ARI TECHNICAL REPORT
TR-76-A4

The Acquisition and Retention of Visual Aircraft Recognition Skills

by
Robert D. Baldwin, Robert E. Cliborn, and Robert J. Foskett

HUMAN RESOURCES RESEARCH ORGANIZATION
300 North Washington Street
Alexandria, Virginia 22314

NOVEMBER 1976

Contract DAHC-19-75-C-0020

Prepared for

U.S. ARMY RESEARCH INSTITUTE
for the BEHAVIORAL and SOCIAL SCIENCES
1300 Wilson Boulevard
Arlington, Virginia 22209

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

	W. C. MAUS
J. E. UHLANER	COL, GS
Technical Director	Commander

Research accomplished
under contract to the Department of the Army

Human Resources Research Organization
Western Division, Fort Bliss Office

NOTICES

DISTRIBUTION: Primary distribution of this report has been made by ARI. Please address correspondence concerning distribution of reports to: U. S. Army Research Institute for the Behavioral and Social Sciences, ATTN: PERI-P, 1300 Wilson Boulevard, Arlington, Virginia 22209.

FINAL DISPOSITION This report may be destroyed when it is no longer needed. Please do not return it to the U. S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER Technical Report 76-A4	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) THE ACQUISITION AND RETENTION OF VISUAL AIRCRAFT RECOGNITION SKILLS		5. TYPE OF REPORT & PERIOD COVERED Final Report 3 Mar 75 — 3 Oct 76
		6. PERFORMING ORG. REPORT NUMBER FR-WD(TX)-76-10
7. AUTHOR(s) Robert D. Baldwin, Robert E. Cliborn, Robert J. Foskett		8. CONTRACT OR GRANT NUMBER(s) DAHC 19-75-C-0020 HumRRO No. S75-14
9. PERFORMING ORGANIZATION NAME AND ADDRESS Human Resources Research Organization 300 North Washington Street Alexandria, Virginia 22314		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 6.27.17A 2Q762717A721
11. CONTROLLING OFFICE NAME AND ADDRESS US Army Research Institute for the Behavioral and Social Sciences 1300 Wilson Boulevard, Arlington, VA 22209		12. REPORT DATE November 1976
		13. NUMBER OF PAGES 61
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

This report is based on research conducted by Human Resources Research Organization, WESTERN Division, Fort Bliss Office for the US Army Research Institute for the Behavioral and Social Sciences under Contract DAHC-19-75-C-0020; HumRRO No. S75-14, Project ALERT.

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

Aircraft Recognition	Self-Paced Training
Aptitudes by Treatments	Training Management
Learning Saturation	Transfer of Training
Perceptual Learning	Unit Training
Retention	Visual Recognition

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

Applied training research measured the effectiveness of training visual aircraft recognition (VAR), evaluated the transfer of training from static to dynamic imagery testing; and evaluated VAR saturation and retention levels in an Army unit training environment. It was found that, although the overall average gains in training were the same for the two training methods, lower aptitude men learned more under lock-stepped instruction, and higher aptitude men gained more under self-paced instruction.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20. The VAR skills developed from training with static images did transfer to a dynamic test situation involving model aircraft which moved.

Men from four batteries participated in a program in which 20, 40, 60, and 80 aircraft had to be learned in successive blocks. All the test personnel from two batteries achieved the final goal of 80 aircraft. None of the members of one of the other batteries achieved the program objectives. Variation in achievement of the batteries was attributed to differences in training management and practice.

FOREWORD

The ARI Field Unit-Fort Bliss is actively engaged in a program of research to develop and evaluate Air Defense training programs responsive to the Army Training and Doctrine Command (TRADOC) policy and guidelines for achieving maximum training effectiveness. This publication involves training in Visual Aircraft Recognition (VAR) a critical human skill required by all forward area air defense gunners.

The research reported here represents an evaluation of an experimental, self-paced, VAR training program tailored to the training requirement of FORSCOM AD units. The program was compared with the current Army VAR training program based on a lock-step training method.

ARI research in this area was conducted under Army Project 2Q762717A745, FY 76 Work Program. The work reported here was accomplished by personnel of the Human Resources Research Organization (HumRRO), El Paso, Texas, Contract No. DAHC19-75-C-0020, under the technical supervision of the ARI Field Unit-Fort Bliss.

J. E. UHLANER,
Technical Director

THE ACQUISITION AND RETENTION OF VISUAL AIRCRAFT RECOGNITION SKILLS

BRIEF

Requirement:

To compare the relative effectiveness of two methods of training visual aircraft recognition (VAR) skills; to evaluate the transfer of training on static imagery to dynamic field conditions; and to determine the VAR saturation and retention levels of Chaparral and Vulcan crewmen in a unit training environment.

Procedure:

A program of applied training research was conducted in three phases. In Phase 1, the classroom program currently specified for VAR training using 35mm slides was compared with a printed VAR program which permitted self-paced, individualized instruction. In Phase 2, soldiers trained by these methods were tested in a miniaturized field test involving dynamic presentations of model aircraft. In Phase 3, soldiers were trained to recognize blocks of 20, 40, 60, and 80 aircraft. The 4th Bn, 1st ADA (C/V), provided the troop support. Contractor personnel administered the training during Phase 1 and administered all official proficiency tests.

Findings:

Although the average increase in VAR proficiency during Phase 1 was the same for both training methods, it was found that the lower aptitude personnel gained more under the lock-step classroom method, whereas persons of high aptitudes learned more under the self-paced, individualized training method.

During Phase 2, it was found that proficiency acquired as a result of training on static imagery did transfer to dynamic testing conditions. Persons who had achieved at least 70 percent accuracy during training, had VAR accuracies of at least 90 percent on the dynamic tests. Persons who had less than 70 percent training averages did not meet the 90 percent field standard. Performance in the dynamic test was not affected by the training method employed.

During Phase 3, personnel from all four batteries of the 4th/1st participated in a program to learn 80 aircraft during a four-month period. The printed training program was used. This permitted batteries to conduct individualized instruction. The results of contractor-administered proficiency tests revealed that all test personnel in two of the batteries could learn 80 aircraft with an average accuracy of about 88 percent.

Three men from a third battery achieved that goal, while the fourth battery suspended training before anyone met the program objectives.

The batteries that achieved the 80-aircraft objective required an average of 4-5 minutes per aircraft for learning. Differences in the battery achievements were attributed to variations in the training management policies and practices employed. The batteries that attained the program objectives used the desired self-paced, individualized approach. The other batteries employed a more instructor-oriented, lock-step method of training. These training management variations were reflected in both the learning rates and the retention scores. Batteries using lock-step methods scheduled training over many calendar days, and experienced large retention losses. The opposite effects occurred for the batteries that used self-paced training.

Utilization of Findings:

Several research findings are of use to TRADOC and FORSCOM training managers. First, training aids which permit individualized instruction produce greater learning achievements for the majority of soldiers, and require less learning time.

Second, the finding that training on static imagery does transfer to skills required in dynamic VAR situations should allay recurrently stated concerns of training managers that static training is ineffective and that dynamic VAR training materials are required.

Finally, the finding that under self-paced conditions, AD crewmen can learn to recognize at least 80 aircraft has significance for the development of proficiency standards used for evaluating MOS proficiency and unit effectiveness.

THE ACQUISITION AND RETENTION OF VISUAL AIRCRAFT RECOGNITION SKILLS

EXECUTIVE SUMMARY

OBJECTIVE

In FY 75, the US Army Research Institute for the Behavioral and Social Sciences (ARI) awarded a contract to the Human Resources Research Organization (HumRRO) to conduct research concerning visual aircraft recognition (VAR). This research had the following objectives.

a. Compare the relative effectiveness of VAR training for (1) the program specified by current Army training literature which employed the 35mm GOAR Kit; and (2) an experimental program developed by the contractor which used printed images and was designed to permit self-paced, individualized instruction.

b. Estimate the transfer of VAR proficiency from training/testing programs using static imagery to a testing environment involving dynamic imagery.

c. Estimate the maximum number (up to 80) of aircraft that can be learned and remembered by Army forward area air defense crewmen.

