



Published in final edited form as:

Adv Exp Med Biol. 2013 ; 782: 1–21. doi:10.1007/978-1-4614-5465-6_1.

Model-Based and Model-Free Mechanisms of Human Motor Learning

Adrian M. Haith and

Department of Neurology, Johns Hopkins University, Baltimore, MD 21205, USA

adrian.haith@jhu.edu

John W. Krakauer

Departments of Neurology and Neuroscience, Johns Hopkins University, Baltimore, MD 21205,

USA jkrakau1@jhmi.edu

Introduction

In laboratory settings, motor learning has typically been studied in the context of adaptation paradigms in which subjects must learn to compensate for a systematic perturbation—either some manipulation of visual feedback (Krakauer et al. 2000) or a change in the dynamics of the motor apparatus, e.g., a force applied to the hand (Shadmehr and Mussa-Ivaldi 1994), Coriolis forces induced by rotation of the body (Lackner and Dizio 1994), or an inertial load attached to the arm (Krakauer et al. 1999). What is typically observed in these tasks is a monotonic improvement in performance that is initially rapid, and then slows to an asymptote close to initial baseline levels of performance. The progress of learning is well described by exponential fits, implying that the amount of improvement on each trial is proportional to the error (Thoroughman and Shadmehr 2000; Donchin et al. 2003). This kind of fast, trial-by-trial reduction in systematic errors is typically referred to as *adaptation*. The term adaptation has been used in some cases to imply a particular mechanism of learning; however, we will adhere to a behavioral definition (as a gradual reduction in error following an abrupt change in conditions) and describe potential underlying learning mechanisms in more computational terms. As we will argue, learning in adaptation paradigms is likely predominantly mediated by a specific learning mechanism that is based on changing an internal forward model.

Not all motor learning falls under our behavioral definition of adaptation. Often one learns to synthesize entirely novel movements even when there is no perturbation, e.g., learning to swing a golf club, hit a tennis serve, balance a pole, or drive a car. Although this kind of learning corresponds more closely to everyday usage of the term “motor learning,” it has hardly been studied in laboratory settings. The few exceptions typically involve learning to manipulate an unfamiliar, possibly complex virtual object (Carmena et al. 2003; Mosier et al. 2005; Nagengast et al. 2009; Sternad et al. 2011). In these kinds of tasks, subjects progress from initial incompetence to a high degree of proficiency, even approaching theoretically optimal behavior. However, performance improvements are far slower than in adaptation paradigms: while tens of trials are usually enough to reach asymptote after a systematic perturbation is introduced, performance in these more complex tasks continues to improve over hundreds of trials and even across days. This slow improvement is not entirely due to the unfamiliarity of the task. Even in much more simple tasks that involve

maneuvering a cursor along a constrained path (Shmuelof et al. 2012) or through a series of via points (Reis et al. 2009), overall variability in task performance reduces substantially over days of practice, even though subjects immediately exhibit near-perfect performance at slow speeds. It appears that a qualitatively different kind of learning may be occurring in these tasks—one that is not reliant on compensating for the highly salient errors that are present in adaptation settings, but instead is associated with incrementally improving the quality of one's movements with practice. We define this long-term reduction in movement variability as *skill learning*. It is not currently clear whether adaptation, skill, and learning to control external objects draw upon identical, overlapping, or entirely different neural mechanisms.

In this review, we argue for the existence of two distinct mechanisms underlying motor learning: (1) a model-based system in which improvements in motor performance occur indirectly, guided by an internal forward model of the environment which is updated based on prediction errors, and (2) a direct, model-free system in which learning occurs directly at the level of the controller and is driven by re-inforcement of successful actions. These distinct learning systems are each suited to different tasks and as such are complementary to one another. Model-based processes are likely to predominate in adaptation paradigms, and model-free processes predominate in skill tasks. However, we argue that both can contribute to learning in any given task.

Theory: Model-Based and Model-Free Approaches to Learning Control Policies

We adopt a general definition of motor learning as the process of improvement in execution of a task according to some chosen measure of performance such as increased chance of success or decreased effort (or potentially a combination of the two). Formally, we describe the state of learning in terms of a control policy π mapping current states, stimuli and time to motor commands u_t ,

$$u_t = \pi(x_t, s_t, t).$$

This general framework can encompass multiple levels of description. A control policy could describe selection of a single action per trial or describe an ongoing stream of motor commands in continuous time according to the instantaneous state. The motor commands u_t could model a high-level decision such as which direction to move the hand or a low-level decision such as which muscles to activate and when. The stimulus s_t would typically correspond to an observed target location and the state x_t would reflect the state of the motor plant. Any systematic, experience-driven change in this control policy can be described as motor learning. The quality of each potential control policy can be quantified in terms of the expected outcome value, i.e., the average performance that would be expected to be obtained when following that control policy for a given task. In studying motor learning, we study the process whereby individuals use experience to improve their control policy.

The optimal policy will depend on two specific things: (1) the structure of the task, i.e., which states are associated with valuable or successful outcomes and what costs may be associated with different states or actions and (2) the dynamics of the motor apparatus and environment, i.e., how do motor commands affect the state. In most motor control paradigms we would generally expect that the structure of the task is unambiguous; however, in general it may be that neither the task structure nor the dynamics dynamics is known precisely.

This general framework and the problem of determining suitable actions in an uncertain environment based on ongoing observations is precisely the theoretical problem studied, at a more abstract level, in the field of *reinforcement learning* (Sutton and Barto 1998). At the heart of reinforcement learning theory is the notion of the *value function* $V(x_t, t)$ which reflects, for a given control policy, the total future reward that can be expected to be gained given the current state and time. The goal of reinforcement learning is to determine the optimal value function—from which the optimal policy follows straightforwardly.

Different variants of reinforcement learning differ in exactly what kind of value function is represented and how this value function is updated based on experience. In particular, two distinct computational strategies have emerged for using experience to update estimated values and thereby determine optimal control policies. The first approach is to use experience to build models of the dynamics of the motor apparatus and environment and the structure of the task, and compute the value function based on these models (Fig. 1.1a). This approach is termed *model-based learning*. Note that model-based learning of this kind is very different from what most people understand intuitively by the term ‘reinforcement learning.’

A second approach, which accords with most people's informal or colloquial use of the term ‘reinforcement learning,’ is to learn the value function directly through a process of trial and error—explore the space of potential actions in each state and keep track of which states and actions lead, either directly or indirectly, to successful outcomes (Fig. 1.1b). This approach is often termed *model-free*, in contrast to model-based approaches. Other learning strategies are clearly possible besides the model-free and model-based approaches described here. However, these represent the most common approaches.

While model-free strategies clearly work and can in certain cases be shown theoretically to be guaranteed to converge upon optimal behavior (Sutton and Barto 1998), learning by trial and error is typically very slow in terms of the number of attempts necessary before a good policy can be acquired, even in relatively simple environments. Model-based learning, by contrast, makes the best possible use of all observations. Any information acquired about the outcome of a particular action is retained and can influence planning of future movements, regardless of whether that action led to success or not. Model-based methods also allow more principled generalization. If the reward structure of the task changes in a known way (e.g., the target moves to a new location), an appropriate new control policy can be computed based on the model of the dynamics that was built in the context of the previous task.

The major disadvantage with a model-based approach is that although the value of any state/action pair can in principle be computed exactly, it can be prohibitively computationally intensive to do so. Existing methods for computing the optimal policy typically involve either dynamic programming—a backward iteration through time to exhaustively compare all possible paths to the target and identify the best ones, or some iterative sequence of approximations to the value function or policy that converge upon a local optimum. Details of these methods are beyond the scope of this chapter (but for excellent introductions see Bertsekas 1996; Sutton and Barto 1998; Todorov 2007).

