Robotic therapy for chronic stroke: general recovery of impairment or improved task-specific skill?

Tomoko Kitago,[1]* Jeff Goldsmith,[2]* Michelle Harran,[1] Leslie Kane,[1] Jessica Berard,[1] Sylvia Huang,[1] Sophia L. Ryan,[1,3] Pietro Mazzoni,[1] John W. Krakauer,[1,4] and Vincent S. Huang[1]

[1]Department of Neurology, Columbia University College of Physicians and Surgeons, New York, New York; [2]Department of Biostatistics, Mailman School of Public Health, New York, New York; [3]Department of Neurology, Massachusetts General Hospital, Boston, Massachusetts; [4]Departments of Neurology and Neuroscience, Johns Hopkins University, Baltimore, Maryland

Kitago T, Goldsmith J, Harran M, Kane L, Berard J, Huang S, Ryan SL, Mazzoni P, Krakauer JW, Huang VS. Robotic therapy for chronic stroke: general recovery of impairment or improved task-specific skill? *J Neurophysiol* 114: 1885–1894, 2015. First published July 15, 2015; doi:10.1152/jn.00336.2015.—There is a great need to develop new approaches for rehabilitation of the upper limb after stroke. Robotic therapy is a promising form of neurorehabilitation that can be delivered in higher doses than conventional therapy. Here we sought to determine whether the reported effects of robotic therapy, which have been based on clinical measures of impairment and function, are accompanied by improved motor control. Patients with chronic hemiparesis were trained for 3 wk, 3 days a week, with titrated assistive robotic therapy in two and three dimensions. Motor control improvements (i.e., skill) in both arms were assessed with a separate untrained visually guided reaching task. We devised a novel PCA-based analysis of arm trajectories that is sensitive to changes in the quality of entire movement trajectories without needing to pre-specify particular kinematic features. Robotic therapy led to skill improvements in the contralesional arm. These changes were not accompanied by changes in clinical measures of impairment or function. There are two possible interpretations of these results. One is that robotic therapy only leads to small task-specific improvements in motor control via normal skill-learning mechanisms. The other is that kinematic assays are more sensitive than clinical measures to a small general improvement in motor control.

stroke; neurorehabilitation; robotic therapy; kinematics; motor control

ROBOTIC THERAPY HAS EMERGED as a promising modality for stroke rehabilitation, as it has the advantage of being able to deliver therapy at a much higher intensity and dosage than conventional therapy. The largest study of robotic therapy conducted to date is the VA ROBOTICS study, a randomized control trial of robotic therapy of the upper limb for patients with chronic stroke using the MIT-Manus device (Lo et al. 2010). It was shown that, after 12 wk of therapy, the patients who received robotic training had slightly larger reductions in arm impairment compared with those who received usual care, but this difference in Fugl-Meyer Upper Extremity (FM) score was a clinically negligible 2.2 points. More recently, a study was conducted with a robot that allows three-dimensional (3D) movements compared with the two-dimensional (2D) planar movements of the MIT-Manus. As in the VA ROBOTICS study, there were significant but negligible effect sizes on the

FM score (Klamroth-Marganska et al. 2014). Several other smaller studies have also shown that patients with chronic stroke undergoing robotic therapy have similar, if not better, gains than conventional therapy (Housman et al. 2009; Liao et al. 2012; Lum et al. 2002; Page et al. 2012; Volpe et al. 2008), with mean FM changes ranging from 2–7 points. These results suggest that robotic therapy does have the potential to elicit improvements at the level of impairment, but the effect is consistently small (Kwakkel et al. 2008).

Although the FM is a widely used measure of motor impairment after stroke, it requires a combination of strength and motor control. This makes it difficult to distinguish between these two different aspects of movement, which may dissociate with respect to recovery time course and respond differently to training (Noskin et al. 2007). Here we distinguish motor control from strength in that, to be skillful, contractions of various muscle groups must be well coordinated. Thus we define motor control, or motor skill, as the ability to make accurate and precise, goal-directed movements without using compensatory movements (Kitago et al. 2013) or reducing movement speed (Reis et al. 2009; Shmuelof et al. 2012), and we define reduced motor control as the loss of this ability.

In the present study, we sought to identify improvements in movements that occur with robotic therapy by performing kinematic analysis on visually guided gravity-supported planar reaching movements, which allowed us to isolate changes in motor control from recovery of strength (Kitago et al. 2013). Kinematic analyses also offer more objective measures of motor performance, compared with clinical assessments such as the FM, which is subject to considerable variance (See et al. 2013). To date, very few robotics rehabilitation studies have used kinematic analysis to look at the effects of robotic training. Krebs and colleagues (1998) showed that patients with subacute stroke had improved ability to draw circles after robotic therapy. Kahn and colleagues (2006) reported that, after 24 sessions of robotic therapy in chronic stroke, patients had straighter reaching trajectories and made fewer submovements. Another study (Lum et al. 2002) showed that, after 2 mo of robotic training, patients had longer reaching extent, along with significant improvements in FM scores.

Another advantage of using kinematic analysis is that it is sensitive to potential changes in the "unaffected arm," which would go undetected by measures like the FM. It is well documented that the ipsilesional arm of patients with stroke also has deficits in motor control but without the confounding

---

* T. Kitago and J. Goldsmith contributed equally to this work.

Address for reprint requests and other correspondence: T. Kitago, 710 W. 168th St., New York, NY 10032 (e-mail: tk2229@columbia.edu).

effect of weakness (Noskin et al. 2007; Schaefer et al. 2007; Yarosh et al. 2004). Recent studies in animals have shown that skill training is not limb specific and has effects on both the trained and untrained sides. Repetitive use of the impaired limb in rats following unilateral ischemic stroke triggered reorganization of the cortex in the intact hemisphere (Barbay et al. 2013). Studies of motor learning in humans also show significant degrees of improvement in the untrained arm and hand (Grafton et al. 2002; Parlow and Kinsbourne 1989; Wiestler et al. 2014). Thus the unaffected arm can be viewed as a unique opportunity to assay for a motor-learning effect on the affected side.

To assay motor control, we use a gravity-supported, planar reaching task that minimizes both the requirement for antigravity strength and the use of compensatory strategies (Kitago et al. 2013). Unlike the robotic training protocol for which 3D training was employed, the 2D planar assay does not have joint angle redundancy amenable to compensation. This makes it an ideal choice for assessing motor control and can be considered the proximal analog of the finger individuation task (Kitago et al. 2013; Lang and Schieber 2004). By using this planar assay, we also sought to test motor control on a task separate from those performed during the training sessions. We applied a novel kinematic analysis to the visually guided reaching trajectories and compared this approach with standard functional and impairment assessments, the Action Research Arm Test (ARAT; Lyle 1981) and the FM, respectively.

The first objective of the present study was to investigate whether motor control in the affected arm changes in response to robotic training in patients with chronic stroke. The second objective was to investigate whether training the paretic limb caused changes in motor control of the ipsilesional (nonparetic) limb, which has not been previously examined in studies of robotic arm training for patients with chronic stroke.

