Vision Research xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Long-term motor learning: Effects of varied and specific practice

Chéla R. Willey *, Zili Liu

Department of Psychology, University of California Los Angeles, 1285 Franz Hall, Box 951563, Los Angeles, CA 90025, USA

A R T I C L E I N F O

Article history:
Received 5 December 2016
Received in revised form 26 February 2017
Accepted 20 March 2017
Available online xxxx
Number of reviewers = 2

Keywords:
Motor learning
Varied training
Specific training
Generalization
Transfer

A B S T R A C T

According to the schema theory, variability during practice allows for a larger repertoire of movements to help form a generalized motor program for similar motor skills. Varied training is thought to enhance long-term retention of the motor program due to the heightened difficulty presented. In a highly cited study on this topic, Kerr and Booth (Perceptual and Motor Skills 46 (1978), 395–401) trained two groups of children for 10 weeks to throw a beanbag towards either one central target (specific group) or two targets that were ±1 foot away from the central target (varied group). They found that the varied group performed significantly better than the specific group when both groups were tested at the central target. We, following the same paradigm, trained 30 adults on a similar beanbag throwing task and tested them at various target distances. Our results suggested that after 5–7 weeks of training, the specific groups tended to undershoot at longer distances and overshoot at shorter distances while the varied group tended to center their throws around the target at all distances. However, the overall magnitude of error (regardless of over- or undershooting) was similar across groups. We found some support for the hypothesis that the varied group could better generalize to untrained distances, but this advantage was found mainly for the longest distance and disappeared by a posttest held two weeks after practice.

1. Introduction

Motor learning has been a crucial aspect to our survival as a species. As we evolved, we have mastered complex motor commands and have become more efficient at a variety of complex skills and movement, such as hunting and, more recently, sports. Through practice, our skills can improve with increasing accuracy and speed, while decreasing variability in the execution of the motor command. When performing an action, the goal is to minimize errors to home in on the desired target and then learn from those errors via feedback for subsequent movements. In ballistic movements, such as throwing a ball, success of the movement is assessed after making the movement to inform subsequent throws. The thrower consequently associates the particular movement of their limbs with the success of that movement. After repeated throws to the target, the thrower is expected to become more consistent and accurate.

Researchers have long been interested in the development of motor skills and have looked for more effective and efficient ways to structure practice in order to achieve an expert level of consistency and accuracy. The schema theory offers up some parameters that one can manipulate in order to influence the rate of learning, the generalizability of the skill, and the ability to retain the skill over a long retention period (Schmidt, 1975; Schmidt & Bjork, 1992; Shea & Wulf, 2005). Usually, large improvements in performing a skill can be seen after just one practice session, but multiple sessions are needed if the skill is to be maintained over long term (Schmidt & Bjork, 1992). Spacing out practice sessions has long been thought to help consolidate learning (Janiszewski, Noel, & Sawyer, 2003). Given enough time after the previous practice session the participant will have to relearn parts of the motor skill, effectively strengthening and modifying what was learned in the last session (Schmidt, 1975; Wymbs, Bastian, & Celnik, 2016). Along these same lines, introducing spacing within a training session seems also beneficial, even when the time between practicing one skill is filled with practicing another skill. Interleaving practice between different skills or different aspects of the same skill has been shown to decrease short-term performance but tends to promote longer-term learning of both skills (Birnbaum, Kornell, Bjork, & Bjork, 2012; Schmidt & Bjork, 1992). This type of practice, also known as contextual interference, introduces contextual shifts between practice trials that theoretically allow for better learning. Additionally, if the interleaved tasks are similar, contextual interference can better generalize to similar but untrained skills or conditions. For example, a basketball player must master shooting a basketball from the free-throw line. During practice, the player can introduce contextual interference by alternating between

* Corresponding author.
E-mail addresses: cwilley@ucla.edu (C.R. Willey), zili@psych.ucla.edu (Z. Liu).

http://dx.doi.org/10.1016/j.visres.2017.03.012
0042-6989/

Please cite this article in press as: Willey, C. R., & Liu, Z. Long-term motor learning: Effects of varied and specific practice. *Vision Research* (2017), http://dx.doi.org/10.1016/j.visres.2017.03.012

shooting from the free-throw line and the three-point line. According to the schema theory, over a long period, the player should improve their shooting at both distances more than if the player focused their shooting practice at only one distance, with the same amount of practice. Furthermore, the schema theory hypothesizes that if this variability is introduced during practice, the player can also better transfer this skill to shooting from unpracticed distances.

This improved transfer effect presumably is due to the broader experience gained from varying their movements to shoot from different distances. However, the schema theory does not provide specific hypotheses about how improvements should manifest. In principle, if a general-purpose motor program has been learned via efficient training, then this learning can transfer to all novel conditions equally well, regardless how similar the novel condition is to the trained conditions. Alternatively, if no generalizable rules can be abstracted from training with concrete examples, then no transfer is expected beyond the trained examples. However, even in this no transfer situation, one should still expect limited transfer that tapers off as the novel condition is less similar from the trained. For example, classical visual perceptual learning has been found to be specific to training stimulus. But even in this exemplary case of specific learning, transfer still tapers off gradually, not abruptly. For example, in visual motion direction discrimination learning, partial transfer was found 45° away, but not 90° away, from the trained direction (Ball & Sekuler, 1987). In this so called "stimulus specific learning," one cannot say whether no rule is learned or only a limited version of the rule is learned. What is empirically testable is whether varied training can better generalize than specific training, which we would test in the current study.

