



Published in final edited form as:

Annu Rev Neurosci. 2017 July 25; 40: 479–498. doi:10.1146/annurev-neuro-072116-031548.

The Role of Variability in Motor Learning

Ashesh K. Dhawale¹, Maurice A. Smith², and Bence P. Ölveczky^{1,#}

¹Department of Organismic and Evolutionary Biology and the Center for Brain Science, Harvard University

²John Paulson School of Engineering and Applied Sciences and the Center for Brain Science, Harvard University

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 both in 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.

Keywords

Reinforcement learning; motor control; songbird; motor adaptation

Introduction

Improving performance, whether on the tennis court or in front of 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 and Machens, 2014; Stein et al., 2005). This ‘noise’ results from stochastic events at the level of ion channels (White et al., 2000), synapses (Calvin and Stevens, 1968), and neurons (Mainen and Sejnowski, 1995), and further from instabilities in the dynamics of neural networks (Babloyantz et al., 1985; Vreeswijk and 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 and Wolpert, 1998; Todorov and Jordan, 2002).

[#]Address of corresponding author: 52 Oxford Street, Room 219.30, Cambridge, MA 02138, olveczky@fas.harvard.edu.

However, there is a complementary view of motor variability that suggests that it may be a feature of how sensorimotor circuits operate and learn (Herzfeld and Shadmehr, 2014; Tumer and 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 and Barto, 1998), steer the motor system towards 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 and Wolpert, 1998; Izawa and Shadmehr, 2008; Izawa et al., 2008; Kording and Wolpert, 2004; Newell, 1993; Todorov, 2004; Todorov and Jordan, 2002), less attention has been given to how motor variability could augment the process of 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 to these questions, discuss current views of motor variability, and identify research directions that may further advance our understanding of the role of motor variability in 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 such variability if it can reproduce neural activity patterns that generate successful outcomes. However, little is known about whether and how variability at various levels of motor planning and control 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 and Machens, 2014). At the cellular level, ‘noise’ comes from stochastic biophysical and chemical events that underlie processes such as spike initiation (van Rossum et al., 2003; Schneidman et al., 1998; White et al., 2000) and propagation (Faisal and Laughlin, 2007; Horikawa, 1991), synaptic transmission (Calvin and Stevens, 1968; Katz and 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 and Doiron, 2012; London et al., 2010; Vogels et al., 2005), while pooling across them can enhance correlated signals and reduce the effects of noise (Bruno and Sakmann, 2006; Diesmann et al., 1999).

Variability in the motor system (Figure 2) has perhaps been most extensively characterized at the motor periphery (van Beers et al., 2004; Clamann, 1969; Hamilton et al., 2004; Jones et al., 2002). Studies have shown signal dependent noise in force production, i.e. trial-to-trial fluctuations that scale linearly with mean force, reflect a fundamental property of muscle function (van Beers et al., 2004; Clamann, 1969; Hamilton et al., 2004; Jones et al., 2002). 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 and Wolpert, 1998) that can predict kinematic features of a range of movements, from eye saccades to limb reaches (van Beers, 2007; van Beers et al., 2004; Harris and Wolpert, 1998).

In contrast, variability in central planning circuits may be better suited for learning-related motor exploration (Figure 2). These circuits have more ready access to reinforcement signals (Björklund and Dunnett, 2007; Schultz, 1998), and show more experience-dependent plasticity (Doyon and Benali, 2005; Nudo et al., 1996; Sanes and 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 and 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 and Shenoy (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.

Besides reflecting cellular and network ‘noise’, trial-to-trial variability in motor output can also be deterministic. For example, 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; van Beers, 2009; Scheidt et al., 2001; Smith et al., 2006). 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 reinforcing states that lead to favorable outcomes forms the basis of reinforcement learning (Kaelbling et al., 1996; Sutton and Barto, 1998) - a computational framework that has illuminated a variety of different learning and decision making processes (Lee et al., 2012; Niv, 2009) as well as inspired algorithms for machine learning (Alpaydin, 2014; Sutton and 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, 1948, 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 while reinforcement learning has proven 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 those 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 (Bernshteĭn, 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 and 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 virtually any type of machine learning (Bellman, 1957). The success of deep learning networks in solving complex decision and classification problems have, in large part, been due to the use of convolutional network architectures that dramatically reduce the dimensionality of the solution space by enforcing highly symmetric patterns in the weights to be learned (LeCun et al., 2015; Simonyan and Zisserman, 2014).

Another key to the success of deep learning networks has been the use of unsupervised methods to pre-train networks based on the statistics of the input data (Hinton et al., 2006; LeCun et al., 2015; Lee et al., 2009). This pre-training can serve to get the network into a fertile part of solution space before the exploratory phase of training starts. For reinforcement learning problems, a similar narrowing of solution space can be achieved by

using imitation as a preamble to reinforcement learning (Kormushev et al., 2010; Price and Boutilier, 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 (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 and 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. While 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 showed that the song circuit modularizes song learning by implementing separate reinforcement learning processes for spectral and temporal aspects of song (Ali et al., 2013).

Songbird studies linking motor variability and learning

Despite a powerful theoretical framework linking variability and reinforcement learning, the degree to which motor variability is conducive for motor learning remains debated (Cohen and 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 and Brainard’s experiments in songbirds (Tumer and Brainard, 2007). Though 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, was delivered following certain pitch variants, the song gradually and persistently shifted 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 and Brainard, 2007).

Taking it a step further, Andalman and Fee (Andalman and Fee, 2009) examined the neural mechanisms underlying this reinforcement learning process and found that the error-correcting signal that drives learning is contributed by the output of the anterior forebrain pathway (AFP), LMAN, i.e. the same brain region that introduces much of the exploratory motor variability (Figures 3C–E) (see also (Warren et al., 2011)). This suggests a two-stage learning process, where reinforcement learning in the AFP produces an error-correcting pre-motor 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 and Fee, 2009; Tesileanu et al., 2016; Warren et al., 2011).