The complexity associated with computing optimal value functions and policies need not preclude biological systems from utilizing some form of model-based control. In certain very simple scenarios, it can be trivial to compute the optimal policy given a specification of the task and plant e.g., if the action on a given trial is simply the aiming direction for a particular movement, then a model-based solution to a rotational perturbation simply amounts to subtracting an estimate of the rotation angle from an observed target angle. The

feasibility of the model-based approach therefore largely depends on the nature of the task. Even if the computations are simple, however, errors may still arise from accumulation of noise that inevitably accompanies computations in biological systems (McGuire and Sabes 2009).

Model-free approaches, by contrast, require only relatively trivial computations because experiences lead directly to changes in the controller. Unlike a model-based approach, there is no intermediate forward model representation and no calculation required to transform a forward model into a control policy. In the long-run, model-free approaches tend to deliver superior performance on a particular task because they do not rely so heavily on noisy computations each and every time a movement must be made. The disadvantage is that the scope of the learned control policy is restricted to the task performed during learning. Even if the reward structure of the task changes in a known way, one must start from scratch (or at least from some previous but incorrect control policy). This is in sharp contrast to the flexibility offered by model-based learning.

In summary, if one wants to learn a good control policy in an uncertain environment, a model-based learning strategy is, in principle, the most powerful and flexible approach but requires unwieldy computations. Direct, model-free approaches rely only on simple computations but can require far more training (exploration) before they lead to a competitive policy. What learning strategy do animals use when placed in a situation where they must learn what to do? The contrasting ways in which model-based and model-free learning mechanisms should be expected to generalize to novel scenarios can act as hallmarks that potentially allow us dissociate an animal's learning strategy based on observing its behavior.

Model-Based and Model-Free Learning in Operant Conditioning Paradigms

In a situation where an animal must learn what actions will lead to reward, such as a rat navigating through a maze to find food, it seems that animals adopt both model-based and model-free learning mechanisms in parallel (Daw et al. 2005). Although any given control policy could be arrived at by either model-based or model-free strategies, these two modes of control can be dissociated by changing the reward structure of the task. In rodents this is typically achieved by stimulus revaluation. For instance, imagine examining the behavior of thirsty rats in a maze that they had learned while they were hungry and seeking food. Under a model-free approach, the thirsty rat will have no way of knowing how to obtain water and will likely either behave like a naïve rat, or rely on the same policy that led to reward while hungry. A model-based approach, by contrast, will enable the rat to flexibly change its behavior immediately in line with its new objective of finding water instead of food (provided, of course that it had previously explored the maze sufficiently to have found the location of the water). In practice, rewards are typically revalued either by satiating the animal prior to the task or, more drastically, pairing a familiar food with a strongly aversive stimulus (e.g., poison).

Behavior in such devaluation paradigms has been studied extensively, leading to a classical division between *goal-directed* behavior, in which animals are sensitive to reward devaluation, and *habitual* behavior in which they are not (Killcross and Coutureau 2003; Balleine and O'Doherty 2010; van der Meer and Redish 2011). Behavior tends to be goal-directed early in learning but becomes more habitual later on (Balleine and O'Doherty 2010). These differences in behavior can be interpreted in terms of reinforcement learning: goal-directed behavior can be understood as model-based, while habitual behavior is model-free (Daw et al. 2005; Dayan 2009). The transition from goal-directed to habitual with experience can even be explained as an evolving, intelligent trade-off between the advantages of each strategy.

Remarkably, these alternative model-based and model-free strategies are neurally dissociable. Lesions to distinct regions of prefrontal cortex can isolate one pattern of behavior or another in hungry rats (Balleine and Dickinson 1998; Killcross and Coutureau 2003). Sequential decision-making tasks in humans have revealed that their behavior can similarly be decomposed into model-based and model-free components (Fermin et al. 2010; Gläscher et al. 2010), while fMRI reveals that these components have distinct underlying neural substrates (Gläscher et al. 2010).

The kinds of control tasks that we are primarily concerned with in this review are quite different from the problem that a rat faces in a maze. In decision-making tasks it is the high-level choice of which path to follow at a junction that is of interest. The low-level movements that register this decision are considered incidental. In motor control, however, it is precisely these low-level movements that are of interest. Critically, control of movements can be cast within the same broad theoretical framework used to describe decision-making. The only differences are that movements of the eyes and limbs occupy a space of potential states and actions that is continuous and potentially high-dimensional, and decisions must be made in continuous time. Nevertheless, the same considerations for solving the general problem apply as in more discrete domains. In particular, both model-based and model-free learning strategies are possible and have similar advantages and disadvantages as in discrete domains. We will argue that, as in the case of rodent decision-making, both strategies are employed by the motor system for continuous control of movement. The underlying neurophysiology may, however, be quite different for the motor system as compared with the discrete action selection paradigms studied in rodents.

Model-Based Motor Learning

Forward models—neural networks which generate predictions about future states of the motor system given a current state and an outgoing motor command—have long been posited to be utilized by the motor system (Wolpert and Miall 1996). Model-based learning has become a dominant framework for understanding human motor learning, with arguably the majority of theories of motor learning assuming a model-based perspective (Shadmehr and Krakauer 2008; Shadmehr et al. 2010). The proposed advantages of maintaining a forward model are twofold: (1) A forward model allows for faster and more precise estimation of the state of the body and/or environment, and (2) Forward models may aid in planning future movements by directing changes in the controller itself, i.e. they may participate in model-based control. While (1) has by now become a relatively uncontroversial claim, (2) is much more difficult to establish.

Before assessing the case for model-based learning in the motor system, we briefly disambiguate model-based learning from learning involving *inverse models*. The simplest kind of controller considered in motor learning theories is a static mapping from a desired outcome to a single action. Such controllers have been referred to as *inverse models* since they are the direct inverse of the forward model. However, inverse models are not really “models” in the true sense of the word—they do not provide an internal representation of any process occurring in the outside world. It is more accurate to think of inverse models as simple control policies. An inverse model control policy can be arrived at in a model-based manner by first learning a forward model and then inverting it (Jordan 1992). Alternatively, changes to an inverse model could be driven directly by task errors (Thorouhman and Shadmehr 2000). We would not describe such learning as model-based, however, since the learning occurs directly at the level of the controller rather than via a forward model representation of the task or plant. Learning of this kind is only really feasible in simple, single-time step scenarios.

Theories based on the notion of inverse models are fairly limited in scope. More generally, motor control is described in terms of time-dependent feedback control policies (Todorov and Jordan 2002). In this context, there is no way to directly update the control policy based on performance errors. By contrast, model-based learning is a very general approach to obtaining a good control policy that is applicable to any problem that can be framed as a Markov decision process. The only limitation to model-based learning is being able to gather enough information to build the model.

Nothing is presently known about the neural computations that underlie the translation of knowledge about the environment in the form of a forward model into a control policy. However, even though the potential mechanisms underlying model-based control processes are poorly understood, this understanding is not necessary to establish whether or not it occurs. Here, we focus on reviewing the evidence at the behavioral level for the existence of forward models and their involvement in motor learning.

The Cerebellum and Forward Models

The cerebellum has long been implicated in motor control and coordination and has emerged as the most likely neural substrate of putative internal models (Bastian 2006; Shadmehr and Krakauer 2008; Wolpert et al. 1998). Patients with hereditary cerebellar ataxia or lesions to the cerebellum have general difficulties in coordinating movement and are grossly impaired in adaptation tasks (Martin et al. 1996; Maschke et al. 2004; Smith and Shadmehr 2005; Tseng et al. 2007; Synofzik et al. 2008; Rabe et al. 2009; Criscimagna-Hemminger et al. 2010; Donchin et al. 2011). There are many potential roles for the cerebellum in learning that might give rise to such an adaptation deficit in cerebellar ataxia. The cerebellum may, for instance, compute an inverse model that directly maps desired outcomes to actions (Medina 2011). We argue here, however, that the adaptation deficit following cerebellar damage stems from an inability to learn forward models.

