in the first place, or are they stochastic outcomes of the complex interactions among sensorimotor strategy, noise, musculoskeletal dynamics, and environment, like any other feature of individual movements? Our previous analysis (Todorov and Jordan, 2002a,b) showed that choosing desired values for movement parameters that are not explicitly specified by the task is suboptimal, no matter how the choice is made. This answers question 2 and renders question 1 irrelevant. Instead of satisfying self-imposed constraints, we propose that the CNS relies on sensorimotor strategies optimized for composite cost functions. In the present experiments, the relevant cost components encourage energetic efficiency, endpoint positional accuracy (measured as bias and variance), endpoint stability (defined as bringing the movement to a complete stop), and movement speed (avoiding the time-out errors incurred when duration exceeds a threshold). We show that, as the relative importance of these components is varied by the experimenter, subjects modify their strategy in agreement with our theory. As in previous stochastic optimal control models (Harris and Wolpert, 1998; Todorov, 2002; Todorov and Jordan, 2002b), taking into account the empirically established signal-dependent nature of motor noise (Sutton and Sykes, 1967; Schmidt et al., 1979; Todorov, 2002; Hamilton et al., 2004) turns out to be important.

Although our focus is on effects taking place before movement, the effects in question correspond to changes in a control strategy with both open-loop and closed-loop components, which in turn is best studied using perturbations. Perturbing the target of a reaching movement in an unpredictable direction has been a productive paradigm for investigating the mechanisms of on-line visuomotor corrections (Pelisson et al., 1986; Prablanc and Martin, 1992; Desmurget and Grafton, 2000). Most previous studies have introduced perturbations around the time of movement onset, and found that the hand path is smoothly corrected to reach the displaced target, in agreement with multiple models of motor control (Hinton, 1984; Flash and Henis, 1991; Hoff and Arbib, 1993; Torres and Zipser, 2002). However, perturbations introduced late in the movement may be more informative because they are not fully corrected, in contradiction with alternative models and, somewhat paradoxically, in agreement with optimal feedback control. Such phenomena have been observed with both visual target perturbations (Komilis et al., 1993) and mechanical limb perturbations (Popescu and Rymer, 2000). In the case of limb perturbations, the correction reflects both neural feedback and musculoskeletal impedance, which are seamlessly integrated (Nicols and Houk, 1976) and difficult to disentangle. Therefore, we focus on target perturbations. We first design a two-dimensional reaching experiment to rule out the trivial explanation that the incomplete correction is simply attributable to lack of time. We then replicate the phenomenon in our model, and find that it reflects a previously unknown trade-off between endpoint accuracy and stability. This yields a novel prediction: if stability requirements are decreased, then accuracy should increase. The prediction is confirmed in three-dimensional obstacle avoidance and interception experiments. The latter experi-

ments give rise to rich motor behavior, allowing us to make a number of additional observations consistent with our theory. These include shaping variability patterns to buffer noise in redundant dimensions, adjusting movement duration to take advantage of temporal error margins, exploiting target impedance and surface friction to achieve endpoint stability, reallocating corrective action among redundant actuators to balance signal-dependent noise and inertial constraints, and correcting for task-relevant perturbations before having reached the task-irrelevant subgoals hypothesized by alternative models.

Materials and Methods

Experiment 1. Seven subjects made planar reaching movements on a table positioned at chest level. A 21 inch flatscreen monitor was mounted above the table facing down and was viewed in a see-through horizontal mirror. In this way, computer-generated images could be physically aligned with the hand workspace. Movement kinematics was recorded with an Optotrak 3020 infrared sensor at 100 Hz. A small pointer, which had an Optotrak marker and a light-emitting diode (LED) attached near its tip, was held in the dominant right hand. The task was to move the LED to a starting position, wait for a target to appear, and move to the target when ready. Movement onset was detected on-line using a 1 cm threshold on the distance between the pointer and the starting position. Analysis of speed profiles (see Fig. 1*b*) revealed that the actual movement started ~100 ms before the distance threshold was reached. Therefore we define the origin of the time axis to be 100 ms before the on-line detection of movement onset and report all times relative to this corrected origin.

The end of the movement was defined as the first point in time when the hand speed had remained <0.5 cm/s for 40 ms. The LED was turned off at movement onset, turned on at the end of the movement, and remained on in the repositioning phase. The room was dark. The target was always visible. Thus, movements were made without visual feedback of the hand, although subjects could see their endpoint error as soon as the movement ended. Movement duration was required to be between 600 and 800 ms. If the duration on any trial fell outside these boundaries, the computer displayed a “slow down” or “speed up” message, respectively. Movement amplitude was 30 cm. The main movement was in the lateral direction from right to left (although in Figs. 1*a* and 2*a*, we plot the movements from left to right, for consistency with the space–time plots).

After a brief familiarization session, every subject performed 240 trials. Within each trial, the target could either remain stationary or jump 5 cm forward or backward, orthogonal to the main movement direction. Subjects were instructed that jumps may occur, and asked to always move to the final target position and stop there within the allowed time interval. The jumps occurred at 100, 200, or 300 ms. There were 180 perturbed trials (30 for every possible latency–direction combination) and 60 non-perturbed trials, presented in random permutation order.

Experiment 2. Eight subjects made three-dimensional arm movements around a horizontal obstacle while aiming for the center of a vertical target (see Fig. 4*a*). Movement dimensions are illustrated in Figure 4*b*. The target was a 20 × 5 cm wooden board with a bull’s-eye pattern, and was mounted on a 3DOF robot (Delta Haptic Device; Force Dimension, Lausanne, Switzerland). Subjects held in their right hand a 7 cm wooden pointer with an electromagnetic Polhemus (Colchester, VT) Liberty sensor attached to it. The sensor measured three-dimensional position and orientation (at 240 Hz), making it possible to compute the position of the tip of the pointer. The latter is referred to as the “hand.” Another sensor was attached to the target (which was also tracked via the encoders of the robot at 1000 Hz). Before each trial, the robot moved the target away and waited for the subject to initiate the trial, by inserting the tip of the pointer in a small receptacle mounted below the obstacle and remaining stationary for 100 ms. Then, the robot “presented” the target by moving toward the subject, at which time the subject was free to move when ready. In case of an anticipation error, the computer played a sound and aborted the trial. Hand movement onset was detected with a 1 cm position threshold on the distance from the starting position. Subsequent analysis of speed profiles revealed that the movement started ~50 ms

before it was detected. Therefore, we define the origin of the time axis to be 50 ms earlier.

Movement end was detected when the hand speed remained <10 cm/s for 40 ms, or when the target was displaced because of impact with the hand by >0.4 cm. The maximum allowed movement duration was 900 ms. Time-out errors were signaled by moving the target away and playing a loud sound. During the hand movement, the target was either stationary or rapidly displaced by the robot 9 cm left or right. The robot trajectory is illustrated in Figure 4, *c* and *f*; it was generated by a model-based controller with both open-loop and closed-loop components. The target jump could be initiated at 50 ms (early) or 350 ms (late).

After brief familiarization, subjects performed 20 trials without perturbations, followed by two experimental conditions/sessions with perturbations, 120 trials each. Each experimental session included 40 early jumps, 40 late jumps, and 40 baseline trials presented in random permutation order. Left and right jumps were equally probable. Subjects were instructed that jumps may occur and asked to always move to the final target position within the allowed time interval. The two experimental sessions differed in the stopping requirements of the task. In the stop condition, subjects were asked to slow down their hand movement before contact and touch the target gently, so that the impact would not displace the target by >0.4 cm. If it did, the computer played an explosive sound and the target was moved away rapidly. This indicated to the subject that the target has been hit too hard. In the hit condition, hitting the target hard was no longer an error. The robot was placed in high-gain servo mode (with carefully chosen nonlinearities to avoid instability) and was able to absorb the impact with the hand. The maximum force output of the robot is 25 N. Subjects were not explicitly asked to hit the target harder, but they quickly discovered the benefits of using such a strategy. One-half of the subjects performed the hit condition first and then the stop condition; the order was reversed for the other one-half of the subjects.

Experiment 3. Ten subjects performed a task similar to experiment 2, with the following modifications. A pressure sensor (FSR; Interlink Electronics, Camarillo, CA) was installed inside the starting position receptacle and used to detect movement onset earlier and more reliably. An ATI Mini-40 six-axis force-torque sensor (2000 Hz sampling) was mounted behind the target. It allowed more reliable detection of contact (which was now defined as the movement end) as well as direct measurement of contact force. Subjects still initiated the movement when ready. As soon as hand movement was detected, the robot began to move downward at a constant speed (6.67 cm/s). This motion continued until the trial ended because of hand–target contact, or until the target hit the horizontal edge of a wooden board mounted underneath. This happened 900 ms after movement onset and was defined as a time-out error. The downward motion was repeatable and easily predictable, providing subjects with an explicit representation of allowed movement duration. When the target jumped (9 cm left or right), the rapid lateral motion was superimposed on the slow downward motion. Instead of the bull’s-eye pattern, the target now had a pattern of vertical stripes (5×1 cm each), with gray levels increasing with lateral distance from the center stripe (which was white). Subjects were asked to make contact with the target as close to the center stripe as possible. In the stop condition, the threshold for hitting too hard was now defined in terms of force rather than displacement (0.8 N in the first 8 ms after contact). The bookshelf obstacle from Figure 4*a* was now replaced with a horizontal bar, and moved backward to induce a more curved hand movement (dimensions are illustrated in Fig. 5*c*). Early jumps were triggered at movement onset; late jumps were triggered at 400 ms after movement onset; allowed movement duration was 900 ms. Every subject still participated in two conditions, hit and stop, in counterbalanced order. In each session, we scheduled 25 baseline trials, 25 early jumps, and 25 late jumps in random permutation order. Any failed trials (time-out errors or hitting-hard errors) were now rescheduled at a random time later in the same session, yielding a more balanced database of analyzable trials. Instead of 20 no-perturbation trials before the experiment, we now had 10 trials before and 10 trials after the experiment. These were used to compute variability in the absence of perturbations (see Fig. 8, “baseline”).

In experiments 2 and 3, the wrist was immobilized with an orthopedic brace to avoid corrective movements using the wrist. There were two

reasons for this restriction. First, pilot experiments revealed different involvement of the wrist in early versus late corrections (see Fig. 6*c*), making the comparison between conditions difficult. Second, our models assume point-mass dynamics and do not capture wrist movements.

Statistical analysis. In experiments 2 and 3, both the time-out errors and the hitting-hard errors were signaled immediately, in a way that disrupted the behavior, and therefore error trials could not be included in the analysis. Only no-error trials were analyzed. They constituted 79% of all trials in experiment 2 and 58% in experiment 3. The higher overall error rate in experiment 3 is because we repeated failed trials, and so subjects performed more trials in the more difficult conditions (late/stop in particular). In experiment 1, errors were signaled by the computer only after the movement had stopped. Thus, the error signals could not disrupt the behavior, allowing us to include trials whose duration was slightly over the time limit (up to 100 ms). Presumably these movements were generated by the same underlying mechanism and the longer duration was simply attributable to trial-to-trial variability. Ninety-three percent of all trials in experiment 1 were analyzed.

All statistical tests were based on *n*-factor ANOVA (“anovan” in the Matlab Statistics Toolbox). We avoided averaging to the extent possible. In the comparisons of undershoot and duration (see Figs. 1*d,g*, 4*d,e,g,h*) and wrist contribution (see Fig. 6*c*), individual trials were treated as repeated measures, and the factors were the experimental conditions (perturbation time, stop vs hit when applicable) as well as the subject identity. Thus, we had two factors in experiment 1 and three factors in experiments 2 and 3 and the pilot experiment in Figure 6*c*. The subject identity was modeled as a factor with random effects because subjects are drawn randomly from the population. In the comparisons of SDs (see Figs. 3*e*, 6*a*), time-out error rates (see Fig. 1*h*), and lateral velocities (see Fig. 6*b*), all trials that a subject performed in a given condition were combined to obtain a single number. All comparisons of means were based on Tukey’s criterion for *post hoc* hypothesis testing. Differences are reported as significant when $p < 0.05$. The error bars shown in the figures correspond to ± 1 pooled SE of the mean, as computed by the multiple-comparison function used to perform Tukey’s test (“multcompare” in the Matlab Statistics Toolbox).

