



Review in Advance first posted online on May 10, 2017. (Changes may still occur before final publication online and in print.)

# The Role of Variability in Motor Learning

Ashesh K. Dhawale,<sup>1,2</sup> Maurice A. Smith,<sup>2,3</sup> and Bence P. Ölveczky<sup>1,2</sup>

<sup>1</sup>Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138; email: olveczky@fas.harvard.edu

<sup>2</sup>Center for Brain Science, Harvard University, Cambridge, Massachusetts 02138

<sup>3</sup>John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138

Annu. Rev. Neurosci. 2017. 40:479–98

The *Annual Review of Neuroscience* is online at [neuro.annualreviews.org](http://neuro.annualreviews.org)

<https://doi.org/10.1146/annurev-neuro-072116-031548>

Copyright © 2017 by Annual Reviews.  
All rights reserved

## Keywords

reinforcement learning, motor control, songbird, motor adaptation

## Abstract

Trial-to-trial variability in the execution of movements and motor skills is ubiquitous and widely considered to be the unwanted consequence of a noisy nervous system. However, recent studies have suggested that motor variability may also be a feature of how sensorimotor systems operate and learn. This view, rooted in reinforcement learning theory, equates motor variability with purposeful exploration of motor space that, when coupled with reinforcement, can drive motor learning. Here we review studies that explore the relationship between motor variability and motor learning in both humans and animal models. We discuss neural circuit mechanisms that underlie the generation and regulation of motor variability and consider the implications that this work has for our understanding of motor learning.

**Contents**

INTRODUCTION..... 480  
 NEURAL SOURCES OF MOTOR VARIABILITY ..... 481  
 REINFORCEMENT LEARNING: A NATURAL FRAMEWORK FOR  
 LINKING MOTOR VARIABILITY AND LEARNING ..... 483  
 SONGBIRD STUDIES LINKING MOTOR VARIABILITY AND LEARNING .. 485  
 THE ROLE OF VARIABILITY IN HUMAN MOTOR LEARNING..... 487  
 LEARNING-DEPENDENT REGULATION OF MOTOR VARIABILITY..... 489  
 CONTEXT-DEPENDENT REGULATION OF MOTOR VARIABILITY ..... 490  
 REGULATING THE STRUCTURE OF MOTOR VARIABILITY ..... 491  
 CONCLUSIONS AND OUTLOOK..... 492

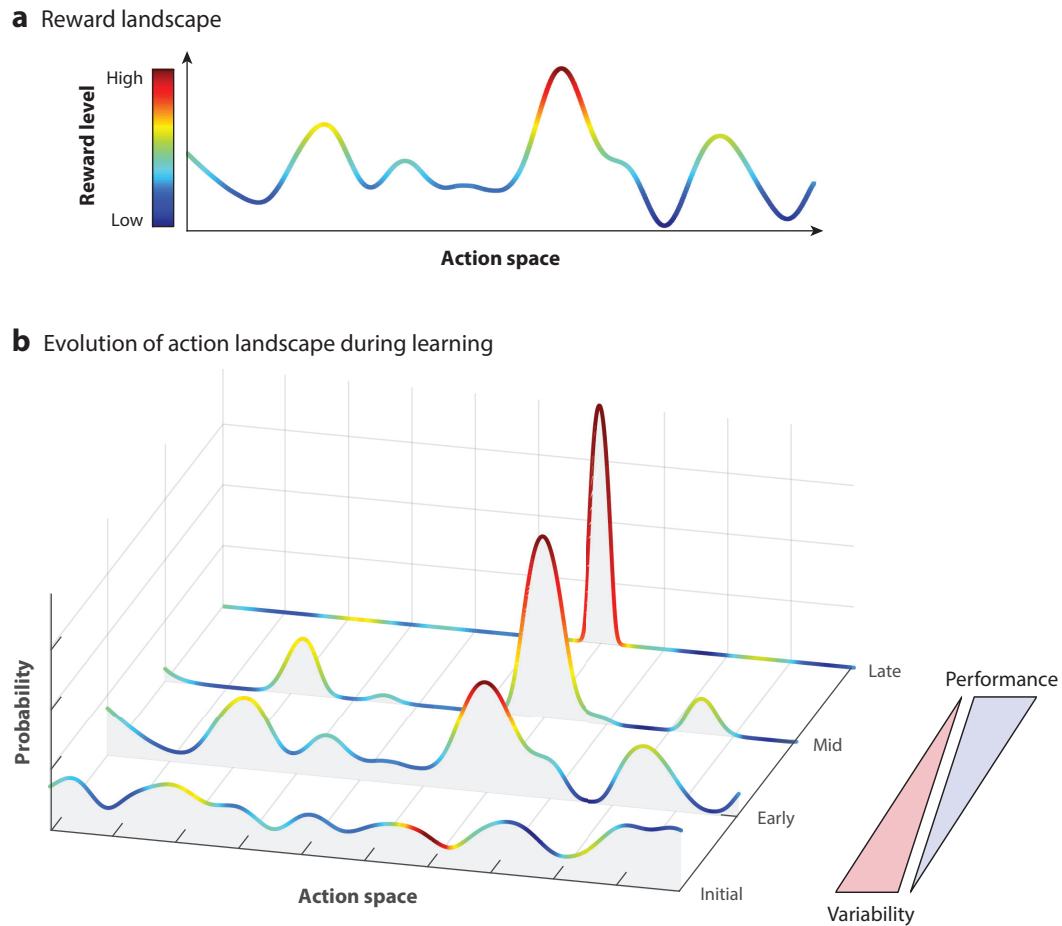
**INTRODUCTION**

Improving performance, whether on the tennis court or at the piano, often means reducing the variability of our actions. Yet, no matter how hard we practice, generating identical movements on successive trials is virtually impossible. But why is it so hard to tame performance variability? One reason is that our actions are generated by an inherently noisy nervous system (Faisal et al. 2008, Renart & Machens 2014, Stein et al. 2005). This noise results from stochastic events at the level of ion channels (White et al. 2000), synapses (Calvin & Stevens 1968), and neurons (Mainen & Sejnowski 1995) and further from instabilities in the dynamics of neural networks (Babloyantz et al. 1985, Vreeswijk & Sompolinsky 1996). These processes combine to add uncertainty and randomness to how the brain operates and generates movements. It is widely believed that motor control is optimized for current performance and that variability that interferes with this goal should be minimized or countered (Harris & Wolpert 1998, Todorov & Jordan 2002).

However, a complementary view of motor variability suggests that it may be a feature of how sensorimotor circuits operate and learn (Herzfeld & Shadmehr 2014, Tumer & Brainard 2007, Wu et al. 2014). This view is best appreciated from the perspective of the novice who has to acquire a new task and for whom variability in motor output can be construed as a means of exploring motor space (Figure 1). Through a process of trial and error, such exploration could, in line with reinforcement learning theory (Kaelbling et al. 1996, Sutton & Barto 1998), steer the motor system toward new control policies and patterns of motor activity that improve performance and reduce costs (Shadmehr et al. 2016). In this view, motor variability is to skill learning what genetic variation is to evolution: an essential component of a process that, through selection by consequence, shapes adaptive behaviors (Skinner 1981).

This review is not about whether motor variability is good, bad, a feature, or a bug, nor are we implying that there is a dichotomy to be resolved. Indeed, there is little doubt that in many situations and for many facets of motor control, uncertainty and noise in neural function, and the motor variability it gives rise to, is undesirable. But while this perspective has been treated and discussed extensively in the literature (Braun et al. 2009, Franklin et al. 2012, Harris & Wolpert 1998, Izawa & Shadmehr 2008, Izawa et al. 2008, Kording & Wolpert 2004, Newell 1993, Todorov 2004, Todorov & Jordan 2002), less attention has been given to how motor variability could augment motor learning (Figure 1). To what extent, and under what circumstances, can motor variability be harnessed to improve future performance? Does the nervous system regulate and shape motor variability to promote learning, and if so, how does it do it? Here we review recent studies related





