Neurorehabilitation and
Neural Repair
2017, Vol. 31(2) 178–189

DOI: 10.1177/1545968316675432
journals.sagepub.com/home/nnr
SAGE

Motor Learning in Stroke: Trained Patients Are Not Equal to Untrained Patients With Less Impairment

Robert M. Hardwick, PhD[1], Vikram A. Rajan, MS[1], Amy J. Bastian, PhD[1,2], John W. Krakauer, MD[1], and Pablo A. Celnik, MD[1]

Abstract

Background and Objective: Stroke rehabilitation assumes motor learning contributes to motor recovery, yet motor learning in stroke has received little systematic investigation. Here we aimed to illustrate that despite matching levels of performance on a task, a trained patient should not be considered equal to an untrained patient with less impairment. *Methods:* We examined motor learning in healthy control participants and groups of stroke survivors with mild-to-moderate or moderate-to-severe motor impairment. Participants performed a series of isometric contractions of the elbow flexors to navigate an on-screen cursor to different targets, and trained to perform this task over a 4-day period. The speed-accuracy trade-off function (SAF) was assessed for each group, controlling for differences in self-selected movement speeds between individuals. *Results:* The initial SAF for each group was proportional to their impairment. All groups were able to improve their performance through skill acquisition. Interestingly, training led the moderate-to-severe group to match the untrained (baseline) performance of the mild-to-moderate group, while the trained mild-to-moderate group matched the untrained (baseline) performance of the controls. Critically, this did not make the two groups equivalent; they differed in their capacity to improve beyond this matched performance level. Specifically, the trained groups had reached a plateau, while the untrained groups had not. *Conclusions:* Despite matching levels of performance on a task, a trained patient is not equal to an untrained patient with less impairment. This has important implications for decisions both on the focus of rehabilitation efforts for chronic stroke, as well as for returning to work and other activities.

Keywords

motor skills, motor skills disorders, motor disorders, motor activity, speed-accuracy trade-off, motor control, hemiparesis, motor impairment

Introduction

Stroke is a leading cause of adult disability, leaving 30% to 66% of patients with lasting motor impairment.[1,2] It has long been proposed that motor recovery following stroke is a form of relearning[3,4] and that there is considerable overlap between the brain regions involved in both processes.[5-7] However, while acquiring skill at a task may allow a patient to perform at the same level as an individual with lesser impairment, this does not necessarily make them equal. For example, well-recovered stroke patients can match the performance of healthy controls on a motor task, but differences exist in the neural networks that underlie performance for each group.[8] Furthermore, matched performance does not necessarily imply that both groups have the same ability to continue improving given the opportunity for practice. These differences can complicate judgments regarding patients' capacity to return to work and other activities,[9] and which rehabilitation activities they should focus on. In this article, we propose that acquiring skill through motor training raises a similar issue—a patient who has trained on a task may "appear better," masking categorical differences in his or her abilities. Consider two hypothetical patients—Patient A, who has mild motor impairment, and Patient B, who is more severely impaired. Patient A

[1]Johns Hopkins University, Baltimore, MD, USA
[2]Kennedy Krieger Institute, Baltimore, MD, USA

Supplementary material for this article is available on the Neurorehabilitation & Neural Repair website at http://nnr.sagepub.com/content/by/supplemental-data.

Corresponding Author:
Pablo Celnik, MD, Department of Physical Medicine and Rehabilitation, Johns Hopkins University, 600 N Wolfe St, Phipps 181, Baltimore, MD 21287, USA.
Email: pcelnik@jhmi.edu

performs better in a movement task than Patient B. Patient B then trains at the task, reaching the same performance level as Patient A. If Patient B is now equal to Patient A, he or she should have a similar capacity for further improvement with training. If this is not the case (eg, if Patient B has reached a performance plateau beyond which further training has a limited effect), then a categorical difference remains between these patients despite their matching task performance.

In comparison to healthy individuals, stroke patients select slower voluntary movement speeds when performing movement tasks.[10] As speed and accuracy are inherently linked,[11] a confound arises when comparing the accuracy of movements performed at different speeds. This limitation makes it difficult to interpret previous results, such as cases where patients improve their accuracy yet decrease their speed.[12] In such cases, it is impossible to determine whether a patient improved his or her ability to perform the task (through skill acquisition) or whether he or she simply changed the aspect of performance on which they focused (eg, sacrificed speed for accuracy while remaining at the same overall level of ability). The only way to disambiguate these alternatives is to first derive the speed-accuracy trade-off function (SAF[13]) for a given task; participants are required to complete the task in a fixed time, allowing accuracy to be measured without the confounding effects of differences in speed. Once derived, skill represents a shift in the SAF.[13-15]

Here we introduce a serial voluntary isometric elbow force task, a modified version of the serial voluntary isometric pinch task (SVIPT). This task is based on an established laboratory-based model of motor learning in which participants learn to control a cursor by producing isometric forces.[13-19] In the task used in the present study, participants controlled a cursor by exerting forces with their elbow flexor muscles, allowing comparisons of performance across participants with greater ranges of impairment than would be possible with the standard (hand controlled) SVIPT paradigm. To control for differences in movement speeds across groups, performance was assessed by comparing the speed-accuracy trade-off pre and post training, using measures of task-level performance (ie, binary success/failure to complete all specified aspects of the task)[13-18] and trial-level measures of endpoint error and variability.[20] We predicted that the severity of a participant's motor impairment would limit his or her ability to perform the task and that training may allow him or her to achieve a similar level of performance as an individual with lesser impairment. However, we hypothesized that despite their matching performance, there would be a categorical difference between these individuals; the previously untrained participant with lesser impairment would be able to make large, rapid improvements through training, while the trained participant would not.

Methods

Participants

A total of 30 participants took part in the study (see Table 1). Participants were required to successfully complete a Mini-Mental Status Examination with a score ≥27/30, excluding participants with potentially confounding cognitive deficits.[21] Stroke survivors with cerebellar lesions and/or ataxia were excluded from the study. All participants had a minimum biceps voluntary contraction strength of 44 N to provide a suitable range for measurement with the experimental apparatus.

Stroke patients were split into 2 groups according to their upper extremity Fugl-Meyer score (ueFMS).[22] In accordance with previously defined groupings,[22-24] "mild-to-moderate" patients had a ueFMS ≥50/66 (n = 10; average ueFMS: 60/66), and "moderate-to-severe" patients had a ueFMS <50/66 (n = 10; average ueFMS: 31/66). The control group (n = 10) consisted of able-bodied participants with no neurological impairments. There were no statistically significant differences between groups for participant age (one-way ANOVA, $F_{2,27} = 0.37$, $P = .70$), or handedness, defined as the dominant hand before the stroke in patients (Fisher's exact test, $P = .31$). The study was approved by the Johns Hopkins Institutional Review Board, and all subjects gave written informed consent in accordance with the declaration of Helsinki.

Apparatus

Participants sat in a robotic exoskeleton[25] that supported their (affected) arm in the horizontal plane (Figure 1A), with the shoulder in 45° of transverse flexion and the elbow in 90° of flexion. Participants controlled the task by exerting isometric forces with the elbow flexor muscles, measured using a force transducer.[26] Participants were required to control the horizontal position of an on screen cursor (Figure 1B). Contracting the elbow flexors moved the cursor rightwards, while relaxing moved it back toward the "Home" position. Participants aimed to stop the cursor within each target in the sequence "Home-1–Home-2–Home-3–Home-4–Home-5." The task was designed as a modified version of the SVIPT.[13-19] Controlling the task with the elbow flexor muscles allowed inclusion of patients with poor hand control.