Neurophysiological recordings from the cerebellum show that Purkinje cell simple spike activity reflects the kinematics of movement, and not the motor commands required to achieve the kinematics (Pasalar et al. 2006). This finding clearly demonstrates that the output of the cerebellum is not directly related to motor output, as would be predicted if the cerebellum were computing an inverse model or otherwise contributing directly to control. Furthermore, Purkinje cell activity during movement precedes the actual kinematic state of the limb (Roitman et al. 2005). So this activity in the cerebellum does not simply reflect a reporting of sensory feedback—instead it appears that the cerebellum implements an internal forward model that predicts the kinematic or sensory consequences of motor commands before that information actually becomes available from the periphery.

Numerous studies have argued from a behavioral standpoint that an estimate of state from a forward model underlies state-specific feedback corrections during movement (Ariff et al. 2002; Chen-Harris et al. 2008; Wagner and Smith 2008; Munuera et al. 2009). This process appears to be cerebellar-dependent (Miall et al. 2007; Xu-Wilson et al. 2009). Together with the above-mentioned neurophysiological findings, these studies make a strong case that the cerebellum generates predictions about future motor states on the basis of outgoing motor commands, and that these predicted states are made available to an already-learned feedback controller that guides ongoing execution of a movement. While this constitutes model-based control of sorts (Mehta and Schaal 2002), in this article we are more interested in the question of whether a forward model brings about changes in the controller, rather than influencing control only through estimates of state. Nevertheless, if forward models exist and can be used to guide online feedback control, it perhaps makes it more likely that the same forward models might participate in planning feedforward control.

Evidence for Forward Model Involvement in Feedforward Control

An often-cited instance where predictions of a forward model are claimed to influence feedforward control (as opposed to only feedback control) is in compensating for the consequences of one effector's actions on another—for instance stabilizing one hand holding a load while removing that load with the other hand or increasing grip force on an object to prevent slippage when accelerating it upward. It has often been argued that such *anticipatory control* is possible because of a forward model that predicts adverse consequences of an upcoming action before it has happened, enabling an appropriate compensation to be planned and executed concomitantly (Wolpert and Miall 1996; Flanagan and Wing 1997; Wolpert et al. 2011). Although the use of a forward model could, in principle, enable this kind of anticipatory control, coordination *per se* is no proof of the existence of forward models. Anticipatory control is simply a feature of a good control policy and there is no way of knowing how this controller may have been arrived at simply by observing it in action. Good coordination could have been learned via model-free mechanisms through trial and error.

Studies of anticipatory control in cerebellar ataxia patients offer some clues as to the nature of anticipatory control. Interestingly, cerebellar ataxic patients demonstrate intact coordination in manual unloading tasks (Diedrichsen et al. 2005) and exhibit intact modulation of grip force with varying load forces (although baseline grip forces are abnormal) (Rost et al. 2005), suggesting that forward models are not at all a prerequisite for performing coordinated movement. Cerebellar patients do, however, show impairment in learning novel anticipatory adjustments (Nowak et al. 2004; Diedrichsen et al. 2005). This suggests that initial acquisition of anticipatory control is facilitated by a forward model that can predict the consequences of the actions of one effector on the goals of another but, with prolonged practice, coordinated control eventually becomes independent of the forward model. To put it another way, there may be a transition from model-based to model-free mechanisms.

The notion of model-based learning implies that improvements in performance are driven by errors in the prediction of a forward model. One plausible alternative to this idea is that adaptation is driven by the feedback corrections one makes to correct errors, rather than by the errors themselves (Kawato and Gomi 1992). This does not appear to be the case for reaching movements, however: learning rates in adaptation tasks are identical whether or not feedback corrections are allowed during movement (Tseng et al. 2007). Similarly, corrective saccades do not appear to be necessary to adapt saccade amplitude (Wallman and Fuchs 1998).

Although not driven by corrective movements, adaptation may not necessarily be driven by prediction errors of a forward model. If control is mediated by an inverse model, changes to a control policy could be driven directly by task errors, without any need for a forward model. In most cases, task errors and prediction errors are closely aligned. In certain cases, however, performance errors and prediction errors can be dissociated. For instance, saccades to visual targets usually tend to fall slightly short of the target, but this shortfall does not lead to an increase in saccadic gain as one would expect if it were induced through a target jump. In fact, if the target is surreptitiously jumped mid-saccade such that the eye lands perfectly on the target every time, then saccadic gain actually begins to decrease despite the absence of performance errors (Wong and Shelhamer 2011). Indeed it is even possible for adaptation to occur in the opposite direction from a task error. This provides compelling evidence that prediction errors and not task errors are what drive motor adaptation.

A similar, even more striking result can be found for reaching movements. In a study by Mazzoni and Krakauer (2006) (Fig. 1.2a), subjects were exposed to a 45° rotation of visual

feedback but were also provided an explicit strategy to counter the rotation: simply aim to an adjacent target deliberately spaced at a 45° separation from the true target. Initially, subjects were able to flawlessly implement the strategy and hit the target. However, performance rapidly began to drift away from the target in the direction of the perturbation despite the fact that the task was being performed without errors. It therefore does not seem to be task error *per se* that drives adaptation, but discrepancies between predicted and observed behavior. Interestingly, this drift effect does not persist indefinitely—after prolonged exposure, subjects begin to reduce their errors again, suggesting that there is some component of learning that acts to close task errors rather than prediction errors (Taylor and Ivry 2011). When patients with cerebellar ataxia are given an explicit strategy, they are able to successfully maintain performance without undergoing any drift in performance (Taylor et al. 2010) (Fig. 1.2b). Thus the adaptation deficit in cerebellar ataxic patients is due to a reduced sensitivity to prediction errors not task errors.

The idea that adaptation is mediated by changes in predictions about the consequences of one's actions can be tested more directly through paradigms that ask subjects to estimate where they perceived their hand to have moved during a reach. Although such assays inevitably contaminate forward model-based predictions with actual visual and proprioceptive sensory experiences, a number of interesting results have been obtained using this approach. Following exposure to rotated visual feedback, healthy subjects undergo a corresponding change in their perceived hand path during movement (Synofzik et al. 2006). Cerebellar ataxic patients show no such perceptual changes (Synofzik et al. 2008; Izawa et al. 2011). These results support the idea that changes in a forward model, which presumably lead to the changes in predicted hand position, are a prerequisite for adaptation.

In summary, adaptation is driven by prediction errors and not by task errors or online motor responses to correct those errors. Exposure to rotated visual feedback leads to a shift in perceived hand location during movement. In patients with cerebellar ataxia, sensory prediction errors do not result in changes in feedforward control in future trials and do not lead to changes in perceived hand position. We believe that the most parsimonious explanation for all of these results is that the cerebellum computes an internal forward model that predicts the consequences of motor commands and that this forward model influences feedforward control of future movements.

Generalization of Learning Across Tasks

A final thread of evidence that has been cited in support of model-based control frameworks concerns generalization. Human subjects exhibit a high degree of generalization of learned compensation for a perturbation to a new movement (Shadmehr and Mussa-Ivaldi 1994; Krakauer et al. 2000). While this generalization is consistent with the idea of model-based control, it is important to bear in mind that model-free learning will also be expected to exhibit some degree of generalization—only in this case the generalization will be of a learned control policy, rather than of an internal model. The amount of generalization across states will be entirely determined by the underlying representation. There is no specific reason why one should expect model-based learning to generalize more broadly across states than model-free. However, subjects trained on a visuomotor rotation with full vector error (presumably engaging primarily model-based mechanisms) do generalize more broadly than subjects who learned to compensate the same perturbation but were given only binary feedback about the success or failure of their movements (presumably relying on model-free learning) (Izawa and Shadmehr 2011).

