MAINTAINING MOTOR SKILL PERFORMANCE (U)

JOSEPH D. HAGMAN, Ph.D.
U.S. ARMY RESEARCH INSTITUTE
ALEXANDRIA, VA 22333

The Army's primary peacetime mission is to maintain combat readiness (5). To be combat ready, soldiers must first become proficient in their performance of assigned job tasks, and then retain this proficiency over what can be prolonged periods of no practice. One way to enable soldiers both to reach and maintain combat readiness is through the use of training methods that promote effective task acquisition and retention. But first, these methods must be identified and compared.

Most training methods have been investigated within the context of laboratory experiments. In these experiments, training of motor tasks has involved the execution of presentation (p) and test (t) trials. At p-trials persons perform the criterion movement to be learned, whereas during t-trials they attempt to reproduce (i.e., recall) the criterion from memory. P- and t-trials can be distinguished procedurally by the presence or absence of a movement constraint. During the training of straight arm (i.e., linear positioning) movements, for example, persons perform p-trials by moving a sliding mechanism along a horizontal track until contacting a stop positioned by the experimenter to define criterion movement distance (extent) and location (end position). At t-trials, an attempt is made to recall the criterion with the stop removed.

The number and sequential arrangement of p- and t-trials performed during training has differed across experiments. Some investigators (16) have alternated p- and t-trials during training, whereas others (2, 4) have emphasized one type of trial over another through repetition. Unfortunately, for those of us responsible for training motor tasks in the operational environment, the relative acquisition and retention benefits associated with each method have not been identified. Thus, it is difficult to select the method(s) that is most effective.

The question of which method(s) is most effective remains an empirical one because of uncertainty regarding the relative contribution of p- and t-trials to motor acquisition and retention. While most researchers and trainers would agree that p-trials contribute to both acquisition and retention, either because they provide reinforcement (2) or because they gen-
create movement feedback information needed to form a reference for movement recall (1), the contribution of t-trials is still a matter for debate. On the one hand, t-trials may be detrimental because they are different from the criterion as a result of error in recall. Thus, they are capable of causing interference (6) and depressing performance. From this perspective, training methods that avoid t-trials and emphasize p-trials should be most effective. On the other hand, t-trials may be beneficial to movement acquisition and retention because they provide trainees with an opportunity to define their own movements during training. Researchers have shown that movements performed under a learner-defined execution mode, i.e., unconstrained by a stop, are retained better than those performed under an experimenter-defined mode, i.e., constrained by a stop (7, 8). Thus, training methods that include learner-defined t-trials may encourage trainees to rely on their retention of t-rather than p-trials, and thereby, promote movement recall. From this perspective, training methods should not only include but emphasize t-trials to be most effective.

The present paper describes three laboratory experiments. The first two compare the effects of p- and t-trials on movement distance (Experiment 1) and location (Experiment 2) within the context of three different training methods. The general approach in these experiments was to repeat or alternate p- and t-trials during training and to examine the effect of this manipulation on movement recall during acquisition as well as after short- and long-term retention intervals. The third experiment examines the validity of alternative interpretations offered to explain the results obtained in the first two experiments. The laboratory task chosen for training was linear positioning.

**Experiment 1**

**Method**

**Subjects.** Forty-five employees (33 men and 12 women) of the Army Research Institute (ARI) volunteered to serve as subjects in the experiment.

**Apparatus.** Subjects performed movements from their left to right using a metal slide that ran along a linear track consisting of two stainless steel rods 35 inches (88.9 cm) in length. Two Thompson Ball Bushings supported the slide on the rods which were mounted in parallel on a metal frame 4.25 inches (11 cm) apart and 11 inches (27.94 cm) above the base of the frame. The base rested on a standard table top 31 inches (78.74 cm) from the floor. A second slide was used by the experimenter to stop the first slide along the track during p-trials. A pointer attached to the experimenter's side of each slide ran along a meter stick to indicate respective slide position. Additional apparatus included; a chin rest to control head and body position, earphones through which subjects received tape-recorded procedural commands, and a blindfold to eliminate visual cues.