This research was performed in a unit training (UT) environment rather than Advanced Individual Training (AIT). A UT environment was desired for two reasons: (1) the effectiveness of VAR training in units had never been investigated and (2) the seven-week AIT schedule could not accommodate the project's four- to six-month training requirement.

METHODS AND RESULTS

Phase 1. Compare GOAR and Printed VAR Training

Method. In April 1975, 205 persons representing all 4 batteries of the 4th/1st ADA, were given a pretest by contractor personnel covering the 25 aircraft on the battalion's VAR priority list. This test consisted of a random sequencing of the views of each of 25 aircraft. The two batteries having the lowest average accuracy scores subsequently were selected to participate in the training program comparison.

The training programs for both batteries consisted of six skill acquisition sessions and three review sessions. Each skill acquisition session involved instruction concerning four to six aircraft which had been assigned to instructional modules on the basis of their apparent visual similarity as judged by members of the contractor's staff. Review sessions were conducted after the third and sixth learning sessions and included all aircraft presented during the preceding three learning sessions. A final review of all 25 aircraft was given at the end of training.

All training was administered by contractor personnel and was scheduled by the two batteries along with other training and operational requirements. In the design of battery training schedules, the VAR training was not afforded a special or high priority.

Battery X received training with the GOAR Kit. The training method employed was based on official Army documentation concerning the use of the GOAR Kit for VAR instruction. Training was instructor-oriented and lock-stepped. Individual trainees were called upon to state name and numerical designation for each image shown. All trainees were given an achievement test at the end of each skill acquisition section.

Battery Y was instructed with a printed version of the GOAR imagery. Students were given multi-view cards and flashcards of each aircraft. The multi-view cards had the name, alpha numerical designation, country of origin and major distinguishing characteristics on one side and five different aspects of the aircraft on the other. Each flashcard presented one aspect of an aircraft on the front with its nomenclature on the back. The training system designed for Battery Y was based on the "Learning Center" concept. Instructional personnel were available at times specified in the battery's training schedule; the soldiers were free to come and go anytime during these scheduled sessions. The training materials were controlled by the contractor and were not available to the men between the scheduled sessions.

Upon completion of training in both batteries, a posttraining proficiency test was given. This test was identical to the pretraining test.

Results. The average VAR accuracy in each battery increased from an accuracy level of about 56 percent before initiation of training to a posttraining level of about 66 percent. This overall low increase in VAR proficiency was surprising. However, a review of training attendance records revealed that only 1 of the 132 men had attended all 9 training and review sessions, and over 50 percent of the trainees had attended less than 50 percent of these sessions. Additional analyses, nevertheless, did suggest that there was an interaction between a trainee's aptitude and the increase in VAR proficiency associated with the two training programs. Persons with low learning aptitudes tended to learn more if they participated in the lock-step, GOAR methods; persons with intermediate and higher abilities learned more if they used the individualized, printed medium. VAR proficiency also was found to be directly related to enlisted rank and also to the number of review sessions attended.

Task 2: Transfer from Static to Dynamic Situations

Method. Previous research has shown that valid simulation of a full-scale VAR environment can be produced by using scale models of aircraft which are artificially placed in dynamic movement and presented to ground observers. Such a miniaturized dynamic field test was designed to test for the transfer of training from static images to dynamic imagery.

The dynamic test was administered about two weeks after the batteries had completed training with the GOAR and printed materials. The dynamic test was given to about 100 men, equally apportioned between the two training programs. Thirteen models of the twenty-five aircraft included in the static training were available for use the dynamic test. Each model was presented five times, providing a test consisting of sixty-five trials. The models were transported toward the observers at a scaled speed of 400 knots and at a scaled altitude of 500 meters. No optical aids were provided. Instrumentation was established to provide means of recording (1) the scaled range at the time of a recognition judgment and (2) the accuracy of that judgment. The observers were provided with a list of the 13 aircraft that would be presented.

Results. Measures of VAR accuracy and the average recognition range (adjusted for the miniaturized test's scale factor) were obtained for each observer. The average VAR accuracy in the dynamic test was 82 percent. The average distance (adjusted) of the aircraft at the time of a correct VAR response was about 1270 meters. The correlation between proficiency levels on the posttraining test involving static images and the dynamic image test was 0.67. This statistically significant correlation indicated that VAR proficiency did transfer from static to dynamic situations. No difference was found in the VAR accuracies of men trained by the two training methods. The average recognition ranges for representatives of both batteries also were not significantly different.

At first glance, an average recognition range of 1270 meters would seem to comprise relatively poor performance. However, the research measure of recognition range was for unaided visual recognition. Previous research has indicated that if optical aids are employed, the average recognition range increases about 1,000 meters over unaided observation.

Task 3: VAR Saturation and Retention

Method. Representatives of all four batteries of the 4th Bn, 1st ADA, participated in the final task of the research program. Since it was anticipated that the duration of a training program which included 80 aircraft would extend for several (4 to 6) months, each battery designated key personnel who were expected to be available during that time period. One hundred such men were assigned to this training task. Of these, 83 were still available for training and evaluation when the data collection had to be terminated on 30 April 1976, because of other battalion commitments.

The objective of this phase of the research was to estimate the proportions of AD crewmen who could learn (and remember) to recognize 20, 40, 60, and 80 aircraft. It was also desired to obtain such data in an operational based training environment. That is, the units would conduct the training, the contractor would monitor their progress, and would administer periodic VAR proficiency tests. In this phase, HumRRO had no control over the conduct or scheduling of training, nor did HumRRO give any training.

HumRRO, however, did provide the training materials to the batteries. These materials consisted of the printed training aids, which included the five-image, multi-view cards and single-image flashcards. These materials were produced for 80 aircraft, which were grouped into 4 blocks of 20 aircraft. Each block was further partitioned into groups of 4 to 5 aircraft.

The training was scheduled and administered by battery personnel. HumRRO provided guidance to each battery concerning the method of instruction to be used and recommended that personnel not be scheduled for the end-of-block proficiency tests until each individual demonstrated a VAR accuracy level of 70 percent based on tests using the training imagery.

The HumRRO-administered tests employed GOAR imagery and included aircraft views not included in the training program. The end-of-block tests were given to groups of 3 to 20 men, as soon as the battery training officer decided that individuals had acquired an adequate level of proficiency.

The end-of-block test scores were cumulative in content. As each crewman progressed through the program, he was retested on previous blocks and then tested on the most recent block of aircraft he had learned. No requirement was made for refresher training. The need for refresher training, if any, was decided by the battery's training officer. This system of cumulative testing provided a means of evaluating VAR retention levels concurrently with testing to determine saturation levels.

Results. Two of the batteries completed the objective of training their men to recognize 80 aircraft. Three men from a third battery achieved this goal, while the fourth battery suspended VAR training after the third block of testing (60 aircraft). The average VAR accuracy for the 31 men who completed the 80-aircraft program was 84 percent.

The differential achievement of the four batteries was attributed to policies and practices associated with training management. The two batteries that fully attained the training objectives employed the desired self-paced, individualized approach to training. The other batteries employed a more instructor-oriented, lock-step approach to training and requested contractor testing only when a majority of the men were judged to be prepared for such evaluation. For example, Battery 01, which completed the program first, requested 14 testing sessions, involving groups of 4-7 men. In contrast, Battery 03, which suspended participation after 60 aircraft, only requested 3 testing sessions over an interval of 115 calendar days.

Learning rates were quite high for the two batteries which reported this data to the contractor. Battery 01 required an average of 4 minutes per aircraft; Battery 02 required 5 minutes per aircraft. Training time data were not available for the two batteries that used a lock-step method.

Differences among battery achievements were not related either to variation in average GT scores or to the amount of training time expended. Differences in VAR accuracy levels among batteries was related to the amount of elapsed time in the program. The batteries having the higher VAR averages spent less time completing the program than the other batteries.

