Temporal structure of motor variability is dynamically regulated and predicts motor learning ability

Howard G Wu[1,4], **Yohsuke R Miyamoto**[1,4], **Luis Nicolas Gonzalez Castro**[1], **Bence P Ölveczky**[2,3], and **Maurice A Smith**[1,3]

[1]School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts, USA

[2]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts, USA

[3]Center for Brain Science, Harvard University, Cambridge, Massachusetts, USA

Abstract

Individual differences in motor learning ability are widely acknowledged, yet little is known about the factors that underlie them. Here we explore whether movement-to-movement variability in motor output, a ubiquitous if often unwanted characteristic of motor performance, predicts motor learning ability. Surprisingly, we found that higher levels of task-relevant motor variability predicted faster learning both across individuals and across tasks in two different paradigms, one relying on reward-based learning to shape specific arm movement trajectories and the other relying on error-based learning to adapt movements in novel physical environments. We proceeded to show that training can reshape the temporal structure of motor variability, aligning it with the trained task to improve learning. These results provide experimental support for the importance of action exploration, a key idea from reinforcement learning theory, showing that motor variability facilitates motor learning in humans and that our nervous systems actively regulate it to improve learning.

In 2009, Brendon Todd became the first golfer to hit two consecutive holes in one on the same hole during a professional tournament. Even if the vagaries of wind conditions, humidity and pin placement were controlled, precisely repeating the action that led to the hole in one would still be an amazing feat, as anyone who has swung a golf club can attest. But why should it be difficult to repeat a hole in one or any other action? The answer is that there is an ever-present variability in motor execution that makes it virtually impossible to exactly repeat actions.

On one hand, motor variability is widely thought of as the inevitable consequence of stochastic nervous system function, arising from noise in sensory or motor processing, or sensorimotor integration[1–5]. Several theories of motor control posit that actions are planned specifically to minimize the extent to which variability affects performance on the task at hand, either alone[6–8] or in combination with minimizing effort[9–11]. These theories generally treat motor variability as inevitable signal-dependent noise, which varies proportionally to the size of the motor output. The idea that noise in motor output is primarily signal dependent, however, is based largely on isometric force generation studies[2–4]. Much less is known about how motor output variability evolves during active movement.

On the other hand, there is evidence that the nervous system specifically regulates, and indeed amplifies, variability instead of minimizing its effects. Recent studies in songbirds have shown that variability in motor performance and motor learning ability are both markedly reduced after inactivating the cortical output nucleus of a basal ganglia circuit, the lateral magnocellular nucleus of the anterior neostriatum (LMAN)[12–14]. These findings suggest that LMAN, which projects directly to a motor cortex analog brain area involved in singing, generates variability in motor output to promote learning[13,15].

Why would variability promote learning? Motor variability can be equated with action exploration, an essential component of reinforcement learning, where the exploration necessary to gather knowledge must be balanced with exploitation of the knowledge that has been accrued[16,17]. Consider the process of learning a golf swing: at first the motion is highly variable, but with practice it becomes increasingly precise as performance improves. This can be interpreted as a progression from exploring different swing motions early on when rapid learning is most beneficial to exploiting the best of these motions later on. Similar learning-related regulation of variability has been observed in other animal models and contexts[18,19].

Thus, it remains unclear whether high initial variability stands in the way of effective performance or whether it facilitates the motor system's ability to learn. To test whether movement variability indeed promotes motor learning in humans, we measured baseline motor variability before participants engaged in different types of motor learning tasks and studied whether the structure of this variability could predict the rate at which these individuals learned. We then investigated whether the motor system could leverage the relationship between variability and learning ability by examining whether it actively reshapes the structure of motor output variability to guide learning.

RESULTS

Inter-individual differences in reward-based learning

Motivated by reinforcement learning theory, which emphasizes action exploration as a key ingredient for learning, we examined the relationship between variability and learning rate in a reward-based motor learning task. In this task, we trained subjects ($n = 20$) to produce hand trajectories with specific shapes using trial-and-error learning, where reward was based on performance (**Fig. 1a–e**). We gave the subjects instructions to repeatedly trace a subtly curved guide shape shown on a monitor with rapid 20-cm point-to-point reaching arm

movements (**Fig. 1a,b**). Participants received no visual feedback of their actual hand paths, and in the training period, they received scores based on an unrelated but also subtly curved shape. Each experiment began with a 250-trial baseline period during which we estimated motor variability (**Fig. 1b**). In this baseline period, we did not reward movements based on a baseline shape. Instead, we withheld the scores and gave subjects (binary) feedback about their movement speed to enforce rapid movements. By the end of the training period, 95% of movements were acceptably fast (peak speed >0.45 m s^{-1}).

During the training period, we instructed participants to maximize a numerical performance score between 0 and 1000 that we displayed after each trial. This score was based on the similarity between the actual path and a rewarded shape that was, similarly to the guide shape, subtly curved but was otherwise independent from the guide shape (examples are shown in **Fig. 1c**, but note the exaggerated lateral scaling). The rewarded shapes were individualized in that they corresponded to a predetermined deflection from each subject's mean baseline trajectory. The performance score was a function of the normalized dot product between the vectors of x-axis deflections for the actual and rewarded hand paths (examples are shown in **Fig. 1d,e**, the Online Methods and **Supplementary Fig. 1**). Notably, participants received no visual feedback about their trajectories or the rewarded shape, ensuring that no error-correcting information was available to guide learning. Moreover, each group was split randomly into two subgroups that learned deflections that were the same in shape but opposite in direction (the thin versus thick lines in **Fig. 1c**).

If the variability produced by the motor system is indeed harnessed during trial-and-error learning to improve performance, then greater task-relevant variability before training should predict higher learning rates during training. This is because greater variability in the task-relevant dimension of motor output immediately before the onset of training would lead to greater task-relevant variability during early training, resulting in greater exploration during training, which should promote faster learning. Perhaps the most direct test of the hypothesis that motor variability promotes faster learning is to examine the relationship between task-relevant variability and learning rate in the same initial exposure period. But in the early training period, directed learning and random motor variability occur alongside one another, and thus changes in motor output from rapid directed learning may contaminate measures of motor output variability. When we compared variability and learning ability in the early training period by attempting to independently estimate them, we found a strong positive relationship, providing support for the aforementioned hypothesis (**Supplementary Figs. 2** and **3**). However, to eliminate any possible contamination between learning and variability, we focused our analysis on the relationship between task-relevant variability in a baseline period before training and learning rates observed during early training. Because we did not introduce the learning task in the baseline period and we balanced the polarity of the task across the subjects, as described below, performance biases in the direction of task-related learning could not systematically bias our measures of motor output variability.

We computed the task-relevant component of the baseline variability by projecting each mean-subtracted hand path onto the deflection that would be rewarded during the subsequent shape-specific training, as this projection isolates the aspect of the motor output that affects the score in the training period. We characterized the amount of baseline task-relevant

variability by computing the standard deviation of the size of these projections over the last 160 baseline trials. Analogously, we characterized the rate of initial learning on the basis of the hand paths during the first 125 trials of the training period after subtracting out the mean hand path during the last 160 baseline trials. We computed learning level estimates (**Fig. 1f**) as the mean size of the projections of these baseline-subtracted hand paths onto the rewarded deflection normalized by the size of the rewarded deflection. Thus, baseline performance would yield a value of 0, and ideal performance would yield a value of 1.

In experiment 1, we trained subjects on shape 1 (shown in **Fig. 1c**). To look at the effect of baseline motor variability on subsequent learning, we stratified individuals into subgroups on the basis of the amount of shape 1 variability they displayed in the baseline period (**Fig. 1f**; $P = 0.0168$, $t(5.42) = 2.83$). We found that subjects with higher than average task-relevant variability learned considerably faster than the subjects with low variability (**Fig. 1f**; $P = 0.0091$, $t(9.78) = 2.83$; **Supplementary Fig. 4** shows a median-based stratification), suggesting that higher task-relevant baseline variability is associated with faster learning rates. In line with this idea, we found that task-relevant variability was positively correlated with learning rate across individuals (**Fig. 1h**; $r = +0.75$, $P < 0.0001$, $t(18) = 4.74$). The subject with the highest task-relevant baseline variability epitomized this correlation by displaying the highest learning rate. However, even when we removed this data point, the correlation between task-relevant variability and learning rate remained ($r = +0.61$, $P = 0.0027$, $t(17) = 3.19$). To our knowledge, this is the first time individual differences in learning rate have been predicted from baseline performance characteristics.

Inter-task differences in reward-based learning

The correlation between task-relevant variability and motor learning rate suggests that higher levels of task-relevant variability might lead to faster learning. However, when we examined the relationship between total variability and learning rate, we also found a significant, albeit weaker, correlation (**Fig. 1g**; $r = +0.43$, $P = 0.03$, $t(18) = 1.99$), raising the possibility that it is the total amount of variability rather than the task-relevant component of variability that matters for learning. Here we calculated total variability as the sum of the variance at each point in the hand path. But in our data, total and task-specific baseline variability were themselves correlated ($r = +0.71$). Thus, to determine which of these variables drives inter-subject differences in learning rate, we performed a second experiment in which we trained two groups of subjects on two different rewarded deflections (**Fig. 1c**; $n = 29$ for shape 1, and $n = 32$ for shape 2), each of which was associated with different amounts of task-relevant variability at baseline. Shape 1 was identical to the shape that we used in the first experiment, whereas we chose shape 2 to account for a smaller amount of the total baseline variability and to be orthogonal to shape 1 ($P < 0.0001$, $t(32.72) = 11.76$). Because the subjects learning shape 1 and shape 2 should display similar total variability but different task-relevant variability, comparing their learning rates should dissociate the effects of task-relevant and total variability.