**Design.** The experiment contained an acquisition and a retention segment, as shown in Figure 1. The acquisition segment consisted of 18 train-
<table>
<thead>
<tr>
<th>GROUPS</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6/7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12/13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>3 min</th>
<th>24 hr</th>
</tr>
</thead>
<tbody>
<tr>
<td>STANDARD</td>
<td>P</td>
<td>T</td>
<td>P</td>
<td>T</td>
<td>P</td>
<td>T</td>
<td>P</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
<tr>
<td>TEST</td>
<td>P</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>P</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
</tr>
</tbody>
</table>

**Figure 1.** Acquisition and retention trial sequences for each training method group.
Fifteen trials divided into three cycles of six trials each. Each cycle contained p- and t-trials. P-trials were experimenter-defined movements constrained by the stop. The distance between movement starting position and the stop defined the criterion distance of 9.84 inches (250 mm). T-trials were learner-defined recall movements performed with the stop removed. The sequence of p- and t-trials performed within cycles differed for each of three training method groups. For the STANDARD group a cycle consisted of three p- and three t-trials administered in an alternating sequence. For the PRESENTATION group, the first five trials of each cycle were p-trials and the sixth was a t-trial. For the TEST group, the first trial was a p-trial and the next five were t-trials. The retention segment of the experiment consisted of a single t-trial performed at both 3 minutes and 24 hours after the final training trial.

Fifteen subjects were assigned randomly to each of the three training method groups with the restriction that each group contain the same proportion of men and women. Including the two retention trials, subjects performed a total of 20 movements during the experiment. Each movement began from a different starting position that varied between 20 and 30 mm, in increments of 20 mm, from the subject's left end of the apparatus. Three random sequences were developed for the 20 movement starting positions.

Procedure. At the beginning of the experiment, subjects were instructed to learn and remember movement distance and shown how distance was separated procedurally from location through changes in movement starting position. They were then shown a written copy of the entire p- and t-trial command sequence appropriate to their training group and told the meaning of each command. "Movement" was the p-trial command and "Recall Movement" was the t-trial command. Each was preceded by "Ready" and followed by "Rest." At "Ready," the experimenter grasped the subject's right hand and placed it on the handle of the slide. Five seconds later, subjects heard either "Movement" or "Recall Movement" depending on their training group membership and the specific training trial. At "Movement," they moved the slide along the track at a moderate (approximately 125 mm/second) pace until contacting the stop. At "Recall Movement," the stop was removed and subjects moved the slide along the track and stopped it when they felt they had moved it the criterion distance. Five seconds were allowed for p- and t-trial execution. During this interval, white noise was delivered through the earphones to eliminate auditory cues resulting from displacement of the slide. "Rest" marked the start of a 10-second rest interval during which subjects removed their hand from the slide and placed it in a predetermined resting position on the table. Concurrently, the experimenter recorded recall accuracy (only for t-trials) and repositioned either the slide alone or both it and the stop in preparation for the next trial. After "Rest," subjects heard "Ready" and the sequence of commands began for the next trial. During the retention segment of the experiment, intervals of 3 minutes and 24 hours were inserted between "Rest" and "Ready." Before training began, subjects were asked not to count while moving the slide. They
then donned their blindfold and earphones and received an opportunity to move the slide to get a feel for its movement characteristics.

Results and Discussion

Although both algebraic (signed) and absolute (unsigned) error were recorded for all t-trials, algebraic error revealed no significant effects of interest. Therefore, only absolute error results are reported.