In accordance with previous studies using this paradigm, a logarithmic transformation of force to cursor movement increased task difficulty.[13] The relationship between cursor position and applied force was scaled to the maximum voluntary contraction (MVC) of each participant, as calculated on the day of the pretraining skill assessment. Movement was scaled such that a contraction of 30% of the individual's MVC would displace the cursor by 30 cm from the home position (ie, to the far right of the visual display; see Supplementary Figure III for further illustration of the

Table 1. Participant Information[a].

Group	Age	Sex	ueFMS	Arm Tested	Years Poststroke	Stroke Type	Lesion Location
Healthy control	28	Female	66	Dominant [R]	N/A	N/A	N/A
	70	Male	66	Dominant [R]	N/A	N/A	N/A
	66	Male	66	Dominant [L]	N/A	N/A	N/A
	61	Female	66	Nondominant [L]	N/A	N/A	N/A
	58	Female	66	Nondominant [L]	N/A	N/A	N/A
	58	Male	66	Nondominant [L]	N/A	N/A	N/A
	51	Female	66	Nondominant [L]	N/A	N/A	N/A
	64	Female	66	Nondominant [L]	N/A	N/A	N/A
	55	Female	66	Nondominant [L]	N/A	N/A	N/A
	47	Male	66	Dominant [L]	N/A	N/A	N/A
Mild-to-Moderate	74	Male	66	Dominant [R]	5	Hemorrhagic	L Frontal
	66	Male	66	Nondominant [L]	9	Hemorrhagic	R Frontal/parietal
	60	Male	65	Dominant [R]	2	Ischemic	L Frontal
	75	Male	64	Dominant [R]	4	Ischemic	L Frontal
	43	Female	61	Nondominant [L]	8	Hemorrhagic	R Parietal
	57	Male	59	Nondominant [L]	1	Ischemic	R Medulla
	64	Male	58	Dominant [R]	4	Ischemic	L Putamen
	21	Male	58	Nondominant [L]	2	Ischemic	Frontal
	66	Male	57	Nondominant [L]	3	Ischemic	R Pons
	58	Female	56	Nondominant [L]	4	Hemorrhagic	R Frontal
Moderate-to-severe	28	Female	43	Dominant [R]	4	Ischemic	L Pons
	69	Male	43	Dominant [R]	3	Ischemic	L Frontal/parietal
	54	Male	40	Nondominant [L]	15	hemorrhagic	R Frontal/putamen
	67	Male	33	Dominant [R]	3	Ischemic	L Thalamus
	43	Male	30	Nondominant [L]	2	Ischemic	L Temporal/putamen
	28	Female	30	Dominant [R]	10	Hemorrhagic	L Frontal
	55	Female	21	Nondominant [L]	5	Hemorrhagic	R Thalamus
	51	Female	21	Nondominant [L]	12	Hemorrhagic	R Frontal/parietal
	70	Female	21	Dominant [R]	11	Ischemic	L Frontal/parietal/thalamus
	63	Male	13	Nondominant [L]	11	Hemorrhagic	Rt Putamen

Abbreviations: ueFMS, upper extremity Fugl-Meyer score (maximum score = 66); L, left; R, right.
[a]There were no significant differences between groups for age, sex, or handedness as assessed prior to stroke (see main text).

positions of the targets and their boundaries). This scaling procedure controlled for differences in participant strength, ensuring that failure to complete the task would be due to control (rather than strength) deficits.

Procedure

The study comprised 3 phases (Figure 1C): a pretraining skill assessment (determining the baseline speed-accuracy trade-off), a training intervention (allowing participants to become skilled at the task), and a posttraining skill assessment (determining whether training changed the speed-accuracy trade-off).

Skill Assessments

To assess skill we determined the speed-accuracy trade-off, measuring accuracy at fixed execution speeds. Participants were instructed to aim to move the cursor to targets in time with a metronome, with the instruction that each beat of the metronome should correspond to hitting one of the targets. One block of the task comprised completing 10 trials at a fixed tempo (the experiment involved blocks completed at 9 different tempos, 24/30/38/45/60/80/100/110/120 BPM, corresponding to approximate trial durations of 12.5/10.0/7.9/6.7/5.0/3.8/3.0/2.7/2.5 seconds, respectively). Blocks were completed in a pseudorandom order (to prevent order effects) until 10 trials were collected for each of the 9 tempos. This procedure was repeated twice, providing a total of 20 trials per tempo.

Training

Participants completed 4 training sessions on consecutive days, each comprising 5 blocks of 30 trials. Each training session took participants approximately 30 minutes to complete. In contrast to the skill assessments, we instructed participants to complete training trials at a self-selected pace,