METHODS

Study participants. Nine patients with chronic stroke were recruited from the outpatient rehabilitation clinics at Columbia University Medical Center between June 2009 and May 2011. Patients were included with the following criteria: *1*) ischemic stroke at least 6 mo before the start of the first assessment session, *2*) motor deficit of one arm with at least 20° wrist extension and 10° finger extension, *3*) ability to sit and be active for an hour on a chair without cardiac, respiratory, and/or pain disturbances. We chose to test patients with moderate impairment of the upper extremity because we were specifically interested in motor control changes rather than weakness, and we reasoned that severely impaired patients may have greater masking of motor control capacity by concomitant weakness. Patients were excluded if they *1*) were unable to understand and/or follow instructions, *2*) had pain in shoulder or arm (visual analog scale ≥4), *3*) had other neurological or musculoskeletal issues affecting the upper limb, *4*) were unable to give informed consent, or *5*) were under 18 yr of age. Kinematic data from 14 neurologically healthy control subjects (9 women, 5 men, mean age 60.8 ± 9.4 yr) recruited from the local community were used as the reference group for the kinematic analysis.

Ethics statement. The study was conducted in accordance with the principles expressed in the Declaration of Helsinki. All participants signed a written consent form that was approved by the Columbia University Medical Center Institutional Review Board. The study is registered with ClinicalTrials.gov (NCT02331407).

Intervention. Patients trained with their affected arm using a robotic device (ReoGo) (Fig. 1) that assisted them in various goal-directed

Fig. 1. ReoGo robotic device. Picture provided courtesy of Motorika Medical. Photograph by Eli Gross.

reaching exercises. They received training 3 days a week for 3 consecutive weeks. On each day of training, there were two training sessions, each averaging 1.5 h in duration, separated by a break. Each session consisted of the patient sitting next to the ReoGo system, a robotic guide that extended near the subject's arm. The patient's arm was then attached to the robotic guide through a brace with Velcro straps. On the first day of testing, patients were fitted specifically to the device. During the therapy, the patients controlled a cursor on a computer monitor by moving the robotic guide and were instructed to move the cursor to targets that were highlighted.

The therapy was conducted according to predetermined protocols for low- and medium-functioning patients. Patients who were only able to abduct their arm <30° were considered low functioning, and those who had between 30 and 60° of arm abduction were considered medium functioning. The exercises included goal-directed tasks that required reaching to targets in 2D (controlling anterior/posterior and medio/lateral directions) or 3D (which had the addition of a vertical movement component) space. Nine different tasks were performed during the training sessions: forward thrust, horizontal reaching, forward reaching in 2D and 3D, horizontal abduction, making a circle, reaching in a star pattern, reaching in a zigzag pattern, and mimicking bringing a cup to the mouth.

The ReoGo has the capacity to operate in five modes of interaction: *1*) Guided: the patient is 100% assisted by the robot but must be attentive to the passive movement; *2*) Initiated: the patient correctly initiates the movement and then is 100% assisted to complete the task; *3*) Step Initiated: the patient correctly initiates the movement, then is assisted for a short distance along a predefined trajectory, then is required to initiate the next segment of the trajectory; *4*) Follow Assist: the patient is allowed to move along the correct predetermined trajectory while the robot assists to prevent deviation from this trajectory; and *5*) Free: the patient completely controls the movement.

Separate protocols were followed for low- and medium-functioning patients, with prespecified tasks, number of repetitions, and degree of assistance for each session. Each week, the therapy protocol emphasized progressively more challenging movements, with an increase in

the number of arm movement repetitions and a decrease in the amount of assistance from the robotic device.

Study design. Clinical and kinematic assessments were performed at four time points: *1*) 3 wk before therapy, *2*) 1 wk before therapy, *3*) after completion of therapy, and *4*) 3 wk after therapy. The two assessments performed before the robotic intervention were to confirm that the patients had a stable baseline and to examine the contribution of practice effects, particularly for the planar reaching task.

Clinical outcome measures. All clinical assessments were performed by a single occupational therapist blinded to the patients' performance during therapy. The primary clinical outcome measures were the FM (Fugl-Meyer et al. 1975) and the ARAT (Lyle 1981). Both of these tests are frequently used and have been shown to have good reliability, validity, and responsiveness to motor change in patients with chronic stroke (Gladstone et al. 2002; Hsieh et al. 2009; Lang et al. 2006; Yozbatiran et al. 2008).

The FM is a measure of impairment that considers movement arm, wrist, hand, and coordination. Each of the 22 items is scored on a three-point ordinal scale for a maximum score of 66. The ARAT tests hand and arm function and consists of 19 items in 4 domains: grasp, grip, pinch, and gross movement. Each domain contains items arranged into hierarchical order of difficulty such that success at the most difficult item of a specific subclass assumes success for all items lower in the hierarchy of the same class. Each item is scored on a four-point ordinal scale with a maximum score of 57.

Kinematics and motor control. To assess the effect of robotic training on motor control of the upper extremity, patients were tested on an untrained planar reaching task, which has been previously described (Kitago et al. 2013). Subjects sat at a glass-surfaced table with their trunk securely belted to a high-backed chair. Table height was adjusted so that the shoulder and elbow were planar. The wrist, hand, and fingers were immobilized with a splint, which only allowed for movements of the shoulder and elbow. The forearm was supported using an air-sled system, which created a frictionless surface for movements. A mirror reflecting a computer display was placed just above the upper extremity so that the subject was unable to see his or her arm. Hand position was tracked in real time using Flock of Birds (Ascension Technology, Burlington, VT) magnetic movement recording system at a frequency of 120 Hz used to provide visual feedback.

Kinematic data from the hand, elbow, and shoulder were calculated and recorded using custom-written routines in RealBasic (Real Software, Austin, TX). The target set consisted of eight radially arrayed circles with a 1-cm radius, 45° apart, 8 cm from a center start circle. Each trial began after the subject held the cursor inside the start circle for 750 ms. Patients were instructed to make straight, out-and-back movements with a sharp reversal within the target. To ensure that movements were made quickly and to minimize online corrections, cursor feedback stopped after 200 ms, and the reversal point was indicated by a white square. Patients were given 1 or 2 practice runs of 88 movements for each arm to become familiar with the task. Patients then completed 2 experimental runs, each comprising 11 cycles of 8 targets, for each arm.

Data analysis. Hand position data were analyzed using custom routines in MATLAB (Mathworks, Natick, MA). Position time series were low pass filtered (Butterworth filter) at 8 Hz. The first velocity peak above a threshold of 10 cm/s was identified for each trial. This threshold was chosen to exclude small movements made by some patients who had difficulty stabilizing their hands within the start circle. The start of the movement was defined as either the point at which the velocity crossed 1 cm/s or the first velocity minimum before the first velocity peak above 10 cm/s, whichever was later. The end point of the outward movement was defined as the reversal point, that is, the point where distance from the origin stopped increasing.

The following types of movements were excluded from analysis: movements that did not reach 30% of the distance to the target,

movements without reversals, and spatial outliers (in which the movement direction was >90° from the target direction).

