6th IEEE RAS/EMBS International Conference on
Biomedical Robotics and Biomechatronics (BioRob)
June 26-29, 2016. UTown, Singapore

Improving the Retention of Motor Skills after Reward-Based Reinforcement by Incorporating Haptic Guidance and Error Augmentation

Dylan P. Losey, *Student Member, IEEE*, Laura H. Blumenschein *Student Member, IEEE*, and Marcia K. O'Malley, *Senior Member, IEEE*

***Abstract*— There has been significant research aimed at leveraging programmable robotic devices to provide haptic assistance or augmentation to a human user so that new motor skills can be trained efficiently and retained long after training has concluded. The success of these approaches has been varied, and retention of skill is typically not significantly better for groups exposed to these controllers during training. These findings point to a need to incorporate a more complete understanding of human motor learning principles when designing haptic interactions with the trainee. Reward-based reinforcement has been studied for its role in improving retention of skills. Haptic guidance, which assists a user to complete a task, and error augmentation, which exaggerates error in order to enhance feedback to the user, have been shown to be beneficial for training depending on the task difficulty, subject ability, and task type. In this paper, we combine the presentation of reward-based reinforcement with these robotic controllers to evaluate their impact on retention of motor skill in a visual rotation task with tunable difficulty using either fixed or moving targets. We found that with the reward-based feedback paradigm, both haptic guidance and error augmentation led to better retention of the desired visuomotor offset during a simple task, while during a more complex task, only subjects trained with haptic guidance demonstrated performance superior to those trained without a controller.**

I. Introduction

Motor learning relates to motor skill acquisition [1], and the process of making accurate goal-directed movements [2]. Because of continuous changes in the body and environment, as well as the delayed and noisy nature of sensory feedback, motor learning is traditionally thought to be based upon adaptive internal predictions of input/output relationships [2]. For this reason, motor learning is central to neurorehabilitation, particularly in the first few months after injury; during this period, subjects may relearn altered sensorimotor mappings before developing novel compensatory strategies [3]. Of course, motor learning—and its underlying principles—should also be considered when designing control algorithms for robotic rehabilitation after a neurological injury, since these controllers seek to promote the subject's recovery by facilitating learning of desired motions [4]. As a result, examination of contemporary research on both motor learning and robotic control suggests means with which to enhance neurorehabilitation.

Recent studies have revealed that multiple distinct processes are likely responsible for motor learning. Huang *et al*. [5] hypothesized that fast-adapting internal models and slower improvement via model-free memories comprise motor learning, and experimentally demonstrated that combining movement repetition with reward-based adaption is key to retention. Similarly, Izawa and Shadmehr [6] found that reward prediction errors—-qualitative measures of a motion's utility—and model prediction errors, i.e., differences between expected and actual experiences during the movement, cause disparate types of learning. Both results were supported by Shmuelof *et al*. [7], who introduced a period of reward-based reinforcement after the model adaption process so as to prevent subjects from rapidly forgetting a novel motor mapping. Furthermore, Galea *et al*. [8] discovered that positive and negative reinforcement independently affect motor learning; punishments induced faster adaption, but rewards increased retention of the acquired behavior. To summarize, it seems that reward-based reinforcement may aid in ingraining model-free memories, desirably leading to better recall of learned motor commands.

Alternatively, controllers for rehabilitation robots frequently attempt to incite motor learning through one of two opposite interaction strategies: haptic guidance, where convergent forces help the subject accurately complete motions, or error augmentation, where divergent forces make the subject's mistakes more pronounced. Since haptic guidance and error augmentation render movements easier or harder, respectively, they may be implemented according to challenge point theory [9], which dictates that optimal learning occurs when task difficulty is suited to participant proficiency. Research on motion timing with healthy young adults [10] and elderly subjects [11] validated this concept, as less-skilled participants learned more using haptic guidance, while, conversely, better-skilled subjects benefited from error augmentation. Marchal-Crespo *et al*. [12] argue that even task characteristics, such as rhythmicity and duration, manipulate motor learning during robot-assisted training; likewise, Heuer and Lüttgen [13] detail a taxonomy of learning tasks, and then review the theoretical and experimental influences of haptic guidance and error augmentation within each category. Hence, to best promote motor learning under robotic rehabilitation, we should consider subject ability, task difficulty, and task type when determining the level of controller assistance.

A primary goal of robotic rehabilitation is to ensure subjects retain the learned movements. With robotic intervention, motions may be successfully completed; however, once the device is removed, users often revert to baseline behavior

This work was funded in part by grants NSF GRFP-0940902 and CNS-1135916.