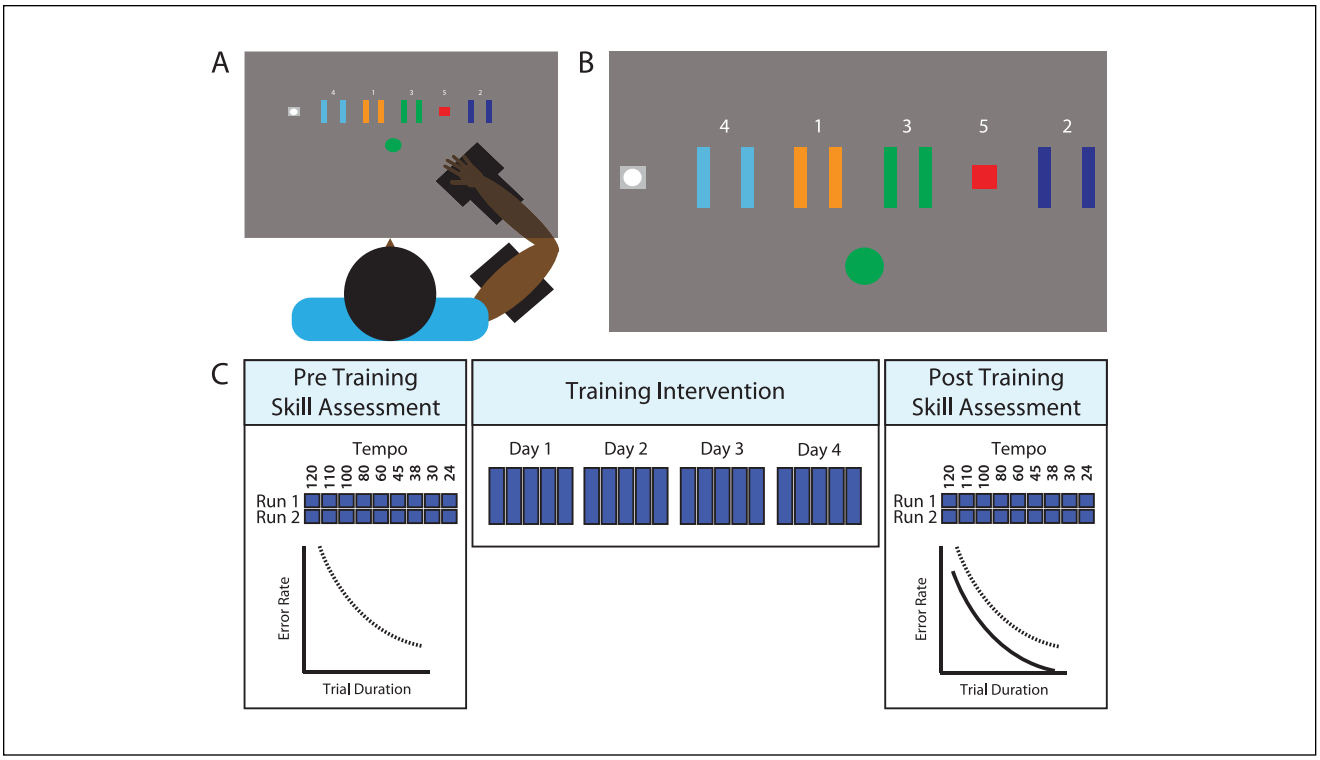

Figure 1. Experimental setup and procedure. (A) Participants sat with their (affected) arm supported by a robotic exoskeleton. A force transducer measured contractions of their elbow flexors. (B) On screen display. Contracting the elbow flexors moved the cursor (white circle) to the right, while relaxing moved the cursor to the home position (grey square). A "go" indicator (used in training trials) indicated to participants that they could begin a trial when ready (illustrated here as a green circle). Each trial involved navigating the cursor through the sequence Home-1–Home-2–Home-3–Home-4–Home-5. Target positions and sequence order remained fixed throughout the study. (C) Procedure. Participants first completed a pretraining skill assessment, performing the task at trial durations set by an auditory metronome (indicated by tempos presented in beats per minute—see main text for further detail). One "run" of the task involved completing 10 trials at each tempo in a pseudorandom order. This procedure was repeated to generate 2 runs of data (ie, a total of 20 trials for each tempo). Participants later trained to perform the task over consecutive days, aiming to complete the sequence as quickly and as accurately as possible. Finally, on a separate day, participants completed a posttraining skill assessment.

with the aim of completing each trial as quickly and as accurately as possible.[13] Participants were verbally encouraged to attempt to improve their speed and accuracy prior to each block, and rested for 2 minutes between blocks to prevent fatigue.

Analysis

Previous studies using the paradigm on which the current task is based have predominantly employed binary performance measures; success is achieved by hitting all targets,[13-18] and an average error-rate is calculated by dividing the number of successful trials by the total number of trials in the block. This provides a useful task-level index of performance, but it has limited sensitivity. For example, when attempting to hit the bullseye on a dartboard, a binary measure (hitting the target) would not detect improvement if a participant missed by an average of *one meter* before training, then by an average of *one centimeter* after training. We therefore performed both a task-level assessment of success using a binary success/failure metric (empirically measuring the speed-accuracy trade-off) and also examined more detailed measures including error magnitude, cursor endpoint variability, and the number of targets the participant attempted to hit, providing more sensitive trial-level metrics.[20]

Skill Assessments. Task-level success was defined as moving the cursor inside each of the 5 targets.[13-18] We used a mixed-design ANOVA examining 3 factors: the 2 within-participant factors of *duration and session*, and the between-participant factor of *group*. Between-group effects were compared using separate $_{RM}$ANOVAs for each session (factors *duration* and *group*). Within-group comparisons were conducted using separate $_{RM}$ANOVAs (factors *session* and *duration*). We conducted planned comparisons on trial *duration* for each group using paired samples *t* tests.

Mean Group Skill. We examined whether each group of participants represented a separate population in terms of skill performance pre and post training. We calculated the mean error rate (ie, averaged across all 9 trial durations) for each participant and conducted a mixed-design ANOVA with factors of *session* and *group*. In a further analysis, we bootstrapped 10 000 mean group resamples, with replacement. Data from each participant contributing to a resample was entered into both the pre and post training sample estimates, providing a repeated-measures test. For example, when resampling the control group, we took the 10 original participants from the group and randomly selected 10 participants that would contribute to the resample (allowing the same participant to be present multiple times within a resample). We then calculated the mean error rate both pre and post training for this resampled group. This process was repeated until we had generated a pool of 10 000 mean values from such resamples. We calculated 95% confidence intervals by taking 2.5 and 97.5 percentiles from this process.[27] This resampling analysis provides the additional benefit of controlling for the presence of participants that are not statistical outliers, yet may be considered to represent extreme values within their groups. For example, the mild-to-moderate group contained 2 stroke patients with ueFMS values of 66/66. While these scores would appear to indicate these patients have no motor impairment, an alternate view is that these patients may have motor impairments to which the ueFMS is not sensitive, but may be revealed by more sensitive examinations of their motor capabilities (such as the present task). A key benefit of resampling with replacement is that this process includes samples wherein these potentially extreme values are both overrepresented (samples in which they were randomly selected for inclusion >2/10) or absent (samples in which they were not randomly selected for inclusion), allowing us to determine whether their presence or absence changed the results of the study.

Trial Endpoint Errors. Magnitudes of errors made when attempting to hit each target were measured as the shortest distance from the cursor endpoint to the outer boundary of the corresponding target (consistent with errors participants observed during the task). Attempts falling within the target boundaries thus had 0 cm error. If a participant omitted a target then 0 cm error was assigned; we took this conservative approach as another analysis examined omission errors (see supplementary materials), and because omitting a target is fundamentally different from attempting to move to it and missing. Error within each trial was summed and analyzed using a mixed-design ANOVA with factors *duration*, *session*, and *group*. We conducted similar analyses for variability (see supplementary materials). A further analysis considered the magnitude of errors for each individual target (see supplementary materials).

Training Data. Previous studies have combined speed and accuracy data to assess "skill" throughout training.[13-18] However, such a comparison here would not be valid due to differences in the baseline SAF for each group (see discussion and supplementary materials). We instead conducted separate mixed-design ANOVAs on durations and error rates, with a within-subject factor *day* and a between-subject factor *group*. Planned *t* tests compared mean performance for subsequent days. We made no comparisons between group "deltas" (ie, comparisons of percentage or absolute change from baseline across groups) to avoid confounding differences in baseline performance.[28]

Results

Task Skill Assessments

The mixed-design ANOVA comparing error rates revealed a significant *duration* × *session* × *group* interaction, $F_{16,216} = 3.49$, $P < .001$. Between-group comparisons were conducted using separate mixed-design ANOVAs for each skill assessment.

Participants With Greater Motor Impairment Had Less Skill Pre and Post Training. In the pretraining skill assessment (Figure 2A), error rates were greater when participants completed trials with shorter durations (mixed-design ANOVA, main effect of duration, $F_{8,16} = 145.01$, $P < .001$). Participants with greater motor impairment also made more errors that varied depending on trial duration (*group* × *duration* interaction, $F_{16,216} = 12.58$, $P < .001$). Control participants were more accurate than mild-to-moderate patients for durations from 3.8 to 12.5 seconds (all $P < .05$), and more accurate than moderate-to-severe patients for durations from 3.0 to 12.5 seconds (all $P < .05$). Mild-to-moderate patients were more accurate than moderate-to-severe patients for durations from 3.0 to 12.5 seconds (all $P < .05$).

The posttraining skill assessment identified similar effects. Error rates were greater when trial durations were shorter (mixed-design ANOVA, main effect of *duration,* $F_{8,16} = 147.17$, $P < .001$). Participants with greater motor impairment again made more errors dependent on the trial duration (*duration* × *group* interaction, $F_{16,216} = 4.58$, $P < .001$). Control participants were more accurate than mild-to-moderate patients for durations from 3.75 to 12.5 seconds and from 2.5 to 2.7 seconds (all $P < .05$), and more accurate than moderate-to-severe patients for all durations (all $P < .05$). Mild-to-moderate patients were more accurate than moderate-to-severe patients for durations from 3 to 10 seconds (all $P < .05$).