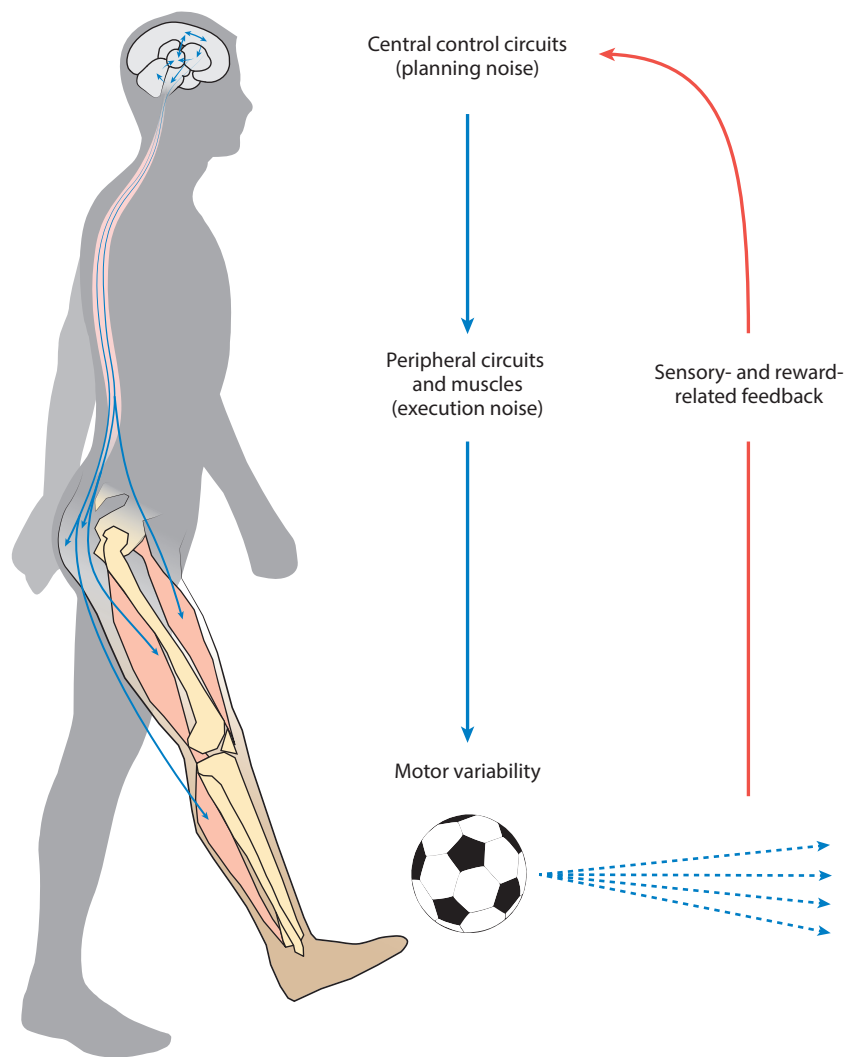
**Figure 1**

Illustration of how variability can be conducive to motor learning. (a) Each task is associated with a reward landscape in action space. (b) When the reward landscape is not known, trial-and-error reinforcement learning offers a powerful strategy for finding appropriate solutions. This requires initial exploration of action space coupled with a process that reinforces rewarded actions. Variability is initially high, but as reinforcement learning proceeds (early to late), variability is reduced as the motor system hones in on action variants associated with high reward. Note that early in learning, there may be motor biases unrelated to the task-specific reward landscape.

to these questions, discuss current views of motor variability, and identify research directions that may further advance our understanding of how motor variability contributes to learning.

## NEURAL SOURCES OF MOTOR VARIABILITY

Movements are the result of tightly choreographed patterns of muscle activity generated by a network of hierarchically organized motor controllers (Lemon 2008). In principle, motor variability can arise at any level of the motor pathway, from variation in movement planning by central circuits to noise in force production by muscles (**Figure 2**). The motor system could take advantage of variability as long as successful variants can be reinforced and reproduced. However, little is known about whether and how variability at various levels of motor planning and control



**Figure 2**

Sources of motor variability. Motor variability can arise at all levels of the motor systems. Here we distinguish variability in central planning and control circuits, referred to as planning noise, from variability in the motor periphery, referred to as execution noise. Variability conducive to learning is more likely to originate in central circuits, which receive performance-related feedback, as opposed to peripheral circuits where variability may be more difficult to reinforce and reproduce.

is harnessed to drive improvements in behavior. Below, we review possible sources of motor variability and discuss their relevance to motor learning.

Broadly speaking, variability in sensorimotor systems can originate at both cellular and network levels (Faisal et al. 2008, Renart & Machens 2014). At the cellular level, noise comes from stochastic biophysical and chemical events that underlie processes such as spike initiation (Schneidman et al. 1998, van Rossum et al. 2003, White et al. 2000) and propagation (Faisal & Laughlin 2007, Horikawa 1991), synaptic transmission (Calvin & Stevens 1968, Katz & Miledi 1970), and muscle



activation (Clamann 1969, Hamilton et al. 2004, Jones et al. 2002). Different neural network architectures can then either amplify or dampen such noise (Faisal et al. 2008). For instance, instabilities in recurrent network dynamics can magnify variability caused by noisy neurons (Babloyantz et al. 1985, Litwin-Kumar & Doiron 2012, London et al. 2010, Vogels et al. 2005), whereas pooling across them can enhance correlated signals and reduce the effects of noise (Bruno & Sakmann 2006, Diesmann et al. 1999).

Variability in the motor system (**Figure 2**) has perhaps been characterized most extensively at the motor periphery. Studies have shown signal-dependent noise in force production (i.e., trial-to-trial fluctuations whose standard deviation scales linearly with mean force) reflects a fundamental property of muscle function (Clamann 1969, Hamilton et al. 2004, Jones et al. 2002, van Beers et al. 2004). Because of its uncontrollable nature, such peripherally derived variability, often referred to as execution noise (van Beers et al. 2004) (**Figure 2**), may not be well suited for learning-related motor exploration. Rather, the motor system may have evolved strategies to decrease it in order to increase movement accuracy. This very idea has inspired an influential set of theories and models (Fitts 1954, Harris & Wolpert 1998) able to predict kinematic features of a variety of movements, including eye saccades and limb reaches (Harris & Wolpert 1998; van Beers et al. 2004, 2007).

In contrast, variability originating in central planning circuits may be better suited for driving learning-related motor exploration (**Figure 2**). These circuits have more ready access to reinforcement signals (Björklund & Dunnett 2007, Schultz 1998) and show more experience-dependent plasticity (Doyon & Benali 2005, Nudo et al. 1996, Sanes & Donoghue 2000). But gauging how higher-order motor circuits contribute to movement variability can be difficult because activity patterns related to motor planning are often intermixed with reafferent signals reflecting past (or ongoing) movements and task performance (Flament & Hore 1988, Kakei et al. 1999, Lauwereyns et al. 2002) (**Figure 2**).

One way to reduce contamination from reafference is to record in delayed response tasks, in which subjects have to withhold a prepared movement until a go cue is presented. Using such a behavioral paradigm, Churchland et al. (2006) found that a significant fraction—up to half—of the trial-to-trial variability in reach velocity could be explained by the firing rates of cortical neurons in the period prior to movement initiation, even in the case of well-practiced movements that were earlier thought to be subject primarily to execution noise (Harris & Wolpert 1998, van Beers et al. 2004).

Besides reflecting cellular and network noise, trial-to-trial variability in motor output can also be deterministic. For example, in uncertain environments, suboptimal inference can add behavioral variability even in the absence of network noise (Beck et al. 2012). Furthermore, task structure and fluctuations in reward expectation can contribute predictable variability in kinematics and motor timing (Haith et al. 2012, Kawagoe et al. 1998, Marcos et al. 2013, Opris et al. 2011, Takikawa et al. 2002). Error-correcting motor learning processes can similarly generate predictable changes in motor output (Baddeley et al. 2003, Scheidt et al. 2001, Smith et al. 2006, van Beers 2009). Whether such deterministic changes can double as exploratory motor variability that drives motor learning (**Figure 1**) remains to be understood.

## REINFORCEMENT LEARNING: A NATURAL FRAMEWORK FOR LINKING MOTOR VARIABILITY AND LEARNING

The process of updating a system by repeating states that lead to favorable outcomes forms the basis of reinforcement learning (Kaelbling et al. 1996, Sutton & Barto 1998), a computational framework that has illuminated a variety of different learning and decision making-processes (Lee et al. 2012, Niv 2009) and has inspired algorithms for machine learning (Alpaydin 2014, Sutton & Barto



1998). Reinforcement learning also provides the theoretical foundation for operant conditioning, a powerful training method that is predicated on the idea that reinforced behaviors become more frequently expressed (Skinner 1938, 1963; Thorndike 1898).

A major difference with other forms of learning is that reinforcement learning explicitly requires exploration. The agent probes the consequences of various actions and registers or updates their values, a process that allows it to adaptively and contextually regulate the expression of the probed actions. There is increasing evidence that the brain implements the computations predicted by reinforcement learning theory (Daw et al. 2006, Eshel et al. 2015, Lee et al. 2012, Niv 2009, Schultz et al. 1997, Wunderlich et al. 2009).

But although reinforcement learning has proved a powerful framework for decision making (i.e., the process of selecting an action from a discrete set of options), learning and generating the details of the specific actions pose very different challenges. Not only are the neural circuits involved in motor control likely distinct from those that implement higher-order decision making, but the dimensionality and complexity of the decisions that the motor system makes are of a different magnitude. This is because motor decisions are made in the high-dimensional and continuous space of possible movement patterns and implemented by a highly redundant motor system with many degrees of freedom (Bernshtein 1967, Lashley 1933).

Standard reinforcement learning algorithms, well suited to low-dimensional tasks such as choosing between discrete options, may not scale well to more complex tasks (Parr 1998, Peters & Schaal 2008), raising the question of whether they represent effective strategies for motor learning. Indeed, the oft-discussed curse of dimensionality, which describes the fact that the size of the solution space explodes as the complexity of a task or its control increases, presents a formidable challenge not only for reinforcement learning algorithms but for virtually any type of machine learning (Bellman 1957). The success of deep learning networks in solving complex supervised decision and classification problems has, in large part, been due to the use of convolutional network architectures that reduce dramatically the dimensionality of the solution space by enforcing highly symmetric patterns in the weights to be learned (LeCun et al. 1998, 2015; Simonyan & Zisserman 2014).

