

Motor Learning Enhances Use-Dependent Plasticity

 **Firas Mawase**,¹  **Shintaro Uehara**,^{1,5}  **Amy J. Bastian**,^{2,4} and  **Pablo Celnik**^{1,2,3}

¹Department of Physical Medicine and Rehabilitation, ²Department of Neuroscience, and ³Department of Neurology, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205-2196, ⁴Kennedy Krieger Institute, Baltimore, Maryland 21205, and ⁵Japan Society for the Promotion of Science, Tokyo 102-0083, Japan

Motor behaviors are shaped not only by current sensory signals but also by the history of recent experiences. For instance, repeated movements toward a particular target bias the subsequent movements toward that target direction. This process, called use-dependent plasticity (UDP), is considered a basic and goal-independent way of forming motor memories. Most studies consider movement history as the critical component that leads to UDP (Classen et al., 1998; Verstynen and Sabes, 2011). However, the effects of learning (i.e., improved performance) on UDP during movement repetition have not been investigated. Here, we used transcranial magnetic stimulation in two experiments to assess plasticity changes occurring in the primary motor cortex after individuals repeated reinforced and nonreinforced actions. The first experiment assessed whether learning a skill task modulates UDP. We found that a group that successfully learned the skill task showed greater UDP than a group that did not accumulate learning, but made comparable repeated actions. The second experiment aimed to understand the role of reinforcement learning in UDP while controlling for reward magnitude and action kinematics. We found that providing subjects with a binary reward without visual feedback of the cursor led to increased UDP effects. Subjects in the group that received comparable reward not associated with their actions maintained the previously induced UDP. Our findings illustrate how reinforcing consistent actions strengthens use-dependent memories and provide insight into operant mechanisms that modulate plastic changes in the motor cortex.

Key words: noninvasive brain stimulation; reinforcement; skill learning; TMS; use-dependent learning

Significance Statement

Performing consistent motor actions induces use-dependent plastic changes in the motor cortex. This plasticity reflects one of the basic forms of human motor learning. Past studies assumed that this form of learning is exclusively affected by repetition of actions. However, here we showed that success-based reinforcement signals could affect the human use-dependent plasticity (UDP) process. Our results indicate that learning augments and interacts with UDP. This effect is important to the understanding of the interplay between the different forms of motor learning and suggests that reinforcement is not only important to learning new behaviors, but can shape our subsequent behavior via its interaction with UDP.

Introduction

Repetitions of simple motor actions lead to plastic reorganizational changes in the primary motor cortex (M1). This phenomenon, known as use-dependent plasticity (UDP), has been

demonstrated by the presence of changes in transcranial magnetic stimulation (TMS)-evoked movement direction after training (Classen et al., 1998; Bütefisch et al., 2000; Celnik et al., 2006; Stefan et al., 2008). Similarly, behavioral investigations have shown that consistent repetition of movements induces directional biases toward the repeated direction (Diedrichsen et al., 2010; Huang et al., 2011; Verstynen and Sabes, 2011). This form of learning has been interpreted as the result of Hebbian changes in the motor cortex (Orban de Xivry et al., 2011). Although UDP seems to decay quickly with time, it has been suggested to be helpful to reduce movement variability (Verstynen and Sabes, 2011). How then can we enhance or prolong this use-dependent memory?

An important clue might be the observation from recent reports showing that repetition of successful movements when dealing with perturbations elicits larger directional biases—the hallmark of UDP—than movement repetitions alone (Diedrich-

Received Oct. 25, 2016; revised Jan. 24, 2017; accepted Jan. 26, 2017.

Author contributions: F.M., S.U., A.B., and P.C. designed research; F.M. performed research; F.M. analyzed data; F.M., S.U., A.B., and P.C. wrote the paper.

This work was supported by the National Institutes of Health/National Institute of Child Health and Human Development Grants R01HD053793 and R01HD073147 and The Rothschild Fellowship of the Yad Hanadiv Foundation and Grants-in-Aid for Scientific Research (JSPS KAKENHI) Grants 25-4917. We thank Adrian M. Haith, Lior Shmuelof, Nicholas Wymbs, and Vikram S. Chib for their helpful comments on the manuscript, and Tziporah Thompson for illustrations in Figure 1.

The authors declare no competing financial interests.

Correspondence should be addressed to Pablo Celnik, Department of Physical Medicine and Rehabilitation, Johns Hopkins School of Medicine, 600 North Wolfe Street, Phipps 160, Baltimore, MD 21287. E-mail: pcelnik@jhmi.edu.

DOI:10.1523/JNEUROSCI.3303-16.2017

Copyright © 2017 the authors 0270-6474/17/372673-13\$15.00/0

sen et al., 2010; Huang et al., 2011; Bernardi et al., 2015). Additionally, evidence from human neuropharmacology studies suggests that administration of levodopa, which increases the presynaptic availability of endogenous dopamine, enhances UDP (Flöel et al., 2005; Floel et al., 2008). These results suggest a hidden process, likely driven by success-related reinforcement signals that might modulate UDP. In contrast to UDP, reinforcement is another form of learning based on the presence of success or failure (Knowlton et al., 1996; Schultz, 2006; Ramayya et al., 2014), has longer-lasting effects (Shmuelof et al., 2012b; Therrien et al., 2016), and is mediated by circuits involving the basal ganglia and motor cortex (Huntley et al., 1992; Ziemann et al., 2001; Luft and Schwarz, 2009; Hosp et al., 2011; Kawai et al., 2015). In this manner, learning new motor behaviors is driven in part by reward-prediction errors that allow the selection of actions based on their likelihood of reinforcement (Sutton and Barto, 1998; Lee et al., 2012). Although the foregoing reports suggested potential interactions between the different forms of learning (Diedrichsen et al., 2010; Abe et al., 2011; Huang et al., 2011; Bernardi et al., 2015), these studies did not completely clarify the effects of learning (i.e., improvement in performance) on UDP. Therefore, the extent to which UDP and learning per se interact both at the behavioral and neural level remains poorly understood.

The aim of the current study was to better understand whether successful goal achievement interacts with and enhances UDP. To do so, we designed two experiments. In Experiment 1, we hypothesized that action repetition while learning a motor task would enhance the magnitude of UDP. Thus, we examined the effect of task success during motor skill learning on UDP. In this experiment, we applied single pulses of TMS over the thumb area on M1 before and immediately after participants trained an isometric pinch-force skill task. We manipulated the amount of learning by controlling trial by trial the mapping between the force transducer and the visual display of the cursor. One group of participants was trained on a consistent map, which allowed performance improvement, whereas a second group of participants was exposed to trial-by-trial randomized maps, which prevented accumulation of learning. In Experiment 2, we manipulated explicit reinforcement signals during repetitions to inform subjects about their performance while controlling the variability of kinematics and magnitude of reward. Here, participants in one group were provided a binary reward feedback associated with the performance, whereas a second group received comparable feedback that did not correspond to their actions. Understanding the interplay between these forms of learning is important for designing better training paradigms that can lead to more efficient and longer-lasting training effects in healthy people and in the context of neurorehabilitation.

Materials and Methods

Subjects

We recruited 52 healthy adult subjects for the study [23.02 ± 4.9 years old (mean \pm SD), including 34 females] who were right-handed and naive to the purpose of the study. Subjects gave written consent to participate in the study, which was approved by the Johns Hopkins University School of Medicine Institutional Review Board. None of the subjects had a history of neurological disease or psychological disorders. All subjects denied the use of acute or chronic CNS-acting medication, said they refrained from drinking alcohol in the previous 24 h, and reported to have had ≥ 5 h of sleep the previous night.

Recording and brain stimulation

Subjects sat comfortably in a chair while their right forearm was restrained in a molded armrest in a semipronated position. Each subject's

four fingers were in a slightly extended position and his or her thumb was left entirely unconstrained. Use of the armrest allowed for consistent positioning of the arm and hand across the different sessions of the experiment. We recorded thumb movements using a highly sensitive three-dimensional accelerometer (Kistler Instruments) positioned on the distal interphalangeal joint of the thumb (Fig. 1*a*). We determined the direction and amplitude of each TMS-evoked thumb movement and calculated the first peak acceleration vector in both the horizontal (abduction/adduction) axis and vertical (flexion/extension) axis, as done in previous studies (Classen et al., 1998; Bütefisch et al., 2000; Galea and Celnik, 2009). We recorded, amplified, and filtered surface electromyography (EMG) activity (Bortec Biomedical) from the right flexor pollicis brevis (FPB) and right extensor pollicis brevis (EPB) muscles (we identified FPB as agonist and EPB as antagonist), both known to be activated during thumb movements.

To elicit focal, isolated, and consistent right thumb movements (Fig. 1*a*), we delivered TMS using a 70 mm figure-eight coil (Magstim) over an optimal scalp position of the left M1. The coil was placed tangentially on the scalp with the handle pointing backward and laterally at a 45° angle away from the midline, approximately perpendicular to the central sulcus. To ensure accurate positioning and consistency of the TMS coil throughout training and test sessions, we used a frameless neuronavigation system (RRID:SCR_009539, BrainSight, Rogue Research) and coregistered the subject's head to a default Talairach template provided by BrainSight software. In this optimal spot, we determined the resting motor threshold, a measure of neuronal excitability, as the minimum TMS intensity that evoked motor-evoked potential (MEPs) of 50 μ V in 5 of 10 trials at rest (Pascual-Leone et al., 1994; Rossini et al., 1994). The intensity of the TMS required to elicit isolated and mild thumb movements in a consistent direction was set above the resting motor threshold (Table 1). We kept the stimulation intensities similar across conditions, and there were no statistical differences between the TMS intensities of all groups ($p > 0.33$). We sampled all signals at 1 kHz. These were displayed on-line for the experimenter and analyzed off-line with custom-built Matlab R2014b program (RRID:SCR_001622, MathWorks).