Though 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 (Charlesworth et al., 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 adaptively update its output. This is important because it 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 showed 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 and Brainard, 2007) (Figure 3). Picking up this thread, Wu and colleagues set out to test whether this may generalize to human motor learning (Wu et al., 2014). They argued that if motor variability is conducive for 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, with higher task-relevant variability predicting faster learning rates both across different tasks and across individuals within a single task (Figure 4). These results demonstrate 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 generalize also 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' (van Beers et al., 2013; Haith and Krakauer, 2013; Krakauer and Mazzoni, 2011). Surprisingly, the results were similar to the reinforcement learning paradigm, with higher task-relevant variability predicting faster learning rates both across subjects and tasks. This suggests a broader role for variability in motor learning, including also in error-based paradigms. 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). Though 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 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 (van Beers et al., 2013; Scholz and Schönner, 1999). In the absence of feedback, such ‘errors’ could accumulate, leading to slow drift in motor output (van Beers, 2009; 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 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; van Beers et al., 2013). This means that estimates of motor variability made with task-relevant feedback, as in He et al., may emphasize execution noise over planning noise, while estimates made without feedback, as in Wu et al, may primarily reflect planning noise because it allows drift in central planning circuits to contribute more to total motor variability.

If differential access to task-relevant feedback 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 for 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 and 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 an intuition for 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 (Box 1), song variability in juvenile birds is, in large part, generated by the AFP, through LMAN's projection to RA neurons (Figures 3C–E). This projection dominates and drives the RA motor program, and consequently the song, early in learning (Aronov et al., 2008; Ölveczky et al., 2011). Since the activity patterns of LMAN neurons vary across renditions (Hessler and Doupe, 1999; Kao et al., 2008; Ölveczky et al., 2005), this results in variable song in juveniles (Figures 3C–E).

Box 1

What songbirds tell us about motor variability and learning (with Figure 3)

The question of whether and how variability can be harnessed for learning has been most thoroughly examined 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) (Figures 3A,B). Vocal control circuits in songbirds are organized into two main pathways: the descending motor pathway, comprising nuclei HVC and a downstream motor cortex analogue RA (Simpson and Vicario, 1990; Yu and Margoliash, 1996), and the anterior forebrain pathway (AFP), a song-specialized basal ganglia-thalamo-cortical circuit that indirectly connects HVC and RA (Perkel, 2004) (Figure 2A). The song, learned early in life by imitating a tutor, is encoded and generated by the motor pathway (Fee and Scharff, 2010). The AFP is important for song learning, but not for producing the song in the adult bird (Bottjer et al., 1984; Scharff and Nottebohm, 1991). Its output nucleus, 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,D), 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 3E). These results suggest that variability is not simply due to intrinsic noise in the descending motor pathway, but introduced into RA by a dedicated, parallel circuit that is required for song learning.

As learning proceeds, control of the RA motor program gradually shifts 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 et al 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 (Garst-Orozco et al., 2015). 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

selectively reduces variability in action elements that have been mastered, while allowing continued exploration in others (Ravbar et al., 2012). Intriguingly, learning-related synaptic changes akin to those described in songbirds, have been observed also 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 and Barto, 1998). There is evidence that the nervous system regulates variability in such a context-dependent manner, producing more reproducible output in high-reward situations. Experiments in a wide range of species, including rodents (Gharib et al., 2001, 2004), pigeons (Stahlman and 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 also 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 and Doupe, 2011). The circuit mechanisms for this social context-dependent regulation of variability involves 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 and 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. (Pekny et al., 2015) trained human subjects in a reinforcement learning task to reach towards 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 (van Beers et al., 2013; Chaisanguanthum et al., 2014) that make it difficult to establish causal relationships between reward history and motor variability. In other words, it could have been an increase in motor variability that led to decreased reward rates, rather

than vice versa. Overcoming these confounds requires controlling for long-term performance history, which is easier done with datasets containing large number of trials. Although collecting such large datasets 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 underway (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 and 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 further be determined and shaped by the nature and reliability of task-related sensory feedback (Osborne et al., 2005). For instance, Izawa and Shadmehr (Izawa and Shadmehr, 2011) found that trial-to-trial motor variability increased substantially when humans were learning from binary reward prediction errors as compared to more informative sensory (visual) prediction errors. 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 (Box 2).

Box 2

Internal estimates of sensorimotor noise can influence motor variability

While 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 and 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 variability in the sensorimotor realm is introduced because a movement is physically perturbed or visual information about it is distorted, the gains of corrective feedback responses are reduced to match the high uncertainty, thereby maximizing the precision of the generated action (Franklin et al., 2012; Kording and Wolpert, 2004). The rate of trial-to-trial motor adaptation has also

been shown to decrease when variability is experimentally added (Wei and Koording, 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).

Regulating the structure of motor variability

Given the high degree of redundancy in how the motor system controls movements and in how tasks can be executed (Bernshteĭn, 1967), certain forms of motor variability may have small effect on performance, while others may be more consequential. But to what extent does the nervous system distinguish task-relevant and task-irrelevant variability? A number of studies have shown that task-relevant variability is systematically reduced after repeated practice, while task-irrelevant variability can remain high (van Beers et al., 2013; Kang et al., 2004; Latash and Anson, 2006; Scholz and Schöner, 1999). These results are consistent with optimal feedback control theory, which posits that movements are specifically planned and shaped to optimize fidelity in performance (Harris and Wolpert, 1998) and reduce effort (Braun et al., 2009; Izawa and Shadmehr, 2008; Izawa et al., 2008; Todorov, 2004; Todorov and Jordan, 2002). It has even been suggested that larger amounts of task-irrelevant variability can afford reduced task-relevant variability (Todorov, 2004; Todorov and Jordan, 2002), though evidence for this is not as clear.

Shaping the structure of motor variability with such specificity allows the motor system to exploit along dimensions that are deemed task-relevant, while enabling continued exploration (and learning) in others. In agreement with this principle, Wu and colleagues showed that the nervous system can reshape the structure of motor variability as a function of learning to increase the task-relevant component (Wu et al., 2014). Interestingly, this reshaping persisted even after training terminated and the learning-related gains in performance had washed out, suggesting a lasting and experience-dependent modification of the structure of motor variability that could promote more efficient exploration.

Interestingly, neural correlates of such task-specific regulation of motor variability have been observed in the activity of cortical neurons during visuomotor adaptation, where learning related increases in trial-to-trial spiking variability are principally seen 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

Though noise in nervous system function can often be detrimental for optimal performance, the studies we have reviewed here suggest that neural variability may also be conducive for motor learning, in-line with reinforcement learning theory (Kaelbling et al., 1996; Sutton

and 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 is actively regulating and shaping motor variability to augment learning.