Another key to the success of deep learning networks has been the use of unsupervised methods to pretrain networks based on the statistics of the input data (Hinton et al. 2006, LeCun et al. 2015, Lee et al. 2009). This pretraining can serve to get the network into a fertile part of solution space before the main training begins. For reinforcement learning problems, a similar narrowing of solution space can be achieved by using imitation, or other ways of instructing the system, as a preamble to reinforcement learning (Kormushev et al. 2010, Price & Boutillier 2003). After emulating the behavior of an expert tutor, local trial-and-error learning can start off in the neighborhood of an approximate solution. For example, AlphaGo, Google's recently unveiled agent for playing the popular board game Go, was created using this general approach by first imitating expert human players and then learning by reinforcement from playing against itself (Silver et al. 2016).

Other solutions for making reinforcement learning algorithms work with more complex problems, such as hierarchical reinforcement learning algorithms (Botvinick 2012, Parr 1998), policy gradient methods (Peters & Schaal 2008), and value decomposition (Gershman et al. 2009), also aim to break down the complexity of the learning problem into smaller, more manageable chunks or policies. Although the extent to which the nervous system implements such strategies remains to be understood, a recent study on songbirds provides some intriguing clues. Using a reinforcement learning paradigm to change temporal and spectral features of the song independently, Ali and colleagues (2013) showed that the song circuit modularizes song learning by implementing separate reinforcement learning processes for spectral and temporal aspects of song.



## WHAT SONGBIRDS TELL US ABOUT MOTOR VARIABILITY AND LEARNING

The question of whether and how variability can be harnessed for learning has been examined most thoroughly in the zebra finch, a songbird that learns its courtship vocalization early in life by first listening to a tutor, then engaging in trial-and-error motor learning to copy the memorized song (Immelmann 1969, Tchernichovski et al. 2001) (**Figure 3a,b**). Vocal control circuits in songbirds are organized into two main pathways. The descending motor pathway, which comprises nuclei HVC and the downstream motor cortex analogue robust nucleus of the arcopallium (RA) (Simpson & Vicario 1990, Yu & Margoliash 1996), and the anterior forebrain pathway (AFP), a song-specialized basal ganglia–thalamocortical circuit that indirectly connects HVC and RA (Perkel 2004) (**Figure 2a**). The song is encoded and generated by the motor pathway (Fee & Scharff 2010). The AFP is important for song learning but not for producing the song in adult birds (Bottjer et al. 1984, Scharff & Nottebohm 1991). Its output nucleus, the lateral magnocellular nucleus of the anterior nidopallium (LMAN), projects to RA and introduces variability into the RA motor program and, consequently, the song (Kao et al. 2005; Ölveczky et al. 2005, 2011). If activity in LMAN is silenced, the otherwise variable juvenile song becomes highly stereotyped (Ölveczky et al. 2005) (**Figure 3c**), and if LMAN is lesioned in juvenile birds, the song learning process stalls (Bottjer et al. 1984). Furthermore, LMAN neurons generate activity patterns that vary from song to song, consistent with a role for LMAN in inducing motor variability (Ölveczky et al. 2005, 2011) (**Figure 3d**). These results suggest that variability is not simply due to intrinsic noise in the descending motor pathway but is introduced into RA by a dedicated circuit that is required for song learning.

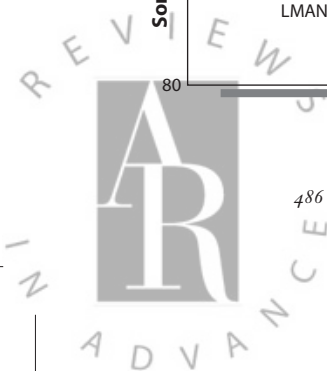
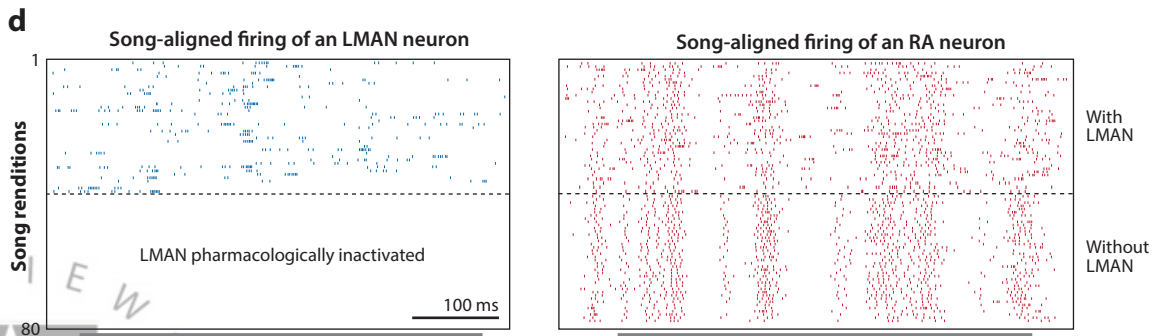
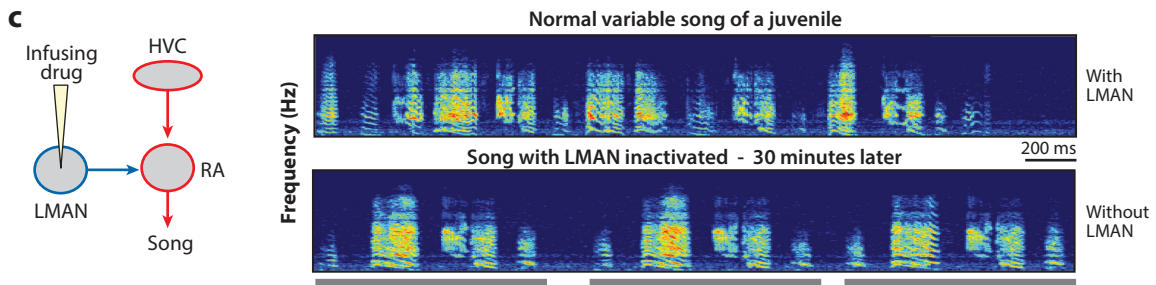
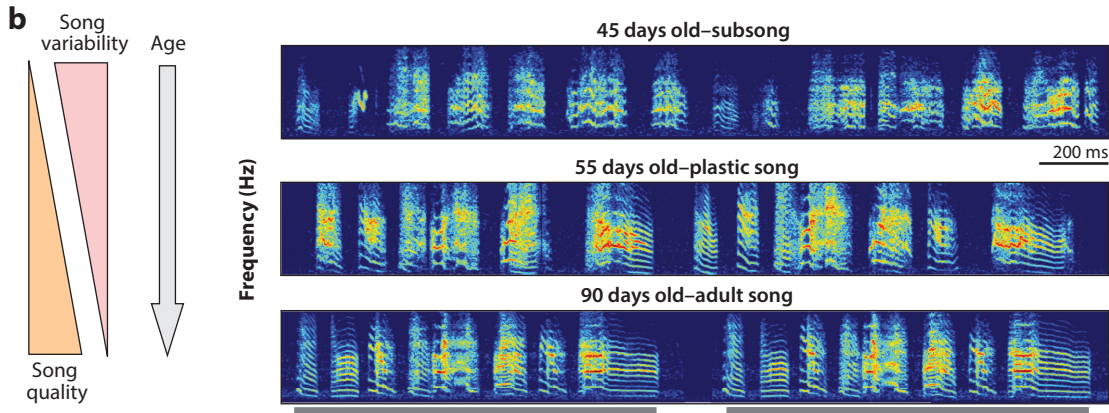
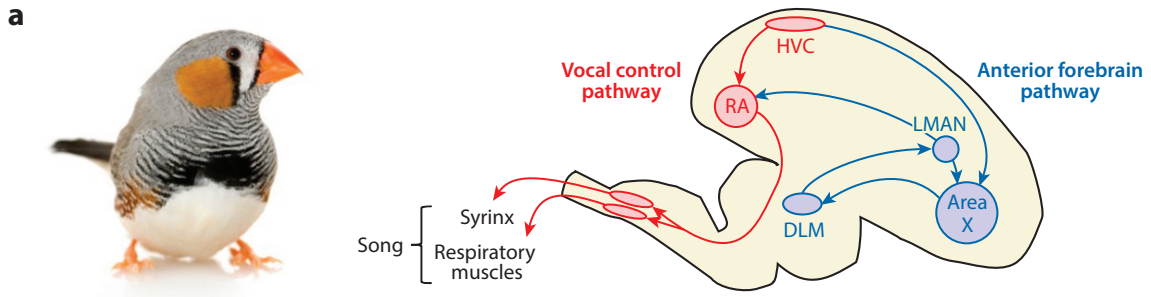
## SONGBIRD STUDIES LINKING MOTOR VARIABILITY AND LEARNING

Despite a powerful theoretical framework linking variability and reinforcement learning, researchers continue to debate the degree to which motor variability is conducive to motor learning (Cohen & Sternad 2008, He et al. 2016, Wu et al. 2014). A direct demonstration of how variability can be used as a substrate for motor learning came from Tumer & Brainard's (2007) experiments in songbirds (see the sidebar titled What Songbirds Tell Us About Motor Variability and Learning). Although adult birds sing highly stereotyped songs, small rendition-to-rendition variability in, for example, the pitch of their vocalizations can be detected in real time. When a negative reinforcer, in the form of a loud noise burst, is delivered following certain pitch variants, the song gradually and persistently shifts away from those. This suggests that variability even in well-learned skills can reflect meaningful motor exploration that supports continuous learning and optimization of performance (Tumer & Brainard 2007).