In the literature, a well-cited study on varied learning was by Kerr and Booth (1978). They, by recruiting children who participated in a 10-week recreational program, trained two groups of participants to throw to either one (the specific group) or two (the varied group) target distances before testing them at the specific group's practiced target distance. They found that children who practiced throwing at 3 ft and 5 ft had significantly fewer errors when throwing at the test target distance (4 ft) than children who practiced throwing only at this distance (4 ft). This finding suggests that varied practice was beneficial to generalization in untrained conditions.

The purpose of the current study was to replicate Kerr and Booth's original study using adult participants with a novel throwing motion. Per the schema theory and Kerr and Booth's findings, the varied group (trained at two distances) should perform better at all distances than the specific group (trained at one distance), including the distance the specific group specifically trained at. That is to say, participants in the varied group should develop a more general motor program using the provided visual and verbal feedback. However, it is possible that improvements may only be limited to a range around which participants practiced. Nevertheless, varied practice should still allow for better performance at untrained target distances.

2. Experiment 1

2.1. Methods and procedure

All methods were approved by UCLA's institutional review board and were in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Informed consent was obtained from all participants before participating in the current experiments.

We recruited 20 undergraduates from the University of California Los Angeles to participate in the current experiment. All participants were right-handed and had normal or corrected-to-normal vision as determined by an acuity test at the start of the experiment.

Participants threw a 9.05 oz. beanbag to various distances during one pretest session, 20 practice sessions, and two posttest sessions. During all throws, the participant threw the beanbag over their shoulder with their back facing the target distances. All target distances were clearly marked on the ground and could be seen from the participant's standing position. The targets were 3, 5, 7, 9, and 11 feet from the participant. After each throw, a trained research assistant measured the distance thrown from the center of the beanbag to the closest inch label, which was marked on the floor.

2.1.1. Pretest

During pretest, participants threw a set of 12 beanbags (one at a time), in a blocked design, to each of the 5 distances (randomly ordered). For each distance, participants only viewed once the distance that they aimed for at the beginning of the 12 trials and received no feedback between trials. A research assistant would indicate when the participant could make their next throw. Participants used only their non-dominant hand.

2.1.2. Practice

Participants were randomly assigned to practice in either the varied or specific condition. During practice, participants in the specific group aimed for a target 7 ft away, while participants in the varied group alternated between aiming for targets 5 ft and 9 ft away, switching targets every 12 trials.

Over 5–7 weeks ($M = 6.15$ weeks), each participant completed 20 practice sessions containing 60 trials each, thus completing a total of 1200 practice throws. After randomization into groups, participants were paired across the two groups. Practice sessions for each pair were scheduled to occur together week to week. Participants completed 3–4 practice sessions each week and the order of practice sessions between individuals of each pair was counterbalanced across weeks. As much as schedules would allow, sessions were spread out throughout the week. During some weeks, some pairs of participants had to perform more than one session in one day. However, no more than two sessions were performed in one day and this did not occur more than twice per participant. If participants completed more than one session in the same day, they took, at minimum, a 10-min break between sessions. Unlike pretest, participants received visual and explicit verbal feedback that included the exact distance thrown after every trial. All participants also switched between hands every 6 trials in order to reduce fatigue throughout practice.

2.1.3. Posttests

Posttest 1 occurred one week after completion of the last practice session and posttest 2 occurred one week after that. This second posttest allowed us to investigate the stability of the effects over a longer retention interval. Both posttests were performed in the same order as pretest. Like pretest, participants did not receive verbal or visual feedback after each test trial and only viewed the distance they were aiming for at the beginning of each set of 12 trials for each distance. They used their non-dominant hand for all throws in both posttests. We were unable to run 2 participants (one from each group) through their second posttests.

2.1.4. Analyses

We used signed errors, absolute errors, and variances to analyze performance. A signed error is defined as the signed distance from the landing spot to the target. A positive signed error indicates an overshot. An absolute error is the absolute value of the signed error. These types of measurements are used throughout the literature to describe errors and can convey important information about performance accuracy.

We should expect that both groups learn and thus reduce errors from pretest to posttests at most distances. Additionally, longer distances are more difficult and thus we would expect a general trend in which errors and variance are greater at longer distances, according to Weber's Law. Per the schema theory, we would predict that the varied group should have overall greater improvement than the specific group at all distances, but especially at 7 ft. We also expect the specific group to have lower variance in errors as compared to the varied group at posttest at 7 ft, due to the specificity of training.

2.2. Results

2.2.1. Signed errors

We first verified that there was no group difference at pretest. A 2 (Group) × 5 (Distance) ANOVA found neither a main effect of group ($p = 0.78$), nor an interaction effect ($p = 0.35$).

