ARI Research Note 86-55

AD-A168 572

NONIMPOSED OVERPRACTICE AND SKILL RETENTION

Marshall B. Jones
The Milton S. Hershey Medical Center
The Pennsylvania State University

for

Contracting Officer's Representative
M. Drillings

BASIC RESEARCH LABORATORY
Milton R. Katz, Director

U. S. Army
Research Institute for the Behavioral and Social Sciences

April 1986

86 6 9 11

U. S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency under the Jurisdiction of the Deputy Chief of Staff for Personnel

EDGAR M. JOHNSON
Technical Director

WM. DARRYL HENDERSON
COL, IN
Commanding

Technical Review by
Steve Kronheim

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: ARI Research Note 86-55	2. GOVT ACCESSION NO.: ADA 168572
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle): NONIMPOSED OVERPRACTICE AND SKILL RETENTION	5. TYPE OF REPORT & PERIOD COVERED: Final Report June 83 - December 85
	6. PERFORMING ORG. REPORT NUMBER: - -
7. AUTHOR(s): Marshall B. Jones	8. CONTRACT OR GRANT NUMBER(s): MDA 903-83-K-0246
9. PERFORMING ORGANIZATION NAME AND ADDRESS: The Milton S. Hershey Medical Center, The Pennsylvania State University, Hershey, Pennsylvania 17033	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 2Q161102B74F
11. CONTROLLING OFFICE NAME AND ADDRESS: U.S. Army Research Institute for the Behavioral and Social Sciences, 5001 Eisenhower Avenue, Alexandria, VA 22333-5600	12. REPORT DATE: April 1986
	13. NUMBER OF PAGES: 28
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): - -	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE: - -

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

- -

18. SUPPLEMENTARY NOTES

M. Drillings, contracting officer's representative

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Training
Video-Computer Skills
Retention
Overpractice
Refresher Training

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The question at issue in this research note is the role of "nonimposed overpractice" in predicting individual performance in skill retention or reacquisition. "Nonimposed overpractice" refers to the shape of individual performance curves late in acquisition. The flatter a curve is, and the longer it remains flat, or nearly so, the more that the individual has overpracticed the task. The overpractice is "nonimposed" because there is no external constraint or scoring convention that prevents further improvement. Performance flattens out

(over)

ARI Research Note 86-55

20. Abstract (continued)

(shows little or no further improvement) because the individual has reached his or her internal limits.

NONIMPOSED OVERPRACTICE AND SKILL RETENTION

BRIEF

Requirement:

To determine the role of "nonimposed overpractice" in predicting individual performance at reacquisition or retention. "Nonimposed overpractice" refers to the shape of individual performance curves late in acquisition. The flatter a curve is and the longer it remains flat or nearly so the more that individual has overpracticed the task. The overpractice is "nonimposed" because there is no external constraint or scoring convention that prevents further improvement. Performance flattens out (shows little or no further improvement) because the individual has reached his or her internal limits.

Procedure:

Three experiments were carried out. In the first, 27 Navy enlisted volunteers practiced each of six video-computer tasks for 15 daily sessions. Practice was resumed after a no-practice interval lasting 4-6 months for two of the tasks, 10-12 months for two others, and 16-18 months for the remaining two tasks. Reacquisition consisted of five sessions of daily practice, with the same number of trials per day as was used in acquisition. For all six tasks the measure of performance appeared on the viewing screen throughout practice; it was the quantity a sailor was trying to maximize.

Three of these same video-computer tasks were used in Experiment 2. Three groups of 50, 56, and 53 college students practiced one of the three tasks for five sessions, each session lasting 16 minutes. After a no-practice interval lasting 4 months, each student was given three reacquisition sessions.

The third experiment utilized two of the three tasks in Experiment 2. With one task (N=56) practice conditions were identically the same as in Experiment 2; with the other task (N=95) practice was extended to nine sessions. All other conditions were the same as in Experiment 2.

Findings:

In all three experiments unequivocal evidence was obtained that nonimposed overpractice improves retention. The longer a subject continues to practice a task, even though performance in acquisition is no longer improving, the better retention is likely to be. This effect is additional to level of performance; that is, nonimposed overpractice improves performance even among individuals who perform well in acquisition and who would, therefore, be expected to perform better than others at retention.

Utilization of Findings:

Considerable periods of time usually intervene between the end of training and the operational use of acquired skills. As a result, soldiers who are proficient at the end of training are not necessarily proficient when first called upon to use what they have learned. Selective refresher training is a major and relatively inexpensive remedy to this loss of proficiency. It depends, however, on accurately predicting which soldiers will not be proficient after a given interval of time if refresher training is not given. Hence, improving prediction by taking nonimposed overpractice into account has direct benefits in making selective refresher training more effective.

NONIMPOSED OVERPRACTICE AND SKILL RETENTION

CONTENTS

LIST OF TABLES

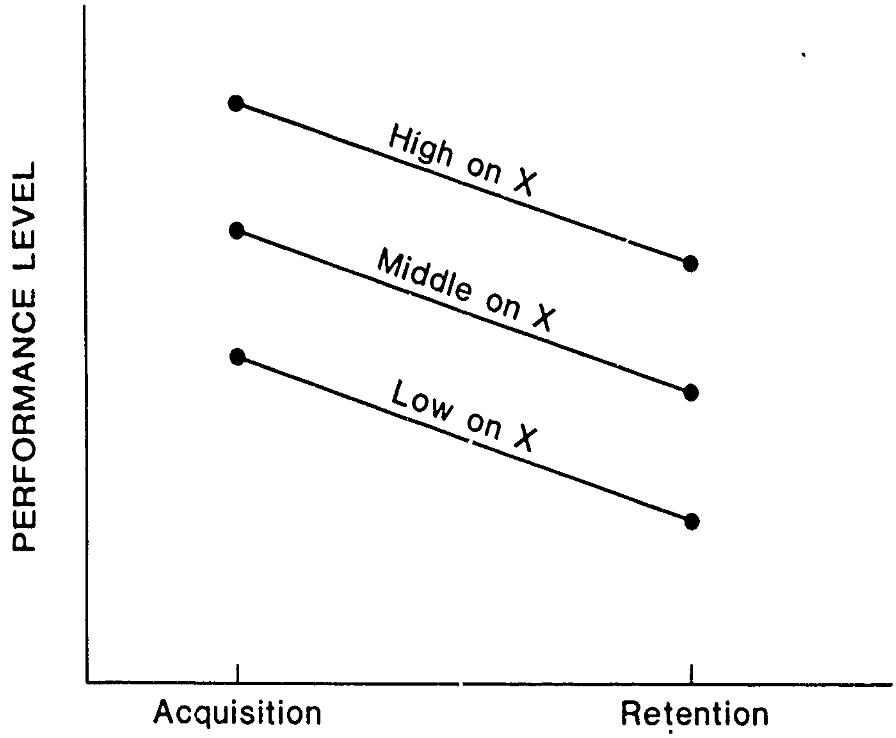
LIST OF FIGURES

NONIMPOSED OVERPRACTICE AND SKILL RETENTION

INTRODUCTION

Performance at the end of acquisition is the most important predictor of performance after a lengthy period of time without practice (Fleishman & Parker, 1962; Naylor & Briggs, 1961; Schendel, Shields, & Katz, 1978). In recent years this well-established proposition has been recast in much stronger form, namely, that performance at the end of acquisition is the <u>only</u> predictor of performance at retention.