Functional principal component analysis. Traditional kinematic analysis of limb trajectories relies on examining specific kinematic variables, such as directional error, smoothness, and end-point accuracy. The choice of variables is dictated by the hypotheses of the study. In the case of recovery from stroke, the range of potential changes in limb trajectory during recovery is not known a priori. Although some specific variables have been shown to change (e.g., number of submovements) (Rohrer et al. 2002), these were preselected based on specific questions about recovery. It would be desirable to examine trajectories at a global level with a method that is sensitive to any changes in overall movement quality, without preselecting a long list of variables and incurring the risk of unnecessary multiple comparisons. We therefore devised a novel method of trajectory analysis, based on functional principal component analysis (FPCA) (Goldsmith et al. 2013; Yao et al. 2005), to characterize reaching kinematics. This approach extends principal component analysis to time series data. Its main benefit is that it examines the entire trajectory and has sensitivity for changes undetected by conventional analyses that focus only on preselected measures such as end-point accuracy or peak velocity. It should be emphasized that FPCA will detect and incorporate all these kinematic features anyway.

We represent kinematic data as $[X_i(t), Y_i(t)]$, where $X_i(t)$ and $Y_i(t)$ are the X and Y position of the hand at time t. FPCA expresses each motion as the combination of population-level components, selected to capture the major features of the kinematic data, and motion-specific weights or scores:

$$X_i(t) = \mu_X(t) + \sum_{k=1}^{3} c_{ik}^X \phi_k^X(t) \text{ and } Y_i(t) = \mu_Y(t) + \sum_{k=1}^{3} c_{ik}^Y \phi_k^Y(t) \quad (1)$$

Here $\mu_X(t)$ and $\mu_Y(t)$ are population mean functions, $\phi_k^X(t)$ and $\phi_k^Y(t)$ are shared components, and the c_{ik}^X and c_{ik}^Y are the motion-specific scores. The mean and shared components are estimated using all curves, and, given these, scores are estimated from individual trajectories. Interpretatively, $\phi_k^X(t)$ and $\phi_k^Y(t)$ give a data-driven summary of the directions that reaching motions differ in the population, and the scores c_{ik}^X and c_{ik}^Y quantify how these directions appear in a particular motion. Focusing on the scores c_{ik}^X and c_{ik}^Y effectively reduces the dimension of the kinematic data; three scores for X and Y suffice to explain >99% of the observed variance in the kinematic data.

The distribution of reaching trajectories can be understood through the distribution of FPCA scores; differences comparing groups are apparent in shifts in the mean or changes in the variance of the scores. We compute the squared Mahalanobis distances $[MD_i^2 = (c_i - \bar{c})^T \sum^{-1}(c_i - \bar{c})]$ for each score to measure the distance of each motion from the population average. MD^2 is computed with respect to a reference population, which in this case is a collection of reaching trajectories from the dominant arm of a group of age-matched healthy older adults using the same kinematic task. Intuitively, MD^2 is a generalization of the squared Z-score and quantifies the difference between motions made by healthy controls and subjects in this study. Subject-specific average squared Mahalanobis distances (AMD²) were computed for each subject at each target for each time point. FPCA analyses were carried out using R version 3.1.1 (R Development Core Team 2014).

Statistical analysis. For each of the outcome variables (FM, ARAT, AMD²), we used repeated-measure ANOVA to detect differences between adjacent time points assuming constant variance; separate analyses were conducted for the affected and unaffected arm. Our primary focus was comparing the difference between the first two pretests to the difference between the second pretest and the posttest.

For the kinematic variable (AMD²), observations were made at each of eight targets at all time points. There were several analysis options: *1*) to treat all targets as independent within subjects, *2*) to treat all targets as uniformly correlated within subjects, and *3*) to treat targets as uniformly correlated within subgroups of targets.

The first analysis is the least restrictive but may neglect any within-subject correlation. The second analysis is the most restrictive in the sense of the assumed correlation structure. The final analysis balances these by assuming a reasonable correlation structure within target subgroups, according to whether they required movements at single or multiple joints. Previous studies have shown that multijoint reaching movements are particularly difficult for patients with stroke because of the need to account for interaction torques at the shoulder and elbow (Beer et al. 2000; Cirstea et al. 2003). Targets where either shoulder or elbow joint excursion accounted for >70% of the combined joint excursion were considered single-joint targets, whereas the remainder were considered multijoint targets. Each analysis is conducted using a generalized estimating equation framework for the assumed correlation structure and target subgroup.

An a priori power analysis based on effect size of 3.8 and SD of 4.3 in previously reported FM changes in chronic stroke (Kitago et al. 2013) was conducted, and we determined that 9 subjects would be sufficient to yield an observed power of 76%. We also conducted a power analysis on the trajectory variables to assess the probability of detecting a true alternative hypothesis. Effect sizes, variances, and correlation across targets for the trajectory analysis were determined from the comparison of healthy subjects to affected patients with stroke (Kitago et al. 2012). Using these values and still assuming 9 subjects, our expected power to detect a training effect on AMD^2 of size 13.5 in the affected arm was 98.8%, assuming that targets are independent.

RESULTS

Table 1 shows a summary of the demographic and lesion characteristics of the nine patients who were enrolled in the study. All patients were able to complete the robotic therapy with no adverse events, and no patients were lost to follow-up. Over the 3 wk of therapy, both low- and medium-functioning patients performed an increasing number of repetitions with progressively less assistance from the robotic device and progressively more complex tasks, as seen in Fig. 2.

Robotic therapy had no demonstrable effect on clinical measures of impairment and function. There were no significant differences between the two baseline FM values (pretest 1 = 30.3, pretest 2 = 32.8; $P = 0.07$), nor were there any significant differences before and after training (pretest 2 = 32.8, posttest 1 = 36.0; $P = 0.07$) (Table 2). Similarly, there were no significant differences in the baseline ARAT measures (pretest 1 = 19.6, pretest 2 = 21.3; $P = 0.38$), nor were significant effects of training (pretest 2 = 21.3, posttest = 23.7; $P = 0.29$) (Table 2).

Movement times. Movement times in the affected arm decreased from pretest 2 to posttest 1 (mean difference −33 ms,

$P = 0.02$), but there was no significant change in movement times between pretest 1 to pretest 2 (mean difference −31 ms, $P = 0.11$) or between posttest 1 and posttest 2 (mean difference 8 ms, $P = 0.55$). In the unaffected arm, there was a significant decrease in movement times from pretest 1 to pretest 2 (mean difference −26 ms, $P = 0.01$) and between posttest 1 and posttest 2 (mean difference −15 ms, $P = 0.00001$), but there was no significant change in movement time with training between pretest 2 and posttest 1 (mean difference −2.9 ms, $P = 0.70$).

Reaching trajectories. Figure 3 shows the reaching trajectories from a representative patient for both the affected (trained) and unaffected (untrained) arms. Mean AMD^2 values for each testing session are shown in Fig. 4. Data for healthy control subjects are included in this figure for reference, but no statistical tests involving these values were conducted; the apparent heteroscedasticity comparing patients to controls does not affect the validity of our tests, which consider only within-group changes over time.

We analyzed the AMD^2 outcome using the three methods described above: *1*) independence across targets, *2*) uniform correlation across all targets, and *3*) uniform correlation in target subgroups (Fig. 5). The target subgroups we considered were single-joint and multijoint targets, defined by the degree of interjoint coordination used by healthy control subjects in reaching to these target directions. Decreases in AMD^2 indicate that trajectories are more similar to those of healthy controls and represent improvements in control.

A total of 251 movements [2% of total movements, 3.5% (range 0–10.7%) of affected arm movements and 0.5% (range 0–1%) of unaffected arm movements] were rejected by our prespecified criteria. We also performed an analysis including all movements, without rejections, and obtained similar results (not shown).

