40 Computational Neuroanatomy of Voluntary Motor Control

REZA SHADMEHR AND JOHN W. KRAKAUER

ABSTRACT We review some of the impairments in motor control, motor learning and higher-order motor control in patients with lesions of the cerebellum, parietal cortex, and basal ganglia. We attempt to explain some of these impairments in terms of computational ideas such as state estimation, optimization, prediction, cost, and reward. We suggest that a function of the cerebellum is system identification: to built internal models that predict sensory outcome of motor commands and correct motor commands through internal feedback. A function of the parietal cortex is state estimation: to integrate the predicted proprioceptive and visual outcomes with sensory feedback to form a belief about how the commands affected the states of the body and the environment. A function of basal ganglia is related to optimal control: learning costs and rewards associated with sensory states and estimating the "cost-to-go" during execution of a motor task.

Over the last 25 years, a large body of experimental and theoretical work has been directed toward understanding the computational basis of motor control, particularly visually guided reaching. Roboticists and engineers largely initiated this work, with the aim of deriving from first principles some of the strikingly stereotypical features of movements observed in people and other primates. That is, they aimed to understand why we move the way that we do. The theories began to explain why in reaching to pick up a cup or in moving the eyes to look at an object, there was such consistency in the detailed trajectory of the hand and the eyes. In many ways, the approach was reminiscent of physics and its earliest attempts to explain regularity in motion of celestial objects except that the regularity was in our movements, and the search was for theories that explained our behavior. Here, we will summarize these theories and then link them to experimental findings in healthy subjects and in patients with neurological disease.

The computational problem of motor control

In 1954, Fitts published a short paper in which he reported that there were regularities in people's movements (Fitts, 1954). He asked volunteers to move a pen from one "goal region" to another as fast and accurately as they could. He found that the movement durations grew logarithmically as a function of the distance between the goals (figure 40.1). This relationship was modulated by two factors. One factor was the size of the goal region. As the goal region became smaller, movements slowed down. A second factor was the mass of the pen. People slowed their movements when they moved a heavier pen. To explain these results, consider that the target box was surrounded by two penalty regions, so it seems rational to aim for the center of the target box. What if the penalty region was only on one side? Now one should aim for a point farther away from the penalty region and not at the center of the target box (Trommershauser, Gepshtein, Maloney, Landy, & Banks, 2005). This is because movements have variability, and one will maximize reward (in terms of sum of hits and misses) if one take into account this variability. This variability explains the speed of movements in Fitt's experiment and the sensitivity to pen weight: Rapid movements are more variable than slow movements, so one should slow down if there is a need to be accurate. Moving heavier objects tends to increase movement variability, again requiring a reduced speed to maintain accuracy. Therefore in planning our movements, our brain takes into account movement variability because variability affects accuracy, which in turn affects our ability to acquire reward.

Harris and Wolpert (1998) began formalizing these ideas by linking variability and movement planning. They noted that larger motor commands required larger neural activity, which in turn produced larger variability owing to a noise process that grew with the mean of the signal. Therefore, motor commands carried an accuracy cost because the larger the command, the larger the standard deviation of the noise that rides on top of the force produced by the muscles (Jones, Hamilton, & Wolpert, 2002). Noise makes movements inaccurate.

In a sense, the theory restated the purpose of movements using language of mathematics: Be as fast as possible,

---

REZA SHADMEHR Laboratory for Computational Motor Control, Department of Biomedical Engineering, Johns Hopkins School of Medicine, Baltimore, Maryland.
JOHN W. KRAKAUER The Motor Performance Laboratory, Department of Neurology, Columbia University College of Physicians and Surgeons, New York, New York.

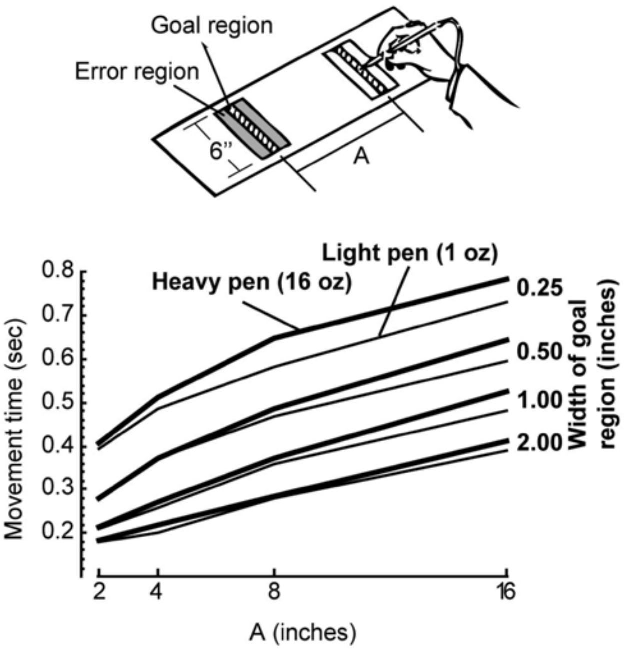

Figure 40.1 Accuracy constraints affect control of reaching. Volunteers were instructed to tap the two goal regions with a pen as many times as possible during a 15-s period. Movement time increased as the accuracy requirements increased (width of target region decreased) and as the weight of the hand-held pen increased. (Figure constructed from data in Fitts, 1954.)

while trying to be as accurate as the requirements imposed by the task. However, by doing so, it forced the theorists to think how one would actually achieve this optimality. Certainly, the solution to the problem could not be "hard wired."

First, costs and rewards of tasks are not constant. Take the simple saccade task in which an animal is given more reward for certain visual targets and less for others. Hikosaka and colleagues (Takikawa, Kawagoe, Itoh, Nakahara, & Hikosaka, 2002) examined eye trajectories when a monkey was asked to make saccades to various target locations. They noted that peak speeds tended to be higher and less variable when saccades were to rewarded target locations. Therefore, when the expected rewards of the task change, movement planning responds to these changes.

Second, the brain alters movement planning as the dynamics of the body or a tool change (e.g., the light versus heavy pens in figure 40.1). That is, the nervous system cannot rely on a motor plant that is time-invariant. Rather, it seems more reasonable that the nervous system should monitor these changes and form an *internal model* of the plant and/or the tool (Shadmehr & Mussa-Ivaldi, 1994). Indeed, maintaining performance in something as simple as a saccade or a reach probably requires constant adjustment of this internal model (Smith, Ghazizadeh, & Shadmehr, 2006; Kording, Tenenbaum, & Shadmehr, 2007).