Acquisition. Mean absolute error scores for acquisition t-trials are shown on the left side of Figure 2. The acquisition curves for the STANDARD and PRESENTATION groups were negatively accelerated and similar in appearance when considering only those t-trials that coincided temporally for both groups, i.e., Trials 6, 12, and 18. The TEST group's curve was above those of the other two groups and had a serrated appearance. A Groups x Trials analysis of variance (ANOVA) performed on the scores for t-trials common to all three groups (i.e., Trials 6, 12, 18) revealed a significant main effect of trials, $F(2,34) = 4.74$, $MSe = 529.24$, and groups $F(2,42) = 4.48$, $MSe = 663.45$, with the rejection region for this and all other analyses being .05. Individual comparisons, using the least significant difference method (3), showed that the trials effect was caused by a decrease in error between Trials 6 and 12 for all groups and that the groups effect resulted from TEST group error scores being greater than those of the STANDARD and PRESENTATION groups, while the scores for the latter two groups did not differ significantly.

Visual inspection of all TEST group acquisition t-trial scores shown in Figure 2 revealed that a build up error within cycles was responsible for the difference found between the TEST and the other two groups. A subsequent Cycles (1-3) x Trials (1-5) ANOVA on all TEST group t-trial scores revealed a significant effect of trials, $F(4,56) = 4.24$, $MSe = 292.54$, with individual comparisons indicating that error was greater on the last t-trial of each cycle than on the first t-trial of each cycle. This within-cycle error increase is consistent with past motor research findings (4) and indicates that t-trial repetition during training adversely affects acquisition of movement distance information when measured within cycles.

Further inspection of TEST group scores revealed that error decreased considerably between cycles. This decrease offset the error increase that occurred within cycles and eliminated group differences found at common t-trials. A separate Groups x Trials ANOVA comparing STANDARD and PRESENTATION group scores on the last t-trial of each cycle, i.e., Trials 6, 12 and 18, with TEST group scores on the first t-trial of each cycle, i.e., Trials 2, 8 and 14, still showed a significant effect of trials, $F(2,84) = 4.78$, $MSe = 593.49$, but no significant effect of groups. Failure to find a groups effect indicated that TEST group performance at Trial 14, for example, after only 3 p-trials did not differ from Trial-18 performance of the STANDARD and PRESENTATION groups after 9 and 15 p-trials, respectively. Thus, t-trials appear to have potentiated or increased the effectiveness of p-trials and indirectly contributed to the acquisition of movement...
For all days after the training, the acquisition data demonstrate that: (a) p-trials contribute positively to acquisition; (b) t-trials produce both positive between-cycle and negative within-cycle effects on acquisition; and that (c) p-trial repetition and p- and t-trial alternation during training produce a higher level of acquisition than t-trial repetition, but that this difference depends on the specific trials at which training methods are compared.

**Retention.** Mean absolute error retention scores for distance are shown on the right side of Figure 2. Scores were analyzed using a Groups (STANDARD, PRESENTATION, TEST) x Retention Interval (Immediate, 3 minutes, 24 hours) mixed factorial ANOVA with Trial-18 scores included to reflect immediate recall at the end of training. This ANOVA revealed a significant Groups x Retention Interval interaction, F(4,84) = 4.10, MSE = 456.68, resulting from an increase in error over time for the STANDARD and PRESENTATION groups and a decrease in error for the TEST group. Simple effect comparisons indicated that Trial-18 error for the TEST group was greater than that for both the STANDARD and the PRESENTATION groups, and that the error for these latter two groups did not differ (i.e., TEST > PRESENTATION = STANDARD). Three minutes after training, there was no difference between TEST and PRESENTATION group error scores, although their combined mean error was greater than the error for the STANDARD group (i.e., TEST = PRESENTATION > STANDARD). At 24 hours after training, TEST group error was less than STANDARD which, in turn, was less than PRESENTATION group error (i.e., TEST < STANDARD < PRESENTATION).

In summary, the retention results indicate that: (a) p-trial repetition during training results in both rapid and extensive forgetting; (b) t-trial repetition prevents forgetting and promotes increased recall proficiency over time; and (c) alternation of p- and t-trials prevents short-term forgetting and reduces long-term forgetting.

**Experiment 2**

The second experiment was a replication of the first except that location rather than distance was the movement cue examined. Separate investigation of location was done to test whether the results obtained for distance would generalize to another cue thought to underlie movement acquisition and retention.