The authors are with the Mechatronics and Haptic Interfaces Laboratory, Department of Mechanical Engineering, Rice University, Houston, TX 77005. (e-mail: dlosey@rice.edu)

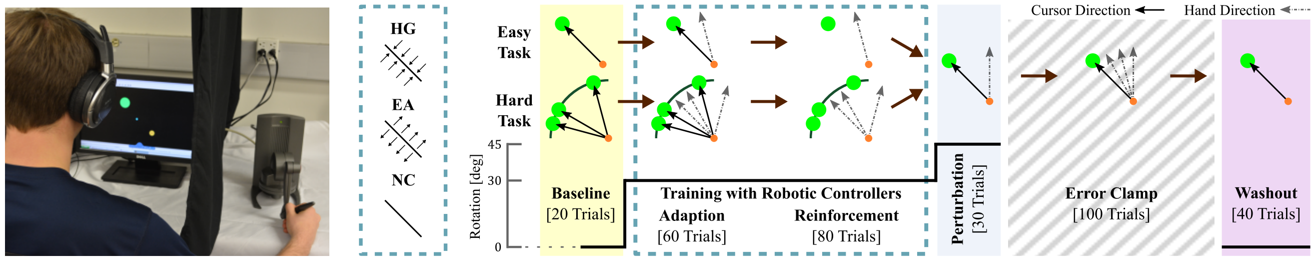

Fig. 1. Experimental setup. (Left) Experimental apparatus, where a user grasps the haptic device while observing visual feedback on a computer monitor. (Right) Experimental protocol, where subjects perform 330 reaching trials under a series of visuomotor rotations: baseline ($0°$), training ($30°$), perturbation ($45°$), error clamp, and washout ($0°$). During the experiment, participants were assigned either Easy or Hard tasks while training with HG, EA, or NC. Both levels of task difficulty are shown above; for the Easy task, the target's position was constant, but for the Hard task, a target was randomly placed in the second quadrant. Note that the Easy and Hard tasks are identical after the training block. Force fields associated with HG, EA, and NC are depicted within the dashed box; HG helps users maintain the desired path, whereas EA exacerbates deviations. Controllers were only applied during training.

[13]. For example, Patton *et al*. [14] found that 600 training movements with error augmentation only produced desired outcomes for 30-60 unimpeded movements. By contrast, Hasson *et al*. [15] recently implemented a reinforcement scheme—without robotic platforms—which improved overnight retention of an unnatural gait pattern. We are here interested in applying reward-based reinforcement together with different control strategies for robotic rehabilitation: what, if any, impact will the combination of these two techniques have on the recall of motor learning? In this paper, we investigate the relative effects of controller strategies on reinforcement learning by performing experiments similar to those described in [7], while additionally incorporating haptic guidance, error augmentation, and variable task difficulty. We found that both haptic guidance and error augmentation led to better retention of the desired visuomotor offset during a simple task, while during a more complex task only subjects trained with haptic guidance improved performance as compared to those trained without a controller.

II. Methods

Subjects. Twenty-eight healthy, right-handed individuals (aged 22.6 ± 4.5 years, 8 females) were recruited for this study. All participants were naïve to the experiment's purpose, received course credit, and signed a written consent form approved by the Rice University Institutional Review Board. Twenty-four subjects were randomly assigned to one of three equal-sized experimental groups: haptic guidance (HG), error augmentation (EA), or no controller (NC). The four additional subjects were recruited to the HG group after anomalous data were observed for four of the original HG subjects (see later discussion of treatment of outliers). Subjects completed two separate experiments, one with the Easy task and one with the Hard task. These experiments were spaced apart by a minimum of three days; half of the participants completed an experiment with the Easy task first.

Experimental Apparatus. Subjects sat at a table while grasping the end-effector of a Touch X haptic robot (Geomagic), which was used to track hand motions and apply controller forces at a frequency of 1 kHz. Participants wore a wrist brace to restrict motion, and were instructed to move by sliding their arm across the table. A curtain occluded view of the arm and hand; visual feedback corresponding to movement of the Touch X end-effector was instead provided on a computer monitor placed directly in front of the subjects. The visualization updated at a rate of 10 fps, and could faithfully depict motions with a $1:1$ ratio. Both visualization and robot were controlled using Matlab/Simulink (Mathworks).

Reaching Task. At the beginning of each trial, subjects had to move their cursor into a start circle (5 mm radius) while aided by robotic forces. After maintaining that position for a variable time interval (1.5 ± 0.5 s), a target circle appeared 80 mm from the start; participants were instructed to perform quick and accurate "slashing" motions through this target. Whenever the subject's current position was also 80 mm from the start, their cursor was replaced by a small, fixed dot, which denoted the user's achieved position. To be considered a successful movement, the cursor must have intersected the target circle—if so, the subject heard a pleasant tone (otherwise no noise was played). We additionally provided participants with the amount of time taken to complete each motion so they could self-regulate for speed and consistency. Finally, the target and marker were erased from the screen after another variable time interval (1.5 ± 0.5 s), and robotic forces returned the user—in the absence of visual feedback—to the start circle, where their cursor reappeared. Timing variability was enforced so as to prevent subjects from developing a rhythm, which might have discouraged active participation. Users initially performed 40 unrecorded movements to become familiar with our setup.