Taking it a step further, Andalman & Fee (2009) examined the neural mechanisms underlying this reinforcement learning process and found that the output of the anterior forebrain pathway (AFP), LMAN (i.e., the same brain region that introduces much of the exploratory motor variability), contributes the error-correcting signal that drives learning (see also Warren et al. 2011). This suggests a two-stage learning process, in which reinforcement learning in the AFP produces an error-correcting premotor signal at the level of LMAN, which then biases the motor output (here, the pitch) in a more favorable direction. This error-correcting bias then becomes incorporated into the motor pathway over time (Andalman & Fee 2009, Tesileanu et al. 2016, Warren et al. 2011).

Although the AFP is essential for song learning, its output, LMAN, is only responsible for about half of the rendition-to-rendition variability in adult birdsong (Kao et al. 2005), raising the question of whether the reinforcement learning algorithm implemented in the AFP can make use of motor variability originating elsewhere in the song circuit. Using a pharmacological strategy to reduce LMAN's contribution to variability while keeping the AFP circuit otherwise unperturbed,







Charlesworth and colleagues (2012) argued that the error-correcting learning signal mediated by LMAN can be built from variability contributed by other parts of the circuit. This implies that the AFP has an efference or sensory copy of this variability and uses it to update its output adaptively. Importantly, this suggests that exploratory variability need not be generated by the circuits implementing the reinforcement learning, as long as information about the variability is relayed to those circuits.

## THE ROLE OF VARIABILITY IN HUMAN MOTOR LEARNING

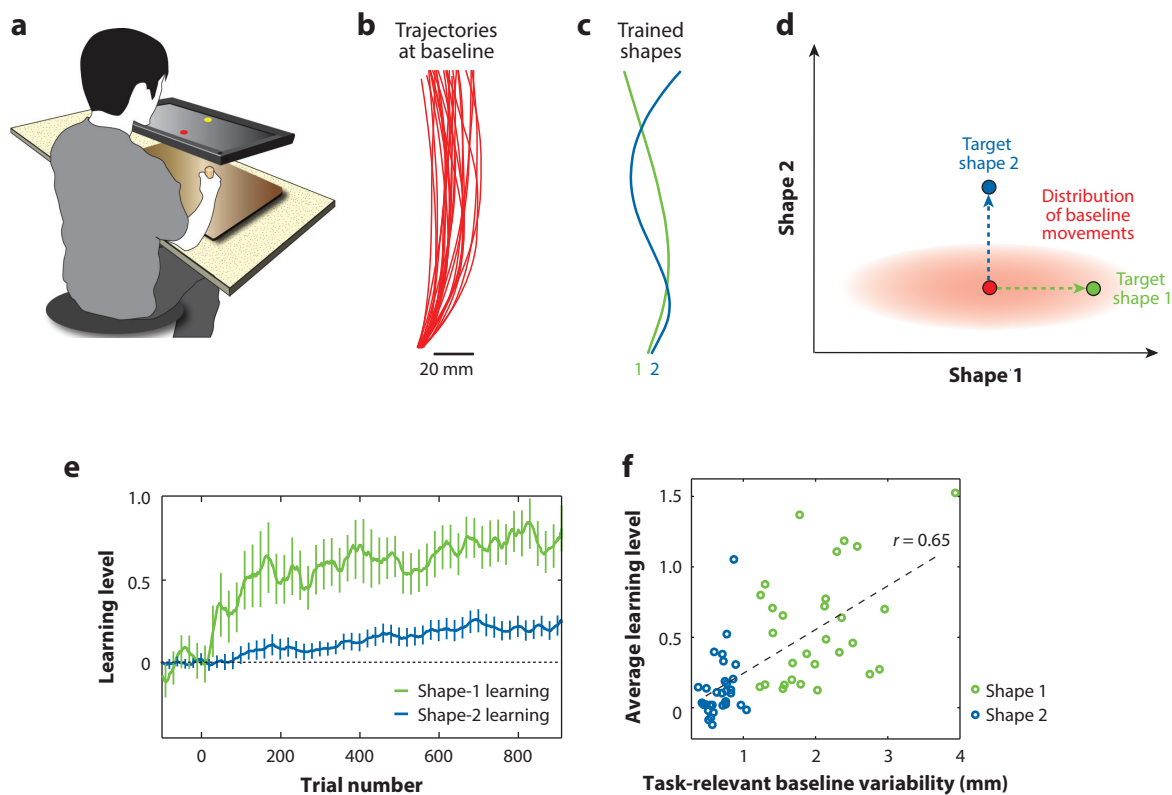
Work in songbirds established that the brain can actively generate and make use of motor variability for the purpose of learning (Kao et al. 2005, Ölveczky et al. 2005, Tumer & Brainard 2007) (Figure 3). Picking up this thread, Wu and colleagues (2014) set out to test whether this may generalize to human motor learning. They argued that if motor variability is conducive to learning, then its structure and magnitude should predict learning ability across individuals and tasks. They first tested their hypothesis in a reinforcement learning–based paradigm, in which subjects were trained to modify the trajectories of ballistic reaching movements to better approximate one of two predefined shapes (Figure 4a). The subjects were not aware of the shapes but received a numerical reward after each trial reflecting how well they had done.

The authors found that learning rates depended strongly on the degree to which the subjects' baseline motor variability aligned with the prescribed shapes; higher task-relevant variability predicted faster learning rates both across different tasks and across individuals within a single task (Figure 4). These results demonstrated that the human brain can make use of trial-to-trial motor variability to update control policies and motor output in a reinforcement learning paradigm.

To probe whether their findings also generalize to other forms of learning, the authors similarly probed the relationship between variability and learning in an error-based motor adaptation paradigm. Here, subjects were tasked with modifying target-specific reaching movements in response to external force-field perturbations. This form of learning is typically framed as an example of optimal feedback control and believed to rely on sensory prediction errors updating an internal model that is used in generating motor output (i.e., deterministic processes that are corruptible by noise) (Haith & Krakauer 2013, Krakauer & Mazzoni 2011, van Beers et al. 2013). Surprisingly, the results were similar to the reinforcement learning paradigm, with higher task-relevant variability predicting faster learning rates across both subjects and tasks. This suggests a broader role for variability in motor learning, including in error-based paradigms.

### Figure 3

Vocal variability in songbirds is generated by a basal ganglia–like circuit. Research on the courtship song of zebra finches has informed the link between motor variability and learning. (a) The song is generated by the vocal control pathway (red) comprising HVC, RA, and brainstem motor regions. The anterior forebrain pathway (blue), a basal ganglia–thalamo–cortical circuit, is important for song learning but not for producing learned song. (b) Spectrograms of zebra finch song at different stages of song learning, showing that learning is associated with a gradual decrease in song variability and an increase in song quality, as defined by the similarity to the song model being imitated (not shown). Grey lines denote the song motif of the bird, which crystallizes to the same syllable sequence late in learning. (c) Inactivating LMAN (left) causes a dramatic reduction in song variability in juvenile birds (right). Song spectrograms from the same bird before and immediately after LMAN inactivation. Data from Ölveczky et al. (2005). (d) Inactivating LMAN reduces the rendition-to-rendition variability of RA neurons. (Left) Activity patterns of an LMAN neuron in a juvenile bird, aligned to a recognizable song motif (i.e., syllable sequence). Each row of spikes represents the activity during one rendition of the song motif. Note the high degree of rendition-to-rendition variability. (Right) Recording from the same RA neuron in a juvenile singing bird with and without pharmacological inactivation of LMAN. Rendition-to-rendition variability in the RA neurons with LMAN silencing is reduced dramatically. Data from Ölveczky et al. (2011). Abbreviations: LMAN, lateral magnocellular nucleus of the anterior nidopallium; RA, robust nucleus of the arcopallium.



**Figure 4**

Structure of motor variability predicts learning rates in a reinforcement-based task. (a) Subjects were asked to move a manipulandum between two points on a screen (red and yellow). (b) Example baseline movements from one participant showing the pattern of trial-to-trial variability. (c) The subjects were rewarded based on how well their movements reflected predefined shapes (two shapes were used in different experiments). The shapes were never made explicit to the subjects, making it a trial-and-error learning task. (d) Schematic showing baseline variability projected into the space defined by the two shapes. The target shapes were chosen to make sure that, on average, task-relevant variability was higher for Shape 1. (e) Average learning curves showing that subjects generally learned Shape 1 faster than Shape 2. (f) Task-relevant variability is correlated with learning level both across tasks (different colors) and individuals (circles). Figure adapted from Wu et al. (2014).

Whether this reflects a contribution of reinforcement learning processes to learning driven by sensory prediction errors remains to be better understood (Huang et al. 2011, Wu et al. 2014).

Confusing matters a bit, a recent study using a visuomotor adaptation paradigm did not find a clear relationship between motor variability and the rate of motor adaptation (He et al. 2016). Although there were several methodological differences between this and the previous study (Wu et al. 2014), the difference in how baseline variability was estimated may help resolve the discrepancy and further illuminate the relationship between variability and learning. In contrast to the earlier study, He and colleagues (2016) measured baseline variability with task-relevant feedback available to subjects (i.e., they could see how their movements deviated from the desired trajectory). This is pertinent because such task-relevant feedback allows errors in the brain's internal model for generating movements, also referred to as planning noise (Figure 2), to be corrected (Scholz & Schöner 1999, van Beers et al. 2013). Central planning circuits, which have ready

access to task-related feedback and whose neural activity patterns exhibit slow drift correlated with behavior (Chaisanguanthum et al. 2014), are likely to be the main source of this variability. In the absence of feedback, such errors could accumulate like the excursions in a random walk process, leading to slow drift in motor output (van Beers 2009, 2013).