Retention levels were good. When averaged over all batteries, accuracy scores only declined between four and five percentage points from the initial testing on Block 7 to the final retesting on that block. Battery 03, which only requested three testings, evidenced the greatest decline: a 14-point decrease in accuracy over time to a final retest average of 64 percent. The retention losses for Battery 03 were significantly greater than those experienced by the other units.

OTHER DATA

The report presents a discussion of the psychometric properties of the VAR tests and compares VAR accuracy on familiar versus unfamiliar views. Data concerning the most common recognition confusions also is presented. Analyses are discussed of alternative procedures for identifying those men who possess high potential for success on VAR training. The report concludes with a section concerning observations about training management practices and their impact upon training achievements.

CONCLUSIONS

The Effectiveness of a Printed Training Aid

The results of (1) the direct comparison of the GOAR and Printed Training Aids, and (2) the VAR achievements of persons in the saturation study indicate that VAR proficiency can be effectively and efficiently accomplished when a printed training medium is employed. The results do suggest that the less able learners achieve more learning when a highly structured training method is employed. Persons with intermediate learning abilities achieve more when more individualized training methods are used, and the highest ability groups learn equally well, at least, under either method of instruction.

Transfer of VAR Proficiency to Dynamic Tests

VAR proficiency does transfer from classroom instruction to dynamic field situations. Differences between the two training methods were not reflected in the field test proficiency levels of trainees. Routine use of miniaturized field testing to supplement classroom training was a beneficial component of the training system.

Saturation and Retention

Soldiers did learn to recognize 80 aircraft with an average accuracy of about 85 percent. The learning achievements of crewmen is highly dependent upon the priorities, scheduling, and training methods employed by a unit. If a self-paced, learner-oriented, criterion-referenced, and individualized approach to training is employed, VAR training time can require as little as four to five minutes per aircraft.

THE ACQUISITION AND RETENTION OF VISUAL AIRCRAFT RECOGNITION SKILLS

CONTENTS

	Page
SECTION 1 - BACKGROUND AND OVERVIEW	1
THE MILITARY PROBLEM	1
Machine-Ascendant Systems	1
Man-Ascendant Systems	1
AREAS IN NEED OF RESEARCH	3
Static Images	3
Media, Methods, and Materials	4
Saturation of Recognition Skills	4
RESEARCH OBJECTIVES	4
SECTION 2 - A COMPARISON OF TWO TRAINING METHODS	7
OBJECTIVES	7
GOAR Training Method	7
Experimental Training Method	7
RESEARCH METHOD	7
Trainees	7
Pretesting	11
Training	11
Posttraining Test	12
RESULTS	12
Test Performances	12
Attendance	13
Factors Affecting VAR Proficiency	14
SECTION 3 - TRANSFER OF VAR PROFICIENCY FROM STATIC TO DYNAMIC IMAGERY	23

Page

PROBLEM 23

METHOD 23

Miniaturized Field Test 23
Observers 23
Performance Measures 25
Data Reduction 26

RESULTS 26

Average Scores 26
Intercorrelations 26
Transfer Effects 26
GOAR versus Printed Training 27

DISCUSSION 27

SECTION 4 - SATURATION AND RETENTION 29

PURPOSE 29

METHOD 29

Personnel 29
Training Materials 29
Self-Paced Training 29
Training Standards 29
Independent Variables 30

RESULTS AND DISCUSSION 30

Unit Variations 30
Learning Rate 32
Retention 37
Data Analysis 40
Overall Losses 40
Loss as a Function of Time 40

SECTION 5 - PREDICTING VAR PROFICIENCY 43

SCOPE 43

TEST RELIABILITY 43

Page

REDUCING TEST LENGTH 43

PROFICIENCY FOR FAMILIAR VERSUS UNFAMILIAR VIEWS 44

Phase I Analysis 44
Phase III Analysis 44
Accuracy as a Function of View 45
The Relative Difficulty of Various Aircraft 46
Response Confusion 47
ACB versus VAR Proficiency 48
Prediction Using a Screening Test 48

SECTION 6 - OBSERVATIONS ON TRAINING MANAGEMENT AND COST EFFECTIVENESS 51

LITERATURE CITED 53

APPENDICES 55

TABLES

1 Mean VAR Proficiency Scores 13

2 Intercorrelations Between VAR Posttraining Score and Other Factors 15

3 R^2 and Raw Score Weights for Selected Regression Equations Predicting Posttest Score 16

4 Average Pre/Post Test Scores and Standard Deviations for Subgroups 1, 2, and 3 17

5 Summary of Veldman's ANCOVA for Two Factors and One Covariate 18

6 Aptitude Area Scores for the Three Homogeneous Subgroups 19

7 *F*-Ratios for Between Group Variations in ACB Scores 20

8 *T*-Tests of Differences Between ACB Average Scores of Subgroups 21

9 Intercorrelations Among Posttraining and Field Test Measures 26

10 General Results of Saturation Study 31

11 Retention Data for the Initial Block of Aircraft 39

12 VAR Accuracy for Familiar and Unfamiliar Views: GOAR versus Printed (Mean Percent Correct) 44

13 Recognition Accuracy for Ten Aircraft Views 45

14 Recognition Accuracy Scores for Twenty-Five Aircraft 46

15 Response Confusions for Twenty-Five Aircraft 47

16 Frequencies of Confusions for Aircraft of Different National Origins 48

FIGURES

		Page
1	Five-View Cards	8-9
2	Multi-View Flashcards	10
3	1/72d Scale Field Test	24
4	Acquisition Curves for Battery 01	33
5	Acquisition Curves for Battery 02	34
6	Acquisition Curves for Battery 03	35
7	Acquisition Curves for Battery 04	36
8	Program Completion in Terms of Mean Number of Aircraft Trained	38

THE ACQUISITION AND RETENTION OF VISUAL AIRCRAFT RECOGNITION SKILLS

SECTION 1

BACKGROUND AND OVERVIEW

THE MILITARY PROBLEM

The defense of US Army ground combat and support units against aerial attack, reconnaissance, and surveillance depends upon the capabilities of a wide variety of air defense weapons systems, as well as upon some ground combat weapons which have an anti-aircraft capability.

The air defense weapons currently deployed with the field Army can be divided into two classes: (a) machine-ascendant and (b) man-ascendant.

Machine-Ascendant Systems. The machine-ascendant weapons consist of the Nike Hercules guided missile system and the Hawk guided missile system, which are designed to engage high- and medium-altitude aircraft, such as strategic bombers. Although these two systems have different capabilities with respect to warhead types and lethality, and engagement ranges and altitudes, they are similar in that detection, identification, and tracking (engagement) are accomplished electronically through the use of machine-ascendant aids, such as radars, computers, and electronic identification--Friend or Foe (IFF). Although soldiers operate these systems, they interact with radar-generated displays and other electronically-generated indicators. In a machine-ascendant system, the majority of the crewmen cannot witness or directly view the aircraft that is under attack. When people who are generally familiar with the armed services think about air defense weapons, they tend to focus their attention on these highly sophisticated machine-ascendant guided missile systems, since the Hercules and Hawk systems have been in the air defense spotlight since the early 50s.

Man-Ascendant Systems. Currently, the US Army man-ascendant air defense (AD) weapon systems consist of the Vulcan, Chaparral, and Redeye weapon systems, which are employed for defense against aircraft flying at low altitudes (less than 3 kilometers) in the forward area of the combat zone. These weapons are specifically designed to engage low-flying aircraft. The Chaparral and Vulcan weapons are organized into Chaparral/Vulcan (C/V) Battalions, such as the 4th Bn, 1st ADA, and are assigned to support other

FORSCOM ground forces. A Vulcan system employs a multi-barrel "Gatling Gun," originally designed for use on USAF Aircraft. The weapon may be self-propelled or towed. The self-propelled version has the chassis and power train of an armored personnel carrier. The Vulcan is equipped with a radar for determination of gun-to-aircraft distance and an optical computing sight, which is a modification of the MARK XIV computing sight originally used by the USN in WWII.