A more concrete dissociation between model-based and model-free learning mechanisms is the extent to which learning should transfer across tasks within the same workspace—for instance tracking a cursor along a curved path versus making point-to-point reaches. This

form of generalization across tasks is directly analogous to the reward devaluation protocols that dissociate model-based from model-free action selection processes in rodents (Daw et al. 2005)—in both cases the reward structure of the task is altered but the consequences of actions remain the same. A number of studies have examined generalization of learning from a redundant task, in which the perturbation is task-irrelevant, to a nonredundant task. For example, Schaefer et al. (2012) had subjects make reaching movements to a point anywhere on a circular arc while imposing a rotation of visual feedback. This rotation did not compromise task performance, since subjects still easily landed on the arc as required. The rotation, however, did lead to sensory prediction errors. In subsequent catch trials toward a single target, subjects showed significant aftereffects, supporting the idea that learning was driven by sensory prediction errors rather than by task errors. An identical pattern was found for subjects who adapted to a visual amplification of movement extent (gain increase) while performing an analogous task in which reach direction mattered, but reach extent did not.

Interestingly, although perturbations led to significant aftereffects even when they produced errors only along task-irrelevant dimensions, the aftereffects found by Schaefer and colleagues were significantly smaller than when the same perturbations were task-relevant. A similar study by Synofzik and colleagues (2006) found similar partial transfer of a visual rotation from reaching to an arc to reaching to a point. In that study, additional probe trials measured perceptual changes accompanying learning. Significant changes in perceived hand position during reaching were found. However, although the mean perceptual changes and mean generalization were of similar magnitude, individual subjects' perceptual changes were not predictive of the reaching behavior. Incomplete transfer of adaptation across tasks is not limited to these two examples. Adaptation to a shift in visual feedback during a manual tracking task generalizes only partially to subsequent reaching movements (Simani et al. 2007). In force field adaptation, subjects who learn a force field while performing a series of point-to-point reaching movements show partial but incomplete generalization of their learning when subsequently asked to make circular movements (Conditt et al. 1997).

The finding that learning is consistently seen to transfer across tasks is consistent with model-based learning. However, the variation in the extent of learning under task-relevant and task-irrelevant conditions suggests that learning might not be purely driven by changes in a forward model. On the one hand, one could interpret these results in terms of task-specific internal models mediating task-specific model-based controllers. However, the notion of task-specific internal models rather defeats one of the primary benefits of a model-based approach to control: the ability to flexibly generalize knowledge about the environment across tasks. It perhaps seems more parsimonious to suggest that task specificity of learning arises because of the task specificity of components of learning that are independent of internal models.

Evidence for Model-Free Learning

Model-based learning implies that the motor system learns to compensate for systematic perturbations by first identifying the dynamics of the system being controlled through a forward model (likely in the cerebellum), then somehow translating this knowledge into a control policy in the motor cortex. We have argued that this kind of mechanism can parsimoniously account for a variety of experimental results. However, not all aspects of motor learning are well explained by such a model-based framework. Here we outline the evidence that the motor system also relies on direct, model-free learning of actions.

As we have described above, patients with cerebellar ataxia are severely impaired in compensating for systematic perturbations. According to a model-based interpretation, this

inability arises from a primary deficit in the ability to learn or update an internal forward model describing the perturbation (either explicitly, or as an adjustment to baseline forward models). Perhaps surprisingly, however, cerebellar ataxic patients are able to learn to compensate for perturbations if the perturbation is introduced sufficiently gradually (Criscimagna-Hemminger et al. 2010; Izawa et al. 2011) (Fig. 1.3). Successful learning in this case is not accompanied by a change in the perceived consequences of actions (Izawa et al. 2011). This spared learning ability of cerebellar patients therefore does not appear to be associated with any change in an internal model. Interestingly, healthy subjects who learn to compensate for a rotation given only binary feedback about the success or failure of their movements are similarly able to successfully learn to counter the rotation without updating their predictions about the outcome of their movements (Izawa and Shadmehr 2011). These findings suggest that cerebellar ataxia patients, as well as healthy controls deprived of being able to extract prediction errors from a movement, are able to compensate by engaging an alternative learning mechanism that relies solely on the degree of success of a movement and not on the directionality or magnitude of errors. This also explains why learning is only possible in a gradual paradigm: natural variability in movements ensures occasional success when perturbations are small, sufficient to allow the patients to shift their control strategy and in turn enabling them to find movements that successfully counter larger perturbations later on.

The clearest example where model-based and model-free learning can be clearly dissociated is in explaining the ability to relearn a perturbation faster the second time around (savings). Huang and colleagues (2011) had subjects reach to a series of targets, while giving them rotated visual feedback. Crucially, the rotation changed pseudo-randomly from trial to trial, so that subjects achieved little success and never repeatedly made the same movement with success. Subjects did, however, adapt to the average imposed perturbation, which was a clockwise rotation of 20°. Following washout of this learning, subjects were given a further test block in which they made reaching movements to a single target under a constant rotation. In this test block, these subjects adapted no faster than naïve subjects, i.e., they did not exhibit savings despite having previously adapted their movements in the direction appropriate to counter the test perturbation. By contrast, a second group of subjects was faced with a nonrandom perturbation that was designed such that the same action would lead to success on every trial, regardless of the target location. After washout, subjects in this group adapted substantially faster to a subsequent test rotation, i.e., they exhibited strong savings. These results suggest that learning an internal model alone does not suffice to achieve savings—some degree of repetition of a successful movement is also necessary. Even so, the savings might have been due to a model-based process that is modulated by task success. A further experiment established that this is not the case by showing that savings can arise from prior learning of an *opposite rotation*, provided that both rotations required the same hand movement to achieve success (Fig. 1.4).

These results suggest that savings arises through recall of previously successful actions, rather than recall of a learned internal model. It may even be that the recall of successful actions is prompted by reward prediction errors rather than by sensory prediction errors. In force field learning, withholding a previously given reinforcement signal triggers partial recall of previously learned actions, even when kinematic performance errors are mechanically clamped at zero (Pekny et al. 2011).

Dopamine-dependence of Model-free Learning

From a theoretical perspective, model-free learning relies on errors in predicted reward (unlike model-based learning, which relies on errors in predicted sensory feedback). Dopamine neurons have been consistently linked with reward prediction errors (Montague et al. 1996; Schultz et al. 1997). Parkinson's disease (PD), in which there is widespread death

of dopaminergic neurons, is known to lead to learning deficits in operant learning tasks that rely on reward prediction error signals (Frank et al. 2004; Shohamy et al. 2005; Avila et al. 2009). Thus the study of learning in PD patients may offer crucial insights into which components of motor learning are subserved by model-free processes. PD patients show no impairment in learning a visuomotor rotation compared with age-matched control subjects. However, savings upon re-adaptation is almost absent in patients with PD (Marinelli et al. 2009; Bedard and Sanes 2011; Leow et al. 2012). This remarkable finding clearly demonstrates the importance of reinforcement and reward for savings and strongly accords with the results of Huang and colleagues (2011). Dopamine is also known to play a pivotal role in skill acquisition in rats. Blocking dopaminergic innervation of M1 from the ventral tegmental area (VTA) abolishes the ability of rats to improve performance in a grasping task. Blocking dopamine did not, however, impair performance of previously acquired skills, suggesting that dopamine plays a key role in learning but not execution of motor skill (Hosp et al. 2011).

Use-Dependent Learning

A significant line of observations that is difficult to reconcile with purely model-based frameworks is the fact that, even in the absence of perturbations, current movements appear to be influenced by previous movements—a phenomenon that has been termed *use-dependent learning*. Repeated movements toward a particular target lead future movements to be biased toward that movement direction (Huang et al. 2011; Verstynen and Sabes 2011). Point-to-point reaching movements around an obstacle lead to a distinct trajectory bias once the obstacle has been removed (Jax and Rosenbaum 2007). In redundant tasks, subjects are biased toward solutions that they were led to on previous attempts, even though these may be far from energetically optimal (Diedrichsen et al. 2010). These history-dependent biases in behavior may originate in primary motor cortex: motor responses elicited by transcranial magnetic stimulation (TMS) over M1 tend to be biased toward movements that were practiced immediately beforehand (Classen et al. 1998). Long-term practice over years seems to have a similar effect—TMS of M1 in expert musicians is more likely to elicit the same hand movements that occur while playing their instrument compared to nonmusician controls, or musicians that play other instruments (Gentner et al. 2010).