Reaching skill improved in the robotically trained arm. In the contralesional arm, there was a significant decrease in AMD^2 with training (from pretest 2 to posttest 1) assuming independence across targets (mean difference −16.31, $P = 0.0073$). There was a decrease assuming uniform correlation although this failed to reach significance (mean difference −16.31, $P = 0.077$). Within the single-joint target subgroup, the decrease in AMD^2 was significant (mean difference −24.28, $P = 0.030$) although, in the multijoint target subgroup, the decrease was not significant (mean difference −8.34, $P = 0.348$). Comparisons of the first and second pretest time points and of the first and second posttest time points were not significant under any analysis strategy. The former indicates that testing itself was not the cause of the improvements, and the latter indicates that the small robotic treatment effect on the control of reaching trajectories was sustained for 3 wk. In summary, there was a robust but small effect of robotic training on the control of visually guided reaches in the affected arm, which was more apparent with single-joint than multijoint reaches.

Reaching skill improved in the untrained arm with initial practice but not with robotic training. In the ipsilesional untrained arm, there was a significant decrease in AMD^2 between the first and second pretest assessments, assuming independence across targets (mean difference −3.05, $P = 0.0001$) and uniform correlation (mean difference −3.05, $P = 0.042$). Within the single-joint target subgroup, the decrease in

Table 1. *Description of patients*

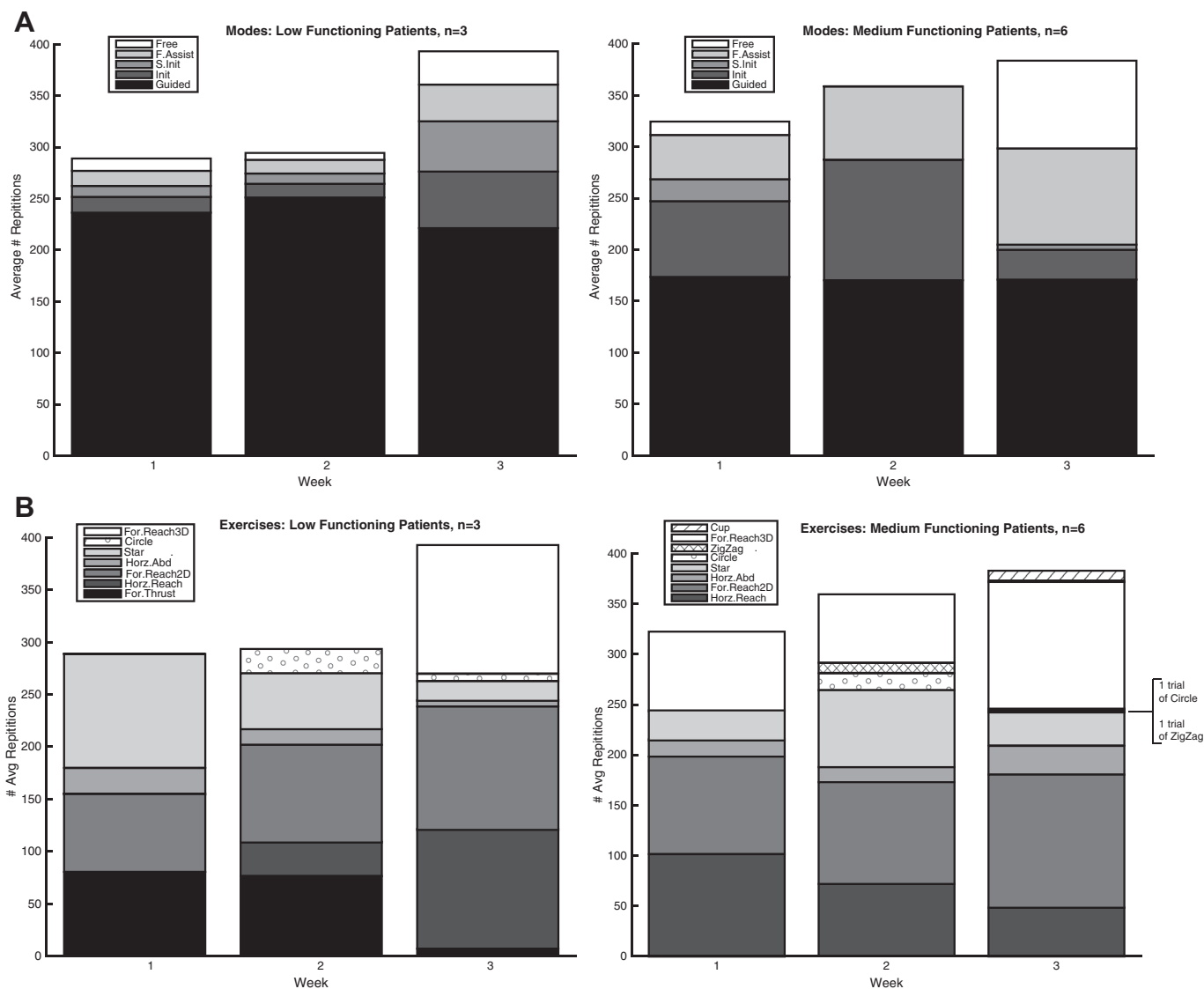| Subject	Sex	Age, yr	Months Since Stroke	Stroke Location	Stroke Side (Brain)
1	M	60	14	Pons	R
2	F	45	12	MCA (fronto-temporal-parietal)	L
3	M	79	8	Internal capsule and corona radiate	L
4	M	35	29	MCA	L
5	M	63	49	Pons	L
6	M	75	15	Internal capsule	L
7	M	73	7	MCA (parietal)	L
8	M	54	15	Internal capsule	L
9	M	47	9	Internal capsule	R

MCA, middle cerebral artery.

Fig. 2. Average number of repetitions per session, divided into the 5 modes of interaction (*A*) and the different exercises performed (*B*). Separate prespecified protocols were followed by low- and medium-functioning patients. Each week, there were progressively more challenging movements, with an increase in the number of repetitions, a decrease in the amount of assistance from the robotic device, and incorporation of more complex movements. Init, initiated; S. Init, step initiated; F. Assist, follow assist; For., forward; Horz., horizontal; Abd. abduction.

AMD^2 was significant (mean difference -3.17, $P = 0.015$) although, in the multi-joint target subgroup, the decrease was not significant (mean difference -2.93, $P = 0.11$). This pattern of improvement in the control of single-joint movements but not multijoint movements is similar to what we observed in the contralesional arm after robotic training. However, for the untrained arm, there was no significant change in AMD^2 with robotic training (from pretest 2 to posttest 1) or between the first and second posttest time points under any analysis strategy. In summary, the untrained "unaffected" arm showed improvement with initial practice during testing on the kinematic task but did not further benefit from robotic training despite not being at ceiling (the skill level of healthy control subjects) (Fig. 4).

DISCUSSION

We sought to detect an effect of robotic therapy on motor control in patients with chronic hemiparesis after stroke. We

found that there were improvements in the control of visually guided reaching trajectories in a planar task with the effects of gravity eliminated. There were no concomitant improvements in measures of impairment (FM) or function (ARAT) in the contralesional arm. These results illustrate that robotic training can lead to improvements in a motor skill in the absence of changes in standard clinical assessments.