The Chaparral weapon system is also hybrid. It consists of a modified, heat-seeking Sidewinder missile, originally developed for USAF use, and a light tank chassis. It has a greater range and hit probability than the Vulcan. The Chaparral does not have any radar components. Neither Vulcan nor Chaparral have an on-board electronic IFF capability.

Redeye is a visually-aimed, shoulder-fired weapon, similar in basic concept to a "bazooka." Redeye fires a heat-seeking missile that has a limited range. It's gunners normally are deployed in two-man teams, each man equipped with one weapon. Redeye teams do not possess any capability for electronic IFF, although they may receive early warning and IFF data by means of radio communications.

Vulcan, Chaparral, and Redeye are considered man-ascendant systems because they are wholly or partly dependent upon the soldier's ability to (a) detect aircraft, (b) identify it as friend or foe, and (c) decide when to engage it based upon range estimation using gunsight profiles as aids. Similar human skill requirements are placed on soldiers in the ground combat branches when rifles and machine guns, for example, are employed against aircraft.

A considerable amount of field and laboratory research and development has been performed concerning the visual detection, identification, and engagement of aircraft by operators of man-ascendant systems (*1*). The majority of this research was conducted by HumRRO under contracts with the Army Research Office. In brief summary, this previous HumRRO research, which was conducted during 1963-1972, produced information concerning the capabilities of ground observers to detect and recognize aircraft, and judge the gun-to-aircraft ranges. Information about a technique for engaging aircraft with small arms was developed. Some preliminary studies concerning visual search techniques and the utility of optical aids for detection were also accomplished.

Although additional research appears needed concerning visual surveillance methods and aids, as well as the use of small arms in the air defense role, there appears to be a more pressing need for research on methods of monitoring (*i.e.*, measuring) and maintaining the aircraft recognition skills of men after they are assigned to air defense units. Since the final decision to engage or not engage an aircraft, for a man-ascendant system, is dependent upon the perceptual abilities of air defense crewmen or NCOs, cost-effective techniques for monitoring and maintaining these skills are essential to minimize identification errors.

AREAS IN NEED OF RESEARCH

Recognition[1] of aircraft on the basis of visual information is a complex skill which requires extensive training to develop an adequate level of proficiency. HumRRO research during the past decade has produced more efficient training methods and the first systematic, research-based set of training images ever developed for use in aircraft recognition training (*2,3,4*). These training images, presented via 35mm slides, have been issued by the Army as the Ground Observer Aircraft Recognition (GOAR) Kit. It is the primary vehicle through which current training and assessment in aircraft recognition is accomplished. Although current Army standards for the performance of AD crewmen in AIT require a proficiency of 90 percent correct recognitions, some of the testing conditions to be used in defining this level of proficiency are currently unspecified; *e.g.*, the number and kinds of views of each aircraft which should be included in the test. The prior research (*2,3*) which led to the development of the GOAR Kit imagery provides a substantial amount of evidence suggesting that this imagery is valid for use in aircraft recognition training and testing. However, a number of questions regarding the use of this imagery remain unanswered.

Static Images. Training programs in aircraft recognition have typically used static images, as found in the GOAR Kit, to represent aircraft in flight. Such images are frequently more convenient and cheaper to use in both training and testing than are dynamic (moving) images. Static images also have certain methodological advantages. For example, they permit the occurrence of more aircraft recognition responses per unit time in either training or testing situations. However, little objective information is available about the adequacy of static imagery for training observers to recognize the dynamic images they will encounter under actual operational conditions. Obtaining reliable information about this issue cannot be accomplished under peacetime conditions because of the obvious lack of access to hostile aircraft. However, proximate answers can be obtained through the examination of closely related problems in simulated situations. For example, the extent of recognition accuracy as measured by a static image in the GOAR Kit can be correlated with recognition accuracy and range as measured by the dynamic image presented by moving, small-scale model aircraft.

[1]The use of the terms "recognition" and "identification" in the Army is at variance with psychological usage. In the Army, an aircraft is identified when it is classified as Friend or Foe. Similarly, an aircraft is recognized when the observer can assign a name or alpha-numeric designation.

Media, Methods, and Materials. There remain a host of unanswered questions about aircraft recognition training, such as the optimal mix of various training media, methods, and materials. Past HumRRO research has provided preliminary evaluation of alternative printed media such as flash-cards and other printed materials (*4*). However, no studies have been accomplished which directly compared the training effectiveness of projected versus printed imagery. Until more definite information is available on the general issues of the role of media in aircraft recognition, only educated guesses about the feasibility of individualized instruction can be provided with respect to aircraft recognition training.

Saturation of Recognition Skills. A third research area is concerned with saturation levels. It is reasonable to assume that there is a practical limit (saturation level) to the number of aircraft that can be remembered, and that the saturation level will vary among individuals. However, the relationship between retention and the number of aircraft to be remembered is currently unknown and considerable divergence of opinion exists among trainers concerning the saturation level. For example, a recent survey of the four C/V batteries of the 4th/1st ADA (4/1) revealed that the number of aircraft included in each battery's training program varied between 40 and 82. In contrast, FM 44-30, *Visual Aircraft Recognition*, specified that ADA crewmen should be able to "expertly" discriminate 20-30 aircraft. "Expertly" usually is defined as achieving a minimum recognition accuracy of 90 percent.

There is obviously a need to provide guidance to Army trainers concerning the saturation limits of the average ADA crewman so that (a) realistic proficiency standards can be established and (b) a basis for estimating when to provide refresher training can be provided.

RESEARCH OBJECTIVES

In Fy 75, the Army Research Institute (ARI) awarded a contract to Human Resources Research Organization (HumRRO) to conduct applied research concerning visual aircraft recognition (VAR). The objectives of the research program were as follows:

a. Compare the relative effectiveness of VAR training as (1) specified in current Army training literature using the 35mm GOAR Kit and (2) an experimental program using printed images as the training materials.

b. Estimate the transfer of VAR proficiency from training/testing using static imagery to a situation involving dynamic imagery.

c. Estimate the maximum numbers of aircraft that can be learned and remembered by typical AD forward area weapons crewmen.

This report presents the results of research concerning these objectives. This report is divided into six sections corresponding to the research objectives. Section 2 describes an experimental comparison of the two training approaches; VAR training using projected versus printed imagery for skill acquisition. Section 3 describes the comparison of VAR proficiency as measured using static versus dynamic imagery. Section 4 presents the results of research concerning "saturation" levels and retention. Section 5 discusses psychometric characteristics such as test reliability and performance on familiar and unfamiliar views. Section 6 discusses certain aspects of cost/effectiveness with respect to options available for accomplishing VAR training.

SECTION 2

A COMPARISON OF TWO TRAINING METHODS

OBJECTIVES

The objective of the initial phase of the research was to compare the effectiveness of two methods of providing VAR training: (1) a "control" program specified by current Army training literature, in which 35mm slides are used with lock-step classroom method of instruction; and (2) an experimental program involving printed images which permit self-paced and individualized instruction.

GOAR Training Method. At present, the Army's approved method of training visual aircraft recognition (VAR) involves a classroom display of 35mm slides of aircraft, using the GOAR Kit. The training is given in the so-called "lock-step" fashion, in which personnel are shown views of a specific group of aircraft during a one- to two-hour period. One training session typically is scheduled each week, for a total of about four to eight hours of VAR instruction each month.

Guidance for the conduct of such training is contained in DA FM 44-30 and TC 44-30. In this method, an instructor controls the rate of display of the aircraft images and calls upon individual students to name the aircraft and describe its recognition features.

Experimental Training Method. In the late 60s, HumRRO developed a prototype training program which used printed aircraft images for VAR training. In this method, soldiers worked in "buddy-pairs" and proceeded to learn to recognize (name) aircraft images under self-paced conditions. In this situation, the instructor served as a training manager rather than an active contributor or imparter of information. Figure 1 presents five-view cards used as training aids in the experimental program; Figure 2 presents the multi-view flashcards.