Todorov and Jordan (2002) recognized that a key component of the problem was presence of feedback. One type of feedback is from sensory receptors that monitor the state of the body and the world. Another type of feedback is from internal models that monitor the motor output and predict their sensory consequences, effectively providing a form of internal feedback. Internal predictions can be made long before sensory feedback, making some very rapid movements such as saccades depend entirely on internal feedback (Chen-Harris, Joiner, Ethier, Zee, & Shadmehr, 2008). However, for longer movements, the two kinds of information would need to be combined to form a belief about the state of the body. Todorov and Jordan (2002) suggested that a more appropriate mathematical approach was to first describe the constraints of the task in terms of a function that included explicit terms for gains and losses and then maximize that function in the framework of *feedback control*. This new formulation was a breakthrough because it formally linked motor costs, expected rewards, noise, sensory feedback, and internal models into a single, coherent mathematical framework (see chapter 42 for a thorough introduction).

We summarize this framework in figure 40.2*A*. At the heart of the approach is the idea that we make movements to achieve a rewarding state. The rewards we expect to get and the costs we expect to pay determine the trajectory we choose to execute and how we will respond to sensory feedback. To make the "best" movement, our brain needs to solve three kinds of problems: We need to be able to accurately predict the sensory consequences of our motor commands (this is called *system identification*), we need to combine these predictions with actual sensory feedback to form a belief about the state of our body and the world (called *state estimation*), and then given this belief about the state of our body and the world, we have to adjust the gains of the sensorimotor feedback loops so that our movements maximize some measure of performance (called *optimal control*).

Here, we will suggest a specific computational neuroanatomy of the motor system (figure 40.2*B*). In this framework, the basal ganglia help to form the expected costs of the motor commands and the expected rewards of the sensory states. The cerebellum plays the role of predicting the sensory consequences of motor commands, that is, the expected changes in proprioceptive and visual feedback. The parietal cortex combines the expected sensory feedback with the actual sensory feedback, computing a belief about the current proprioceptive and visual states. Given the motor costs and expected rewards of the sensory states, the premotor and the primary motor cortex assign "feedback gains" to the visual and proprioceptive states, respectively, resulting in sensorimotor maps that transform the internal belief about states into motor commands.

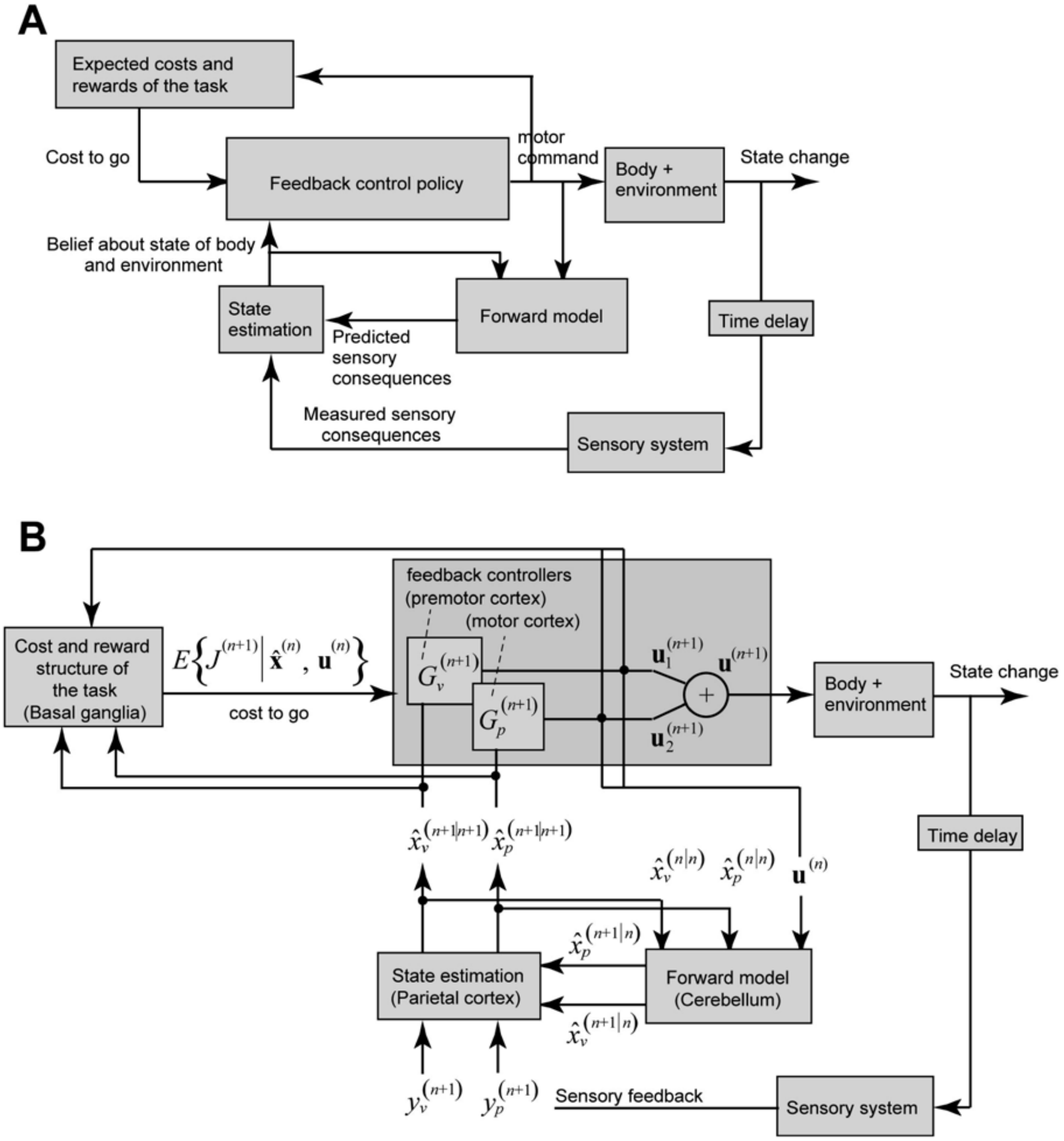

Figure 40.2 A schematic model for generating goal-directed movements. See the text for explanation of variables and box labels.