Figure 1 presents the strong proposition graphically. "X" is an ability variation operationally independent of performance in either acquisition or retention, for example, score on the Armed Forces Qualification Test (AFQT). High scorers on the AFQT perform better than middle scorers and middle scorers perform better than low scorers in both acquisition and retention; but the differences remain the same. The external ability X has no effect on performance in retention it does not have in acquisition; it contributes nothing to the prediction of performance in retention not mediated by performance in acquisition.



<u>Figure 1</u>. Theoretical Mean Performance Levels at the End of Acquisition and at the Beginning of Reacquisition for Three Groups of Subjects, One with High Scores on an External Ability or Ability-related Variation (X), Another with Middling Scores on X, and the Third with Low Scores on X.

Several studies spanning a variety of task materials and subjects have obtained this result (Carron, 1971; Grimsley, 1967; Holmgren et al., 1979; Purdy & Lockhart, 1962; Mengelkoch, Adams, & Gainer, 1971; Vineberg, 1975). Summarizing these studies, Schendel, Shields, and Katz (1978, p. 12) state that "the weight of the evidence indicates that the rate at which motor proficiency is lost is not related to a performer's initial ability level." Hagman and Rose (1983) come to the same conclusion; so do Slamecka and McElree (1983). The last two investigators varied amount of practice and, therefore, level of performance at the end of acquisition. They found associated differences in performance at retention but no interaction between amount of practice and acquisition-retention; the effects on retention were prefigured by those on acquisition.

If one does not group subjects on X, the same result (no effect on retention not prefigured in acquisition) can be expressed correlationally. The relevant variables are performance at the end of acquisition (A), performance at the start of reacquisition or retention (R), and an "external" ability or ability-related measure (X), that is, any variation over individuals other than A, R, or R-A. If the strong proposition holds, then the partial correlation between X and R controlling for A will vanish. X will share no variance in common with R it does not share with A. An answer for one X does not, of course, prejudice the answer for any other. The strong proposition is a generalization; it holds only so long as no exception to it has been found, that is, no X which makes an independent contribution to R. The principal burden of the present paper is to report such an exception and to elaborate its consequences.

Nonimposed Overpractice

The term "overpractice" is ordinarily used in connection with tasks that have a reachable ceiling. Either the task has a physical limit that many subjects actually achieve or it is scored in such a way that measured performance cannot improve past a certain point. The Army, for example, analyzes the majority of tasks that soldiers perform into component steps, each of which is scored "Go" or "No Go." When a soldier has received a "Go" on all component steps, his performance cannot improve. If practice nevertheless continues past this point, it becomes "overpractice" or, as the Army usually calls it, "mastery training." In the psychological literature overpractice is often called "overlearning."

The effects of overpractice on retention are both long and well established

(Hammerton, 1963: Krueger, 1929; Luh, 1922; Melnick, 1971; Goldberg, Drillings, & Dressel, 1981): overpractice improves retention. At the level of measured performance, this proposition flatly contradicts the idea that performance at the end of acquisition is the only predictor of performance in retention. During overpractice the performance curve is perforce flat; hence, retention improves in the absence of any corresponding improvement in performance at the end of acquisition. The objection may be made, however, that performance in acquisition fails to improve only because it cannot or because no such improvement registers in performance as measured. Slamecka and McElree, for example, exclude from discussion tasks where performance is "artifactually constricted" by floor or ceiling effects.

Consider, then, tasks which do not have a reachable physical limit and are not scored in such a way as to limit measured performance. Performance on most such tasks improves over very long periods of practice, albeit at ever slower rates. Improvement on these tasks is slowed or stopped not by any external limit or by artifactually constricted scoring but, rather, by the subjects' own internal limits. The performance curve looks much as it does in conventional overpractice, except in being not quite flat and in slowing down gradually; but the reasons are different and not describable as artifactual.

The term "nonimposed overpractice" refers to the slopes of individual performance curves late in acquisition. The flatter these curves are and the longer they remain flat or nearly so the more a subject will have practiced the response that is to be retained over the no-practice interval and, hence, the more that response is overpracticed. Overpractice in these cases is not imposed. Nevertheless, the upshot from a learning standpoint is the same. If a subject continues to practice without appreciable improvement, then performance at the end of acquisition has been overpracticed in essentially the same sense as when further improvement has been prevented by external constraint or artifactual scoring. If, in contrast, the performance curve is steep late in acquisition, practice ends with the subject performing at a level he or she has only recently attained and there is little or no nonimposed overpractice.

The hypothesis to be tested is that nonimposed overpractice has the same effects on retention as conventional overpractice.

EXPERIMENT 1

The first experiment was carried out in a small sample (N=27) of Navy enlisted volunteers. Each subject practiced the same six microcomputer-video tasks and in the same order. Video games were used because the alternatives for psychomotor testing are much less attractive in terms of ruggedness, the space they occupy, weight, cost, ease of replacement, maintenance, and availability (Kennedy, Bittner, Harbeson, & Jones, 1982). The six tasks selected for study were chosen on the basis of stabilization with practice and task definition (Jones, 1972). Other video-computer tasks also studied in the volunteer population either did not stabilize with practice or did so with poor task definition (Jones, 1981). Reacquisition was begun on the six tasks after intervals of no practice ranging from 4 to 18 months.

Method

Subjects. The 27 volunteers were all males between 19 and 24 years of age and with 20/20 corrected vision. The sailors were paid for their participation in accordance with Navy guidelines.