We found that task-relevant variability during baseline explained differences in learning rate both within (**Fig. 1k**; shape 1, $P = 0.012$, $t(29) = 2.39$; shape 2, $P = 0.022$, $t(32) = 2.11$) and across ($P < 0.0001$, $t(61) = 6.70$) groups, with slower learning rates for shape 2 compared to

shape 1 (**Fig. 1i**; $P < 0.0001$, $t(41.92) = 5.20$). In contrast, total variability failed to explain the between-group differences in learning rate (**Fig. 1j**), suggesting that task-relevant rather than total variability drives changes in learning rates. Moreover, simultaneous regression of learning rate onto both types of variability across groups revealed a highly significant effect of task-relevant variability (partial $R^2 = 0.37$, $P < 0.0001$, $F(1,58) = 34.7062$) but no effect of total variability (partial $R^2 = 0.017$, $P = 0.316$, $F(1,58) = 1.02$). Correspondingly, a single relationship between task-relevant variability and learning rate (**Fig. 1k**) explained the individual differences in learning ability for both groups just as accurately ($R^2 = 0.432$) as did a composite relationship individualized for each group ($R^2 = 0.437$), which can be visualized as a model combining the blue (shape 2) and green (shape 1) lines in **Figure 1k** (partial $R^2 = 0.009$, $P = 0.77$, $F(2,57) = 0.25$). This result indicates that task-relevant variability provides a unifying explanation for both inter-individual and inter-task differences in learning rate across the tasks that we studied.

Inter-individual differences in error-based force-field learning

We next examined the generality of the relationship between motor variability and motor learning rates on a task that is thought to be learned through error-based learning. Reward-based tasks that lack error signals are generally learned more slowly than tasks in which dynamic or kinematic perturbations produce error signals. Reinforcement learning theory posits that motor exploration is essential for reward-based learning, and although recent studies have suggested that learning in error-based tasks may arise from multiple mechanisms[20], error-based adaptation need not be contingent on such exploration.

To examine whether motor variability is associated with improved learning rates in error-based motor adaptation tasks, we performed a third experiment in which participants ($n = 40$) adapted point-to-point reaching movements in an environment with altered physical dynamics (**Fig. 2a**). We exposed subjects to a velocity-dependent force field in which the force vector perturbing the hand was proportional in magnitude and lateral in direction to the velocity of the hand[21–23] (**Fig. 2c**). The experimental design incorporated an extended baseline period so that we could accurately measure baseline variability. We used randomly interspersed error-clamp trials[21–23] (**Fig. 2d,e**) to measure the lateral force profiles produced during both baseline (**Fig. 2b**) and training (**Fig. 2c**). We quantified task-relevant variability at baseline and task-specific adaptation during training by projecting these force profiles onto the ideal velocity-dependent force patterns that were associated with each movement (**Supplementary Fig. 5b–e**). When we stratified individuals on the basis of the level of velocity-dependent variability they displayed during the baseline period (**Fig. 2f**), we found that participants with above-average variability showed faster velocity-dependent force-field learning than those with below-average variability (**Fig. 2g**; $P = 0.011$, $t(32.36) = 2.40$ when comparing the average learning rate over the first ten trials; **Supplementary Fig. 6** shows a median-based stratification). Moreover, individuals with variability that was at least one standard deviation above average had more than twofold higher learning rates than individuals with variability levels that were one standard deviation below average (**Fig. 2g**; $P = 0.0021$, $t(6.76) = 4.23$). Correspondingly, we found a significant positive correlation between the amount of velocity-dependent baseline variability and the initial learning rate across subjects (**Fig. 2h**; $r = +0.46$, $P = 0.0013$, $t(38) = 3.21$). It is notable that the

variability-related changes in learning appear to be transient, peaking at about trial eight, in a pattern that resembles the fast component from a two-rate model of force-field adaptation[23–25], raising the possibility that this component may underlie the observed relationship between variability and learning. These findings show that individuals with higher levels of task-relevant variability learn faster in both error-based and reward-based motor learning tasks, suggesting that the relationship between motor variability and learning may reflect a general principle of motor learning.

The temporal structure of motor variability during reaching

If indeed there is a general relationship between the structure of motor variability and learning ability, variability in force production during movement may explain why some types of force dynamics are learned more quickly than others[22,26]. However, little is known about movement-related variability in the production of force, as force output variability has generally been studied under isometric conditions[2,3,11]. To examine whether the temporal structure of motor output variability can indeed explain differences in how individuals adapt to different types of dynamics, we designed an experiment to measure force variability during voluntary point-to-point reaching movements. As in the analysis of the individual differences in error-based learning described above, we used occasional error-clamp trials to measure lateral force directly during these reaching movements.

Even during the null-field baseline period, we found a surprising degree of trial-to-trial variability, although we asked the subjects to repeat the same reaching movement precisely (**Fig. 3a**). We performed principal component analysis on the structure of this variability, which is a data-driven method that can identify the temporal patterns that contribute most to the total variance. Notably, we found that the first principal component (PC_1), i.e., the force pattern that best characterizes the total motor variability, alone accounted for $40 \pm 2\%$ (mean $\pm$ s.e.m.) of the total variance, which is over three times as much as any other component (**Fig. 3b**). Inspired by previous work showing that new dynamics are learned as a function of motion state rather than time[22,27–29], we examined the extent to which each principal component explained motion-related variability specifically (**Fig. 3c**), which we define as the portion of variability that can be explained as a linear combination of the position, velocity and acceleration of the motion. PC_1 accounted for $72 \pm 2\%$ of the variability associated with motion state, which is more than twice as much as all the other components combined. This suggests that PC_1 may account for nearly three-quarters of learning-related variability. In line with this idea, the shape of PC_1 was itself strongly related to motion, as it was well approximated ($R^2 = 0.95$) by a linear combination of the position, velocity and acceleration of the hand (**Fig. 3d**; **Supplementary Fig. 7** shows the shapes of PC_2–PC_5). Notably, the position and velocity contributions that account for the majority of the shape of PC_1 ($R^2 = 0.85$) appear in positive combination, closely resembling the pattern of viscoelastic dynamics that are known to be learned fastest[22].

The temporal structure of baseline variability predicts which types of dynamics are learned faster

The resemblance between the component of the temporal structure that accounts for the greatest fraction of the baseline variability (PC_1) and the type of dynamics that is learned

most rapidly suggests a link between motor output variability and motor learning ability across different dynamic environments, which is analogous to what we found for different reward-based learning tasks (**Fig. 1i–k**). To explore the nature of this connection, we examined the relationship between baseline motor variability and learning rates in four different dynamic environments, each consisting of a different combination of position and velocity contributions (**Fig. 3e**).

Projecting the overall baseline variability onto the force patterns that are associated with each of the perturbations revealed that negative combination dynamics, which are the most difficult to learn (**Fig. 3f**), accounted for the smallest portion ($12 \pm 1\%$) of the baseline variability (**Fig. 3g**). In contrast, the positive combination dynamics, which are the easiest to learn, accounted for the largest portion ($34 \pm 2\%$). Across all four force fields, we found that the amount of task-relevant variability during baseline strongly predicted single-trial learning rates (**Fig. 3h**; $r = +0.94$), similarly to what we observed across reward-based learning tasks (**Fig. 1i–k**). These results show that task-relevant variability can predict motor learning ability both across individuals and across tasks in both reward-based and error-based learning.

Training paradigms that alter learning ability

Having established a relationship between motor variability and learning ability, we wondered whether the motor system is capable of capitalizing on this relationship to improve learning by modulating the structure of variability. Recent work in songbirds has suggested that the circuits generating motor variability can promote learning by directing exploration toward more rewarding regions of motor output space[30,31]. But can the motor system do more to promote efficient exploration than adaptively re-center motor output[30–32]? We considered the possibility that the structure of motor variability can also be reshaped around its mean, thus allowing for a more efficient pattern of exploration. Such an adaptive reshaping would require a specific increase in variability along the task-relevant dimension of motor output space.

To explore this hypothesis, we measured motor variability before and after a training paradigm that was designed to increase motor learning ability in a fourth experiment. We induced increases in learning rates for position-dependent or velocity-dependent force fields by exposing subjects to these force fields repeatedly in short blocks of seven trials interleaved with slightly longer blocks of unperturbed trials (**Fig. 4a**). This created environments that were highly consistent from one trial to the next, as characterized by high lag-1 autocorrelations[33] (Online Methods). We repeatedly exposed one group to the same position-dependent force field and exposed the other group to the same velocity-dependent force field; both fields were interleaved with identical unperturbed trial blocks. We first examined how these high-consistency environments (HCEs) affected learning ability and then determined whether they reshaped motor variability.