Use-dependent learning appears to be difficult to reconcile with the model-based view that movements are planned to be optimal according to a current model of the environment. Use-dependent learning also does not seem to be due to the same model-free learning mechanisms that give rise to savings. Movement execution biases can be induced without giving rise to savings (Huang et al. 2011), and savings occurs when target directions are distributed uniformly around a single start position, in which case no use-dependent learning occurs (Verstynen and Sabes 2011). It can be argued that these kinds of history-dependent effects reflect a learned model of the structure of the task, with biases in movement direction reflecting the influence of prior expectations about the location of the target (Verstynen and Sabes 2011). In this case, use-dependent learning can be thought of as a form of unsupervised learning, being driven not by prediction errors or rewards but by the statistics of previous actions.

An alternative way to frame use-dependent learning is as a form of motor *habit*—i.e. an insensitivity to changes in task goals (which is a generalization of the notion of reward revaluation). Habit and model-free learning have been equated in the realm of rodent decision-making (Daw et al. 2005). It may be possible, however, to dissociate them in the domain of motor control where, critically, independent assays of habit (movement biases) and model-free learning (savings) are available.

In summary, there are an increasing number of experimental observations in motor learning that are difficult to describe within model-based frameworks. We do not wish to suggest that these findings negate the idea of model-based learning. Rather, we propose that these phenomena occur due to additional learning mechanisms that operate independently of internal models. Furthermore, these mechanisms constitute more than just a curious nuisance that contaminates behavior in adaptation paradigms. We suggest that they are equally if not more critical than model-based learning. Indeed, acquisition of entirely novel motor behaviors may depend upon model-free learning mechanisms in which successful control policies in motor cortex are dopaminergically reinforced.

Combining Model-Based and Model-Free Learning Mechanisms

In sequential decision-making contexts, model-based and model-free learning systems are often conceived as operating independently, interacting only at the action selection stage. Both model-based and model-free strategies yield estimates of the value of choosing a particular action in a particular state, and these independent estimates must be combined to guide the ultimate choice of action. The confidence in each estimate plays a crucial role in this arbitration process—a concept that can be formalized in a Bayesian sense (Daw et al. 2005). The ultimate course of action could be based on either choosing the most reliable estimate (Daw et al. 2005) or combining the estimated values weighted by their relative reliabilities (Glascher et al. 2010). Early in learning, when there are few observations available, model-based approaches tend to be more reliable than model-free ones. Imprecise computations place a limit on the reliability of model-based methods, however, so that, when data are abundant, the direct approach of model-free learning becomes more reliable. Consequently, behavior tends to rely more on model-based mechanisms early on and more on model-free mechanisms after extensive practice.

In the context of control of the eyes and limbs, there is a continuous, possibly high-dimensional space of potential actions. When choosing among a discrete set of actions, an animal can exhaustively sample all available actions in each state—the only constraints on learning are the confidence that the animal has in its own observations and in the constancy of the environment. When the action space is continuous-valued, however, exhaustive sampling is not possible. In these scenarios, the model-based system becomes essential to rapidly guide behavior toward promising control policies, effectively guiding exploration for model-free learning. As is the case in discrete domains, abilities that are initially learned model-based should, with experience, become model-free behaviors. This transfer of responsibility explains the fact that anticipatory control remains intact in patients with cerebellar ataxia (who will have initially learned such coordination long before the onset of ataxia (Diedrichsen et al. 2005)) but new patterns of coordination cannot be learned.

Conclusions and Outlook

We have contrasted two distinct approaches to learning new control policies: model-based learning, in which sensory prediction errors indirectly drive updates to a control policy by updating a forward model, and model-free learning in which reward prediction errors drive changes to a control policy directly. We have argued that the motor system utilizes both kinds of learning. A parallel learning architecture lends the motor system robustness and redundancy; learning a given task can be achieved in many different ways, so that if one fails either through circumstance (such as if impoverished sensory feedback precludes learning a forward model) or disease, other mechanisms are still in place to ensure that overall performance can still be maintained or improved.

It has long been argued that the different structures within the brain appear to be well suited to learning from experience in different ways (Doya 1999). The complementary roles of

different brain regions for model-based and model-free learning are particularly nicely exhibited in a recent study by Galea and colleagues (2011). They found that anodal transcranial direct current stimulation (tDCS) over primary motor cortex did not affect the rate of adaptation to a visuomotor rotation. However, it did lead to a large effect on retention of learning. Conversely, anodal tDCS of the cerebellum substantially increased the rate of adaptation, but did not influence retention. Although the exact mechanisms by which brain stimulation is able to modulate learning in this way are far from clear at present, we suggest a general way of viewing this result as tDCS of the cerebellum promoting model-based learning and tDCS of M1 promoting model-free learning.

It is likely that the brain utilizes many different forms of learning, not restricted to the specific strategies and mechanisms we have focused on here. For instance, although we have described model-based learning as being dependent on an implicit internal model in the cerebellum, it is likely that other brain areas such as prefrontal cortex may contribute alternative, explicit task models that may give rise to a form of model-based control that might be considered more strategic (Mazzoni and Krakauer 2006; Taylor and Ivry 2011). Use-dependent learning appears to be neither model-based nor model-free in the sense that we have described. Many error-driven learning strategies have been proposed in which vector performance errors directly drive updates to a controller (Thoroughman and Shadmehr 2000; Franklin et al. 2008). Such strategies are neither model-based in the sense we have described (since the error directly influences the controller, rather than going through a forward model) nor model-free (since learning is error-driven rather than reward-driven).

A deeper understanding of the multiple component processes that support motor learning is paramount to advance the efficacy of neuro-rehabilitation following brain injury. The most promising path to achieving this goal is through coupling theoretical insights with carefully designed experiments, study of specific patient groups, and the use of brain stimulation techniques. Dissecting motor learning into its constituent components is a clear and important goal for future motor control research. We believe that computational levels of description offer a sound basis by which to begin this classification (Fig. 1.5).

References

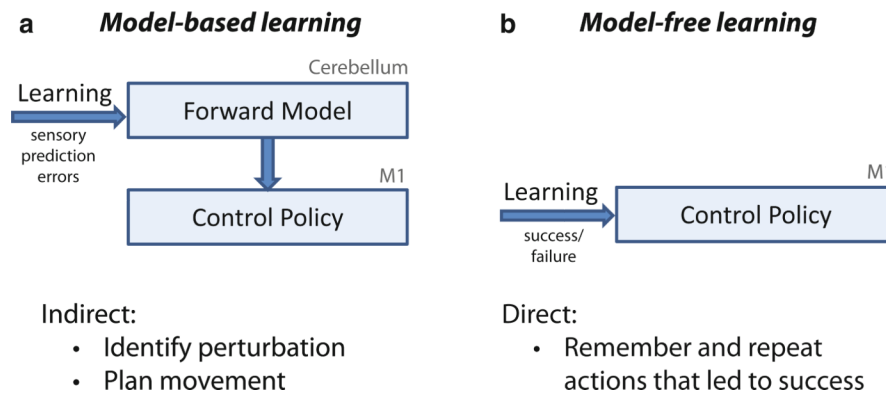
- Ariff G, Donchin O, Nanayakkara T, Shadmehr R. A real-time state predictor in motor control: study of saccadic eye movements during unseen reaching movements. *J Neurosci*. 2002; 22:7721–7729. [PubMed: 12196595]
- Avila I, Reilly MP, Sanabria F, Posadas-Sanchez D, Chavez CL, Banerjee N, Killeen P, Castaneda E. Modeling operant behavior in the Parkinsonian rat. *Behav Brain Res*. 2009; 198:298–305. [PubMed: 19073222]
- Balleine BW, Dickinson A. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology*. 1998; 37:407–419. [PubMed: 9704982]
- Balleine BW, O'Doherty JP. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology*. 2010; 35:48–69. [PubMed: 19776734]
- Bastian AJ. Learning to predict the future: the cerebellum adapts feedforward movement control. *Curr Opin Neurobiol*. 2006; 16:645–649. [PubMed: 17071073]
- Bedard P, Sanes JN. Basal ganglia-dependent processes in recalling learned visual-motor adaptations. *Exp Brain Res*. 2011; 209:385–393. [PubMed: 21311878]
- Bertsekas, DP. *Dynamic programming and optimal control*. Athena Scientific; 1996.
- Carmena JM, Lebedev MA, Crist RE, O'Doherty JE, Santucci DM, Dimitrov DF, Patil PG, Henriquez CS, Nicolelis MA. Learning to control a brain-machine interface for reaching and grasping by primates. *PLoS Biol*. 2003; 1:E42. [PubMed: 14624244]