2.2.1.1. Pretest vs. Posttest 1. A mixed 2 (Group) × 2 (Time) × 5 (Distance) ANOVA revealed a main effect of time, where participants reduced their underestimation from pretest ($M = -0.83$ ft, $SE = 0.29$) to posttest ($M = -0.13$ ft, $SE = 0.25$), F (1, 18) = 8.57, $p = 0.009$, $\eta_p^2 = 0.32$. There was also a main effect of target distance, $F(2.59, 46.53) = 19.73$, $p < 0.001$, $\eta_p^2 = 0.87$ (corrected for the violation of the assumption of sphericity, χ^2 (9) = 22.62, $p = 0.007$, $\varepsilon = 0.65$), such that at longer distances, participants undershot more, as expected from Weber's Law.

More important to our hypothesis testing, there was a three-way interaction between time, distance, and group, F (3.12, 56.18) = 3.08, $p = 0.021$, $\eta_p^2 = 0.15$, suggesting that groups reduced errors differently across distances from pretest to posttest 1 (Fig. 1). In order to look into this interaction in detail and, in particular, the effect involving group, the time × group relationship was analyzed at each distance. It turned out that the groups differed most in signed errors at the 11 ft target. This effect between groups can be better illustrated when the pretest data at 11 ft were controlled for. An ANCOVA with the pretest data as a covariate, revealed that there was a significant group difference at 11 ft during posttest 1, $F(1, 17) = 9.23$, $p = 0.007$, $\eta_p^2 = 0.35$. The second largest difference between groups was found using a similar ANCOVA at 3 ft, F (1, 17) = 6.43, $p = 0.021$, $\eta_p^2 = 0.27$. In order to compare with the previous literature, we found that the group difference at 7 ft was not significant using an ANCOVA, F (1, 17) = 1.66, $p = 0.215$, $\eta_p^2 = 0.09$. We found no group differences at any other distances from pretest to posttest 1.The general trend captured by the overall significant three-way interactions between pretest and posttest suggested that, while both groups tended to similarly improve their accuracies from pretest to posttest 1, the varied group outperformed the specific group at the farthest distance of 11 ft, and nearest distance of 3 ft. Additionally, we found the expected negative linear trend of distance within the specific group's posttest 1 signed errors in which participants in the specific group tended to undershoot at longer distances and overshoot at shorter distances, F (1, 9) = 87.71, $p < 0.001$. This trend did not reach significance within the varied group's posttest 1 signed errors, suggesting that they did not tend to undershoot as much at longer distances than did the specific group, F (1, 9) = 3.42, $p = 0.097$. We found similar effects in posttest 2.

2.2.1.2. Posttest 1 vs. Posttest 2. A mixed 2 (Group) × 2 (Time) × 5 (Distance) ANOVA revealed a significant main effect of distance (corrected for violation of sphericity, χ^2 (9) = 18.34, $p = 0.032$, $\varepsilon = 0.76$), $F(3.04, 48.55) = 11.98$, $p < 0.001$, $\eta_p^2 = 0.43$. Importantly, no effect involving time was significant. This indicated that no significant changes occurred between the two posttests. Therefore, the two sets of posttest data could be considered together in the future to consolidate the overall posttest data.

The same ANOVA also revealed an interaction effect between distance and group, $F(3.04, 48.55) = 7.65$, $p < 0.001$, $\eta_p^2 = 0.32$, suggesting that the two groups performed differently based upon the distance tested. In order to look into this further, we collapsed across time and used a 2 (Group) × 5 (Distance) ANOVA to examine the group differences at each distance. Along with the expected main effect of distance, we found a significant interaction between distance and group, reconfirming our interaction effect reported above, F (2.87, 51.723) = 6.89, $p < 0.001$, $\eta_p^2 = 0.28$. Using t-tests, we found significant differences between groups at 3 ft, t (18) = 3.22, $p = 0.005$, $d = 1.44$; and at 11 ft, t (18) = 3.11, $p = 0.006$, $d = 1.39$. These effects suggest that the varied group outperformed the specific groups at both 3 ft and 11 ft target distances, but both groups performed similarly at trained distances and nearby untrained distances (5, 7, and 9 ft). Specifically, they deviated from each other at the two extreme distances where the specific group undershot at the longest distance and overshot at the shortest distance as compared to the varied group.

We also looked at the absolute errors, but found no group differences at any target distance at either posttest (Fig. 2). Since

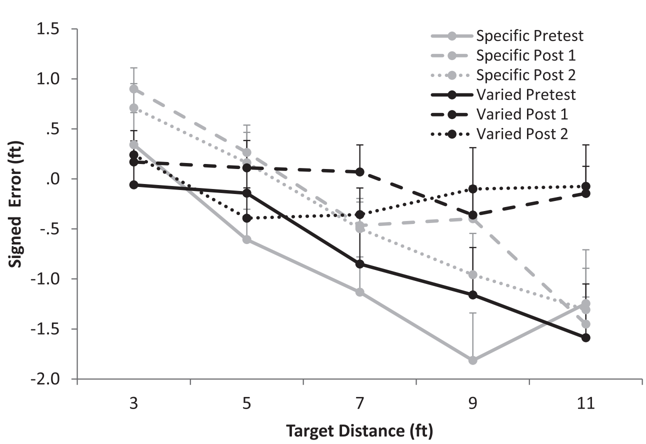

Fig. 1. Signed errors across target distances. Average signed errors on pretest, posttest 1, and posttest 2 of the specific (practice target = 7 ft) and varied groups (5 ft & 9 ft) at the 5 target distances. Positive scores are overthrows while negative scores are underthrows. Error bars represent the positive standard error about the mean (as they do in all subsequent figures).