We found that prolonged and repeated exposure to the force fields in these environments resulted in single-trial adaptation that was not only larger in amplitude but also more specific to the trained environment compared to early exposure, where the adaptation was small in amplitude and largely nonspecific in shape[22,33]. To measure single-trial adaptation, we

presented error-clamp trials immediately before and after single-trial force-field exposures, and we compared the differences in the lateral force output to the ideal (full) adaptation, similarly to in previous studies[22]. We found that single-trial adaptation became larger and more specific for both velocity-dependent and position-dependent HCEs (**Fig. 4b,c**; the dark red and blue traces based on late exposure (blocks 40–48) to the velocity-dependent and position-dependent HCEs, respectively, can be compared to the lighter traces based on early exposure (block 1)). To quantify these effects, we used linear regression to determine the components of the single-trial adaptive responses that were associated with hand velocity and position[22]. We found a more than twofold increase in the velocity-dependent component of adaptation ($P < 0.0001$) and a more than fourfold increase in the position-dependent component ($P < 0.0001$) after training to the corresponding HCEs, with no significant changes in the untrained components (**Fig. 4d,e**; $P = 0.35$ and $P = 0.21$ and for position and velocity, respectively). This selective increase in learning rates resulted in increased task specificity in the adaptive response, which was readily apparent when projecting it into a position-velocity gain space (**Fig. 4f**). The adaptive response to the velocity-dependent force field displayed increased velocity specificity (i.e., is closer to the y axis in **Fig. 4f**, as quantified in **Fig. 4g**; $P < 0.0001$), whereas the position response showed increased position specificity (i.e., is closer to the x axis in **Fig. 4f**; $P = 0.016$) (**Fig. 4f,g**).

The temporal structure of motor output variability is reshaped to promote motor learning

We next examined whether these experience-dependent changes in learning ability were paralleled by changes in the temporal structure of motor variability. To quantify changes in the amplitude and structure of movement-related force variability, we scaled the unit vector characterizing the main axis of variability (i.e., the direction of PC_1) by the amount of variability it explained (**Fig. 4h,i**). We compared the scaled first principal component of motor output variability before and after exposure to the velocity-dependent (**Fig. 4h**) and the position-dependent (**Fig. 4i**) HCEs. The shape of PC_1 was well characterized by a linear combination of position, velocity and acceleration both before ($R^2 = 0.94$) and after ($R^2 = 0.96$ (velocity HCE); $R^2 = 0.94$ (position HCE)) exposure to the HCEs, which is in line with what we found for an independent data set (**Fig. 3d**; $R^2 = 0.95$).

Remarkably, we found that training in the velocity-dependent HCE induced a $78 \pm 21\%$ increase in the velocity-dependent component of PC_1 (**Fig. 4h,j**; $P < 0.0001$) without affecting the position component (**Fig. 4h,k**; $P = 0.34$). This effect is evidenced by a marked increase in the velocity-related component of PC1 (**Fig. 4h**, bold pink line) alongside essentially no change in the position-dependent component (**Fig. 4h**, thin light blue line). In contrast, exposure to the position-dependent HCE led to an $82 \pm 26\%$ increase in the amplitude of the position-dependent component of PC_1 (**Fig. 4i,k**; $P < 0.0001$) without affecting the velocity-component (**Fig. 4i,j**; $P = 0.19$). This effect is evidenced by a marked increase in the position-related component of PC_1 (**Fig. 4i**, bold light blue line) alongside essentially no change in the velocity-dependent component (**Fig. 4i**, thin pink line). Correspondingly, the position-velocity gain-space projections of these principal components showed that motor output variability changed in an environment-specific fashion (**Fig. 4l,m**). Both velocity- and position-dependent HCEs lead to increased task specificity of the motor

variability (**Fig. 4l,m**; $P < 0.0001$ in both cases), which is analogous to the changes we observed in motor learning ability (**Fig. 4f,g**).

An alternative analysis that did not rely specifically on PC_1 but instead considered all components of motor output variability yielded similar results. In particular, the overall changes in velocity-related and position-related variability also displayed environment-specific reshaping of the structure of motor variability (**Fig. 4n,o**). We found that the velocity-related variability was greater after velocity-dependent HCE training than after position-dependent HCE training or during baseline ($P < 0.0001$, $t(43.78) = 6.74$ and $P < 0.0001$, $t(54.24) = 4.58$, respectively). Analogously, we found that the position-related variability was greater after position-dependent HCE training than after velocity-dependent HCE training or during baseline ($P < 0.0001$, $t(42.84) = 3.34$ and $P < 0.0001$, $t(35.33) = 4.64$, respectively). Notably, over half of the specificity increases induced during the 90-min training sessions in the HCEs persisted into the next day (**Fig. 4l** and **Supplementary Fig. 8**; $P = 0.047$), suggesting that training-induced changes in motor output variability can be long lasting. Together these findings indicate that the motor system can reshape the temporal structure of motor output variability to align it with the environment and that this realignment that can persist from one day to the next.

DISCUSSION

We have shown that the temporal structure of motor variability at baseline predicts learning rates across both individuals and tasks. Remarkably, we found that individuals with higher task-relevant variability at baseline learned faster than those with lower baseline variability and that tasks associated with higher baseline variability in task-relevant dimensions elicited faster learning. Interestingly, we found that neither the inter-individual nor the inter-task effects of variability were specific to reward-based learning, as we also observed them in an error-based force-field adaptation paradigm. Taken together these results suggest a general principle whereby increased variability enables faster learning.

When we examined whether the motor system might exploit the relationship between variability and learning, we found that subjects consistently reshaped the structure of their motor variability in a manner that would promote learning in the trained environment. In particular, we found specific increases in the components of variability that were aligned with the trained environment, and these increases persisted from one day to the next. In sum, our findings suggest that motor variability should not be considered merely the inevitable consequence of neural noise in the motor system but should instead be viewed as a key ingredient of learning, which the motor system leverages by actively reshaping its structure.

Previous studies have identified genetic[34], structural[35,36] and neural activity markers[37–39] that correlate with learning rate; however, to our knowledge, our findings provide the first demonstration of the ability to predict differences in learning ability from baseline performance characteristics. Moreover, our results provide an explanation for why individual differences in motor learning ability may be task specific, as particular individuals may display above-average task-relevant variability for some tasks but below-average variability for others.

Motor variability and motor learning in the songbird

Our work was inspired largely by recent studies that link motor variability and learning in songbirds[13,15]. Previous studies showed that the cortical outflow of a song-specialized basal ganglia circuit (LMAN) essential for vocal learning[12–14] injects variability into the song-control circuits in a manner that can be regulated[13,14]. This provides a neural circuit mechanism for actively controlling motor output variability to promote learning. For example, during song learning, juvenile birds express substantially greater vocal variability than adults, which is in line with the idea that increased variability promotes rapid learning[40]. This observation cannot be explained simply by juvenile birds having a more immature motor system, as inactivations of LMAN yield song precision similar to that in adult crystallized song[13]. Moreover, birds significantly reduce their variability when performing for a potential mate as compared to solitary practice[14], demonstrating short-term and context-specific regulation of motor variability that, again, suggests a link between increased variability and learning situations. These observations are consistent with the need for high-variability exploration when rapid learning is called for and low-variability exploitation when precision is desired[41]. Consistent with these findings, birdsong learning is often cast in the framework of reinforcement learning, and most successful attempts at modeling the song learning process employ it[42,43]. Indeed recent paradigms in songbirds have used external reinforcers to shape the song, similarly to our experiments[15,30,32].

However, the leap from songbirds to humans is not a trivial one to make. Not only are the motor control circuits in humans and songbirds markedly different, but in addition, the highly evolved and specialized vocal behavior of songbirds may not speak to the diverse set of learned motor behaviors in humans. Now, by demonstrating a link between motor variability and learning in humans, we suggest that results from songbirds may indeed reflect a general principle of motor learning. We found that humans, like songbirds, can actively modulate motor output variability in the context of learning in the sense that the motor system contributes variability beyond the level of inevitable noise. Yet our current findings go beyond what has been found in songbirds in several ways. By demonstrating that motor variability predicts individual differences in learning ability, our findings strengthen the link between variability and learning and suggest that even subtle differences in the amount and structure of variability can affect learning. Moreover, we show that the modulation of variability is not limited to the overall amount, but rather its structure can be reshaped to increase the task-relevant component so that the variability in dimensions that are not relevant to learning can remain low. The ability to produce finely sculpted changes in the temporal structure of motor variability thus allows the motor system to improve the efficiency of learning by guiding exploration to the relevant parts of motor output space.

Exploration and exploitation in reward-based learning

A key concept in reinforcement learning theory is the idea that the learning system must both explore the environment to gain better knowledge about it and exploit current, albeit imperfect, knowledge. The processes of exploration and exploitation are often thought of as being in opposition to one another and are thus discussed in terms of a tradeoff between the two (the so-called 'exploration-exploitation tradeoff') such that more of one necessarily means less of the other. If this were the case, learners would have to sacrifice the degree to

which they exploit to increase exploration, and our results would thus need to imply that the benefit accrued from greater exploration outweighs the loss endured from commensurately reduced exploitation.