- Chen-Harris H, Joiner WM, Ethier V, Zee DS, Shadmehr R. Adaptive control of saccades via internal feedback. *J Neurosci*. 2008; 28:2804–2813. [PubMed: 18337410]
- Classen J, Liepert J, Wise SP, Hallett M, Cohen LG. Rapid plasticity of human cortical movement representation induced by practice. *J Neurophysiol*. 1998; 79:1117–1123. [PubMed: 9463469]
- Conditt MA, Gandolfo F, Mussa-Ivaldi FA. The motor system does not learn the dynamics of the arm by rote memorization of past experience. *J Neurophysiol*. 1997; 78:554–560. [PubMed: 9242306]
- Criscimagna-Hemminger SE, Bastian AJ, Shadmehr R. Size of error affects cerebellar contributions to motor learning. *J Neurophysiol*. 2010; 103:2275–2284. [PubMed: 20164398]
- Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci*. 2005; 8:1704–1711. [PubMed: 16286932]
- Dayan P. Goal-directed control and its antipodes. *Neural Netw*. 2009; 22:213–219. [PubMed: 19362448]
- Diedrichsen J, Verstynen T, Lehman SL, Ivry RB. Cerebellar involvement in anticipating the consequences of self-produced actions during bimanual movements. *J Neurophysiol*. 2005; 93:801–812. [PubMed: 15356182]
- Diedrichsen J, White O, Newman D, Lally N. Use-dependent and error-based learning of motor behaviors. *J Neurosci*. 2010; 30:5159–5166. [PubMed: 20392938]
- Donchin O, Francis JT, Shadmehr R. Quantifying generalization from trial-by-trial behavior of adaptive systems that learn with basis functions: theory and experiments in human motor control. *J Neurosci*. 2003; 23:9032–9045. [PubMed: 14534237]
- Donchin O, Rabe K, Diedrichsen J, Lally N, Schoch B, Gizewski ER, Timmann D. Cerebellar regions involved in adaptation to force field and visuomotor perturbation. *J Neurophysiol*. 2011; 107(1): 134–47. [PubMed: 21975446]
- Doya K. What are the computations of the cerebellum, the basal ganglia and the cerebral cortex? *Neural Netw*. 1999; 12:961–974. [PubMed: 12662639]
- Fermin A, Yoshida T, Ito M, Yoshimoto J, Doya K. Evidence for model-based action planning in a sequential finger movement task. *J Mot Behav*. 2010; 42:371–379. [PubMed: 21184355]
- Flanagan JR, Wing AM. The role of internal models in motion planning and control: evidence from grip force adjustments during movements of hand-held loads. *J Neurosci*. 1997; 17:1519–1528. [PubMed: 9006993]
- Frank MJ, Seeberger LC, O'Reilly RC. By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science*. 2004; 306:1940–1943. [PubMed: 15528409]
- Franklin DW, Burdet E, Tee KP, Osu R, Chew CM, Milner TE, Kawato M. CNS learns stable, accurate, and efficient movements using a simple algorithm. *J Neurosci*. 2008; 28:11165–11173. [PubMed: 18971459]
- Galea JM, Vazquez A, Pasricha N, de Xivry JJ, Celnik P. Dissociating the roles of the cerebellum and motor cortex during adaptive learning: the motor cortex retains what the cerebellum learns. *Cereb Cortex*. 2011; 21:1761–1770. [PubMed: 21139077]
- Gentner R, Gorges S, Weise D, aufm Kampe K, Buttman M, Classen J. Encoding of motor skill in the corticomuscular system of musicians. *Curr Biol*. 2010; 20:1869–1874. [PubMed: 20951047]
- Glascher J, Daw N, Dayan P, O'Doherty JP. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*. 2010; 66:585–595. [PubMed: 20510862]
- Hosp JA, Pekanovic A, Rioult-Pedotti MS, Luft AR. Dopaminergic projections from midbrain to primary motor cortex mediate motor skill learning. *J Neurosci*. 2011; 31:2481–2487. [PubMed: 21325515]
- Huang VS, Haith A, Mazzoni P, Krakauer JW. Rethinking motor learning and savings in adaptation paradigms: model-free memory for successful actions combines with internal models. *Neuron*. 2011; 70:787–801. [PubMed: 21609832]
- Izawa J, Shadmehr R. Learning from sensory and reward prediction errors during motor adaptation. *PLoS Comput Biol*. 2011; 7:e1002012. [PubMed: 21423711]
- Izawa J, Criscimagna-Hemminger SE, Shadmehr R. Cerebellar Contributions to Learning Sensory Consequences of Action. *J Neurosci*. 2011; 32(12):4230–4239. [PubMed: 22442085]

- Jax SA, Rosenbaum DA. 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.* 2007; 33:425–441.
- Jordan MI, R. DE. Forward models: Supervised learning with a distal teacher. *Cognitive Sci.* 1992; 16:307–354.
- Kawato M, Gomi H. A computational model of four regions of the cerebellum based on feedback-error learning. *Biol Cybern.* 1992; 68:95–103. [PubMed: 1486143]
- Killcross S, Coutureau E. Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb Cortex.* 2003; 13:400–408. [PubMed: 12631569]
- Krakauer JW, Ghilardi MF, Ghez C. Independent learning of internal models for kinematic and dynamic control of reaching. *Nat Neurosci.* 1999; 2:1026–1031. [PubMed: 10526344]
- Krakauer JW, Pine ZM, Ghilardi MF, Ghez C. Learning of visuomotor transformations for vectorial planning of reaching trajectories. *J Neurosci.* 2000; 20:8916–8924. [PubMed: 11102502]
- Lackner JR, Dizio P. Rapid adaptation to Coriolis force perturbations of arm trajectory. *J Neurophysiol.* 1994; 72:299–313. [PubMed: 7965013]
- Leow LA, Loftus AM, Hammond GR. Impaired savings despite intact initial learning of motor adaptation in Parkinson's disease. *Exp Brain Res.* 2012; 2:295–304. [PubMed: 22430183]
- Marinelli L, Crupi D, Di Rocco A, Bove M, Eidelberg D, Abbruzzese G, Ghilardi MF. Learning and consolidation of visuo-motor adaptation in Parkinson's disease. *Parkinsonism Relat Disord.* 2009; 15:6–11. [PubMed: 18424221]
- Martin TA, Keating JG, Goodkin HP, Bastian AJ, Thach WT. Throwing while looking through prisms. I. Focal olivocerebellar lesions impair adaptation. *Brain.* 1996; 119(4):1183–1198. [PubMed: 8813282]
- Maschke M, Gomez CM, Ebner TJ, Konczak J. Hereditary cerebellar ataxia progressively impairs force adaptation during goal-directed arm movements. *J Neurophysiol.* 2004; 91:230–238. [PubMed: 13679403]
- Mazzoni P, Krakauer JW. An implicit plan overrides an explicit strategy during visuomotor adaptation. *J Neurosci.* 2006; 26:3642–3645. [PubMed: 16597717]
- McGuire LM, Sabes PN. Sensory transformations and the use of multiple reference frames for reach planning. *Nat Neurosci.* 2009; 12:1056–1061. [PubMed: 19597495]
- Medina JF. The multiple roles of Purkinje cells in sensori-motor calibration: to predict, teach and command. *Curr Opin Neurobiol.* 2011; 21:616–622. [PubMed: 21684147]
- Mehta B, Schaal S. Forward models in visuomotor control. *J Neurophysiol.* 2002; 88:942–953. [PubMed: 12163543]
- Miall RC, Christensen LO, Cain O, Stanley J. Disruption of state estimation in the human lateral cerebellum. *PLoS Biol.* 2007; 5:e316. [PubMed: 18044990]
- Montague PR, Dayan P, Sejnowski TJ. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J Neurosci.* 1996; 16:1936–1947. [PubMed: 8774460]
- Mosier KM, Scheidt RA, Acosta S, Mussa-Ivaldi FA. Remapping hand movements in a novel geometrical environment. *J Neurophysiol.* 2005; 94:4362–4372. [PubMed: 16148276]
- Munuera J, Morel P, Duhamel JR, Deneve S. Optimal sensorimotor control in eye movement sequences. *J Neurosci.* 2009; 29:3026–3035. [PubMed: 19279239]
- Nagengast AJ, Braun DA, Wolpert DM. Optimal control predicts human performance on objects with internal degrees of freedom. *PLoS Comput Biol.* 2009; 5:e1000419. [PubMed: 19557193]
- Nowak DA, Hermsdorfer J, Rost K, Timmann D, Topka H. Predictive and reactive finger force control during catching in cerebellar degeneration. *Cerebellum.* 2004; 3:227–235. [PubMed: 15686101]
- Pasalar S, Roitman AV, Durfee WK, Ebner TJ. Force field effects on cerebellar Purkinje cell discharge with implications for internal models. *Nat Neurosci.* 2006; 9:1404–1411. [PubMed: 17028585]
- Pekny SE, Criscimagna-Hemminger SE, Shadmehr R. Protection and expression of human motor memories. *J Neurosci.* 2011; 31:13829–13839. [PubMed: 21957245]
- Rabe K, Livne O, Gizewski ER, Aurich V, Beck A, Timmann D, Donchin O. Adaptation to visuomotor rotation and force field perturbation is correlated to different brain areas in patients with cerebellar degeneration. *J Neurophysiol.* 2009; 101:1961–1971. [PubMed: 19176608]