All Groups Performed Better in the Posttraining Skill Assessment. Figure 2B presents pre versus post comparisons of performance for each group. Control participants were able

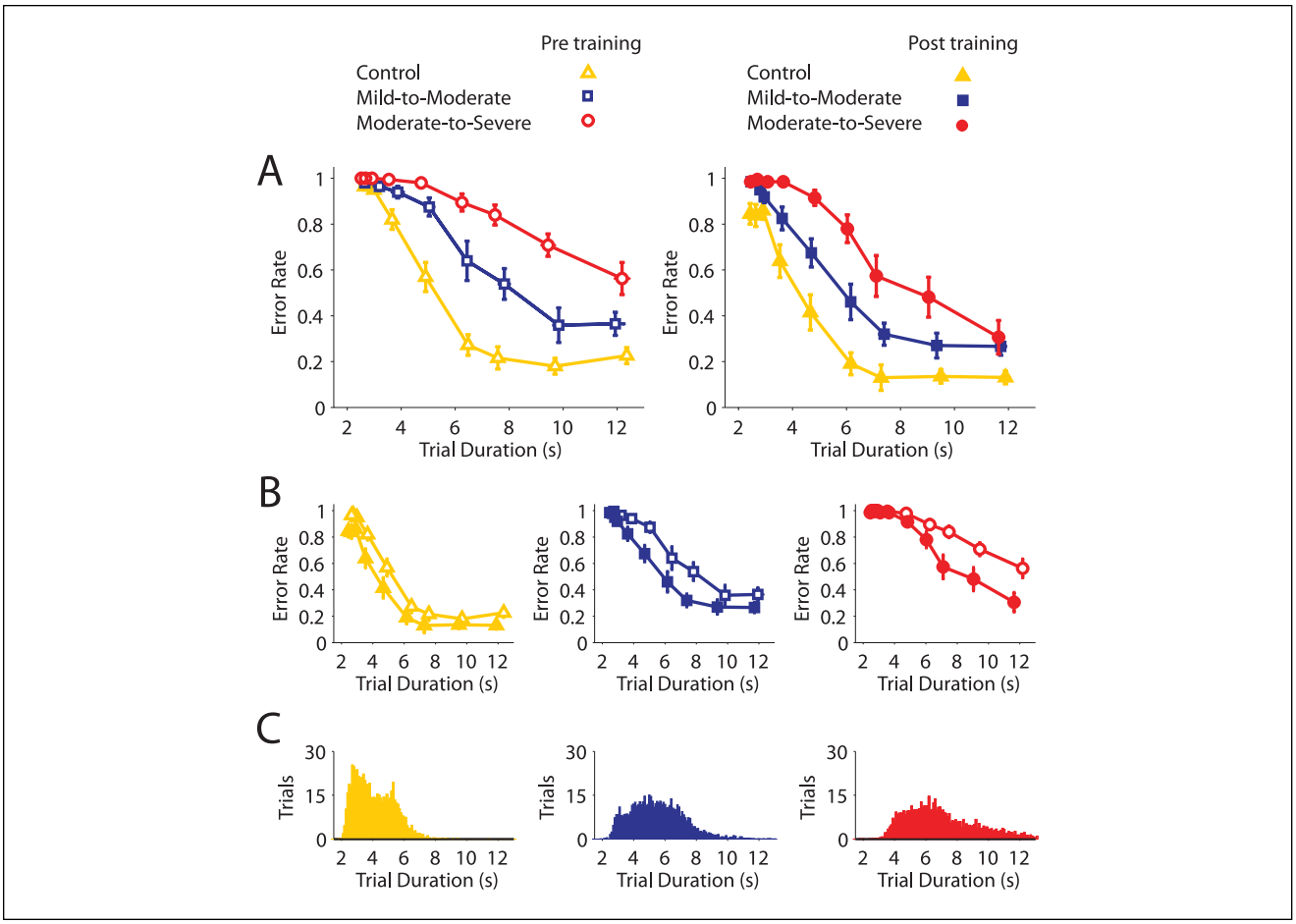

Figure 2. Task-level skill assessment results and training data distributions. Open and closed shapes show group average performance before and training, respectively. (A) Between-group comparisons of performance before (left) and after (right) training, highlighting significant differences in group performance for both sessions. (B) Within-group comparisons of performance pre and post training for the control, mild-to-moderate impairment, and moderate-to-severe impairment groups, respectively. (C) Histograms showing the average number of times that participants completed training trials at a particular speed. Error bars present SEM.

to reduce their error rates across the entire range of trial durations examined ($_{RM}$ANOVA, effects of *session,* $F_{1,9}$ = 16.69, $P < 0.01$, and *duration,* $F_{8,72}$ = 95.53, $P < 0.001$, but no significant *session* × *duration* interaction, $F_{8,72}$ = 1.43, P = .20; this indicates differences were not driven by any particular duration—planned pre vs post comparisons for each trial duration were all $P < .05$).

Participants with mild-to-moderate impairment were able to improve their performance across some (but not all) trial durations ($_{RM}$ANOVA, significant *session* × *duration* interaction, $F_{8,72}$ = 2.46, $P < .05$). Mild-to-moderate patients improved over durations from 2.7 to 10.0 seconds (all $P < 0.05$), but not at the fastest or slowest durations tested (2.5 seconds/12.5 seconds, both $P > .18$).

Participants with moderate-to-severe impairment improved their performance over a limited range of trial durations ($_{RM}$ANOVA, *session* × *duration* interaction, $F_{8,72}$ = 7.00, $P < .001$). Moderate-to-severe patients improved their performance only at the slower movement durations from 5.0 to 12.5 seconds (all $P < .05$), but not the faster durations from 2.5 to 3.75 seconds (all $P > .3$).

Training Increased Patient Skill to the Level of an Untrained Group With Less Impairment. The mixed-design ANOVA examining mean error rates revealed significant main effect of *session,* $F_{1,27}$ = 52.23, $P < .001$, but no *session* × *group* interaction, $F_{2,27}$ = 0.16, P = .99. Figure 3 presents the resampling analysis of mean error rates in skill assessments. Distributions indicate each group represented a separate population before and after training (Figure 3A and B). Notably, there was considerable overlap between posttraining performance of some groups with pretraining performance of others (see Figure 3C). Posttraining moderate-to-severe patients did not differ from pretraining mild-to-moderate patients. Similarly, posttraining mild-to-moderate patients did not differ from the pretraining healthy controls.