In contrast, variability from execution noise, which probably originates in the motor periphery (**Figure 2**), is not expected to accumulate, even in the absence of corrective feedback (van Beers 2009, 2013). This means that experimental measurements of overall motor variability made with task-relevant feedback, as in He et al. (2016), may emphasize execution noise over planning noise, whereas measurements made without feedback, as in Wu et al. (2014), may primarily reflect planning noise because it allows drift in central planning circuits to contribute more to total motor variability.

If the effect of task-relevant feedback on measurements of baseline motor variability indeed explains the discrepancy between the two studies, and this remains to be rigorously tested, it would support the idea that motor variability originating from central circuits (i.e., planning noise) is the main substrate for learning-related motor exploration.

## LEARNING-DEPENDENT REGULATION OF MOTOR VARIABILITY

Given that motor variability can be beneficial for learning new motor skills but detrimental to expert performance, it would be desirable to regulate it in a way that optimizes its utility. In the context of reinforcement learning, how much to explore (i.e., vary motor output) relates to the exploration-exploitation dilemma (Kaelbling et al. 1996, Sutton & Barto 1998). Simply put, the dilemma is whether to explore new options (e.g., actions or movement patterns) or exploit those with known values. How the nervous system deals with this dilemma has been studied extensively in the context of decision making (Cohen et al. 2007) but less so in the realm of motor control. However, reinforcement learning theory gives researchers intuition into how variability should be regulated. First, exploration should decrease with practice as more information becomes available about the values of various actions (**Figure 1**). Second, the reward-context in which actions are generated should influence the relative amount of variability, with more exploitation in high-stakes situations. This is because there is more to lose from exploring when more is on the line. Third, if the relative reward of an action is reduced, it could signal that the overall reward landscape has changed and that the system should explore (i.e., increase motor variability) to find better solutions.

In agreement with the first point, trial-to-trial variability does generally decrease with practice (“practice makes perfect”). Much of what we know about the neural circuit mechanisms underlying such learning-related decreases in motor variability comes, yet again, from research in songbirds. As discussed above (see the sidebar titled *What Songbirds Tell Us About Motor Variability and Learning*), song variability in juvenile birds is, in large part, generated by the AFP, through LMAN’s projection to RA neurons (**Figure 3c,d**). This projection dominates and drives the RA motor program, and consequently the song, early in learning (Aronov et al. 2008, Ölveczky et al. 2011). Because the activity patterns of LMAN neurons vary across renditions (Hessler & Doupe 1999, Kao et al. 2008, Ölveczky et al. 2005), this results in variable song in juveniles (**Figure 3c,d**).

As learning proceeds, control of the RA motor program shifts gradually away from LMAN, which is variable, to HVC, which produces stereotyped activity patterns (Hahnloser et al. 2002), resulting in less variable song late in learning (Aronov et al. 2008). Probing the synaptic connectivity in RA over the course of song learning, Garst-Orozco et al. (2015) showed that this shift in control happens not by changing the overall strength of HVC and LMAN input to RA, but rather



by strengthening and pruning HVC to RA connections. Such learning-related synaptic reorganization renders the variable LMAN input to RA less effective, leading to more stereotyped song.

Regulating motor variability through learning-related reorganization of action-specific synapses lends considerable flexibility to the process of motor sequence learning, as it reduces variability selectively in action elements that have been mastered while allowing continued exploration in others (Ravbar et al. 2012). Intriguingly, researchers have also observed learning-related synaptic changes akin to those described in songbirds in mammalian motor cortex (Fu et al. 2012, Wang et al. 2011, Xu et al. 2009), suggesting that it may be a general mechanism for regulating motor variability as a function of learning.

### CONTEXT-DEPENDENT REGULATION OF MOTOR VARIABILITY

As discussed above, reinforcement learning theory favors exploitation when stakes are high and exploration when they are not (Sutton & Barto 1998). Evidence suggests that the nervous system also regulates variability in such a context-dependent manner, producing more reproducible output in high-reward situations. Experiments in a wide variety of species, including rodents (Gharib et al. 2001, 2004), pigeons (Stahlman & Blaisdell 2011, Stahlman et al. 2010), and monkeys (Takikawa et al. 2002), have shown that animals generate more variable actions when they have been cued to expect less reward. In motor learning, however, reward expectations are typically set by performance history, not sensory cues. To investigate the influence of reward history on motor variability, a recent study in humans (Pekny et al. 2015) manipulated the reward probability for reaching movements. Increasing or decreasing reward probabilities caused arm movements to become less or more variable, respectively, suggesting that reward context can have a causal effect on the degree of motor variability.

A similar form of context-dependent regulation of motor variability is seen in songbirds, where high-stakes situations equate to those in which songs are directed to potential partners (directed singing). Songs are significantly less variable during directed singing than when birds sing undirected song (Kojima & Doupe 2011). The circuit mechanisms for this social context-dependent regulation of variability involve a dopamine-dependent switch in AFP circuit dynamics (Leblois 2013), which results in less bursty and more regular firing in LMAN neurons when a female (and dopamine) is present (Hessler & Doupe 1999, Kao et al. 2008, Woolley et al. 2014). This mechanism is distinct from learning-related reduction in variability, which involves reorganization of synaptic connectivity within the descending motor pathway (Garst-Orozco et al. 2015), suggesting (at least) two independent ways of regulating motor variability for the same behavior.

Compared with songbirds, less is known about context-dependent regulation of motor variability in mammals. An important context for any action is its reward landscape (**Figure 1**). Detecting changes in reward contingencies requires subjects to compare past and present performance, yet it is unclear over what timescales the brain tracks performance history, how changes in reward landscape are assessed, and how these computations ultimately regulate motor variability. In their recent study, Pekny et al. (2015) trained human subjects in a reinforcement learning task to reach toward a hidden target. A trial-by-trial analysis suggested that motor variability is modulated by the outcome of the past 2–3 trials. However, the analysis may have been confounded by trial-to-trial correlations in task performance (Chaisanguanthum et al. 2014, van Beers et al. 2013) that make it difficult to establish causal relationships between reward history and motor variability. In other words, an increase in motor variability could have led to decreased reward rates, rather than vice versa. Overcoming these confounds requires controlling for long-term performance history, which is more easily done with data sets containing many trials. Although collecting such large data sets can be cumbersome in human subjects, it is becoming increasingly feasible in

rodents (Poddar et al. 2013). Efforts to interrogate the effect of reward history on motor variability in rodents are currently under way (Miyamoto et al. 2015).

If the mammalian brain monitors the reward landscape and regulates motor variability based on it, where are these computations implemented? Studies on decision making have implicated basal ganglia circuits (Hamid et al. 2016, Hikosaka et al. 2014, Samejima et al. 2005, Schultz et al. 2003, Tai et al. 2012, Wang et al. 2013) as well as prefrontal regions (Matsumoto et al. 2003, Roesch & Olson 2004, Wunderlich et al. 2009) in encoding action values. The basal ganglia are also thought to be involved in invigorating movements associated with greater reward (Hamid et al. 2016, Kawagoe et al. 1998, Lauwereyns et al. 2002, Wang et al. 2013). Whether similar neural substrates are involved in regulating motor variability in a reward-history dependent manner remains to be understood.

Motor variability can be determined and shaped further by the nature and reliability of task-related sensory feedback (Osborne et al. 2005). For instance, Izawa & Shadmehr (2011) found that trial-to-trial motor variability increased substantially when humans were learning from binary reward feedback as compared to more informative sensory (visual) feedback. More generally, the sensorimotor system is thought to weight distinct sources of sensory input streams based on the amount of information that these sources provide. The degree of confidence in the sensory evidence can then feed back into the motor system to influence variability in motor output (see the sidebar titled Internal Estimates of Sensorimotor Noise Can Influence Motor Variability).

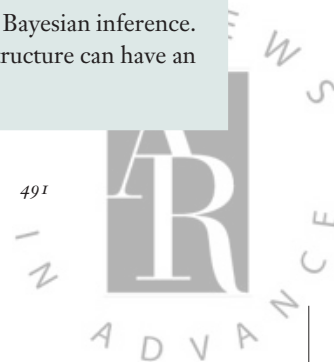
## REGULATING THE STRUCTURE OF MOTOR VARIABILITY

Given the high degree of redundancy in how the motor system controls movements and how tasks can be executed (Bernshtein 1967), certain forms of motor variability may have small effects on performance, whereas others may be more consequential. But to what extent does the nervous system distinguish task-relevant and task-irrelevant variability? Several studies have shown that

### INTERNAL ESTIMATES OF SENSORIMOTOR NOISE CAN INFLUENCE MOTOR VARIABILITY

Although this review focuses on how and why motor variability is generated, it should be noted that the motor system also monitors sensorimotor variability and uses internal estimates of it to optimize motor control strategies and performance. One salient example is in setting safety margins for grip forces, whose misestimation can have grossly asymmetric consequences. Whereas grasp can be maintained by overgripping, undergripping can lead to catastrophic failures. A recent study showed that this safety margin is determined by an adaptable internal estimate of environmental variability (Hadjiosif & Smith 2015), analogous to maintaining a greater safety margin when driving near an erratically behaving vehicle than a predictable one.