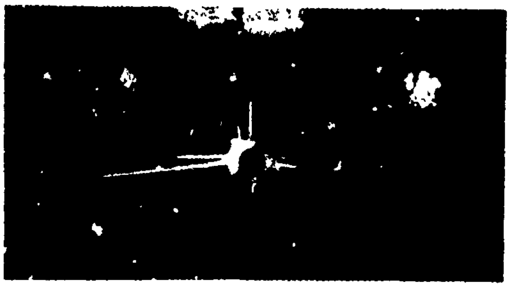
RESEARCH METHOD

Trainees. The personnel designated to support this research were provided by the 4th Bn, 1st ADA. Personnel from a tactical (FORSCOM) unit were requested for several reasons:

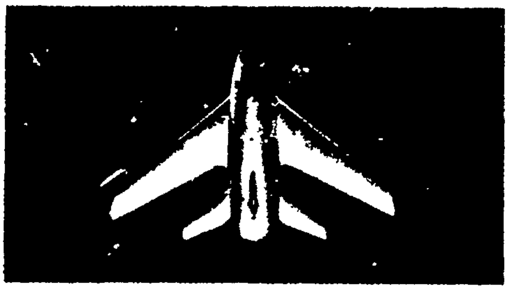
1. It was desired that the supporting units view the training research as directly relevant to their operational mission.

2. It was desired that the individuals serving as "subjects" meet the aptitude requirements for forward area crewmen (MOS 16P & 16R).

MIG-19 FARMER USSR

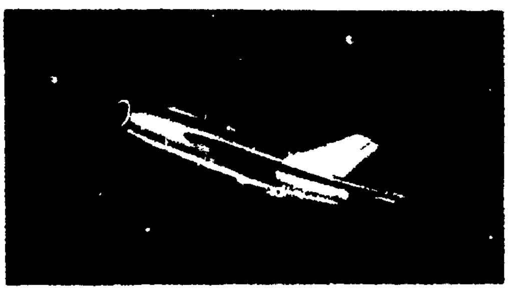


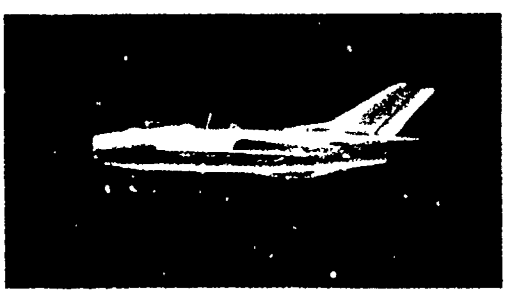
WINGS: mid; swept; has fences; slant down



TAIL FLATS: high on body; swept

INTAKE: 1; in nose; small; has vertical divider





BODY: short; fat; raised canopy

Figure 1. Five-View Cards

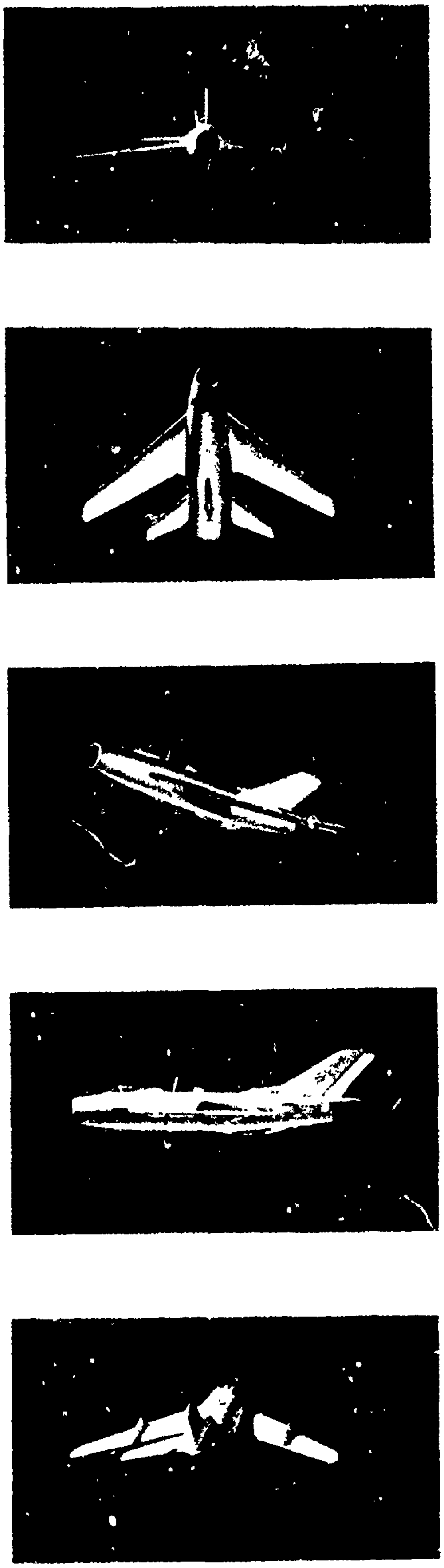


Figure 1 continued.

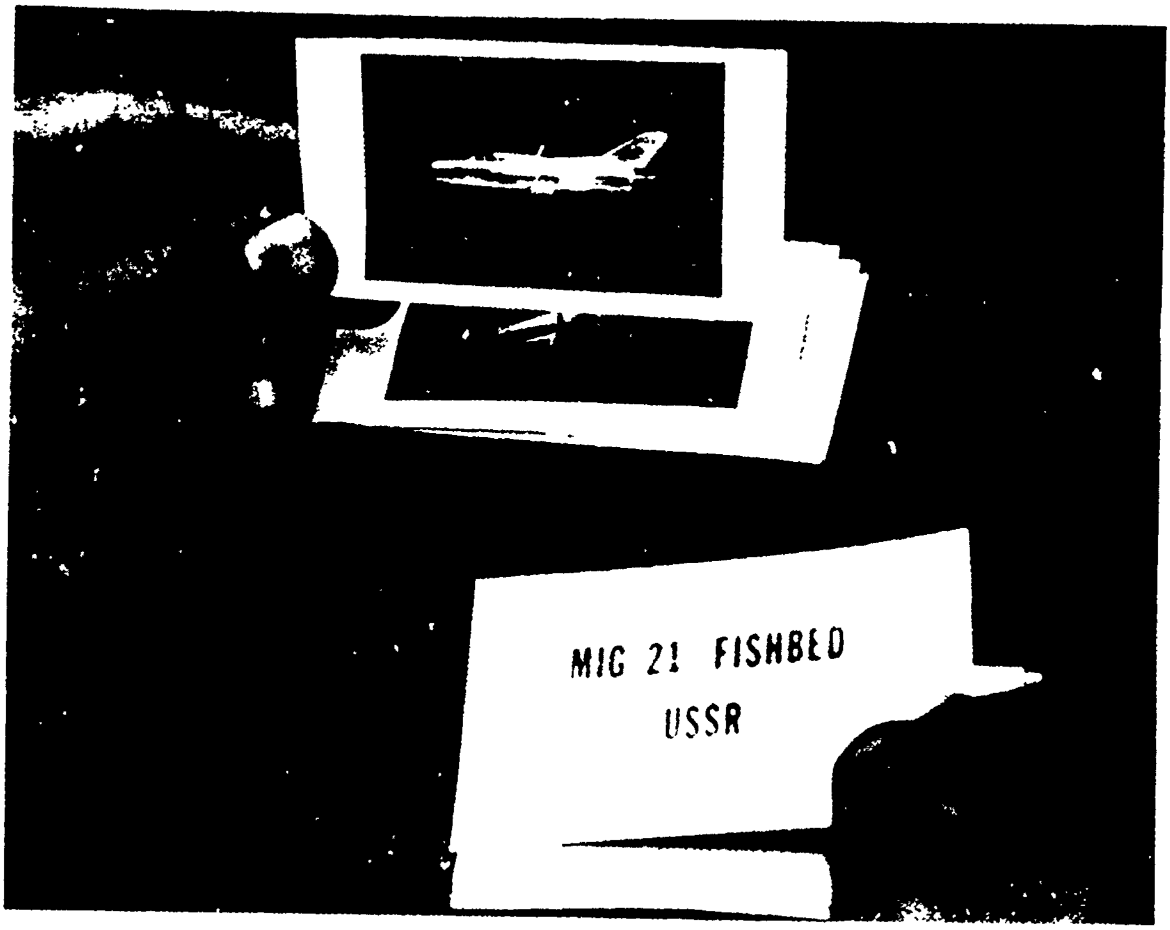


Figure 2. Multi-View Flashcards

3. Since the complete training research effort would encompass 80 aircraft, the personnel to be trained would have to be stabilized in their job assignments for several months.