However, a tradeoff may not be required, at least for early learning. A reasonable framework for a task in a continuous output space is that exploitation sets the mean action, and exploration controls the variability around it. In this framework, greater exploration (i.e., greater variability) would reduce the expected reward if the reward surface had a local maximum near the mean (exploited) action (or, more generally, was predominantly concave down in shape), as increased exploration would result in actions that are farther away from the local maximum, leading to a lower expected reward. However, if the reward surface instead resembles a constant gradient (i.e., a line or a plane) in the neighborhood around the currently exploited action, then increased variability would be just as likely to lead to an increase in reward when compared to the exploited action as it would be to lead to a decrease in it. Thus, there would be no tradeoff between the amount of exploration (variability) and the ability to exploit the knowledge gained so far (i.e., to obtain the expected reward). In addition, if the reward surface is predominantly concave up in the neighborhood of the current action, exploration would systematically increase the expected reward. The reward surface in our shape-learning task (experiments 1 and 2) was essentially a constant gradient around the initial (baseline) actions, and we would argue that it is very reasonable to believe that during early learning, the reward surface in many reinforcement learning tasks also displays an approximately constant gradient. Insofar as this is the case, increased exploration would not generally reduce the average performance and thus the ability to exploit, allowing exploration to proceed independently of exploitation.

A Bayesian viewpoint

Bayes law, and thus Bayesian inference, allows for optimal integration of information in combining the probability distributions that are associated with prior expectations (the prior) with the probability distributions that are associated with the current sensory information (the likelihood). The alternating environments in experiment 4 could be viewed as broadening the prior in the task-relevant dimension, which would result in an increased learning rate for a Bayesian learner. A broader prior would also result in a broader posterior but would have no direct effect on the motor output variability for a Bayesian learner, such as a Kalman filter, that uses maximum *a posteriori* estimates. However, a Bayesian learner that generates output by sampling from the posterior[44] would display increased task-specific variability, which is consistent with our results.

Can motor adaptation ability affect motor output variability?

A limitation of the first three experiments is that the relationships we uncovered between variability and learning were correlational in nature. Therefore our results, although intriguing, fall short of establishing a causal link, as other variables may underlie the observed associations (**Supplementary Fig. 9**). One possibility we investigated is that high baseline variability might itself be a manifestation of increased learning ability. However, analysis of the baseline data from our experiments suggested that this is unlikely to be the case, as the positive autocorrelations present in this data would actually lead to a somewhat

negative relationship between learning rate and baseline variability (**Supplementary Fig. 10**).

Motor output variability and optimal control

Our present findings challenge current models of optimal control on two different levels. First, the experience-dependent changes we observed in motor variability showed that variability during active movement cannot be explained as simply signal-dependent noise. Current optimal control models generally assume this, and their predictions depend strongly on this assumption[6–8]. However, data supporting the signal-dependent noise hypothesis have been generated predominantly in isometric-force control experiments without regard to how force variability evolves during the course of active movement[2–4], raising the need for improved noise models that take into account how variability evolves during active movement.

Second, the idea that the motor system actively regulates motor variability to promote learning challenges the viewpoint that motor behavior should be modeled generally as optimally controlled performance. Theories of optimal control for motor planning hold that actions are planned to minimize an objective function describing a combination of the effort and motor error involved in the execution of a given plan. Such optimal planning incorporates the motor system's current knowledge about the specifics of the responses generated from the motor commands[9–11] and thus amounts to exploiting the current knowledge about a system to maximize performance. In contrast, our findings suggest that the human motor system does not simply exploit what it currently knows but instead actively engages in motor exploration, possibly sacrificing accurate performance in lieu of facilitating learning. Critically, our findings dispel the view that motor variability is nothing but a source of error to be overcome, demonstrating that current models of optimal control are relevant only in cases where exploitation dominates over exploration.

Motor variability improves error-based learning

The positive relationships we observed between task-relevant variability at baseline and subsequent motor learning ability in experiments 1 and 2 are in line with the idea from reinforcement learning theory that exploration enables reward-based learning. However, we found it somewhat surprising that the same relationship was present in error-based learning. One possible explanation for this finding involves the effects of motor variability on internal estimates of the gradient function between motor errors and motor commands. During reinforcement learning, exploration is required for discovering more rewarding actions because there is generally no *a priori* information about how motor output should be adjusted to obtain greater reward. Theories of error-based learning, however, posit that the motor system is equipped with not only with the ability to measure motor errors but also the knowledge of how adjustments in motor output reduce them[23,45,46]. This knowledge corresponds to the existence of an internal model of the gradient between motor errors and changes in motor commands. However, the internal representation of this gradient function is likely imperfect. We suggest that motor exploration provides information that is useful for improving the fidelity of the internal representation of the gradient function and the

confidence in that representation. This may explain the higher motor learning rates that we observed in error-based learning.

Taken together, our results support the view that motor variability, rather than being an unwanted consequence of noisy nervous system function, is an essential feature of motor learning[13–17] that is centrally driven[1,13,47] and actively regulated[14,48]. This view emerged from work in songbirds, in which motor variability and learning are coupled[13–15], and is further supported by experimental evidence in both songbirds[14,41] and primates[48,49] showing that motor variability is actively reduced when motor precision is crucial, such as when reward is at stake[14,41,49], and that motor variability is increased during learning[14,48,50]. Our current findings extend these observations by demonstrating that (i) learning ability is linked to motor variability in humans, (ii) variability can predict individual differences in learning ability, and (iii) the motor system does not merely modulate the overall amount of motor variability but instead actively reshapes its structure to direct exploration for more efficient learning. Elucidating the relationship between variability and learning not only enhances our basic understanding of learning in the motor system but also provides potential avenues for the rational design of new training procedures for improving motor learning and rehabilitation.

ONLINE METHODS

Subjects

All participants gave informed consent for the experimental procedures, which were approved by Harvard's Committee on the Use of Human Subjects. Eighty-four naive, neurologically intact subjects (age range 18–55 years, 39 male, 45 female, all right handed) participated in experiments 1 and 2, with 20 subjects in experiment 1 and 64 subjects in experiment 2 (29 subjects in subgroup A and 35 subjects in subgroup B). Forty right-handed, neurologically intact subjects (20 female, 20 male, age range 18–31 years) participated in experiment 3, and 24 naive neurologically intact right-handed individuals (16 female, 8 male, age range 18–58 years) participated in experiment 4. The sample sizes for experiments 1 and 2 were determined based on pilot data, and the sample sizes for experiments 3 and 4 were based on previous experience from similar experiments and the existing literature[22,23,28].

Methods for the reward-based learning experiments (experiments 1 and 2)

Participants and experimental paradigm—Subjects performed rapid 200-mm point-to-point reaching movements while grasping a handle, whose position was recorded on a high-resolution digitizing tablet (Wacom Intuos3) at a sampling rate of 200 Hz. A monitor, horizontally mounted above the tablet, obstructed vision of the hand (**Fig. 1a**) and displayed the start location, the target and a guide shape connecting the two positions, all of which remained identical throughout the experiment. Subjects were instructed to trace the guide shape quickly, moving from the starting location to the target; however, subjects received no visual feedback about their hand positions during each trial. Movements that failed to reach the target within 450 ms of movement onset (defined by a velocity threshold of 12.7 mm s^{-1}) were discouraged with immediate negative auditory feedback.

A 250-trial baseline period was followed by the training period (500 trials for experiment 1 and 1,000 trials for experiment 2), which was identical to the baseline except that each trial was awarded with a numerical score between 0 and 1,000 that was displayed above the target. Subjects were instructed to maximize their score by more accurately tracing the displayed curve. However, unbeknownst to the subjects, each score was calculated based on the normalized projection of the hand path onto the shape of the rewarded deflection, which was unrelated to the displayed path, as described in the next section. Notably, the lack of visual feedback eliminated any obvious error signal that could consistently improve performance on this task.

During the training period, the threshold movement time was changed from 450 ms to a new value based on each subject's baseline movement durations. The new threshold was set to be 2 s.d. above the mean baseline movement duration and averaged 429 ms with a range of 289–529 ms. Movements that failed to reach the target within this threshold time during training were discouraged with negative feedback and withholding of the score. We excluded less than 4% of the overall training trials based on this threshold after omitting 3 of the 84 subjects, who did not consistently complete trials in the allotted movement time (in all three, over 50% of trials were too slow).