The sensorimotor system also weighs different sources of sensory information depending on their reliability, allowing it to dynamically modify the information it extracts from variable and noisy sensory inputs. For example, if greater uncertainty in the sensorimotor realm is introduced and a movement is physically perturbed or visual information about it is distorted, the gains of the corrective feedback responses are reduced. By accounting for such sensorimotor uncertainty, the reduced feedback response can improve the precision of the generated action (Franklin et al. 2012, Kording & Wolpert 2004). The rate of trial-to-trial motor adaptation has also been shown to decrease when variability is added experimentally (Wei & Koerding 2010), in line with optimal Bayesian inference. However, having an estimate of the persistence of environmental variability and its statistical structure can have an even greater effect on learning rates (Gonzalez Castro et al. 2014).



task-relevant variability is systematically reduced after repeated practice, whereas task-irrelevant variability can remain high (Kang et al. 2004, Latash & Anson 2006, Scholz & Schöner 1999, van Beers et al. 2013). These results are consistent with optimal feedback control theory, which posits that movements are planned and shaped specifically to optimize fidelity in performance (Harris & Wolpert 1998) and reduce effort (Braun et al. 2009, Izawa & Shadmehr 2008, Izawa et al. 2008, Todorov 2004, Todorov & Jordan 2002). Researchers have even suggested that larger amounts of task-irrelevant variability can afford reduced task-relevant variability (Todorov 2004, Todorov & Jordan 2002), although evidence for this is not as clear.

Shaping the structure of motor variability with such specificity can allow the motor system to exploit along dimensions that are deemed task-relevant, while enabling continued exploration (and learning) in others. Conversely, the nervous system can also selectively increase task-relevant variability when directed exploration along a particular dimension is conducive for fast learning, such as when the same task (or control policy) has to be reacquired. Wu and colleagues (2014) exposed subjects repeatedly to the same adaptation paradigm and observed a significant overall increase in the task-relevant component of their variability. This reshaping persisted long after training had terminated, suggesting a lasting and experience-dependent modification of the structure of motor variability that could promote more efficient exploration.

Interestingly, researchers have observed neural correlates of such task-specific regulation of motor variability in the activity of cortical neurons during visuomotor adaptation, where learning-related increases in trial-to-trial spiking variability are seen principally in the subset of neurons that are tuned to the movement directions being trained (Mandelblat-Cerf et al. 2009). Taken together, these results suggest that the nervous system shapes the structure of motor variability in sophisticated ways to adapt it to the specific task demands. Further follow-up studies will be required to better understand how and under what circumstances the brain sculpts motor variability and how the underlying computations are implemented in neural circuitry.

## CONCLUSIONS AND OUTLOOK

Although noise in nervous system function can often be detrimental to optimal performance, the studies we have reviewed here suggest that neural variability may also be conducive to motor learning, in line with reinforcement learning theory (Kaelbling et al. 1996, Sutton & Barto 1998). Random fluctuations (or noise) in the activity of neurons could plausibly underlie such motor exploration, but recent findings suggest that the nervous system is more deliberate and sophisticated than that and may be regulating and shaping motor variability actively to augment learning.

The link between variability and motor learning has been established, but the specifics of this relationship remain to be worked out (see the Future Issues section). Importantly, furnishing our understanding with mechanistic insight will require animal models with suitable experimental paradigms. Songbirds have proved powerful in this regard (see the sidebar titled *What Songbirds Tell Us About Motor Variability and Learning*) and have offered valuable clues. Rodent models also hold significant promise (Ölveczky 2011), given the feasibility of high-throughput and longitudinal studies (Miyamoto et al. 2015, Poddar et al. 2013) and the increasingly sophisticated ways in which their neural circuits can be manipulated (Luo et al. 2008).

A deeper understanding of the link between variability and learning will be further helped by detailed descriptions of the structure of motor variability. Increasingly sophisticated methods for tracking the movements of experimental animals at high spatiotemporal resolution (Anderson & Perona 2014, Egnor & Branson 2016) will fuel progress and allow trial-to-trial motor variability to be used and appreciated as an important tool in our quest to understand how the nervous system operates and learns.

## FUTURE ISSUES

1. Which form (or forms) of motor variability, in terms of both statistical structure and neural origin, can be harnessed for motor learning? For example, to what extent can the motor system learn from centrally versus peripherally generated motor variability?
2. What are the neural circuit mechanisms that underlie the generation of learning-related motor variability? Are there dedicated circuits in mammals akin to those described in songbirds? If so, what are these circuits and how do they function?
3. How is motor variability regulated? How are the reward landscape and other relevant contextual cues computed and monitored, and how does this information influence the amount and structure of trial-to-trial motor variability?
4. How does the nervous system implement reinforcement learning in the motor domain? Specifically, how does it reduce the dimensionality of the solution space?
5. How are action variants that improve performance reinforced, reproduced, and ultimately consolidated for long-term improvements in motor output?

## DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

## ACKNOWLEDGMENTS

We thank Jesse Goldberg, Naoshige Uchida, Sam Gershman, Reza Shadmehr, and members of the Ölveczky lab for comments on the manuscript.

## LITERATURE CITED

- Ali F, Otchy TM, Pehlevan C, Fantana AL, Burak Y, Ölveczky BP. 2013. The basal ganglia is necessary for learning spectral, but not temporal, features of birdsong. *Neuron* 80:494–506
- Alpaydin E. 2014. *Introduction to Machine Learning*. Cambridge, MA: MIT Press
- Andalman AS, Fee MS. 2009. A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. *PNAS* 106:12518–23
- Anderson DJ, Perona P. 2014. Toward a science of computational ethology. *Neuron* 84:18–31
- Aronov D, Andalman AS, Fee MS. 2008. A specialized forebrain circuit for vocal babbling in the juvenile songbird. *Science* 320:630–34
- Babloyantz A, Salazar JM, Nicolis C. 1985. Evidence of chaotic dynamics of brain activity during the sleep cycle. *Phys. Lett. A* 111:152–56
- Baddeley RJ, Ingram HA, Miall RC. 2003. System identification applied to a visuomotor task: near-optimal human performance in a noisy changing task. *J. Neurosci.* 23:3066–75
- Beck JM, Ma WJ, Pitkow X, Latham PE, Pouget A. 2012. Not noisy, just wrong: the role of suboptimal inference in behavioral variability. *Neuron* 74:30–39
- Bellman R. 1957. *Dynamic Programming*. Mineola, NY: Dover Publ.
- Bernshstein NA. 1967. *The Co-ordination and Regulation of Movements*. Oxford, UK/New York: Pergamon Press
- Björklund A, Dunnett SB. 2007. Dopamine neuron systems in the brain: an update. *Trends Neurosci.* 30:194–202
- Bottjer SW, Miesner EA, Arnold AP. 1984. Forebrain lesions disrupt development but not maintenance of song in passerine birds. *Science* 224:901–3



- Botvinick MM. 2012. Hierarchical reinforcement learning and decision making. *Curr. Opin. Neurobiol.* 22:956–62
- Braun DA, Aertsen A, Wolpert DM, Mehring C. 2009. Learning optimal adaptation strategies in unpredictable motor tasks. *J. Neurosci.* 29:6472–78
- Bruno RM, Sakmann B. 2006. Cortex is driven by weak but synchronously active thalamocortical synapses. *Science* 312:1622–27
- Calvin WH, Stevens CF. 1968. Synaptic noise and other sources of randomness in motoneuron interspike intervals. *J. Neurophysiol.* 31:574–87
- Chaisanguanthum KS, Shen HH, Sabes PN. 2014. Motor variability arises from a slow random walk in neural state. *J. Neurosci.* 34:12071–80
- Charlesworth JD, Warren TL, Brainard MS. 2012. Covert skill learning in a cortical-basal ganglia circuit. *Nature* 486:251–55
- Churchland MM, Afshar A, Shenoy KV. 2006. A central source of movement variability. *Neuron* 52:1085–96
- Clamann HP. 1969. Statistical analysis of motor unit firing patterns in a human skeletal muscle. *Biophys. J.* 9:1233–51
- Cohen JD, McClure SM, Yu AJ. 2007. Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philos. Trans. R. Soc. B* 362:933–42
- Cohen RG, Sternad D. 2008. Variability in motor learning: relocating, channeling and reducing noise. *Exp. Brain Res.* 193:69–83
- Daw ND, O’Doherty JP, Dayan P, Seymour B, Dolan RJ. 2006. Cortical substrates for exploratory decisions in humans. *Nature* 441:876–79
- Diesmann M, Gewaltig M-O, Aertsen A. 1999. Stable propagation of synchronous spiking in cortical neural networks. *Nature* 402:529–33
- Doyon J, Benali H. 2005. Reorganization and plasticity in the adult brain during learning of motor skills. *Curr. Opin. Neurobiol.* 15:161–67
- Egnor SER, Branson K. 2016. Computational analysis of behavior. *Annu. Rev. Neurosci.* 39:217–36
- Eshel N, Bukwich M, Rao V, Hemmelder V, Tian J, Uchida N. 2015. Arithmetic and local circuitry underlying dopamine prediction errors. *Nature* 525:243–46
- Faisal AA, Laughlin SB. 2007. Stochastic simulations on the reliability of action potential propagation in thin axons. *PLOS Comput. Biol.* 3:e79
- Faisal AA, Selen LPJ, Wolpert DM. 2008. Noise in the nervous system. *Nat. Rev. Neurosci.* 9:292–303
- Fee MS, Scharff C. 2010. The songbird as a model for the generation and learning of complex sequential behaviors. *ILAR J.* 51:362–77
- Fitts PM. 1954. The information capacity of the human motor system in controlling the amplitude of movement. *J. Exp. Psychol.* 47:381–91
- Flament D, Hore J. 1988. Relations of motor cortex neural discharge to kinematics of passive and active elbow movements in the monkey. *J. Neurophysiol.* 60:1268–84
- Franklin S, Wolpert DM, Franklin DW. 2012. Visuomotor feedback gains upregulate during the learning of novel dynamics. *J. Neurophysiol.* 108:467–78
- Fu M, Yu X, Lu J, Zuo Y. 2012. Repetitive motor learning induces coordinated formation of clustered dendritic spines in vivo. *Nature* 483:92–95
- Garst-Orozco J, Babadi B, Ölveczky BP. 2015. A neural circuit mechanism for regulating vocal variability during song learning in zebra finches. *eLife* 3:e03697
- Gershman SJ, Pesaran B, Daw ND. 2009. Human reinforcement learning subdivides structured action spaces by learning effector-specific values. *J. Neurosci.* 29:13524–31
- Gharib A, Derby S, Roberts S. 2001. Timing and the control of variation. *J. Exp. Psychol. Anim. Behav. Process.* 27:165–78
- Gharib A, Gade C, Roberts S. 2004. Control of variation by reward probability. *J. Exp. Psychol. Anim. Behav. Process.* 30:271–82
- Gonzalez Castro LN, Hadjiosif AM, Hemphill MA, Smith MA. 2014. Environmental consistency determines the rate of motor adaptation. *Curr. Biol.* 24:1050–61
- Hadjiosif AM, Smith MA. 2015. Flexible control of safety margins for action based on environmental variability. *J. Neurosci.* 35:9106–21