4. The Air Defense School had established a requirement for printed VAR materials for issue to units. It was considered most desirable to evaluate the effectiveness of these materials in a unit training context. However, in order to achieve these goals, it was recognized that the soldiers who would receive the training already had received VAR training as part of (1) Advanced Individual Training (AIT), and (2) as part of their previous unit training. In other words, the men would already posses unknown levels of VAR proficiency prior to the initiation of the research program.

Three hundred and eighteen men participated in the first phase of the project at one time or other. Of these, 132 took a posttraining test and 61 took both the pre- and posttraining tests.

Pretesting. A pretraining test was administered to 205 men from all 4 batteries of the 4th Battalion. The purpose of the pretest was to obtain data concerning the *status quo* of VAR proficiency before initiating the experimental training. This test consisted of 10 views of 25 aircraft which had been designated as "most critical" by the battalion. About one month before the pretest was given, the batteries were advised by the Battalion Commander that a battalion-level test would be given and they were told which aircraft would be included. This test was given on 8-9 April, 1975.

The personnel were tested in groups of 15-20 in a conference room. They were seated at three-feet intervals horizontally from a front projection screen. The viewing distance from the screen to the men varied between 12 and 24 feet. The size of the image varied between 4 inches in width for the head-on view of the Cobra helicopter to 14 inches for the side view of the Ute aircraft. Viewing angles for individuals varied between 5 and 45 degrees off the perpendicular. Thus, the visual angles of the images varied between 0.8 and 5.4 degrees for the viewer, depending upon the latter's location in the room and the horizontal extent of the image on the screen. The images were displayed for 5 seconds with a 5-second interval between presentations. The men wrote the answers on a test form. The 250 items required about 50 minutes to administer, including a short break in the middle of the test.

Training. The two batteries which had the lowest pretest scores were designated to participate in the comparison of the GOAR and Printed Image training programs. Battery X was assigned to receive the GOAR training and Battery Y received the experimental training. The training encompassed the 25 aircraft used in the pretest.

In each case, training was scheduled to be conducted in nine sessions: Six initial skill acquisition sessions and three review sessions. Each skill acquisition session provided instruction on 4 to 6 aircraft which had been grouped by the research staff on the basis of judgments of visual similarity. Review sessions were scheduled after the third and sixth skill acquisition periods. A final review of all 25 aircraft was given at the end of the training program and about 1 week before administration of a posttraining test. The duration of each training period varied between one to two hours for the GOAR method and between two to three hours for the printed media program, depending on the number of aircraft included and, in case of the printed media, the number of persons attending training. The printed media sessions were longer for two reasons: (1) since training was individualized and self-paced, the men worked at different rates and the test given at the end of each skill acquisiti. session was not lock-stepped; (2) the training was conducted in an "open classroom" or learning center environment--the men could come and go at any time during the period, depending upon their schedule. In contrast, the GOAR training was conducted according to the traditional classroom schedule: all persons began at the same time and were tested at one time.

The training in both batteries was conducted (GOAR) or supervised (Printed) by a member of the research staff. This training covered only five of the ten views available in the GOAR slide kit. The five views were selected to correspond with those specified by the Air Defense School in its requirement for procurement of a printed image training aid. These views included the following:

a. Head on, or 0-degree climb and 0-degree heading.

b. Incoming-overhead, or 35-degree climb and 0-degree heading.

c. Over-the-shoulder, or 15-degree climb and 45-degree heading.

d. Crossing, or 0-degree climb and 90-degree heading.

e. Receding, or 15-degree climb and 190-degree heading.

Limiting training to less than the complete set of 10 views is prescribed in DA TC 44-10 and also permitted analysis of test data to evaluate transfer of VAR proficiency from familiar to unfamiliar views.

Posttraining Test. A final examination was given about one week after the final review session. This test was identical with the pretest and was given under testing conditions like those described for the pretest. This test was given on 1-2 September, 1975 to 130 men.

RESULTS

Test Performances. The average pre-and posttraining scores btained by each battery are presented in Table 1.

Table 1

Mean VAR Proficiency Scores

Method	Pre	TEST Post	Gain
GOAR (N = 30)	56.2	65.9	9.7
Printed (N = 31)	56.8	66.5	9.7

Attendance. Although the average gain in proficiency was statistically significant ($F = 26.20$; $p < .001$), its relatively low magnitude (10%) was unexpected. In an effort to identify possible reasons for such a low gain, the attendance records during training for all personnel who took the post-training test were examined. The results of the analysis of attendance records is presented below:

No. of Training Sessions	Number of Men	Percent of Total Men
0	1	0.8
1	7	5.3
2	9	6.8
3	18	13.5
4	39	29.3
5	21	15.9
6	20	15.0
7	9	6.8
8	7	5.3
9	1	0.8

To summarize this data, it is apparent that more than 50 percent of those given the postraining test attended less than 50 percent of the training and review sessions. Their posttraining proficiency thus was influenced by a lack of skill acquisition as well as possibly low skill retention.

Factors Affecting VAR Proficiency. Additional analyses were conducted to identify factors that may have contributed to the proficiency levels shown by the 61 men for which pre- and posttraining records were available. The following variables were evaluated.

a. pretraining test scores,

b. number of learning sessions attended (ranged from 1 to 6),

c. number of review sessions attended (1, 2, or 3),

d. training methods used,

e. individual differences in aptitudes, and

f. rank.

Over 200 analyses were performed using covariance and factor analysis methods. The major results of these analyses are described below. The simple relationship between each of the independent variables and VAR proficiency will be summarized first. The more complex analyses of interactions between aptitude factors and the other independent factors will be presented.

Simple effects. Table 2 presents the product moment correlation between VAR posttraining proficiency and each of the independent variables, including ACB Area Scores.

For the sample of 61 men who took both tests, the major factor influencing posttraining achievement was the pretraining proficiency (the product moment correlation between pre- and postraining scores was 0.88).

Whereas, the simple effect of the number of learning (skill acquisition) sessions attended was not significantly related to VAR proficiency (r = .13), attendance at review sessions was significantly related (r = .42). An individual's rank also correlated with VAR skill in a positive manner--the higher ranks achieved higher VAR scores (r = .33).

The frequency distribution of the 61 personnel by military rank was as follows:

E-2 = 5
E-3 = 17
E-4 = 17
E-5 = 11
E-6 = 11
E-7 = 2
0-1 = 1

Table 2

Intercorrelations
Between
VAR Posttraining Score
and
Other Factors

ACB Scores (N = 48)

ACB Scores	*r*	*p*
CO	.44	.01
FA	.34	.05
EL	.29	.05
OF	.36	.05
GM	.38	.01
MM	.31	.05
CL	.22	NS
SA	.17	NS
GT	.27	NS
SC	.34	.05

Other Factors (N = 61)

Factor No.	Name	2	3	4	5	6
1.	Pretest Score	.28*	.11	.06	.09	.88*
2.	Rank		.11	-.03	.27	.33*
3.	Training Method			.02	.25	.17
4.	Number of Training Sessions				-.06	.14
5	Number of Review Sessions					.42*
6	Posttraining Score					

*Statistically significant at $p \leq .05$.