Experimental Procedure. Consistent with the procedure outlined by [7], experiments were broken into five blocks, each containing multiple trials of the reaching task described above (see Fig. 1). Subjects practiced this reaching task (baseline), learned a visuomotor rotation (training), briefly experienced a second rotation (perturbation), and reverted to some motor mapping (error clamp), before returning to their original behavior (washout). The three types of robotic controllers were crossed factorially with two levels of task difficulty. During the Easy task the target orientation was constant at $135°$ throughout the experiment; on the other hand, for the Hard task the target angle was uniformly randomly placed between $90°$ and $180°$ during baseline and training blocks. The target radius was 10 mm for both

Easy and Hard tasks—this offered a successful movement range of $14°$. Robotic controllers were only implemented in the training block, and involved a proportional gain $K_P = diag(K_{P_x}, K_{P_y}, K_{P_z})$ multiplied by the 3×1 vector connecting the user's current position (x) to the closest point on the line between start and goal circles (x_c)

$$u = K_P(x_c - x) \tag{1}$$

For the HG case $K_{P_x} = K_{P_y} = 20$ N/m, for the EA case $K_{P_x} = K_{P_y} = -20$ N/m, and for the NC case $K_{P_x} = K_{P_y} = 0$ N/m. In every case K_{P_z} was 500 N/m in order to keep subject movements in the xy plane.

The first block (baseline) simply consisted of 20 trials with veridical feedback. A $30°$ CCW visuomotor rotation was introduced during the second block (training); to help subjects learn this rotation, both visual and auditory feedback were provided throughout 60 trials. For 66 of the next 80 trials, however, continuous visual feedback of the robot's position was removed, forcing users to rely upon endpoint auditory reinforcement when determining movement success. The remaining 14 trials with visual feedback were pseudo-randomly interspersed so that participants could maintain the desired behavior. Within the third block (perturbation), subjects experienced a novel $45°$ CCW visuomotor rotation for 30 trials. In order to examine retention of desired motor learning (the $30°$ CCW rotation) after this perturbation, 100 error clamp trials were conducted during the fourth block. Here participants controlled the radial distance of their cursor from the start circle, but the cursor's angular position was fixed to a line between start and goal markers with some uniformly random variability ($135 \pm 2.5°$), ensuring perceived success regardless of actual arm motion [16]. Because subjects always received artificial visual and auditory feedback for successful trajectories, we used these trials to determine what movements people associated to the given target direction—i.e., whether they retained a $30°$ CCW rotation or reverted to their natural, unrotated, baseline motions. The last block (washout) again entailed veridical feedback for 40 trials. Subjects were given short breaks at predefined points during experiments; they were not informed of the visuomotor rotations, error clamp, or robotic control strategy.

Learning Models. We used a single-state state-space model [5], [8], [17] to quantify learning during the training block. Given z_n, the subject's estimate of the visuomotor rotation at trial n, we can write

$$\begin{aligned} z_{n+1} &= Az_n + B(r_n - z_n) \\ y_n &= -z_n \end{aligned} \tag{2}$$

where r is the visuomotor rotation, and y is the measured hand direction, $A \in [0, 1]$ is a forgetting factor, and $B \in [-1, 1]$ is a learning rate. Assuming $z_0 = 0°$, we solved for A and B by using the MATLAB function `fmincon` to minimize the total squared error between predicted $(\hat{y})$ and measured (y) hand directions. Similarly, to study how rapidly subjects reverted to their baseline behavior after the $45°$ perturbation, we fit an exponential function $y = C_1 exp(-\lambda t) + C_0$ to measured hand directions during the error clamp block. Here C_0 and C_1 are constants, t is the trial number, and λ is the decay rate.

Data Analysis. Although twenty-eight participants completed both experiments, the data belonging to four participants from the HG group were excluded, leaving equal-sized groups with eight subjects each for subsequent analysis. One of the excluded individuals consciously noticed the presence of the error clamp, and we were therefore unable to examine their underlying behavior. The error clamp movements of three other HG subjects converged in the opposite direction of their trained behavior (error clamp hand direction: $-68.3°$, $-54.8°$, $-53.9°$, decay rate: $\lambda < 0$); consistent with prior work [8], all data from these subjects were excluded from further analysis. It is interesting to note that all three of these cases occurred when participants trained using HG while performing the Easy task first.

Statistical analysis was performed using SPSS (IBM). A general linear model (GLM) was used to analyze main effects of our within subjects factor (task difficulty) and between subjects factors (controller type and order of task presentation). Outcome measures were defined based on the block being analyzed, as we separately treated the training and error clamp blocks to independently examine learning and retention. Performance of the Easy and Hard tasks were then analyzed separately; contrasts were used to interpret statistically significant interactions among the between subjects factors of controller and presentation order.

For the training blocks, we used six outcome measures: directional error, intra-subject variability, RMSE, learning rate, forgetting factor, and percent success. For the error clamp blocks, we analyzed directional error, intra-subject variability, RMSE, and decay rate. These metrics are discussed with the presentation of results. The GLM allows us to test the effect of each between and within subjects factor on overall performance as defined by these sets of metrics. Our subsequent analysis focused on investigating the effects of controller type on particular outcome measures, and interactions between controller type and other factors; these investigations were separately conducted for the Easy and Hard tasks. Finally, we analyzed task difficulty and order of presentation.