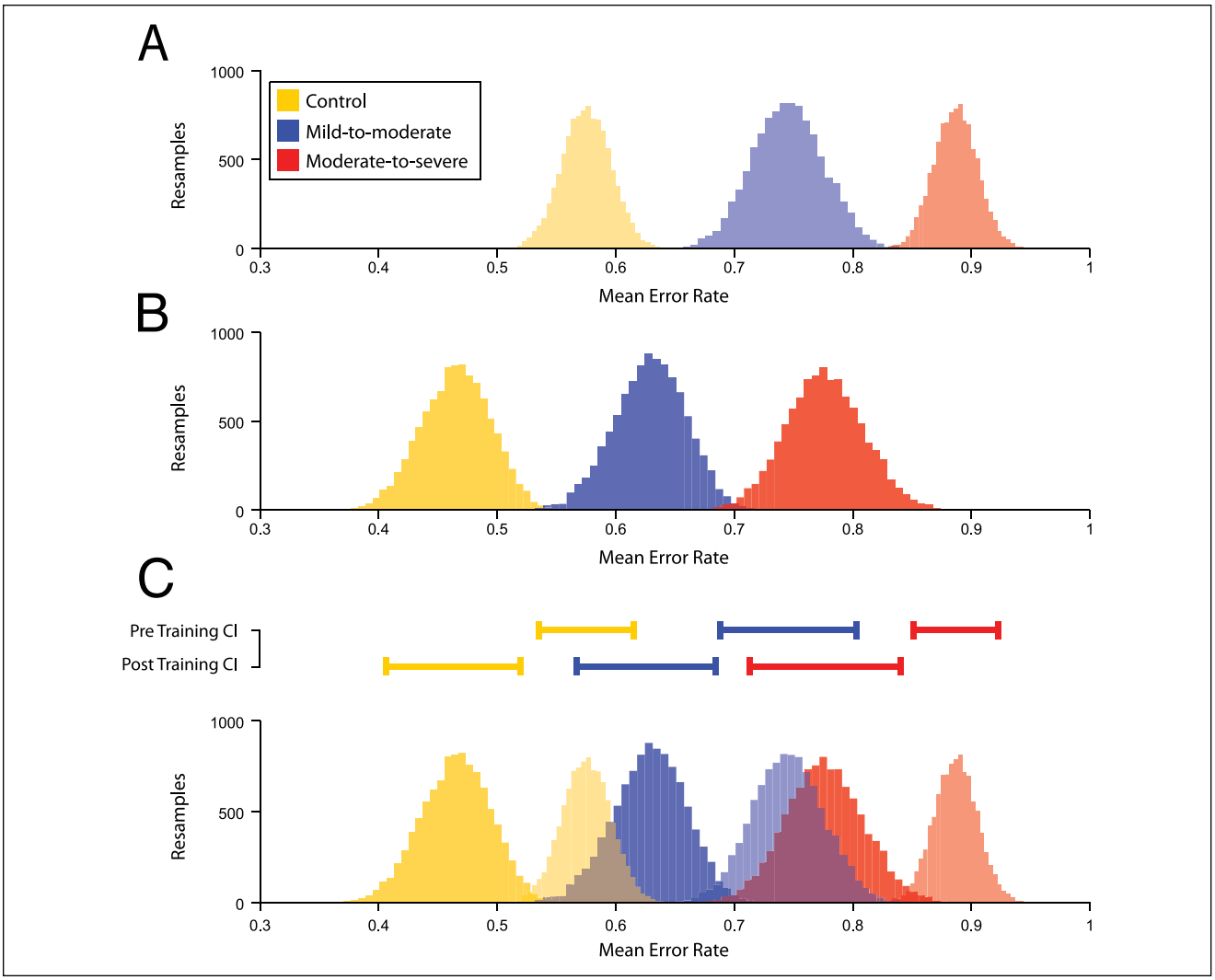

Figure 3. Estimates of group mean error rates during SAF assessments. Estimates were generated from 10 000 resamples of the data for each group. (A) Pretraining estimates of mean error rates. (B) Posttraining estimates of mean error rates. Note that A and B illustrate that each group are separate populations both before and after training. (C) Overlays of the pre and post training data; 95% confidence intervals are presented above the distributions, illustrating significant differences between groups and sessions. Note that posttraining performance for the moderate-to-severe group overlaps pretraining performance for the mild-to-moderate group, and that the posttraining performance for the mild-to-moderate group overlaps pretraining performance for the healthy control group.

Training Reduced Cursor Endpoint Error. Figure 4A illustrates the effects of training on cursor endpoint errors. Participants made larger errors on faster trials (mixed design ANOVA, main effect *duration,* $F_{8,216} = 38.10$, $P < 1 \times 10^{-36}$), and greater impairment exacerbated this effect (*duration* × *group,* $F_{16,216} = 5.099$, $P < 1 \times 10^{-8}$). All participants made smaller errors following training (main effect of *session,* $F_{1,27} = 9.55$, $P < .005$) across a range of trial durations (*duration* × *session* interaction, $F_{8,216} = 2.312$, $P < .05$), though error magnitudes were still higher in participants with greater impairment (*session* × *group,* $F_{2,27} = 3.93$, $P < .05$), particularly when completing faster trials(*duration* × *session* × *group* interaction, $F_{16,216} = 2.578$, $P < .01$).

Figure 4B illustrates the pattern of endpoint positions across the 3 groups. A post hoc analysis (see supplementary materials) indicated that participants made larger errors when attempting to hit targets closer to the home position (ie, those targets that required the execution of smaller, more precise isometric forces). Both speed and impairment exacerbated the magnitude of these errors, while training reduced them.

Training Data

Participants Reached a Plateau in Performance During Training. Training performance is presented in Figure 5. During training sessions participants completed the task at

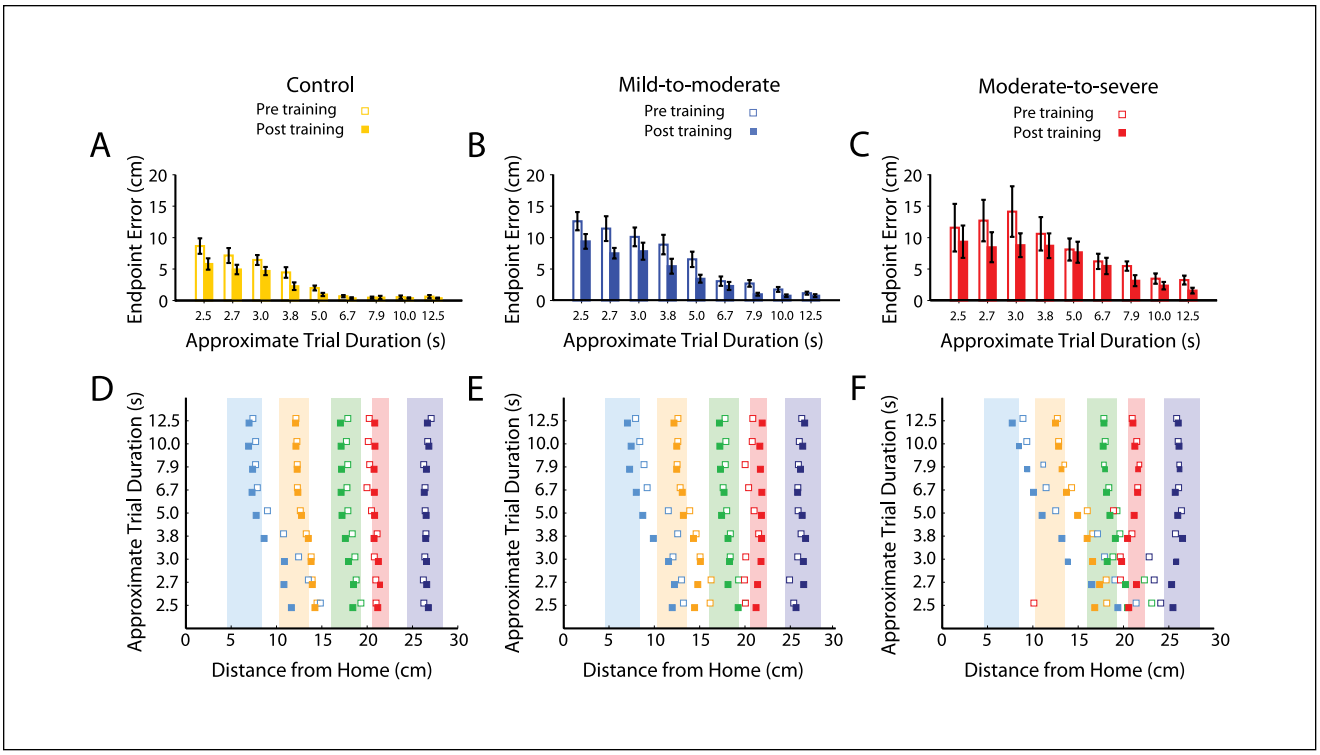

Figure 4. Endpoint error magnitudes and cursor endpoint positions. (A) Within-group comparisons of pre and post training endpoint errors, illustrating the significant effects of trial duration and session on performance. Error bars present SEM. (B) Average cursor endpoint positions for the 3 groups. Squares show group average endpoints for the different targets, represented by the correspondingly colored shaded regions. Numbers above shaded regions denote target number.

self-selected speeds. With training all groups started to complete trials at faster speeds, though this did not differ across groups (mixed-model ANOVA, main effect of *day,* $F_{1,3} = 6.12$, $P < .01$, but no *day* × *group* interaction, $F_{2,6} = 1.61$, $P = .15$). There was a significant reduction in trial duration between days 1 and 2, $t_{29} = 2.35$, $P < .05$, a trend for reduction between days 2 and 3, $t_{29} = 2.04$, $P = .051$, and no significant change between days 3 and 4, $t_{29} = 0.18$, $P = .86$. There was also a difference in speed between groups (main effect *group*, $F_{2,27} = 10.29$, $P < .001$). Trial durations for the control group were significantly shorter than those of the mild-to-moderate group ($P < .05$) and the moderate-to-severe group ($P < .001$). The mild-to-moderate group had significantly shorter trial durations than the moderate-to-severe group ($P < .05$).