Learning level—A dimensionless learning level (z) for each trial was calculated by projecting the vector of x-positions of the hand path (from y-positions of 15–190 mm of each 200-mm movement as seen in **Fig. 1b**) onto the vector of x-positions of one of the rewarded deflections (**Fig. 1c**):

$$z = \underset{c}{\operatorname{argmin}}\left(\left(\vec{x}(y) - \vec{x}_0(y)\right) - c \cdot \vec{r}(y)\right)^2 = \frac{\left(\vec{x}(y) - \vec{x}_0(y)\right)}{\|r\|} \cdot \frac{\vec{r}(y)}{\|r\|} = \frac{\left(\vec{x}(y) - \vec{x}_0(y)\right) \cdot \vec{r}(y)}{\|r\|}$$

Here $\vec{x}(y)$ denotes the x-positions of the hand during one trial as a function of y-positions and $\vec{x}_0(y)$ denotes the average x-positions of the hand from the last 160 trials of the baseline period. Thus $\vec{x}(y) - \vec{x}_0(y)$ corresponds to the actual deflection with respect to baseline in each trial, which is the quantity being rewarded. $\vec{r}(y)$ denotes the rewarded deflection—i.e., the deflection that is being trained. Argmin of an expression denotes the value of the scalar c that minimizes that expression. Thus, z corresponds to the scaling of $\vec{r}(y)$ that would make it most similar to $\vec{x}(y) - \vec{x}_0(y)$—i.e., the scaling of $\vec{x}(y)$ that minimizes the squared error between $z\vec{r}(y)$ and $\vec{x}(y) - \vec{x}_0(y)$. To generate $\vec{x}(y)$ and $\vec{x}_0(y)$, we linearly interpolated the x-positions of the hand path onto a vector of y-positions every 0.254 mm to align the hand path measurements across trials. A learning level of 1 corresponds to the ideal amount of deflection relative to baseline performance. Scores awarded in the training trials were calculated as a function of learning level (**Supplementary Fig. 1b,c**).

Rewarded deflections—The two rewarded deflections (shape 1 and shape 2) were orthogonal, with identical root mean square amplitudes (3.6 mm). This amplitude was based on pilot data indicating that it allowed for robust learning in a reasonable amount of time (~1

h of training). Of note, the shapes of the rewarded deflections were unrelated to the guide shape shown to subjects. Experiment 1 trained subjects on shape 1, and experiment 2 trained subjects on either shape 1 or shape 2. All subjects were split randomly into two groups: 45 randomly chosen subjects were trained with positive deflections, and the remaining 40 subjects were trained with negative deflections (thick versus thin lines in **Fig. 1c**), thus ensuring that any drifts that might occur could not consistently promote or impede learning.

Measuring baseline variability and learning rate for experiments 1 and 2—We quantified task-relevant variability at baseline using the standard deviation of the learning levels of the last 160 trials from baseline. We quantified total baseline variability using the square root of total variance summed across all points in the path $x(y)$ with units of mm RMS to facilitate an intuitive interpretation of the magnitude. On average, shape 1 and shape 2 accounted for $9.2 \pm 1.1\%$ and $0.66 \pm 0.12\%$, respectively, of the total variability during the baseline period.

To compute the average learning level shown in **Figure 1f–h**, we used the average learning level during the first 125 trials for experiment 1 (**Fig. 1f–h**) and during the first 800 trials for experiment 2 (**Fig. 1i–k**). We used a longer window in experiment 2 because the learning curves, particularly those for shape 2, were slower. Had we used a 125-trial window for experiment 2, the slower curve for shape 2 learning would dictate an average learning level that was not significantly different from zero, precluding us from examining inter-individual differences in learning.

Force-field experiments (experiments 3 and 4)

Setup and instructions—Subjects were instructed to grasp the handle of a robotic manipulandum while making rapid (500-ms), point-to-point reaching arm movements (10 cm in length) (**Fig. 2a**). Experiments consisted of three different trial types: null-field trials during which no active forces were applied to the subject's arm (**Fig. 2b**), error-clamp trials to measure the lateral forces produced (**Fig. 2d**) and force-field perturbation trials (**Fig. 2c**). The position and velocity of the hand and the forces generated by the manipulandum were recorded at a sampling rate of 200 Hz.

Force-field environments—The force fields used in the current experiments were composed of a linear combination of position and velocity dependence (**Fig. 3e** with the form

$$\begin{bmatrix} F_x \\ F_y \end{bmatrix} = K \cdot \vec{p} + B \cdot \vec{v} = \begin{bmatrix} 0 & -K \\ K & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 0 & -B \\ B & 0 \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix}$$

where x and y and denote the x and y positions $(\vec{p})$ of the hand and $\dot{x}$ and $\dot{y}$ denote the x and y velocities $(\vec{v})$ with axes as illustrated in **Figure 2a**. Velocity-dependent force fields had values of $K = 0$ N m^{-1} and $B = \pm15$ Ns m^{-1}, position-dependent force fields had values of $K = \pm45$ N m^{-1} and $B = 0$ Ns m^{-1}, and positive-combination and negative-combination force fields had values of $K = \pm21.2$ N m^{-1} and $B = \pm13.2$ Ns m$^-$, and $K = \pm35$ N m$^-$ and $B =$

± 9.4 Ns m^{-1}, respectively. To account for biases from biomechanical effects, clockwise and counterclockwise versions of force fields were balanced in all experiments.

Error-clamp trials—Error-clamp trials were designed to measure the feed-forward motor output produced during a reaching movement. Actions made during reaching movements result from feedforward motor output and online feedback error correction. Error-clamp trials[21–23] restricted the lateral deviations during movement below 1 mm, effectively eliminating the lateral error signal and allowing for isolation of feedforward motor output (**Fig. 2d**). In error-clamp trials, we applied a damped high-stiffness elastic force ($K = 6,000$ N m^{-1}, $B = 250$ Ns m^{-1}) to counteract the subjects' lateral forces, essentially clamping movements within a straight channel toward the target. Although the movement duration was generally between 500 and 600 ms, we examined the force output generated in an 860-ms window centered at the peak speed point to ensure that we captured the entire movement. Force data were smoothed with a second-order Butterworth filter with a cutoff frequency of 10 Hz to remove high-frequency noise generated by force sensors and motor actuators.

Experiment 3 design—Experiment 3 examined the relationship between person-to-person differences in baseline task-relevant variability and initial force-field learning rates. To obtain accurate measurements of baseline variability, we designed a prolonged baseline period of 150 trials with 20 errorclamp trials that were interspersed among 130 null-field trials (**Fig. 2e**), which followed a 100-trial familiarization period during which variability was not measured. After the baseline period, subjects experienced a 150-trial training period during which a velocity-dependent force-field environment was applied (**Fig. 2e**). In this training period, 80% of trials were force-field trials and 20% were error-clamp trials intermixed to measure the learning level during the course of training.

Measuring baseline variability and initial learning rate for experiment 3— Because subjects would later be exposed to a velocity-dependent force field, we analyzed what would become the task-relevant component of the variability by calculating the amount of velocity-dependent variability present during baseline error-clamp trials. We computed the velocity-dependent component of variability by projecting each force trace onto its corresponding velocity profile (**Supplementary Fig. 11**) and calculating the standard deviation of the magnitudes of these projections.

As in previous studies[22,23,28], we performed per-subject baseline subtraction using the last five error clamps in the baseline, $F_0(t)$, as a reference for learning-related changes in force output measurements in error-clamp trials in the training period, $F(t) - F_0(t)$, which were regressed onto a linear combination of each trial's motion states: position, velocity and acceleration (**Supplementary Fig. 11d,e**). We determined the learning level for the pure velocity-dependent force-field environment by normalizing the velocity regression coefficient so that a value of 1 indicated full learning. An initial learning rate was calculated for each subject by finding the average rate of increase in learning level over the first ten trials of training, a period that included two error-clamp trials.

Analyzing the temporal structure of baseline variability—To examine the overall structure of variability, we performed principal component analysis[51] on the aggregated

baseline lateral force traces for all subjects in experiment 3 after performing subject-by-subject baseline subtraction to prevent individual differences in mean behavior from influencing the calculation of the principal components. We computed the principal components of the force variability by performing eigenvalue decomposition on the covariance matrix of the aggregated force data. **Figures 3** and **4** scale each principal component vector by the square root of its eigenvalue to visualize the amount of variability it explains. The fraction of overall variance (**Fig. 3b**) was calculated as the eigenvalue (variance) for a particular principal component as a fraction of the sum of the eigenvalues for all principal components.

We then regressed each scaled eigenvector onto a linear combination of position, velocity and acceleration. We operationally define motion-related variance as variance that can be explained by a linear combination of position, velocity and acceleration. We then scaled each eigenvalue by the fraction of its corresponding eigenvector's variance related to its motion state to obtain the motion-related eigenvalue. To determine the fraction of motion-related variance each principal component contained (**Fig. 3c**), we found the ratio of each motion-related eigenvalue to the sum of the motion-related eigenvalues for all principal components.

Explaining learning-rate differences across different dynamic environments—
We found the variance associated with four different force-field environments (**Fig. 3e**) during baseline by projecting each baseline error-clamp force trace from experiment 3 onto motion-dependent traces that were representative of each force-field environment. These representative traces were generated by linearly combining the position and velocity traces of the hand in the associated error-clamp trial after scaling by the K and B values that were associated with each environment. The variance of the magnitudes of the projections was divided by the total variance of the force traces to compute the fraction of variance associated with each type of force field (**Fig. 3g**).