Optimal feedback control model (linear-quadratic-Gaussian). We model the hand as an $m = 1$ kg point mass moving in a horizontal plane, with viscosity $b = 10$ Ns/m approximating intrinsic muscle damping. The point mass is driven by two orthogonal force actuators that can both pull and push (approximating two pairs of agonist–antagonist muscles). The actuators act as muscle-like first-order low-pass filters of the control signals, with time constant $\tau = 0.05$ s. These settings of m , b , and τ were chosen to be compatible with biomechanics and were not adjusted to fit the data.

Let $\mathbf{p}(t)$, $\mathbf{v}(t)$, $\mathbf{a}(t)$, $\mathbf{u}(t)$ be the two-dimensional hand position, velocity, actuator state, and control signal, respectively. The corresponding units are meters, meters/second, newtons, and newtons. The time index is $t \in [0, t_f]$. The final time t_f is specified (taken from the experimental data in Fig. 1*g*). The plant dynamics in continuous time are modeled as follows:

$$d\mathbf{p}(t) = \mathbf{v}(t)dt$$

$$m d\mathbf{v}(t) = (\mathbf{a}(t) - b\mathbf{v}(t))dt$$

$$\tau d\mathbf{a}(t) = (\mathbf{u}(t) - \mathbf{a}(t))dt + M(\mathbf{u}(t))d\mathbf{w}(t).$$

$\mathbf{w}(t)$ is standard Brownian motion. $M(\mathbf{u}(t))$ represents control-multiplicative or signal-dependent motor noise, and is given by the following:

$$M(\mathbf{u}(t)) = \begin{bmatrix} c_1 u_1(t) & c_2 u_1(t) \\ -c_2 u_2(t) & c_1 u_2(t) \end{bmatrix}.$$

$c_1 = 0.15$ corresponds to two-dimensional noise in the same direction as the control vector $\mathbf{u}(t)$, whereas $c_2 = 0.05$ corresponds to two-dimensional noise in the direction orthogonal $\mathbf{u}(t)$. The parallel noise component is larger because muscles pulling in the direction of net muscle force are more active and therefore more affected by signal-dependent noise. The parameters c_1 , c_2 , as well as the sensory noise magnitude σ

described below, are adjusted so that the baseline variability predicted by the model (see Fig. 2e) is similar to the experimental data (see Fig. 1e). Note that the shape of these curves cannot be fully captured by three scalar parameters, and so the good fit mostly reflects the quality of the model. Denoting the target position $\mathbf{p}^*(t)$, we can assemble all variables into an eight-dimensional state vector as follows:

$$\mathbf{x}(t) = [\mathbf{p}(t); \mathbf{v}(t); \mathbf{a}(t); \mathbf{p}^*(t)],$$

and write its dynamics in general first-order form as follows:

$$d\mathbf{x}(t) = (A\mathbf{x}(t) + B\mathbf{u}(t))dt + C(\mathbf{u}(t))d\mathbf{w}(t),$$

with A , B , and C obtained from the above equations.

To define an optimal control problem, we also need a cost function. As in our previous work (Todorov and Jordan, 2002b), we use a mixed cost function defined as follows:

$$\|\mathbf{p}^*(t_f) - \mathbf{p}(t_f)\|^2 + w_{\text{stop}}(\|\mathbf{v}(t_f)\|^2 + \|s_a \mathbf{a}(t_f)\|^2) + w_{\text{energy}} \int_0^{t_f} \|\mathbf{u}(t)\|^2 dt.$$

The three cost terms encourage endpoint positional accuracy, stopping at the target, and energetic efficiency, respectively. The activations $\mathbf{a}(t)$ are scaled by $s_a = 0.1$ because their numerical values turn out to be an order of magnitude larger than positions and velocities. The weight $w_{\text{energy}} = 0.00005$ of the control term is a free parameter (it is not clear how to estimate this parameter independently). The weight w_{stop} determines the relative importance of coming to a complete stop (i.e., achieving zero velocity and acceleration) at the end of the movement. We use $w_{\text{stop}} = 1$ for the stop condition and $w_{\text{stop}} = 0.01$ for the hit condition. These values are chosen to capture the qualitative differences between the stop and hit conditions. Note that we do not model the three-dimensional experiments explicitly.

The state of the plant $\mathbf{x}(t)$ is not directly observable but has to be inferred from noisy observations whose time integral $\mathbf{y}(t)$ satisfies the following:

$$d\mathbf{y}(t) = \mathbf{x}(t)dt + Gd\mathbf{z}(t).$$

The sensory noise covariance is diagonal: $G = \sigma \text{diag}(1, 1, 1, 1, 1/s_a, 1/s_a, 0, 0)$ with $\sigma = 0.015$ adjusted to reproduce the observed movement variability. $\mathbf{z}(t)$ is standard Brownian motion. In comparison with our previous model (Todorov and Jordan, 2002b), the present model is simpler in that here we use first-order rather than second-order muscle filters and do not explicitly represent sensory delays. We do not model explicitly the visuomotor delays or uncertainty in detecting the target observations. To obtain correct reaction times, we simply model each target perturbation as occurring 120 ms later than the corresponding experimental perturbation.

With these definitions, we discretize the time axis (with 1 ms time step) and obtain a discrete-time linear-quadratic-Gaussian (LQG) optimal control problem. The reason for formulating the model in continuous time and then discretizing the time axis, as opposed to working in discrete time all along, is that in a discrete-time formulation the model parameters are affected by the time step. If one were to change the time step, it would not be obvious how the model parameters should scale. Continuous-time formulations have the advantage of being independent of discretization time steps. For details on how to discretize a continuous-time system, see Li and Todorov (2007).

The presence of signal-dependent noise complicates matters; however, we derived an efficient algorithm for solving such problems previously (Todorov, 2005). That algorithm is applied here to yield a modified Kalman filter for computing the optimal state estimate, $\hat{\mathbf{x}}(t)$, and an optimal feedback controller of the following form:

$$\mathbf{u}(t) = -L(t)\hat{\mathbf{x}}(t).$$

Once the filter and controller are available, the state is initialized with the experimentally defined starting position and $\mathbf{v}(0) = \mathbf{a}(0) = 0$, and the system is simulated until the final time t_f .

The time-varying matrix of feedback gains $L(t)$ is 2×8 and is in

Table 1. Parameters of linear-quadratic-Gaussian model

Parameters	Criteria
$m = 1 \text{ kg}, b = 10 \text{ Ns/m}, \tau = 0.05 \text{ s}$	Compatibility with biomechanics
$s_a = 0.1$	Order-of-magnitude normalization
$c_1 = 0.15, c_2 = 0.05, \sigma = 0.015$	Fit to variability of unperturbed trials
$w_{\text{energy}} = 0.00005$	Overall fit to data
$w_{\text{stop}} = 1 \rightarrow 0.01$	Qualitative predictions regarding stop versus hit

principle described by 16 numbers. However, in the present problem, it turns out to have a lot of structure that can be captured by only three independent parameters. In particular, the control law can be written as follows:

$$\mathbf{u}(t) = k_p(t)(\mathbf{p}^*(t) - \hat{\mathbf{p}}(t)) - k_v(t)\hat{\mathbf{v}}(t) - k_a(t)\hat{\mathbf{a}}(t),$$

where k_p , k_v , and k_a are time-varying scalar gains illustrated in Figure 3, a and c .

When a perturbation is introduced in the model, the final time t_f is adjusted according to Figure 1g, and the optimal estimator and controller for the remainder of the movement are recomputed (given the new final time and target position). This is necessary because the optimal feedback gains are originally scheduled up to the duration of the unperturbed movement. Consequently, the predicted feedback corrections are not generated by exactly the same feedback gains as shown in Figure 3, a and c .

However, the change attributable to the recomputation is small, and so our intuitive analysis of feedback gains is valid. Similar recomputation is involved in the minimum-jerk feedback controller. A more general optimal feedback control model capable of predicting the changes in movement duration is described later.

The model parameters, and the criteria for choosing their values, can be summarized as shown in Table 1.

Modified minimum-jerk model. The original minimum-jerk model (Flash and Hogan, 1985) postulates that the hand moves from a starting position \mathbf{p}_0 to a target position \mathbf{p}^* along a trajectory that minimizes the time integral of the squared jerk (third derivative of position) as follows:

$$\min_{\mathbf{p}(\cdot)} \int_0^{t_f} \|\ddot{\mathbf{p}}(t)\|^2 dt.$$

To make this optimization problem well posed, one has to specify the velocity and acceleration at the endpoints. Let \mathbf{v}_0 and \mathbf{a}_0 be the initial velocity and acceleration (possibly nonzero) and suppose the final velocity and acceleration are 0. Then the constraints are as follows:

$$\mathbf{p}(0) = \mathbf{p}_0, \dot{\mathbf{p}}(0) = \mathbf{v}_0, \ddot{\mathbf{p}}(0) = \mathbf{a}_0, \mathbf{p}(t_f) = \mathbf{p}^*, \dot{\mathbf{p}}(t_f) = 0, \ddot{\mathbf{p}}(t_f) = 0.$$

One can find the solution to this minimization problem using the calculus of variations (Flash and Hogan, 1985). The expression for the optimal trajectory $\mathbf{p}(t)$ is a somewhat complicated function of t , t_f , \mathbf{p}_0 , \mathbf{v}_0 , \mathbf{a}_0 , \mathbf{p}^* . Differentiating that function with respect to t three times yields the following:

$$\ddot{\mathbf{p}}(t) = \frac{60}{(t_f - t)^3}(\mathbf{p}^* - \mathbf{p}_0) - \frac{36}{(t_f - t)^2}\mathbf{v}_0 - \frac{9}{(t_f - t)}\mathbf{a}_0,$$

where $t_f - t$ is the remaining movement time. The rationale for working with third derivatives is that we are given all derivatives up to second order at the initial time, and if we have a way of computing the third derivative at each time, we could simply integrate and obtain the entire trajectory. This suggests a feedback-control formulation (Hoff and Arbib, 1993) in which $\ddot{\mathbf{p}}$ is only computed at the current time and is treated as an instantaneous control signal \mathbf{u} as follows:

$$\mathbf{u}(t) = \ddot{\mathbf{p}}(t) = \frac{60}{(t_f - t)^3}(\mathbf{p}^*(t) - \mathbf{p}(t)) - \frac{36}{(t_f - t)^2}\dot{\mathbf{p}}(t) - \frac{9}{(t_f - t)}\ddot{\mathbf{p}}(t).$$

The three scalar coefficients in the above expression are time-varying feedback gains for a third-order system with state vector

$$\mathbf{x}(t) = [\mathbf{p}(t); \dot{\mathbf{p}}(t); \ddot{\mathbf{p}}(t)].$$

Note that the feedback-control formulation allows us to make the target position time varying.

In the absence of perturbations, the trajectory predicted by this modified minimum-jerk model is identical with the prediction of the original minimum-jerk model (Hoff and Arbib, 1993). The advantage of the modified formulation is that it can generate feedback corrections and thus serve as a model of perturbation experiments. Note that the modified minimum-jerk model reflects a very different philosophy compared with the original model, because a trajectory plan for the rest of the movement is no longer needed. In that sense, it is closer to our optimal feedback control model (Todorov and Jordan, 2002b). The empirical success of jerk minimization has often been interpreted as evidence for trajectory planning. Such interpretations are unjustified given that the same predictions can be made without the assumption of trajectory planning.