In contrast, error rates did not differ with training (mixed-design ANOVA, no effect of *day,* $F_{1,3} = 0.73$, $P = .54$), or across groups (no main effect of *group,* $F_{2,27} = 1.94$, $P = 0.16$, and no *day* × *group* interaction, $F_{2,6} = 0.68$, $P = .67$). Planned comparisons between subsequent days indicated no differences in error rates (all $t < 1.8$). Planned between-group comparisons indicated no significant differences in error rates between healthy controls and patients with mild-to-moderate impairment ($P = .86$), the healthy controls and the moderate-to-severe group ($P = .12$), or the mild-to-moderate and the moderate-to-severe groups ($P = .08$).

Discussion

The present study examined the ability of chronic stroke patients to learn a motor task with their paretic arm, and compared their performance with healthy controls. We controlled for changes in performance attributable to moving along baseline SAFs (eg, improving accuracy by decreasing speed) by assessing performance at fixed trial durations in skill assessments. Pretraining assessments indicated all 3 groups had different SAFs, with greater impairment associated with an inferior SAF. All groups improved their performance with training as evidenced by shifts in the SAF; however, posttraining comparisons still indicated clear between-group differences in performance. More detailed trial-level measures identified that all groups improved their performance in the same manner—by reducing the endpoint error and variability of their movements. A resampling analysis indicated that trained patients were able to match the untrained baseline performance of patients with lesser impairment. Critically, data from training indicated that

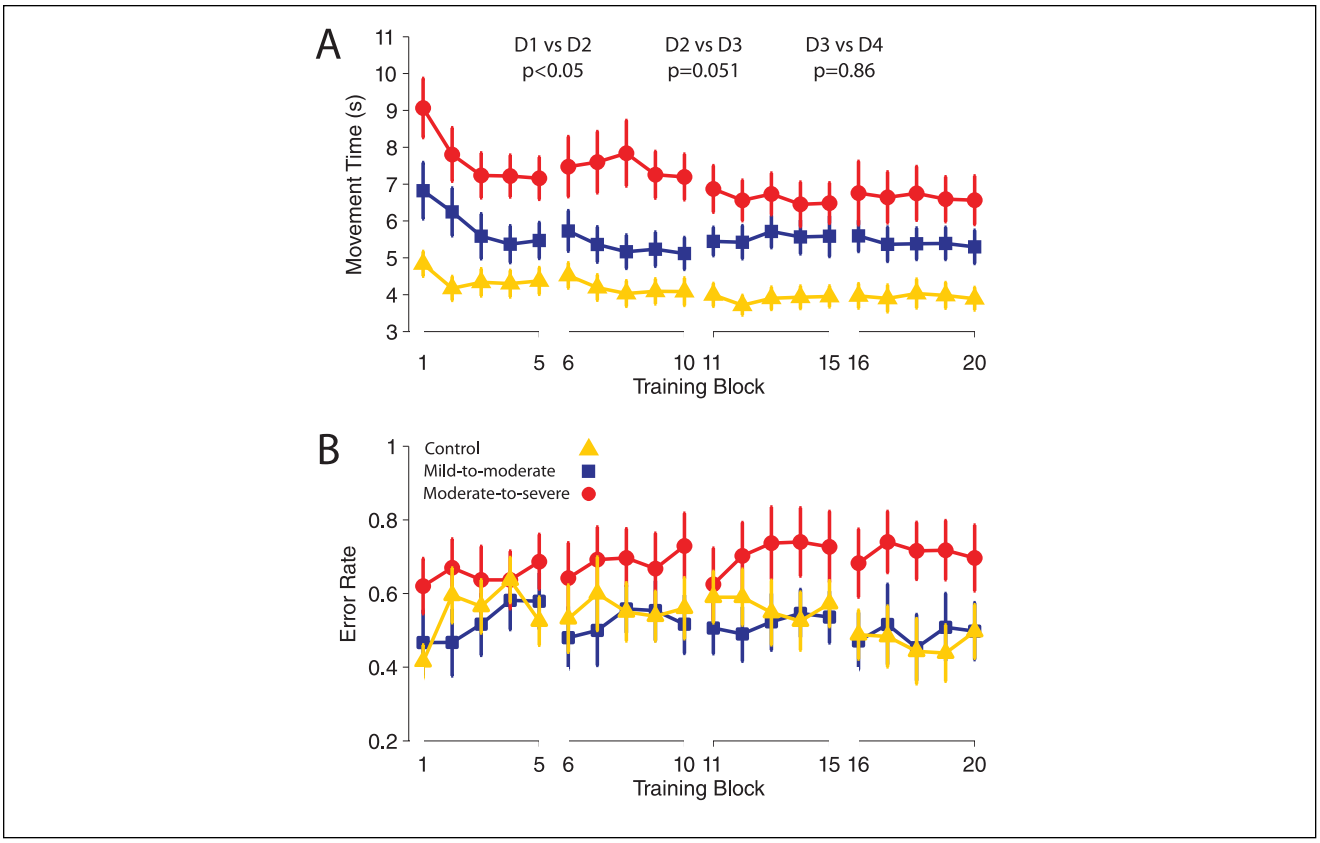

Figure 5. Performance during training. Each maker presents the group mean for a block (30 trials) of training data. (A) Group average movement times over training indicate movement times plateaued on day 3 of training (ie, no significant difference between days 3 and 4). (B) Mean group error rates showed no change over the course of training. Error bars present SEM.

trained patients had reached a plateau in their performance. Therefore, despite their matched performance, a categorical difference remained between these groups; the untrained groups could make large, rapid improvements with training that the trained groups could not.

The present study operationalized skill as a SAF.[13-19] During training participants improved their speed in completing the task; during skill assessments, where trial duration was fixed, this translated to increased accuracy after training. Comparisons of performance across training indicate participants reached a performance plateau during the study (large reductions in trial durations between days 1 and 2 of training, smaller changes from days 2 to 3, and no differences between days 3 and 4). This plateau is consistent with previous studies that have found training leads to initial rapid gains in performance, but that as participants gain experience with a task further training only leads to smaller, gradual improvements.[29,30] The level of performance at which this plateau in learning occurred was affected by the severity of motor impairment; healthy controls reached a plateau with a greater level of performance than mild-to-moderate stroke patients, who in turn reached a plateau with a greater level of performance than moderate-to-severe patients. Notably, the *posttraining* performance of the moderate-to-severe group did not differ from *pretraining* baseline performance of the mild-to-moderate group (see overlapping dark red and light blue distributions in Figure 3C), and the posttraining performance of the mild-to-moderate group did not differ from the pretraining performance of healthy controls (see overlapping dark blue and light yellow distributions from Figure 3C). From this it could be argued that training led the moderate-to-severe group to be equal to the mild-to-moderate group, and that further training would allow both groups to reach the same level of performance as healthy controls. However, the *trained* patients had reached a *performance plateau* by day 3 of the study, whereby further training had a negligible effect on performance. In contrast, the *groups that had not yet trained* still had the capacity to make rapid improvements through motor learning. This categorical difference between the trained and untrained groups indicates that training to acquire the same level of performance as a group with less impairment does not make the groups equivalent. This is because the mechanisms underlying performance—being

impaired and acquiring skill as opposed to having less impairment—are not the same.