III. Results

We sought to investigate how haptic guidance (HG), error augmentation (EA), or no controller (NC) affected reinforcement learning of a visuomotor rotation when the target location was fixed (Easy task) and when the target location varied (Hard task). Shmuelof *et al.* [7] recently demonstrated that providing only binary auditory feedback of trial success during training led to better retention of the desired visuomotor rotation, while both continuous visual feedback and binary auditory feedback impeded learning. We used a protocol similar to that introduced by [7]; participants adapted to a $30°$ visuomotor rotation while experiencing controller forces and trials with only reward-based reinforcement. After a brief perturbation, we examined each subject's retention of the $30°$ rotation by using an error clamp.

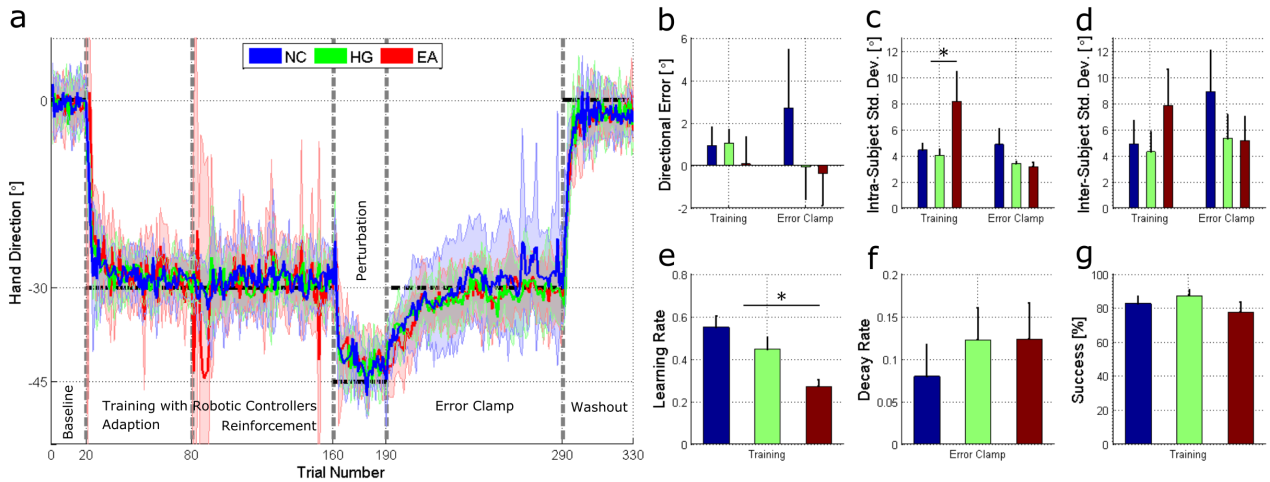

Fig. 2. Easy Task ($n = 24$); NC ($n = 8$), HG ($n = 8$), and EA ($n = 8$). (a) Average hand direction (solid colored lines) and inter-subject standard deviation (shaded areas) for the controller groups during each trial. The horizontal black lines indicate the visuomotor offset, as well as the desired hand direction for the error clamp trials. (b-d) Averaged directional error and standard deviation during the final 20 trials of the training and error clamp blocks. (e) Average learning rate B computed from the adaption portion of the training block. (f) Average decay rate λ during the error clamp trails. (g) Average success percentage during the last 20 trials of the training block. In (b-g), error bars show the standard error of the mean and * denotes $p < .05$.

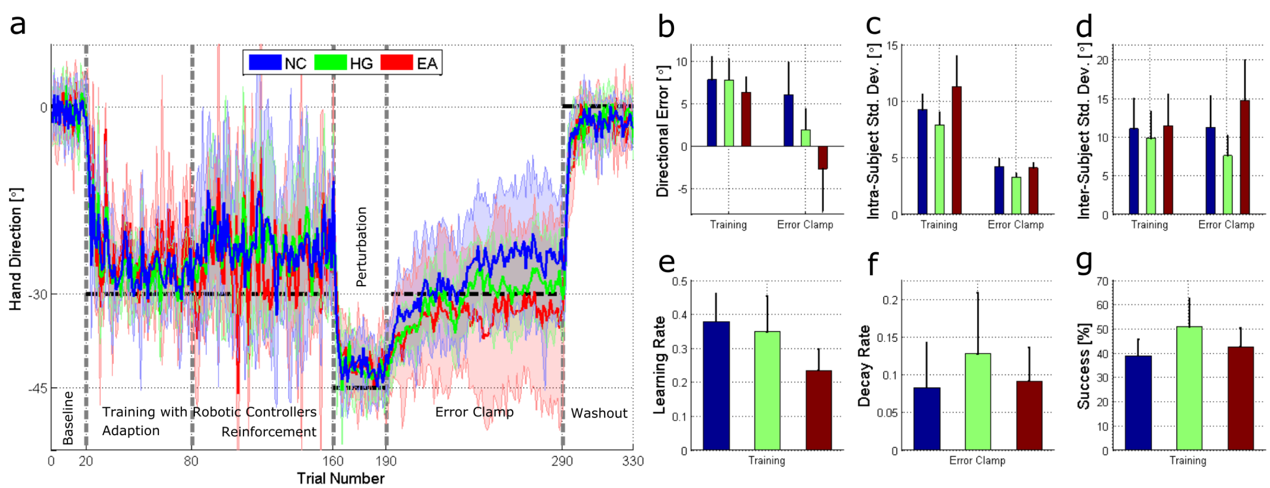

Fig. 3. Hard Task ($n = 24$); NC ($n = 8$), HG ($n = 8$), and EA ($n = 8$). We here follow the same conventions as in Fig. 2. Larger directional errors and standard deviation during the training block as well as lower learning and success rates suggest that it was more difficult for subjects to adapt to a visuomotor rotation when the target position was randomized (Hard task) as compared to when the target position was fixed (Easy task).