**Method**

**Subjects.** An additional 45 ARI employees (33 men and 12 women) volunteered to serve as subjects in this experiment.

**Apparatus.** The same linear positioning and support apparatus were used as in Experiment 1.

**Design.** The general design was identical to that of Experiment 1. Movement starting positions varied between 20 and 380 mm, in increments of 20 mm, from the subject's left of the criterion location, set at 520 mm...
from the subject's left end of the apparatus. Three random starting position sequences were generated; five subjects from each group were trained under one of the three sequences.

Procedure. The procedure was identical to that of Experiment 1 except that subjects were instructed to learn location rather than distance.

Results and Discussion.

Analysis of algebraic error revealed no significant differences of interest. Thus, only the results of absolute error analyses are reported.

Acquisition. Mean absolute error acquisition scores are shown on the left in Figure 3. The pattern of performance found for location was similar to that found earlier for distance. A Groups x Trials ANOVA performed on error scores at common t-trials revealed significant main effects of trials, \( F(2,84) = 6.16, \text{MSE} = 272.02 \), and groups, \( F(2,42) = 8.07, \text{MSE} = 539.72 \). Individual comparisons revealed that the trials effect was caused by a decrease in error between Trials 6 and 18 for all groups and that the groups effect resulted from error being greater for the TEST group than for the STANDARD and PRESENTATION groups, with no difference in error present between these latter two groups. A separate Cycle x Trials ANOVA on all TEST group t-trial scores showed that a significant increase in error occurred within cycles, \( F(4,56) = 5.28, \text{MSE} = 178.64 \). Individual comparisons revealed that the combined mean error for the last four t-trials of each cycle differed from that of the first t-trial of each cycle.

Location TEST group scores also showed a decrease in between-cycle error. This decrease offset the increase in error within-cycles and eliminated group differences found at common t-trials. A Groups x Trials ANOVA comparing STANDARD and PRESENTATION group scores on the last t-trial of each cycle with TEST group scores on the first t-trial of each cycle showed a significant effect of trials, \( F(2,84) = 13.83, \text{MSE} = 217.68 \), but no effect of groups. Thus, the p-trial potentiation effect of t-trials found for distance was also evident for location.

In summary, the acquisition data for location demonstrate that: (a) p-trials contribute positively to acquisition; (b) t-trials produce both positive between-cycle and negative within-cycle effects on acquisition; and that (c) p-trial repetition and p- and t-trial alternation during training produce a higher level of acquisition than that produced by t-trial repetition, but that this difference depends on the specific trials at which training methods are compared.

Retention. Mean absolute error retention scores for location are shown on the right side of Figure 3. A Groups \( \times \) Retention Interval mixed factorial ANOVA revealed a significant main effect of retention interval, \( F(2,84) = 12.53, \text{MSE} = 312.86 \), and a significant Retention Interval \( \times \) Group interaction, \( F(4,84) = 7.03, \text{MSE} = 312.86 \). Individual comparisons of simple effects confined to the interaction revealed that at Trial 18 TEST group error was greater than either STANDARD or PRESENTATION group error, and that no error difference existed between these latter two groups (i.e.,...
Figure 3. Mean location absolute error at acquisition and retention t-trials for each training method group.
TEST > PRESENTATION = STANDARD). Three minutes after training, no differences were present among groups (i.e., TEST = PRESENTATION = STANDARD). At 24 hours after training, TEST group error was less than STANDARD and PRESENTATION group error, while the error for these latter two groups did not differ (i.e., STANDARD = PRESENTATION > TEST).

In summary, the retention results indicate that: (a) p-trial repetition during training results in both rapid and extensive forgetting over time; (b) t-trial repetition prevents forgetting; and (c) p- and t-trial alternation produces the same effects as p-trial repetition.