Procedure. Each task was practiced one session a day for 15 consecutive working days. The six tasks were practiced in the following order: Air Combat Maneuvering (ACM), Breakout, Race Car, Pong, Basketball, and Anti-Aircraft, the last two concurrently. Table 1 presents details on each task, including trial length, number of trials per day, and dependent measure (score). Race Car, for example, is game #2 in the Atari Indy 500 cartridge (CX-2511). In playing any of these games the player has the option of setting a difficulty switch on the game console. "A" is hard and "B" easy. In the case of Race Car difficulty level controls the maximum speed at which the car travels. When the switch is set at "A", the car travels at higher speeds making it more difficult to control. The purpose of the game is to complete as many laps as possible in the 60 seconds allowed; whenever the car crashes into a boundary or barrier, time is lost. The control devices are a special knob which functions as a steering wheel and a button which functions as an accelerator. Each trial lasted 60 seconds and each sailor received 15 trials a day. Intertrial interval was not strictly controlled but never lasted more than one minute and usually much less. The dependent measure was number of laps completed.

Practice was resumed on Race Car and Pong after an interval of 4-6 months, on Basketball and Anti-Aircraft after 10-12 months, and on ACM and Breakout after 16-18 months. All retention intervals were measured from the 15th day of acquisition on that particular task to the first day of reacquisition. Each task was repracticed for five consecutive working days, with the same number of trials per day as was used in acquisition. The two tasks in the same retention

TABLE 1

Information Regarding the Six Video-Computer Tasks in Experiment 1

NAME		CASSETTE, GAME	DIFFI-CULTY	TASK DESCRIPTION	CONTROL DEVICE	TRIAL LENGTH	TRIALS/DAY	SCORE
I)	Race Car	Indy 500 (CX-2511), Game 2	Left=A* Right=A	Car must drive around course & attempt to avoid hitting wall (hitting wall slows car down)	special knob, button	60 sec	15 trials/day	# if laps completed around the course
II)	Pong	Video Olympics (CX-2621), Game 2	Left=B Right=A	Ball is rallied back and forth from computer to S	paddle, button	1st to score 2	3 trials/day 15 days	player's percentage of total points
III)	Basketball	Basketball (CX-2624), Game 2	Left=A Right=A	Player must score baskets while pre-venting a computer controlled cursor from scoring	joystick, button	4 min	4 trials/day	# of baskets x 2
IV)	Anti-Aircraft	Air Sea Battle	Left=A Right=A	Shoot down over-passing jets from a stationary platform	joystick, button	2 min 16 sec	7 trials/day	# of planes shot down
V)	Air Combat Maneuvering	Combat (CX-2601), Game 24	Left=B Right=B	Attack and shoot down opposing jet from one's own jet	joystick, button	2 min 16 sec	10 trials/day	# of times opposing jet is hit
VI)	Breakout	Breakout (CX-2622), Game 1	Left=B Right=B	Ball ricochets from brickwall and is returned by paddle	paddle, button	5 balls	10 trials/day	# of bricks knocked down (each "wall" of bricks is worth dif-ferent # points)

*The "Left" difficulty control applies to the player and the "Right" control to the computer or second player. In many one-player games the "Right" control is irrelevant.

interval, for example, Race Car and Pong, were repracticed concurrently. That is, both tasks were practiced on the same five consecutive working days. All subjects were instructed not to practice any of the six tasks during the no-practice interval.

The number of subjects who completed the experiment varied from retention interval to retention interval. Seventeen sailors completed work on Race Car and Pong, 16 on Basketball and Anti-Aircraft, and 13 on ACM and Breakout. The subjects who completed the longer intervals were not all nested among those who completed the shorter ones. Three sailors, for example, were already more than a year past acquisition on Basketball and Anti-Aircraft when the decision to carry out a retention study was made. These three sailors could be repracticed only on ACM and Breakout. Other sailors were transferred after repracticing two or four but not all six tasks.

In all tasks the dependent measure appeared on the viewing screen throughout play; it was the quantity the subject was trying to maximize. A player's score was the value of the dependent measure when the game ended. A subject's score on any given day (session) was the mean score of the games he played that day. Thus for each subject and task, analysis begins with 20 data points--15 in acquisition and 5 in reacquisition.

Results and Discussion

Average performance. Figure 2 presents means and standard deviations in both acquisition and reacquisition for the six tasks. The average-performance curves for Race Car, Pong and Basketball decelerate more strongly than those for Anti-Aircraft, ACM, and Breakout. The standard-deviation curves are essentially flat after the first few sessions, except possibly for ACM and Breakout where there appears to be a tendency for variability to increase slightly with the mean.

Since the units for the six tasks are different, comparisons among them must be made in terms of unit-free ratios. Table 2 contains two such ratios. The first is the gain for the second half (days 8 through 15) relative to that for the first half (days 1 through 8) of acquisition. For example, mean performance on Race Car increased from 8.96 laps on day 1 to 11.95 laps on day 8 to 12.36 laps on day 15. Hence, the gains during the first and second halves of acquisition were 2.99 and 0.41 laps respectively. The ratio of the latter to the former is 0.14; that is, average performance increased roughly one seventh as