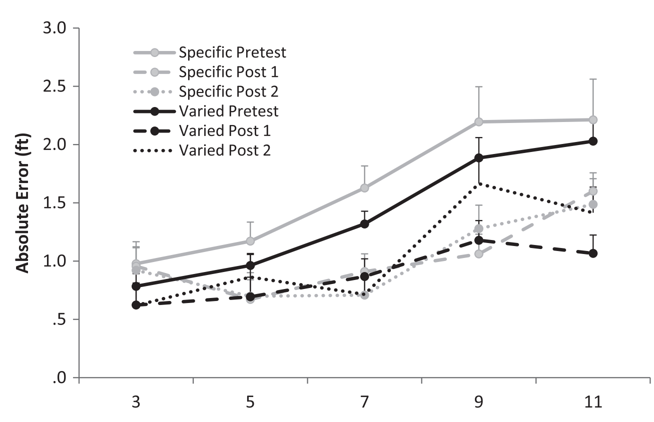

Fig. 2. Absolute errors across target distances. Average absolute errors on pretests (solid line), posttest 1 (dashed lines) and posttest 2 (dotted lines) between the specific (gray) and varied (black) groups at the 5 target distances.

Please cite this article in press as: Willey, C. R., & Liu, Z. Long-term motor learning: Effects of varied and specific practice. *Vision Research* (2017), http://dx.doi.org/10.1016/j.visres.2017.03.012

the absolute error is closely related to variance, we will focus on the analysis on variance instead.

2.2.2. Variance. Using a 2 (Group) × 5 (Distance) ANOVA with pretest variances, we found neither a main effect of group ($p = 0.38$), nor an interaction between distance and group, ($p = 0.58$). There was, however, the expected main effect of distance, in which greater target distances tended to have greater variance (corrected for violation of sphericity, χ^2 (9) = 48.78, $p < 0.001$, $\varepsilon = 0.43$), F (1.73, 31.23) = 4.80, $p = 0.019$.

2.2.2.1. Pretest vs. Posttest 1. A mixed model 2 (Group) × 2 (Time) × 5 (Distance) ANOVA revealed a main effect of time in that, not surprisingly, the pretest had a larger variance ($M = 1.13$ ft^2, $SE = 0.14$) than posttest 1($M = 0.67$ ft^2, $SE = 0.07$), F (1, 18) = 9.53, $p = 0.006$, $\eta_p^2 = 0.35$, indicating greater consistency after practice. We also found the expected effect of distances mentioned previously (corrected for violation of sphericity, χ^2 (9) = 36.12, $p < 0.001$, $\varepsilon = 0.50$), F (2.02, 36.30) = 9.64 $p < 0.001$, $\eta_p^2 = 0.35$.

There was neither an overall main effect of group nor any interactions, however. Nevertheless, since there was a theoretical reason to compare between the two groups at 7 ft, the trained distance for the specific group, we went ahead and compared between the two groups at this distance. Using the same ANCOVA as above, we confirmed that the specific group did not significantly differ from the varied group in consistency of throws, F (1, 17) = 2.33, $p = 0.145$, $\eta_p^2 = 0.12$ (Fig. 3).

2.2.2.2. Posttest 1 vs. Posttest 2. A mixed model 2 (Group) × 2 (Time) × 5 (Distance) ANOVA suggested again that longer distances had greater variances, as expected (corrected for violation of sphericity, χ^2 (9) = 37.92, $p < 0.001$, $\varepsilon = 0.45$), F (1.80, 28.78) = 6.49, $p = 0.006$, $\eta_p^2 = 0.29$. Overall, the specific group did not differ in variability from the varied group. However, at certain distances, we found marginally significant effects of variance between groups across both posttests. Using a 2 (Group) × 2 (Time), the specific group tended to have lower variances than the varied group at 5, 7, and 9 ft (F (1, 16) = 4.14, $p = 0.059$, $\eta_p^2 = 0.21$); but after using a Bonferroni corrected critical p-value these did not reach statistical significance.

Overall, our results suggest that both groups decreased their errors from pretest to posttests due to their training. Between groups at posttests, the varied group tended to center the signed errors around zero while the specific group tended to underestimate, particularly at the farthest target distances. This may give some evidence for the idea that varied practice enabled generalization to these farther distances, corresponding to predictions made

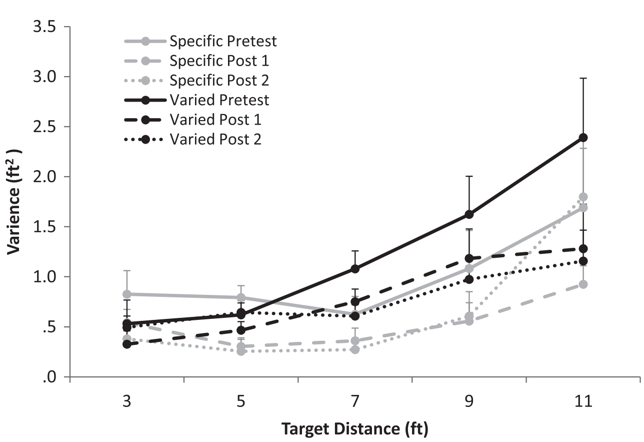

Fig. 3. Variances across target distances. Variance of pretest, posttest 1, and posttest 2 throws for the specific and varied groups at the five distances.