Plots of averaged hand direction during the Easy task are shown in Fig. 2a, where hand direction is defined as the user's achieved position relative to the target. Plots of average hand direction during the Hard task are shown in Fig. 3a. From the GLM analysis, we investigated the overall effect of task difficulty, controller type, and order of task presentation on performance in training and error clamp blocks. We found that there was a statistically significant difference in overall training behavior (assessed across all training block outcome measures) based on task difficulty ($F_{6,13} = 12.23,\ p < .01$). The effect of task difficulty on error clamp behavior (assessed across all error clamp block outcome measures) trended towards significance ($F_{4,15} = 2.53,\ p = .09$). There was a statistically significant difference in training behavior based on controller type ($F_{12,26} = 2.18,\ p = .05$; Wilks' $\lambda = 0.25$), but this did not extend to error clamp behavior ($F_{8,30} = 1.01,\ p = .45$; Wilks' $\lambda = 0.62$). Similarly, the order of task presentation had a statistically significant impact on training behavior ($F_{6,13} = 3.33,\ p = .03$), but not on error clamp behavior ($F_{4,15} = 0.46,\ p = .76$).

Next, we analyzed the effects of controller type on training and EC block performance metrics during the easy task. We found that there were no statistically significant differences between average hand directions at the end of training or error clamp blocks for any of the controller groups (Group $\times$ Block interaction: $F_{2,18} = 0.50,\ p = .62$), indicating the learned behavior was retained regardless of robotic intervention level (Fig. 2b). As can be seen in Fig. 2c, however, subjects with EA exhibited significantly more variability during the training block than those with NC and HG ($F_{1,18} = 7.43,\ p = .01$). We used RMSE in the error clamp block to assess deviations from the desired $-30°$ hand direction, and conclude that subjects with EA experienced significantly worse learning ($F_{1,18} = 6.76,\ p = .02$; NC+HG,

$4.66 \pm 1.37°$; EA, $8.42 \pm 6.53°$), but users trained with HG and EA better retained the desired behavior than those trained with NC ($F_{1,18} = 6.48,\ p = .02$; HG+EA, $4.72 \pm 2.18°$; NC, $8.23 \pm 5.26°$). The same trends applied to inter-subject variability (Fig. 2d), indicating that training with robotic controllers led to more consistent retention among subjects during error clamp trials. We next observed a significantly lower learning rate (Fig. 2e) for EA relative to NC ($F_{2,18} = 7.68,\ p < .01$) after applying a Tukey post-hoc comparison. The decay rate (Fig. 2f) was not significantly different for any of the controller groups ($F_{2,18} = 0.40,\ p = .68$). To summarize, we found that the addition of robotic controllers during the Easy task benefited retention; although the EA group experienced a lower learning rate and greater training standard deviation when contrasted with the NC group, those trained with HG or EA outperformed the NC group during error clamp trials in terms of RMSE. It is interesting to note that these differences in performance cannot be attributed to changes in success rates (Fig. 2g; $F_{2,18} = 1.46,\ p = .26$), demonstrating that the value of controller involvement is not limited to helping users accurately reach the target.

For the hard task, we observed that there were no statistically significant differences between average hand directions or intra-subject variability at the end of training trials (mean, $F_{2,18} = 0.22,\ p = .81$; std. dev., $F_{2,18} = 0.98,\ p = .39$) or error clamp trials (mean, $F_{2,18} = 1.20,\ p = .32$; std. dev., $F_{2,18} = 1.05,\ p = .37$), demonstrating that learning and retention were to some extent unaffected by robotic interaction (Fig. 3b-c). When considering our RMSE performance metric, however, we found an indication that HG subjects better retained the desired visuomotor rotation than those trained with NC or EA ($F_{1,18} = 3.13,\ p = .09$; NC+EA, $11.6 \pm 6.86°$; HG, $6.54 \pm 3.97°$), even though all groups had similar performance levels at the end of training ($F_{2,18} = 0.33,\ p = .72$; NC, $13.25 \pm 5.44°$; HG, $11.48 \pm 6.58°$; EA, $13.69 \pm 6.91°$). This trend was also borne out for inter-subject variability, as can be seen in Fig. 3d. Moreover—although not statistically significant—average success rates for the HG group exceeded NC and EA groups by 10% (Fig. 3g), suggesting that assistive interaction led to more training trials with reward. Variations in learning rate (Fig. 3e; $F_{2,18} = 0.97,\ p = .40$) and decay rate (Fig. 3f; $F_{2,18} = 0.15,\ p = .86$) were insignificant across controllers. We therefore conclude that robotic intervention may have slightly benefited retention during the Hard task; more specifically, HG subjects had higher average success rates during training, as well as more accurate and consistent performance during error clamp trials.