- Reis J, Schambra HM, Cohen LG, Buch ER, Fritsch B, Zarahn E, Celnik PA, Krakauer JW. Noninvasive cortical stimulation enhances motor skill acquisition over multiple days through an effect on consolidation. *Proc Natl Acad Sci.* 2009; 106:1590–1595. [PubMed: 19164589]
- Roitman AV, Pasalar S, Johnson MT, Ebner TJ. Position, direction of movement, and speed tuning of cerebellar Purkinje cells during circular manual tracking in monkey. *J Neurosci.* 2005; 25:9244–9257. [PubMed: 16207884]
- Rost K, Nowak DA, Timmann D, Hermsdorfer J. Preserved and impaired aspects of predictive grip force control in cerebellar patients. *Clin Neurophysiol.* 2005; 116:1405–1414. [PubMed: 15978503]
- Schaefer SY, Shelly IL, Thoroughman KA. Beside the point: motor adaptation without feedback-based error correction in task-irrelevant conditions. *J Neurophysiol.* 2012; 107(4):1247–56. [PubMed: 22157120]
- Schultz W, Dayan P, Montague PR. A neural substrate of prediction and reward. *Science.* 1997; 275:1593–1599. [PubMed: 9054347]
- Shadmehr R, Mussa-Ivaldi FA. Adaptive representation of dynamics during learning of a motor task. *J Neurosci.* 1994; 14:3208–3224. [PubMed: 8182467]
- Shadmehr R, Krakauer JW. A computational neuroanatomy for motor control. *Exp Brain Res.* 2008; 185:359–381. [PubMed: 18251019]
- Shadmehr R, Smith MA, Krakauer JW. Error correction, sensory prediction, and adaptation in motor control. *Annu Rev Neurosci.* 2010; 33:89–108. [PubMed: 20367317]
- Shmuelof L, Krakauer JW, Mazzoni P. How is a motor skill learned? Change and invariance at the levels of task success and trajectory control. *J Neurophysiol.* 2012; 08(2):578–94. [PubMed: 22514286]
- Shohamy D, Myers CE, Grossman S, Sage J, Gluck MA. The role of dopamine in cognitive sequence learning: evidence from Parkinson's disease. *Behav Brain Res.* 2005; 156:191–199. [PubMed: 15582105]
- Simani MC, McGuire LM, Sabes PN. Visual-shift adaptation is composed of separable sensory and task-dependent effects. *J Neurophysiol.* 2007; 98:2827–2841. [PubMed: 17728389]
- Smith MA, Shadmehr R. Intact ability to learn internal models of arm dynamics in Huntington's disease but not cerebellar degeneration. *J Neurophysiol.* 2005; 93:2809–2821. [PubMed: 15625094]
- Squire LR. Memory and the hippocampus: a synthesis of findings with rats, monkeys and humans. *Psychol Rev.* 1992; 99:195–231. [PubMed: 1594723]
- Sternad D, Abe MO, Hu X, Muller H. Neuromotor noise, error tolerance and velocity-dependent costs in skilled performance. *PLoS Comput Biol.* 2011; 7:e1002159. [PubMed: 21966262]
- Sutton, RS.; Barto, AG. Reinforcement learning: An introduction. Cambridge Univ Press; 1998.
- Synofzik M, Thier P, Lindner A. Internalizing agency of self-action: perception of one's own hand movements depends on an adaptable prediction about the sensory action outcome. *J Neurophysiol.* 2006; 96:1592–1601. [PubMed: 16738220]
- Synofzik M, Lindner A, Thier P. The cerebellum updates predictions about the visual consequences of one's behavior. *Curr Biol.* 2008; 18:814–818. [PubMed: 18514520]
- Taylor JA, Ivry RB. Flexible cognitive strategies during motor learning. *PLoS Comput Biol.* 2011; 7:e1001096. [PubMed: 21390266]
- Taylor JA, Klemfuss NM, Ivry RB. An explicit strategy prevails when the cerebellum fails to compute movement errors. *Cerebellum.* 2010; 9:580–586. [PubMed: 20697860]
- Thoroughman KA, Shadmehr R. Learning of action through adaptive combination of motor primitives. *Nature.* 2000; 407:742–747. [PubMed: 11048720]
- Todorov, E. Bayesian brain: probabilistic approaches to neural coding. MIT Press; Cambridge: 2007. Optimal control theory.; p. 269-298.
- Todorov E, Jordan MI. Optimal feedback control as a theory of motor coordination. *Nat Neurosci.* 2002; 5:1226–1235. [PubMed: 12404008]

- Tseng YW, Diedrichsen J, Krakauer JW, Shadmehr R, Bastian AJ. Sensory prediction errors drive cerebellum-dependent adaptation of reaching. *J Neurophysiol.* 2007; 98:54–62. [PubMed: 17507504]
- van der Meer MA, Redish AD. Ventral striatum: a critical look at models of learning and evaluation. *Curr Opin Neurobiol.* 2011; 21:387–392. [PubMed: 21420853]
- Verstynen T, Sabes PN. How each movement changes the next: an experimental and theoretical study of fast adaptive priors in reaching. *J Neurosci.* 2011; 31:10050–10059. [PubMed: 21734297]
- Wallman J, Fuchs AF. Saccadic gain modification: visual error drives motor adaptation. *J Neurophysiol.* 1998; 80:2405–2416. [PubMed: 9819252]
- Wagner MJ, Smith MA. Shared internal models for feedforward and feedback control. *J Neurosci.* 2008; 28:10663–10673. [PubMed: 18923042]
- Wolpert DM, Miall RC. Forward Models for Physiological Motor Control. *Neural Netw.* 1996; 9:1265–1279. [PubMed: 12662535]
- Wolpert DM, Miall RC, Kawato M. Internal models in the cerebellum. *Trends Cogn Sci.* 1998; 2:338–347. [PubMed: 21227230]
- Wolpert DM, Diedrichsen J, Flanagan JR. Principles of sensorimotor learning. *Nat Rev Neurosci.* 2011; 12:739–751. [PubMed: 22033537]
- Wong AL, Shelhamer M. Sensorimotor adaptation signals are derived from realistic predictions of movement outcomes. *J Neurophysiol.* 2011; 105(3):1130–40. [PubMed: 21123665]
- Xu-Wilson M, Chen-Harris H, Zee DS, Shadmehr R. Cerebellar contributions to adaptive control of saccades in humans. *J Neurosci.* 2009; 29:12930–12939. [PubMed: 19828807]