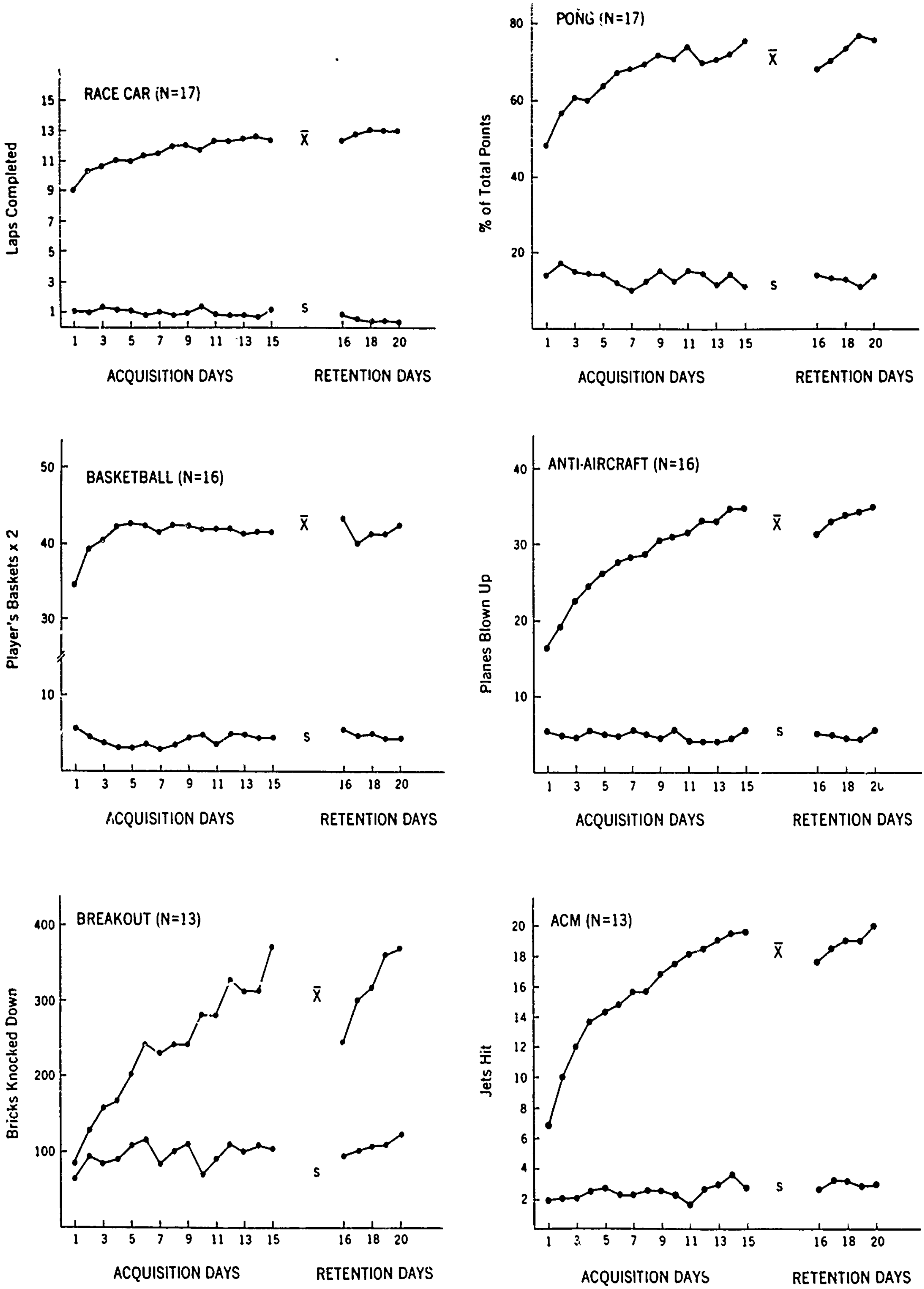


Figure 2. Means and Standard Deviations in Acquisition and Retention for Race Car, Pong, Basketball, Anti-Aircraft, ACM, and Breakout in Experiment 1.

much in the second as in the first half of acquisition. This first ratio is an inverse measure of nonimposed overpractice.

The second ratio is the change from the last day of acquisition to the first day of reacquisition, relative to the gain in the first half of acquisition. Mean performance on Race Car on the first day of reacquisition was 12.24, only 0.12 less than mean performance on the last day of acquisition. Hence, the second ratio for Race Car is 0.12 divided by 2.99 or 0.04. This second ratio is an inverse measure of relative retention. It makes use of the same baseline, that is, the increase from day 1 to day 8, as the first ratio.

The tasks in Table 2 are ordered by increasing slope in the second half of acquisition. As the table makes clear, the fall-off over the no-practice interval increases in almost perfect correspondence. The correlation between the two ratios is 0.90, and even with 4 degrees of freedom this value is significant at the .02 level, $t(14)=4.13$. On the group level, therefore, the positive relationship between nonimposed overpractice and relative retention appears to hold. The steeper the slope of the mean performance curve late in acquisition the worse retained the skill tends to be.

TABLE 2

Two Ratios, One Indicating Nonimposed Overpractice and the Other Relative Retention, for the Six Video-Computer Tasks in Experiment 1.

Task	Retest Interval (months)	$\frac{\bar{X}(15)-\bar{X}(8)}{\bar{X}(8)-\bar{X}(1)}$	$\frac{\bar{X}(15)-\bar{X}(16)}{\bar{X}(8)-\bar{X}(1)}$
Basketball	11	-0.09	-0.21
Race Car	5	0.14	0.04
Pong	5	0.27	0.34
ACM	17	0.43	0.21
Anti-Aircraft	11	0.74	0.41
Breakout	17	0.75	0.77

Since the control devices used in the six games were not systematically varied independently of other possible determinants of average retention, definite conclusions regarding these devices are not possible. Nevertheless, they do not seem to make a large difference. Although it has different functions, the "special knob" used in Race Car is the same device as the "paddle" used in Breakout and Pong; yet these three tasks run the gamut from good to relatively poor relative retention. Similarly, ACM, Basketball, and Anti-Aircraft all use a joystick, again with a wide range of differences in how well these tasks are retained on the average over a no-practice interval.

The length of the no-practice interval, on the other hand, does appear to be a factor in retention. In the last column of Table 2 the larger values are weakly associated with the longer retention intervals. Unfortunately, this association is badly confounded with overpractice. Two of the three most overpracticed tasks were assigned to the shortest retention interval, and two of the three least overpracticed tasks were assigned to the longest retention interval. On the basis of the present evidence, therefore, the effect of how long the no-practice interval is on skill retention is moot.

Individual performance. In Experiment 1 practice continued long enough that a subject's performance on day 15, the last in acquisition, was not always or even usually his best; on all tasks most subjects performed at least as well on an earlier day as they did on day 15 and, therefore, had at least one day of overpractice. If one takes a subject's performance on day 15 as his "criterion level," then most subjects reached criterion before day 15 and, therefore, overpracticed at that level for at least one day (session). Accordingly, the measure of individual overpractice used in this first experiment was "15 minus the number of the session on which the subject first reached or exceeded his score on session 15." A subject who performed better on day 15 than on any earlier day received a score of 0 on nonimposed overpractice. A subject who first reached or exceeded his final level of performance in acquisition on day 10 received a score of 5 on nonimposed overpractice. On all tasks amount of overpractice varied widely. On Race Car, for example, nonimposed overpractice ranged from 0 to 13 with a mean of 3.7 and a standard deviation of 4.5. ACM was least overpracticed but even it had scores ranging from 0 to 9 with a mean of 1.7 and a standard deviation of 2.5.

A subject's level of performance at the end of acquisition (A) was his score on day 15. Similarly, his level of performance on the first day of reacquisition (R) was his score on session 16. Retention was measured both absolutely (R) and relatively (R-A). The first-order correlations between an external ability or ability-related measure X and these two measures may, of course, be different; but the partial correlations between X and absolute and relative retention controlling for A are always the same because the partial correlation between absolute and relative retention controlling for A is unity.

Two external ability measures were considered: performance on day 1 and nonimposed overpractice, that is, number of additional sessions at or beyond a subject's level of performance on day 15 as discussed above. The first column of Table 3 contains the partial correlations between nonimposed overpractice and absolute or relative retention, controlling for A. All six partial correlations are positive and three of them significantly so at or beyond the .05 level. The second and third columns in Table 3 contain the partial correlations between level (A) and absolute and relative retention, holding overpractice constant. The correlations between A and relative retention (R-A) are generally negative, as is to be expected since A appears with reversed sign in (R-A). The partial correlations for absolute retention are more interesting. All six are positive, and five of them are significant at the .05 level or beyond. On the whole, these correlations are somewhat but not much larger than the corresponding correlations for overpractice. In this data set nonimposed overpractice is almost as large a factor in retention as performance at the end of acquisition.