Turning our attention to the effects of task difficulty, we found that EA subjects' retention was the most negativity affected by switching from the Easy task to the Hard task. While the error clamp RMSE of NC and HG subjects increased by 37.9% and 38.4%, respectively, the error clamp RMSE of EA subjects increased by 148.5% between Easy and Hard tasks. The decay rate of those trained with EA decreased by 26.0%, but the decay rates of NC (4.2%) and HG (3.3%) groups marginally increased. Finally, inter-

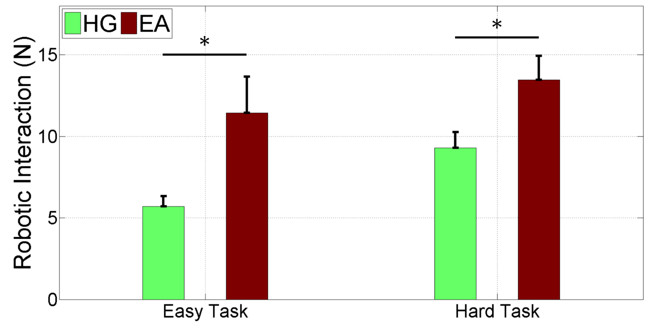

Fig. 4. Averaged L_2 norm of controller forces, u in (1). The Hard task resulted in significantly larger errors, and thus a larger amount of robotic interaction as compared to the Easy task. * denotes $p < .05$.

subject variability during the error clamp increased by 26.8% for NC users, 42.9% for HG users, and 185.4% for EA users. We were also interested in comparing the controller forces applied during Easy and Hard tasks. Accordingly, we present the averaged L_2 norm of controller forces (u) in Fig. 4; the NC group is here omitted since they underwent no robotic interaction. We found that EA subjects experienced a larger amount of controller forces during both the Easy task ($F_{1,18} = 11.2,\ p < .01$) and Hard task ($F_{1,18} = 13.3,\ p < .01$). As expected, the Hard task elicited more robotic involvement than the Easy task (Task interaction: $F_{1,18} = 5.92,\ p = .03$), suggesting that the relative impacts of different control strategies may have been more pronounced during the Hard task.

Finally, we sought to determine how the order in which subjects performed the experiments affected their resulting behavior; we concluded that ordering did not alter retention, but did impact some training metrics. In particular, success rates during the Hard task ($F_{1,18} = 7.14,\ p = .02$), and both hand direction ($F_{1,18} = 14.31,\ p < .01$) and intra-subject variability ($F_{1,18} = 5.11,\ p = .04$) during the training portion of the Easy task were influenced by ordering.

IV. Discussion

In this study we examined interactions between haptic guidance, error augmentation, and reward-based reinforcement during visuomotor rotation tasks with variable difficulty. We found that subjects were able to successfully integrate these controllers without negative effects on the retention of motor learning. More precisely, training with either haptic guidance and error augmentation desirably reduced movement variability during retention trials for a simple task. Increasing task difficulty led to greater disparities in retention among subjects trained using the different robot controllers, with the strongest and most consistent learning occurring in subjects assisted by haptic guidance.

A. Using haptic intervention with reward-based reinforcement leads to retention

The examined haptic guidance and error augmentation controllers provided a source of kinesthetic feedback not present in the described reward-based reinforcement paradigms [7], [15]. This persistent kinesthetic feedback during training had the possibility to adversely effect motor learning, especially since its direction and magnitude was derived from the user's error vector (1). Results in [7] revealed that those trained with visual vector error (continuous visual feedback of hand position) in addition to reward-based reinforcement demonstrated significantly worse retention than those trained with just reward-based reinforcement. We might therefore expect haptic guidance and error augmentation groups to consistently exhibit larger hand direction errors during error clamp trials; however, our findings indicate additional kinesthetic feedback based on vector error did not categorically harm retention. Subjects—regardless of the presence or type of controller—did not have significantly different directional errors during retention, suggesting that kinesthetic vector error has a different underlying impact on motor learning than do visual vector errors.

Haptic guidance and error augmentation introduced a force field to which subjects had to adapt. Since motor learning of visuomotor rotations and force fields appear to rely on similar mechanisms [18], adaptation to the kinesthetic feedback may have actually interfered with adaptation to the correct hand direction. Indeed, we found that learning rates were lower for subjects trained using haptic guidance or error augmentation than for subjects trained without a controller, particularly in the Easy task. It should be noted, however, that (a) these differences in learning rate did not appear to have a long-term influence on other aspects of training, such as hand direction, variability, or success rates, and (b) lower learning rates did not necessarily lead to worse retention during error clamp trials. Hence, we conclude that the reduced learning rates attributed to robotic controllers did not preclude retention. Moreover, while haptic guidance and error augmentation users trained with both a force field and visuomotor rotation, they maintained the desired hand direction during error clamp trials despite the absence of a robotic controller. Although motor learning is often restricted to the training environment [13], we here found that users retained the trained behavior even when assistive or resistive force fields were removed.