Analysis of the correlations between aptitude area scores and VAR proficiency indicated that 7 of the 10 area scores were significantly correlated with VAR performance. It should be noted that the correlation of VAR proficiency and area score OF is reduced because the crewmen were preselected for AIT on the basis of OF; thus, the range of this aptitude score was reduced in the sample of men. A discussion of the use of other ACB scores for selection of persons for VAR training in units is presented in Section 5.

Multiple regression. Multiple regression analyses were also performed to provide estimates of the amount of variance in VAR posttraining that may be predicted by various combinations of factors (excluding ACB scores). Table 3 presents the value of R^2 and the raw-score weights for a sample of regression equations.

Table 3

R^2 and Raw Score Weights
for
Selected Regression Equations
Predicting Posttest Score

R^2	Pretest Score	Rank	No. Learning Sessions	Review Session			Constant
				1	2	3	
.77	.88						16.60
.79	.87	.55					15.03
.79	.87	.55	.00				15.03
.80	.85	.52	.00	3.78			14.33
.80	.83	.53	.00	3.05	3.01		12.80
.84	.82	.60	.41	1.78	5.90	10.14	6.99

Comparison of the predicted variance for each equation suggested that the main factors contributing to posttraining proficiency were the individual's pretest score, his rank, and his attendance at the final session which reviewed all 25 aircraft. It is of particular interest that the R^2 for the combination of pretest score and attendance at the final review was .83.

The final regression equation reflects a very interesting effect. Whereas the number of learning sessions attended carried no weight for the reduced equations, when the final review session is included, an interaction of this factor and attendance at learning sessions emerged. As a result, learning sessions carried a weight of 0.41 in predicting posttest score and a constant of 10.4 points for attendance at the final review session is added to the predicted posttest score. In other words, the number of learning sessions attended carried no predictive weight if a person did not attend the final review.

Treatment by aptitude interactions. Analyses were performed to examine the hypothesis that learners reacted differently to the two training methods. Factor analytic techniques (cluster analysis) were employed to identify subgroups of learners who possessed similar pre- and posttraining proficiency levels. This cluster analysis was performed separately for each training method. The analysis yielded three groups of persons for each training method. The cluster analysis minimized the variance in VAR proficiency within each subgroup and maximized the difference in average VAR scores between subgroups. Table 4 presents the VAR scores for each of these subgroups.

Table 4

Average Pre/Post Test Scores and Standard Deviations for Subgroups 1, 2, and 3

			Subgroup 1	Subgroup 2	Subgroup 3	Average
PRINTED	Pre	$\overline{X}$	30%	58%	87%	
		SD	11%	8%	6%	
		N	10	12	9	
	Post	$\overline{X}$	34%	76%	94%	
		SD	10%	10%	5%	
		N	10	12	9	
		$\overline{X}$ Gain	4%	18%	7%	10%
GOAR	Pre	$\overline{X}$	28%	58%	83%	
		SD	12%	12%	7%	
		N	11	13	6	
	Post	$\overline{X}$	41%	69%	88%	
		SD	14%	10%	6%	
		N	11	13	6	
		$\overline{X}$ Gain	13%	11%	5%	10%

Inspection of the gain scores for the subgroups suggests that there was an interaction between training method and VAR proficiency. Those persons with low initial proficiency achieved greater gains using the GOAR method (13% *vs.* 4%); whereas, persons with intermediate initial proficiency gained more using the printed materials (18% *vs.* 11%). Only a 2 percent difference between training methods occurred for high ability men. An analysis of covariance (ANCOVA) of the posttraining scores was performed to evaluate the effects of training methods and subgroups. The results of this ANCOVA, which used pretraining score as the covariate, tended to confirm the existance of an interaction between subgroups (ability levels) and training methods; the F-ratio for the interaction of training methods and subgroups was 2.79 (p = .06). The complete ANCOVA summary is presented in Table 5. Supplementary comparisons of the means for the two training methods at each subgroup level did not reveal any significant differences.

Table 5

Summary of Veldman's[1] ANCOVA
for
Two Factors and One Covariate

Factor	MS	df	F	p
Training Method (A)	0.04	1	.0003	.98
3 Subgroups (B)	337.4	2	3.66	.03
AB Interaction	288.1	2	2.79	.06
Within	103.1	55	--	--

[1]Veldman, D. J. *Fortran Programming for the Behavioral Sciences.* New York: Holt, Rinehart & Winston, 1967.

The ACB aptitude area scores for the three subgroups were analyzed to determine the extent to which the variation in proficiency of the subgroups could be attributed to aptitude differences. Table 6 presents the means and standard deviations of the aptitude area scores for each subgroup. The three area scores displaying the greatest variation among the

Table 6

Aptitude Area Scores
for the
Three Homogeneous Subgroups

Subgroup		Pretest	Posttest	ACB Aptitude Area									
				CO	FA	EL	OF	GM	MM	CL	ST	GT	SC
1	$\overline{X}$	29%	38%	98	98	99	104	97	101	101	96	96	97
	SD	11%	12%	12	6	12	8	12	18	8	8	10	10
	N	21	21	12	11	15	11	15	15	16	6	15	9
2	$\overline{X}$	58%	72%	109	102	105	105	105	104	105	100	106	106
	SD	10%	10%	15	14	14	13	13	15	15	14	15	14
	N	22	25	15	10	19	10	19	20	20	7	20	10
3	$\overline{X}$	85%	92%	121	114	113	120	111	116	110	106	108	115
	SD	6%	6%	15	13	18	9	14	13	15	14	15	11
	N	15	15	7	6	13	5	13	13	13	2	13	6

three subgroups were CO, GM, and FA. Aptitude area OF undoubtedly showed less between-group variation since the crewmen had been preselected for AIT on the basis of this aptitude area score. The matrix of ACB scores for the subgroups were analyzed by ANOVA techniques to evaluate the statistical reliability of the between-group variations. The obtained F-ratios and probability levels are presented in Table 7. Supplementary t-tests were

Table 7

F-Ratios for Between Group Variations
in
ACB Scores

ACB Area	F	p
CO	4.49	.02
FA	3.55	.04
EL	3.23	.05
OF	2.74	.07
GM	3.91	.03
MM	3.28	.05
CL	1.48	.24
ST	0.35	.71
GT	3.27	.05
SC	3.21	.05

performed to evaluate differences between pairs of subgroup means for each aptitude area. The obtained t-values are presented in Table 8. In general, the significant differences occurred between high- and low-aptitude groups. Only in the case of GT score was the intermediate-ability group statistically different from the low-ability group.

Table 8

T-Tests of Differences
Between
ACB Average Scores of Subgroups

ACB Area	*df*	Subgroup Comparisons 1 *vs.* 2	2 *vs.* 3	1 *vs.* 3
CO	31	1.95	1.78	3.30**
FA	24	0.79	1.96	2.65*
EL	44	1.15	1.48	2.44*
OF	23	0.21	2.48*	2.69*
GM	44	1.73	1.24	2.90**
MM	45	0.52	2.10*	2.47*
CL	46	1.15	1.04	1.78
ST	12	0.54	0.56	0.92
GT	45	2.07*	0.40	2.47*
SC	22	1.53	1.37	2.67*

*Statistically significant at $p \leq .05$.
**Statistically significant at $p \leq .01$.

SECTION 3

TRANSFER OF VAR PROFICIENCY FROM STATIC TO DYNAMIC IMAGERY

PROBLEM

Currently all VAR training and testing by the Armed Forces requires personnel to recognize (name) aircraft presented as static (still photograph) images. Although some research has been done concerning the recognition of dynamic imagery, these previous studies have either (a) been restricted to a small number of full-scale aircraft, or (b) used observers who probably were not representative of typical crewmen in terms of VAR training, proficiency, and motivation. Thus, there is a need for data concerning the transfer of proficiency from training/testing based on static imagery to the recognition of aircraft images in a dynamic (moving) environment.