While the link between variability and motor learning has been established, the specifics of this relationship remain to be worked out (see future questions). Importantly, furnishing our understanding with mechanistic insight will require animal models with suitable experimental paradigms. Songbirds have proven powerful in this regard (Box 1), 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).

Future questions

1. Which form(s) of motor variability, both in terms of 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?

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 and Perona, 2014; Egnor and 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.

Acknowledgments

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

Bibliography

- Ali F, Otchy TM, Pehlevan C, Fantana AL, Burak Y, Olveczky BP. The Basal Ganglia is necessary for learning spectral, but not temporal, features of birdsong. *Neuron*. 2013; 80:494–506. [PubMed: 24075977]
- Alpaydin E. *Introduction to Machine Learning*. MIT Press; 2014.
- Andalman AS, Fee MS. A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. *Proceedings of the National Academy of Sciences*. 2009; 106:12518–12523.
- Anderson DJ, Perona P. Toward a Science of Computational Ethology. *Neuron*. 2014; 84:18–31. [PubMed: 25277452]
- Aronov D, Andalman AS, Fee MS. A Specialized Forebrain Circuit for Vocal Babbling in the Juvenile Songbird. *Science*. 2008; 320:630–634. [PubMed: 18451295]
- Babloyantz A, Salazar JM, Nicolis C. Evidence of chaotic dynamics of brain activity during the sleep cycle. *Physics Letters A*. 1985; 111:152–156.
- Baddeley RJ, Ingram HA, Miall RC. System Identification Applied to a Visuomotor Task: Near-Optimal Human Performance in a Noisy Changing Task. *J Neurosci*. 2003; 23:3066–3075. [PubMed: 12684493]
- van Beers RJ. The Sources of Variability in Saccadic Eye Movements. *J Neurosci*. 2007; 27:8757–8770. [PubMed: 17699658]
- van Beers RJ. Motor Learning Is Optimally Tuned to the Properties of Motor Noise. *Neuron*. 2009; 63:406–417. [PubMed: 19679079]
- van Beers RJ, Haggard P, Wolpert DM. The Role of Execution Noise in Movement Variability. *Journal of Neurophysiology*. 2004; 91:1050–1063. [PubMed: 14561687]
- van Beers RJ, Brenner E, Smeets JBJ. Random walk of motor planning in task-irrelevant dimensions. *Journal of Neurophysiology*. 2013; 109:969–977. [PubMed: 23175799]
- Bellman R. *Dynamic Programming*. Mineola, N.Y: Dover Publications; 1957.
- Bernshteĭn NA. *The co-ordination and regulation of movements*. Oxford: New York: Pergamon Press; 1967.
- Björklund A, Dunnett SB. Dopamine neuron systems in the brain: an update. *Trends in Neurosciences*. 2007; 30:194–202. [PubMed: 17408759]
- Bottjer SW, Miesner EA, Arnold AP. Forebrain lesions disrupt development but not maintenance of song in passerine birds. *Science*. 1984; 224:901–903. [PubMed: 6719123]
- Botvinick MM. Hierarchical reinforcement learning and decision making. *Current Opinion in Neurobiology*. 2012; 22:956–962. [PubMed: 22695048]
- Braun DA, Aertsen A, Wolpert DM, Mehring C. Learning Optimal Adaptation Strategies in Unpredictable Motor Tasks. *J Neurosci*. 2009; 29:6472–6478. [PubMed: 19458218]
- Bruno RM, Sakmann B. Cortex Is Driven by Weak but Synchronously Active Thalamocortical Synapses. *Science*. 2006; 312:1622–1627. [PubMed: 16778049]
- Calvin WH, Stevens CF. Synaptic noise and other sources of randomness in motoneuron interspike intervals. *Journal of Neurophysiology*. 1968; 31:574–587. [PubMed: 5709873]
- Chaisanguanthum KS, Shen HH, Sabes PN. Motor Variability Arises from a Slow Random Walk in Neural State. *J Neurosci*. 2014; 34:12071–12080. [PubMed: 25186752]
- Charlesworth JD, Warren TL, Brainard MS. Covert skill learning in a cortical-basal ganglia circuit. *Nature*. 2012; 486:251–255. [PubMed: 22699618]
- Churchland MM, Afshar A, Shenoy KV. A Central Source of Movement Variability. *Neuron*. 2006; 52:1085–1096. [PubMed: 17178410]
- Clamann HP. Statistical Analysis of Motor Unit Firing Patterns in a Human Skeletal Muscle. *Biophys J*. 1969; 9:1233–1251. [PubMed: 5824412]
- Cohen RG, Sternad D. Variability in motor learning: relocating, channeling and reducing noise. *Exp Brain Res*. 2008; 193:69–83. [PubMed: 18953531]
- Cohen JD, McClure SM, Yu AJ. Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*. 2007; 362:933–942. [PubMed: 17395573]

- Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ. Cortical substrates for exploratory decisions in humans. *Nature*. 2006; 441:876–879. [PubMed: 16778890]
- Diesmann M, Gewaltig MO, Aertsen A. Stable propagation of synchronous spiking in cortical neural networks. *Nature*. 1999; 402:529–533. [PubMed: 10591212]
- Doyon J, Benali H. Reorganization and plasticity in the adult brain during learning of motor skills. *Current Opinion in Neurobiology*. 2005; 15:161–167. [PubMed: 15831397]
- Egnor SER, Branson K. Computational Analysis of Behavior. *Annual Review of Neuroscience*. 2016; 39:217–236.
- Eshel N, Bukwich M, Rao V, Hemmelder V, Tian J, Uchida N. Arithmetic and local circuitry underlying dopamine prediction errors. *Nature*. 2015; 525:243–246. [PubMed: 26322583]
- Faisal AA, Laughlin SB. Stochastic Simulations on the Reliability of Action Potential Propagation in Thin Axons. *PLoS Comput Biol*. 2007; 3
- Faisal AA, Selen LPJ, Wolpert DM. Noise in the nervous system. *Nat Rev Neurosci*. 2008; 9:292–303. [PubMed: 18319728]
- Fee MS, Scharff C. The songbird as a model for the generation and learning of complex sequential behaviors. *ILAR J*. 2010; 51:362–377. [PubMed: 21131713]
- Fitts PM. The information capacity of the human motor system in controlling the amplitude of movement. *J Exp Psychol*. 1954; 47:381–391. [PubMed: 13174710]
- Flament D, Hore J. Relations of motor cortex neural discharge to kinematics of passive and active elbow movements in the monkey. *Journal of Neurophysiology*. 1988; 60:1268–1284. [PubMed: 3193157]
- Franklin S, Wolpert DM, Franklin DW. Visuomotor feedback gains upregulate during the learning of novel dynamics. *Journal of Neurophysiology*. 2012; 108:467–478. [PubMed: 22539828]
- Fu M, Yu X, Lu J, Zuo Y. Repetitive motor learning induces coordinated formation of clustered dendritic spines in vivo. *Nature*. 2012; 483:92–95. [PubMed: 22343892]
- Garst-Orozco J, Babadi B, Ölveczky BP. A neural circuit mechanism for regulating vocal variability during song learning in zebra finches. *eLife Sciences*. 2015; 3:e03697.
- Gershman SJ, Pesaran B, Daw ND. Human Reinforcement Learning Subdivides Structured Action Spaces by Learning Effector-Specific Values. *J Neurosci*. 2009; 29:13524–13531. [PubMed: 19864565]
- Gharib A, Derby S, Roberts S. Timing and the control of variation. *J Exp Psychol Anim Behav Process*. 2001; 27:165–178. [PubMed: 11296491]
- Gharib A, Gade C, Roberts S. Control of Variation by Reward Probability. *Journal of Experimental Psychology: Animal Behavior Processes*. 2004; 30:271–282. [PubMed: 15506853]
- Gonzalez Castro LN, Hadjiosif AM, Hemphill MA, Smith MA. Environmental Consistency Determines the Rate of Motor Adaptation. *Current Biology*. 2014; 24:1050–1061. [PubMed: 24794296]
- Hadjiosif AM, Smith MA. Flexible Control of Safety Margins for Action Based on Environmental Variability. *J Neurosci*. 2015; 35:9106–9121. [PubMed: 26085634]
- Hahnloser RH, Kozhevnikov AA, Fee MS. An ultra-sparse code underlies the generation of neural sequences in a songbird. *Nature*. 2002; 419:65–70. [PubMed: 12214232]
- Haith AM, Krakauer JW. Model-Based and Model-Free Mechanisms of Human Motor Learning. In: Richardson MJ, Riley MA, Shockley K, editors *Progress in Motor Control*. Springer; New York: 2013. 1–21.
- Haith AM, Reppert TR, Shadmehr R. Evidence for Hyperbolic Temporal Discounting of Reward in Control of Movements. *J Neurosci*. 2012; 32:11727–11736. [PubMed: 22915115]
- Hamid AA, Pettibone JR, Mabrouk OS, Hetrick VL, Schmidt R, Vander Weele CM, Kennedy RT, Aragona BJ, Berke JD. Mesolimbic dopamine signals the value of work. *Nat Neurosci*. 2016; 19:117–126. [PubMed: 26595651]
- de Hamilton AFC, Jones KE, Wolpert DM. The scaling of motor noise with muscle strength and motor unit number in humans. *Exp Brain Res*. 2004; 157:417–430. [PubMed: 15014922]
- Harris CM, Wolpert DM. Signal-dependent noise determines motor planning. *Nature*. 1998; 394:780–784. [PubMed: 9723616]

- He K, Liang Y, Abdollahi F, Bittmann MF, Kording K, Wei K. The Statistical Determinants of the Speed of Motor Learning. *PLOS Comput Biol*. 2016; 12:e1005023. [PubMed: 27606808]
- Herzfeld DJ, Shadmehr R. Motor variability is not noise, but grist for the learning mill. *Nat Neurosci*. 2014; 17:149–150. [PubMed: 24473260]
- Hessler NA, Doupe AJ. Social context modulates singing-related neural activity in the songbird forebrain. *Nature Neuroscience*. 1999; 2:209–211. [PubMed: 10195211]
- Hikosaka O, Kim HF, Yasuda M, Yamamoto S. Basal Ganglia Circuits for Reward Value-Guided Behavior. *Annual Review of Neuroscience*. 2014; 37:289–306.
- Hinton GE, Osindero S, Teh YW. A Fast Learning Algorithm for Deep Belief Nets. *Neural Comput*. 2006; 18:1527–1554. [PubMed: 16764513]
- Horikawa Y. Noise effects on spike propagation in the stochastic Hodgkin-Huxley models. *Biol Cybern*. 1991; 66:19–25. [PubMed: 1768709]
- 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]
- Immelmann K. Song development in the zebra finch and other estrildid finches. In: Hinde RA, editor *Bird Vocalizations*. Cambridge University Press; 1969. 61–74.
- Izawa J, Shadmehr R. On-Line Processing of Uncertain Information in Visuomotor Control. *J Neurosci*. 2008; 28:11360–11368. [PubMed: 18971478]
- Izawa J, Shadmehr R. Learning from Sensory and Reward Prediction Errors during Motor Adaptation. *PLoS Comput Biol*. 2011; 7:e1002012. [PubMed: 21423711]
- Izawa J, Rane T, Donchin O, Shadmehr R. Motor Adaptation as a Process of Reoptimization. *J Neurosci*. 2008; 28:2883–2891. [PubMed: 18337419]
- Jones KE, de Hamilton AFC, Wolpert DM. Sources of Signal-Dependent Noise During Isometric Force Production. *Journal of Neurophysiology*. 2002; 88:1533–1544. [PubMed: 12205173]
- Kaelbling LP, Littman ML, Moore AW. Reinforcement Learning: A Survey. 1996 Eprint arXiv:cs/9605103 arXiv:cs/9605103.
- Takei S, Hoffman DS, Strick PL. Muscle and Movement Representations in the Primary Motor Cortex. *Science*. 1999; 285:2136–2139. [PubMed: 10497133]
- Kang N, Shinohara M, Zatsiorsky VM, Latash ML. Learning multi-finger synergies: an uncontrolled manifold analysis. *Exp Brain Res*. 2004; 157:336–350. [PubMed: 15042264]
- Kao MH, Doupe AJ, Brainard MS. Contributions of an avian basal ganglia-forebrain circuit to real-time modulation of song. *Nature*. 2005; 433:638–643. [PubMed: 15703748]
- Kao MH, Wright BD, Doupe AJ. Neurons in a Forebrain Nucleus Required for Vocal Plasticity Rapidly Switch between Precise Firing and Variable Bursting Depending on Social Context. *J Neurosci*. 2008; 28:13232–13247. [PubMed: 19052215]
- Katz B, Miledi R. Membrane Noise produced by Acetylcholine. *Nature*. 1970; 226:962–963. [PubMed: 5445879]
- Kawagoe R, Takikawa Y, Hikosaka O. Expectation of reward modulates cognitive signals in the basal ganglia. *Nat Neurosci*. 1998; 1:411–416. [PubMed: 10196532]
- Kojima S, Doupe AJ. Social performance reveals unexpected vocal competency in young songbirds. *Proceedings of the National Academy of Sciences*. 2011
- Kording KP, Wolpert DM. Bayesian integration in sensorimotor learning. *Nature*. 2004; 427:244–247. [PubMed: 14724638]
- Kormushev P, Calinon S, Caldwell DG. Robot motor skill coordination with EM-based Reinforcement Learning. 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); 2010. 3232–3237.
- Krakauer JW, Mazzoni P. Human sensorimotor learning: adaptation, skill, and beyond. *Current Opinion in Neurobiology*. 2011; 21:636–644. [PubMed: 21764294]
- Lashley KS. Integrative Functions of the Cerebral Cortex. *Physiological Reviews*. 1933; 13:1–42.
- Latash ML, Anson JG. Synergies in Health and Disease: Relations to Adaptive Changes in Motor Coordination. *Physical Therapy*. 2006; 86:1151–1160. [PubMed: 16879049]