Performance on day 1 was not a factor. The six partial correlations between performance on day 1 and absolute or relative retention, controlling for A, were variable as to sign and not significant. The largest of these six correlations, the one for Pong, fell short of significance at the .05 level, $t(14)=1.95$, $p=.08$.

EXPERIMENT 2

Experiment 1 involved a single sample of subjects with different tasks repracticed after different retention intervals. The one sample was small, each task was extensively practiced in acquisition, and retention intervals were long, ranging up to 18 months; furthermore, no information relative to possibly contributory variations external to the practiced tasks themselves was collected. Experiment 2 differed in all these respects. The new design called for three samples of approximately 50 subjects each, with each sample to be practiced on one task only, the task varying from sample to sample, and repracticed after a single relatively short interval of time, four months. The amount of practice in acquisition was reduced, and 13 reference tests were administered to all subjects as possibly predictive external variations.

The reduction in amount of practice implies, of course, a reduction in amount of overpractice also. Hence, the relationships between overpractice and retention ought to be smaller in Experiment 2 than in Experiment 1. At the same time, the power of the design has been increased with the increase in sample size and the provision of separate, independent tests for the three tasks.

TABLE 3

Partial Correlations in Experiment 1 between Overpractice and Level as Predictors and Absolute and Relative Retention as Criteria, Controlling for the Other Predictor.

Task	Overpractice	Level	
	Absolute/Relative	Absolute	Relative
Race Car	+.473	+.603[1]	-.226
Pong	+.608[1]	+.612[1]	-.098
Basketball	+.238	+.325	-.245
Anti-Aircraft	+.767[3]	+.872[3]	+.091
ACM	+.742[2]	+.870[3]	-.200
Breakout	+.263	+.700[2]	-.443

[1] Significant of the .05 level.

[2] Significant of the .01 level.

[3] Significant of the .001 level.

Method

Subjects. The subjects were all college students at three Central Pennsylvania colleges: Elizabethtown College, Lebanon Valley College, and the Capitol Campus of Pennsylvania State University. All subjects were between 18 and 28 years old with 20/20 corrected vision. The numbers of subjects at the three campuses were:

Elizabethtown, 33 females and 17 males

Lebanon Valley, 27 females and 29 males

Capitol Campus, 19 females and 34 males.

Each student was paid $50 for his or her participation in the study.

Procedure. All subjects in all three samples were first administered the following sequence of 13 paper-and-pencil tests, the first 12 from the ETS Kit of Factor-Referenced Cognitive Tests (Ekstrom, French, & Harman, 1976): Hidden Patterns (CF-2), Copying (CF-3), Gestalt Completion (CS-1), Addition and Subtraction (N-4), Finding A's (P-1), Identical Pictures (P-3), Nonsense Syllogisms (RL-1), Card Rotations (S-1), Cube Comparisons (S-2), Map Planning (SS-3), Advanced Vocabulary (V-5), Paper Folding (VZ-2), and Reversed Printing. These particular tests were chosen because they seemed most likely on the basis of content to relate to video-game performance. The 12 factor-referenced tests were all administered as prescribed in the ETS Manual (Ekstrom, French, & Harman, 1976). In Reversed Printing the subjects were asked "to print the letters of the alphabet as fast as you can, in sequence, but upside down and backwards." The test lasted three minutes. None of the 13 reference tests was practiced.

Each subject was given five practice sessions on one of the three video-computer tasks: ACM at Elizabethtown, Race Car at Lebanon Valley, and Anti-Aircraft at Capitol Campus. In the cases of ACM and Anti-Aircraft each session consisted of seven 2-minute-and-16-second games, a total of 15 minutes and 52 seconds of playing time. Each session of Race Car consisted of 16 1-minute games. In all three cases the five practice sessions were completed within a 10-day period, with no more than two sessions taking place on a given day.

Reacquisition was begun approximately four months after acquisition was completed. Each subject was given three practice sessions with the same number of games per session and the same conditions as to distribution as in acquisition.

Analysis. Nonimposed overpractice in Experiment 2 could not be measured in the same way as in Experiment 1 because few subjects in any of the three samples performed as well in an earlier session as they did in session 5. With less practice and performance curves that were still rising sharply when acquisition ended, nonimposed overpractice could be indexed only by the slopes of individual performance curves "late" in practice. Two major possibilities presented themselves in this connection. One was to index overpractice by the simple algebraic difference between a subject's performance in sessions 3 and 5, that is, $(X_5 - X_3)$. The other was to fit a regression line to X_3, X_4, and X_5 and take the slope of this regression line as the index of overpractice. Different as these two approaches sound, they come to exactly the same thing. With equally spaced points on the abscissa and N=3, the slope of the regression line

$$b_{y.x} = (X_5 - X_3)/2.$$

Overpractice in Experiment 2 is indexed inversely, by $(X_5 - X_3)$; but this index may be interpreted, if one wishes, as the slope of the regression line over the last half of acquisition, that is, sessions 3 through 5.

Results and Discussion

Average performance. The mean-performance curves for the three tasks followed much the same courses in acquisition as are depicted for the same tasks in Figure 2, provided that the latter three curves are truncated after day 5. The standard-deviation curves are also much the same in acquisition, with the same proviso. In retention, however, there are some differences. In Experiment 1 relative retention $(\overline{X}_{16} - \overline{X}_{15})$ for Race Car equalled -0.12 laps completed. In Experiment 2 relative retention was on the same order, -0.18 laps. For Anti-Aircraft and ACM, however, relative retention was much better than it was in Experiment 1. Anti-Aircraft dropped 3.28 planes downed from session 15 to session 16 in Experiment 1 but only 1.07 planes downed from session 5 to session 6 in Experiment 2 ($t(67)=2.49$; $p<.02$). Similarly, ACM dropped 1.88 jets hit over the no-practice interval in Experiment 1 but increased by 0.04 jets hit in Experiment 2 ($t(61)=3.54$; $p<.001$).

Experiments 1 and 2 differed in several respects: amount of practice, the inclusion of women in the second experiment, and the subjects' educational background. These differences, however, applied equally to all three tasks whereas length of the retention interval was essentially the same in the two experiments for Race Car but much shorter for Anti-Aircraft and ACM in Experiment 2. The results are still not conclusive but retention of these particular video-computer tasks and, perhaps, others like them would seem to degrade somewhat with increasing length of the no-practice interval.