Optimal feedback control model (Markov decision process). Here, we describe a different optimal feedback control model in which movement duration is no longer predefined and task constraints are enforced more explicitly. It is constructed using the more general but less efficient methodology of Markov decision processes: the continuous state and action spaces are discretized (Kushner and Dupuis, 2001) and the resulting discrete optimization problem is solved via dynamic programming (Bertsekas, 2001). Discretization methods suffer from the curse of dimensionality and only apply to low-dimensional problems. This necessitates a simplification of the dynamics model: the arm is now modeled as a fully observable second-order plant with state vector containing hand position \mathbf{p} and velocity \mathbf{v} and control vector \mathbf{u} corresponding to hand acceleration. All quantities are expressed in units of centimeters and seconds. The initial state is $\mathbf{p}(0) = \mathbf{v}(0) = [0; 0]$. The default target position is $\mathbf{p}^* = [20; 0]$ but can be perturbed to either $[20; 5]$ or $[20; -5]$. Instead of perturbing the target, we perturb the hand in the opposite direction (without changing hand velocity) and then correct the hand and target positions in the subsequent analysis. In this way, the target can be treated as constant and omitted from the state vector.

Each trial ends when the horizontal hand position exceeds 20 cm (i.e., the hand reaches the target plane) or when the duration exceeds a maximum allowed duration of 0.6 s, whichever comes first. Let t_f denote the duration of a given trial. The total cost to be minimized is defined as follows:

$$\text{total_cost} = \text{final_cost} + w_{\text{energy}} \int_0^{t_f} \|\mathbf{u}(t)\|^2 dt.$$

The final cost, computed at the end of the movement, is defined as follows:

$$\text{final_cost} = \begin{cases} \|\mathbf{p}^* - \mathbf{p}(t_f)\|^2 + w_{\text{time}} t_f, & \text{if } t_f < 0.6 \text{ and } \|\mathbf{v}(t_f)\| < v_{\text{max}} \\ 100, & \text{otherwise} \end{cases}.$$

The endpoint velocity threshold v_{max} is 5 cm/s in the stop condition and 20 cm/s in the hit condition. These values were chosen to match the observed endpoint velocities. The main movement amplitude (20 cm), perturbation amplitude (5 cm), maximum movement duration (0.6 s), constraint violation cost (100), and other simulation parameters described below were chosen in advance and were not adjusted. The only parameters that were adjusted to fit the data were $w_{\text{energy}} = 0.00003$ and $w_{\text{time}} = 20$. This was done by solving the problem multiple times for different points in $w_{\text{energy}} - w_{\text{time}}$ space. The qualitative pattern of results shown in Figures 7 and 8 depended weakly on w_{time} but was sensitive to w_{energy} . The latter parameter, denoted r in other studies, has proven to be important in almost every optimal control model we ever constructed.

The sizes of the discretization grids were 101×61 points in position (20×12 cm; step size, 0.2 cm), 25×25 points in velocity (80×80 cm/s; step size, 3.33 cm/s), 11×11 points in acceleration (1666.67×1666.67

cm/s²; step size, 166.67 cm/s²), and 31 points in time (0.6 s; step size, 0.02 s). These numbers were carefully balanced so that the grid density was sufficient to allow accurate approximation, the grid range was sufficient to cover the optimal state-control trajectories, and yet the number of grid points was not intractably large. The noise was uniform and additive, and perturbed the hand velocity by up to ± 2 grid points in each time step.

The feedback control law is in the following general form:

$$\mathbf{u} = \pi(\mathbf{p}, \mathbf{v}, t),$$

where \mathbf{u} , \mathbf{p} , \mathbf{v} , and t are constrained to the corresponding grids. Unlike the LQG framework in which the function π is linear and can be represented with a small number of feedback gains, here we do not know in advance the form of π . Instead, we represent it as a lookup table that specifies the value of \mathbf{u} for every possible combination of \mathbf{p} , \mathbf{v} , and t . This table consists of ~ 220 million numbers computed by the dynamic programming algorithm in about one-half an hour of CPU time. To speed up the computation and be able to explore the effects of various parameters, we also implemented the algorithm on an nVidia GeForce 8800 GTX videocard with 128 parallel processors. This reduced the running time of the algorithm by about a factor of 30. Because the videocard only supports single-precision floating point arithmetic, we used it for model exploration and run the final model on an Intel CPU with double precision. Once the control laws for the stop and hit conditions were obtained (the only difference being the value of v_{max}), we applied them to the stochastic plant and simulated 3000 movement trajectories per condition: 1000 without perturbation, 1000 with perturbation at 0.1 s, and 1000 with perturbation at 0.3 s.

Although perturbations were applied in the testing phase to characterize the response of the optimal feedback control laws, the control laws themselves were optimized for an environment without perturbations. To study possible adaptation effects, we computed a second pair of feedback control laws optimized for a perturbed environment. In the latter environment, the target could jump either up or down (at 0.2 s) or remain stationary. The three types of trials had equal probability. Target perturbations were taken into account in the optimization process by incorporating an appropriate position noise term (with trimodal distribution) in the dynamics model.

Results

Undershoot in reaching to perturbed targets

We define “undershoot” as endpoint error in the direction in which the target was displaced. In reaching movements, undershoot (or incomplete correction) for late target perturbations has already been demonstrated (Komilis et al., 1993). However, a key question remains unanswered: is the effect simply attributable to lack of time, or is there a more subtle reason? Experiment 1 was designed to rule out the first possibility. Subjects made lateral reaching movements on a horizontal table, without vision of the hand, while the target was displaced in an orthogonal direction (forward or backward relative to the subject) at different times during movement. Reach duration was experimentally controlled to ensure that even the latest perturbation could have been fully corrected if that was the only objective of the underlying control strategy. More precisely, the remaining time after the onset of the latest correction was substantially larger than the time necessary to make the same movement in isolation.

Figure 1*a* shows the average hand paths for different perturbation times as well as for baseline (unperturbed) movements. Note the undershoot for 300 ms perturbations. In the rest of the analysis, the backward-perturbed trials are mirrored around the horizontal axis and pooled with the corresponding forward-perturbed trials. Figure 1*b* shows the tangential speed profiles. The early correction is incorporated so smoothly that its effect on the speed profile is hardly visible. The late correction, in contrast, causes a clear deviation from the bell-shaped baseline profile.

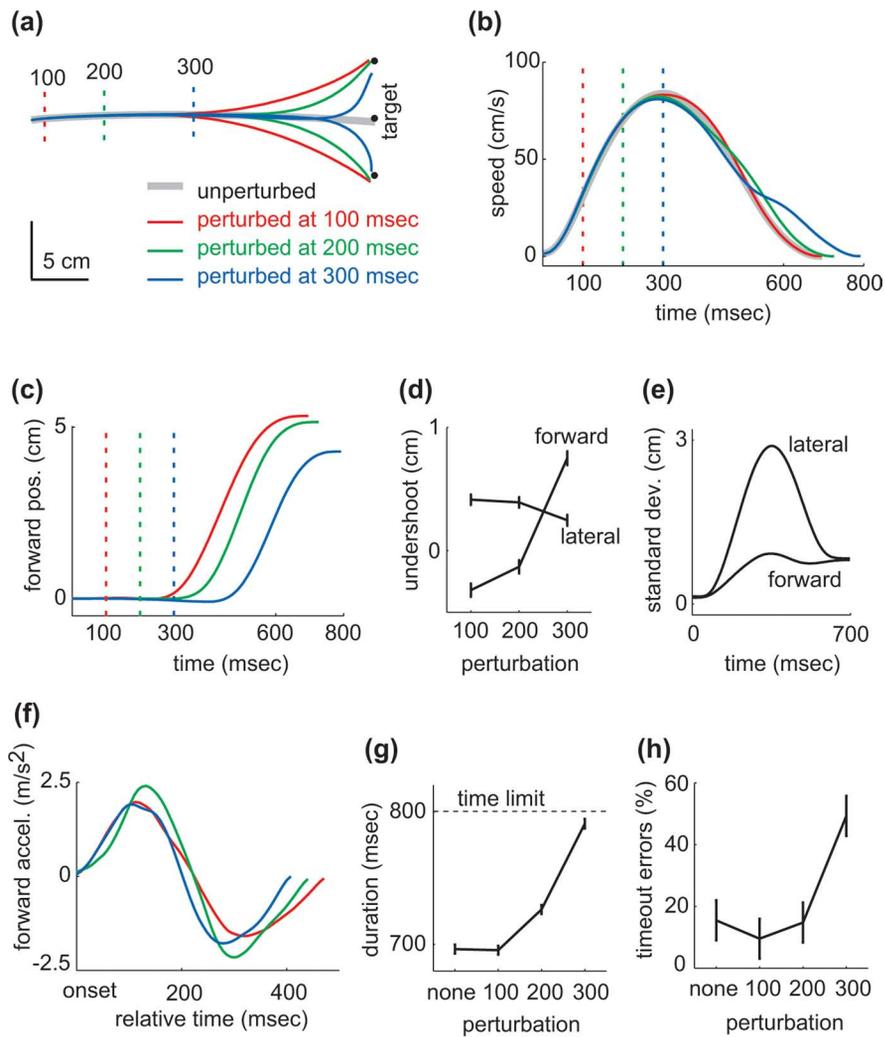


Figure 1. *a*, Average hand paths in experiment 1. The vertical marks show where the hand was at each perturbation time. Trajectory averaging was done as follows. The trajectory data from each individual trial were smoothed with a cubic spline (“csaps” function in the Matlab Spline Toolbox, smoothing parameter 0.001), and resampled at 100 points equally spaced in time. Analytical derivatives of the cubic spline were also computed at these 100 points, yielding velocities and accelerations. The resampled data were averaged separately in each condition. *b*, Tangential speed profiles for the hand paths shown in *a*. *c*, Corrective (forward) movement. The backward-perturbed trials have been mirrored around the horizontal axis and pooled with the corresponding forward-perturbed trials. The color code is the same as given in the legend in *a*. *d*, Undershoot, defined as endpoint error in the direction indicated in the plot. SEs are computed as described in Materials and Methods. *e*, Positional variance of the hand trajectories in unperturbed trials. Variances at each point in time are computed separately for each subject (from the resampled data), and then averaged over subjects, and the square root is plotted. *f*, Acceleration in the forward direction. For each perturbation time, the corresponding curve is aligned on the time when forward acceleration reached 5% of peak forward acceleration. *g*, Movement duration. *h*, Percentage of time-out errors, as signaled during the experiment. Note that for data analysis purposes, we increased the threshold on movement duration by 100 ms.

One could interpret this as a discrete submovement superimposed on the main movement; however, we will see below that the same effect can arise from a continuous optimal controller. The corrective movement, defined as movement in the forward direction, is shown in Figure 1*c*. The undershoot for the 300 ms perturbation is significantly larger than the undershoot for the 200 and 100 ms perturbations (Fig. 1*d*). Note that subjects are moving without visual feedback of the hand, and therefore some misalignment between vision and proprioception (Van Beers et al., 1999) should be expected. This may be the cause for the slight overshoot in the 100 ms conditions (indeed no overshoot was observed in the remaining experiments that were performed with visual feedback). Such misalignment should not depend on the

time of the target perturbation, and so the comparison between conditions is meaningful. There is also some systematic endpoint error in the lateral direction, although it shows a weaker and opposite trend (Fig. 1*d*); we return to it later.

Figure 1*f* shows the acceleration profile of each corrective movement, aligned on the time when forward acceleration first exceeds 5% of peak forward acceleration. The correction for the late perturbation lasts ~400 ms, which is more than sufficient to make a movement with the amplitude needed for complete correction. Indeed, an accurate 5 cm movement to a 1-cm-diameter target should take a little under 300 ms according to Fitts’ Law (Jeannerod, 1988). Thus, the lack of complete correction is not simply attributable to lack of time. Yet it has something to do with time: the overall movement duration was significantly increased in late-perturbation trials, all the way up to the 800 ms time limit (Fig. 1*g*), and consequently the percentage of time-out errors was increased (Fig. 1*h*). Although we asked subjects to treat the time limit as a hard constraint, they treated it on equal footing with the instruction to reach as close as possible to the center of the target and found a balance between these two task requirements. It is reasonable to assume that if we had convinced subjects to avoid time-out errors at all cost, the undershoot would have been even larger.