- Hahnloser RH, Kozhevnikov AA, Fee MS. 2002. An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature* 419:65–70
- Haith AM, Krakauer JW. 2013. Model-based and model-free mechanisms of human motor learning. In *Progress in Motor Control*, ed. MJ Richardson, MA Riley, K Shockley, pp. 1–21. New York: Springer
- Haith AM, Reppert TR, Shadmehr R. 2012. Evidence for hyperbolic temporal discounting of reward in control of movements. *J. Neurosci.* 32:11727–36
- Hamid AA, Pettibone JR, Mabrouk OS, Hetrick VL, Schmidt R, et al. 2016. Mesolimbic dopamine signals the value of work. *Nat. Neurosci.* 19:117–26
- Hamilton AF de C, Jones KE, Wolpert DM. 2004. The scaling of motor noise with muscle strength and motor unit number in humans. *Exp. Brain Res.* 157:417–30
- Harris CM, Wolpert DM. 1998. Signal-dependent noise determines motor planning. *Nature* 394:780–84
- He K, Liang Y, Abdollahi F, Bittmann MF, Kording K, Wei K. 2016. The statistical determinants of the speed of motor learning. *PLOS Comput. Biol.* 12:e1005023
- Herzfeld DJ, Shadmehr R. 2014. Motor variability is not noise, but grist for the learning mill. *Nat. Neurosci.* 17:149–50
- Hessler NA, Doupe AJ. 1999. Social context modulates singing-related neural activity in the songbird forebrain. *Nat. Neurosci.* 2:209–11
- Hikosaka O, Kim HF, Yasuda M, Yamamoto S. 2014. Basal ganglia circuits for reward value-guided behavior. *Annu. Rev. Neurosci.* 37:289–306
- Hinton GE, Osindero S, Teh Y-W. 2006. A fast learning algorithm for deep belief nets. *Neural. Comput.* 18:1527–54
- Horikawa Y. 1991. Noise effects on spike propagation in the stochastic Hodgkin-Huxley models. *Biol. Cybern.* 66:19–25
- Huang VS, Haith A, Mazzoni P, Krakauer JW. 2011. Rethinking motor learning and savings in adaptation paradigms: model-free memory for successful actions combines with internal models. *Neuron* 70:787–801
- Immelmann K. 1969. Song development in the zebra finch and other estrildid finches. In *Bird Vocalizations*, ed. RA Hinde, pp. 61–74. Cambridge, UK: Cambridge Univ. Press
- Izawa J, Shadmehr R. 2008. On-line processing of uncertain information in visuomotor control. *J. Neurosci.* 28:11360–68
- Izawa J, Shadmehr R. 2011. Learning from sensory and reward prediction errors during motor adaptation. *PLOS Comput. Biol.* 7:e1002012
- Izawa J, Rane T, Donchin O, Shadmehr R. 2008. Motor adaptation as a process of reoptimization. *J. Neurosci.* 28:2883–91
- Jones KE, Hamilton AF de C, Wolpert DM. 2002. Sources of signal-dependent noise during isometric force production. *J. Neurophysiol.* 88:1533–44
- Kaelbling LP, Littman ML, Moore AW. 1996. Reinforcement learning: a survey. arXiv:cs/9605103
- Kakei S, Hoffman DS, Strick PL. 1999. Muscle and movement representations in the primary motor cortex. *Science* 285:2136–39
- Kang N, Shinohara M, Zatsiorsky VM, Latash ML. 2004. Learning multi-finger synergies: an uncontrolled manifold analysis. *Exp. Brain Res.* 157:336–50
- Kao MH, Doupe AJ, Brainard MS. 2005. Contributions of an avian basal ganglia–forebrain circuit to real-time modulation of song. *Nature* 433:638–43
- Kao MH, Wright BD, Doupe AJ. 2008. Neurons in a forebrain nucleus required for vocal plasticity rapidly switch between precise firing and variable bursting depending on social context. *J. Neurosci.* 28:13232–47
- Katz B, Miledi R. 1970. Membrane noise produced by acetylcholine. *Nature* 226:962–63
- Kawagoe R, Takikawa Y, Hikosaka O. 1998. Expectation of reward modulates cognitive signals in the basal ganglia. *Nat. Neurosci.* 1:411–16
- Kojima S, Doupe AJ. 2011. Social performance reveals unexpected vocal competency in young songbirds. *PNAS* 108:1687–92
- Kording KP, Wolpert DM. 2004. Bayesian integration in sensorimotor learning. *Nature* 427:244–47
- Kormushev P, Calinon S, Caldwell DG. 2010. Robot motor skill coordination with EM-based reinforcement learning. *2010 IEEE/RSJ Int. Conf. Intell. Robot. Syst. (IROS)*, pp. 3232–37



- Krakauer JW, Mazzoni P. 2011. Human sensorimotor learning: adaptation, skill, and beyond. *Curr. Opin. Neurobiol.* 21:636–44
- Lashley KS. 1933. Integrative functions of the cerebral cortex. *Physiol. Rev.* 13:1–42
- Latah ML, Anson JG. 2006. Synergies in health and disease: relations to adaptive changes in motor coordination. *Phys. Ther.* 86:1151–60
- Lauwereyns J, Watanabe K, Coe B, Hikosaka O. 2002. A neural correlate of response bias in monkey caudate nucleus. *Nature* 418:413–17
- Leblois A. 2013. Social modulation of learned behavior by dopamine in the basal ganglia: insights from songbirds. *J. Physiol.* 107:219–29
- LeCun Y, Bengio Y, Hinton G. 2015. Deep learning. *Nature* 521:436–44
- LeCun Y, Bottou L, Bengio Y, Haffner P. 1998. Gradient-based learning applied to document recognition. *Proc. IEEE* 86:2278–324
- Lee D, Seo H, Jung MW. 2012. Neural basis of reinforcement learning and decision making. *Annu. Rev. Neurosci.* 35:287–308
- Lee H, Pham P, Largman Y, Ng AY. 2009. Unsupervised feature learning for audio classification using convolutional deep belief networks. In *Advances in Neural Information Processing Systems 22*, ed. Y Bengio, D Schuurmans, JD Lafferty, CKI Williams, A Culotta, pp. 1096–104. Vancouver, Can.: Neural. Inf. Proc. Syst.
- Lemon RN. 2008. Descending pathways in motor control. *Annu. Rev. Neurosci.* 31:195–218
- Litwin-Kumar A, Doiron B. 2012. Slow dynamics and high variability in balanced cortical networks with clustered connections. *Nat. Neurosci.* 15:1498–505
- London M, Roth A, Beeren L, Hausser M, Latham PE. 2010. Sensitivity to perturbations in vivo implies high noise and suggests rate coding in cortex. *Nature* 466:123–27
- Luo L, Callaway E, Svoboda K. 2008. Genetic dissection of neural circuits. *Neuron* 57:634–60
- Mainen ZF, Sejnowski TJ. 1995. Reliability of spike timing in neocortical neurons. *Science* 268:1503
- Mandelblat-Cerf Y, Paz R, Vaadia E. 2009. Trial-to-trial variability of single cells in motor cortices is dynamically modified during visuomotor adaptation. *J. Neurosci.* 29:15053–62
- Marcos E, Pani P, Brunamonti E, Deco G, Ferraina S, Verschure P. 2013. Neural variability in premotor cortex is modulated by trial history and predicts behavioral performance. *Neuron* 78:249–55
- Matsumoto K, Suzuki W, Tanaka K. 2003. Neuronal correlates of goal-based motor selection in the prefrontal cortex. *Science* 301:229–32
- Miyamoto YR, Dhawale AK, Smith MA, Ölveczky BP. 2015. Investigating reward-based regulation of task-relevant motor variability in rats. Presented at Soc. Neurosci. Conf., Oct. 21, Chicago
- Newell KM. 1993. *Variability and Motor Control*. Champaign, IL: Hum. Kinet. Publ.
- Niv Y. 2009. Reinforcement learning in the brain. *J. Math. Psychol.* 53:139–54
- Nudo RJ, Milliken GW, Jenkins WM, Merzenich MM. 1996. Use-dependent alterations of movement representations in primary motor cortex of adult squirrel monkeys. *J. Neurosci.* 16:785–807
- Ölveczky BP. 2011. Motoring ahead with rodents. *Curr. Opin. Neurobiol.* 21:571–78
- Ölveczky BP, Andalman AS, Fee MS. 2005. Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLOS Biol.* 3:e153
- Ölveczky BP, Otchy TM, Goldberg JH, Aronov D, Fee MS. 2011. Changes in the neural control of a complex motor sequence during learning. *J. Neurophysiol.* 106:386–97
- Opris I, Lebedev M, Nelson RJ. 2011. Motor planning under unpredictable reward: modulations of movement vigor and primate striatum activity. *Front. Neurosci.* 5:61
- Osborne LC, Lisberger SG, Bialek W. 2005. A sensory source for motor variation. *Nature* 437:412–16
- Parr R. 1998. Reinforcement learning with hierarchies of machines. *Adv. Neural Inf. Proc. Syst.* 10:1043–49
- Pekny SE, Izawa J, Shadmehr R. 2015. Reward-dependent modulation of movement variability. *J. Neurosci.* 35:4015–24
- Perkel DJ. 2004. Origin of the anterior forebrain pathway. *Ann. N. Y. Acad. Sci.* 1016:736–48
- Peters J, Schaal S. 2008. Reinforcement learning of motor skills with policy gradients. *Neural Netw.* 21:682–97
- Poddar R, Kawai R, Ölveczky BP. 2013. A fully automated high-throughput training system for rodents. *PLOS ONE* 8:e83171