- Lauwereyns J, Watanabe K, Coe B, Hikosaka O. A neural correlate of response bias in monkey caudate nucleus. *Nature*. 2002; 418:413–417. [PubMed: 12140557]
- Leblois A. Social modulation of learned behavior by dopamine in the basal ganglia: Insights from songbirds. *Journal of Physiology-Paris*. 2013; 107:219–229.
- LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015; 521:436–444. [PubMed: 26017442]
- Lee D, Seo H, Jung MW. Neural Basis of Reinforcement Learning and Decision Making. *Annual Review of Neuroscience*. 2012; 35:287–308.
- Lee H, Pham P, Largman Y, Ng AY. Unsupervised feature learning for audio classification using convolutional deep belief networks. In: Bengio Y, Schuurmans D, Lafferty JD, Williams CKI, Culotta A, editors *Advances in Neural Information Processing Systems 22*. Curran Associates, Inc; 2009. 1096–1104.
- Lemon RN. Descending Pathways in Motor Control. *Annual Review of Neuroscience*. 2008; 31:195–218.
- Litwin-Kumar A, Doiron B. Slow dynamics and high variability in balanced cortical networks with clustered connections. *Nat Neurosci*. 2012; 15:1498–1505. [PubMed: 23001062]
- London M, Roth A, Beeren L, Hausser M, Latham PE. Sensitivity to perturbations in vivo implies high noise and suggests rate coding in cortex. *Nature*. 2010; 466:123–127. [PubMed: 20596024]
- Luo L, Callaway E, Svoboda K. Genetic dissection of neural circuits. *Neuron*. 2008; 57:634–660. [PubMed: 18341986]
- Mainen ZF, Sejnowski TJ. Reliability of spike timing in neocortical neurons. *Science*. 1995; 268:1503. [PubMed: 7770778]
- Mandelblat-Cerf Y, Paz R, Vaadia E. Trial-to-Trial Variability of Single Cells in Motor Cortices Is Dynamically Modified during Visuomotor Adaptation. *J Neurosci*. 2009; 29:15053–15062. [PubMed: 19955356]
- Marcos E, Pani P, Brunamonti E, Deco G, Ferraina S, Verschure P. Neural Variability in Premotor Cortex Is Modulated by Trial History and Predicts Behavioral Performance. *Neuron*. 2013; 78:249–255. [PubMed: 23622062]
- Matsumoto K, Suzuki W, Tanaka K. Neuronal Correlates of Goal-Based Motor Selection in the Prefrontal Cortex. *Science*. 2003; 301:229–232. [PubMed: 12855813]
- Miyamoto YR, Dhawale AK, Smith MA, Ölveczky BP. Investigating reward-based regulation of task-relevant motor variability in rats. *Neuroscience Meeting Planner*. 2015
- Newell KM. *Variability and Motor Control*. Champaign IL: Human Kinetics Pub; 1993.
- Niv Y. Reinforcement learning in the brain. *Journal of Mathematical Psychology*. 2009; 53:139–154.
- Nudo RJ, Milliken GW, Jenkins WM, Merzenich MM. Use-dependent alterations of movement representations in primary motor cortex of adult squirrel monkeys. *J Neurosci*. 1996; 16:785–807. [PubMed: 8551360]
- Ölveczky BP. Motoring ahead with rodents. *Curr Opin Neurobiol*. 2011; 21:571–578. [PubMed: 21628098]
- Ölveczky BP, Andalman AS, Fee MS. Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. *PLoS Biol*. 2005; 3:e153. [PubMed: 15826219]
- Ölveczky BP, Otchy TM, Goldberg JH, Aronov D, Fee MS. Changes in the neural control of a complex motor sequence during learning. *J Neurophysiol*. 2011; 106:386–397. [PubMed: 21543758]
- Opris I, Lebedev M, Nelson RJ. Motor Planning under Unpredictable Reward: Modulations of Movement Vigor and Primate Striatum Activity. *Front Neurosci*. 2011; 5
- Osborne LC, Lisberger SG, Bialek W. A sensory source for motor variation. *Nature*. 2005; 437:412–416. [PubMed: 16163357]
- Parr R. Reinforcement learning with hierarchies of machines. *Advances in Neural Information Processing Systems*. 1998; 10:1043–1049.
- Pekny SE, Izawa J, Shadmehr R. Reward-Dependent Modulation of Movement Variability. *J Neurosci*. 2015; 35:4015–4024. [PubMed: 25740529]
- Perkel DJ. Origin of the Anterior Forebrain Pathway. *Annals of the New York Academy of Sciences*. 2004; 1016:736–748. [PubMed: 15313803]