The computational problem in reaching

Let us use the well-studied reach adaptation paradigm to formulate the problem in the framework outlined in figure 40.2. What are the costs and rewards of a reaching task? Suppose that we are instructed to hold a tool and move it so that a cursor displayed on a monitor arrives at a target. If we accomplish this in a specific time period, we are provided a monetary reward, or juice, or perhaps a "target explosion." We can sense the position of the cursor $\mathbf{y}_v$ and the target $\mathbf{r}$ via vision and position of our arm $\mathbf{y}_p$ via proprioception. Through experience in the task, we learn that the objective is to minimize the quantity $(\mathbf{y}_v^{(t)} - \mathbf{r})^T (\mathbf{y}_v^{(t)} - \mathbf{r})$ at time $t = N$ after the reach starts (e.g., this is the time that the movement is rewarded if the cursor is in the target). Superscript T is the transpose operator. To denote the fact that this cost is zero except for time N, we write it as

$$\sum_{t=1}^{N} \left(\mathbf{y}_v^{(t)} - \mathbf{r}\right)^T Q^{(t)} \left(\mathbf{y}_v^{(t)} - \mathbf{r}\right)$$

where the matrix Q is a measure of our cost at each time step (which may be zero except at time N). That is, matrix Q specifies how important it may be for us to put the cursor in the target. If we value the reward, then we set this variable to be large.

There is also a cost associated with motor commands $\mathbf{u}$. This cost may reflect a desire to be as frugal as possible with our energy expenditure, or it may reflect the fact that the larger the motor commands, the larger the noise in the forces that are produced by the muscles, resulting in variability. This variability increases the difficulty in controlling the movement. As a result, we want to produce the smallest amount of motor commands possible. Now the total cost becomes

$$\mathcal{J} = \sum_{t=1}^{N} \left(\mathbf{y}_v^{(t)} - \mathbf{r}\right)^T Q^{(t)} \left(\mathbf{y}_v^{(t)} - \mathbf{r}\right) + \mathbf{u}^{(t)T} L \mathbf{u}^{(t)} \qquad (1)$$

where matrix L is a measure of the costs associated with the motor commands. The relative weight of Q and L is an internal measure of expected value of achieving the goal versus expected motor costs.

To be successful in this task (consistently arrive at the target in time), we need to find the motor commands that, on the one hand, are as small as possible and, on the other hand, are large enough to get the cursor to the target in time. To do so, we need some way to relate motor commands to their outcomes. This is called an internal model. For example, through observation, we learn that moving the tool moves the cursor on the screen. In particular, motor commands $\mathbf{u}^{(t)}$ are expected to produce proprioceptive and visual feedback $\hat{\mathbf{y}}^{(t)} = [\hat{\mathbf{y}}_v^{(t)}, \hat{\mathbf{y}}_p^{(t)}]$. These are the expected sensory consequences of our action. Here, we write this "internal model" as a linear function of motor commands:

$$\begin{aligned} \hat{\mathbf{x}}^{(t+1|t)} &= \hat{A}\hat{\mathbf{x}}^{(t|t)} + \hat{B}\mathbf{u}^{(t)} \\ \hat{\mathbf{y}}^{(t)} &= \hat{H}\hat{\mathbf{x}}^{(t)} \end{aligned} \qquad (2)$$

where $\hat{\mathbf{x}}^{(t|t)}$ represents the predicted state (of our body and the world) at time t given the sensory feedback up until that time, H is a transformation of those states to expected sensory feedback $\hat{\mathbf{y}}^{(t)}$ (i.e., proprioception and vision), and $\hat{\mathbf{x}}^{(t+1|t)}$ is predicted state at time $t + 1$ given the state and motor command at time t. Equation 2 describes an internal model of the dynamical system that we are trying to control. The actual dynamics of that system may be more complicated. For example, the motor commands may carry signal-dependent noise $\mathring{\mathbf{a}}_u^{(t)}$, that is, a noise in which the standard deviation grows with the size of the motor command. In general, there may be similar signal-dependent noises on our sensory system, $\mathring{\mathbf{a}}_y^{(t)}$. In sum, a reasonable representation of the stochastic system that we are trying to control might be written as

$$\begin{aligned} \mathbf{x}^{(t+1)} &= A\mathbf{x}^{(t)} + B\left(\mathbf{u}^{(t)} + \mathring{\mathbf{a}}_u^{(t)}\right) \\ \mathbf{y}^{(t)} &= H\left(\mathbf{x}^{(t)} + \mathring{\mathbf{a}}_y^{(t)}\right) \end{aligned} \qquad (3)$$

As motor commands are generated, we receives a continuous stream of sensory feedback $\mathbf{y}$. We combine the predicted sensory feedback with the observed quantities to form a *belief* about states:

$$\hat{\mathbf{x}}^{(t+1|t+1)} = \hat{\mathbf{x}}^{(t+1|t)} + K^{(t+1)}\left(\mathbf{y}^{(t+1)} - \hat{\mathbf{y}}^{(t+1)}\right) \qquad (4)$$

In this equation, the term $\hat{\mathbf{x}}^{(t+1|t+1)}$ is the belief state at time $t + 1$, given that we have acquired sensory information at that time. K is a mixing gain (or a Kalman gain) that determines how much we should change our belief on the basis of the difference between what we predicted and what we observed. Therefore, equation 2 describes how we make predictions about sensory feedback, and equation 4 describes how we combine the actual sensory observations with predictions to update beliefs about states.

Our task is to perform the movement in a way that maximizes our chances for reward. If equation 2 is an accurate model of how motor commands produce changes in the states, then we can use it as a set of constraints with which to minimize equation 1. Because there is noise in our system, the cost $\mathcal{J}$ in equation 1 is a stochastic variable. At each time point during a movement, the best that we can do is minimize the expected value of this cost, given the state that we believe to be in and the motor commands that we have produced: $E\{\mathcal{J}^{(t)} | \hat{\mathbf{x}}^{(t-1)}, \mathbf{u}^{(t-1)}\}$. The term $E\{\mathcal{J}^{(t)}\}$ reflects the expected value of the cost-to-go, that is, the total cost remaining in the current trial. The result is a feedback control "gain":

$$\begin{aligned} \mathbf{u}^{(t)} &= -G^{(t)}\hat{\mathbf{x}}^{(t|t-1)} \\ &= -G_p^{(t)}\hat{\mathbf{x}}_p^{(t|t-1)} - G_v^{(t)}\hat{\mathbf{x}}_v^{(t|t-1)} \end{aligned} \qquad (5)$$

The new variable G is a matrix that changes with time during a movement. It tells us how at time t, we can transform beliefs in sensory states (in terms of proprioception and vision) into motor commands so that we maximize performance in the remaining task time.