Optimal feedback control versus alternative models

The undershoot phenomenon is inconsistent with all previous models of motor control we are aware of. One such model (Flash and Henis, 1991) is an extension of the minimum-jerk model of trajectory planning (Flash and Hogan, 1985) to the domain of feedback corrections. It postulates that the hand tracks a planned minimum-jerk trajectory, and if the target is displaced, another minimum-jerk trajectory connecting the original and displaced target positions is added vectorially to the original plan. Naturally, that model predicts full correction in all cases (Fig. 2*f*). A related model (Hoff and Arbib, 1993), discussed in more detail below, is a feedback-control version of the minimum-jerk model. In our task, it makes the same predictions as the additive model (Flash and Henis, 1991) in Figure 2*f*. Another incompatible class of models consists of equilibrium-point control (Feldman and Levin, 1995; Bizzi et al., 1992) as well as other schemes (Hinton, 1984; Hoff and Arbib, 1993; Torres and Zipser, 2002) in which the hand is drawn to the target by some virtual spring. In such models, stopping can only occur when the hand reaches the target, which is in general contradiction with systematic endpoint errors that cannot be attributed to sensory-motor misalignment. In addition to the undershoot

studied here, phenomena that are problematic for these models include the undershoot of primary saccades (Harris, 1995), the overshoot of rapid wrist movements (Hoffman and Strick, 1999; Haruno and Wolpert, 2005), and the lack of equifinality (or failure to reach the target) in certain adaptation paradigms (Lackner and Dizio, 1994; Hinder and Milner, 2003).

Our optimal feedback control model (Todorov and Jordan, 2002b), which we have previously used to explain a number of unrelated phenomena, turns out to be compatible with the undershoot. One might have thought (as we did) that an optimal feedback controller should make larger on-line corrections and get closer to the displaced target than any other controller. This is not so, for interesting reasons explained in the next section. Here, we only present the simulation results. The model uses a linear approximation to arm dynamics. This is justifiable because the detailed nonlinearities of the arm are unlikely to influence the response to visual perturbations significantly (they are much more relevant when it comes to resisting mechanical perturbations). The model incorporates signal-dependent motor noise (Sutton and Sykes, 1967; Schmidt et al., 1979; Harris and Wolpert, 1998; Todorov, 2002) as well as sensory noise, and optimizes a mixed cost encouraging endpoint positional accuracy, endpoint stability (stopping), and energetic efficiency.

The model predictions (Fig. 2*a–e*) are plotted in the same format as the data (Fig. 1*a–e*). Note the close correspondence and, in particular, the undershoot for the late perturbation. When a target perturbation occurs in the model, we increase the remaining movement time as in the experimental data (Fig. 1*g*) and recompute the optimal controller (see Materials and Methods). If we use the unmodified optimal controller, which always ends the movement at the same time, the predicted undershoot is greatly increased (Fig. 2*c*, dashed lines). Thus, increasing movement duration is essential for avoiding a much larger undershoot in late perturbations, which may be why subjects were so reluctant to finish the movement on time as instructed.

In Figure 2*d*, we also see endpoint error in the lateral direction, but its magnitude decreases with increasing perturbation time, as in the experimental data. This is because the target is not perturbed in the lateral direction, and yet the overall movement time is increased, so the control costs that cause this endpoint error (energy consumption, and reduced accuracy because of signal-dependent noise) are effectively smaller. Note also the secondary speed bump in Figure 2*b*, which could be mistaken for a discrete submovement.

In addition to reproducing average behavior, the model faithfully captures the positional variability pattern of the hand trajectories in unperturbed trials (Figs. 1*e*, 2*e*). The larger variability in the lateral (main movement) direction reflects signal-dependent

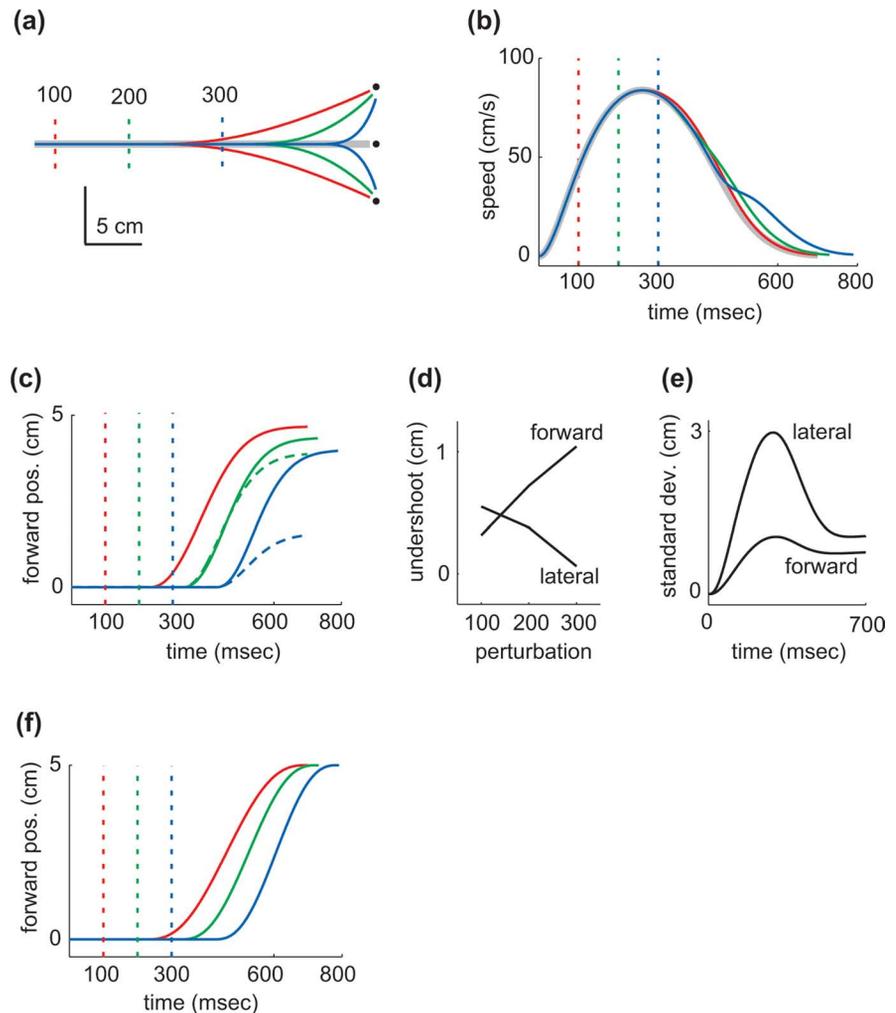


Figure 2. *a–e*, Same as the corresponding subplots of Figure 1, but for data generated by our optimal feedback control model. The dashed lines in *c* show predictions of a different optimal control model, in which movement duration is not adjusted when a perturbation arises. There is no dashed line for the 100 ms perturbation (red), because in that condition subjects did not increase the movement duration. *f*, Corrective movements predicted by the modified minimum-jerk model.

motor noise, which is larger in actuators that are more active. The reduction seen toward the end of the movement is an example of structured movement variability consistent with the minimal intervention principle (Todorov and Jordan, 2002b). Another manifestation of signal-dependent noise is the increased variability of the undershoot for late perturbations (Fig. 3*e*). This phenomenon is observed in the model and all experiments, and is only present in the corrective movement direction. Although any corrective movement incurs signal-dependent noise in that direction, the feedback controller has less time to suppress it when the noise is introduced late.

Analysis of feedback gains and new predictions

The above simulation results show that the optimal thing to do is make an incomplete correction. Why is this seemingly paradoxical strategy optimal? Such questions are often meaningless: the solution to a complex optimization problem is what it is, and the relationship between the problem and its solution does not have to be intuitive. Nevertheless, analysis of the model yields intuitive answers here. Key to our analysis is the fact that the optimal feedback gains are time varying.

As explained in Materials and Methods, the optimal feedback controller can be written as follows:

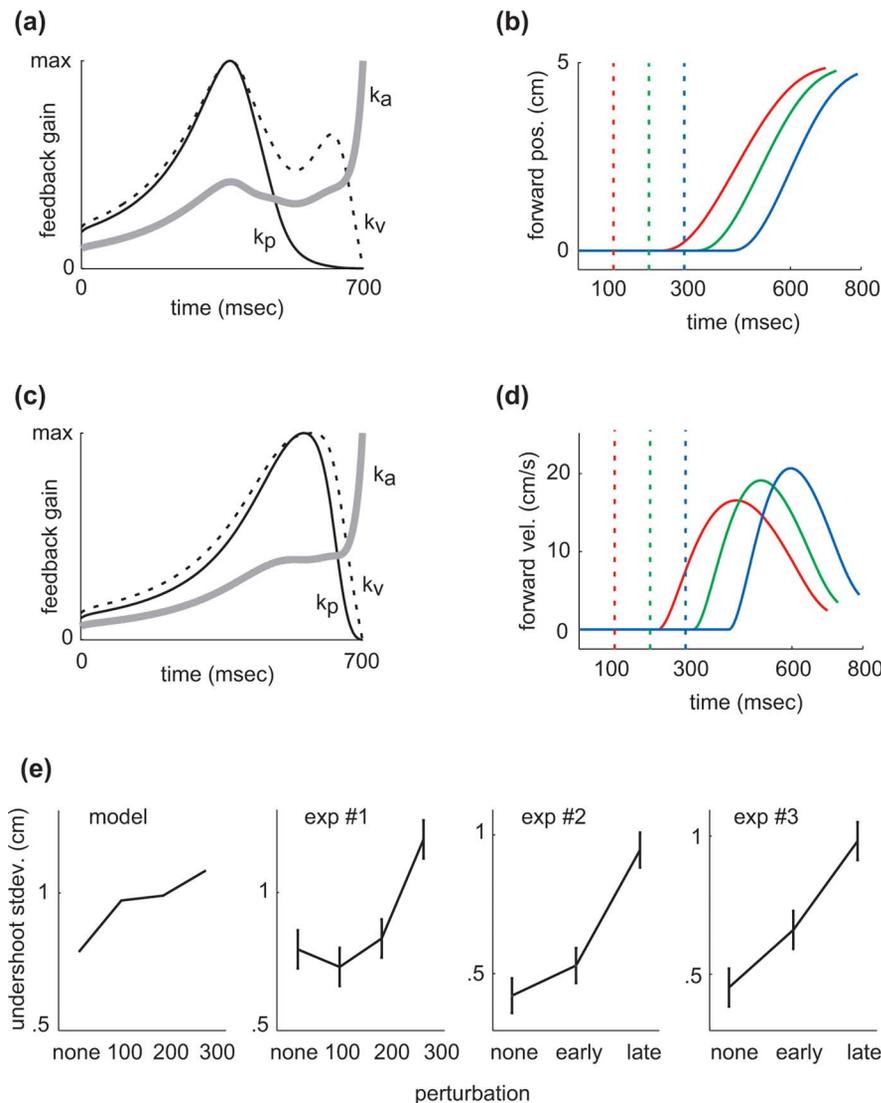


Figure 3. *a, c*, Optimal feedback gains, each scaled by its maximum value. The stop condition is shown in *a*; the hit condition is shown in *c*. *b*, Corrective movements predicted by the optimal feedback controller in the hit condition. Note that velocity is not reduced to zero at the end of the movement, especially for the 300 ms perturbation. *d*, Velocity of the corrective movements predicted in the hit condition. Note that velocity is not reduced to zero at the end of the movement, especially for the 300 ms perturbation. *e*, SD of the undershoot in the model and all three experiments. The SD was computed separately for each subject and perturbation time, and then averaged over subjects (by the ANOVA procedure) (see Materials and Methods). In unperturbed trials (“none”), we computed variability along the perturbation axis for the corresponding experiment, although these trials were unperturbed.

$$\mathbf{u}(t) = k_p(t)(\mathbf{p}^* - \hat{\mathbf{p}}(t)) - k_v(t)\hat{\mathbf{v}}(t) - k_a(t)\hat{\mathbf{a}}(t),$$

where k_p , k_v , and k_a are the optimal feedback gains; $\hat{\mathbf{p}}$, $\hat{\mathbf{v}}$, and $\hat{\mathbf{a}}$ are the optimal estimates of hand position, velocity, and muscle activation state (obtained by a modified Kalman filter); \mathbf{p}^* is the target position; and \mathbf{u} is the optimal control signal. The optimal feedback gains are illustrated in Figure 3*a*. We see that the positional gain k_p peaks early and then decreases in the last phase of the movement. In that phase, the velocity gain k_v , as well as the acceleration gain k_a , which can be thought of as force feedback, are large. Although these gain fluctuations are hard to understand quantitatively, qualitatively they have a simple interpretation: near the end of the movement the optimal controller enters a regime in which it is less sensitive to positional errors and instead aims to stop the movement in a stable manner. In retrospect, this is not surprising. If we think of a mass–spring–damper system, a large spring constant (or positional gain) will make the system

underdamped and cause oscillations, which is in conflict with the requirement to stop. The optimal controller effectively makes the system overdamped, achieving stability while compromising its ability to fully respond to last-minute positional errors. Thus, our analysis has uncovered a trade-off between endpoint stability and positional accuracy.