The single-trial learning rates for each force-field environment (**Fig. 3f**) were determined on the basis of data reported in a previous publication[22]. In this previous experiment, error-clamp trials were presented before and after a single-trial force-field exposure, and the differences in the force output were used to assess single-trial learning rates. Because the previous study used a different window size, we recomputed the learning rates in the same 860-ms window that we used to assess motor output variability.

Experiment 4 design—The training paradigm of experiment 4 (**Fig. 4a**) was designed to increase the learning rate for either velocity-dependent or position-dependent force fields by creating HCEs. HCEs, in which consistency is operationally defined as the correlation between the force field in the current trial versus that in the next trial (i.e., the lag-1 autocorrelation), have been shown to increase learning rate[33] and were created here by exposing subjects to 48 runs of 7 force-field trials (either position or velocity dependent), each of which was followed by a run of 8–12 null-field trials.

Motor variability was assessed during two epochs on each day: immediately preceding training and after an extended washout period (26–33 null-field trials) after training. Each

assessment consisted of 50 error-clamp trials interspersed among 150 null-field trials. Half of the subjects were trained on the HCE for the position-dependent force field on day 1 and for the velocity-dependent force field on day 2. The other half of the subjects experienced these environments in the opposite order. The initial assessment of variability on day 2 was also used to probe the degree to which changes in variability induced on day 1 persisted to the next day. Consequently, variability data from immediately before and after exposure were available for all environmental exposures we studied, whereas next-day retention was available for the day 1 environment data (half the data set) only.

Measuring changes in the amount and specificity of motor learning—The single-trial learning rate was assessed in 21 of the 48 force-field runs using error-clamp triplets[22] centered around the first force-field trial in each of these pseudorandomly chosen runs. In these triplets, subjects performed an error-clamp trial, followed by a force-field trial and then another error-clamp trial. Pretraining learning data were obtained in the very first triplet before any exposure to the HCE. Because each subject thus contributed only a single triplet, we supplemented this data with first-exposure single-trial learning data from other experiments (58 additional velocity learning trials and 12 additional position learning trials) that were essentially identical to the current experiment up to the point of this first exposure[22,33], with the only difference being the number of baseline trials experienced beforehand. Post-exposure single-trial learning was computed from the averaged learning data from the last nine training blocks, which included five measurement triplets. We measured task-specific adaptation for position- and velocity-dependent force fields by regressing the learning-related changes in the force profile onto the motion-state profiles.

We quantified the specificity of single-trial adaptation using a position-velocity gain-space analysis. Task-specific adaptation for velocity-dependent learning is shown on the vertical axis in the relevant figures (**Fig. 4f,l**), and task-specific adaptation for position-dependent learning is shown on the horizontal axis. In this gain space, velocity-specific learning would correspond to an adaptation vector with a 90° angle, whereas position-specific would correspond to a 0° angle.

Measuring changes in the amount and specificity of motor output variability—We examined the effects of the HCE on the structure of motor variability by looking at the position- and velocity-dependent changes in PC_1. Data were aggregated across all subjects, as in the analysis of the temporal structure of baseline variability in experiment 3 detailed above. We computed PC_1 and its motion-state fits in three different epochs: before training, after velocity training and after position training. Similarly to the changes in single-trial force-field adaptation, we can examine the changes in PC_1 using a position-velocity gain-space analysis. An alternate way to look at changes in the structure of motor variability is to look at changes in position- and velocity-related variability. We used linear regression to find the position- and velocity-dependent contributions of the force traces in each epoch and determined the fractions of variance accounted for by position and velocity as the ratios of the variance of these regression coefficients to the total variance in the data.

Statistical methods and randomization of subjects

Subjects were assigned to experiments based only on when they responded to advertisements about participation and their availability for scheduling. When different experiments were running concurrently, subjects were assigned randomly for participation. In most cases, data were successively collected for different experiments. Although subjects were unaware of their experimental group assignments, investigators were aware of these assignments during the analysis.

Statistical testing for group learning rates (**Figs. 1f,i** and **2g** and **Supplementary Figs. 4** and **6**) and group variability (**Figs. 1f,i** and **4n,o**) were performed by using two-sample one-sided Student's t tests without assuming equal variances. Linear regression analyses based on the t statistic for the slope of a best-fit straight line were used to test for positive correlations between variability and learning rate (**Figs. 1g,h,j,k** and **2h** and **Supplementary Figs. 2**, **3** and **9**). An F test was used to test (i) whether separate linear models of variability versus learning for shape 1 and shape 2 in experiment 2 accounted for greater variance than a single linear model of variability versus learning across these two tasks (**Fig. 1j,k**) and (ii) whether task-relevant variability and total variability individually contributed to a bivariate regression of early learning onto these two quantities for the experiment 2 data. The data we analyzed using F and t statistics were inspected and found to appear grossly normal in their distributions. Because we were computing the principal components of the lateral force variability over the population (**Figs. 3b–d** and **4h–m**), confidence intervals around PC_1 (**Figs. 3d** and **4h,i**) were computed using bootstrap analysis in which the population of participants was resampled 10,000–100,000 times to determine the confidence bounds for PC_1. In this bootstrap analysis, PC_1 was computed for each iteration using the covariance of the mean-subtracted force profiles from the randomly sampled subpopulation. A similar bootstrap was performed to compute confidence intervals for single-trial learning in experiment 4 (**Fig. 4b,c**) by resampling adaptation force profiles from the population. Correspondingly, statistical testing on the changes in the position and velocity contributions to adaptation and variability (**Fig. 4d,e,j,k**) were performed with a bootstrap, as were the changes in position-velocity gain-space angle (**Fig. 4g,m**).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

References

1. Churchland MM, Afshar A, Shenoy KV. A central source of movement variability. Neuron. 2006; 52:1085–1096. [PubMed: 17178410]

2. Jones KE, Hamilton AFC, Wolpert DM. Sources of signal-dependent noise during isometric force production. J. Neurophysiol. 2002; 88:1533–1544. [PubMed: 12205173]

3. Schmidt RA, Zelaznik H, Hawkins B, Frank JS, Quinn JT Jr. Motor-output variability: a theory for the accuracy of rapid motor acts. Psychol. Rev. 1979; 47:415–451.

4. Stein RB, Gossen ER, Jones KE. Neuronal variability: noise or part of the signal? Nat. Rev. Neurosci. 2005; 6:389–397. [PubMed: 15861181]

5. Osborne LC, Lisberger SG, Bialek W. A sensory source for motor variation. Nature. 2005; 437:412–416. [PubMed: 16163357]

6. Harris CM, Wolpert DM. Signal-dependent noise determines motor planning. Nature. 1998; 394:780–784. [PubMed: 9723616]

7. van Beers RJ, Baraduc P, Wolpert DM. Role of uncertainty in sensorimotor control. Phil. Trans. R. Soc. Lond. B. 2002; 357:1137–1145. [PubMed: 12217180]

8. Scholz JP, Schöner G. The uncontrolled manifold concept: identifying control variables for a functional task. Exp. Brain Res. 1999; 126:289–306. [PubMed: 10382616]

9. Todorov E. Optimality principles in sensorimotor control. Nat. Neurosci. 2004; 7:907–915. [PubMed: 15332089]

10. Todorov E, Jordan MI. Optimal feedback control as a theory of motor coordination. Nat. Neurosci. 2002; 5:1226–1235. [PubMed: 12404008]

11. O'Sullivan I, Burdet E, Diedrichsen J. Dissociating variability and effort as determinants of coordination. PLoS Comput. Biol. 2009; 5:e1000345. [PubMed: 19360132]

12. Charlesworth JD, Warren TL, Brainard MS. Covert skill learning in a cortical-basal ganglia circuit. Nature. 2012; 486:251–255. [PubMed: 22699618]

13. Ölveczky BP, Andalman AS, Fee MS. Vocal experimentation in the juvenile songbird requires a basal ganglia circuit. PLoS Biol. 2005; 3:e153. [PubMed: 15826219]

14. 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]

15. Tumer EC, Brainard MS. Performance variability enables adaptive plasticity of 'crystallized' adult birdsong. Nature. 2007; 450:1240–1244. [PubMed: 18097411]

16. Sutton, RS., Barto, AG. Introduction to Reinforcement Learning. MIT Press; 1998.

17. Kaelbling LP, Littman ML, Moore AW. Reinforcement learning: a survey. J. Artif. Intell. Res. 1996; 4:237–285.

18. Roberts S, Gharib A. Variation of bar-press duration: where do new responses come from? Behav. Processes. 2006; 72:215–223.