- Peters J, Schaal S. Reinforcement learning of motor skills with policy gradients. *Neural Networks*. 2008; 21:682–697. [PubMed: 18482830]
- Poddar R, Kawai R, Ölveczky BP. A Fully Automated High-Throughput Training System for Rodents. *PLoS ONE*. 2013; 8:e83171. [PubMed: 24349451]
- Price B, Boutilier C. Accelerating Reinforcement Learning Through Implicit Imitation. *J Artif Int Res*. 2003; 19:569–629.
- Ravbar P, Lipkind D, Parra LC, Tchernichovski O. Vocal Exploration Is Locally Regulated during Song Learning. *J Neurosci*. 2012; 32:3422–3432. [PubMed: 22399765]
- Renart A, Machens CK. Variability in neural activity and behavior. *Current Opinion in Neurobiology*. 2014; 25:211–220. [PubMed: 24632334]
- Roesch MR, Olson CR. Neuronal Activity Related to Reward Value and Motivation in Primate Frontal Cortex. *Science*. 2004; 304:307–310. [PubMed: 15073380]
- van Rossum MCW, O'Brien BJ, Smith RG. Effects of Noise on the Spike Timing Precision of Retinal Ganglion Cells. *Journal of Neurophysiology*. 2003; 89:2406–2419. [PubMed: 12740401]
- Samejima K, Ueda Y, Doya K, Kimura M. Representation of Action-Specific Reward Values in the Striatum. *Science*. 2005; 310:1337–1340. [PubMed: 16311337]
- Sanes JN, Donoghue JP. Plasticity and primary motor cortex. *Annual Review of Neuroscience*. 2000; 23:393–415.
- Scharff C, Nottebohm F. A comparative study of the behavioral deficits following lesions of various parts of the zebra finch song system: implications for vocal learning. *The Journal of Neuroscience*. 1991; 11:2896–2913. [PubMed: 1880555]
- Scheidt RA, Dingwell JB, Mussa-Ivaldi FA. Learning to Move Amid Uncertainty. *Journal of Neurophysiology*. 2001; 86:971–985. [PubMed: 11495965]
- Schneidman E, Freedman B, Segev I. Ion channel stochasticity may be critical in determining the reliability and precision of spike timing. *Neural Comput*. 1998; 10:1679–1703. [PubMed: 9744892]
- Scholz JP, Schöner G. The uncontrolled manifold concept: identifying control variables for a functional task. *Exp Brain Res*. 1999; 126:289–306. [PubMed: 10382616]
- Schultz W. Predictive Reward Signal of Dopamine Neurons. *J Neurophysiol*. 1998; 80:1–27. [PubMed: 9658025]
- Schultz W, Dayan P, Montague PR. A Neural Substrate of Prediction and Reward. *Science*. 1997; 275:1593–1599. [PubMed: 9054347]
- Schultz W, Tremblay L, Hollerman JR. Changes in behavior-related neuronal activity in the striatum during learning. *Trends in Neurosciences*. 2003; 26:321–328. [PubMed: 12798602]
- Shadmehr R, Huang HJ, Ahmed AA. A Representation of Effort in Decision-Making and Motor Control. *Current Biology*. 2016; 26:1929–1934. [PubMed: 27374338]
- Silver D, Huang A, Maddison CJ, Guez A, Sifre L, van den Driessche G, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M, et al. Mastering the game of Go with deep neural networks and tree search. *Nature*. 2016; 529:484–489. [PubMed: 26819042]
- Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. 2014 arXiv:1409.1556 [Cs].
- Simpson HB, Vicario DS. Brain pathways for learned and unlearned vocalizations differ in zebra finches. *J Neurosci*. 1990; 10:1541–1556. [PubMed: 2332796]
- Skinner BF. *The Behavior of Organisms*. Copley Publishing Group; 1948.
- Skinner BF. Operant behavior. *American Psychologist*. 1963; 18:503–515.
- Skinner BF. Selection by Consequences. *Science*. 1981; 213:501–504. [PubMed: 7244649]
- Smith MA, Ghazizadeh A, Shadmehr R. Interacting Adaptive Processes with Different Timescales Underlie Short-Term Motor Learning. *PLOS Biol*. 2006; 4:e179. [PubMed: 16700627]
- Stahlman WD, Blaisdell AP. The modulation of operant variation by the probability, magnitude, and delay of reinforcement. *Learning and Motivation*. 2011; 42:221–236. [PubMed: 21731111]
- Stahlman WD, Roberts S, Blaisdell AP. Effect of reward probability on spatial and temporal variation. *J Exp Psychol Anim Behav Process*. 2010; 36:77–91. [PubMed: 20141319]