In general, the results of both experiments demonstrate that: (a) both p- and t-trials affect motor task acquisition and retention, and (b) training method effectiveness depends on the type of trial repeated. P-trial repetition improves acquisition. This improvement presumably is derived from at least two sources; prevention of errors and repetition of the criterion movement. By preventing errant movement at t-trials, the possibility of interference during training is avoided. At the same time, an accurate memorial representation of the criterion is developed through repeated practice.

In contrast to the consistent beneficial effects of p-trial repetition, t-trial repetition produces mixed acquisition effects in the form of increased error within cycles and decreased error between cycles. Increases in within-cycle error result from interference generated by repetition of t-trials. These t-trials are different from the criterion because of imperfect recall by the subject, and therefore possess interference capabilities (6). In contrast, the decreases in between-cycle error are the result of p-trial potentiation following t-trial repetition. Following t-trial repetition, subjects presumably know more about their recall than after p-trial repetition. Thus, they are better able to discriminate their recalled t-trial movement from the p-trial criterion movement, and are more capable of making appropriate adjustments needed to increase post p-trial recall proficiency. Although speculative, the general notion that movement discrimination ability increases with repeated recall is consistent with past findings (10).

P- and t-trials also had different effects on retention. P-trial repetition resulted in rapid and extensive forgetting, whereas t-trial repetition prevented forgetting of both distance and location. T-trials may benefit retention for at least two reasons. First, they may be easier to remember than p-trials because of superior encoding characteristics resulting from being performed under a learner-defined movement mode, i.e., unconstrained by a stop (7,14). Learner-defined movements allow subjects to predict or anticipate movement consequences (e.g., proprioception) via corollary discharge prior to movement initiation. Prediction provides a context for receiving movement produced cues which enhances their encoding and subsequent retrieval (13). In contrast, prediction is more difficult at p-trials because movements are experimenter-defined i.e., constrained by a stop, and are not under total control of the subject. Of course, even though t-trials may be easier to remember than p-trials, subjects still
A second reason why t-trials aid retention relies on the notions of movement variability and motor schema. The motor schema is an abstraction of task and environmental characteristics that develops through repeated and varied movement during training (12), and serves as a rule or concept for movement generation. As the schema develops, the abstracted information becomes more resistant to forgetting than that resulting from specific movement instances (11). In the first two experiments, variability was generated at t-trials because subjects were variable in their recall. In contrast, no variability was generated at p-trials because every p-trial was identical. Thus, it could be argued that schema strength was greater for the TEST group than for the other two groups and this promoted superior TEST group retention.

Experiment 3

This experiment was performed to determine the relative validity of the two aforementioned reasons for enhanced retention following t-trial repetition during training. It involved adding yoked p-trial groups to the t-trial distance and location TEST groups of Experiments 1 and 2. Yoking required the use of a stop to ensure that yoked groups performed p-trials that were identical to the t-trials performed by the two TEST groups. Thus, yoking afforded the means of equating p- and t-trials for variability during training, but allowed the distinction between p- and t-trial execution modes (i.e., experimenter- versus learner-defined) to remain. If variability per se during training is the key to enhanced retention, then one would expect no difference in retention scores between the two t-trial TEST groups and their yoked p-trial counterparts. If, on the other hand, movement execution mode is the key, then the two TEST groups should display better retention than that of the two yoked groups.

Method

Subjects. Thirty additional ARI employees (22 men and 8 women) volunteered to serve as subjects in the experiment. Fifteen subjects were randomly assigned to the yoked distance group and the yoked location group with the restriction that each group contain the same proportion of men and women.

Apparatus. The apparatus were identical to those used in Experiment 1 and 2.

Design. Four training method groups were included in the design, i.e., yoked distance, yoked location, and the two TEST groups from Experiments 1 and 2. The only difference between the yoked and TEST groups was that yoked groups performed p-trials identical to the t-trials performed by the TEST groups. Data from the two yoked groups were collected in this experiment.
Procedure. Subjects were instructed to learn either movement distance or location depending on their group. Those in the yoked groups were told a yoking procedure. The rest of the procedure was identical to that in Experiments 1 and 2.