- Price B, Boutillier C. 2003. Accelerating reinforcement learning through implicit imitation. *J. Artif. Intell. Res.* 19:569–629
- Ravbar P, Lipkind D, Parra LC, Tchernichovski O. 2012. Vocal exploration is locally regulated during song learning. *J. Neurosci.* 32:3422–32
- Renart A, Machens CK. 2014. Variability in neural activity and behavior. *Curr. Opin. Neurobiol.* 25:211–20
- Roesch MR, Olson CR. 2004. Neuronal activity related to reward value and motivation in primate frontal cortex. *Science* 304:307–10
- Samejima K, Ueda Y, Doya K, Kimura M. 2005. Representation of action-specific reward values in the striatum. *Science* 310:1337–40
- Sanes JN, Donoghue JP. 2000. Plasticity and primary motor cortex. *Annu. Rev. Neurosci.* 23:393–415
- Scharff C, Nottebohm F. 1991. A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. *J. Neurosci.* 11:2896–913
- Scheidt RA, Dingwell JB, Mussa-Ivaldi FA. 2001. Learning to move amid uncertainty. *J. Neurophysiol.* 86:971–85
- Schneidman E, Freedman B, Segev I. 1998. Ion channel stochasticity may be critical in determining the reliability and precision of spike timing. *Neural Comput.* 10:1679–703
- Scholz JP, Schönner G. 1999. The uncontrolled manifold concept: identifying control variables for a functional task. *Exp. Brain Res.* 126:289–306
- Schultz W. 1998. Predictive reward signal of dopamine neurons. *J. Neurophysiol.* 80:1–27
- Schultz W, Dayan P, Montague PR. 1997. A neural substrate of prediction and reward. *Science* 275:1593–99
- Schultz W, Tremblay L, Hollerman JR. 2003. Changes in behavior-related neuronal activity in the striatum during learning. *Trends Neurosci.* 26:321–28
- Shadmehr R, Huang HJ, Ahmed AA. 2016. A representation of effort in decision-making and motor control. *Curr. Biol.* 26:1929–34
- Silver D, Huang A, Maddison CJ, Guez A, Sifre L, et al. 2016. Mastering the game of Go with deep neural networks and tree search. *Nature* 529:484–89
- Simonyan K, Zisserman A. 2014. Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556 [cx.CV]
- Simpson HB, Vicario DS. 1990. Brain pathways for learned and unlearned vocalizations differ in zebra finches. *J. Neurosci.* 10:1541–56
- Skinner BF. 1938. *The Behavior of Organisms*. New York: Appleton Century Crofts
- Skinner BF. 1963. Operant behavior. *Am. Psychol.* 18:503–15
- Skinner BF. 1981. Selection by consequences. *Science* 213:501–4
- Smith MA, Ghazizadeh A, Shadmehr R. 2006. Interacting adaptive processes with different timescales underlie short-term motor learning. *PLOS Biol.* 4:e179
- Stahlman WD, Blaisdell AP. 2011. The modulation of operant variation by the probability, magnitude, and delay of reinforcement. *Learn. Motiv.* 42:221–36
- Stahlman WD, Roberts S, Blaisdell AP. 2010. Effect of reward probability on spatial and temporal variation. *J. Exp. Psychol. Anim. Behav. Process.* 36:77–91
- Stein RB, Gossen ER, Jones KE. 2005. Neuronal variability: noise or part of the signal? *Nat. Rev. Neurosci.* 6:389–97
- Sutton RS, Barto AG. 1998. *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press
- Tai L-H, Lee AM, Benavidez N, Bonci A, Wilbrecht L. 2012. Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. *Nat. Neurosci.* 15:1281–89
- Takikawa Y, Kawagoe R, Itoh H, Nakahara H, Hikosaka O. 2002. Modulation of saccadic eye movements by predicted reward outcome. *Exp. Brain Res.* 142:284–91
- Tchernichovski O, Mitra PP, Lints T, Nottebohm F. 2001. Dynamics of the vocal imitation process: how a zebra finch learns its song. *Science* 291:2564
- Tesileanu T, Olveczky B, Balasubramanian V. 2016. Matching tutor to student: rules and mechanisms for efficient two-stage learning in neural circuits. *bioRxiv* 71910. <http://dx.doi.org/10.1101/071910>
- Thorndike EL. 1898. *Animal Intelligence: An Experimental Study of the Associative Processes in Animals*. New York: Macmillan

- Todorov E. 2004. Optimality principles in sensorimotor control. *Nat. Neurosci.* 7:907–15
- Todorov E, Jordan MI. 2002. Optimal feedback control as a theory of motor coordination. *Nat. Neurosci.* 5:1226–35
- Tumer EC, Brainard MS. 2007. Performance variability enables adaptive plasticity of “crystallized” adult birdsong. *Nature* 450:1240–44
- van Beers RJ. 2007. The sources of variability in saccadic eye movements. *J. Neurosci.* 27:8757–70
- van Beers RJ. 2009. Motor learning is optimally tuned to the properties of motor noise. *Neuron* 63:406–17
- van Beers RJ, Haggard P, Wolpert DM. 2004. The role of execution noise in movement variability. *J. Neurophysiol.* 91:1050–63
- van Beers RJ, Brenner E, Smeets JBJ. 2013. Random walk of motor planning in task-irrelevant dimensions. *J. Neurophysiol.* 109:969–77
- van Rossum MCW, O’Brien BJ, Smith RG. 2003. Effects of noise on the spike timing precision of retinal ganglion cells. *J. Neurophysiol.* 89:2406–19
- Vogels TP, Rajan K, Abbott LF. 2005. Neural network dynamics. *Annu. Rev. Neurosci.* 28:357–76
- van Vreeswijk C, Sompolinsky H. 1996. Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science* 274:1724–26
- Wang AY, Miura K, Uchida N. 2013. The dorsomedial striatum encodes net expected return, critical for energizing performance vigor. *Nat. Neurosci.* 16:639–47
- Wang L, Conner JM, Rickert J, Tuszynski MH. 2011. Structural plasticity within highly specific neuronal populations identifies a unique parcellation of motor learning in the adult brain. *PNAS* 108:2545–50
- Warren TL, Tumer EC, Charlesworth JD, Brainard MS. 2011. Mechanisms and time course of vocal learning and consolidation in the adult songbird. *J. Neurophysiol.* 106:1806–21
- Wei K, Koerding K. 2010. Uncertainty of feedback and state estimation determines the speed of motor adaptation. *Front. Comput. Neurosci.* 4:11
- White JA, Rubinstein JT, Kay AR, White JA, Rubinstein JT, et al. 2000. Channel noise in neurons. *Trends Neurosci.* 23:131–37
- Woolley SC, Rajan R, Joshua M, Doupe AJ. 2014. Emergence of context-dependent variability across a basal ganglia network. *Neuron* 82:208–23
- Wu HG, Miyamoto YR, Gonzalez Castro LN, Ölveczky BP, Smith MA. 2014. Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. *Nat. Neurosci.* 17:312–21
- Wunderlich K, Rangel A, O’Doherty JP. 2009. Neural computations underlying action-based decision making in the human brain. *PNAS* 106:17199–204
- Xu T, Yu X, Perlik AJ, Tobin WF, Zweig JA, et al. 2009. Rapid formation and selective stabilization of synapses for enduring motor memories. *Nature* 462:915–19
- Yu AC, Margoliash D. 1996. Temporal hierarchical control of singing in birds. *Science* 273:1871–75