Another intuitive explanation for the undershoot observed in the model is the control cost associated with large and rapid last-minute corrections. Large control signals are penalized in two ways. One is a direct energy cost; the other is an indirect accuracy cost resulting from the signal-dependent nature of motor noise. Increased noise is particularly undesirable near the end of the movement when the feedback loop no longer has time to correct for it (Fig. 3*e*). In agreement with the latter interpretation, the undershoot predicted by the model increases when either the energy cost or the signal-dependent noise magnitude are increased (results not shown).

Analysis of feedback gains is also illuminating with regard to the modified minimum-jerk model and its failure to predict the undershoot (Fig. 2*f*). Consider the following feedback-control formulation (Hoff and Arbib, 1993) of the minimum-jerk model. At each point in time, a new minimum-jerk trajectory is formed, starting at the current hand position, velocity, and acceleration, and ending at the target with zero velocity and acceleration. The initial portion of this trajectory is used to control the movement, and then the procedure is repeated, making it possible to correct for on-line disturbances. More precisely, the hand is treated as a third-order system in which the position \mathbf{p} , velocity \mathbf{v} , and acceleration \mathbf{a} are state variables, and the control signal \mathbf{u} is defined as the derivative of acceleration (or jerk). It can be shown (see Materials and Methods) that the minimum-jerk feedback controller has the same general form as the optimal feedback controller, but with different feedback gains as follows:

$$k_p(t) = \frac{60}{(t_f - t)^3}, \quad k_v(t) = \frac{36}{(t_f - t)^2}, \quad k_a(t) = \frac{9}{(t_f - t)}.$$

We now see something that is in retrospect obvious: the only way the minimum-jerk feedback controller can always make a full correction, regardless of how late the perturbation arises, is to use infinite feedback gains at the end of the movement. As the time t approaches the final time t_f , all three feedback gains go to infinity, with k_p increasing faster than k_v , and k_a . Note that we could apply this minimum-jerk controller to a partially observable system, in which state estimates are obtained by a Kalman filter, and obtain a control scheme that overall is very similar to optimal feedback

control. The only important difference is in the sequence of time-varying feedback gains being used. The optimal feedback gains (Fig. 3*a*) not only predict behavior that better corresponds to the experimental data, but also guarantee minimal expected cost, and are finite rather than infinite (which is more biologically plausible).

We now return to the stability–accuracy trade-off, and use the model to obtain novel predictions reflecting this trade-off more directly. In the above analysis, the reason for the reduced sensitivity to positional errors was the need to stop at the target. What would happen if the importance of stopping decreases relative to the importance of reaching the target? In the model, stopping is enforced with a cost term quadratic in the final velocity and activation. If we scale down this cost term, then the optimal feedback gains change as shown in Figure 3*c*. Note that the positional gain k_p , now peaks much later. Consequently, the predicted undershoot is almost eliminated (Fig. 3*b*). The optimal controller for the modified cost function takes advantage of the relaxed stopping requirement, and no longer brings the velocity to zero at the specified final time, particularly for late perturbations (Fig. 3*d*).

Experimental confirmation of model predictions

The above predictions were tested in experiment 2, which compared two conditions: asking subjects to stop at the target versus allowing them to hit the target. Subjects made three-dimensional movements around a horizontal obstacle and aimed for a physical target attached to a three-dimensional robot (see Materials and Methods) (Fig. 4*a*). The obstacle was introduced to increase movement duration (so that we no longer had to impose a lower limit) and also to test the different predictions of optimal feedback control and alternative models with regard to obstacle avoidance (see next section). On randomly chosen trials, the robot rapidly displaced the target, left or right, either 50 or 350 ms after movement onset. Exceeding the maximum allowed duration (900 ms) resulted in a time-out error. Average movement trajectories are shown in Figure 4*b*.

Each subject was now tested in two conditions. In the stop condition, subjects were required to slow down their movement and touch the target gently. The robot used a low-gain servo controller so that the target could be easily displaced by the hand; a displacement >0.4 cm resulted in a “hitting-hard” error. In the hit condition, subjects were allowed to hit the target, although they were not instructed to do so. The robot used a high-gain servo controller and could absorb the impact with the hand; displacing the target no longer resulted in an error. The difference in target impedance made the distinction between the two conditions more ecologically valid.

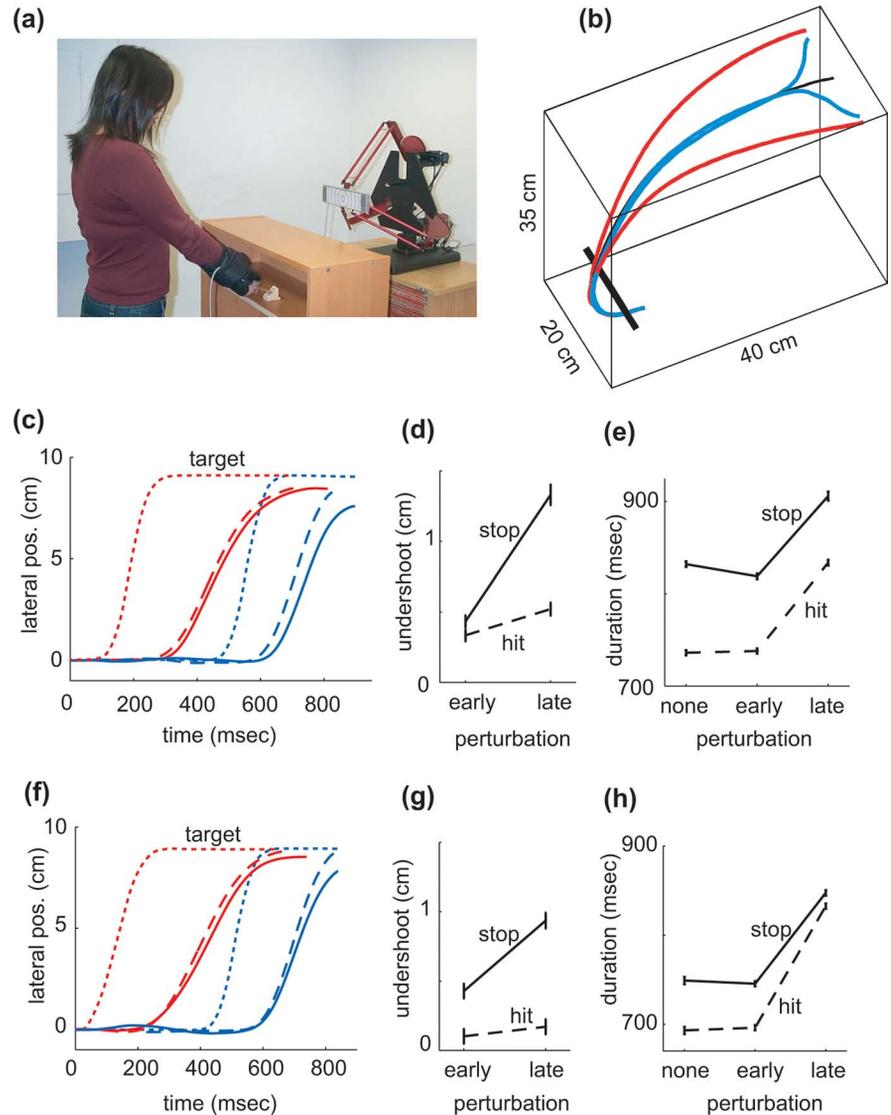


Figure 4. *a*, Setup for experiment 2. Subjects make a movement from the starting position receptacle to a target attached to the robot, while clearing a horizontal obstacle (bookshelf). The robot may displace the target by 9 cm left or right during the movement. *b*, Average hand paths in the stop condition of experiment 2. Trajectory averaging was done in a way similar to experiment 1, except that we now used a zero-phase-lag fourth-order Butterworth filter. The color code is the same as before: black, baseline; red, early perturbation; blue, late perturbation. *c*, Corrective movements in experiment 2. Dashed lines, Hit condition; solid lines, stop condition. *d*, Undershoot in experiment 2. *e*, Movement duration in experiment 2. *f*, Corrective movements in experiment 3. *g*, Undershoot in experiment 3. *h*, Movement duration in experiment 3.

The experimental results confirmed our model predictions. The undershoot for late perturbations was still present (Fig. 4*c,d*), but it was significantly smaller in the hit condition compared with the stop condition. For early perturbations, the undershoot was smaller (compared with late perturbations) and the difference between the hit and stop conditions was not significant. As before, late perturbations caused an increase in movement duration (Fig. 4*e*), and a substantial percentage of time-out errors in late/stop trials (40%). Movement duration in the stop condition was larger compared with the hit condition, and yet the correction was smaller (i.e., the undershoot was larger). Thus, as in experiment 1, subjects could have made a larger correction in the stop condition if that was their only objective.

Although at this point we had a convincing story, we remained puzzled by subjects' reluctance to treat the time limit as a hard constraint, although in experiment 2 we made time-out errors

much more salient. We reasoned that this may be because the time limit does not correspond to any physical property of the environment, and instead is signaled by the computer on the basis of an (invisible) timer. Could the outcome change if we provided an explicit and ecologically valid time cue? More importantly, could such a time cue reduce the uncertainty in the task and somehow enable subjects to eliminate the undershoot? These issues were addressed in experiment 3 in which we used an interception task rather than a pointing task (see Materials and Methods). The main change was that, as soon as hand movement was detected, the robot began to move the target downward at a low constant speed. Lateral target jumps were superimposed on this downward motion. Subjects were instructed to make contact with the target before it hit the horizontal edge of a board mounted underneath. The downward motion was repeatable and easily predictable, providing an explicit representation of allowed movement duration.

The results from experiment 3 (Fig. 4*f–h*) were similar to those of experiment 2 and in agreement with our model predictions. The undershoot in the stop condition was again larger than in the hit condition; the difference was now significant even for early perturbations (perhaps because we modified the method for detecting hitting-hard errors, making the threshold effectively smaller). Movement duration was again increased in late-perturbation trials and was larger in the stop condition compared with the hit condition. The time-out error rate in late/stop trials was reduced to 31%, indicating that the explicit time cue had an effect, but this rate was still higher than what would be expected if subjects treated the time limit as a constraint. The time-out error rate in late/hit trials was much lower (12% in experiment 2 and 7% in experiment 3).