Experimental procedure

Experiment 1: UDP in skill learning task. The first experiment included 22 participants and was designed to examine the effect of the task's success during motor skill learning on UDP. The behavioral task and stimulation procedure have been previously described (Classen et al., 1998; Bütefisch et al., 2000; Galea and Celnik, 2009; Reis et al., 2009; Cantarero et al., 2013). Briefly, subjects participated in three consecutive sessions: pre-training (baseline), motor training, and post-training. During pre-training, 65 TMS single pulses were delivered at 0.2 Hz to the optimal scalp position that evoked isolated thumb movements. Although subjects might feel some contractions of their thumb muscles following each stimulus, they could not determine the direction of the movements and subjects had no direct view of their hand. Baseline-consistent TMS-evoked movement direction in the abduction direction was identified (Fig. 1*b*). After this, subjects were trained to perform the sequential visual isometric pinch skill task (Reis et al., 2009; Cantarero et al., 2013). To this end, subjects were seated in front of a computer monitor and were instructed to pinch against an isometric force transducer using their right thumb (adduction movement direction) to perform the sequential visual isometric pinch task (SVIPT). In this manner, the direction of pinching force during training was carefully controlled by the experimenter and was set to a direction almost opposite to the baseline direction elicited by TMS. At baseline, the participants displayed TMS-evoked movements in an abduction direction. In this skill task, subjects' applied forces were mapped to lateral on-screen cursor displacement. In each trial, subjects were instructed to navigate the cursor as accurately and quickly as possible between a HOME position and five different targets in ascending order (i.e., HOME-1, HOME-2, HOME-3, HOME-4, HOME-5) by alternating the pinch force exerted onto the transducer, and thus learning the sequence of contractions to move the cursor as fast and accurately as possible (Fig. 1*b*). The order and position of the targets did not change during the experiment, and subjects were able to see the cursor on the screen. The training session lasted ~ 30 min. During post-training, 65

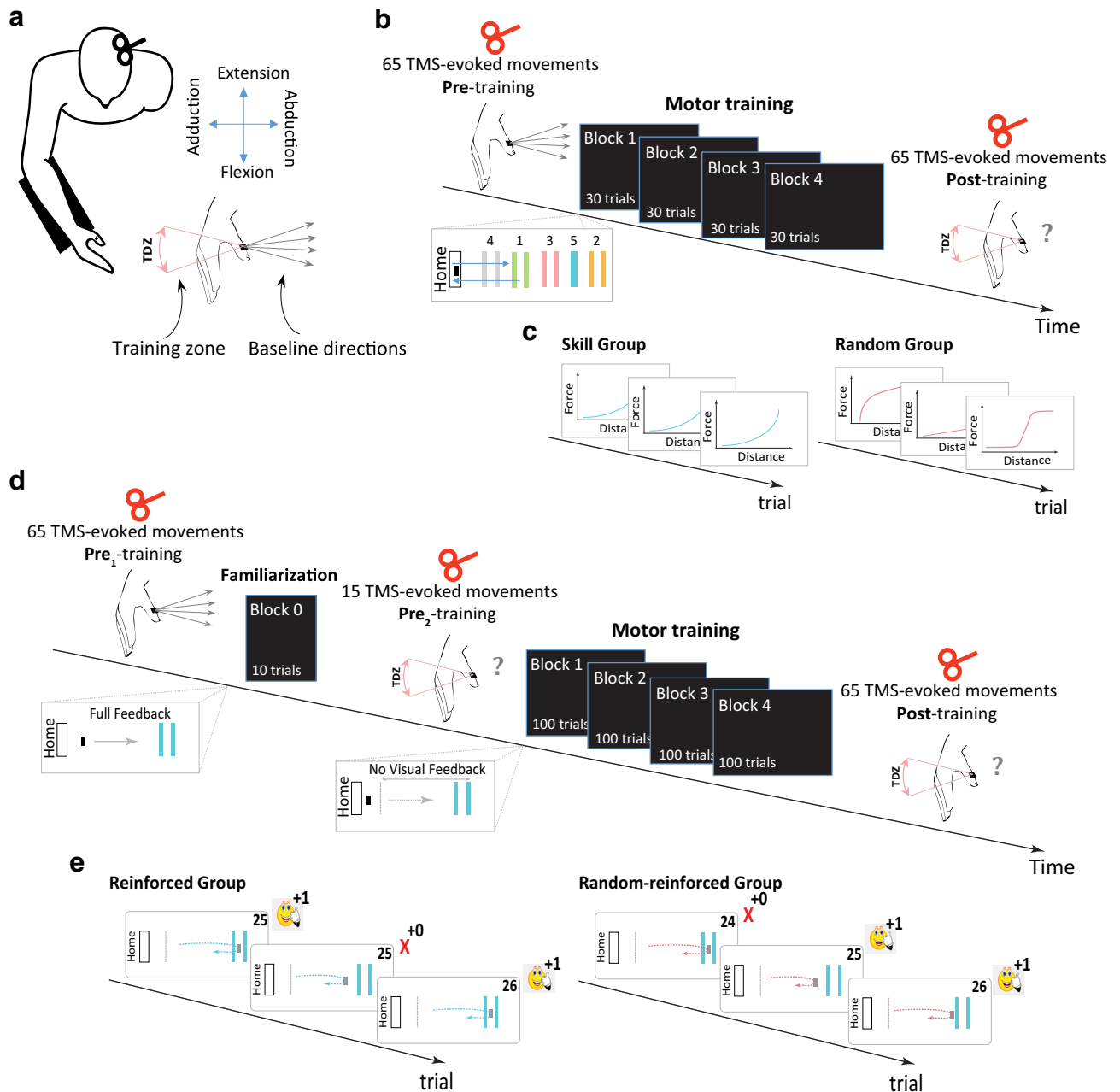


Figure 1. Experimental setups. **a**, We delivered TMS over the left M1 to elicit thumb movements before and after training and measured the percentage of TMS-evoked thumb movements falling in the training direction zone (TDZ). **b**, Protocol of Experiment 1. For each subject, we first delivered 65 TMS single pulses to assess baseline mean direction of the thumb movement (pre-training session). Subjects then performed four blocks of 30 trials of the SVIPT task. We then redelivered 65 TMS-evoked isolated thumb movements as done in the pre-training session. **c**, The Skill group was trained in a consistent map that logarithmically transfers forces to actual cursor displacement (left), whereas the Random group was exposed to trial-by-trial randomized force transducer-visual display mapping (right). **d**, Protocol of Experiment 2. The pre₁-training and post-training procedures were similar to those in Experiment 1. Following the pre₁-training session, subjects performed 10 trials of the single-target SVIPT task with full feedback of the cursor position. We then delivered 15 TMS-evoked thumb movements to ensure similar level of performance and use-dependent plasticity (UDP; pre₂-training session). In the training session, subjects performed four blocks of 100 trials of the single-target SVIPT task with identical transformation as the familiarization session. The feedback of the cursor was switched off and only reinforcement binary visual feedback was provided. **e**, The Reinforced group performed the task and was reinforced by the explicit binary feedback associated with task success (left), whereas the Random-reinforced group was presented with random reinforcement independent of task success (right). The visual feedback was represented either by a smiley face image accompanied with a positive score of one point (+1), indicating target hit, or a red cross accompanied with zero score (0), indicating target missed. Accumulated score across the trial was also presented to the subjects.

TMS single pulses at 0.2 Hz were delivered over the same area as in the pre-training session.

We randomly assigned participants to one of two groups: the Skill group or the Random group. Each group was exposed to a different mapping between the force transducer and the visual display. In the Skill group ($n = 11$; age: mean \pm SD, 23.7 ± 4.4 years; seven females), subjects trained on a consistent map that logarithmically transferred forces to

actual cursor displacement, allowing performance improvement (Fig. 1c, left), as previously described (Reis et al., 2009; Cantarero et al., 2013). In the Random group ($n = 11$; age: mean \pm SD, 22.6 ± 5.3 years; eight females), subjects were exposed to a trial-by-trial randomized force transducer-visual display mapping that markedly reduced the number of successful trials and thus prevented accumulation of learning (Fig. 1c, right; Cantarero et al., 2013). The random maps included different

Table 1. Resting motor threshold, TMS intensity, and cortical excitability measures for the pre-training and pre₁-training sessions (i.e. baseline) in Experiments 1 and 2, respectively

	Skill group Mean ± SEM	Random group Mean ± SEM	Group difference	
			<i>t</i> ₍₂₀₎	<i>P</i> value
Experiment 1				
Motor threshold (%)	44.09 ± 2.74	46.09 ± 2.52	0.26	0.61
TMS intensity (%)	58.63 ± 2.52	60.27 ± 2.51	0.19	0.66
MEP agonist (mV)	1.38 ± 0.28	1.5 ± 0.2	0.12	0.73
MEP antagonist (mV)	1.4 ± 0.18	1.11 ± 0.15	1.32	0.26
	Reinforced group Mean ± SEM	Random-reinforced group Mean ± SEM	Group difference	
			<i>t</i> ₍₂₂₎	<i>p</i> value
Experiment 2				
Motor threshold (%)	41.5 ± 1.55	45.58 ± 2.03	2.33	0.14
TMS intensity (%)	56.25 ± 2.02	59.25 ± 2.03	1.00	0.33
MEP agonist (mV)	1.74 ± 0.28	2.06 ± 0.44	0.33	0.57
MEP antagonist (mV)	1.32 ± 0.19	0.99 ± 0.14	1.69	0.21

shapes of linear, logarithmic, exponential, sigmoid, and double-sigmoid functions in which the cursor movement to the right increases monotonically as a function of the force (Fig. 2*d*). All random maps included transfer functions of consistent direction of force with regard to the cursor movement. We carefully chose those functions to control for mean force across trials. Both groups performed four blocks of 30 trials (120 trials in total). During the training session, the between-trials interval was 2 s and the between-blocks interval was 10 s. All subjects—but not the experimenter—were blinded to their group assignment. Overall, the first experiment examined the effect of learning on UDP. Although subjects in this experiment did not receive explicit reinforcement feedback, it is possible that observing trial success or failure may have served a similar role as an implicit reinforcement (Izawa and Shadmehr, 2011; Krakauer and Mazzoni, 2011).

Experiment 2: UDP in reinforcement learning task. The second experiment included 24 participants and was designed to investigate the effect of learning a reinforcement task on UDP and to control for the magnitude of reward and for the potential kinematic changes associated with the skill performance of Experiment 1. Here, subjects participated in five consecutive sessions: pre₁-training (baseline), task familiarization, pre₂-training, motor training, and post-training. In pre₁-training, we identified the baseline direction of the thumb by applying 65 TMS single pulses at 0.2 Hz to the optimal scalp position (hotspot) that evoked isolated thumb movements. The familiarization session consisted of 10 trials in a form simpler than that of the original skill task described above (i.e., SVIPT). Here, the sensor had a linear transduction of pinch force to cursor movement. In this experiment, we instructed the subjects to pinch against the isometric force transducer to navigate the cursor as accurately as possible between a HOME position and a single target, located 20 cm to the right of the HOME position (Fig. 1*d*). As in Experiment 1, the direction of the pinching force during the training session was carefully controlled by the experimenter and was set to a direction almost opposite to the baseline (i.e., pre₁-training session) direction elicited by TMS. Full feedback of the cursor was provided in this session. To test whether the familiarization session induced UDP changes, we delivered 15 TMS single pulses at 0.2 Hz to the same hotspot (i.e., pre₂-training session) while subjects were at rest. There were two reasons for including the second baseline test (i.e., pre₂-training). First, it allowed us to determine whether very brief exposure to the task induced TMS-evoked thumb movement in the repeated direction. Second, we needed to rule out the possibility that any differences we might observe between the groups after training (during the post-training) was due to differences elicited during the brief familiarization. In the motor training session, subjects performed four blocks of 100 trials of the single-target modified SVIPT task with transformation identical to the familiarization session but with the following changes: once the cursor had moved a quarter of the distance (5 cm), the continuous cursor feedback was switched off and only binary visual feedback was provided for 1.5 s at the end of the trial. Trials were terminated when the force amplitude descended below 0.01 N. The binary visual feedback was

represented either by a smiley face image accompanied by a positive score of one point (+1), indicating target hit, or a red sign accompanied with zero score (0), indicating a target miss. The accumulated score across trials was also presented to the subjects. After the training session, which lasted ~30 min, participants underwent the post-training session, which was identical to the baseline session in which 65 TMS pulses were delivered. During the brain stimulation parts (i.e., pre₁-training, pre₂-training, and post-training), we monitored muscle relaxation by EMG signals.