from the schema theory. The varied group's practice also allowed for generalization to 7 ft, while the specific group's practice still allowed for generalization to 5 ft and 9 ft, suggesting that generalization can occur in nearby untrained distances regardless of training schedule. Both groups produced similar absolute errors across all targets. Furthermore, the specific group had slightly greater consistency between throws. However, these effects were not significant after controlling for multiple comparisons. Using Kerr and Booth's published data, we attempted to extract their effect sizes to compare with our own. Kerr and Booth found a main group effect at the specific group's trained target distance when absolute errors were compared. Kerr and Booth's main effect of group had an approximate Cohen's d of 0.60 using signed errors. After transforming our η_p^2 into a Cohen's d, we found an approximate d of 0.13 between groups at 7 ft using signed errors and a Cohen's d of less than 0.01 using absolute errors.

Beyond the three middle-range distances, at 3 ft and at 11 ft, only the varied group had significantly reduced errors. Additionally, while both groups had about the same variance at 11 ft, the varied group centered around the target distance while the specific group undershot. The between-group differences at target distances 3 ft and 11 ft may not reflect better generalization of the varied group to unpracticed distances, however. Recall that the varied group's trained distances, 5 ft and 9 ft, were respectively closer to 3 ft and 11 ft than the specific group's trained distance, 7 ft. One may suspect that the varied group's better performance at 3 ft and 11 ft was simply due to proximity, rather than the training method. Indeed, the specific group seemed to also have the ability to generalize to nearby distances (5 ft and 9 ft) from their own trained target distance. In Experiment 2, we tested whether the varied group's better performance than the specific group at 11 ft was due to varied training or due to the fact that the 11 ft target distance was closer to one of the trained distances for the varied group (9 ft) than for the specific group (7 ft).

3. Experiment 2

If the varied group's better performance in Experiment 1 at 11 ft was due to this distance being closer to a trained target distance, then we should expect that the specific group, if trained at the same target distance of 9 ft, should also improve similarly at 11 ft. We recruited 10 participants for a new specific training who practiced solely at the 9 ft target distance during practice trials. This new specific group allowed us to directly compare the effects of practice schedule with the varied group. Participants completed 1200 practice trials over 5–6 weeks. In addition to the five tested target distances used in Experiment 1, we also tested this specific group at 13 ft during pretest and posttests. The addition of this target distance allowed us to test whether this specific group could similarly extend its generalization to a distance 4 ft away, or not, as in Experiment 1. All other procedures were identical to Experiment 1. All participants were right-handed undergraduates with normal or corrected to normal vision.

3.1. Results

3.1.1. Exp. 2 results from the specific group

A 2 (Time) × 6 (Distance) ANOVA on signed errors found that participants in the new specific group only marginally reduced their signed error across all distances from pretest to posttest 1, F (1, 9) = 4.13, $p = 0.073$, $\eta_p^2 = 0.32$. Similar to our results in Experiment 1, we found a significant effect of distance, F (5, 45) = 33.10, $p < 0.001$, $\eta_p^2 = 0.79$. Again, we found that this effect indicated a significant negative linear trend of distance within posttest 1 errors, F (1, 9) = 78.20, $p < 0.001$, such that participants tended to

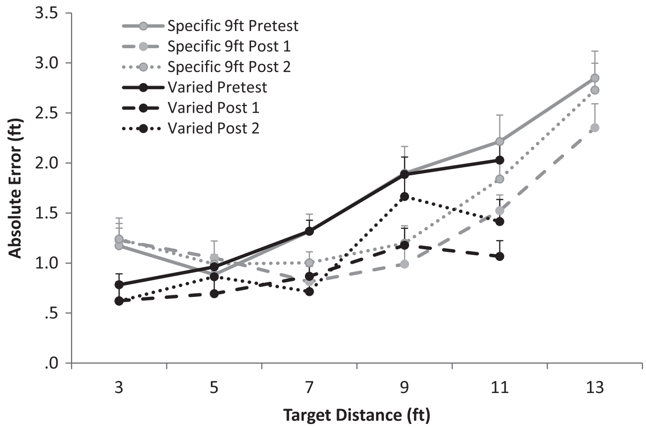

Fig. 4. Absolute errors across target distances. Average absolute errors on pretest (solid lines), posttest 1 (dashed lines), and posttest 2 (dotted lines) between the specific (Exp. 2 – gray lines) and the varied group (Exp. 1 – black lines) at 6 target distances.

undershoot to farther distances. There was no interaction between time and distance (see Fig. 5A). Additionally, between posttest 1 and posttest 2, there were no effects of time, but a similar effect of distance.