Some examples

As an example, consider a simple task first described by Uno, Kawato, and Suzuki (1989) and shown in figure 40.3*A*. The objective is to reach from point T1 to T2. In one condition, the subject is holding a lightweight tool that moves freely in air. In a second condition, the tool is attached to a spring that pulls the hand to the right. Without the spring, people reach in a straight line. This is the path that minimizes the cost. However, once the spring is attached, the straight path incurs substantially more motor costs than a curved path. The curved path is the one that subjects choose (Uno et al., 1989).

In our second example, the task is to move one's hand from one point to another in a given amount of time (450 ms), but now instead of a spring, there is a velocity-dependent field that pushes the hand perpendicular to its direction of motion. Before the field is imposed, the motion that mini-

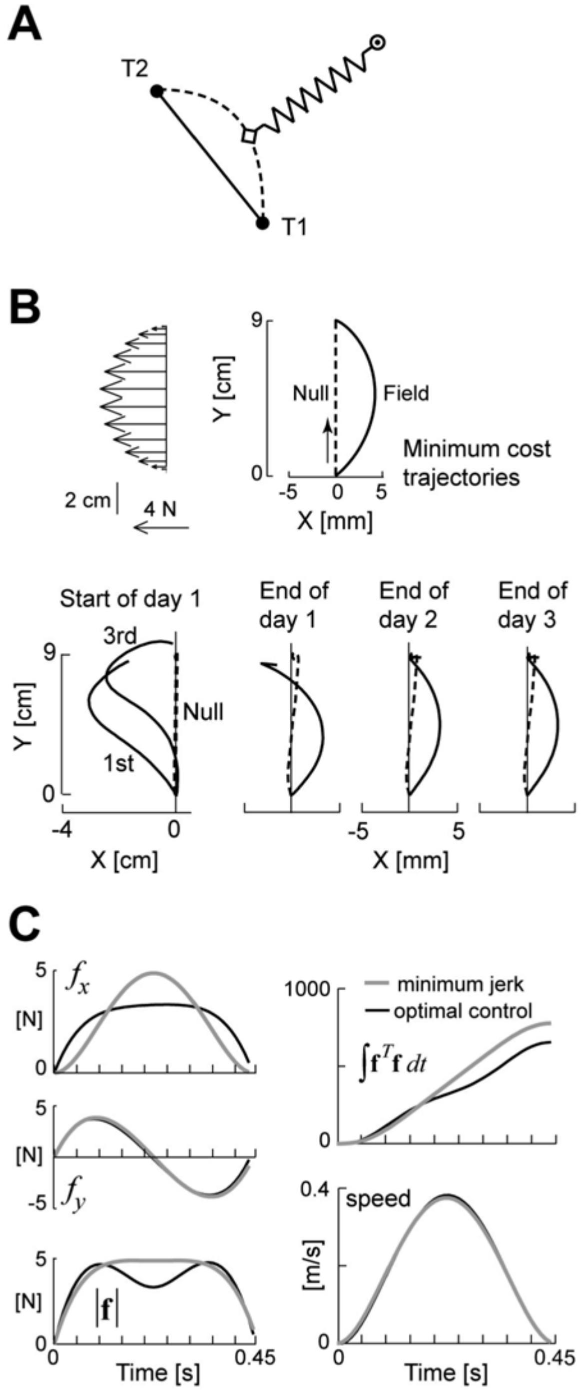

Figure 40.3 Task dynamics affect reach trajectories. (*A*) The task is to reach from point T1 to T2. In one condition, the reach takes place in free space (straight line). In another condition, a spring is attached to the hand. In this case, the subject chooses to move the hand along an arc. (*B*) A velocity-dependent force field pushes the hand perpendicular to its direction of motion. For example, for an upward movement, the forces push the hand to the left. The motion that minimizes cost of equation 1 is not a straight line but one that has a curvature to the right. The data show hand paths for a typical subject at start of training on day 1 and then at the end of training each day. Except for the first and third trials, all other trajectories are average of 50 trials. (*C*) A rationale for why a curved movement is of lower cost. The curves show simulation results on forces that the controller produces and speed of movement in the optimal control scenario of equation 1 and in a scenario where the objective is to minimize jerk. (*A* is redrawn from Uno et al., 1989. Data in parts *B* and *C* are from Izawa et al., 2008.)

mizes the cost (and maximizes probability of reward) is simply a straight line with a bell-shaped velocity profile. However, when the field is imposed, the solution is no longer a straight line (Izawa, Rane, Donchin, & Shadmehr, 2008). For example, if the field pushes the hand to the left, the policy that produces the least cost in terms of equation 1 is one that moves the hand slightly to the right of a straight line, resulting in a curved movement that appears to overcompensate for the forces (figure 40.3*B*). As subjects train, their hand paths converge to this curved trajectory. To see the rationale for this behavior, figure 40.3*C* plots the forces produced by the optimal controller and compares it to forces that must be produced if a mass is moving along a "minimum-jerk" trajectory. By moving the hand along a curved path, the optimal controller produces less total force: It overcompensates early into the movement when the field is weak but undercompensates at peak speed when the field is strongest. Therefore, the curved path actually produces less total force than a straight trajectory does. People produce similarly curved trajectories when they move in such fields (Thoroughman & Shadmehr, 2000).

The cerebellum: predicting sensory consequences of motor commands

According to the theory, we generate motor commands the basis of on beliefs about the state of our body and the environment (equation 5). This state estimate depends on two quantities: a prediction and an observation. The prediction comes from an internal model that uses a copy of the motor commands to estimate the state change that is expected to occur. The observation comes from the sensory system that provides a measure of those state changes. That is, our beliefs are not based on our observations alone. Rather, our beliefs are a combination of what we predicted and what we observed (Kording & Wolpert, 2004a; Vaziri, Diedrichsen, & Shadmehr, 2006).