To dissociate reinforced repeated actions from nonreinforced repeated actions, we divided subjects into two groups. The Reinforced group (*n* = 12; age: mean ± SD, 24.1 ± 5.1 years; eight females) performed the task while receiving the explicit binary feedback-associated task success (Fig. 1*e*, left). The Random-reinforced group (*n* = 12; age: mean ± SD, 22.4 ± 4.8 years; seven females) was presented with random feedback (both positive and negative) independent of task success (Fig. 1*e*, right). To ensure that this group received the same amount of feedback (i.e., magnitude of reward) as the Reinforced group, we provided the same schedule of reinforcement that individual subjects had in the Reinforced group, but presented the feedback randomly within blocks and independently of task success. To maintain subjects' motivation and to reduce frustration, both groups received occasional refresher trials (randomly inserted into the sequence of trials) with continuous cursor feedback (12 of 100 of the trials). In addition, we included a third group with six subjects (Time group; *n* = 6; age: mean ± SD, 21.3 ± 3.8 years; four females) to control whether simple repetitions are important to maintain UDP level. This group underwent the same protocol as the previous two groups, but instead of training in the task after pre₂-training, subjects rested for 30 min. Following the idle 30 min, participants underwent the post-training session, which was identical to the session for previous groups. All subjects were blinded to the group assignment. In our experiments, we chose independent samples of subjects to avoid generalization and after-effects commonly observed when performing similar motor skill tasks that shared the same behavioral structure (Braun et al., 2009, 2010). In both experiments, the range of forces applied to the force transducer during the training session was 0 ≤ *f*(*t*) ≤ 8 N (Fig. 2*e*).

Data analysis

UDP measure. To explore the effect of the learning on UDP, we first defined the training direction zone (TDZ) as a window of ±20° centered on the mean training direction (Classen et al., 1998; Bütefisch et al., 2000; Galea and Celnik, 2009). We then calculated the proportion of the TMS-evoked thumb movements that fell within the TDZ before and after training. We calculated this primary outcome measure for pre-training and post-training sessions in the two experiments. We then calculated the change in the proportion of the TMS-evoked thumb movement that fell within the TDZ (i.e., bias) as the change between post-training and pre-training in Experiment 1 and post-training and pre₁-training in Experiment 2.

Force direction during training. To rule out differences in the quality of training, we calculated the direction of the isometric contractions used to compress the force transducer during the training session by integrating the accelerometer (obtained from a highly sensitive 3-D piezoelectric accelerometer) and EMG data. During each training session, we recorded EMG and accelerometer data continuously for 2.5 min (10 min total for the training session). For analysis, we first filtered the EMG data using a second-order infinite impulse response (IIR) 60 Hz Notch filter to remove the electrical noise. This filtered signal was then low-pass filtered using fifth-order Butterworth filter with 10 Hz cutoff frequency. We then calculated the root mean square of the net EMG signal to detect the peaks of the contractions. Our assumption was that the actual direction of the force applied to compress the force transducer would be best estimated by calculating the direction of the net acceleration at the time point when the contraction reached the peak. Mathematically, the direction of the action, θ_{training} , was calculated in each trial using the following formula:

$$\angle Acc = \theta_{\text{training}} = \tan^{-1} \left(\frac{Acc_y(t_{\text{max}})}{Acc_x(t_{\text{max}})} \right),$$

$$|Acc| = 2 \sqrt{Acc_y(t_{\text{max}})^2 + Acc_x(t_{\text{max}})^2}$$

where $Acc_y(t_{\text{max}})$ represents the *y* component of the acceleration at the peak contraction and $Acc_x(t_{\text{max}})$ represents the *x* component of the ac-

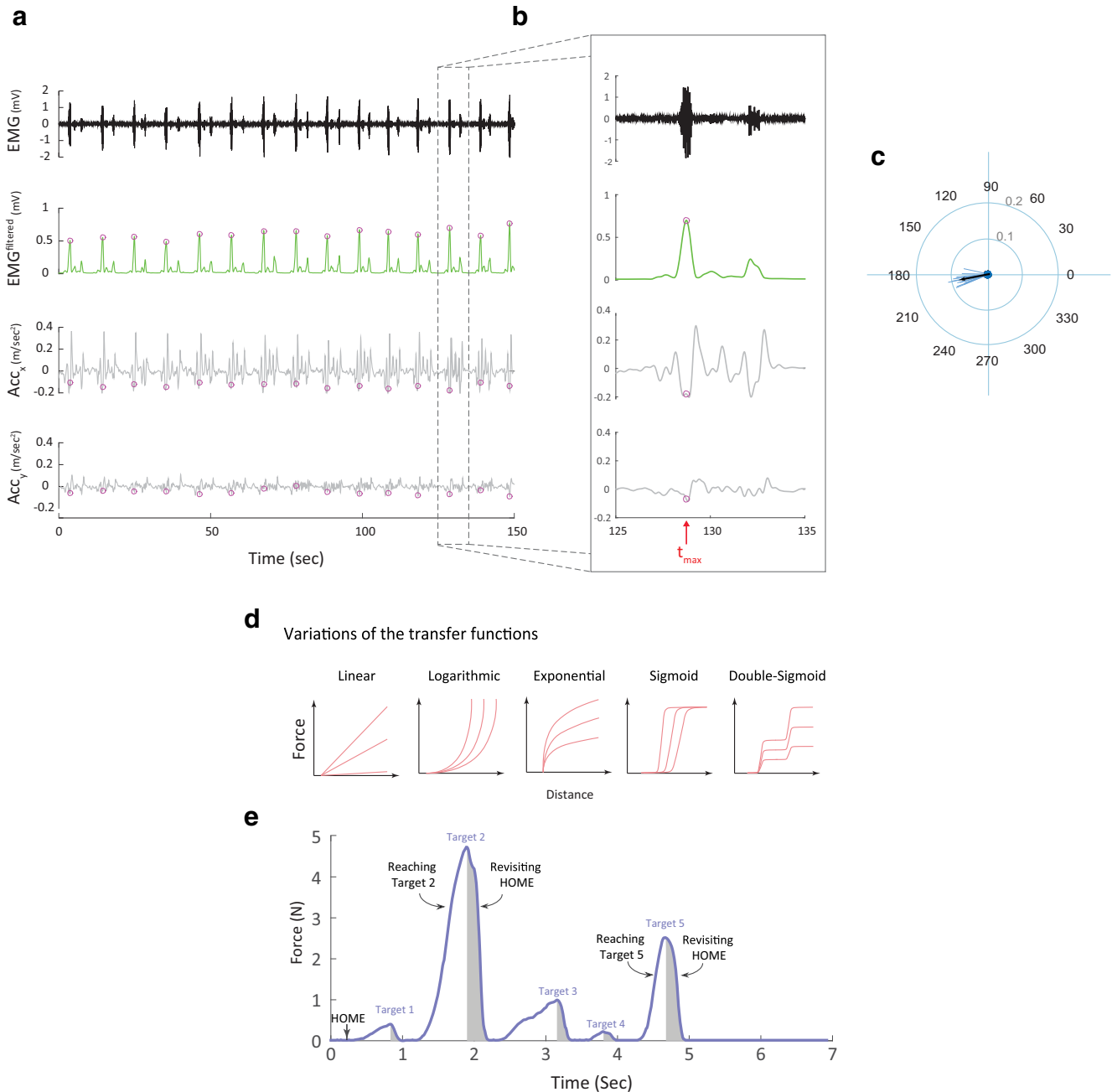


Figure 2. Direction of the actions during the training session, force-cursor transfer functions, and force profile. **a**, Top, Raw EMG data for one subject during one block of the training session. Second from top, Filtered EMG data (green) and the peaks of the contractions (magenta circles). Third from top, Component *x* of the acceleration signal. Bottom, Component *y* of the acceleration signal. **b**, Zoom in for one trial during the training session. t_{max} indicates the time point of maximum contraction at that trial. **c**, Direction of actions in the training. Each blue line represents the direction and the amplitude of a single trial. Black line represents the average of the direction vector across all trials. **d**, Variations of the force-cursor transfer functions for the Random group in Experiment 1. These random maps included different shapes of linear, logarithmic, exponential, sigmoid, and double-sigmoid functions. **e**, Example of force profile in a single trial in the Skill group. Shaded areas illustrate the actual data when participants revisited the HOME position.

celeration at the same time point. $|Acc|$ represents the amplitude of the acceleration. Although the magnitude of the signal was small, due to the nature of the isometric force task, we were able to accurately detect the direction of the actions during the training session (Fig. 2).

Neurophysiological measures. We assessed cortical excitability by measuring the peak-to-peak amplitudes (in mV) of the MEPs for the agonist and antagonist muscles in all stimulation sessions.

Kinematic measures. In addition to measuring the force direction during the training session, we also measured the amplitudes of the first-peak acceleration of the involuntary TMS-evoked thumb movements during the stimulation portions of the experiment and the mean force applied to the force transducer in each trial during the learning session.

Skill learning measure. In Experiment 1, we quantified the amount of skill learning by determining the change in the speed-accuracy trade-off function (SAF). In particular, we used the proposed estimate of this change, α , as follows:

$$\alpha = \frac{1 - \text{error rate}}{\text{error rate} * (\ln(\text{movement time})^b)}$$

where error rate was calculated as the proportion of unsuccessful trials (unsuccessful trial being defined as the trials with ≥ 1 overshooting or undershooting movement), and movement time was calculated as the time interval between movement onset and the time

point when the cursor reached target five. We averaged these two measures over period of 30 consecutive trials. The parameter b was set to be 5.424, a value that was found and confirmed in independent samples of subjects who performed the same skill task (Reis et al., 2009; Schambra et al., 2011; Cantarero et al., 2013; Dayan et al., 2014a,b). Additionally, we calculated the improvement in learning (i.e., learning gain) for each subject as the difference in the skill measure between the last (fourth) and first block. This measure focused on the overall shifts of SAF function, rather than changes in each component independently (i.e., speed or accuracy; Dayan et al., 2014b).