Using a 2 (Time) × 6 (Distance) ANOVA on absolute errors, we found that participants significantly improved from pretest to posttest 1, $F(1, 9) = 10.26$, $p = 0.011$, $\eta_p^2 = 0.53$. We also found a significant effect of distance, $F(5, 45) = 18.70$, $p < 0.001$, $\eta_p^2 = 0.68$. There was no interaction between time and distance (see Fig. 4). A marginal effect of time was found between posttest 1 and posttest 2, $F(1, 9) = 3.60$, $p = 0.090$, $\eta_p^2 = 0.286$, suggesting a small decrease in performance from pretest 1 to pretest 2. There was a similar effect of distance but no interaction, suggesting this was a relatively uniform decrease in performance across time.

Similarly, using a 2 (Time) × 6 (Distance) ANOVA on the variances, we found that participants significantly improved from pretest to posttest 1, $F(1, 9) = 6.65$, $p = 0.030$, $\eta_p^2 = 0.43$. We also found a significant effect of distance, $F(5, 45) = 6.02$, $p < 0.001$, $\eta_p^2 = 0.40$. However, there was no interaction between time and distance. Similar effects were found within posttest comparisons as we found with signed errors.

3.1.2. Exp. 1 vs Exp. 2

We next tested for differences in performance between the new specific group in Experiment 2 and the varied and specific groups (separately) from Experiment 1. All subsequent analyses will focus on between-group effects and interactions. The time and distance main effects can be characterized by the analyses reported above. The 13 ft target distance was not included in the following analyses since only participants in Experiment 2 were tested at this distance.

There were no group effects at pretest between the three groups. A 3 (Group) × 5 (Distance) ANOVA found neither a main effect of group ($p = 0.95$), nor an interaction effect ($p = 0.46$). Similar null effects were found at pretest for absolute errors and variances.

3.1.2.1. Pretest vs. Posttest 1. A mixed 2 (Group) × 2 (Time) × 5 (Distance) ANOVA was used to compare signed errors between the varied group from Experiment 1 and the specific group from Experiment 2. This analysis revealed a two-way interaction between distance and group, $F(2.47, 44.51) = 4.24$, $p = 0.015$, $\eta_p^2 = 0.19$, suggesting that groups performed differently across distances (Fig. 5A). Since we previously found no differences in pretest, we tested for this difference between groups at posttest 1. A 2 (Group) × 5 (Distance) ANOVA revealed that the interaction between group and distance persisted within posttest 1, $F(4, 72) = 7.99$, $p < 0.001$, $\eta_p^2 = 0.31$. Using t-tests, we compared groups at each of the five distances at posttest. We found that this group difference was only significant for the 11 ft target, $t(18) = 3.06$, $p = 0.007$, $d = 1.37$, and marginally significant for the 3 ft target, $t(18) = 2.79$, $p = 0.012$, $d = 1.25$, after correcting for multiple comparisons. This significant effect suggested that the varied group had better performance ($M = -0.15$, $SE = 0.35$) at the 11 ft target distance at posttest 1 than the specific group from Experiment 2 ($M = -1.37$, $SE = 0.23$), see Fig. 5A. Additionally, an analysis comparing the two specific groups from Experiments 1 and 2 found no differences between their signed errors (see Fig. 5B).

3.1.2.2. Posttest 1 vs. Posttest 2. A mixed 2 (Group) × 2 (Time) × 5 (Distance) ANOVA comparing signed errors between the varied group from Experiment 1 and the specific group from Experiment 2 revealed an interaction effect between distance and group, $F(4, 68) = 12.40$, $p < 0.001$, $\eta_p^2 = 0.42$, suggesting that the two groups performed differently across distances tested. In order to test for retention of this generalization, we used t-tests to compare differences in performance between groups at each distance within posttest 2. After correcting for multiple comparisons ($p_{crit} = 0.01$), we did not find any significant group differences at any of the five distances at posttest 2 (see Fig. 5A). It may be the case that while the varied group is able to generalize better in the short-term, this generalization is at the cost of long-term retention without additional practice. Since there was no effect of time, indicating that