Some movements are so fast that there is no time for the sensory system to play a role. A prominent example is control of saccades (rapid eye movements that move the eyes to a new location typically within 50–80 ms). Such movements are too brief for visual feedback to influence saccade trajectory. In fact, the brain actively suppresses visual processing during saccades to reduce the perception of motion (Thiele, Henning, Kubischik, & Hoffmann, 2002). Furthermore, proprioceptive signals from the eyes do not play any significant role in controlling saccade trajectories (Keller & Robinson, 1971; Guthrie, Porter & Sparks, 1983). Thus the brain must guide saccade trajectories in the absence of sensory feedback. How is this accomplished? A plausible solution is for the brain to use an internal estimate of the state of the eye, derived from a copy of ongoing motor commands (Robinson, 1975). This internal feedback probably accounts for the fact that variability at saccade initiation is partially corrected as the saccade progresses (Quaia, Pare, Wurtz, & Optican, 2000). That is, saccades are steered midflight via an internal feedback system (Chen-Harris et al., 2008).

What are the neural substrates of this internal feedback? The available evidence points to the cerebellum (Optican & Quaia, 2002; Optican, 2005). That is, the cerebellum appears to act as a forward model of the plant to produce midflight corrections. A simple experiment can test whether the cerebellum plays a role in predicting consequences of self-generated motor commands. Nowak, Timmann, and Hermsdorfer (2007) asked subjects to hold a force transducer that measures grip force, and then they attached a basket to the transducer. The experimenter dropped a ball into the basket. When the ball dropped, it exerted a downward force on the hand. The subject responded by squeezing the transducer so that it would not slip out of his or her hand. Because there are delays in sensing the impact of the ball, the grip response came about 100 ms after the ball's impact. Nowak and colleagues (2007) described patient HK, who did not have a cerebellum, owing to a very rare developmental condition. When the experimenter dropped the ball into the basket, both the healthy individuals and HK showed the delayed response. Therefore, the sensory feedback pathways appeared to be intact. In a subsequent trial, the subject (rather than the experimenter) dropped the ball. In a healthy individual, the brain can predict that the release of the ball will soon result in an impact that will increase the downward load. In anticipation of this event, the healthy individual squeezed the basket's handle harder around the time when the ball was released. HK, however, could not make this anticipatory adjustment. Rather, she responded to the perturbation in the same way that she responded when the ball was dropped by the experimenter. Therefore, the cerebellum appears to be required for the ability to predict the sensory consequences of motor commands (Wolpert, Miall, & Kawato, 1998).

The cerebellum and construction of internal models

It is not easy to make accurate predictions about the sensory consequences of motor commands; our muscles respond differently depending on their fatigue state, and our limbs move differently depending on whether we are holding a light or heavy object. To maintain accuracy of the predictions, our brain needs to learn from the sensory feedback and adapt its internal model. This adaptation can be simple, such as changing parameter values of a known structure (changing A, B, or H in equation 2), or complex, such as identifying the structure de novo (replacing the linear form of equation 2 with some nonlinear function). The cerebellum appears to be one of the crucial sites of this process.

Cerebellar damage often prevents individuals from learning how to use novel tools. For example, when subjects are asked to move the handle of a robotic tool to manipulate cursor positions, they may not be able to learn to compensate for forces generated by the robot (Maschke, Gomez, Ebner, & Konczak, 2004; Smith & Shadmehr, 2005) or to compensate for the novel visual feedback through a mirror (Sanes, Dimitrov, & Hallett, 1990). If the cerebellum is the crucial site for learning internal models, then it probably makes its contribution to control of reaching via its outputs to the thalamus, which in turn projects to the cerebral cortex. In humans, it is possible to reversibly disrupt this pathway. Essential tremor patients are occasionally treated with deep-brain stimulators that artificially disrupt the ventrolateral thalamus, improving their tremor. However, these patients learn the reach task better when the stimulator is turned off (Chen, Hua, Smith, Lenz, & Shadmehr, 2006). In contrast, patients with damage to the basal ganglia showed little or no deficit in adaptation with either the robot task (Smith & Shadmehr, 2005) or the mirror task (Agostino, Sanes, & Hallett, 1996; Gabrieli, Stebbins, Singh, Willingham, & Goetz, 1997). Therefore it seems quite likely that the cerebellum is a key structure that allows us to learn tool use.

Experiments show that the cerebellar damage causes abnormalities in adaptation to both kinematic (Tseng, Diedrichsen, Krakauer, Shadmehr, & Bastian, 2007) and force (Smith & Shadmehr, 2005) perturbations. One unifying concept is that the cerebellum may be the site of the internal model that predicts the sensory consequences of motor commands (equation 2). The output of the internal model could be used to generate a prediction error that drives adaptation and also be used to update a previous estimate of limb state. Support for this idea comes from a recent experiment in which transcranial magnetic stimulation was used to disrupt the lateral cerebellum in human subjects while they slowly moved their arm in preparation for a making a rapid reaching movements (Miall, Christensen, Owen, & Stanley, 2007). Reaching errors in initial direction and final finger position suggested that the reaching movements had been made from an estimated hand position that was approximately 140 ms out of date, consistent with a role for the cerebellum in iteratively updating limb state.

Learning the rewarding nature of sensory states

You might expect that a severely amnesic individual who was performing a novel task would have to be regularly reminded of the task's instructions. For example, if it is a reaching task, we might have to repeat "try to move the cursor to the target fast enough so it explodes." However, when we examined the severely amnesic patient HM on the standard reach adaptation task with the robot (Shadmehr, Brandt, & Corkin, 1998), after he had exploded a few targets, he no longer needed verbal reminders. The visual appearance of the target was enough for him to initiate a reaching movement. Strikingly, when he returned a few hours later (or the next day), he voluntarily reached for the robot handle and began preparing for onset of targets by moving the cursor to the center location (naïve individuals avoid touching the machine). It was clear that despite having no conscious recollection of having done the task before, some part of HM's brain recognized that the contraption was a tool that had a particular purpose: to manipulate cursors on a screen. This behavior suggested that during the first session, he implicitly learned the reward basis of the task (equation 1). (For HM, the target explosion triggered a childhood memory of going bird hunting. As he was performing the task and was able to get a target explosion, he would spend the next few minutes describing the memory in detail: the type of gun that he used, the porch in the rear of his childhood home, the terrain of the woods in his backyard, and the kinds of birds that he hunted.) What brain regions were involved in learning the rewarding nature of bringing the cursor to the target?