In recent years, skill has been operationally defined as changes in the SAF when performing a task (Karni et al., 1998; Reis et al., 2009; Dayan and Cohen, 2011; Shmuelof et al., 2012a). This means that skill improvement is the result of changes in both speed and accuracy; while the opposite—improving accuracy because performance slows down or moves faster but less accurately—does not represent skill improvement. Thus, to quantify skill learning, it is important first to characterize the SAF of a given task. To do so, we followed the findings of our previously published study (Reis et al., 2009). Briefly, in different datasets, we measured the SAF during the SVIPT task before and after 5 days of training. We then searched for the mathematical model that best fit the behavioral data (accuracy as a function of the movement time). After testing several functions, we found that the two-parameter logistic function was the best model that satisfied our fitting criteria expressed as follows:

$$\text{error rate} = \frac{1}{1 + \alpha * (\ln(\text{movement time}))^b}$$

Where α and b are the dimensionless free parameters, movement time is the average time it took for the subject to complete each trial, and error rate is the number of errors per block. Fitting this model to data showed that while the least-square estimate of the parameter b was almost unchanged, the corresponding estimate of α markedly increased. We therefore tentatively defined α as the skill parameter and fixed the value of b at the average value of 5.424 [the average of pre-training (5.51) and post-training (5.34)] and from each bivariate observation of movement time and error rate, estimated α as shown in Equation 1. This measure was then used and validated in new datasets from different studies (Reis et al., 2009; Schambra et al., 2011; Cantarero et al., 2013; Dayan et al., 2014b), indicating that using the above equation to estimate α with a fixed value of b is acceptably accurate. Our second experiment was not a good example of skill learning because subjects were not requested to maximize their speed and received only binary feedback concerning hitting the target. Thus, this task aimed only to improve accuracy via reinforcement. Therefore, we analyzed and presented the accuracy and movement time separately for this experiment.

Consistency measure. In Experiment 2, we defined consistency as the lag-1 autocorrelation [$R(1)$] of the peak force sequence during the training session. Sequence of the peak forces is defined as the magnitude (i.e., peak) of the force in each trial across the 400 trials (i.e., 4 blocks \times 100 trials). Consistency measure reflects the expected value of the covariance of the peak force between the current trial and the subsequent trial in the same direction of action, normalized by the overall variance of the peak forces, as follows:

$$\text{Consistency} = R(1) = \frac{E[(\text{force}_n^{\text{max}} - \mu_{\text{force}^{\text{max}}})(\text{force}_{n+1}^{\text{max}} - \mu_{\text{force}^{\text{max}}})]}{\text{variance}(\text{force}^{\text{max}})}$$

where $\text{force}_n^{\text{max}}$ represents the peak force in trial n and $\mu_{\text{force}^{\text{max}}}$ represents the mean of peak forces across all trials. $E(x)$ represents the expected value or the mean value in the long run for many repeated samples of x . We performed this analysis to verify that consistency of the repeated actions across groups could not simply account for the differences observed in our main UDP measure.

Variability measure. We defined within-trial variability as the SD of the force signal for each trial during the training session. The force signal

refers to all data points of the force profile in a given trial. The within-trial variability can be then formulated as follows: $\text{Var}_n = \text{SD}(F_n)$, where F_n represents all data points of the force signal in trial n . Between-trial variability was calculated as the SD of the means of forces across trials. This variability can be formulated as follows: $\text{Var}_{[1:n]} = \text{SD}([M_1, M_2, \dots, M_n])$, where M_n represents the mean of the force signal in trial n ($M_n = \bar{F}_n$).

Statistical analysis

We performed the statistical analysis using Matlab software with Statistics Toolbox (MathWorks) and Prism software (RRID:SCR_002798, GraphPad). To determine whether the groups were significantly different at baseline (i.e., pre-training), we used two-tailed t tests between the different groups to assess differences in corticomotor excitability (i.e., motor threshold, TMS stimulus intensity, and MEP for the agonist and antagonist muscles) and kinematics (i.e., peak acceleration of the thumb movements). We then used separate repeated-measures ANOVA (ANOVA_{RM}) to assess differences in task performance and in the proportion of movements falling in TDZ (main outcome measure of the study) with factors block (Pre-training and Post-training in Experiment 1; Pre₁-training, Pre₂-training, and Post-training in Experiment 2) and group (Skill and Random in Experiment 1; Reinforced and Random-reinforced in Experiment 2). When significant differences were identified, *post hoc* analysis was conducted using the Holm–Sidak t test for multiple comparison. In all comparisons, significance level was set at 0.05.

Results

Experiment 1

Both Skill and Random groups, before training, showed small but comparable TMS-evoked movements within the TDZ (Fig. 3*a–e*) and similar cortical excitability and kinematic measures (Table 1). During training, the Skill group experienced significantly greater improvement in the SVIPT performance compared with the Random group (ANOVA_{RM} , group effect (Skill, Random): $F_{(1,20)} = 20.89, p = 0.002$; time (blocks) \times group (Skill, Random) interaction effect: $F_{(3,60)} = 3.129, p = 0.0322$; Fig. 3*f*). Interestingly, after training, we found a significant difference in the amount of UDP between groups (ANOVA_{RM} , group effect (Skill, Random): $F_{(1,20)} = 6.916, p = 0.0161$, and time (Pre-training, Post-training) \times group (Skill, Random) interaction effect: $F_{(1,20)} = 5.021, p = 0.03$; Figure 3*e*). Specifically, the Skill group showed a significant increase in the proportion of TMS-evoked thumb movement falling in the TDZ relative to the Random group (*post hoc t* test, $p = 0.0069$; Fig. 3*a,c,e*). This suggests that UDP develops when repetitions are associated with skill learning, which contains implicit success-based reinforcement. Meanwhile, there was no change in movement direction bias when the practice did not lead to learning (Random group; Fig. 3*b,d,e*).

The change on UDP could not be explained by differences in force directions or in force magnitudes applied during the training (Fig. 3*g,h*) given that these kinematic measures were comparable across the two groups. We found no significant group (Skill, Random) differences in the training direction ($F_{(1,20)} = 1.03, p = 0.32$; Fig. 3*g*) or in the amount of total forces applied during the training ($F_{(1,20)} = 1.33, p = 0.26$; Fig. 3*h*). It should be noted, however, that not only was the Skill group more accurate than the Random group (group effect: $F_{(1,20)} = 10.39, p = 0.0043$; Fig. 4*a,c*), but it also performed the task faster with shorter movement times relative to the Random group (group effect: $F_{(1,20)} = 103.3, p < 0.0001$; Fig. 4*b,d*).

ANOVA_{RM} on the mean peak acceleration of the TMS-evoked movements before (mean \pm SE: Skill, $0.47 \pm 0.10 \text{ m/s}^2$; Random, $0.40 \pm 0.06 \text{ m/s}^2$) and after (mean \pm SE: Skill, $0.82 \pm 0.17 \text{ m/s}^2$;