19. Stahlman WD, Blaisdell AP. The modulation of operant variation by the probability, magnitude, and delay of reinforcement. Learn. Motiv. 2011; 42:221–236. [PubMed: 21731111]

20. 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]

21. Scheidt RA, Reinkensmeyer DJ, Conditt MA, Rymer WZ, Mussa-Ivaldi FA. Persistence of motor adaptation during constrained, multi-joint, arm movements. J. Neurophysiol. 2000; 84:853–862. [PubMed: 10938312]

22. Sing GC, Joiner WM, Nanayakkara T, Brayanov JB, Smith MA. Primitives for motor adaptation reflect correlated neural tuning to position and velocity. Neuron. 2009; 64:575–589. [PubMed: 19945398]

23. Smith MA, Ghazizadeh A, Shadmehr R. Interacting adaptive processes with different timescales underlie short-term motor learning. PLoS Biol. 2006; 4:e179. [PubMed: 16700627]

24. Sing GC, Smith MA. Reduction in learning rates associated with anterograde interference results from interactions between different timescales in motor adaptation. PLoS Comput. Biol. 2010; 6:e1000893. [PubMed: 20808880]

25. Joiner WM, Smith MA. Long-term retention explained by a model of short-term learning in the adaptive control of reaching. J. Neurophysiol. 2008; 100:2948–2955. [PubMed: 18784273]

26. Bays PM, Flanagan JR, Wolpert DM. Interference between velocity-dependent and position-dependent force-fields indicates that tasks depending on different kinematic parameters compete for motor working memory. Exp. Brain Res. 2005; 163:400–405. [PubMed: 15856202]

27. Diedrichsen J, Criscimagna-Hemminger SE, Shadmehr R. Dissociating timing and coordination as functions of the cerebellum. J. Neurosci. 2007; 27:6291–6301. [PubMed: 17554003]

28. Joiner WM, Ajayi O, Sing GC, Smith MA. Linear hypergeneralization of learned dynamics across movement speeds reveals anisotropic, gain-encoding primitives for motor adaptation. J. Neurophysiol. 2011; 105:45–59. [PubMed: 20881197]

29. Conditt MA, Mussa-Ivaldi FA. Central representation of time during motor learning. Proc. Natl. Acad. Sci. USA. 1999; 96:11625–11630. [PubMed: 10500227]

30. Andalman AS, Fee MS. A basal ganglia-forebrain circuit in the songbird biases motor output to avoid vocal errors. Proc. Natl. Acad. Sci. USA. 2009; 106:12518–12523. [PubMed: 19597157]

31. 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]

32. Ali F, Otchy TM, Pehlevan C, Fantana AL, Burak Y, Ölveczky BP. The basal ganglia is necessary for learning spectral, but not temporal, features of birdsong. Neuron. 2013; 80:494–506. [PubMed: 24075977]

33. Gonzales-Castro LN, Hemphill M, Smith MA. Learning to learn: environmental consistency modulates motor adaptation rates. Proc. Ann. Symp.: Advances in Comp. Motor Control. 2008; 7

34. Frank MJ, Doll BB, Oas-Terpstra J, Moreno F. Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. Nat. Neurosci. 2009; 12:1062–1068. [PubMed: 19620978]

35. Della-Maggiore V, Scholz J, Johansen-Berg H, Paus T. The rate of visuomotor adaptation correlates with cerebellar white-matter microstructure. Hum. Brain Mapp. 2009; 30:4048–4053. [PubMed: 19507158]

36. Tomassini V, et al. Structural and functional bases for individual differences in motor learning. Hum. Brain Mapp. 2011; 32:494–508. [PubMed: 20533562]

37. Rutishauser U, Ross IB, Mamelak AN, Schuman EM. Human memory strength is predicted by theta-frequency phase-locking of single neurons. Nature. 2010; 464:903–907. [PubMed: 20336071]

38. Berry SD, Thompson RF. Prediction of learning rate from the hippocampal electroencephalogram. Science. 1978; 200:1298–1300. [PubMed: 663612]

39. Tamás Kincses Z, et al. Model-free characterization of brain functional networks for motor sequence learning using fMRI. Neuroimage. 2008; 39:1950–1958. [PubMed: 18053746]

40. Tchernichovski O, Mitra PP, Lints T, Nottebohm F. Dynamics of the vocal imitation process: how a zebra finch learns its song. Science. 2001; 291:2564–2569. [PubMed: 11283361]

41. Woolley SC, Doupe AJ. Social context–induced song variation affects female behavior and gene expression. PLoS Biol. 2008; 6:e62. [PubMed: 18351801]

42. Doya K, Sejnowski T. A novel reinforcement model of birdsong vocalization learning. Adv. Neural Inf. Process. Syst. 1995; 8:101–108.

43. Fiete IR, Fee MS, Seung HS. Model of birdsong learning based on gradient estimation by dynamic perturbation of neural conductances. J. Neurophysiol. 2007; 98:2038–2057. [PubMed: 17652414]

44. Daw ND, Niv Y, Dayan P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. Nat. Neurosci. 2005; 8:1704–1711. [PubMed: 16286932]

45. Thoroughman KA, Shadmehr R. Learning of action through adaptive combination of motor primitives. Nature. 2000; 407:742–747. [PubMed: 11048720]

46. Wagner MJ, Smith MA. Shared internal models for feedforward and feedback control. J. Neurosci. 2008; 28:10663–10673. [PubMed: 18923042]

47. Sober SJ, Wohlgemuth MJ, Brainard MS. Central contributions to acoustic variation in birdsong. J. Neurosci. 2008; 28:10370–10379. [PubMed: 18842896]

48. 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]

49. 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]

50. 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]

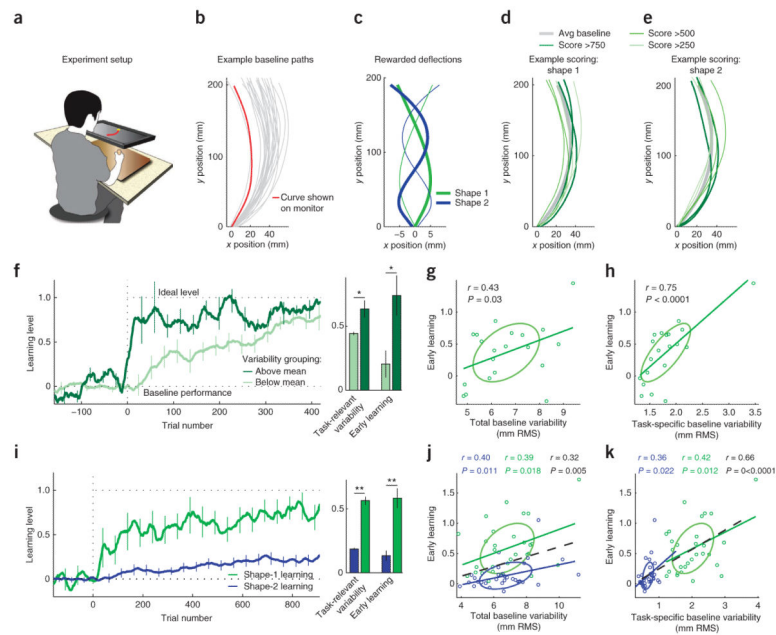51. Jolliffe, I. Principal Component Analysis. John Wiley & Sons Ltd; 2005.

Figure 1.
Task-relevant variability predicts the rate for reward-based learning. (**a**–**c**) Task description for the reward-based learning experiments. (**a**) Basic setup (additional details are provided in the Online Methods). Subjects were asked to follow a trace shown on a screen (red), which was unrelated to the rewarded deflections shown in **c** and remained unchanged for the entire duration of the experiment. (**b**) Example baseline movements from one participant showing the pattern of trial-to-trial variability. Of note, the *x* axis is magnified by a factor of two as compared to the *y* axis. (**c**) The shapes of the rewarded deflections used for experiments 1 and 2. (**d,e**) Example scored movements in comparison to the mean baseline trajectory (gray). The rewarded deflection was either shape 1 (**d**) or shape 2 (**e**). Darker shades of green correspond to higher scored movements. (**f**–**k**) Analysis of experiments 1 and 2. The learning level was computed as the magnitude of the normalized projection (dot product) of the vector of *x* positions of a movement trajectory onto the vector of *x* positions of the shape of the rewarded deflection after aligning the movement trajectories by their *y* coordinates. The scores displayed at the end of the trial were based on this learning level (Online Methods). Early learning was computed as the mean of the learning levels of the first 125 trials for experiment 1 and the first 800 trials for experiment 2, in which learning was considerably slower on average. (**f**) Participants displaying above-average amounts of shape-1 variability ($n = 6$) during the baseline period in experiment 1 exhibit faster learning than participants with below-average variability ($n = 14$). In the inter-individual analyses in **g**, **h**, **i** and **k**, small open circles show data from individual subjects alongside linear fits and 1 s.d. confidence ellipses. (**g,h**) Subject-by-subject comparison of early learning in experiment 1 as a function of total or task-relevant (shape 1) variability in the preceding baseline period. In **h**, the subject with the highest task-relevant baseline variability epitomized the positive relationship ($r = +0.75$, $P < 0.0001$) that we observed between learning and variability by displaying the highest learning rate. However, even if we remove this data point, the correlation between task-relevant variability and learning rate remains (*r*