Experiments on action selection in rodents provide important insights into this question. For example, suppose that a rat is released into a pool of water from some random starting point. A platform is positioned in a specific location just below the water line and cannot be seen. The platform is always at the same location in the pool. Rats dislike being wet and will try to find a way to elevate themselves. The normal rat can learn to locate the platform position by paying attention to the visual cues that surround the pool. This requires learning a spatial map of where the platform is located with respect to the surrounding visual cues. With repeated swims, the animal learns a spatial map. This spatial map is analogous to a reward function that associates places in the pool with the likelihood of the platform (and therefore the likelihood of not having to be wet).

Once the map has been learned, the animal can find the platform regardless of where the rat is released into the water because the map is with respect to the cues on the walls. If the platform is removed, the normal animal will spend most of the time searching in the region where the platform should be. Sometimes, certain cues are rewarding no matter where they are located. Consider a pool where there are two hidden platforms: one that is large enough for the rat to mount and one that is too small. Both have a distinct visual cue associated with them: a little flag attached to each platform, each of a different color, sticking out of the water. Suppose that the flag attached to the large platform is red and the flag attached to the small platform is green. The platforms may be positioned in any part of the pool and will change from trial to trial. Therefore, in this experiment, the animal needs to learn that the red flag indicates the location of the suitable platform and is a rewarding object. In another version of the experiment, the large platform will always be located in a particular spatial location, but the flag on top of it will be a random color. In this version of the experiment, the animal needs to learn that it is not the color of the flag that is important, but the spatial location.

We see that there is a natural competition between the learning systems that might be involved in these two conditions: Is the platform in the same "place" as before (where place refers to a location in the spatial map), or is the platform always where the red flag is located? Packard and McGaugh (1992) performed both experiments by having their animals swim eight times per day for a number of days. They recorded the number of times the animals mounted the small platform and labeled these as errors. In the first experiment, in which reward was associated with the red flag, healthy animals gradually learned to swim to the red flag. Interestingly, animals with damage to the medial temporal lobe learned the task just as well as the healthy controls did. However, animals with damage to the caudate nucleus were much slower in learning the association. After days of training, they continued to attempt to mount the platform under the green flag. Therefore it appears that the ability to associate reward to stimuli regardless of its spatial location depends on the basal ganglia.

In the second experiment, in which reward was associated with a spatial location, healthy animals gradually learned to swim to that location and ignore the color of the flag. Animals with damage to the caudate nucleus performed similarly to the healthy controls. However, animals with damage to the medial temporal lobe were much slower in learning the association. Therefore the ability to associate reward to a spatial location depends on the medial temporal lobe.

Returning to our observations in HM, we would speculate that it was his basal ganglia that learned that if he were to place the cursor in the box on the screen and do so rapidly, a rewarding state would be experienced (explosions, which triggered a pleasant childhood memory). During the later sessions, the visual appearance of the machine and the act of holding its handle likely triggered a recall of this reward structure.

Effects of striatal damage on the assessment of movement costs and rewards

One of the striking features of damage to the human striatum is micrographia, an impairment of writing in which letters become very small and writing speed becomes slow.

This condition is most common in degenerative diseases of the basal ganglia such as Parkinson's disease (Van Gemmert, Teulings, & Stelmach, 2001). However, it can also occur with focal lesions. Consider patient FF, an individual who suffered an ischemic stroke in the left basal ganglia, in the head of the caudate nucleus and the anterior part of the putamen (Barbarulo, Grossi, Merola, Conson, & Trojano, 2007). When FF was asked to copy a four- or eight-letter string of characters, writing with the right hand was much smaller than with the left hand. Micrographia reflects an abnormal choice of speed and amplitude and is one manifestation of generalized slowing of movement (bradykinesia).

In the optimal control framework, there are no desired trajectories for our movements. Rather, the path is a result of a control policy (equation 5), which itself is a result of minimization of a cost (equation 1). The cost depends on two quantities: spatial accuracy (error cost) and required effort (energy cost). Accuracy requirements influence speed selection, due to the signal-dependent noise property of motor commands. The desired accuracy of a movement sets an upper limit on the maximum speed of a movement. The accuracy term of the cost function offers an explanation for the wealth of experimental data demonstrating speed-accuracy tradeoff in reaching movements. Normal movements, however, do not appear to be made at the limits imposed by the speed-accuracy tradeoff: We can reach for an object faster than usual without appreciable loss of accuracy. Although very little experimental data exist on spontaneous speed selection, the effort term of the cost function offers a potential explanation for this phenomenon; that is, perhaps micrographia is an indication of an abnormally high motor cost.

One of us recently tested this idea that in Parkinson's disease, there may be an abnormally high cost associated with motor commands (Mazzoni, Hristova, & Krakauer, 2007). We required healthy control subjects to make accurate reaching movements of specified speeds. As the required speed increased, subjects took longer (required more trials) to accumulate a set number of movements at the required speed. This reluctance to move faster could be explained by the increase in required energy as well as by the degradation of spatial accuracy and thus did not disambiguate the contribution of these two costs. We then compared the performance of patients with Parkinson's disease to that of control subjects in this task. Parkinson's disease patients demonstrated normal spatial accuracy in each condition but required more trials than controls to accumulate the required number of movements in each speed range. The patients' increased reluctance to execute movements requiring greater effort, in spite of preserved spatial accuracy, provided experimental demonstration of the contribution of energy cost to speed selection, independent of spatial accuracy. The implication is that bradykinesia results when striatal dysfunction changes the value of effort minimization (increased sensitivity to effort cost; L in equation 1) relative to that of accuracy optimization (error cost; Q in equation 1). Thus it appears that the basal ganglia either provides the motor motivation signal, which is then used to compute the cost-to-go elsewhere, or is where the cost-to-go is computed.

Parietal cortex damage and state estimation

Sometimes goal states change as the task is being performed. For example, when one reaches to pick up a pen, the pen may start rolling away. Healthy individuals have no problems adjusting their movements to compensate for this change. However, parietal patients show particular difficulties with this task. For example, if parietal damage impairs representation of visual states contralateral to the fixation, then motion of the goal state to this region during a movement impairs the ability to adjust the reach mid-flight. Grea and colleagues (2002) observed this phenomenon in a patient with bilateral posterior parietal cortex damage. The patient had no problems reaching to targets in central fixation. However, when the target shifted to the right at reach onset, the subject continued to reach to the original location of the target as if the target had not moved.

Disruption of the parietal cortex in healthy individuals can produce a similar phenomenon. Desmurget and colleagues (1999) provided a single pulse via a transcranial magnetic stimulator as the reach to the target began. On trials in which the target jumped, most of the participants had hand movements that disregarded the shift in the target location.