Of course, we also identified some negative repercussions from using haptic intervention. Three subjects who experienced haptic guidance during the Easy task increased their visuomotor offset in error clamp trials, and almost mirrored the location of the target about the y-axis. This may indicate poor learning of the intended motion caused by passively relying on haptic guidance during training [13], or it could be that retention also preserves the reflected actions [5]; in either case, this phenomenon merits further study. Another concern is whether using an error clamp is suitable for checking user's retention. For example, by making the error clamp trials more similar to training trials, it may be possible to preserve a visuomotor offset without subjects ever reverting to their actual retained behavior [16]. Recent work, however, demonstrates that decay to retained behavior is inevitable during error clamp trials, and that the onset of this decay to retained behavior does not depend on similarities between error clamp and training trials [19]. Furthermore, Kitago *et al*. [17] found that error clamp trials are an active form of unlearning—unlike letting time elapse or receiving veridical feedback—so retention after error clamp trials should indicate that the trained behavior is well-remembered. Based on this research, we argue that hand directions during error clamp trials accurately reflect a subject's level of retention.

B. Task difficulty influences the effectiveness of reward-based reinforcement with haptic guidance or error augmentation

Viewed together, task difficulty and the amount of controller assistance determine how challenging it is for users to perform a task; there is likely some relationship between this challenge level and subject involvement. For example, using assistive training to artificially increase the reward associated with a movement has been found to lower involvement by discouraging subjects from exploring potentially better alternatives [20], which may even accelerate the rate of unlearning [21]. Lower levels of subject involvement have previously been used to explain the detrimental effects of haptic guidance on motor learning, while error augmentation has been shown to increase a subject's energy expenditure [13]. We therefore expect subjects with error augmentation to be more engaged in the task than those trained with no controller, while the haptic guidance users should exhibit the least involvement. Since it is difficult to directly measure user participation in a non-invasive manner, we instead used interaction forces during training to indicate involvement. As predicted, for both Easy and Hard tasks error augmentation resulted in larger controller forces, suggesting more effort as well as increased exploration. We note, however, that optimal motor learning occurs when the task's challenge matches subject ability, not necessarily when participation is maximized [9]–[11]. Given that error augmentation led to worse retention than haptic guidance during the Hard task, we conclude that the Hard task with error augmentation may have been unsuitably challenging for many subjects.

Task difficulty and controller assistance also influence the trade-off between variability and success during training with reward-based reinforcement. Providing only reward-based reinforcement is more conducive to subject variability than paradigms which display a continuous error signal [6]. Indeed, greater amounts of movement variability at baseline likely correspond to faster learning rates within reward-based reinforcement paradigms [22], and increasing variability through force fields or visuomotor rotations can cause better retention of the trained behavior [18], [21]. Hence, the error augmentation group should exhibit better retention than the haptic guidance group during both Easy and Hard tasks since they had higher intra-subject variability within the training phase. While it is important for subjects to explore multiple trajectories, however, subjects should also consistently reach the target in order to maximize reinforcement [22]. We note that the haptic guidance group experienced greater success rates than the error augmentation group, and accordingly received more reinforcement during training. Results from the Easy task could therefore show a balance point between increasing success but lowering variability (haptic guidance) and increasing variability but lowering success (error augmentation), as the benefits of each seem relatively matched. During the Hard task, it is possible that success rates for subjects trained with error augmentation were beneath some minimum threshold, and, as such, the desired action was not sufficiently reinforced despite having higher variability.

V. Conclusion

In this paper we explored the applicability of reward-based reinforcement (end-of-trial auditory feedback indicating success) together with different control strategies for robotic rehabilitation (no controller, haptic guidance, and error augmentation) on the retention of a visuomotor rotation with two task difficulty levels. For both tasks, haptic guidance produced the best performance within training, and was therefore associated with more reward-based feedback. Subjects training without a controller experienced the fastest learning rates, presumably because other groups had to interpret haptic cues in addition to task cues. Both haptic guidance and error augmentation groups showed superior performance in terms of visuomotor skill retention during the Easy task. For the Hard task, the haptic guidance group again had the highest success rate, and outperformed the error augmentation group in terms of skill retention. We argue that these findings demonstrate robotic rehabilitation control strategies such as haptic guidance and error augmentation can effectively be incorporated in the reward-based reinforcement paradigm, and, in some cases, may improve the subject's retention when compared to training without a controller. This successful combination of haptic guidance and error augmentation with reward-based reinforcement could positively benefit the outcomes of robotic rehabilitation, where robot guidance and augmentation are used to influence motor retraining following neurological injury.

References

[1] M. A. Smith, A. Ghazizadeh, and R. Shadmehr, "Interacting adaptive processes with different timescales underlie short-term motor learning," *PLoS Biol.*, vol. 4, no. 6, p. e179, 2006.

[2] R. Shadmehr, M. A. Smith, and J. W. Krakauer, "Error correction, sensory prediction, and adaptation in motor control," *Annu. Rev. of Neurosci.*, vol. 33, pp. 89–108, 2010.