Figure 3 presents the mean-performance curves for Race Car broken down by sex. The males do better than the females throughout both acquisition and reacquisition. This result is not peculiar to Race Car but is characteristic of almost all video-computer games (Jones, 1984). The two curves, moreover, are not parallel either late in acquisition or over the no-practice interval. The males improve less than the females from session 3 to session 5 and retain what they have learned better over the following no-practice interval. Both differences are statistically significant.

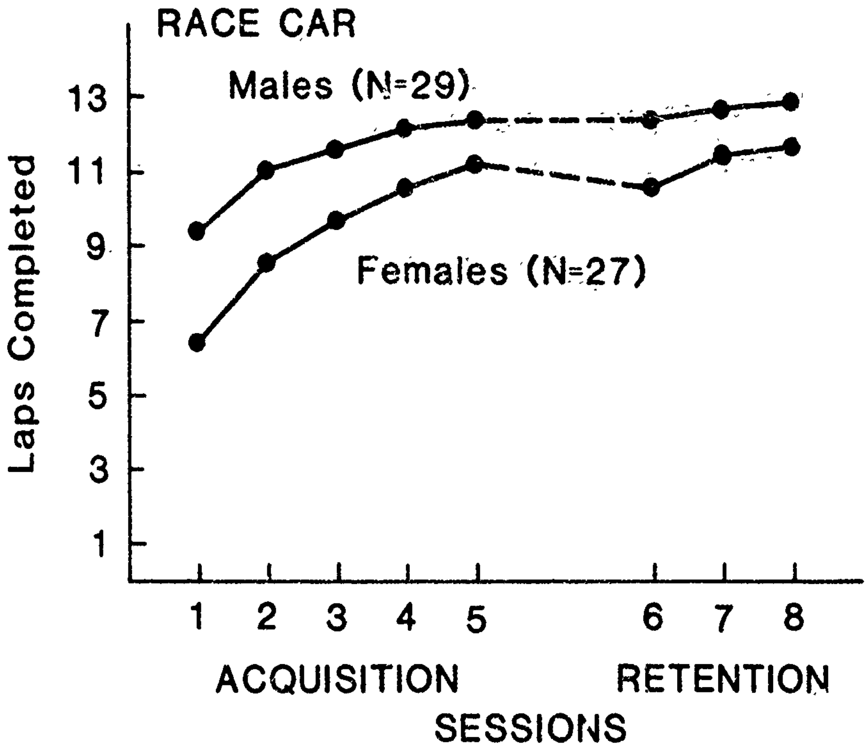


Figure 3. Mean Performance Levels for Males and Females in Acquisition and Retention on Race Car in Experiment 2.

The mean changes are presented in Table 4 along with the corresponding results for Anti-Aircraft and ACM. The difference between the two slopes for Race Car, 0.66 and 1.40, yields a value of $t(54)=5.87$, $p < .001$. Similarly, the two figures for relative retention, 0.02 and 0.36, yield a value of $t(54)=3.19$, $p < .001$.

Males perform better than females on Anti-Aircraft and ACM also. In both cases the males improve more over the last three sessions of acquisition and deteriorate more over the no-practice interval. The differences, moreover, for Anti-Aircraft are significant. For the difference in slope $(\bar{X}_5 - \bar{X}_3)$ $t(51)=2.42$, $p < .02$, and for the difference in relative retention, actually its inverse $(\bar{X}_5 - \bar{X}_6)$, $t(51)=2.30$, $p < .03$.

TABLE 4

Average Increase from Day 3 to Day 5 and Average Decrease from Day 5 to Day 6 for the Three Tasks in Experiment 2, by Sex.

Tasks	Sex	$(\bar{X}_5-\bar{X}_3)$	$(\bar{X}_5-\bar{X}_6)$
Race Car	Male	0.66	0.02
	Female	1.40	0.36
Anti-Aircraft	Male	4.23	1.65
	Female	2.07	0.04
ACM	Male	2.33	0.13
	Female	1.80	-0.13

These results directly contradict the idea that performance at the end of acquisition is the only predictor of performance at retention. First, the curves for the two sexes over the no-practice interval are not parallel and, second, the steeper of the two curves is regularly associated with a steeper slope late in acquisition. By the same token, of course, these results confirm the role of nonimposed overpractice in skill retention. In these cases, moreover, overpractice is not a matter of how long a subject practices at his or her final level of performance in acquisition but purely a matter of slope. The closer a subject is to his or her final level of performance in sessions before the last the better retention tends to be.

Individual performance. The first column of Table 5 contains the partial correlations between slope $(X_5 - X_3)$ and absolute (X_6) or relative $(X_6 - X_5)$ retention, controlling for level (X_5). As already noted, the partial correlations for absolute and relative retention are necessarily the same. The negative correlations with slope reflect positive associations between overpractice and retention. As was expected, the magnitudes of these associations are smaller than in Experiment 1. Nevertheless, one of the three correlations, the one for Race Car, is significant. The second and third columns are also consistent with the results from Experiment 1. Level of performance at the end of acquisition is strongly and positively associated with absolute retention and negatively associated with relative retention.

Several of the first-order correlations between the 13 reference tests and absolute or relative retention were significant at the .05 level; but none of these relationships remained significant when slope and level were statistically controlled. This result does not mean that no external ability or ability-related variation relates to retention in ways not mediated by level of

TABLE 5

Partial Correlations in Experiment 2 between Slope and Level in Acquisition as Predictors and Absolute and Relative Retention, Controlling for the Other Predictor.

Task	Slope	Level	
	Absolute/Relative	Absolute	Relative
Race Car	-.626[2]	+.945[2]	-.268[1]
Anti-Aircraft	-.226	+.821[2]	-.509[2]
ACM	-.266	+.888[2]	-.056

[1] Significant at the .05 level.

[2] Significant at the .001 level.

performance at the end of acquisition or overpractice. It does mean, however, that no such variation has yet been implicated, let alone established.

Experiment 2 strongly confirms the role of nonimposed overpractice in skill retention. Practice need not be continued long enough that most subjects accumulate additional practice time at their final levels of performance in acquisition. Even with relatively little practice, the role of overpractice can still be demonstrated. The sex differences in slope and relative retention were not large; yet even these small differences were consistent with a positive effect of overpractice on retention. Similarly, the partial correlations for overpractice (slope), though reduced in magnitude, were positive and, in one instance, statistically significant.