Results and Discussion

Acquisition. Because yoked groups performed the same movements as TEST group counterparts, no differences in acquisition scores were found.

Retention. Separate Groups x Retention Interval mixed factorial ANOVAs were performed on absolute error retention scores for the two and two distance groups. The means for the distance groups are shown on the left and those for the two location groups are shown on the right in Figure 4.

For distance, a significant Groups x Retention Interval interaction was found F(1,28) = 6.85, MSe = 454.02. Individual comparisons revealed no group differences occurred at 3 minutes after training, but that TEST group was superior to the yoked group 24 hours later. The Groups x Retention Interval interaction for location was not significant F(1,28) = .05 < p < .10. However, a priori expectations about the retention effects of t-trials based on the significant findings distance justified further examination of simple effects. Individual comparisons indicated that group performance did not differ 3 minutes after training, but did differ 24 hours after training with the TEST group being superior to the yoked group. These data are consistent with the interpretation that differences in p- and t-trial movement execution mode, and not increased variability associated with t-trial repetition, caused the l-long-term retention following t-trial repetition in Experiments 1 and 2. Yoking increased p-trial variability during training, yet was unable to prevent forgetting of either distance or location.

If one is to accept the execution mode interpretation of t-trial retention benefits, then it needs to be explained why the error level was so for the distance and location TEST groups 24 hours after training. If t-trial retention is based on the retention of learner-defined t-trials, one might expect an error level equal to that of either the mean error at training t-trials or the error displayed on the last training t-trial of which was the case. Figures 2 and 3 show that the error observed 24 hours after training was similar to that displayed at the first t-trial of Cycle 3 (i.e., Trial 14). To account for this, it is suggested that t-trial repetition produces a form of within-cycle serial learning where subjects attempt to recall each preceding t-trial of the series during training. Because of the strong primacy effect usually found for motor tasks (9), they are able to remember their recall from the t-ttrial of the series better than their recall from later t-trials of
Figure 4. Mean distance and location absolute error on retention trials for TEST and YOKED training method groups.
the series at the time of long-term retention testing. The influence
of these later t-trials, however, appears in the form of increased with-
cycle error during training and retroactive interference during retention
testing. If true, then the unusual decrease in error found between 3 min-
utes and 24 hours after training for the distance TEST group in Experiment
1 may have been caused by dissipation of retroactive interference generated
by the last four training t-trials of Cycle 3 (i.e., Trials 15-18). At the
same time, the rate of dissipation may not have been rapid enough to pro-
duce a decrease in error between the end of training and the 3- minute re-
tention test. It is also possible that differences in the rate of dissi-
pation for distance and location cues may have been responsible for the
lack of a significant decrease in error over time for the location TEST
group of Experiment 2. Finally, with p- and t-trial alternation, retro-
active interference did not build up because t-trials were not repeated
during training. Thus, both short- and long-term retention benefits were
found, at least for distance. Of course the above explanation remains
speculative and requires considerable additional research. However, the
general notion that retroactive interference dissipates over time for motor
tasks has already received support (15).

General Conclusions

The results of the present experiments clarify certain training issues
regarding the relative contributions of presentation and testing to motor
task performance. In doing so, they answer the Army's question of which
training methods most effectively promote the highest levels of skill ac-
quision and retention. First, the results indicate that the goal of
training should dictate the training method used. If the goal is a consist-
ant and high level of acquisition, then training should emphasize repeated
presentation or alternation of presentation with testing, with the latter
method being preferred because of enhanced short- and long-term retention.
If effective long-term retention is the goal, then training should empha-
size repeated testing. Second, testing should not only be viewed merely as
a means of evaluation, but also as a means of improving motor skill perfor-
ance. And third, the enhanced retention associated with repeated testing
can be achieved by merely changing the emphasis of training from presen-
tation to testing. This could be done without the usual negative aspects of
additional expenditures in training time, money and personnel.

References