Absence of imaginary targets in obstacle avoidance

Optimal feedback control differs from most alternative models in that it does not invent arbitrary subgoals, such as desired trajectories or imaginary targets, but instead uses all available resources to pursue the high-level movement goal. In obstacle avoidance tasks, it predicts that the hand should clear the obstacle without aiming for a specific imaginary target to the side of the obstacle. In contrast, deterministic trajectory planning models (Flash and Hogan, 1985; Uno et al., 1989) as well as other models (Rosenbaum et al., 1999) need such imaginary targets to avoid obstacles (and make curved movements in general). Note that stochastic optimal control models can avoid this limitation by taking into account the probability of collision attributable to random deviations from the average trajectory (Hamilton and Wolpert, 2002).

Here, we present two lines of evidence that subjects do not use imaginary targets in obstacle avoidance. First, we analyze the variability pattern of the hand paths in unperturbed trials in experiments 2 and 3. If subjects were aiming for an intermediate target, their hand paths should be less variable in the vicinity of that target, as we showed previously with real targets (Todorov and Jordan, 2002b). The variability pattern is plotted spatially in Figure 5, *a* and *c*, and as a scalar quantity (variability per dimension) in Figure 5, *b* and *d*. For both experiments, and for both the hit and stop conditions, we see that the pattern is bell-shaped. In particular, there is no evidence for a reduction of variability in the middle of the movement in which the imaginary target should be.

Second, we analyze the onset of the lateral correction relative to the time when the hand clears the obstacle and starts moving toward the robot. If subjects were aiming for an imaginary target to clear the obstacle, that target should be close to the reversal

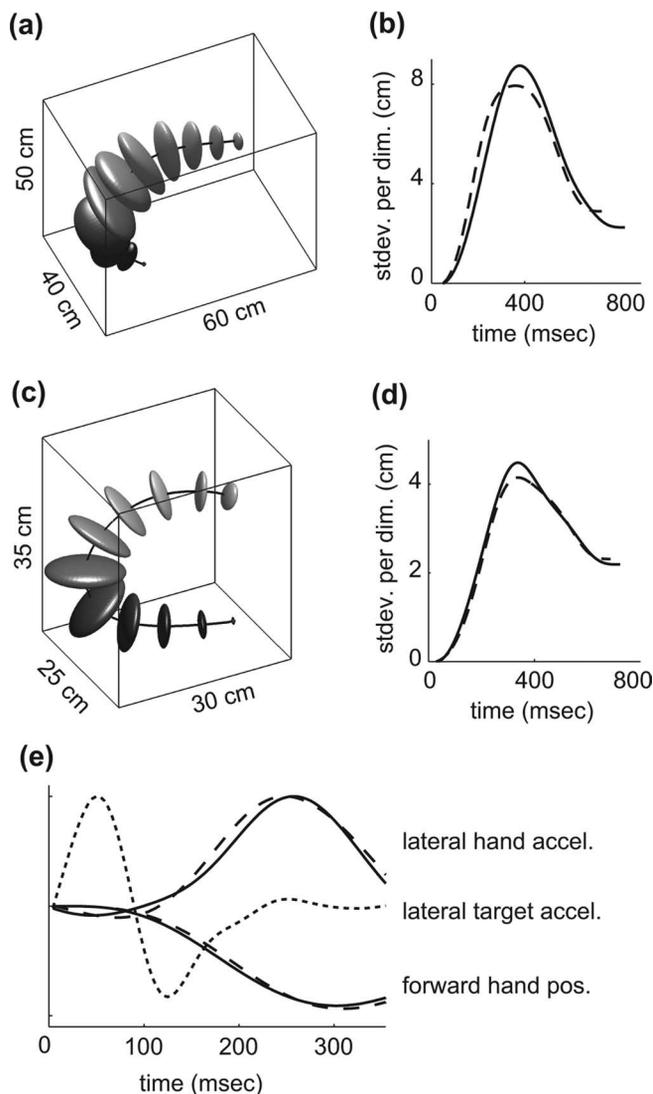


Figure 5. *a*, Spatial variability of unperturbed hand paths in experiment 2. The ellipsoids correspond to ± 2 SDs in each direction. Aligning three-dimensional trajectories for the purpose of computing variance is nontrivial and was done as follows. We first resampled all movements for a given subject at 100 points equally spaced along the path, and found the average trajectory. Then, for each point along the average trajectory, we found the nearest sample point from each individual trajectory. These nearest points were averaged to recompute the corresponding point along the average trajectory, and the procedure was repeated until convergence (which only takes 2–3 iterations). In this way, we extracted the spatial variability of the hand paths, independent of timing fluctuations. That is why the covariance ellipsoids are flat in the movement direction. *b*, Variability per dimension, for the stop (solid) and hit (dashed) conditions in experiment 2. At each point along the path, this quantity was computed as the square root of the trace of the covariance matrix for the corresponding ellipsoid, divided by 3. To plot variability as a function of time, we resampled back from equal-space to equal-time intervals. *c*, *d*, Same as subplots (*a*, *b*) but for experiment 3. *e*, Normalized target acceleration in the lateral direction, lateral hand position, and hand position in the forward direction (positive is toward the robot). Dashed lines, Hit condition; solid lines, stop condition. Note that the onset of hand acceleration occurs before the movement reversal in the forward direction.

point and should not move when the robot displaces the final target. Therefore, the corrective movement should not start before the reversal point. We focus on experiment 3, which was specifically designed to address this question by shifting the obstacle farther away from the starting position (thus delaying the reversal) and detecting the onset of hand movement with a pressure sensor (allowing an earlier perturbation). Figure 5*e* shows that the reversal occurs at ~ 300 ms in both the hit and stop

conditions, whereas the corrective lateral acceleration in early perturbations starts at 100–150 ms. Given the filtering of the musculoskeletal system, the neural command driving the correction must have been generated even earlier. Thus, subjects begin to correct for the target jump before having reached the hypothetical via-point, casting serious doubt on the existence of the latter.

Flexible strategies for opportunistic control

The sensorimotor strategies predicted by optimal feedback control exhibit great flexibility, in the sense that they are adapted to the task, body, and environment, and take advantage of every opportunity for achieving higher performance. This is in sharp contrast with traditional trajectory-planning models (Flash and Hogan, 1985; Uno et al., 1989), which essentially view all tasks as being the same as long as the average trajectory is the same. Here, we make four additional experimental observations illustrating the flexibility inherent in the optimal control framework.

First, in the hit condition in experiments 2 and 3, subjects actually hit the target harder, although they were not instructed to do so (and one-half of them had already performed the task in the stop condition). In both experiments, the forward hand velocity before impact was two to three times larger in the hit condition compared with the stop condition. In experiment 3, in which we used a force sensor, the normal force in the first 50 ms after contact was approximately three times larger in the hit condition compared with the stop condition. Both the model and the experimental results suggest that the stopping requirement causes decreased sensitivity to positional errors late in the movement. Thus, the relaxed stopping requirement in the hit condition is exploited to increase positional accuracy. It should be noted, however, that the stopping requirement was not fully eliminated in the hit condition. In experiment 2 (unperturbed trials, hit condition), subjects reduced their hand speed from 175 cm/s peak to 39 cm/s before contact, a 78% reduction; in experiment 3, this reduction was 84%. For comparison, the speed reduction in the stop condition was 96 and 92%, respectively.

Second, subjects exploited the relaxed accuracy requirement in the vertical direction in experiment 3, in which the target was a vertical stripe rather than a circle. Focusing on unperturbed trials, we see that lateral and vertical endpoint errors are equally variable in experiment 2, but vertical “errors” are significantly more variable than lateral errors in experiment 3 (Fig. 6*a*). Furthermore, in experiment 3, subjects did not fully use feedback to adjust their vertical hand position relative to the falling target. Indeed, variability of the vertical endpoint position in absolute coordinates (relative to the room) was smaller than variability relative to the target. We know that subjects are able to correct in the lateral direction, not completely, but still the undershoot is much smaller than the correction. Thus, the difference between absolute and relative variance in the vertical direction does not reflect an inability to correct, but rather an absence of a need to correct, in agreement with the minimal intervention principle (Todorov and Jordan, 2002b).

Third, subjects found a way to exploit the different methods we used to detect the movement end in experiments 2 and 3. Here, we focus on late perturbation trials and analyze the hand velocity immediately before contact with the target. In experiment 2, we used a speed threshold that required both forward and lateral velocity to be reduced to end the trial. In the hit condition, the necessary velocity reduction could result from contact with the target, attributable to target impedance and friction, respectively. In the stop condition, however, the target could not be

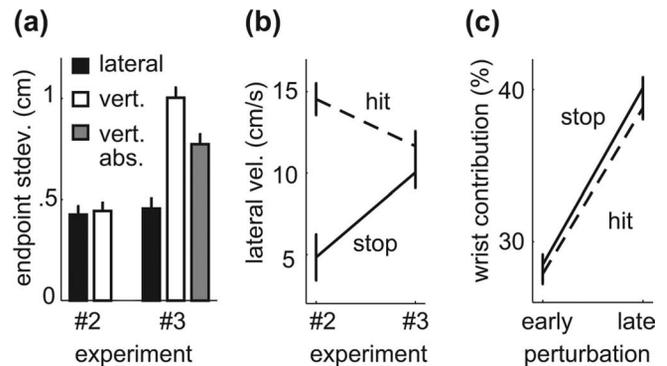


Figure 6. *a*, Endpoint SD in different directions, experiments 2 and 3, unperturbed trials. Black, Lateral direction; white, vertical direction (coordinates relative to the target); gray, vertical direction (absolute coordinates). In experiment 3, the relative and absolute endpoint positions are different in the vertical direction, because the target is falling and the variability in movement duration causes variability in vertical target position at the end of the movement. *b*, Lateral velocity immediately before contact with the robot, in late perturbation trials. *c*, Wrist contribution to the lateral correction, in a pilot experiment with 10 subjects. The main difference from experiment 2 was that the wrist was not braced. The lateral correction could be accomplished with humeral rotation (resulting mostly in translation of the hand-held pointer) or wrist flexion/extension (resulting in rotation of the pointer in the horizontal plane). The pointer was held in such a way that the Polhemus sensor was near the wrist. Therefore, the lateral displacement of the sensor on perturbed trials (relative to the average trajectory on unperturbed trials) can be used as an index of how much humeral rotation contributes to the correction. The displacement of the tip of the pointer is defined as the total correction. The difference between the two is the contribution of the wrist. Dividing the latter by the total correction, and multiplying by 100, we obtain the percentage wrist contribution.

exploited to stop the movement in either direction, and thus the lateral velocity in experiment 2 was small (Fig. 6*b*). In experiment 3, we used a force sensor to detect contact with the target, and defined the time of contact as the movement end, so lateral velocity did not have to be reduced as much. Subjects took advantage of this: the difference in lateral velocity between the stop and hit conditions in experiment 3 was much smaller than in experiment 2, and was not significant (Fig. 6*b*). We already know that the stopping requirement conflicts with positional accuracy. It is then likely that finding a way to partially avoid this requirement (in the lateral direction) afforded improved positional accuracy in experiment 3.

Fourth, subjects exploited the biomechanical redundancy of the arm when they had a chance. In experiments 2 and 3, redundancy was reduced by bracing the wrist; however, we performed a previous pilot experiment in which the wrist was not braced (otherwise, it was similar to experiment 2). In that case, the lateral correction was accomplished with a combination of wrist flexion/extension and humeral rotation. We found that the percentage contribution of the wrist was larger in late perturbations compared with early perturbations, in both the hit and stop conditions (Fig. 6*c*). The fact that the wrist contributes <30% in early corrections suggests that the preferred strategy is to use humeral rotation. This may be because for a given force level larger muscles are less affected by signal-dependent noise (Hamilton et al., 2004). In late corrections, however, it is perhaps more difficult to accelerate and decelerate the entire forearm within the remaining time, and thus the wrist contribution increases.