The effects of robotic therapy on motor control in the current study were small but robust, with sustained improvements 3 wk later. Significant improvements were seen only for single-joint rather than multijoint movements. The same pattern of improvement was seen in the ipsilesional arm after practice of the task, suggesting that similar skill learning is occurring in both arms. This is interesting, as previous work, in healthy subjects and in patients with stroke, has shown that maintaining accuracy and path straightness is more challenging for multijoint compared with single-joint movements because of the need to compensate for interaction torques (Beer et al.

Table 2. *Clinical scores*

Subject	FM Scores				ARAT Scores			
	Pretest 1	Pretest 2	Posttest 1	Posttest 2	Pretest 1	Pretest 2	Posttest 1	Posttest 2
1	33	37	42	45	36	37	36	36
2	50	57	55	59	37	41	42	39
3	9	9	13	13	0	0	1	2
4	13	16	19	20	4	6	5	8
5	45	46	44	46	36	37	34	36
6	21	18	30	31	0	10	12	19
7	37	45	44	50	25	33	36	39
8	48	48	51	51	38	28	46	42
9	17	19	26	30	0	0	1	4
Mean	30.3	32.8	36.0	38.3	19.6	21.3	23.7	25.0
SD	15.8	17.4	14.6	15.5	18.0	17.1	18.6	16.7

FM, Fugl-Meyer Upper Extremity; ARAT, Action Research Arm Test.

2000; Cirstea et al. 2003). Thus it appears that, after stroke, anisotropies attributable to limb dynamics are accentuated.

These improvements in motor control occurred without any significant increase in movement time. Improved trajectory accuracy for movements of the same or higher speed represents an improvement in the speed-accuracy tradeoff. The observed improvements thus imply that a form of motor skill learning has taken place (Reis et al. 2009; Shmuelof et al. 2012). There was no ceiling effect for the planar task, as healthy controls performed markedly better than the patients did even with their ipsilesional arm.

Previous studies of robotic therapy in patients with chronic hemiparesis have shown small effect sizes at the impairment level and none at the functional level (Kwakkel et al. 2008; Lo et al. 2010; Prange et al. 2006). Here we report significant but small effects only at the level of motor control. The magnitude in improvement in the FM in our study was comparable to what was found in two larger studies of robotic arm training in

Fig. 3. Reaching trajectories from a representative subject, before and after training for the affected (*top*) and unaffected (*bottom*) arms.

patients with chronic stroke, the VA ROBOTICS study (Lo et al. 2010) (3.9 points) and the recently published 3D robotics study (Klamroth-Marganska et al. 2014) (3.3 points). Although the lack of significant improvements in our clinical measures may be related to our small sample size, the absolute changes with training were small (3.2 points on the FM and 2.4 points on the ARAT) and unlikely to be clinically meaningful (Gladstone et al. 2002; Van der Lee et al. 2001). Reaching kinematics were not assessed in either of these aforementioned studies, therefore it is not possible to say whether the small improvements seen at the impairment level were accompanied by improvements in motor control.

The absence of significant effects with the ARAT and the FM could be interpreted to mean that the improvements in motor control were not large enough to generalize to either of these measures. That is to say, the motor control assay may be more sensitive to small effects of robotic training. Furthermore, the movements that are tested in ARAT and FM involve a combination of motor control, strength, and the use of compensatory strategies. Thus an improvement in motor control alone may not be sufficient to lead to improvements in these clinical measures.

Our study population encompassed a wide range of severity, with baseline scores ranging from 9 to 50 for the FM and from 0–38 for the ARAT. We explored the relationship between baseline clinical scores and the amount of improvement with training for the affected arm. Subjects with lower FM scores tended to have larger improvements with training, but, for the ARAT, subjects with higher scores tended to improve more with training although these correlations were nonsignificant. This study was not powered for subgroup analyses based on the severity of the motor deficits, but we found little evidence that the heterogeneity in our population was masking an improvement in the lower or higher functioning groups.

For the ipsilesional, untrained arm, no further improvement was seen after the 3 wk of training with the affected arm despite a persistent impairment compared with healthy control subjects. This lack of generalization to the untrained arm does not necessarily signify that motor learning has not occurred for the trained arm. Previous studies that have demonstrated improvements in both trained and untrained limbs have not examined the quality of the movements as we did in this study, with the exception of one primate study in which the animals were trained and tested with a planar reaching task (Georgo-
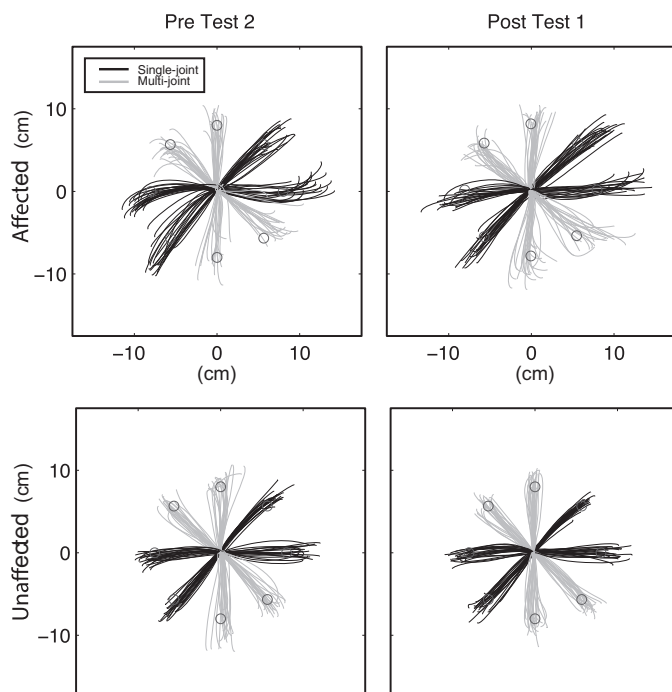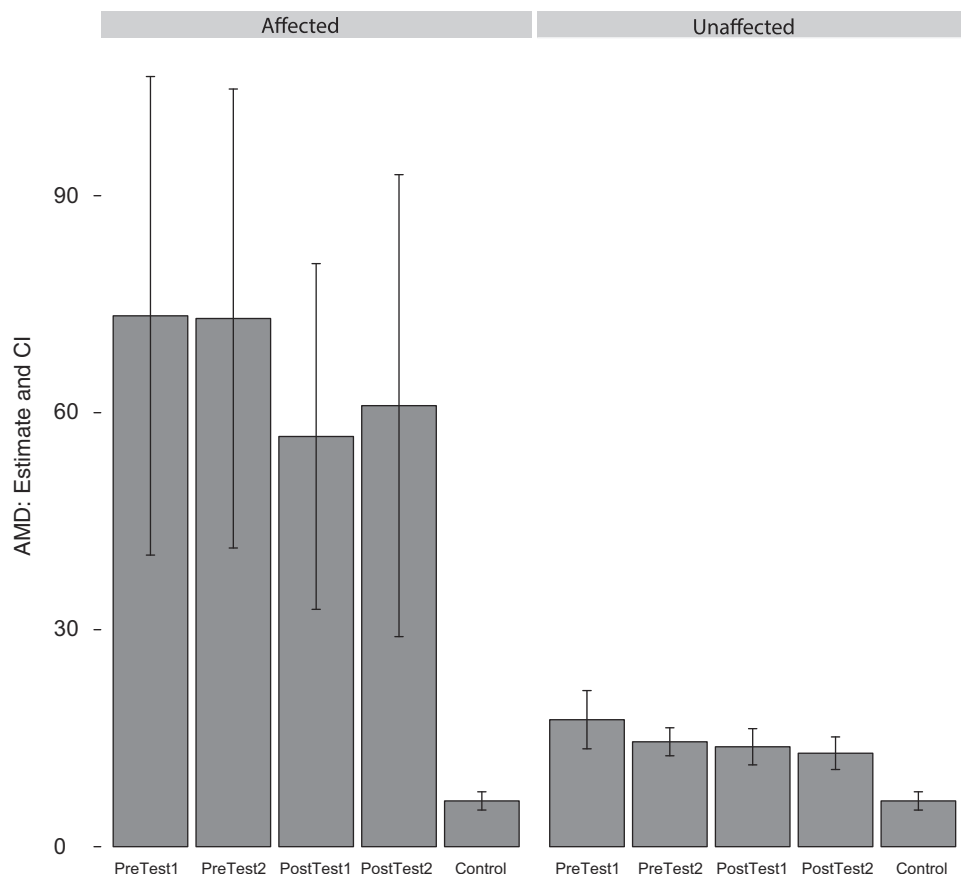
Fig. 4. Average squared Mahalanobis distances (AMD) ($\pm$ confidence intervals) at each time point, for the affected and unaffected arms. Values from a reference population of healthy control subjects are presented for comparison. Even with the unaffected arm, patients do not reach the level of performance of the healthy control subjects.