EXPERIMENT 3

Experiment 3 was designed partly to replicate results obtained in Experiment

2 and partly to test hypotheses suggested by the first two experiments. Race Car was replicated exactly as in Experiment 2, using the same procedures at the same college except that no reference tests were administered (N=56). Students at the other two campuses were all administered ACM, again with no reference tests. Otherwise the only difference with Experiment 2 was that all subjects at both campuses were given nine practice sessions rather than five and two weeks to complete them rather than ten days. At Elizabethtown 15 of the subjects were male and 32 female; at Capitol Campus the figures were almost exactly the reverse, 32 males and 16 females.

ACM was the only task in Experiment 2 not to show significant sex differences in slope and relative retention. Nor were the partial correlations for slope with absolute and relative retention significant. In both cases the results were consistent with an overpractice effect but not strong enough to be significant. With more practice and, therefore, more overpractice, the results ought to become stronger.

Results and Discussion

Race Car. As in Experiment 2, males outperformed females in both acquisition and reacquisition. Also as in Experiment 2, the males improved more from session 3 to session 5 and deteriorated less over the no-practice interval than the females. This time, however, neither sex difference was significant. For the difference in slopes $t(54)=1.62$ and for the difference in relative retention $t(54)=1.89$.

The partial correlation between overpractice (slope with the sign changed) and absolute or relative retention, controlling for performance in session 5, was 0.588, not much different from the corresponding value of 0.626 in Experiment 2 and, like it, significant at the .001 level.

Except for the failure of the sex differences to reach significance, these results replicate all of the key findings regarding Race Car in Experiment 2.

ACM. Combining the samples from Elizabethtown and Capitol Campus yields a total of 95 subjects, 47 males and 48 females. The gains for the two sexes from session 7 to session 9 were 0.86 and 1.00 jets hit for males and females respectively. Similarly, relative retention ($\bar{X}_{10} - \bar{X}_9$) for the two sexes was -1.04 and -0.92 respectively. These two differences are not in accord with an overpractice effect. The males improved slightly less in the last three sessions of acquisition but fell back slightly more over the no-practice interval. The larger fact, however, is that the curves for slope late in acquisition and for retention are both almost parallel. The differences between the two sexes are very small on both counts. The difference in slope, for example, is 0.09 standard-score units; that is,

$$\frac{\bar{X}\text{ female} - \bar{X}\text{ male}}{\sigma} = 0.09,$$

where σ is estimated from the pooled within-sex variances. The difference in relative retention is 0.07 standard-score units. The values of t for the two differences, 0.45 and 0.36, are both small. In brief, with more practice on ACM in acquisition overpractice is almost the same for the two sexes and, as a consequence, relative retention for the two sexes is also almost the same.

The partial correlation between slope ($X_9 - X_7$) and absolute or relative retention, controlling for performance at the end of acquisition (X_9), was -0.458, $t(92)=6.03$, $p<.001$. In Experiment 2, with only five practice sessions, the same correlation was -0.266 and not significant. By itself, this result confirms the role of nonimposed overpractice in facilitating retention. In combination with Experiment 2, it also indicates that up to a point increasing amounts of practice and, hence, overpractice enlarge that role.

With very little practice, less than the amounts used in any of the three experiments reported here, it is probable that no subject is overpracticed and, therefore, that overpractice is a negligible factor in retention. As amount of practice increases, almost all subjects become overpracticed in varying degrees and the role of overpractice in retention increases. If practice were continued long enough, more than any amount used in these experiments, it is possible that all subjects would become heavily overpracticed and that overpractice might again become a negligible factor in how well different individuals retained whatever skill levels they had attained at the end of acquisition. In such a case, mean performance levels would, of course, be strongly sustained by overpractice; but overpractice would no longer differentiate among individuals because overpractice would not itself vary importantly from one individual to the next.

Higher-order partial correlations. Except for sign, the partial correlation between slope ($X_9 - X_7$) and absolute or relative retention, controlling for X_9, necessarily equals the partial correlation between X_7 and absolute or relative retention, controlling for X_9 --the reason being that, if X_9 is controlled, the correlation between X_7 and ($X_9 - X_7$) equals -1. In other words, the results so far obtained for "slope" might equally well have been reported as partial correlations for X_7 alone. If slope makes a significant contribution to retention, then so does X_7 (or X_3 in Experiment 2) and conversely.

These observations suggest another way of looking at the results obtained in these experiments. If performance at the end of acquisition is the only predictor of performance at retention, then skill acquisition-retention is a Markov process; that is, stages of acquisition prior to the last do not affect the probabilities that a subject will perform well, poorly, or in between at retention. A showing, therefore, that performance at retention varies with <u>any</u> session before the last where performance in the last session is controlled, contradicts the idea that acquisition-retention is a Markov process.

Such a showing has, of course, already been made, since slope or X_7 has been shown to relate to retention in ways not mediated by X_9. Perhaps, however, X_7 would no longer contribute independently to retention if X_8 as well as X_9 were controlled. If so, the dependence of retention on practice sessions preceding the last would not go any deeper than X_8. It is also possible that the dependence of retention on earlier sessions in acquisition than the last might reach back further than X_7. To test these various possibilities one calculates first the partial correlation between X_8 and X_{10} controlling for X_9 ($r_{8,10.9}$); then the partial correlation between X_7 and X_{10} controlling for X_8 and X_9 ($r_{7,10.8,9}$); then the partial correlation between X_6 and X_{10} controlling for X_7, X_8, and X_9 ($r_{6,10.7,8,9}$); and so on. Making these calculations for ACM in Experiment 3, one obtains the following results:

$r_{9,10}$ = +0.886, $t(93)$=18.4, $p < .001$;
$r_{8,10.9}$ = +0.366, $t(92)$= 3.7, $p < .001$;
$r_{7,10.8,9}$ = +0.317, $t(91)$= 3.2, $p < .002$;
$r_{6,10.7,8,9}$ = +0.357, $t(90)$= 3.6, $p < .001$;
$r_{5,10.6,7,8,9}$ = -0.045, $t(89)$= 0.4, $p = .680$.

For ACM in Experiment 3 retention depends on performance as far back as X_6, but not X_5. Note, too, that the signs of the significant partial correlations are all positive. The better a subject performs in any of these sessions the better his or her performance at retention is likely to be, even after one controls for all remaining sessions in acquisition. The better, however, performance in any session preceding the last is, while holding performance in the last session constant, the flatter the slope of the performance curve late in acquisition will be. In short, these calculations allow us to give the phrase "late in acquisition" a more precise meaning. In the case at hand, it means sessions X_6 through X_9.