Modeling changes in duration and variability

The LQG framework, which we used in the above model as well as in most of our previous work on optimal feedback control, is computationally efficient but has a number of limitations. In the

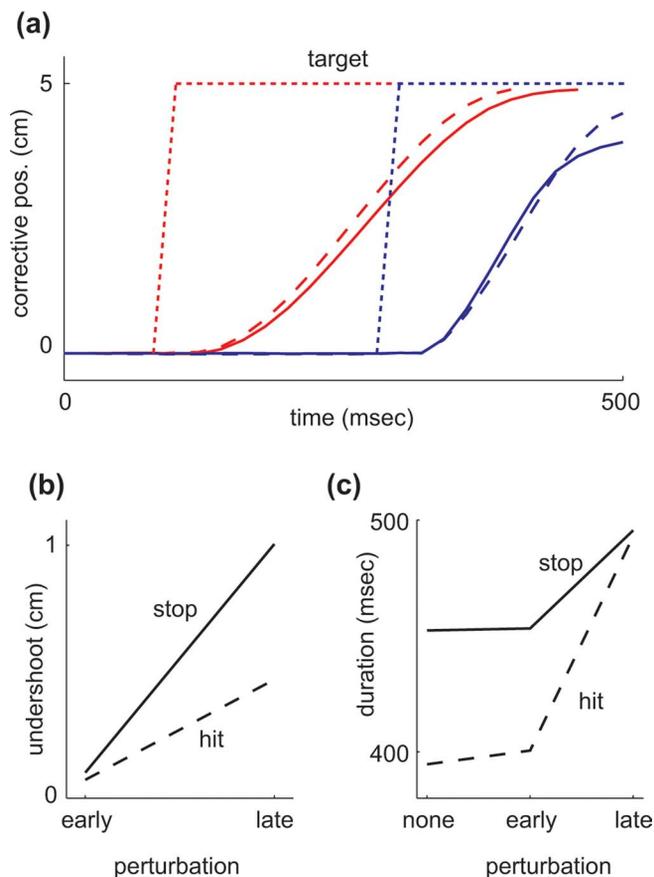


Figure 7. *a*, Corrective movements of the more general optimal feedback control model. The solid and dashed lines correspond to the stop and hit conditions, respectively. The hand is restricted to a grid of discrete states; however, the dynamics are stochastic, and so the average (over 1000 simulated trials) is smooth, although the individual trajectories have a staircase pattern. *b*, *c*, Undershoot and movement duration in the stop and hit conditions for different perturbation times. Same format as the experimental data in Figure 4.

present context, its limitations are as follows: (1) movement duration cannot be modified on-line in response to target perturbations; (2) stopping constraints have to be modeled with quadratic costs instead of more natural step-function costs; (3) the controller cannot be adapted to the statistics of the target perturbations. Here, we present an optimal feedback control model that avoids these limitations.

The new model is constructed using more general but less efficient discretization techniques that require a simpler second-order model of arm dynamics. Movement duration is defined as the point in time when the hand first reaches the target plane. A term proportional to movement duration is included in the cost function. Time-out errors and hitting-hard errors are penalized with step-function costs. The trimodal distribution of final target positions is taken into account in the optimization process. For details, see Materials and Methods.

The new model (Fig. 7) accounts for the salient findings in experiment 2 (Fig. 4*c–e*) and experiment 3 (Fig. 4*f–h*). The speed of the corrective movement (slope of the positional traces in Figs. 7*a* and 4*c,f*) is smaller in early perturbations. The undershoot in late perturbations is larger in the stop versus hit condition. The undershoot in early perturbations is smaller compared with late perturbations, in both stop and hit conditions. Movement duration in baseline and early perturbations is larger in the stop versus hit condition. For late perturbations, movement duration in-

creases in both the stop and hit conditions. Note that the changes in movement duration are now predicted by the model, as opposed to being taken from the data as in the LQG model. This is possible because the new controller can adjust the duration on-line, by modulating the speed in the main movement direction and thus reaching the target plane at different times.

The new model allows us to address an additional phenomenon that is beyond the scope of LQG models. The phenomenon (Fig. 8) is that trajectory variability on unperturbed trials was larger in experimental sessions with perturbations compared with baseline sessions without perturbations. Results from the hit and stop conditions are averaged in this analysis. Figure 8 shows that frequent perturbations lead to some adaptive change in the sensorimotor system, which in turn leads to increased variability on trials without perturbations. The precise time course of such adaptation is difficult to estimate because measuring variability requires many trials.

What could be the nature of this adaptive change? One possible explanation is that, in sessions with perturbations, trial-to-trial adaptation (Thoroughman and Shadmehr, 2000; Donchin et al., 2003) causes the system to be in a different state every time an unperturbed trial is encountered. However, in target perturbation paradigms, trial-to-trial adaptation is negligible (Diedrichsen et al., 2005) (see also below). Another possible explanation is that target perturbations are for some reason misinterpreted as an increase in sensory noise, in which case the “optimal” thing to do is reduce the reliance on sensory feedback, causing suboptimal performance. A third explanation, which we pursue below, is an adaptive change in the feedback controller.

In environments with large unpredictable perturbations, one would expect the optimal feedback controller to be more concerned with correcting the perturbations than the smaller errors attributable to internal noise. This is confirmed by our simulations. In Figure 8*a*, we compare the trajectory variability of two feedback controllers, one optimized for an unperturbed environment and the other one for a perturbed environment matching our task. As expected, the latter controller is better at correcting for perturbations (data not shown); however, it allows higher variability on trials without perturbations. This is broadly consistent with the minimal intervention principle as well as with the idea of Pareto optimality: improving any aspect of the behavior of an optimal controller requires sacrifices elsewhere. Analyzing the specific changes in feedback gains that lead to increased variability is interesting but beyond the scope of the present paper.

Lack of trial-to-trial adaptation

Trial-to-trial adaptation has been found in perturbations of the hand (Thoroughman and Shadmehr, 2000; Donchin et al., 2003) but not the target (Diedrichsen et al., 2005). Here, we replicate the latter finding: in our target perturbation experiments trial-to-trial adaptation turns out to be negligible.

To quantify such adaptation we adopted a version of the state-space approach, namely the following:

$$\mathbf{z}(n+1) = \mathbf{A}\mathbf{z}(n) + \mathbf{B}\mathbf{w}(n) + \boldsymbol{\eta}(n)$$

$$\mathbf{y}(n) = \mathbf{C}\mathbf{z}(n) + \mathbf{D}\mathbf{w}(n) + \boldsymbol{\gamma}(n).$$

The correction on trial n is denoted $\mathbf{y}(n)$ and has two elements: distance in the corrective direction from the average unperturbed trajectory, measured at the time of the late perturbation and at the end of the movement. The perturbation $\mathbf{w}(n)$ has two elements specifying the target position (again in the corrective direction) immediately before the late perturbation and at the end

of the movement. The vector $\mathbf{z}(n)$ is the internal learning state. It also has two elements that the model is free to use in whatever way is needed to fit the data. We quantify the correction and the perturbation using pairs of measurements to allow the model to capture the difference between early and late perturbations. $\eta(n)$, $\gamma(n)$ are independent zero-mean two-dimensional Gaussian random variables with covariances Q , S .

The next learning state $\mathbf{z}(n+1)$ may in general depend on the current learning state $\mathbf{z}(n)$, the perturbation $\mathbf{w}(n)$, and the correction $\mathbf{y}(n)$. However, \mathbf{y} is a linear function of \mathbf{z} , \mathbf{w} , and so we do not include a \mathbf{y} term in the first equation. Note also that if there is any trial-to-trial learning here, it should be related to predicting the final target position (and initiating a normal movement aimed at that position); thus, it makes more sense to learn from \mathbf{w} rather than \mathbf{y} .

The sequences of corrections $\mathbf{y}(n)$ and perturbations $\mathbf{w}(n)$ were measured. Given these measurements, the most likely values of A , B , C , D , Q , S as well as the sequence of learning states $\mathbf{z}(n)$ were computed using the expectation maximization (EM) algorithm (Cheng and Sabes, 2006). Because EM can get trapped in local minima, model fitting was run multiple times with different initial conditions and the best result was used. We fit the model separately for each experiment (2 vs 3), condition (stop vs hit), and subject. The first two-thirds of the data in each experimental session was used for model fitting and the last one-third for model testing.

To evaluate the performance of the learning model, we regressed each component of \mathbf{y} either on \mathbf{z} and \mathbf{w} , or on \mathbf{w} alone, or on \mathbf{z} alone. The first regression measures the performance of the full model, whereas the latter two regressions measure the contribution of feedback correction (\mathbf{w}) and learning component (\mathbf{z}), respectively. The R^2 and p values for the regressions were averaged over subjects, experiments, and conditions. The regressions were first done on all trials, and then separately on the baseline (no perturbation) trials, early jump trials, and late jump trials, because the models are likely to perform differently on different trial types. Table 2 shows the average R^2 values multiplied by 100 (to obtain a measure of variance explained), only for the cases in which $p < 0.05$ on average. In the remaining cases, we found $p > 0.3$; thus, there was clear separation between significant and nonsignificant regression fits.

None of the regressions on \mathbf{z} alone are significant, suggesting lack of trial-to-trial learning. The regressions on \mathbf{w} are as expected: given the target position at the middle and at the end of the movement, one can predict the hand position at the end of the movement (y_2) in both types of perturbations as well as the hand position at the middle of the movement (y_1) in early perturbations. Note that the combined model \mathbf{z} , \mathbf{w} is slightly but systematically better than \mathbf{w} alone, suggesting that there may be a small learning effect. Given how small this effect is, it is not surprising that regressions on \mathbf{z} alone were far from significant.

Discussion

The biological processes that continuously improve behavior closely resemble iterative optimization. This makes optimal control theory a natural framework for studying the neural control of

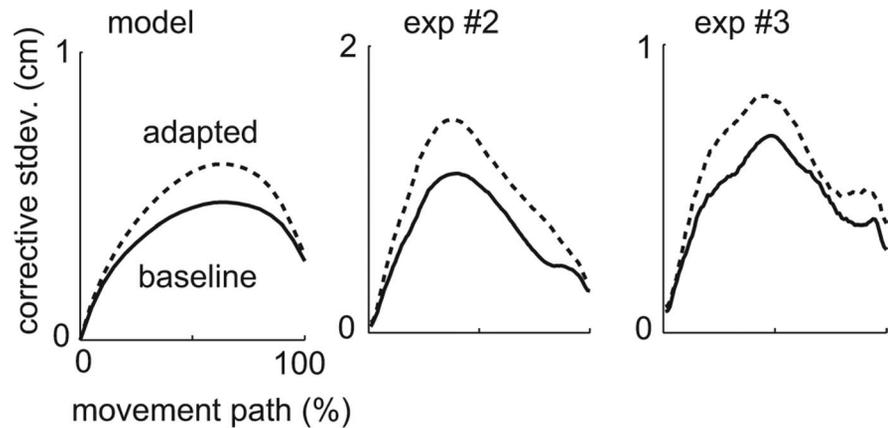


Figure 8. Hand positional variance on unperturbed trials, measured along the perturbation direction. Trajectories are aligned at equal intervals along the movement path to compute variance. The solid line (baseline) is the variance in blocks without perturbations. The dashed line (adapted) is the variance in blocks with 66% perturbations. Data from the hit and stop conditions are averaged.

Table 2. Percentage variance explained

Variable	Regressor	All trials	Baseline	Early jump	Late jump
y_1	\mathbf{z}, \mathbf{w}	74.36		91.13	
y_1	\mathbf{w}	71.50		86.33	
y_1	\mathbf{z}				
y_2	\mathbf{z}, \mathbf{w}	99.34		99.85	98.15
y_2	\mathbf{w}	99.28		99.83	97.34
y_2	\mathbf{z}				

movement. It is also a very successful framework in terms of explaining the details of experimental data (Todorov, 2004). However, one of its most appealing features remains mostly untapped: the ability to predict task-specific sensorimotor strategies and thereby changes in behavior that result from systematic task variation. This is a gap not only in optimal control models but in the field of motor control in general. A substantial number of studies (including most of the literature on motor adaptation) have used a single task: reaching. The emphasis on servo control has created the impression that, as long as a desired trajectory can somehow be planned, motor execution (and all sensorimotor processing during movement) is the same no matter what the organism is trying to accomplish. Planning models have focused on the geometry of limb trajectories and have mostly ignored the context that gives functional meaning to these trajectories. We see this as a substantial gap in the current understanding of sensorimotor function; the present paper is a step toward filling that gap.