poulos et al. 1981). It may be that the magnitude of motor skill improvement in our study was not large enough to generalize to the untrained side or that, in patients with chronic stroke, the type of motor skill learning we tested is limb and task specific (Bavelier et al. 2012). Indeed, it is the assumption that rehabilitation is based on motor learning coupled with task specificity that has led to the emphasis placed on task-oriented

training after stroke (Bayona et al. 2005; French et al. 2010; Hubbard et al. 2009; Schaefer et al. 2013).

The critical question is whether task-oriented training after stroke is merely exploiting normal learning mechanisms to increase skill within the performance envelope available to the patient. This would be comparable to a healthy person getting better at handwriting with their nondominant arm; it has a
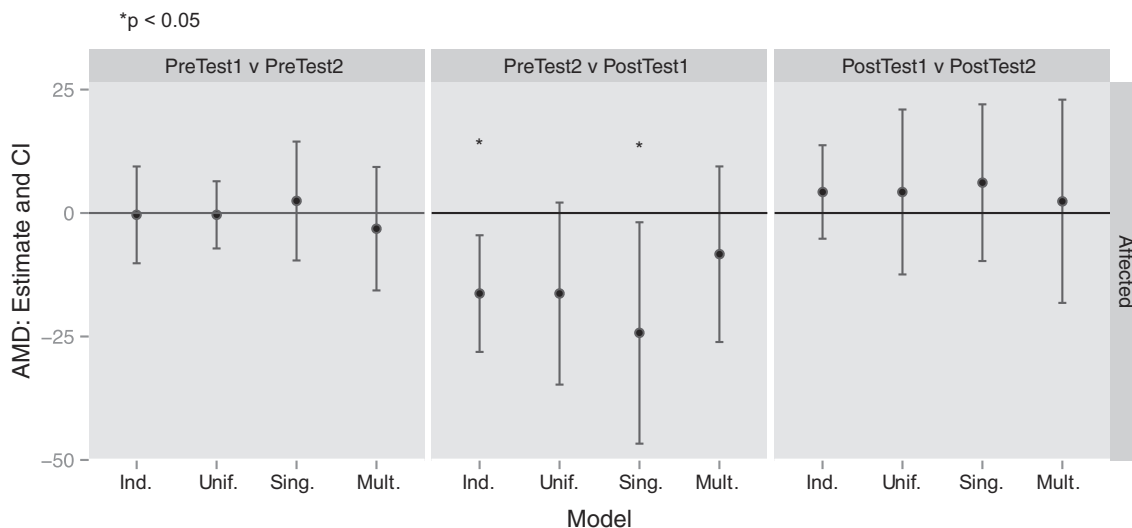

Fig. 5. Changes in average squared Mahalanobis distances ($\pm$ confidence intervals) across time points for 3 methods: assuming independence across targets (Ind.), assuming uniform correlation across targets (Unif.), and assuming correlations within subgroups for single-joint (Sing.) and multijoint (Mult.) targets. Negative values represent an improvement in reaching kinematics. The only changes that reach statistical significance (*) occur between pretest 2 and posttest 1, for the affected/trained arm.

baseline level of performance that can be improved with practice. The core point is that healthy subjects and patients alike can augment performance to some degree on any task with practice. There is no need to invoke repair or reorganization. Thus the fundamental question raised by our results is whether the robotic training led to an increase in task-specific skill or to a general improvement in motor control. One possibility is that there are 2D task elements in the robotic training that are similar to the movements in the planar assay, which would mean that improvements that we observed can be attributed to task-specific skill. Alternatively, as mentioned already, the training could have had a more general effect on motor control, but it was too small to either generalize or be detected by the ARAT and FM. In the VA ROBOTICS study, there was a small effect of planar training on the FM, which might be taken as evidence for generalization from 2D reaches with gravity eliminated to 3D multijoint movements.

The ambiguity in ascertaining precisely what robotic therapy accomplishes arises because of the inherent difficulties of using task A (planar assay) to assess the effect of training with task B (robot). Task A can be viewed both as an assay for general motor control or a specific task that one can become more skilled at through training. As a hypothetical example, imagine two patients with stroke with hemiparesis. The first patient has a hemiparesis of moderate severity and is assessed on the planar reaching task before any rehabilitation is given. The second patient has more severe arm impairment and receives 2 wk of robotic training. This second patient is then assessed posttraining with the planar reaching task and is found to have exactly the same level of performance on it as the first patient. Thus the two patients now look phenotypically identical on the planar reaching task, but one patient required robotic training and the other did not. Are they phenotypically identical because the robotic training has partially reversed the second patient's more severe impairment, or has the robotic training only made the second patient a little more skilled at the planar task but is otherwise unchanged? This emphasizes the often unappreciated fact that pretraining correlation does not imply training covariation (Moreau and Conway 2014).

Ideally, one would have a battery of preestablished compensation-proof tasks with some a priori framework for how they differ from each other and from the training task. The question, however, of generalization of task-specific training is a vexed one because it presupposes that we have an a priori definition of task. If the degree of similarity between task A and task B is established via a transfer assay, then that same assay cannot be taken as evidence for generalization in the extrapolative sense because transfer occurred only to the degree that there was already overlap between tasks A and B within the part of parameter space trained. Thus transfer of learning across tasks can only be considered generalization if there is a separate assay for task similarity. In a recent study that looked at training on three tasks for the arm and hand, kinematics of the proximal limb was used as the measure of similarity across tasks. Interestingly, kinematic similarity did not predict the degree of transfer. The authors postulated that goal similarity might be the feature that matters instead. Goal, however, is a fuzzy concept; at one level, the goal is always to complete the task successfully, but successful task completion necessarily depends on the specifics of the task and its context. For example, in recent studies of split-treadmill adaptation after stroke, the poor generalization to overground walking has been taken as evidence for a context effect preventing transfer of what is a very similar locomotor pattern on and off the treadmill (Reisman et al. 2009; Torres-Oviedo and Bastian 2012). Finally, it needs to be recognized that rehabilitative training can lead to increased cardiovascular fitness and muscular conditioning, in addition to neural changes. The very small amount of generalization to the FM scale seen in the VA ROBOTICS study (Lo et al. 2010) could be attributable to strength increase in the proximal limb rather than motor-learning effects.