Let us examine these results in the framework of figure 40.2. The relevant state variables in this task include position of the limb (in proprioceptive and visual coordinates) and the position of the target (in visual coordinates). As motor commands are generated, the forward model should update its predicted state of the limb. Generally, we expect targets to remain stationary, and therefore the output of the forward model should continue to predict the target position. Together, these predictions represent the prior belief about the state of the body and the world. The sensory feedback from proprioception and vision is integrated with this prediction to make a posterior belief. When the target jumps, the novel sensory information needs to be integrated with the output of the forward model. If it is not, the reach will continue to the prior expectation of its location. The results noted above suggest that either this integration step is affected by damage or stimulation of the parietal cortex or that the sensory information outside the central fixation region cannot reach the integration step.

Limitations in applying the theory to biological motor control

This review of motor control has been written within the framework of optimal feedback control. At the heart of the theory is the conjecture that animals make voluntary movements in order to acquire the most reward while expending the least effort. However, the theory cannot make a behavioral prediction unless we can specify three kinds of information: (1) what the costs and rewards are; (2) what the constraints are, that is, dynamics of the task; and (3) what the mechanisms of state estimation are. In this review, we have chosen a specific set of equations to represent each kind of information. However, it is not difficult to find examples of behavior that are inconsistent with our formulation.

The cost that we wrote in equation 1 is perhaps the simplest possible cost function for goal-directed movements. How seriously can we take this specific representation? As demonstrated by attempts to reverse-engineer the cost (Kording & Wolpert, 2004b), the quadratic cost function should not be taken too seriously.

Consider a set of experiments that highlighted the importance of costs associated with postural stability, a quantity that we did not include in equation 1. Scheidt and Ghez (2007) explored a task in which continuous random noise perturbed the hand at rest. This constraint encouraged increasing the cocontraction levels of muscles. However, the noise was present only during the postural phase of the task and disappeared when subjects made a reaching movement. They found that if a kinematic perturbation required adaptation of the movement, the learning did not generalize to the postural phase at the end of the movement. They suggested that the control processes that moved the limb appeared distinct from control processes that set muscle activity levels during posture. If so, do these processes have separate costs? A recent study suggests that the answer is yes, the weighting of postural cost is flexible and can be determined by task context (Liu & Todorov, 2007).

Finally, consider an experiment by Jax and Rosenbaum (2007) in which they asked subjects to make arm movements to an array of 12 targets positioned in a 16-cm radius circle on a vertical screen. Targets were presented randomly, and in some trials, an obstacle was presented halfway between the start and the target. The same target was never shown twice in a row. Interestingly, whenever a no-obstacle trial followed an obstacle trial, subjects made curved rather than straight trajectories. However, the movements straightened out when a no-obstacle trial followed another no-obstacle trial. Why make a suboptimal curved trajectory when you see that there is no obstacle?

These results highlight a number of important problems with our framework. First, without knowing precisely the costs and rewards of a movement, it will not be possible to make quantitatively reliable predictions of behavior. Without a priori predictions, how can the theory be falsified?

Second, what are the timescales of optimization? Is optimization computed in the reaction time of each trial de novo? The timescale appears to be longer than a single trial, as exemplified by the example from Jax and Rosenbaum (2007). Certainly, new costs can be conjured up. For example, in this case, we can assume that finding feedback control gains that minimize a cost requires neural processing that itself has a cost, so it might be more efficient to allow the solution in one trial to linger on to influence the solution in the next trial. Or perhaps there is a cost in switching control policies.

Third, what is the timescale of system identification? Our body changes over multiple timescales. Muscles fatigue and recover quickly, objects are lifted and replaced rapidly, yet aging can produce gradual loss of motor neurons and transformation of muscle fibers. In other words, the parameters of the constraint equation and perhaps its structure are changing over multiple timescales. Unfortunately, we cannot make optimized movements unless we have an accurate set of constraint equations, that is, an accurate internal model. When we see a suboptimum movement, can we dissociate the effects of an inaccurate internal model from effects of an inaccurate cost function?

Finally, what is the alternative hypothesis to this theory? At this time, the alternative is another cost or constraint, not a fundamentally distinct theory. However, formalization of a theory is the key step that accelerates its evolution toward acceptance or rejection.

Conclusions

The relationship between theories and the neural machinery that implements them is still in the courtship stage, but despite the separation, it has begun to bear modest fruit; theories have informed the neural basis of motor control in patients, while lesion studies have informed the algorithms and representations that implement the computational theories. The result is the functional anatomy of voluntary movements outlined in figure 40.2*B.* In this framework, a role for the cerebellum is system identification, that is, predicting the changes in state that arise as a result of motor commands. A role for the parietal cortex is state estimation, in which predictions about sensory feedback are integrated with visual and proprioceptive observations to form beliefs about states of our selves and objects/people around us. The basal ganglia may play a role in computing a cost-to-go function, estimating value of states and costs of motor commands. Finally, once a goal state has been selected, motor cortical areas minimize this cost function and transform state estimates into motor output by formulating a feedback control policy.

ACKNOWLEDGMENTS The work was supported by National Institutes of Health (NIH) grants K02-048099 and R01-052804 to JWK and R01-037422 to RS.

REFERENCES

Agostino, R., Sanes, J. N., & Hallett, M. (1996). Motor skill learning in Parkinson's disease. *J. Neurol. Sci.*, *139*, 218–226.

Barbarulo, A. M., Grossi, D., Merola, S., Conson, M., & Trojano, L. (2007). On the genesis of unilateral micrographia of the progressive type. *Neuropsychologia*, *45*, 1685–1696.

Chen, H., Hua, S. E., Smith, M. A., Lenz, F. A., & Shadmehr, R. (2006). Effects of human cerebellar thalamus disruption on adaptive control of reaching. *Cerebral Cortex*, *16*, 1462–1473.

Chen-Harris, H., Joiner, W. M., Ethier, V., Zee, D. S., & Shadmehr, R. (2008). Adaptive control of saccades via internal feedback. *J. Neurosci.*, *28*, 2804–2813.

Desmurget, M., Epstein, C. M., Turner, R. S., Prablanc, C., Alexander, G. E., & Grafton, S. T. (1999). Role of the posterior parietal cortex in updating reaching movements to a visual target. *Nature Neurosci.*, *2*, 563–567.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *J. Exp. Psychol.*, *47*, 381–391.