Baseline performance differences across impairment levels prevent comparisons of the amount learned by one group versus another. In our view, such analyses are conceptually fraught. For example, assessing performance changes using the difference between baseline and post-training performance would suggest all groups learned to an equal extent; in contrast, calculating the *percentage* change from the baseline would suggest patients with moderate-to-severe impairment learned more. Thus, arbitrary model selection could lead to differing results.[28] Furthermore, there is considerable evidence that the relationship between training and performance is nonlinear; as people become increasingly skilled they require more training to achieve small gains in performance.[31] Here we avoided this conceptual pitfall by comparing *absolute* levels of performance rather than normalized changes. We conclude that all groups improved their performance of the task through skill acquisition, but the final level of performance each group achieved was still affected by their motor impairment.

The task-level measure used in skill assessments examined performance with a binary metric, insensitive to more subtle improvements in performance. Trial-level analyses revealed improvements in endpoint precision and variability following training; such improved quality of control is typically observed in skill acquisition.[20,32-35] However, despite this improvement in movement control, patients' performance remained proportional to the severity of their impairment. This suggests that chronic patients can learn and improve control within a task,[36] but this is not necessarily enough to change more coarse (ie, task-level success/failure) measures of performance.

Strengths and Limitations

The present study employed principles developed using an established laboratory-based skill task to study motor learning in stroke. Modifying this task to use the elbow flexor muscles (as opposed to the hand) allowed the inclusion of patients with a wide range of neurological impairment. Controlling for differences in the speed-accuracy trade-off also allowed separation of changes in skill from changes in the way the task was performed (ie, prevented participants from slowing down to increase their accuracy). Furthermore, distinguishing between task-level success and more detailed trial-level metrics is an important advance for studying motor control in stroke, and for the interpretation of results of this task. Here, trial-level metrics revealed increases in endpoint precision/decreases in variability that drive changes in task-level measures of success or failure. Such 2-tier analysis allows a greater depth of interpretation of how motor learning changes performance.

The present study used an established laboratory-based paradigm to conduct a finely controlled investigation of motor learning capacity in both chronic stroke patients and healthy controls. Notably, the primary goal of the task was to assay motor learning rather than to approximate functional relevance. However, we anticipate that the skill learning processes studied here would also apply to the learning of more ecologically relevant tasks.

Our patient sample spanned a wide range of impairment, including 2 patients with ueFMS of 66/66. It could be argued that the inclusion of patients with "no impairment" may compromise the study. In contrast, we view this as a strength of the present investigation, as it allowed us to study motor learning in patients with a full range of impairment (ie, a ueFMS of 66/66 does not rule out impairments in learning). Furthermore, the resampling analysis included randomly generated samples in which these patients were completely absent, indicating that their inclusion did not change the empirical result of the study.

The present result examined motor learning over a relatively short (4-day) period. This was sufficient to show that a patient can train to match the performance of an untrained individual with less impairment. Even if subthreshold changes in performance occurred beyond the plateau identified on days 3 to 4, the small magnitude of such changes is consistent with our central point; an untrained patient can make rapid and significant improvements in comparison to a trained patient. It could be proposed that training over several months or years could lead to greater improvements and that the brain may still be undergoing training-induced changes. That said, longer studies based on motor learning principles have shown limited impact in chronic stroke.[37-40] In contrast to the relatively static levels of impairment seen in chronic stroke, there is evidence that spontaneous biological recovery early after stroke can lead to substantial improvements in patient's neurological motor function.[41-45] Recent evidence in a mouse model indicates that initiating training during the period early, but not later, after stroke leads to a greater improvements than would be expected through spontaneous biological recovery alone.[46] Future investigations will determine whether training early after stroke can be similarly beneficial in humans.[47-49]

Clinical Implications

Models of motor rehabilitation propose that motor recovery following stroke is a form of relearning.[3,4] Our results indicate that motor learning allows stroke patients with a wide range of impairment to improve their performance of a trained task with their affected arm. Patients improved in the same manner as healthy controls (ie, through learning to control the cursor with increased precision/reduced variability), indicating the same behavioral processes underlie improved performance in all groups. This illustrates that as

long as a patient can perform a task at baseline, he or she should have a capacity to improve at the task with training. However, despite a preserved ability to learn through motor training, the overall level of performance patients were able to achieve was still affected by their motor impairment. This indicates that studies in patients with chronic stroke should focus on training functionally relevant tasks to the limit allowed by impairment and/or on teaching compensatory strategies.[50]

Previous imaging work has indicated that chronic stroke patients may achieve levels of performance that match healthy individuals, yet differing neural networks underlie this performance.[8] We have previously argued that such differences complicate decisions regarding capacity to return to work and other activities[9] and can affect decision making when focusing the target of rehabilitation interventions. Results of the present study indicate that the acquisition of motor skill can lead to similar problems in assessing the capabilities of patients. A patient who has trained on a task on which he or she is assessed may have performance equal to that of an individual with lesser impairment. However, this does not mean that these patients are now equal, as evidenced by the difference in their abilities to improve with training. This should be considered when determining the rehabilitation activities they should focus on and taken into account when assessing a patient's ability to return to work and other activities.[9]

Conclusions

In conclusion, we compared the effects of motor training in healthy individuals and chronic stroke patients with different levels of impairment. We controlled for the performance confound (ie, reducing speed to improve accuracy) by assessing performance at fixed trial durations, allowing detection of changes in motor skill (ie, shifts in the speed-accuracy trade-off). Patients improved their performance through skill acquisition in a manner analogous to healthy controls (ie, increased precision/reduced variability). Training improved patient performance to the same level as the *untrained baseline* of a group with less impairment. However, over the course of the training sessions each group reached a performance plateau whereby more practice did not lead to further improvement. Our results indicate that chronic stroke patients can improve the control of their paretic arm through motor learning, yet remain categorically different to untrained individuals who have the same level of performance but with lesser impairment.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: NIH Grant R01HD073147.

References

1. Richards L, Pohl P. Therapeutic interventions to improve upper extremity recovery and function. *Clin Geriatr Med*. 1999;15:819-832.
2. World Health Organization. *International Classification of Functioning, Disability and Health (ICF)*. Geneva, Switzerland: World Health Organization; 2010.
3. Carr JH, Shepherd RB. A motor learning model for stroke rehabilitation. *Physiotherapy*. 1989;75:372-380.
4. Horak FB. Assumptions underlying motor control for neurologic rehabilitation. In: Lister M, ed. *Contemporary Management of Motor Control Problems*. Proceedings of the II STEP Conference. Alexandria, VA: Foundations for Physical Therapy; 1991:11-27.
5. Rehme AK, Eickhoff SB, Rottschy C, Fink GR, Grefkes C. Activation likelihood estimation meta-analysis of motor-related neural activity after stroke. *Neuroimage*. 2012;59:2771-2782.
6. Hardwick RM, Rottschy C, Miall RC, Eickhoff SB. A quantitative meta-analysis and review of motor learning in the human brain. *Neuroimage*. 2013;67:283-297.
7. Hardwick RM, Lesage E, Eickhoff CR, Clos M, Fox P, Eickhoff SB. Multimodal connectivity of motor learning-related dorsal premotor cortex. *Neuroimage*. 2015;123:114-128.
8. Sharma N, Baron JC, Rowe JB. Motor imagery after stroke: relating outcome to motor network connectivity. *Ann Neurol*. 2009;66:604-616.
9. Celnik P, Hillis AE. Reconnecting the dots after stroke. *Ann Neurol*. 2009;66:570-571.
10. Turnbull GI, Charteris J, Wall JC. A comparison of the range of walking speeds between normal and hemiplegic subjects. *Scand J Rehabil Med*. 1995;27:175-182.
11. Fitts PM. The information capacity of the human motor system in controlling the amplitude of movement. *J Exp Psychol*. 1954;47:381-391.
12. Platz T, Denzler P, Kaden B, Mauritz KH. Motor learning after recovery from hemiparesis. *Neuropsychologia*. 1994;32:1209-1223.
13. Reis J, Schambra HM, Cohen LG, et al. Noninvasive cortical stimulation enhances motor skill acquisition over multiple days through an effect on consolidation. *Proc Natl Acad Sci U S A*. 2009;106:1590-1595.
14. Cantarero G, Tang B, O'Malley R, Salas R, Celnik P. Motor learning interference is proportional to occlusion of LTP-like plasticity. *J Neurosci*. 2013;33:4634-4641.
15. Cantarero G, Lloyd A, Celnik P. Reversal of long-term potentiation-like plasticity processes after motor learning disrupts skill retention. *J Neurosci*. 2013;33:12862-12869.
16. Reis J, Fischer JT, Prichard G, Weiller C, Cohen LG, Fritsch B. Time- but not sleep-dependent consolidation of tDCS enhanced visuomotor skills. *Cereb Cortex*. 2015;25:109-117.
17. Schambra HM, Abe M, Luckenbaugh DA, Reis J, Krakauer JW, Cohen LG. Probing for hemispheric specialization for