On balance we would tentatively conclude that the improvement in motor control in the planar reaching task after 3 wk of robotic therapy is because of some skill improvement at the task attributable to overlap with some of the planar components of robotic training. The lack of significant improvement in the FM and ARAT further supports our interpretation of a small increment in task-specific skill.

We have recently shown that constraint-induced movement therapy (CIMT) seems to promote functional recovery in chronic stroke mainly, if not exclusively, through compensation. Specifically, we showed that there was no discernible improvement in motor control, using the same planar reaching task as used in the current study, despite a clinically meaningful change in the ARAT (Kitago et al. 2013). Thus for CIMT the results were the opposite of those obtained here for robotic therapy, which suggests a dissociation for these two rehabilitation approaches with respect to gains in motor control vs. function, the latter presumably driven by compensation. This conclusion is not inconsistent with previously reported results, which overall suggest that robotic therapy exerts its small effects at the impairment level (Huang and Krakauer 2009; Kwakkel et al. 2008; Lo et al. 2010; Prange et al. 2006), whereas CIMT exerts its effects at the functional level (Kitago et al. 2013; Wolf et al. 2006).

In this study, patients with chronic hemiparesis who underwent 3 wk of 2D and 3D robotic training showed small but robust improvements in the control of reaching trajectories, which were assayed with an untrained planar task. The same pattern of skill improvement with practice was seen on the untrained ipsilesional side. That patients can show small increases in skill on any given task is hardly surprising, as to date there is no evidence that skill-learning mechanisms are impaired after stroke; hemiparesis is a motor control deficit (Krakauer 2006; Winstein et al. 1999). We conclude that training can eke out small task-specific performance improvements in motor control in patients after stroke through normal motor-learning mechanisms, which cannot reverse damage to the corticospinal tract or trigger reorganization. Generalization will only occur to the degree that the tasks share parts of command space. The much larger and much more general gains seen early after stroke are due to unique plasticity conditions in a limited time window that interact with training to enable reorganization and repair processes that are qualitatively and quantitatively distinct from normal learning (Zeiler and Krakauer 2013). Robotic therapy initiated within this time window might be able to augment such spontaneous biological recovery.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

Author contributions: T.K., L.K., S.H., S.L.R., J.W.K., and V.S.H. conception and design of research; T.K., J.G., M.H., J.B., and V.S.H. analyzed data; T.K., J.G., M.H., J.B., P.M., and V.S.H. interpreted results of experiments; T.K., J.G., M.H., and J.B. prepared figures; T.K. and J.W.K. drafted manuscript; T.K., J.G., J.B., P.M., J.W.K., and V.S.H. edited and revised manuscript; T.K., J.G., M.H., P.M., J.W.K., and V.S.H. approved final version of manuscript; L.K., S.H., S.L.R., and V.S.H. performed experiments.

REFERENCES

Barbay S, Guggenmos DJ, Nishibe M, Nudo RJ. Motor representations in the intact hemisphere of the rat are reduced after repetitive training of the impaired forelimb. *Neurorehab Neural Repair* 27: 381–384, 2013.

Bavelier D, Green CS, Pouget A, Schrater P. Brain plasticity through the life span: learning to learn and action video games. *Annu Rev Neurosci* 35: 391–416, 2012.

Bayona NA, Bitensky J, Salter K, Teasell R. The role of task-specific training in rehabilitation therapies. *Top Stroke Rehabil* 12: 58–65, 2005.

Beer RF, Dewald JP, Rymer WZ. Deficits in the coordination of multijoint arm movements in patients with hemiparesis: evidence for disturbed control of limb dynamics. *Exp Brain Res* 131: 305–319, 2000.

Cirstea MC, Mitnitski AB, Feldman AG, Levin MF. Interjoint coordination dynamics during reaching in stroke. *Exp Brain Res* 151: 289–300, 2003.

French B, Thomas L, Leathley M, Sutton C, McAdam J, Forster A, Langhorne P, Price C, Walker A, Watkins C. Does repetitive task training improve functional activity after stroke? A Cochrane systematic review and meta-analysis. *J Rehab Med* 42: 9–14, 2010.

Fugl-Meyer AR, Jaasko L, Leyman I, Olsson S, Steglind S. The post-stroke hemiplegic patient. 1. A method for evaluation of physical performance. *Scand J Rehabil Med* 7: 13–31, 1975.

Georgopoulos AP, Kalaska JF, Massey JT. Spatial trajectories and reaction times of aimed movements: Effects of practice, uncertainty, and change in target location. *J Neurophysiol* 46: 725–743, 1981.

Gladstone DJ, Danells CJ, Black SE. The Fugl-Meyer assessment of motor recovery after stroke: a critical review of its measurement properties. *Neurorehabil Neural Repair* 16: 232–240, 2002.

Goldsmith J, Greven S, Crainiceanu CM. Corrected confidence bands for functional data using principal components. *Biometrics* 69: 41–51, 2013.

Grafton ST, Hazeltine E, Ivry RB. Motor sequence learning with the nondominant left hand. A PET functional imaging study. *Exp Brain Res* 146: 369–378, 2002.

Housman SJ, Scott KM, Reinkensmeyer DJ. A randomized controlled trial of gravity-supported, computer-enhanced arm exercise for individuals with severe hemiparesis. *Neurorehabil Neural Repair* 23: 505–514, 2009.

Hsieh YW, Wu CY, Lin KC, Chang YF, Chen CL, Liu JS. Responsiveness and validity of three outcome measures of motor function after stroke rehabilitation. *Stroke* 40: 1386–1391, 2009.

Huang VS, Krakauer JW. Robotic neurorehabilitation: a computational motor learning perspective. *J Neuroeng Rehabil* 6: 5, 2009.

Hubbard IJ, Parsons MW, Neilson C, Carey LM. Task-specific training: evidence for and translation to clinical practice. *Occup Ther Int* 16: 175–189, 2009.

Kahn LE, Zygman ML, Rymer WZ, Reinkensmeyer DJ. Robot-assisted reaching exercise promotes arm movement recovery in chronic hemiparetic stroke: A randomized controlled pilot study. *J Neuroeng Rehabil* 3: 12, 2006.

Kitago T, Goldsmith J, Haith A, Huang VS, Crainiceanu CM, Mazzoni P, Krakauer JW. The motor control deficit after unilateral stroke qualitatively mimics the dominant/non-dominant difference. In: *Annual Meeting of the Society for Neuroscience*. New Orleans, LA: Society for Neuroscience, 2012.

Kitago T, Liang J, Huang VS, Hayes S, Simon P, Tenteromano L, Lazar RM, Marshall RS, Mazzoni P, Lennihan L, Krakauer JW. Improvement after constraint-induced movement therapy: recovery of normal motor control or task-specific compensation? *Neurorehabil Neural Repair* 27: 99–109, 2013.