= +0.61, $P = 0.0027$). Units of variability are mm root mean squared (RMS). Hence, a subject's amount of variability can be intuitively interpreted as the amount of lateral spread of the movements in physical space (in mm) on average along the movement. (**i**) Comparison of the curves for shape-1 compared to shape-2 learning in experiment 2 showing that shape-1 learning, for which there was greater task-relevant baseline variability, proceeded at a faster rate. (**j**) Total variability predicts individual differences within tasks but not across tasks. In **j** and **k**, the colored lines depict the regression for each task, whereas the dashed black line depicts regression across both data sets. (**k**) Task-relevant variability predicts individual differences both within and across tasks. Correspondingly, the bivariate linear regression analysis reported in the text shows a significant effect of task-relevant but not total variability. All error bars represent the s.e.m. Asterisks indicate significance (*$P <$ 0.05, **$P <$ 0.005).
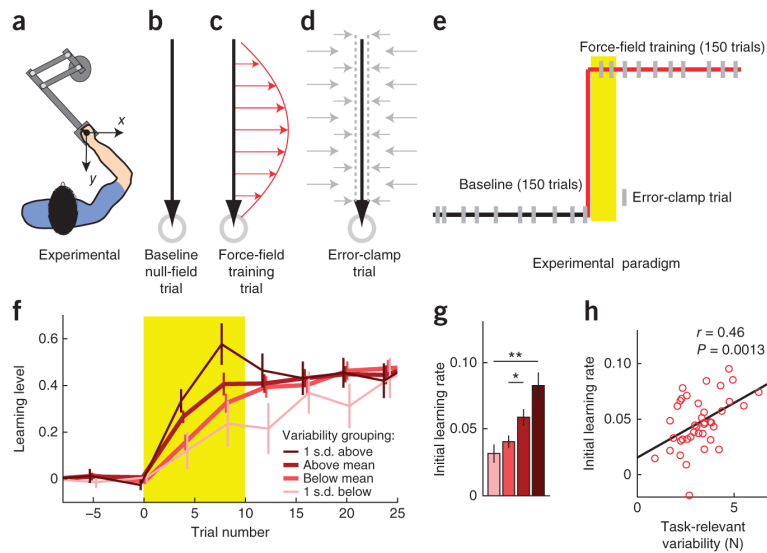
Figure 2.
Task-relevant variability predicts the initial learning rate in error-based learning. (**a**) Basic experimental setup for the velocity-dependent force-field adaptation experiments. (**b–d**) Illustration of the different trial types. (**e**) Trial schedule of the experiments comprising different trial types. In this depiction, the *x* axis corresponds to trial number, and the *y* axis corresponds to the level of perturbation (null versus force field). The different trial types shown in **b–d** are color coded. The first 40 of 150 trials in the training period are illustrated, as are the last 40 of 150 in the baseline period. We defined the initial learning as the period of the first ten trials (highlighted in yellow). (**f,g**) Comparison of the learning curves and initial learning in the first ten trials for participants stratified by task-relevant variability during the baseline period. (*n* = 4, 23, 17 and 5 for the four subgroups shown from left to right in **g**, respectively). (**h**) Subject-by-subject comparison of learning during the first ten trials of the training period, with the amount of total or task-relevant (velocity-dependent) variability displayed in the preceding baseline period in units of Newtons (N). All error bars represent the s.e.m. Asterisks indicate significance (*$P < 0.05$, **$P < 0.005$).
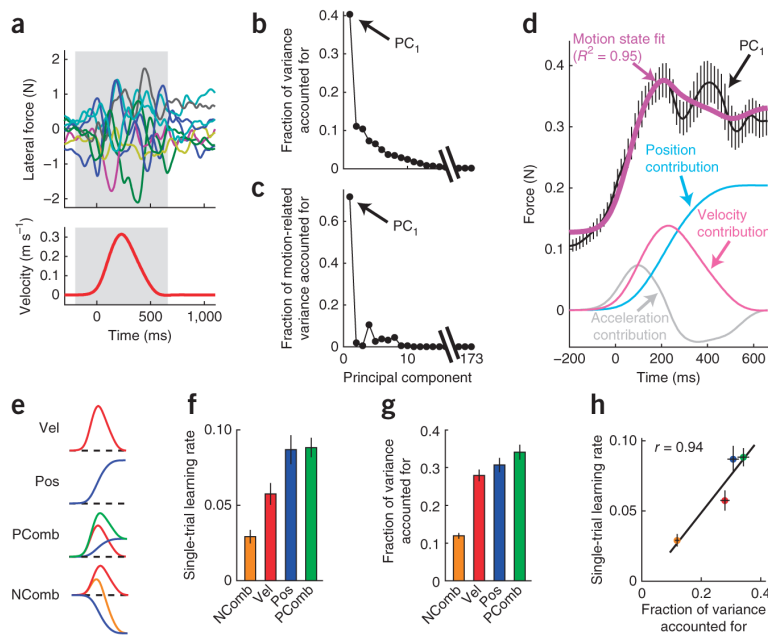
Figure 3.
The structure of force output variability and the prediction of learning rate for different types of dynamics. (**a**) Example lateral force profiles from one participant aligned to the corresponding movement's peak speed point. The mean velocity profile is plotted below as a timeline reference. The gray shaded area indicates the time interval over which the data were analyzed in **b**–**h**). (**b**–**d**) Principal component analysis of baseline lateral force variability. (**b,c**) Fraction of total and motion-related variance accounted for by each principal component. Of note, PC$_1$ accounts for over 40% of the total variance and over 70% of the motion-related variance. (**d**) The shape of PC$_1$ is highly dependent on motion state ($R^2 = 0.95$) and corresponds closely to a positive linear combination of the mean position and velocity traces. (**e**) Illustration of four different patterns of dynamics[22]. Red, velocity (Vel); blue, position (Pos); green, positive combination (PComb) of position and velocity; purple, negative combination (NComb) of position and velocity. (**f**–**h**) The fraction of variance accounted for by projecting the force patterns onto each type of dynamics strongly correlates with the single-trial learning rates ($r = +0.94$). All error bars represent the s.e.m.
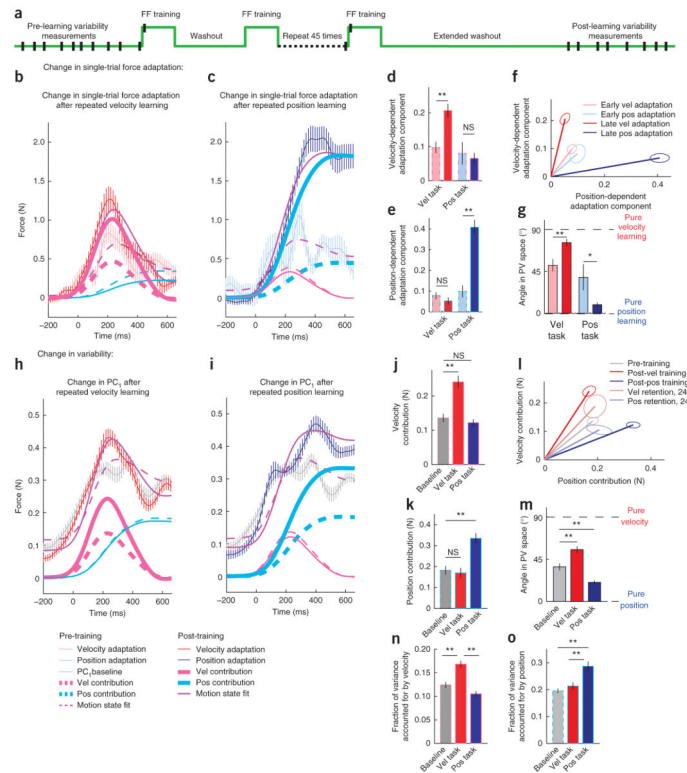
Figure 4.
The structure of motor output variability is reshaped by the nervous system. (**a**) Trial schedule of the experiments. Black dashes denote error-clamp trials. We assessed changes in learning rate and variability induced by repeated cycles of training and washout trials. Each cycle consisted of 7 force-field trials and 8–12 null-field trials. We refer to this cyclic trial schedule as an HCE. (**b,c**) Specific increases in single-trial force adaptation after exposure to HCEs. (**d,e**) The velocity component of adaptation increases only with exposure to the velocity HCE (vel task), whereas the position component of adaptation increases only with exposure to the position HCE (pos task). (**f**) Plotting single-trial adaptation in a position-velocity gain space reveals nonspecific adaptation during the initial exposure but specific adaptation during late exposure to the type of HCE. In **f** and **l**, the ellipses represent 1 s.e.m. (**g**) The angle of single-trial adaptation in this position-velocity (PV) gain space reveals significantly increased specificity after HCE exposure. (**h**–**m**) As in **b**–**g** except that the analysis describes changes in PC$_1$, which characterizes the main axis of the motor variability. (**h,i**) Specific increases in the velocity- and position-dependent components of PC$_1$ result from exposure to the HCEs. (**j,k**) The velocity component of PC$_1$ increases only with exposure to the velocity HCE, whereas the position component of PC$_1$ increases only with exposure to the position HCE. (**l**) Plotting PC$_1$ in a position-velocity gain space reveals that exposure to HCEs aligns PC$_1$ to the trained HCE force field. (**m**) The angle of PC$_1$ in the position-velocity gain space reveals significantly increased specificity after HCE exposure. (**n,o**) An alternative analysis shows specific increases in overall position- and velocity-dependent variability resulting from HCE exposure. All error bars represent the s.e.m. Asterisks indicate significance (*$P < 0.05$, **$P < 0.005$; NS, not significant).