- Stein RB, Gossen ER, Jones KE. Neuronal variability: noise or part of the signal? *Nat Rev Neurosci*. 2005; 6:389–397. [PubMed: 15861181]
- Sutton RS, Barto AG. *Reinforcement Learning: An Introduction*. The MIT Press; 1998.
- Tai LH, Lee AM, Benavidez N, Bonci A, Wilbrecht L. Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value. *Nature Neuroscience*. 2012; 15:1281–1289. [PubMed: 22902719]
- Takikawa Y, Kawagoe R, Itoh H, Nakahara H, Hikosaka O. Modulation of saccadic eye movements by predicted reward outcome. *Exp Brain Res*. 2002; 142:284–291. [PubMed: 11807582]
- Tchernichovski O, Mitra PP, Lints T, Nottebohm F. Dynamics of the vocal imitation process: how a zebra finch learns its song. *Science*. 2001; 291:2564. [PubMed: 11283361]
- Tesileanu T, Ölveczky B, Balasubramanian V. Matching tutor to student: rules and mechanisms for efficient two-stage learning in neural circuits. 2016 bioRxiv 71910.
- Thorndike EL, Edward L. *Animal intelligence : an experimental study of the associative processes in animals*. New York: Macmillan; 1898.
- Todorov E. Optimality principles in sensorimotor control. *Nat Neurosci*. 2004; 7:907–915. [PubMed: 15332089]
- Todorov E, Jordan MI. Optimal feedback control as a theory of motor coordination. *Nat Neurosci*. 2002; 5:1226–1235. [PubMed: 12404008]
- Tumer EC, Brainard MS. Performance variability enables adaptive plasticity of “crystallized” adult birdsong. *Nature*. 2007; 450:1240–1244. [PubMed: 18097411]
- Vogels TP, Rajan K, Abbott LF. *Neural Network Dynamics*. Annual Review of Neuroscience. 2005; 28:357–376.
- van Vreeswijk C, Sompolinsky H. Chaos in Neuronal Networks with Balanced Excitatory and Inhibitory Activity. *Science*. 1996; 274:1724–1726. [PubMed: 8939866]
- Wang AY, Miura K, Uchida N. The dorsomedial striatum encodes net expected return, critical for energizing performance vigor. *Nat Neurosci*. 2013; 16:639–647. [PubMed: 23584742]
- Wang L, Conner JM, Rickert J, Tuszynski MH. Structural plasticity within highly specific neuronal populations identifies a unique parcellation of motor learning in the adult brain. *Proceedings of the National Academy of Sciences*. 2011
- Warren TL, Tumer EC, Charlesworth JD, Brainard MS. Mechanisms and time course of vocal learning and consolidation in the adult songbird. *J Neurophysiol*. 2011; 106:1806–1821. [PubMed: 21734110]
- Wei K, Koerding K. Uncertainty of feedback and state estimation determines the speed of motor adaptation. *Front Comput Neurosci*. 2010; 4:11. [PubMed: 20485466]
- White JA, Rubinstein JT, Kay AR, White JA, Rubinstein JT, Kay AR, White JA, Rubinstein JT, Kay AR, White JA, et al. Channel noise in neurons. *Trends in Neurosciences*. 2000; 23:131–137. [PubMed: 10675918]
- Woolley SC, Rajan R, Joshua M, Doupe AJ. Emergence of Context-Dependent Variability across a Basal Ganglia Network. *Neuron*. 2014; 82:208–223. [PubMed: 24698276]
- Wu HG, Miyamoto YR, Castro LNG, Ölveczky BP, Smith MA. Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. *Nature Neuroscience*. 2014; 17:312–321. [PubMed: 24413700]
- Wunderlich K, Rangel A, O’Doherty JP. Neural computations underlying action-based decision making in the human brain. *PNAS*. 2009; 106:17199–17204. [PubMed: 19805082]
- Xu T, Yu X, Perlik AJ, Tobin WF, Zweig JA, Tennant K, Jones T, Zuo Y. Rapid formation and selective stabilization of synapses for enduring motor memories. *Nature*. 2009; 462:915–919. [PubMed: 19946267]
- Yu AC, Margoliash D. Temporal Hierarchical Control of Singing in Birds. *Science*. 1996; 273:1871–1875. [PubMed: 8791594]

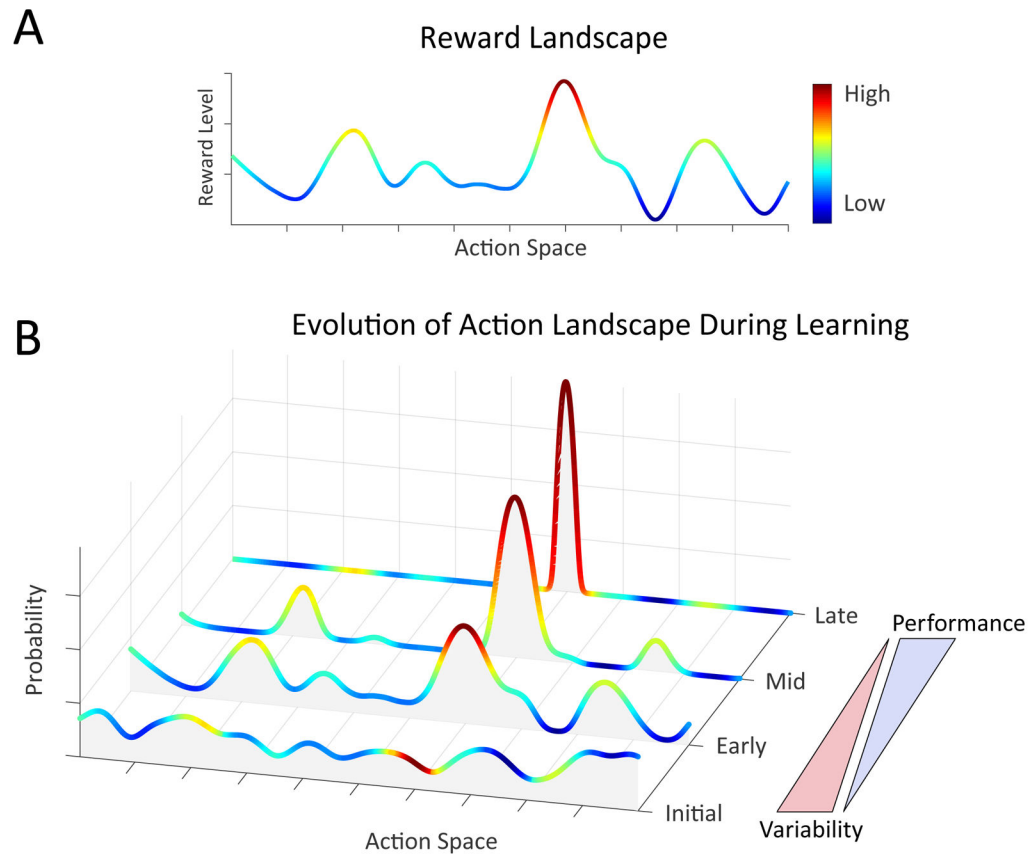


Figure 1. Illustrating how variability can be conducive for motor learning

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. Initially variability is 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.

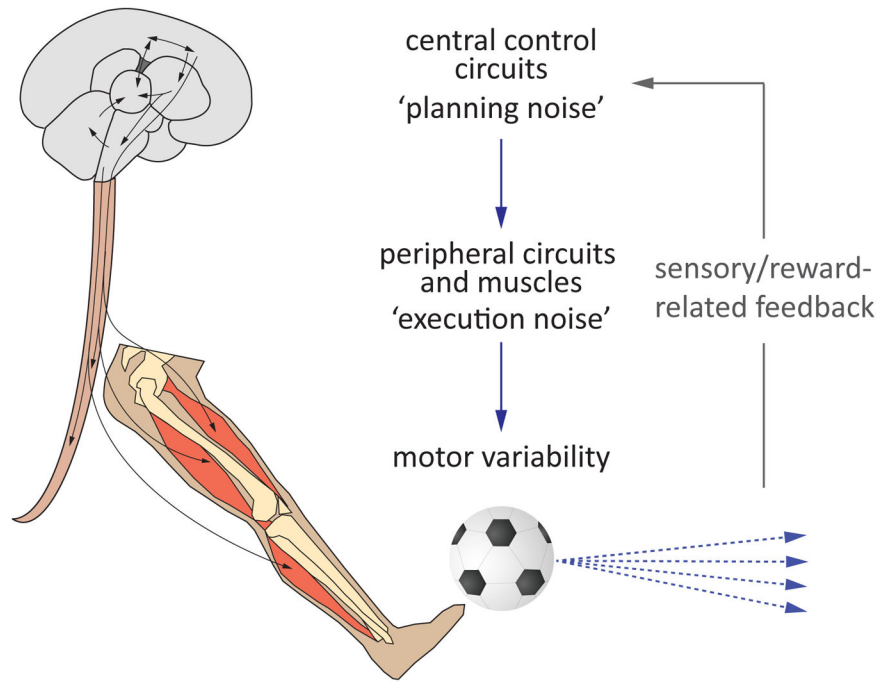


Figure 2. Sources of motor variability

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 for 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 (see text).