[3] J. W. Krakauer, "The applicability of motor learning to neurorehabilitation," in *Oxford Textbook of Neurorehabilitation*, V. Dietz and N. Ward, Eds. Oxford, UK: Oxford U. Press, 2015, pp. 55–63.

[4] L. Marchal-Crespo and D. J. Reinkensmeyer, "Review of control strategies for robotic movement training after neurologic injury," *J. Neuroeng. Rehabil.*, vol. 6, no. 20, 2009.

[5] V. S. Huang, A. Haith, P. Mazzoni, and J. W. Krakauer, "Rethinking motor learning and savings in adaptation paradigms: Model-free memory for successful actions combines with internal models," *Neuron*, vol. 70, no. 4, pp. 787–801, 2011.

[6] J. Izawa and R. Shadmehr, "Learning from sensory and reward prediction errors during motor adaptation," *PLoS Comput. Biol.*, vol. 7, no. 3, p. e1002012, 2011.

[7] L. Shmuelof, V. S. Huang, A. M. Haith, R. J. Delnicki, P. Mazzoni, and J. W. Krakauer, "Overcoming motor "forgetting" through reinforcement of learned actions," *J. Neurosci.*, vol. 32, no. 42, pp. 14 617–14 621, 2012.

[8] J. M. Galea, E. Mallia, J. Rothwell, and J. Diedrichsen, "The dissociable effects of punishment and reward on motor learning," *Nat. Neurosci.*, vol. 18, no. 4, pp. 597–602, 2015.

[9] M. A. Guadagnoli and T. D. Lee, "Challenge point: A framework for conceptualizing the effects of various practice conditions in motor learning," *J. Motor Behav.*, vol. 36, no. 2, pp. 212–224, 2004.

[10] M.-H. Milot, L. Marchal-Crespo, C. S. Green, S. C. Cramer, and D. J. Reinkensmeyer, "Comparison of error-amplification and haptic-guidance training techniques for learning of a timing-based motor task by healthy individuals," *Exp. Brain Res.*, vol. 201, no. 2, pp. 119–131, 2010.

[11] A. E. Bouchard, H. Corriveau, and M.-H. Milot, "Comparison of haptic guidance and error amplification robotic trainings for the learning of a timing-based motor task by healthy seniors," *Front. Syst. Neurosci.*, vol. 9, 2015.

[12] L. Marchal-Crespo, P. Wolf, N. Gerig, G. Rauter, L. Jaeger, H. Vallery, and R. Riener, "The role of skill level and motor task characteristics on the effectiveness of robotic training: First results," in *Proc. IEEE Conf. Rehab. Robot.*, 2015, pp. 151–156.

[13] H. Heuer and J. Lüttgen, "Robot assistance of motor learning: A neurocognitive perspective," *Neurosci. Biobehav. Rev.*, vol. 56, pp. 222–240, 2015.

[14] J. L. Patton, M. E. Stoykov, M. Kovic, and F. A. Mussa-Ivaldi, "Evaluation of robotic training forces that either enhance or reduce error in chronic hemiparetic stroke survivors," *Exp. Brain Res.*, vol. 168, no. 3, pp. 368–383, 2006.

[15] C. J. Hasson, J. Manczurowsky, and S.-C. Yen, "A reinforcement learning approach to gait training improves retention," *Front. Hum. Neurosci.*, vol. 9, 2015.

[16] P. A. Vaswani and R. Shadmehr, "Decay of motor memories in the absence of error," *J. Neurosci.*, vol. 33, no. 18, pp. 7700–7709, 2013.

[17] T. Kitago, S. L. Ryan, P. Mazzoni, J. W. Krakauer, and A. M. Haith, "Unlearning versus savings in visuomotor adaptation: Comparing effects of washout, passage of time, and removal of errors on motor memory," *Front. Hum. Neurosci.*, vol. 7, 2013.

[18] J.-J. O. de Xivry and P. Lefèvre, "Formation of model-free motor memories during motor adaptation depends on perturbation schedule," *J. Neurophysiol.*, vol. 113, no. 7, pp. 2733–2741, 2015.

[19] A. E. Brennan and M. A. Smith, "The decay of motor memories is independent of context change detection," *PLoS Comput. Biol.*, vol. 11, no. 6, p. e1004278, 2015.

[20] A. Sans-Muntadas, J. E. Duarte, and D. J. Reinkensmeyer, "Robot-assisted motor training: Assistance decreases exploration during reinforcement learning," in *Proc. IEEE Conf. Eng. Med. Biol. Soc.*, 2014, pp. 3516–3520.

[21] B. R. Ballester, L. Serra Oliva, A. Duff, and P. Verschure, "Accelerating motor adaptation by virtual reality based modulation of error memories," in *Proc. IEEE Conf. Rehab. Robot.*, 2015, pp. 623–629.

[22] H. G. Wu, Y. R. Miyamoto, L. N. G. Castro, B. P. Ölveczky, and M. A. Smith, "Temporal structure of motor variability is dynamically regulated and predicts motor learning ability," *Nat. Neurosci.*, vol. 17, no. 2, pp. 312–321, 2014.