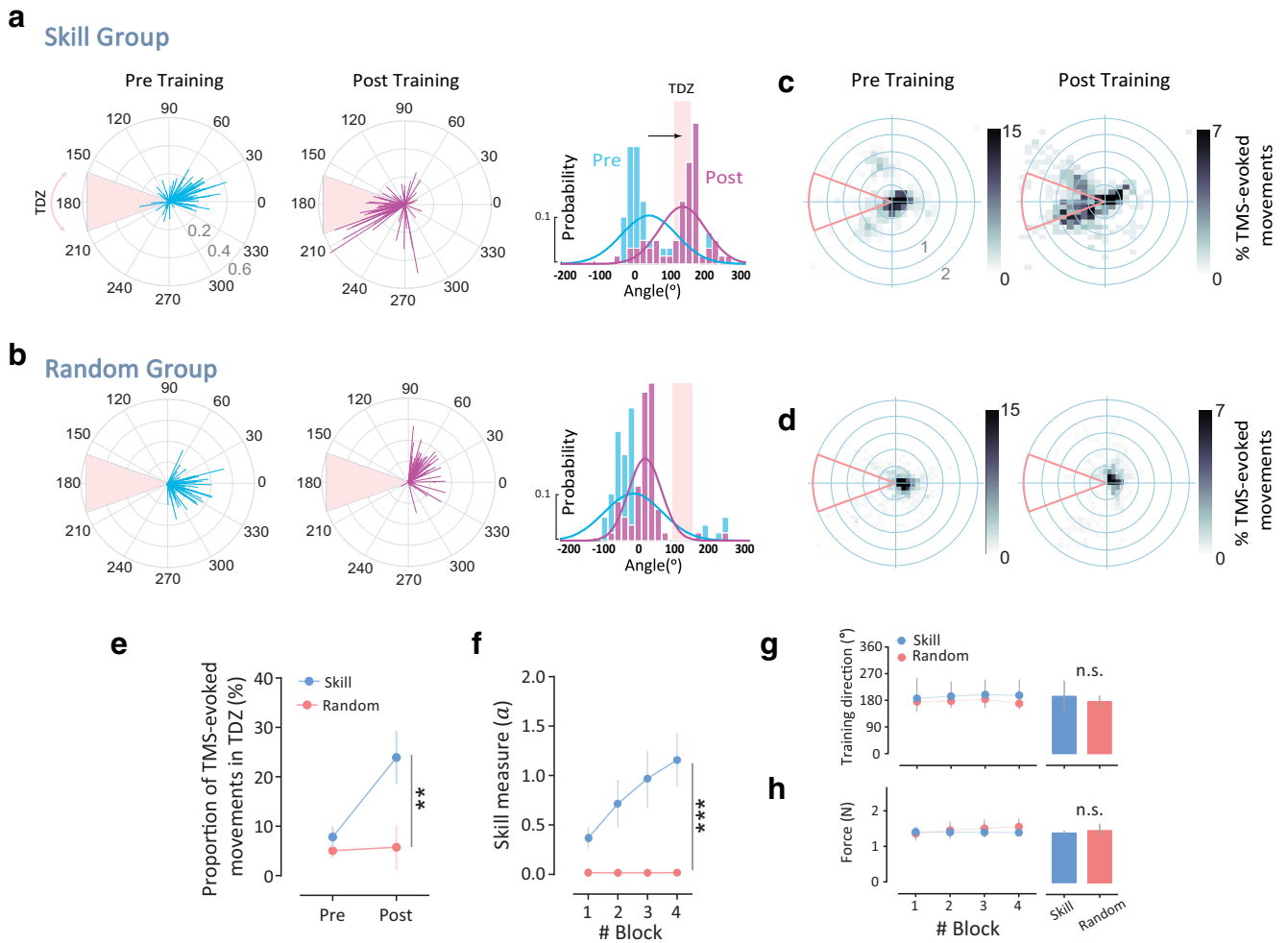


Figure 3. Skill learning and implicit success-based reinforcement enhances UDP. **a**, TMS-evoked thumb movements before (cyan, left) and after (magenta, middle) training for a single subject in the Skill group. Each line represents the direction ($\theta \in [0, 360^\circ]$) and the amplitude ($|Acc| \in [0, 0.6 \text{ m/s}^2]$) of a single thumb movement evoked by a single TMS pulse. Probability distribution of the thumb directions before (cyan lines) and after (magenta lines) performing the skill task (right). Pink shaded area represents the TDZ. **b**, TMS-evoked thumb movements before (cyan, left) and after (magenta, middle) training for a single subject in the Random group. Probability distribution of the thumb directions before (cyan lines) and after (magenta lines) performing the random skill task (right). Probability distribution of the thumb directions before (cyan lines) and after (magenta lines) performing the skill task (right). Note how only the reinforced subject experienced a shift in the TMS-evoked movement directions after the skill training. **c**, Group data: 2-D histogram of the TMS-evoked thumb movement for all subjects in the Skill group before (left) and after the training (right) sessions. Although baseline directions predominantly fell out from the TDZ, a large number of TMS-evoked movements during post-training fell within the TDZ. **d**, Group data: 2-D histogram of the TMS-evoked thumb movement for all subjects in the Random group during pre-training (left) and post-training (right) sessions. Thumb directions after the skill task were not different from the baseline directions; a large percentage of movements fell outside the TDZ. **e**, Proportion of TMS-evoked movements in TDZ (in %) for the two groups (blue and red for the Skill and Random group, respectively). **f**, Performance of the skill task as quantified by the skill measure (α). **g**, To rule out the possibility of different force directions during the training block across groups, we calculated the direction of the isometric forces. This figure shows that the mean direction in each block was similar for each group (bar graphs on the right represent the means of all blocks). **h**, Mean forces applied in the task session. The amount of total forces applied was not different across groups (bar graphs on the right represent the means of all blocks). This indicates that the differences in the TMS-evoked movement direction change (bias) after the training period were not a result of differences in training directions or the amount of force applied during the training. Points in **e–h** represent mean. Error bars represent the SEM. Asterisks indicate significance (** $p < 0.01$; *** $p < 0.001$) and n.s. indicates statistically not significant.

Random, $0.54 \pm 0.16 \text{ m/s}^2$) the training session for the Skill and the Random groups revealed a significant time effect ($F_{(1,20)} = 8.28$, $p = 0.0039$), but no significant difference between the groups ($F_{(1,20)} = 1.208$, $p = 0.28$) and no significant interaction ($F_{(1,20)} = 1.48$, $p = 0.24$). ANOVA_{RM} on the mean MEP for the agonist muscle before (mean \pm SE: Skill, $1.38 \pm 0.29 \text{ mV}$; Random, $1.50 \pm 0.21 \text{ mV}$) and after training (mean \pm SE: Skill, $1.44 \pm 0.19 \text{ mV}$; Random, $1.17 \pm 0.17 \text{ mV}$) for the Skill and the Random groups revealed a significant time (Pre-training, Post-training) \times group (Skill, Random) interaction ($F_{(1,20)} = 5.73$, $p = 0.026$), suggesting that while the Skill group showed increased MEP agonist, the Random group tended to reduce the MEP agonist after training. Nevertheless, there was no significant difference when testing the group or the time factor separately

($p > 0.34$). Data showed that no statistical change ($p = 0.1$) occurred in the MEP antagonist muscle.

Previous UDP investigations described the importance of consistent UDP during the training (Classen et al., 1998; Bütefisch et al., 2000). In our experimental setup, inspection of the behavior during the training session revealed that within-trial variability was higher for the Skill group (group effect: $F_{(1,20)} = 55.26$, $p < 0.0001$; Fig. 4e). This greater variability could be interpreted as additional supporting evidence for successful learning, because to achieve successful task performance, subjects were required to produce different forces to reach different targets within each trial in accordance with the nonlinear logarithmic transformation. Additionally, reduction of between-trial variability would also support evidence of improvement in

task performance. Indeed, between-trial variability was significantly lower in the Skill group (group effect: $F_{(1,20)} = 618.2$, $p < 0.0001$; Fig. 4f). Thus, these results demonstrate that successful repetition of actions, with improvement of consistency across trials during skill learning enhanced UDP. One possible explanation for this result is that learning the skill itself engaged success-related implicit reinforcement and that these reinforcement signals modulated UDP. Nevertheless, to control for potential kinematic changes and magnitude of reward that might be associated with skill performance, we performed a second experiment that tested the effect of learning a reinforcement task during action repetition on UDP.

Experiment 2

The second experiment investigated the effect of learning a reinforcement task during simple action repetition on UDP while controlling the potential kinematic and reward magnitude confounds across groups. Here, in addition to force magnitude and direction, we carefully matched force variability (between-trial and within-trial), consistency, and the magnitude of the reward. Based on current theories of UDP, we predicted that the two groups should show similar magnitudes of UDP since they repeated the same actions, in the same context, and using the same end-effector. However, if reinforcement learning affects UDP, then the Reinforced group should express larger changes in UDP than the non-reinforced participants.

We found that reinforcement learning has a critical role in modulating UDP (ANOVA_{RM}: time effect: $F_{(2,44)} = 17.35$, $p < 0.0001$; time (Pre₁-training, Pre₂-training, Post-training) \times group (Reinforced, Random-reinforced) interaction effect: $F_{(2,44)} = 5.362$, $p = 0.0082$; Fig. 5c). Specifically, the Reinforced group significantly benefited from the appropriate correspondence between cues and actions, showing an ~ 3 -fold change in TMS-evoked thumb movement direction (*post hoc t* test, $p < 0.0001$), while the Random-reinforced group did not increase the movement direction bias after familiarization (*post hoc t* test, $p = 0.99$; Fig. 5a–c). Given that our main analysis compared the proportion of TMS-evoked thumb movements falling in the TDZ at three different timings (pre₁-training, pre₂-training, and post-training) and the pre₂-training test has only 15 TMS pulses (as opposed to 65 pulses in pre₁-training and post-training tests), we also performed an ANOVA_{RM} on post-training and pre₁-training tests. Consistently with the effect reported in the previous result, we found a significant time (Pre₁-training, Post-training) \times group (Reinforced, Random-reinforced) interaction ($F_{(1,2)} = 6.722$, $p = 0.0166$) and time effect ($F_{(1,2)} = 35.91$, $p < 0.0001$). *Post hoc* analysis (between post-training and pre₁-training) showed that the Reinforced group significantly benefited from reinforcement more than the Random-reinforced group (two-tailed *t* test, $p = 0.0129$).

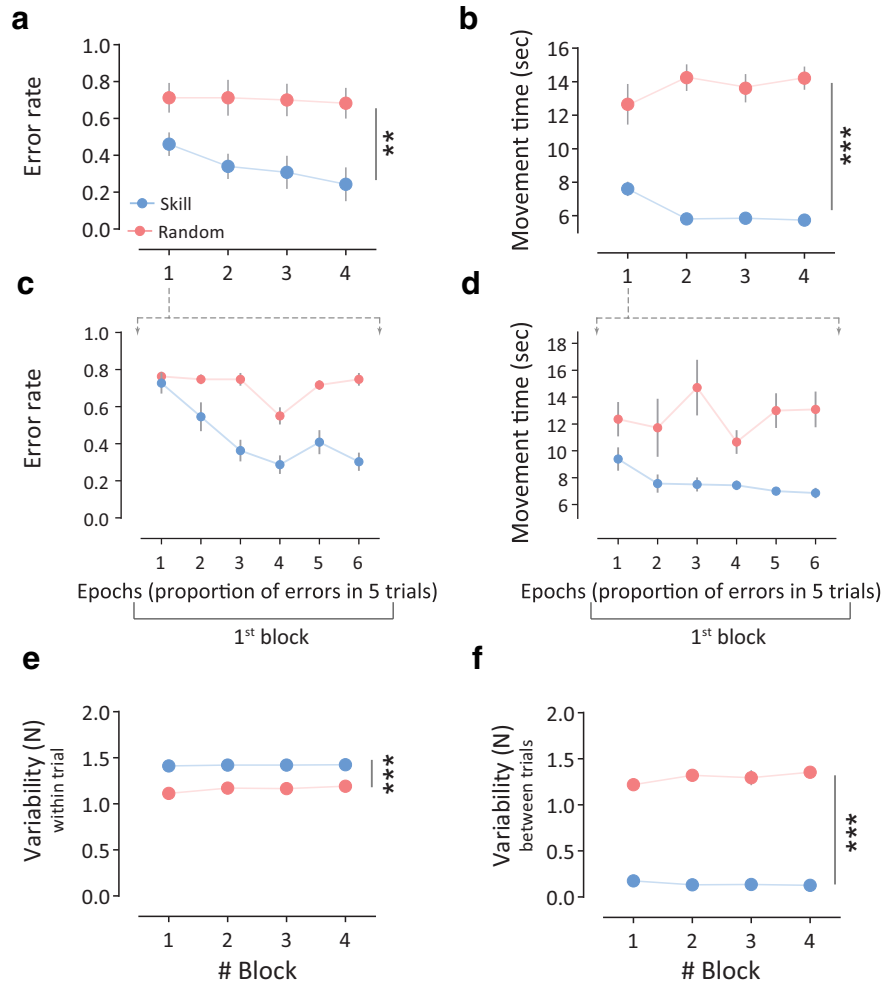


Figure 4. Motor performance during the training session for Experiment 1. *a*, Mean error rate for the two groups. Note that subjects in the Skill group were significantly more accurate during the training (i.e., low error rate) compared with the Random group. *b*, Mean movement time for the different groups. Skill group showed faster performance compared with the Random group. *c*, Mean error rates are shown during the first block of the training session. *d*, Movement time (in seconds) is shown during the first block of the training session. Points in *c* and *d* represent mean of five trial epochs (6 epochs in total for the first block). *e*, Within-trial variability. Here, although subjects in the Skill group performed more accurate and quicker movements across trials, they exhibited greater within-trial variability. *f*, Between-trial variability. Points represent mean, error bars represent \pm SEM, and asterisks indicate significance (** $p < 0.01$, *** $p < 0.001$).