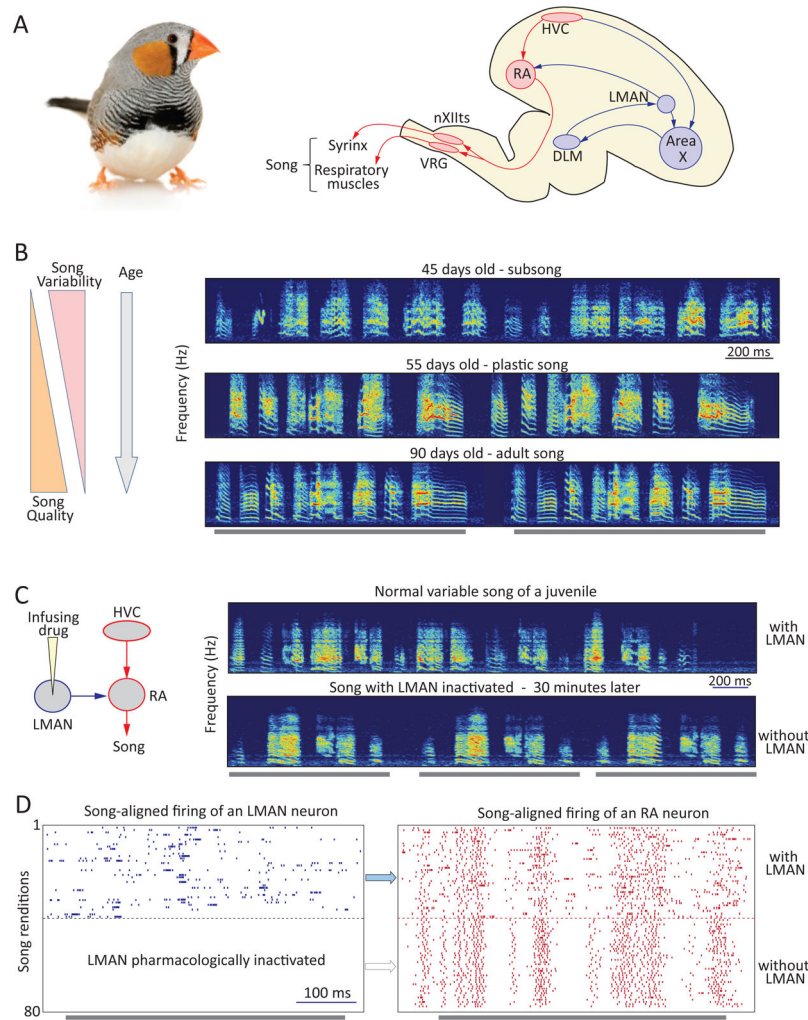


Figure 3. Vocal variability in songbirds is generated by a basal ganglia-like circuit
 Research on the courtship song of zebra finches has informed the link between motor variability and learning. **A.** The song is generated by the vocal control pathway comprising HVC, RA, and brainstem motor regions (red pathway). The Anterior Forebrain Pathway (blue pathway), 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, shows 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, shows a dramatic reduction in rendition-to-rendition variability in the RA neurons with LMAN silencing. Data from (Ölveczky et al., 2011).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

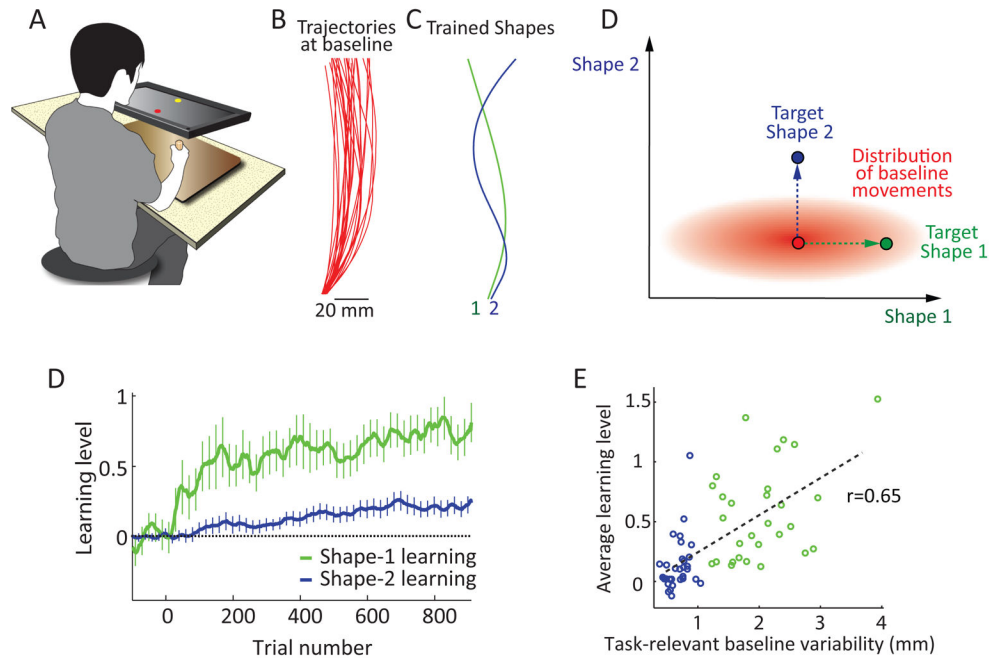


Figure 4.

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, 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 (markers). Adapted from (Wu et al., 2014).