Gabrieli, J. D. E., Stebbins, G. T., Singh, J., Willingham, D. B., & Goetz, C. G. (1997). Intact mirror-tracing and impaired rotary-pursuit skill learning in patients with Huntington's disease: Evidence for dissociable memory systems in skill learning. *Neuropsychology*, *11*, 272–281.

Grea, H., Pisella, L., Rossetti, Y., Desmurget, M., Tilikete, C., Grafton, S., Prablanc, C., & Vighetto, A. (2002). A lesion of the posterior parietal cortex disrupts on-line adjustments during aiming movements. *Neuropsychologia*, *40*, 2471–2480.

Guthrie, B. L., Porter, J. D., & Sparks, D. L. (1983). Corollary discharge provides accurate eye position information to the oculomotor system. *Science*, *221*, 1193–1195.

Harris, C. M., & Wolpert, D. M. (1998). Signal-dependent noise determines motor planning. *Nature*, *394*, 780–784.

Izawa, J., Rane, T., Donchin, O., & Shadmehr, R. (2008). Motor adaptation as a process of reoptimization. *J. Neurosci.*, *28*, 2883–2891.

Jax, S. A., & Rosenbaum, D. A. (2007). Hand path priming in manual obstacle avoidance: Evidence that the dorsal stream does not only control visually guided actions in real time. *J. Exp. Psychol. Hum. Percept. Perform.*, *33*, 425–441.

Jones, K. E., Hamilton, A. F., & Wolpert, D. M. (2002). Sources of signal-dependent noise during isometric force production. *J. Neurophysiol.*, *88*, 1533–1544.

Keller, E. L., & Robinson, D. A. (1971). Absence of a stretch reflex in extraocular muscles of the monkey. *J. Neurophysiol.*, *34*, 908–919.

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

Kording, K. P., & Wolpert D. M. (2004a). Bayesian integration in sensorimotor learning. *Nature*, *427*, 244–247.

Kording, K. P., & Wolpert D. M. (2004b). The loss function of sensorimotor learning. *Proc. Natl. Acad. Sci. USA*, *101*, 9839–9842.

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

Maschke, M., Gomez, C. M., Ebner, T. J., & Konczak, J. (2004). Hereditary cerebellar ataxia progressively impairs force adaptation during goal-directed arm movements: *J. Neurophysiol.*, *91*, 230–238.

Mazzoni, P., Hristova, A., & Krakauer, J. W. (2007). Why don't we move faster? Parkinson's disease, movement vigor, and implicit motivation. *J. Neurosci.*, *27*, 7105–7116.

Miall, R. C., Christensen, L. O. D., Owen, C., & Stanley, J. (2007). Disruption of state estimation in the human lateral cerebellum. *PLoS Biol.*, *5*, e316.

Nowak, D. A., Timmann, D., & Hermsdorfer, J. (2007). Dexterity in cerebellar agenesis. *Neuropsychologia*, *45*, 696–703.

Optican, L. M. (2005). Sensorimotor transformation for visually guided saccades. *Ann. NY Acad. Sci.*, *1039*, 132–148.

Optican, L. M., & Quaia, C. (2002). Distributed model of collicular and cerebellar function during saccades. *Ann. NY Acad. Sci.*, *956*, 164–177.

Packard, M. G., & McGaugh, J. L. (1992). Double dissociation of fornix and caudate nucleus lesions on acquisition of two water maze tasks: Further evidence for multiple memory systems. *Behav. Neurosci.*, *106*, 439–446.

Quaia, C., Pare, M., Wurtz, R. H., & Optican, L. M. (2000). Extent of compensation for variations in monkey saccadic eye movements. *Exp. Brain Res.*, *132*, 39–51.

Robinson, D. A. (1975). Oculomotor control signals. In P. Bachy-Rita & G. Lennerstrand (Eds.), *Basic mechanisms of ocular motility and their clinical implications* (pp 337–374). Oxford, UK: Pergamon.

Sanes, J. N., Dimitrov, B., & Hallett, M. (1990). Motor learning in patients with cerebellar dysfunction. *Brain*, *113*, 103–120.

Scheidt, R. A., & Ghez, C. (2007). Separate adaptive mechanisms for controlling trajectory and final position in reaching. *J. Neurophysiol.*, *98*, 3600–3613.

Shadmehr, R., Brandt, J., & Corkin, S. (1998). Time dependent motor memory processes in H.M. and other amnesic subjects. *J. Neurophysiol.*, *80*, 1590–1597.

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

Smith, M. A., Ghazizadeh, A., & Shadmehr, R. (2006). Interacting adaptive processes with different timescales underlie short-term motor learning. *PLoS Biol.*, *4*, e179.

Smith, M. A., & Shadmehr, R. (2005). Intact ability to learn internal models of arm dynamics in Huntington's disease but not cerebellar degeneration. *J. Neurophysiol.*, *93*, 2809–2821.

Takikawa, Y., Kawagoe, R., Itoh, H., Nakahara, H., & Hikosaka, O. (2002). Modulation of saccadic eye movements by predicted reward outcome. *Exp. Brain Res.*, *142*, 284–291.

Thiele, A., Henning, P., Kubischik, M., & Hoffmann, K. P. (2002). Neural mechanisms of saccadic suppression. *Science*, *295*, 2460–2462.

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

Todorov, E., & Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nat. Neurosci.*, *5*, 1226–1235.

Trommershauser, J., Gepshtein, S., Maloney, L. T., Landy, M. S., & Banks, M. S. (2005). Optimal compensation for changes in task-relevant movement variability. *J. Neurosci.*, *25*, 7169–7178.

Tseng, Y. W., Diedrichsen, J., Krakauer, J. W., Shadmehr, R., & Bastian, A. J. (2007). Sensory prediction errors drive cerebellum-dependent adaptation of reaching. *J. Neurophysiol.*, *98*, 54–62.

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 Gemmert, A. W., Teulings, H. L., & Stelmach, G. E. (2001). Parkinsonian patients reduce their stroke size with increased processing demands. *Brain Cogn.*, *47*, 504–512.

Vaziri, S., Diedrichsen, J., & Shadmehr, R. (2006). Why does the brain predict sensory consequences of oculomotor commands? Optimal integration of the predicted and the actual sensory feedback. *J. Neurosci.*, *26*, 4188–4197.

Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends Cogn. Sci.*, *2*, 338–347.