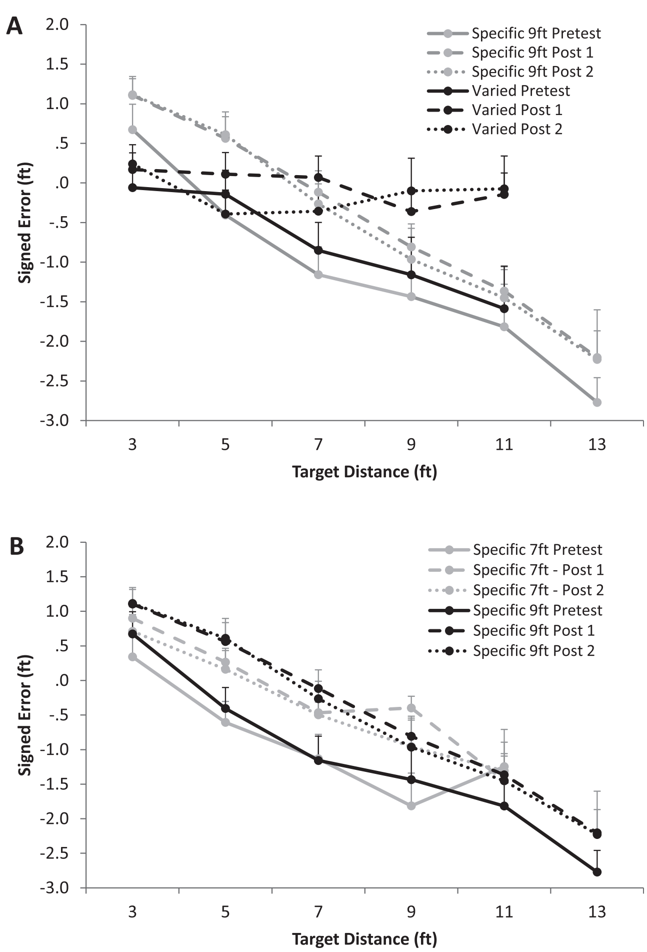

Fig. 5. Signed errors in both experiments replotted. (A) Average signed errors on pretest and posttests for the specific group in Experiment 2 and the varied group from Experiment 1. (B) Average signed errors on pretest and posttests for the specific group in Experiment 2 and posttests for the specific group from Experiment 1.

no significant change occurred between the two posttests, we looked at the interaction effect between distance and group further by collapsing across time and comparing the two groups at each distance. We found a significant effect at 3 ft ($p = 0.007$) and a marginally significant effect at 11 ft ($p = 0.015$), after correcting for multiple comparisons. However, these group effects found after collapsing across posttest seem to mainly be driven by posttest 1 results. Additionally, we found no differences in posttest performances between the two specific groups. We found similar but weaker effects using absolute error and variance in our analyses between groups (see Fig. 4).

To recap, we found that at the 11 ft target distance, the varied group had significantly greater improvement from pretest to posttest 1 as compared to the specific group that practiced solely at the 7 ft distance. This effect persisted when comparing the varied group against the specific group in Experiment 2, who shared the training distance of 9 ft with the varied group. The varied group's advantage at 11 ft was reduced to non-significance by posttest 2 without further practice suggesting that after long-term retention, the varied training schedule was not beneficial.

4. General discussion

The current study investigated the effect of long-term practice and retention of a simple throwing task. Consistent with the previous experiments, we found generalized effects of motor learning to untrained distances. However, we found that this occurred in both the varied and the specific groups. In both experiments, the specific groups, regardless of their trained distance, performed similarly. While the specific groups were able to generalize their training to untrained distances, the varied group outperformed them at the 11 ft target distance. There is also some evidence to suggest that the varied group also had a slight advantage at the 3 ft target at posttest 1. However, any benefits due to the varied practice schedule disappeared by posttest 2. This finding suggests that there was only a relatively short-term benefit to the varied practice schedule without additional training.

These data with adult participants did not support the findings of Kerr and Booth (1978), who suggested that the varied group should outperform the specific group at 7 ft. At 7 ft (analogous to the central target distance in their study), we found that all groups performed comparably. However, our data still support the idea that a varied group may develop a better generalized motor program for more difficult distances than the specific group, at least after a short retention period. Using signed errors, we were able to investigate trends within each groups' performance. Importantly, the specific group tended to consistently underthrow to their target distance, while the varied group tended to center their throws around the target distance, as seen by signed error measurements. The negative linear trend in the specific group's posttest signed errors suggests that there may indeed be a bias towards the trained distance. Within the varied group, no such linear trend was present, suggesting that training at two distances as compared to one reduced the systematic errors across distances at posttest 1. Nevertheless, these results suggest that specific practice still gave rise to a moderately generalized motor program that performed well at nearby untrained distances while maintaining the effects from the specific practice.

Specificity in learning has mostly been demonstrated in perceptual learning paradigms (see reviews in Fahle, 2005; Fahle & Poggio, 2002; Gibson, 1969; Sagi, 2011). In these experiments, learning is specific to the trained attributes with little transfer to untrained attributes. For example, during a motion direction discrimination task, participants improved at discriminating motion directions with a fine difference around the trained direction. If a test direction was different from this trained direction, motion discrimination partially transferred but tapered off (Ball & Sekuler, 1987). Our data here also support this level of specificity of learning. We found that in both groups, generalization of performance was limited to a region around the trained distance. Indeed, this finding is the main message of the current study. Namely, transient differences aside, motor learning in our throwing task partially generalized around the trained target distances, regardless of variations in training schedules.

Recent findings in motor learning suggest initial and induced variabilities during practice influence the rate of learning and relearning. Wu, Miyamoto, Gonzalez Castro, Olveczky, and Smith (2014) found that initial temporal variability in motor learning predicted the rate of learning during a simple reaching task. During the exploratory phase of learning, greater initial variability may expose the participant to a variety of similar movements for the same motor command in order to fine tune motor parameters to achieve the desired outcome. Similarly, recent work in both motor and perceptual learning suggests that exposure to untrained or task-irrelevant conditions can improve learning rates in those untrained conditions (Nishina, Seitz, Kawato, & Watanabe, 2007; Seitz & Watanabe, 2009; Wymbs et al., 2016; Yin, Bi, Yu, & Wei, 2016). While this idea was not the aim of the current experiments, nor was it tested directly, one may suspect that the generalization that we found in both groups may suggest that participants were indirectly exposed to nearby untrained target distances due to variability in throws during practice. We did not find that initial variability in the first few practice blocks related to greater generalizability in posttests in this dataset. However, this idea should further be tested in future studies designed to directly test this.