To vary the task systematically, we need a compact and experimentally accessible representation of the task space. Optimal control provides the perfect tool: composite cost functions. A central argument of this paper is that subjects optimize a composite cost as opposed to a homogeneous cost under multiple hard constraints. Indeed, we did everything we could to enforce a hard constraint on movement duration, and yet subjects never treated it as such. Instead, they always found a balance between undershoot and time-out errors. The changes in our experimental design affected the relative importance of these errors; in particular, the switch to the intercept task (which made the duration threshold explicit) resulted in the lowest percentage of time-out errors. In addition to accuracy and duration, we proposed that the composite cost includes endpoint stability (stopping in par-

ticular) and energy consumption. We showed directly that stability is part of the cost, by allowing subjects to interact with a high-impedance target and finding that they take advantage of it. The only evidence for the energy cost was indirect: it was needed to make our model fit the data. However, other studies have provided more direct evidence: increased muscle cocontraction has been found to yield more accurate movements (Burdet et al., 2001; Gribble et al., 2003), and yet this is not a strategy that subjects normally use, suggesting that they care about energetic efficiency in addition to accuracy. It is also notable that successful optimal control models of full-body movements are predominantly based on energy minimization (Anderson and Pandy, 2001; Pandy, 2001). In obstacle avoidance tasks (experiments 2 and 3), there is likely to be a fifth component of the cost having to do with avoiding the obstacle. Although we did not model this cost, we showed that obstacle avoidance does not rely on hard constraints such as fixed imaginary targets to the side of the obstacle.

The main effect we analyzed, the incomplete correction for late perturbations, reflects the closed-loop component of the sensorimotor strategy. The fact that the effect decreased in the hit condition means that the visuomotor loop operated differently, as predicted by our model. Thus, changes in stopping requirements (as well as target impedance) caused changes in the way visual feedback is used to make on-line corrections. This may be the first demonstration that visuomotor feedback loops are affected by the task and in particular by nonvisual components of the task. In addition to demonstrating task sensitivity, we provided additional evidence that sensorimotor strategies are consistent with the minimal intervention principle of optimal feedback control (Todorov and Jordan, 2002b). We found that positional variability is large during movement (especially three-dimensional movement) and is only reduced near the end, where accuracy is needed. We also found that, when the target is a vertical stripe, endpoint variability is larger in the vertical direction and visual feedback is not fully used to suppress variability in that direction. These results reaffirm the usefulness of looking beyond average trajectories, studying variability patterns and responses to perturbations, and modeling the sensorimotor strategies responsible for such effects.

Motor adaptation is a phenomenon that has not yet been addressed in the optimal control framework, but in principle is easy to address, as we showed in our model of increased variability attributable to frequent perturbations. One can impose any change in the task or environment, compute the new optimal controller, and use it as a model of adapted behavior. Of course, adaptation is rarely complete; thus, the predicted adaptation effect should be somewhere in between the baseline and fully adapted optimal controllers. An interesting open question is how to relate trial-to-trial dynamics of learning to asymptotic predictions regarding optimal adaptation. One way to do this is to model trial-to-trial changes as arising from an iterative optimization algorithm, which in the limit converges to the adapted optimal controller. This approach may yield richer models of learning dynamics than the linear state-space models currently used.

References

- Anderson F, Pandy M (2001) Dynamic optimization of human walking. *J Biomech Eng* 123:381–390.
- Bernstein N (1967) *The coordination and regulation of movements*. Oxford: Pergamon.
- Bertsekas D (2001) *Dynamic programming and optimal control*, Ed 2. Belmont, MA: Athena Scientific.
- Bizzzi E, Hogan N, Mussa-Ivaldi F, Giszter S (1992) Does the nervous system use equilibrium-point control to guide single and multiple joint movements? *Behav Brain Sci* 15:603–613.
- Burdet E, Osu R, Franklin D, Milner T, Kawato M (2001) The central nervous system stabilizes unstable dynamics by learning optimal impedance. *Nature* 414:446–449.
- Cheng S, Sabes P (2006) Modeling sensorimotor learning with linear dynamical systems. *Neural Comput* 18:760–793.
- Cole K, Abbs J (1987) Kinematic and electromyographic responses to perturbation of a rapid grasp. *J Neurophysiol* 57:1498–1510.
- Desmurget M, Grafton S (2000) Forward modeling allows feedback control for fast reaching movements. *Trends Cogn Sci* 4:423–431.
- Diedrichsen J, Hashambhoy Y, Rane T, Shadmehr R (2005) Neural correlates of reach errors. *J Neurosci* 25:9919–9931.
- Domkin D, Laczko J, Jaric S, Johansson H, Latash M (2002) Structure of joint variability in bimanual pointing tasks. *Exp Brain Res* 143:11–23.
- Donchin O, Francis J, Shadmehr R (2003) Quantifying generalization from trial-by-trial behavior of adaptive systems that learn with basis functions: theory and experiments in human motor control. *J Neurosci* 23:9032–9045.
- Feldman A, Levin M (1995) The origin and use of positional frames of reference in motor control. *Behav Brain Sci* 18:723–744.
- Flash T, Henis E (1991) Arm trajectory modification during reaching towards visual targets. *J Cogn Neurosci* 3:220–230.
- Flash T, Hogan N (1985) The coordination of arm movements: an experimentally confirmed mathematical model. *J Neurosci* 5:1688–1703.
- Gribble P, Mullin L, Cothros N, Mattar A (2003) Role of cocontraction in arm movement accuracy. *J Neurophysiol* 89:2396–2405.
- Hamilton A, Wolpert D (2002) Controlling the statistics of action: obstacle avoidance. *J Neurophysiol* 87:2434–2440.
- Hamilton A, Jones K, Wolpert D (2004) The scaling of motor noise with muscle strength and motor unit number in humans. *Exp Brain Res* 157:417–430.
- Harris C (1995) Does saccadic undershoot minimize saccadic flight-time? A Monte-Carlo study. *Vision Res* 35:691–701.
- Harris C, Wolpert D (1998) Signal-dependent noise determines motor planning. *Nature* 394:780–784.
- Haruno M, Wolpert D (2005) Optimal control of redundant muscles in step-tracking wrist movements. *J Neurophysiol* 94:4244–4255.
- Hinder M, Milner T (2003) The case for an internal dynamics model versus equilibrium point control in human movement. *J Physiol (Lond)* 549:953–963.
- Hinton G (1984) Parallel computations for controlling an arm. *J Mot Behav* 16:171–194.
- Hoff B (1992) *A computational description of the organization of human reaching and prehension*. PhD thesis, University of Southern California.
- Hoff B, Arbib M (1993) Models of trajectory formation and temporal interaction of reach and grasp. *J Mot Behav* 25:175–192.
- Hoffman D, Strick P (1999) Step-tracking movements of the wrist. IV. Muscle activity associated with movements in different directions. *J Neurophysiol* 81:319–333.
- Jeannerod M (1988) *The neural and behavioural organization of goal-directed movements*. Oxford: Oxford UP.
- Komilis E, Pelissou D, Prablanc C (1993) Error processing in pointing at randomly feedback-induced double-step stimuli. *J Mot Behav* 25:299–308.
- Kuo A (1995) An optimal control model for analyzing human postural balance. *IEEE Trans Biomed Eng* 42:87–101.
- Kushner H, Dupuis P (2001) *Numerical methods and methods for stochastic optimal control problems in continuous time*, Ed 2. New York: Springer.
- Lackner J, Dizio P (1994) Rapid adaptation to coriolis-force perturbations of arm trajectory. *J Neurophysiol* 72:299–313.
- Li W, Todorov E (2007) Iterative linearization methods for approximately optimal control and estimation of nonlinear stochastic systems. *Int J Control*, in press.
- Loeb G, Levine W, He J (1990) Understanding sensorimotor feedback through optimal control. *Cold Spring Harb Symp Quant Biol* 55:791–803.
- Meyer D, Abrams R, Kornblum S, Wright C, Smith J (1988) Optimality in human motor performance: ideal control of rapid aimed movements. *Psychol Rev* 95:340–370.

- Nelson W (1983) Physical principles for economies of skilled movements. *Biol Cybern* 46:135–147.
- Nicols T, Houk J (1976) Improvement in linearity and regulations of stiffness that result from actions of stretch reflex. *J Neurophysiol* 39:119–142.
- Pandy M (2001) Computer modeling and simulation of human movement. *Annu Rev Biomed Eng* 3:245–273.
- Pelisson D, Prablanc C, Goodale M, Jeannerod M (1986) Visual control of reaching movements without vision of the limb. II. Evidence of fast unconscious processes correcting the trajectory of the hand to the final position of a double-step stimulus. *Exp Brain Res* 62:303–311.
- Popescu F, Rymer W (2000) End points of planar reaching movements are disrupted by small force pulses: an evaluation of the hypothesis of equifinality. *J Neurophysiol* 84:2670–2679.
- Prablanc C, Martin O (1992) Automatic control during hand reaching at undetected two-dimensional target displacements. *J Neurophysiol* 67:455–469.
- Rosenbaum D, Meulenbroek R, Vaughan J, Jansen C (1999) Coordination of reaching and grasping by capitalizing on obstacle avoidance and other constraints. *Exp Brain Res* 128:92–100.
- Schmidt R, Zelaznik H, Hawkins B, Frank J, Quinn J (1979) Motor-output variability: a theory for the accuracy of rapid motor acts. *Psychol Rev* 86:415–451.
- Scholz J, Schoner G (1999) The uncontrolled manifold concept: identifying control variables for a functional task. *Exp Brain Res* 126:289–306.
- Scott S (2004) Optimal feedback control and the neural basis of volitional motor control. *Nat Rev Neurosci* 5:534–546.
- Sutton G, Sykes K (1967) The variation of hand tremor with force in healthy subjects. *J Physiol (Lond)* 191:699–711.
- Thoroughman K, Shadmehr R (2000) Learning of action through adaptive combination of motor primitives. *Nature* 407:742–747.
- Todorov E (2002) Cosine tuning minimizes motor errors. *Neural Comput* 14:1233–1260.
- Todorov E (2004) Optimality principles in sensorimotor control. *Nat Neurosci* 7:907–915.
- Todorov E (2005) Stochastic optimal control and estimation methods adapted to the noise characteristics of the sensorimotor system. *Neural Comput* 17:1084–1108.
- Todorov E, Jordan M (2002a) A minimal intervention principle for coordinated movement. In: *Advances in neural information processing systems 15* (Becker S, Thrun S, Obermayer K, eds), pp 27–34. Cambridge, MA: MIT.
- Todorov E, Jordan M (2002b) Optimal feedback control as a theory of motor coordination. *Nat Neurosci* 5:1226–1235.
- Todorov E, Li W, Pan X (2005) From task parameters to motor synergies: a hierarchical framework for approximately-optimal feedback control of redundant manipulators. *J Robot Syst* 22:691–710.
- Torres E, Zipser D (2002) Reaching to grasp with a multi-jointed arm. I. Computational model. *J Neurophysiol* 88:2355–2367.
- Uno Y, Kawato M, Suzuki R (1989) Formation and control of optimal trajectory in human multijoint arm movement: minimum torque-change model. *Biol Cybern* 61:89–101.
- Van Beers R, Sittig A, Van der Gon J (1999) Integration of proprioceptive and visual position-information: an experimentally supported model. *J Neurophysiol* 81:1355–1364.