GENERAL DISCUSSION

Where practice is continued long enough that many subjects reach their final level of performance sometime before acquisition ends, as in Experiment 1, the explanation of the overpractice effect would seem to be straightforward. Nonimposed overpractice improves retention because the more there is of it the more subjects have practiced the response to be retained during acquisition. When practice is less extended and overpractice has to be indexed exclusively by the slopes of individual performance curves, as in Experiments 2 and 3, the explanation has to be more complicated. One might reason, for example, that the closer to their final levels of performance subjects are at earlier stages of acquisition, the more what they learn in those earlier stages will generalize to the response that will ultimately have to be retained over the no-practice interval. Where overpractice can be measured only by slope, one must resort to generalization, common elements, or some similar idea.

In fact, the explanation may be even more complicated. Under an interference theory of forgetting performance levels do not degrade over a period of no practice simply because time is passing but, rather, because other learned responses interfere with the response to be retained. The more isolated the response to be retained is or the more isolated it can be made by discrimination learning, the less interfered with and the better retained it will be. From this point of view, the overpractice effect is mediated in part or, perhaps, mainly by discrimination learning (Siegel, 1969). The longer a subject performs at his or her final level or the more nearly he or she performs at that level, the greater the likelihood of the subject's discriminating between appropriate responding in the task at hand and appropriate responding in similar but different tasks. As a consequence, the response at hand is better protected against interference effects over the no-practice interval and, therefore, better retained.

To the extent that this latter explanation is true one might expect that distributing practice late in acquisition would facilitate retention, since doing so would increase the likelihood that interfering responses would occur. If practice is massed late in acquisition, interference is, so to speak, deferred until the no-practice interval where it will degrade performance. If practice is distributed late in acquisition, the likelihood is greater that interfering responses and discrimination learning will both take place during the acquisition process.

This latter explanation is easily tested. To do so, a single group of subjects would be practiced on a task, say, ACM, until the mean performance curve had strongly decelerated. The group would then be split at random into two halves. Both halves would practice the task for an additional *n* sessions, one half under massed and the other under distributed conditions. If discrimination learning mediates the overpractice effect, one would expect better retention in the group with distributed practice after the split.

CONCLUSIONS AND MILITARY APPLICATIONS

Considerable periods of time usually elapse between the end of training and the operational use of acquired skills. As a result, soldiers who are proficient at the end of training are not necessarily proficient when first called upon to use what they have learned. One remedy to this loss of proficiency is selective refresher training. Those soldiers who are predicted not to be proficient after a given interval of time are selectively given refresher training. If prediction is poor, selective refresher training is ineffective; soldiers who do not need it are trained and soldiers who do need it are not. In proportion, however, as future individual performance can be accurately predicted, selective refresher training can be given to just those soldiers who need it and not to any others. The results obtained in this research allow improved prediction of individual performance in any continuous performance task, provided that suitable performance measures are available; it is only necessary to measure nonimposed overpractice along with level of performance in predicting future performance. Improved prediction, however, makes the same expenditure in selective refresher training more effective.

REFERENCES

Bilodeau, E.A., Jones, M.B. & Levy, C.M. (1964). Long-term memory as a function of retention time and repeated recalling. Journal of Experimental 67, 303-309.

Carron, A.V. (1971). Effect of ability upon retention of a balance skill after two years. Perceptual and Motor Skills, 33, 527-529.

Corballis, M.C. (1965). Practice and the simplex. Psychological Review, 72, 399-406.

Ekstrom, R.F., French, J.W., & Harman, H.W. (1976). Manual for Kit of Factor-Referenced Cognitive Tests, 1976. Princeton, NJ: Educational Testing Service.

Fleishman, E.A., & Parker, J.R. Jr. (1962). Factors in the retention and relearning of perceptual motor skill. Journal of Experimental Psychology, 64, 205-226.

Goldberg, S.L., Drillings, M. & Dressel, J.D. Mastery training: Effects on skill retention. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences, Technical Report 513.

Grimsley, D.L. (1969). Acquisition, retention, and retraining: Effects of high and low fidelity in training devices. Alexandria, VA: Human Resources Research Organization, Technical Report 69-1.

Hagman, J.D. & Rose, A.M. (1983). Retention of military tasks: A review. Human Factors, 25, 199-213.

Hammerton, M. (1963). Retention of learning in a difficult tracking task. Journal of Experimental Psychology, 66, 108-110.

Holmgren, J.E., Hilligoss, R.E., Sweezey, R.W., & Eakins, R.C. (1979). Training effectiveness and retention of training extension course (TEC) instruction in the combat arms. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences, Research Report 1208.

Jones, M.B. (1972). Individual differences. In R.N. Singer (Ed.), The psychomotor domain. Philadelphia: Lea and Febiger.

Jones, M.B. (1981). Further studies on stabilization with practice in performance tests. New Orleans, LA: U.S. Naval Biodynamics Laboratory, Final Report on Contract No. N00203-80-M-5387.

Jones, M.B. (1984). Video games as psychological tests. Simulation & Games, 15, 131-157.

Kennedy, R.S., Bittner, A.C., Jr. Harbeson, M., & Jones, M.B. (1982). Television-computer games: A 'new look' in performance testing. Aviation, Space, and environmental Medicine, 53, 49-53.

Krueger, W.C.F. (1929). The effects of overlearning on retention. Journal of Experimental Psychology, 12, 71-78.

Luh, C.W. (1922). The conditions of retention. Psychological Monographs, 31 (Whole No. 142).

Melnick, M.J. (1971). Effects of overlearning on the retention of a gross motor skill. Research Quarterly, 42, 60-69.

Mengelkoch, R.F., Adams, J.A., & Gainer, C.A. (1960). The forgetting of instrument flying skills as a function of the level of initial proficiency. Port Washington, MY: U.S. Naval Training Device Center, Technical Report 71-16-18.

Naylor, J.C., & Briggs, G.E. (1961). Long term retention of learned skills: A review of the literature. Canton, OH: Aerospace Medical Laboratory, Wright-Patterson Air Force Base, Technical Report 61-390.

Purdy, B.J., & Lockhart, A. (1962). Retention and relearning of gross motor skills after long periods of no practice. Research Quarterly, 33, 265-272.

Schendel, J.D., Shields, J.L., & Katz, M.S. (1978). Retention of motor skills: Review. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences, Technical Paper 313.

Siegel, S. (1969). Discrimination overtraining and shift behavior. In N.S. Sutherland & R. Gilbert (Eds.) Animal discrimination learning. London: Academic Press.

Slamecka, N.J., & McElree, B. (1983). Normal forgetting of verbal lists as a function of their degree of learning. Journal of Experimental Psychology: Learning, Memory and Cognition, 9, 384-397.

Vineberg, R. (1975). A study of the retention of skills and knowledge acquired in basic training. Alexandria, VA: Human Resources Research Organization, Technical Report 75-10.