motor skill learning: a transcranial direct current stimulation study. *J Neurophysiol*. 2011;106:652-661.
18. Wymbs NF, Bastian AJ, Celnik PA. Motor skills are strengthened through reconsolidation. *Curr Biol*. 2016;26:338-343.
19. Saucedo Marquez CM, Zhang X, Swinnen SP, Meesen R, Wenderoth N. Task-specific effect of transcranial direct current stimulation on motor learning. *Front Hum Neurosci*. 2013;7:333.
20. Shmuelof L, Krakauer JW, Mazzoni P. How is a motor skill learned? Change and invariance at the levels of task success and trajectory control. *J Neurophysiol*. 2012;108:578-594.
21. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12:189-198.
22. Fugl-Meyer AR, Jääskö L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient. 1. A method for evaluation of physical performance. *Scand J Rehabil Med*. 1975;7:13-31.
23. Cirstea MC, Ptito A, Levin MF. Arm reaching improvements with short-term practice depend on the severity of the motor deficit in stroke. *Exp Brain Res*. 2003;152:476-488.
24. Ellis MD, Sukal-Moulton TM, Dewald JPA. Impairment-based 3-D robotic intervention improves upper extremity work area in chronic stroke: targeting abnormal joint torque coupling with progressive shoulder abduction loading. *IEEE Trans Robot*. 2009;25:549-555.
25. Scott SH. Apparatus for measuring and perturbing shoulder and elbow joint positions and torques during reaching. *J Neurosci Methods*. 1999;89:119-127.
26. Charles SK, Okamura AM, Bastian AJ. Does a basic deficit in force control underlie cerebellar ataxia? *J Neurophysiol*. 2013;109:1107-1116.
27. Hardwick RM, Celnik PA. Cerebellar direct current stimulation enhances motor learning in older adults. *Neurobiol Aging*. 2014;35:2217-2221.
28. Kitago T, Krakauer JW. Motor learning principles for neurorehabilitation. *Handb Clin Neurol*. 2013;110:93-103.
29. Newell A, Rosenbloom PS. Mechanisms of skill acquisition and the law of practice. *Cogn Skills Acquis*. 1981;6:1-55.
30. Heathcote A, Brown S, Mewhort DJ. The power law repealed: the case for an exponential law of practice. *Psychon Bull Rev*. 2000;7:185-207.
31. Wymbs NF, Grafton ST. The human motor system supports sequence-specific representations over multiple training-dependent timescales. *Cereb Cortex*. 2015;25:4213-4225.
32. Barden JM, Balyk R, James Raso V, Moreau M, Bagnall K. Repetitive pointing to remembered proprioceptive targets improves 3D hand positioning accuracy. *Hum Mov Sci*. 2005;24:184-205.
33. Madeleine P, Madsen TMT. Changes in the amount and structure of motor variability during a deboning process are associated with work experience and neck-shoulder discomfort. *Appl Ergon*. 2009;40:887-894.
34. Müller H, Sternad D. Motor learning: changes in the structure of variability in a redundant task. *Adv Exp Med Biol*. 2009;629:439-456.
35. Scholz JP, Schöner G. The uncontrolled manifold concept: identifying control variables for a functional task. *Exp Brain Res*. 1999;126:289-306.
36. Winstein CJ, Merians AS, Sullivan KJ. Motor learning after unilateral brain damage. *Neuropsychologia*. 1999;37:975-987.
37. Lo AC, Guarino PD, Richards LG, et al. Robot-assisted therapy for long-term upper-limb impairment after stroke. *N Engl J Med*. 2010;362:1772-1783.
38. Kitago T, Goldsmith J, Harran M, et al. Robotic therapy for chronic stroke: general recovery of impairment or improved task-specific skill? *J Neurophysiol*. 2015;114:1885-1894.
39. Lang KC, Thompson PA, Wolf SL. The EXCITE Trial: reacquiring upper-extremity task performance with early versus late delivery of constraint therapy. *Neurorehabil Neural Repair*. 2013;27:654-663.
40. Kitago T, Liang J, Huang VS, et al. Improvement after constraint-induced movement therapy: recovery of normal motor control or task-specific compensation? *Neurorehabil Neural Repair*. 2013;27:99-109.
41. Prabhakaran S, Zarahn E, Riley C, et al. Inter-individual variability in the capacity for motor recovery after ischemic stroke. *Neurorehabil Neural Repair*. 2008;22:64-71.
42. Cramer SC. Repairing the human brain after stroke: I. Mechanisms of spontaneous recovery. *Ann Neurol*. 2008;63:272-287.
43. Jørgensen HS, Nakayama H, Raaschou HO, Olsen TS. Stroke. Neurologic and functional recovery the Copenhagen Stroke Study. *Phys Med Rehabil Clin N Am*. 1999;10:887-906.
44. Hankey GJ, Spiesser J, Hakimi Z, Bego G, Carita P, Gabriel S. Rate, degree, and predictors of recovery from disability following ischemic stroke. *Neurology*. 2007;68:1583-1587.
45. Duncan PW, Goldstein LB, Matchar D, Divine GW, Feussner J. Measurement of motor recovery after stroke. Outcome assessment and sample size requirements. *Stroke*. 1992;23:1084-1089.
46. Zeiler SR, Hubbard R, Gibson EM, et al. Paradoxical motor recovery from a first stroke after induction of a second stroke: reopening a postischemic sensitive period. *Neurorehabil Neural Repair*. 2016;30:794-800.
47. Zeiler SR, Krakauer JW. The interaction between training and plasticity in the poststroke brain. *Curr Opin Neurol*. 2013;26:609-616.
48. Stinear CM, Byblow WD. Predicting and accelerating motor recovery after stroke. *Curr Opin Neurol*. 2014;27:624-630.
49. Hardwick RM, Celnik PA. Non-invasive brain stimulation in physical medicine and rehabilitation. *Curr Phys Med Rehabil Rep*. 2014;2:300-309.
50. van Kordelaar J, van Wegen EEH, Nijland RHM, Daffertshofer A, Kwakkel G. Understanding adaptive motor control of the paretic upper limb early poststroke: the EXPLICIT-stroke program. *Neurorehabil Neural Repair*. 2013;27:854-863.