The lack of UDP change between post-training and pre₂-training tests in the Random-reinforced group suggests that repetition of actions alone maintains the previous level of UDP. Indeed, if participants do not execute any movements after developing UDP, the directional bias returns to baseline levels. This was demonstrated in an additional group (Time group) that performed the same protocol as the previous two groups, but after pre₂-training subjects had rested for 30 min. Here, the significant time (Pre₁-training, Pre₂-training, Post-training) \times group (Reinforced, Random-reinforced, Time) interaction ($F_{(4,54)} = 5.396$, $p = 0.001$; Fig. 5c) indicates that repetition alone is an important factor in preservation of previous achieved UDP level. The familiarization block induced a small but significant increase in movement direction bias. ANOVA_{RM} with factor time (Pre₁-training, Pre₂-training) and group (Reinforced, Random-reinforced, Time) revealed a significant time effect ($F_{(1,27)} = 9.199$, $p = 0.0053$), but no main effect of group ($F_{(2,27)} = 0.2332$, $p = 0.7935$) nor interaction ($F_{(2,27)} = 0.2007$, $p = 0.8193$).

ANOVA_{RM} on the mean peak acceleration of the TMS-evoked movements before (mean \pm SE: Reinforced, 0.54 ± 0.06 m/s²;

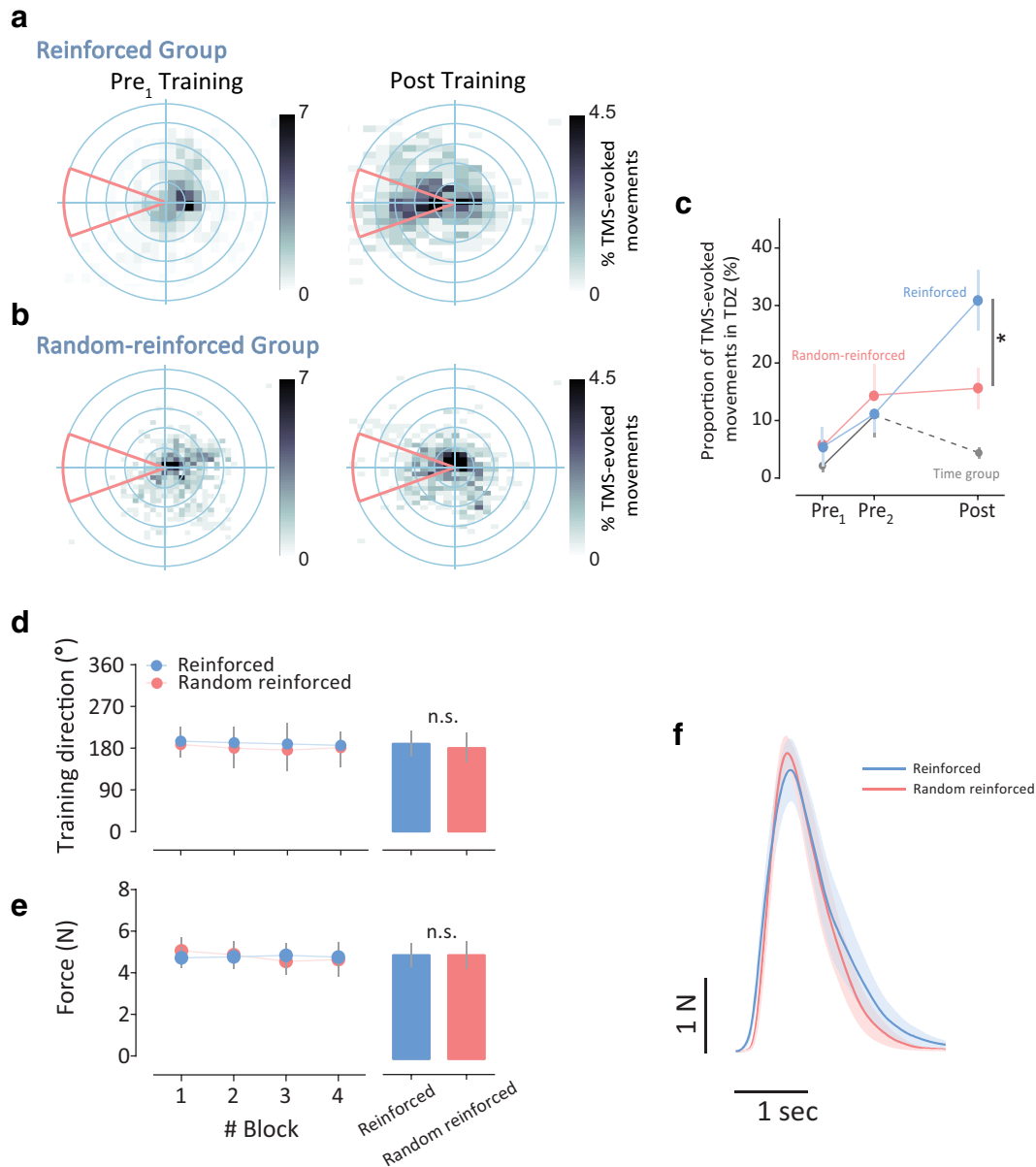


Figure 5. Reinforcement learning modulates UDP. *a*, 2-D histogram of the TMS-evoked thumb movement for all subjects in the Reinforced group pre-training (left) and post-training (right) sessions. The number of TMS-evoked movements fell in the TDZ significantly increased following reinforcement learning. *b*, 2-D histogram of the TMS-evoked thumb movement for all subjects in the Random-reinforcement group before (left) and after (right) the training sessions. *c*, Bias to the TDZ (in %) in the three sessions of the experiment (i.e., pre₁-training, pre₂-training, and post-training) for the two groups (blue and red for the Reinforced and Random-reinforced group, respectively). Gray trace depicts the behavior of additional subjects (Time group) that did not perform the training session and rested for 30 min between pre₂-training and post-training sessions. Points represent mean and error bars indicating the SEM. *d*, Similarly to Experiment 1, to rule out differences in force directions during the training across groups, we calculated the direction of the isometric forces. This figure shows that the mean direction in each block for each group was not different across groups (bar graphs represent the means of all blocks). *e*, Mean forces applied in the task session. The amount of total forces applied were not different across groups (bar graphs represent the means of all blocks). *f*, Actual force profile in the training session for each group. Curves represent mean forces and the shaded area represents the SEM. Asterisks indicate significance (* $p < 0.05$) and n.s. indicates statistically not significant.

Random-reinforced, $0.55 \pm 0.08 \text{ m/s}^2$) and after (mean \pm SE: Reinforced, $0.77 \pm 0.09 \text{ m/s}^2$; Random-reinforced, $0.62 \pm 0.08 \text{ m/s}^2$) the training session for the Reinforced group and the Random-reinforced group revealed a significant time effect ($F_{(1,22)} = 4.363, p = 0.048$), but no significant difference between the groups ($F_{(1,22)} = 0.507, p = 0.48$), and no significant interaction ($F_{(1,22)} = 1.17, p = 0.29$). ANOVA_{RM} on the mean MEP for the agonist muscle before (mean \pm SE: Reinforced, $1.74 \pm 0.30 \text{ mV}$; Random-reinforced, $2.06 \pm 0.45 \text{ mV}$) and after (mean \pm SE: Reinforced, $1.69 \pm 0.39 \text{ mV}$; Random-reinforced, $1.75 \pm 0.29 \text{ mV}$) training for the Reinforced group and the Random-

reinforced group revealed no significant results ($p = 0.4$). Data also showed that no change ($p = 0.11$) occurred in the MEP antagonist muscle in both groups (see Discussion).

The UDP changes observed were not due to differences across groups in baseline corticomotor excitability or TMS intensities (i.e., during pre-training and pre₁-training sessions in Experiments 1 and 2, respectively; Table 1). Most importantly, comparison of various kinematic measures during task training show no significant group effect (Reinforced, Random-reinforced) on force directions ($F_{(1,22)} = 0.59, p = 0.45$; Fig. 5*d*), total forces ($F_{(1,22)} = 6.4 \times 10^{-4}, p = 0.99$; Fig. 5*e,f*), within-trial force