**Fig. 1.1.**

Comparison of model-based and model-free strategies for updating a control policy based on experience. **a** Model-based learning schematic. Changes to the control policy are brought about indirectly through first updating a forward model of the motor apparatus and environment based on sensory prediction errors, then using this knowledge to calculate an appropriate controller for the current task. **b** Model-free learning schematic. The control policy is updated directly based on reward prediction errors

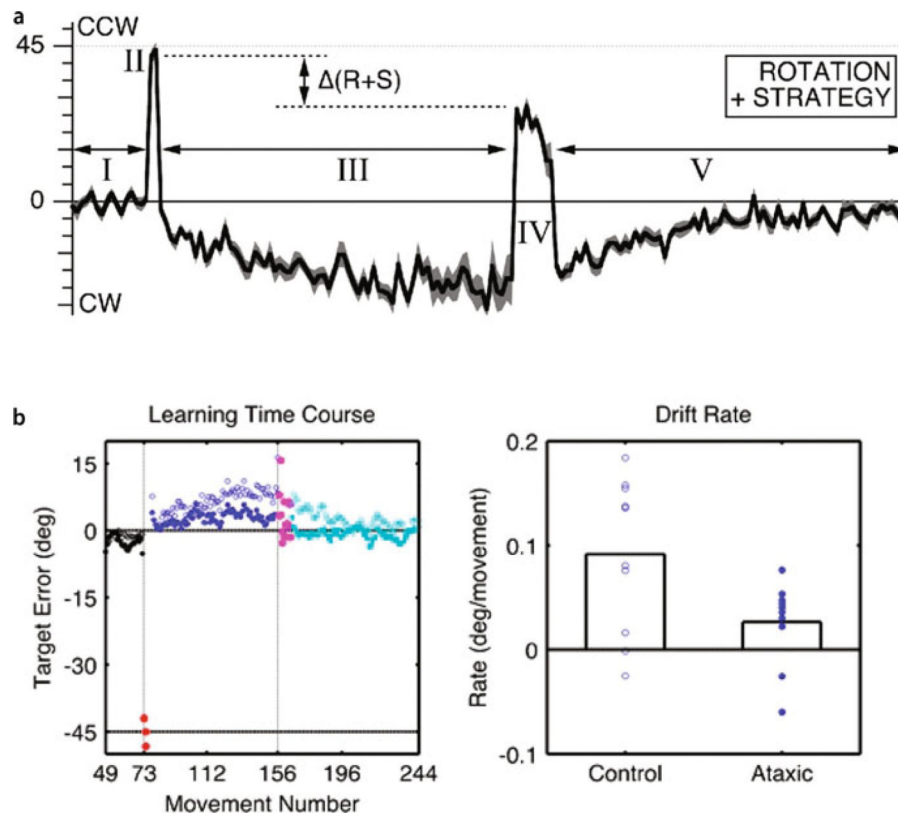


Fig. 1.2.

Motor learning is driven by sensory prediction errors. **a** Healthy subjects that are provided with an explicit strategy to counter a 45° rotation initially counter the perturbation successfully, but performance immediately drifts in the direction opposite the rotation. (Reproduced from Mazzoni and Krakauer 2006). **b** This drift is attenuated in patients with cerebellar ataxia (note that rotation direction is opposite as compared with panel **a**). (Reproduced from Taylor et al. 2010)

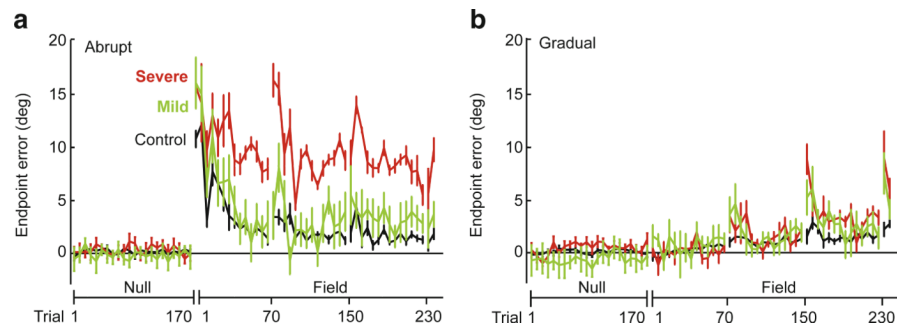


Fig. 1.3.

Force field learning in patients with cerebellar ataxia. **a** Severe patients are grossly impaired in adapting to an abruptly introduced force field. **b** Introducing the force field gradually allows compensation similar to healthy controls. (Reproduced from Criscimagna-Hemminger et al. 2010)

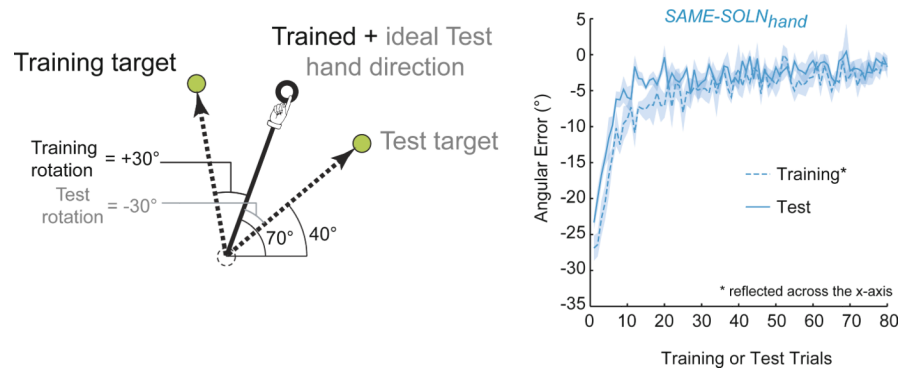


Fig. 1.4. Adaptation to a visual rotation is accelerated by prior learning of an opposite rotation, provided the targets are arranged such that the action required to successfully counter the rotation is the same in both cases. (Reproduced from Huang et al. 2011)

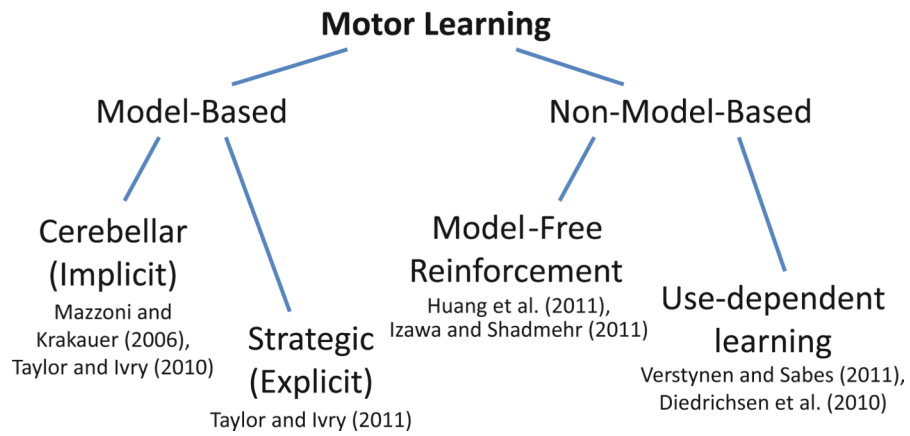


Fig. 1.5. A motor learning taxonomy, by analogy to the famous taxonomy of memory by Squire. (Squire 1992)