Klamroth-Marganska V, Blanco J, Campen K, Curt A, Dietz V, Ettlin T, Felder M, Fellinghauer B, Guidali M, Kollmar A, Luft A, Nef T, Schuster-Amft C, Stahel W, Riener R. Three-dimensional, task-specific robot therapy of the arm after stroke: a multicentre, parallel-group randomised trial. *Lancet Neurol* 13: 159–166, 2014.

Krakauer JW. Motor learning: its relevance to stroke recovery and neurorehabilitation. *Curr Opin Neurol* 19: 84–90, 2006.

Krebs HI, Hogan N, Aisen ML, Volpe BT. Robot-aided neurorehabilitation. *IEEE Trans Rehabil Eng* 6: 75–87, 1998.

Kwakkel G, Kollen BJ, Krebs HI. Effects of robot-assisted therapy on upper limb recovery after stroke: a systematic review. *Neurorehabil Neural Repair* 22: 111–121, 2008.

Lang CE, Schieber MH. Reduced muscle selectivity during individuated finger movements in humans after damage to the motor cortex or corticospinal tract. *J Neurophysiol* 91: 1722–1733, 2004.

Lang CE, Wagner JM, Dromerick AW, Edwards DF. Measurement of upper-extremity function early after stroke: properties of the action research arm test. *Arch Phys Med Rehabil* 87: 1605–1610, 2006.

Liao WW, Wu CY, Hsieh YW, Lin KC, Chang WY. Effects of robot-assisted upper limb rehabilitation on daily function and real-world arm activity in patients with chronic stroke: a randomized controlled trial. *Clin Rehabil* 26: 111–120, 2012.

Lo AC, Guarino PD, Richards LG, Haselkorn JK, Wittenberg GF, Federman DG, Ringer RJ, Wagner TH, Krebs HI, Volpe BT, Bever CT Jr, Bravata DM, Duncan PW, Corn BH, Maffucci AD, Nadeau SE, Conroy SS, Powell JM, Huang GD, Peduzzi P. Robot-assisted therapy for long-term upper-limb impairment after stroke. *N Engl J Med* 362: 1772–1783, 2010.

Lum PS, Burgar CG, Shor PC, Majmundar M, Van der Loos M. Robot-assisted movement training compared with conventional therapy techniques for the rehabilitation of upper-limb motor function after stroke. *Arch Phys Med Rehabil* 83: 952–959, 2002.

Lyle RC. A performance test for assessment of upper limb function in physical rehabilitation treatment and research. *Int J Rehabil Res* 4: 483–492, 1981.

Moreau D, Conway AR. The case for an ecological approach to cognitive training. *Trends Cogn Sci* 18: 334–336, 2014.

Noskin O, Krakauer JW, Lazar RM, Festa JR, Handy C, O'Brien KA, Marshall RS. Ipsilateral motor dysfunction from unilateral stroke: Implications for the functional neuroanatomy of hemiparesis. *J Neurol Neurosurg Psychiatry* 79: 401–406, 2008.

Page SJ, Hill V, White S. Portable upper extremity robotics is as efficacious as upper extremity rehabilitative therapy: a randomized controlled pilot trial. *Clin Rehabil* 27: 494–503, 2013.

Parlow SE, Kinsbourne M. Asymmetrical transfer of training between hands: implications for interhemispheric communication in normal brain. *Brain Cogn* 11: 98–113, 1989.

Prange GB, Jannink MJ, Groothuis-Oudshoorn CG, Hermens HJ, Ijzerman MJ. Systematic review of the effect of robot-aided therapy on recovery of the hemiparetic arm after stroke. *J Rehabil Res Dev* 43: 171–184, 2006.

R Development Core Team. *R*: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, 2014.

Reis J, Schambra HM, Cohen LG, Buch ER, Fritsch B, Zarahn E, Celnik PA, Krakauer JW. Noninvasive cortical stimulation enhances motor skill acquisition over multiple days through an effect on consolidation. *Proc Natl Acad Sci USA* 106: 1590–1595, 2009.

Reisman DS, Wityk R, Silver K, Bastian AJ. Split-belt treadmill adaptation transfers to overground walking in persons poststroke. *Neurorehabil Neural Repair* 23: 735–744, 2009.

Rohrer B, Fasoli S, Krebs HI, Hughes R, Volpe B, Frontera WR, Stein J, Hogan N. Movement smoothness changes during stroke recovery. *J Neurosci* 22: 8297–8304, 2002.

Schaefer SY, Haaland KY, Sainburg RL. Ipsilesional motor deficits following stroke reflect hemispheric specializations for movement control. *Brain* 130: 2146–2158, 2007.

Schaefer SY, Patterson CB, Lang CE. Transfer of training between distinct motor tasks after stroke: implications for task-specific approaches to upper-

extremity neurorehabilitation. *Neurorehabil Neural Repair* 27: 602–612, 2013.

See J, Dodakian L, Chou C, Chan V, McKenzie A, Reinkensmeyer DJ, Cramer SC. A standardized approach to the Fugl-Meyer assessment and its implications for clinical trials. *Neurorehabil Neural Repair* 27: 732–741, 2013.

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* 108: 578–594, 2012.

Torres-Oviedo G, Bastian AJ. Natural error patterns enable transfer of motor learning to novel contexts. *J Neurophysiol* 107: 346–356, 2012.

Van der Lee JH, De Groot V, Beckerman H, Wagenaar RC, Lankhorst GJ, Bouter LM. The intra- and interrater reliability of the action research arm test: a practical test of upper extremity function in patients with stroke. *Arch Phys Med Rehabil* 82: 14–19, 2001.

Volpe BT, Lynch D, Rykman-Berland A, Ferraro M, Galgano M, Hogan N, Krebs HI. Intensive sensorimotor arm training mediated by therapist or robot improves hemiparesis in patients with chronic stroke. *Neurorehabil Neural Repair* 22: 305–310, 2008.

Wiestler T, Waters-Metenier S, Diedrichsen J. Effector-independent motor sequence representations exist in extrinsic and intrinsic reference frames. *J Neurosci* 34: 5054–5064, 2014.

Winstein CJ, Merians AS, Sullivan KJ. Motor learning after unilateral brain damage. *Neuropsychologia* 37: 975–987, 1999.

Wolf SL, Winstein CJ, Miller JP, Taub E, Uswatte G, Morris D, Giuliani C, Light KE, Nichols-Larsen D. Effect of constraint-induced movement therapy on upper extremity function 3 to 9 mo after stroke: the EXCITE randomized clinical trial. *JAMA* 296: 2095–2104, 2006.

Yao F, Muller HG, Wang JL. Functional data analysis for sparse longitudinal data. *J Am Stat Assoc* 100: 577–590, 2005.

Yarosh CA, Hoffman DS, Strick PL. Deficits in movements of the wrist ipsilateral to a stroke in hemiparetic subjects. *J Neurophysiol* 92: 3276–3285, 2004.

Yozbatiran N, Der-Yeghiaian L, Cramer SC. A standardized approach to performing the action research arm test. *Neurorehabil Neural Repair* 22: 78–90, 2008.

Zeiler SR, Krakauer JW. The interaction between training and plasticity in the poststroke brain. *Curr Opin Neurol* 26: 609–616, 2013.