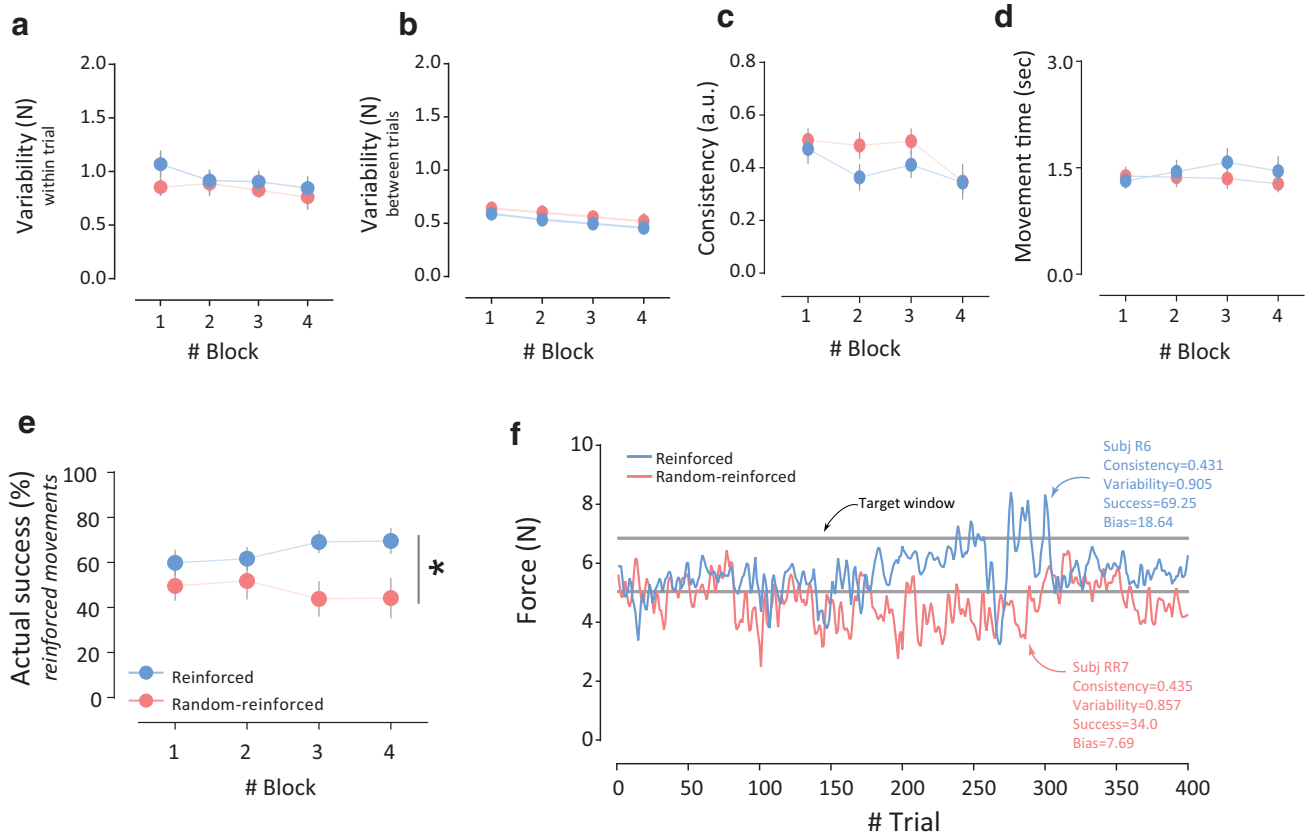


Figure 6. Variability and consistency of the forces cannot explain the biases observed in the Reinforced and Random-reinforced groups. **a**, Within-trial variability of the forces was calculated as the SD of the force sequence across blocks. We found no across-group differences in the variability of the peak forces. **b**, Between-trial variability of the forces. Again, no group effect was found in this kinematic measure. **c**, Consistency of the forces for each group. Peak force consistency across trials was comparable across groups. **d**, Mean movement time for the different groups. No significant difference was found between the groups. **e**, Mean actual success (i.e., accuracy) for each group across the training blocks. Actual success (in %) was calculated as the number of successful trials (e.g., when the unseen cursor hit the target) during the training session. Although actual target accuracy was unbeknownst to all participants, the reinforcement manipulation led to better accuracy in the Reinforced group. **f**, Sequence of the peak forces from two representative subjects from the Reinforced (in blue, subject R6) and the Random-reinforced (in red, subject RR7) groups. Points represent mean, error bars represent \pm SEM, and asterisks indicate significance ($*p < 0.05$).

variability ($F_{(1,22)} = 0.6552, p = 0.4296$; Fig. 6*a*), between-trial force variability ($F_{(1,22)} = 0.852, p = 0.3658$; Fig. 6*b*), force consistency ($F_{(1,22)} = 1.696, p = 0.2063$; Fig. 6*c*), and movement time ($F_{(1,22)} = 0.253, p = 0.6202$; Fig. 6*d*). In addition, the feedback schedule (magnitude of reward) was by design comparable between the two main groups; however, only the Reinforced group experienced correspondence between the executed action and the visual feedback. This, although unbeknownst to the subjects, led to significantly increased target accuracy in the Reinforced group only (i.e., higher actual success; Fig. 6*e,f*). Together, our findings provide evidence that correct association between reinforcement and executed movement modulates UDP, whereas repetition of actions alone maintains the UDP level.

Discussion

In this study, we investigated whether action repetition while learning a motor skill task could modulate UDP. In the first experiment, we found that a group that successfully learned the skill task showed greater UDP than a group that did not accumulate learning, but made repeated actions with comparable magnitude and direction. This might be due to the interaction between movement repetition and improvement in performance (i.e., learning). Although the direction and mean force application was consistent across all aspects of the experiment, the magnitude of the forces applied varied across targets. In addition, the variability

between trials and the magnitude of success-related reward varied across groups. Thus, the task of Experiment 1 does not distinguish between action repetition and alternative potential drivers of cortical plasticity, such as the altered sensorimotor mapping, between-trial variability, and the magnitude of reward. To address these concerns, we designed a second experiment that aimed at testing the reinforcement learning effect on UDP while controlling for reward magnitude and kinematics. We found that providing subjects with a binary reward without visual feedback of the cursor led to increased UDP effects. Subjects that received comparable magnitude of reward which was not associated with their performance maintained the previously induced UDP. Subjects that remained idle for 30 min (the same duration of training) experienced a reduction of the previously encoded UDP. Altogether, these results provide evidence that simple repetition can alter TMS markers of cortical plasticity and that this effect is augmented by reinforcement coupled with successful goal achievement during skill learning or a purer reinforcement task.

Previous investigations have shown that performing repetitive movements can bias the direction of future movements; however, those investigations failed to explain why at times these biases were not present or were of small magnitude (Classen et al., 1998; Bütefisch et al., 2000). These discrepancies could have arisen

because these investigations focused on training consistent movements without controlling for potential success-related reinforcement signals that might have been hidden behind the goal of the task. The finding that repeating actions alone is not sufficient to modulate UDP is consistent with prior studies showing a large degree of UDP when the tasks being trained also involved either goal achievement (Diedrichsen et al., 2010) or successful error reduction (Huang et al., 2011). In addition, past physiological studies describing UDP used verbal reinforcement cues given by the investigator during training to ensure consistency and quality of the repeated directions, albeit this feedback was not consistent, monitored, or carefully controlled (Classen et al., 1998; Celnik et al., 2006). Thus, it is possible that in these studies the verbal cues intended to ensure consistent movement repetition may have served roles similar to success-related reinforcement signals that led to enhancement of the bias toward the repeated action. Recently, Selvanayagam et al. (2011) showed that the direction of TMS-induced twitch force vectors shifted toward the training direction immediately after single isometric strength training with various durations of contraction and rate-of-force development characteristics. Selvanayagam et al. (2016) also showed that force amplitude modulates the extent of directional biases in voluntary aiming. These studies might also speak to the possible involvement of the subcortical dopaminergic system in the modulation of M1 connections—possibly by reinforcing synaptic connections between directionally sensitive neurons active during repeated movement and their downstream targets.

Evidence from animal lesions and imaging work suggests a close association between success-related reinforcement learning, dopaminergic neurons (Hollerman and Schultz, 1998; Waelti et al., 2001; Wise, 2004; Daw et al., 2005), and M1 plasticity (Classen et al., 1998; Bütefisch et al., 2000; Hosp et al., 2011). Dopaminergic terminals in M1 originating in the ventral tegmental area contribute to M1 plasticity (Luft and Schwarz, 2009) and are necessary for skill learning (Hosp et al., 2011; Kawai et al., 2015). Human aging studies that have shown a decrease in dopaminergic function have found reduced ability to elicit UDP changes. The same studies, however, have found that dopaminergic medication in older adults leads to an increase in UDP, an effect correlated with the amount of dopamine released in the caudate region of the basal ganglia (Floel et al., 2008). Although UDP has been interpreted as the result of Hebbian changes in the motor cortex (Orban de Xivry et al., 2011; Huang et al., 2011; Verstynen and Sabes, 2011), this form of plasticity seems to be sensitive to inputs from the basal ganglia. This is supported by the presence of dopamine receptors on cells in M1 (Huntley et al., 1992; Ziemann et al., 2001; Luft and Schwarz, 2009), the fact that dopaminergic medication leads to increase of UDP (Floel et al., 2008), and the finding of deficient motor cortex plasticity in Parkinson's disease patients under off-medication conditions (Morgante et al., 2006).

It is important to note that the repetitions of actions in the Random group of Experiment 1 did not lead to UDP. Three possible explanations can account for this observation. First, it is possible that the larger between-trial variability in this group (note that mean force was matched across conditions but not variance) reduced the ability to establish persistent changes in the cortical representation of the thumb, leading to limited encoding of kinematic details of the practiced actions. As expected, each of the other three groups experienced very low trial-to-trial variability during the training session. This observation supports the view that reduced between-trial variability may be a prerequisite

for UDP. Second, the amount of practice might be an important factor inducing the UDP effect, especially in the absence of any reinforcement or requirement for maximal force exertion. Some previous studies using the TMS-thumb paradigm, showed that UDP needs 30 min, at a rate of 1 Hz, to develop (Classen et al., 1998; Bütefisch et al., 2000). This intensive repetition ends up with >1000 movements. The time needed to develop UDP, however, could be significantly reduced when providing levodopa (≤ 20 min at 1 Hz of repetition). Our training sessions involved fewer practice trials. Thus, the lack of UDP effect in the Random group was perhaps because this group simply did not reach the minimum time or number of repetitions needed. Third, errors during the random skill practice not only reduced the amount of learning but also could have acted as negative signals weakening cortical representational changes that lead to UDP. Future studies are needed to determine the potential effect of these factors on the modulation of UDP.

The UDP paradigm used in our experiments differs from that studied previously (Classen et al., 1998; Bütefisch et al., 2000). We chose to use repetitions of isometric contraction instead of the commonly used finger abduction. This allowed us to assess the presence of UDP in the context of learning a motor skill (Reis et al., 2009), in which we could control the magnitude, timing, variability, and direction of the applied force. Importantly, we found that, besides force, other kinematic details, such as movement direction, acceleration, and movement time were consistent during the training across groups, making them unlikely confounders of our results. Specifically, in the first experiment, to learn the skill task within a trial subjects had to perform a series of isometric contractions of different magnitude. This is the same situation in the Random group, where within trials subjects had to execute a series of contractions of different magnitude, where the distance monotonically increased as a function of force; however, between trials the maps changed randomly. Similarly, in the second experiment, the maps experienced were the same between groups, but only the association between reinforcement signals and actions differed across groups. Thus, it is unlikely that differences in kinematic variables during the performance of Experiment 2 played a role in our results.

We found small effects of training on corticomotor excitability. There are two reasons that could potentially account for this. First, the amount of practice could have influenced cortical excitability. While most of previous studies used numerous contractions of the muscles agonist to the training direction (e.g., 30 min of voluntary brisk thumb movements at a rate of 1 Hz; >1000 movements), our training sessions involved fewer practice trials, with 120 and 400 trials in the first and second experiments, respectively. Second, the type of muscle contraction may have a role in corticomotor excitability changes. In previous studies assessing UDP with TMS, the voluntary movements were executed by applying isotonic movements, which is known to generate force by changing the length of the muscle (Classen et al., 1998; Bütefisch et al., 2000). In our study, however, we used an isometric contraction task, which requires no joint motion. Thus, it is possible that change in muscle length during training may represent a factor that affects change in excitability. Indeed, this is consistent with a previous study that showed larger MEP increases with isotonic contractions when compared with isometric movements (Hayashi et al., 2006).

Our results cannot be attributed to differences in motivation or frustration (Hikosaka et al., 2013; Wang et al., 2013; Wong et al., 2015). If this was the case, we would expect the Random groups from Experiments 1 and 2 to gradually make smaller

and/or more variable and inconsistent actions. However, the direction and the magnitude of the forces did not change during training and were similar to those of the Skill and Reinforced groups. In addition, subjects in the second experiment were not aware of the nature of the reinforcement regimen and received only binary feedback at the end of each trial. Given the similar kinematics across groups, it is unlikely that the effort of participants to optimize performance and attention devoted to the training played a role in our results.

In summary, we found that action repetition while learning a motor task enhances UDP. This finding is important for understanding the interplay between the different forms of motor learning. It also suggests that improvement in skill performance might be in part mediated by UDP. Our results can be used to design better training paradigms for improving the efficiency of motor training in healthy people and in the context of neurorehabilitation following brain damage.