Nevertheless, in visual motion direction discrimination learning, variation did appear to promote generalization, as follows (see also Xiong, Xie, & Yu, 2016). Rather than training along a single reference direction, as was traditionally the case, Liu and Vaina (1998) simultaneously trained motion discrimination along two reference directions. Critically, the training trials along these two reference directions were unbalanced. Two-thirds of the trials were along one reference direction, and the remaining one-third of trials were along the other reference direction. On average, therefore, training along the second reference direction lagged behind the first reference direction. Liu and Vaina (1998) hypothesized that if such design enabled transfer between these two reference directions (where traditionally only specificity had been shown), then the learning rate along the second reference direction would benefit more from the first reference direction. Indeed, they found that the learning rate along the second direction was greater than along the first direction.

To summarize, the current study investigated the generalizability and long-term retention of a simple motor skill in an adult sample using different practice schedules. Given that previous research using the schema theory has suggested that varied schedules of practice yielded better performance than specific schedule of practice, we investigated this claim in a similar motor learning paradigm and found limited support for the schema theory. However, this result was obtained in only one specific task of motor learning. Additional studies in motor and other modality learning with a variety of tasks are needed to assess the applicability of the schema theory in general.

Acknowledgment

We thank the Zili Lab research assistants for collecting and entering data. This research was supported in part by an NSF grant (BCS 0617628) to Zili Liu.

Please cite this article in press as: Willey, C. R., & Liu, Z. Long-term motor learning: Effects of varied and specific practice. *Vision Research* (2017), http://dx.doi.org/10.1016/j.visres.2017.03.012

References

Ball, K., & Sekuler, R. (1987). Direction-specific improvement in motion discrimination. *Vision Research, 27*(6), 953–965.

Birnbaum, M. S., Kornell, N., Bjork, E. L., & Bjork, R. A. (2012). Why interleaving enhances inductive learning: The roles of discrimination and retrieval. *Memory & Cognition, 41*(3). http://dx.doi.org/10.3758/s13421-012-0272-7.

Fahle, M. (2005). Perceptual learning: Specificity versus generalization. *Current Opinion in Neurobiology, 15*, 154–160.

Fahle, M., & Poggio, T. (2002). *Perceptual learning*. Cambridge, MA: MIT Press.

Gibson, E. J. (1969). *Principles of perceptual learning*. New York: Appleton-Century-Crofts.

Janiszewski, C., Noel, H., & Sawyer, A. G. (2003). A meta-analysis of the spacing effect in verbal learning: Implications for research on advertising repetition and consumer memory. *Journal of Consumer Research, 30*(1), 138–149. http://dx.doi.org/10.1086/374692.

Kerr, R., & Booth, B. (1978). Specific and varied practice of motor skill. *Perceptual and Motor Skills, 46*, 395–401.

Liu, Z., & Vaina, L. M. (1998). Simultaneous learning of motion discrimination in two directions. *Cognitive Brain Research, 6*(4), 347–349.

Nishina, S., Seitz, A. R., Kawato, M., & Watanabe, T. (2007). Effect of spatial distance to the task stimulus on task-irrelevant perceptual learning of static gabors. *Journal of Vision, 7*(2), 2.1–2.10. http://dx.doi.org/10.1167/7.13.2.

Sagi, D. (2011). Perceptual learning in vision research. *Vision Research, 51*, 1552–1566.

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review, 82*(4), 225–260.

Schmidt, R., & Bjork, R. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3*(4), 207–217.

Seitz, A. R., & Watanabe, T. (2009). The phenomenon of task-irrelevant perceptual learning. *Vision Research, 49*(21), 2604–2610. http://dx.doi.org/10.1016/j.visres.2009.08.003.

Shea, C. H., & Wulf, G. (2005). Schema theory: a critical appraisal and reevaluation. *Journal of Motor Behavior, 37*(2), 85–101. http://dx.doi.org/10.3200/JMBR.37.2.85-102.

Wu, H. G., Miyamoto, Y. R., Gonzalez Castro, L. N., Olveczky, B. P., & Smith, M. A. (2014). Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. *Nature Neuroscience, 17*(2), 312–321. http://dx.doi.org/10.1038/nn.3616.

Wymbs, N. F., Bastian, A. J., & Celnik, P. A. (2016). Motor skills are strengthened through reconsolidation. *Current Biology, 26*(3), 338–343. http://dx.doi.org/10.1016/j.cub.2015.11.066.

Xiong, Y. Z., Xie, X. Y., & Yu, C. (2016). Location and direction specificity in motion direction learning associated with a single-level method of constant stimuli. *Vision Research, 119*, 9–15. http://dx.doi.org/10.1016/j.visres.2015.11.005.

Yin, C., Bi, Y., Yu, C., & Wei, K. (2016). Eliminating direction specificity in visuomotor learning. *The Journal of Neuroscience, 36*(13), 3839–3847. http://dx.doi.org/10.1523/JNEUROSCI.2712-15.2016.

Please cite this article in press as: Willey, C. R., & Liu, Z. Long-term motor learning: Effects of varied and specific practice. *Vision Research* (2017), http://dx.doi.org/10.1016/j.visres.2017.03.012