References

- Abe M, Schambra H, Wassermann EM, Luckenbaugh D, Schweighofer N, Cohen LG (2011) Reward improves long-term retention of a motor memory through induction of offline memory gains. *Curr Biol* 21:557–562. [CrossRef Medline](#)
- Bernardi NF, Darainy M, Ostry DJ (2015) Somatosensory contribution to the initial stages of human motor learning. *J Neurosci* 35:14316–14326. [CrossRef Medline](#)
- Braun DA, Aertsen A, Wolpert DM, Mehring C (2009) Learning optimal adaptation strategies in unpredictable motor tasks. *J Neurosci* 29:6472–6478. [CrossRef Medline](#)
- Braun DA, Mehring C, Wolpert DM (2010) Structure learning in action. *Behav Brain Res* 206:157–165. [CrossRef Medline](#)
- Bütefisch CM, Davis BC, Wise SP, Sawaki L, Kopylev L, Classen J, Cohen LG (2000) Mechanisms of use-dependent plasticity in the human motor cortex. *Proc Natl Acad Sci U S A* 97:3661–3665. [CrossRef Medline](#)
- Cantarero G, Tang B, O'Malley R, Salas R, Celnik P (2013) Motor learning interference is proportional to occlusion of LTP-like plasticity. *J Neurosci* 33:4634–4641. [CrossRef Medline](#)
- Celnik P, Stefan K, Hummel F, Duque J, Classen J, Cohen LG (2006) Encoding a motor memory in the older adult by action observation. *Neuroimage* 29:677–684. [CrossRef Medline](#)
- Classen J, Liepert J, Wise SP, Hallett M, Cohen LG (1998) Rapid plasticity of human cortical movement representation induced by practice. *J Neurophysiol* 79:1117–1123. [Medline](#)
- Daw ND, Niv Y, Dayan P (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci* 8:1704–1711. [CrossRef Medline](#)
- Dayan E, Cohen LG (2011) Neuroplasticity subserving motor skill learning. *Neuron* 72:443–454. [CrossRef Medline](#)
- Dayan E, Averbach BB, Richmond BJ, Cohen LG (2014a) Stochastic reinforcement benefits skill acquisition. *Learn Mem* 21:140–142. [CrossRef Medline](#)
- Dayan E, Hamann JM, Averbach BB, Cohen LG (2014b) Brain structural substrates of reward dependence during behavioral performance. *J Neurosci* 34:16433–16441. [CrossRef Medline](#)
- Diedrichsen J, White O, Newman D, Lally N (2010) Use-dependent and error-based learning of motor behaviors. *J Neurosci* 30:5159–5166. [CrossRef Medline](#)
- Flöel A, Breitenstein C, Hummel F, Celnik P, Gingert C, Sawaki L, Knecht S, Cohen LG (2005) Dopaminergic influences on formation of a motor memory. *Ann Neurol* 58:121–130. [CrossRef Medline](#)
- Flöel A, Garraux G, Xu B, Breitenstein C, Knecht S, Herscovitch P, Cohen LG (2008) Levodopa increases memory encoding and dopamine release in the striatum in the elderly. *Neurobiol Aging* 29:267–279. [CrossRef Medline](#)
- Galea JM, Celnik P (2009) Brain polarization enhances the formation and retention of motor memories. *J Neurophysiol* 102:294–301. [CrossRef Medline](#)
- Hayashi S, Shimura K, Kasai T (2006) Modulations of use-dependent excitability changes of human motor cortex (M1) by practice condition. *Percept Mot Skills* 103:697–702. [CrossRef Medline](#)
- Hikosaka O, Yamamoto S, Yasuda M, Kim HF (2013) Why skill matters. *Trends Cogn Sci* 17:434–441. [CrossRef Medline](#)
- Hollerman JR, Schultz W (1998) Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat Neurosci* 1:304–309. [CrossRef Medline](#)
- Hosp JA, Pektanovic A, Rioult-Pedotti MS, Luft AR (2011) Dopaminergic projections from midbrain to primary motor cortex mediate motor skill learning. *J Neurosci* 31:2481–2487. [CrossRef Medline](#)
- Huang VS, Haith A, Mazzoni P, Krakauer JW (2011) Rethinking motor learning and savings in adaptation paradigms: model-free memory for successful actions combines with internal models. *Neuron* 70:787–801. [CrossRef Medline](#)
- Huntley GW, Morrison JH, Prikhozhan A, Sealton SC (1992) Localization of multiple dopamine receptor subtype mRNAs in human and monkey motor cortex and striatum. *Mol Brain Res* 15:181–188. [CrossRef Medline](#)
- Izawa J, Shadmehr R (2011) Learning from sensory and reward prediction errors during motor adaptation. *PLoS Comput Biol* 7:e1002012. [CrossRef Medline](#)
- Karni A, Meyer G, Rey-Hipolito C, Jezard P, Adams MM, Turner R, Ungerleider LG (1998) The acquisition of skilled motor performance: fast and slow experience-driven changes in primary motor cortex. *Proc Natl Acad Sci U S A* 95:861–868. [CrossRef Medline](#)
- Kawai R, Markman T, Poddar R, Ko R, Fantana AL, Dhawale AK, Kampff AR, Ölveczky BP (2015) Motor cortex is required for learning but not for executing a motor skill. *Neuron* 86:800–812. [CrossRef Medline](#)
- Knowlton BJ, Mangels JA, Squire LR (1996) A neostriatal habit learning system in humans. *Science* 273:1399–1402. [CrossRef Medline](#)
- Krakauer JW, Mazzoni P (2011) Human sensorimotor learning: adaptation, skill, and beyond. *Curr Opin Neurobiol* 21:636–644. [CrossRef Medline](#)
- Lee D, Seo H, Jung MW (2012) Neural basis of reinforcement learning and decision making. *Annu Rev Neurosci* 35:287–308. [CrossRef Medline](#)
- Luft AR, Schwarz S (2009) Dopaminergic signals in primary motor cortex. *Int J Dev Neurosci* 27:415–421. [CrossRef Medline](#)
- Morgante F, Espay AJ, Gunraj C, Lang AE, Chen R (2006) Motor cortex plasticity in Parkinson's disease and levodopa-induced dyskinesias. *Brain* 129:1059–1069. [CrossRef Medline](#)
- Orban de Xivry JJ, Criscimagna-Hemminger SE, Shadmehr R (2011) Contributions of the motor cortex to adaptive control of reaching depend on the perturbation schedule. *Cereb Cortex* 21:1475–1484. [CrossRef Medline](#)
- Pascual-Leone A, Grafman J, Hallett M (1994) Modulation of cortical motor output maps during development of implicit and explicit knowledge. *Science* 263:1287–1289. [CrossRef Medline](#)
- Ramayya AG, Misra A, Baltuch GH, Kahana MJ (2014) Microstimulation of the human substantia nigra alters reinforcement learning. *J Neurosci* 34:6887–6895. [CrossRef Medline](#)
- Reis J, Schambra HM, Cohen LG, Buch ER, Fritsch B, Zarahn E, Celnik PA, Krakauer JW (2009) Noninvasive cortical stimulation enhances motor skill acquisition over multiple days through an effect on consolidation. *Proc Natl Acad Sci U S A* 106:1590–1595. [CrossRef Medline](#)
- Rossini PM, Barker AT, Berardelli A, Caramia MD, Caruso G, Cracco RQ, Dimitrijević MR, Hallett M, Katayama Y, Lücking CH (1994) Non-invasive electrical and magnetic stimulation of the brain, spinal cord and roots: basic principles and procedures for routine clinical application. Report of an IFCN committee. *Electroencephalogr Clin Neurophysiol* 91:79–92. [CrossRef Medline](#)
- Schambra HM, Abe M, Luckenbaugh DA, Reis J, Krakauer JW, Cohen LG (2011) Probing for hemispheric specialization for motor skill learning: a transcranial direct current stimulation study. *J Neurophysiol* 106:652–661. [CrossRef Medline](#)
- Schultz W (2006) Behavioral theories and the neurophysiology of reward. *Annu Rev Psychol* 57:87–115. [CrossRef Medline](#)
- Selvanayagam VS, Riek S, Carroll TJ (2011) Early neural responses to strength training. *J Appl Physiol* 111:367–375. [CrossRef Medline](#)
- Selvanayagam VS, Riek S, de Rugy A, Carroll TJ (2016) Strength training biases goal-directed aiming. *Med Sci Sports Exerc* 48:1835–1846. [CrossRef Medline](#)
- Shmuelof L, Krakauer JW, Mazzoni P (2012a) How is a motor skill learned? Change and invariance at the levels of task success and trajectory control. *J Neurophysiol* 108:578–594. [CrossRef Medline](#)

- Shmuelof L, Huang VS, Haith AM, Delnicki RJ, Mazzoni P, Krakauer JW (2012b) Overcoming motor “forgetting” through reinforcement of learned actions. *J Neurosci* 32:14617–14621. [CrossRef Medline](#)
- Stefan K, Classen J, Celnik P, Cohen LG (2008) Concurrent action observation modulates practice-induced motor memory formation. *Eur J Neurosci* 27:730–738. [CrossRef Medline](#)
- Sutton RS, Barto AG (1998) Reinforcement learning: an introduction. Cambridge, MA: MIT.
- Therrien AS, Wolpert DM, Bastian AJ (2016) Effective reinforcement learning following cerebellar damage requires a balance between exploration and motor noise. *Brain* 139:101–114. [CrossRef Medline](#)
- Verstynen T, Sabes PN (2011) How each movement changes the next: an experimental and theoretical study of fast adaptive priors in reaching. *J Neurosci* 31:10050–10059. [CrossRef Medline](#)
- Waelti P, Dickinson A, Schultz W (2001) Dopamine responses comply with basic assumptions of formal learning theory. *Nature* 412:43–48. [CrossRef Medline](#)
- Wang AY, Miura K, Uchida N (2013) The dorsomedial striatum encodes net expected return, critical for energizing performance vigor. *Nat Neurosci* 16:639–647. [CrossRef Medline](#)
- Wise RA (2004) Dopamine, learning and motivation. *Nat Rev Neurosci* 5:483–494. [CrossRef Medline](#)
- Wong AL, Lindquist MA, Haith AM, Krakauer JW (2015) Explicit knowledge enhances motor vigor and performance: motivation versus practice in sequence tasks. *J Neurophysiol* 114:219–232. [CrossRef Medline](#)
- Ziemann U, Muellbacher W, Hallett M, Cohen LG (2001) Modulation of practice-dependent plasticity in human motor cortex. *Brain* 124:1171–1181. [CrossRef Medline](#)