METHOD

Miniaturized Field Test. Previous research has shown that valid simulation of a full-scale aircraft recognition environment can be achieved by employing scale models of aircraft which are transported by a motor vehicle in a miniaturized, dynamic field environment (*5*). This type of test situation was designed to evaluate the transfer of VAR skill from a static to a dynamic situation.

Observers. The dynamic test was conducted in September 1975, at Biggs Army Air Field, El Paso, Texas. The personnel tested consisted of approximately 100 men who had received training on 25 aircraft. The dynamic testing was conducted about two weeks after the training had been completed and one week after the men had been given a posttraining proficiency test which used static imagery (projected slides). Approximately equal numbers of men had been trained using the GOAR and the printed imagery.

Of the 25 aircraft included in the training programs, 13 were available as 1/72d-scale models for use in the miniaturized field test. The test consisted of 65 trials in which each of the 13 models was presented 5 times in a randomized sequence. For each trial one of the models was attached to a boom extending from the left top of a van-type motor vehicle. The model was mounted to the boom by means of a right angle metal rod which extended to the rear and down from the model to the boom (see Figure 3.).

Each model was transported toward, and to the left of, the observers. The scale used in designing all aspects of the miniaturized test facility was

PRECEDING PAGE BLANK-NOT FILMED

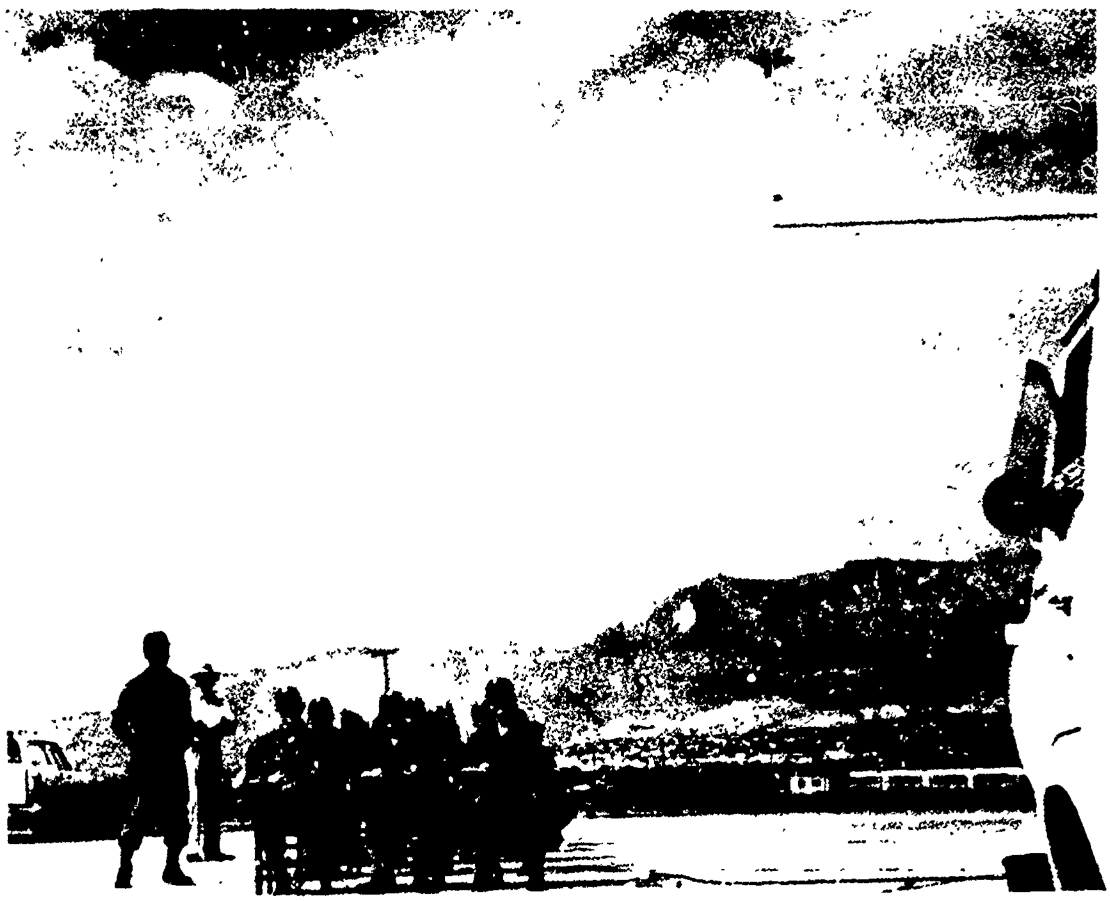


Figure 3. 1/72d Scale Field Test

1/72d, resulting in a scaled velocity of about 400 knots at an altitude of 500 meters. Each trial began with the aircraft model positioned 7,000 meters (scaled) to the front of the observers.

The observers were arranged to the right of the "flight path." They were seated at portable student desks having a writing surface. Six rows of three desks, separated by blinds, permitted simultaneous testing of 18 observers. Each observer was provided with a clipboard which contained (1) an answer sheet, (2) a push button which he used to electrically signal when he made a response, and (3) a list of the names of the 13 aircraft.[1]

Performance Measures. Measures of recognition accuracy (percent correct) and recognition range (adjusted for the scale factor) at the time a recognition judgment occurred were obtained. (Recognition range was obtained from a 20-channel event recorder which displayed the "time point" when each observer signaled a recognition response and another time point when the van drove over a series of pneumatic hoses spaced at even intervals along the vehicle's path.) The range accuracy data subseqently were collated to obtain range for correct judgments. The observers also were allowed to signal when they revised their recognition decisions by merely activating the pushbutton each time they changed their response.

An adjusted range score was also computed for incorrect answers. For all incorrect judgments, a recognition range of minus 1,000 meters was assigned. A penalty of minus 1,000 meters was selected because the aircraft would be beyond engagement range at that distance if the crewman subsequently changed his mind. In other words, if a true "friendly" had been recognized as a "foe," the friendly would have been engaged by the time it reached minus 1,000 meters. Similarly, if a true foe had been recognized (named) as a friendly, no engagement would have occurred by the time the foe reached the minus 1,000-meter limit. Use of the adjusted range score which reflected recognition accuracy presumably would provide a more systems-oriented performance measure than either accuracy or range scores taken alone.

---

[1]The names of the aircraft were provided to observers during the field test for two reasons:

(1) It was assumed that air intelligence would provide such information to units assigned to a specific theater of operations. Perhaps a job aid listing the names would be provided.

(2) During the posttraining test, the names were not provided. A review of those test results indicated many instances in which response confusion seemed to occur; for example, a Fishbed consistently would be named Fishpot and *vice versa,* or an individual would consistently confuse A-4 and F-4 aircraft. (See Section 5 for a more detailed discussion on response confusion.)

Data Reduction. The average (mean) of the individuals' percentage correct scores was computed. In order to evaluate recognition range, the median range for the correct responses was computed for each observer. The arithmetic mean of these median ranges was then computed for describing the average field test results. (This median was used as an individual's score to minimize the effects of response skewness on the overall results.)

RESULTS

Average Scores. The average accuracy for the miniaturized field test was 82 percent (*SD* = 15%). The average range to the aircraft for correct recognitions was 1,268 meters (*SD* = 598 meters). The average range for the adjusted range score was 986 meters (*SD* = 705 meters.). It should be noted that three observers, who made 50 percent or more errors, contributed a majority of this increase in variance.

Intercorrelations. Table 9 presents the intercorrelations among posttraining scores and field test performance measures.

Table 9

Intercorrelations Among Posttraining and Field Test Measures

Variable Number	Measure	Field Test Measure 2	3	4
1	Posttraining Accuracy	.67*	.19	.42*
2	Field Test Accuracy		.12	.63*
3	Field Test Ranges			.69*
4	Adjusted Range Score			

**r* is statistically significant at $p = .05$ or less

Transfer Effects. The correlation between accuracy scores obtained on the static posttraining test and the dynamic test was 0.67. That is, 44 percent of the variance in the dynamic test scores was attributable to the posttraining proficiency levels. The average field test accuracy was 92 percent for those men who scored 70 percent or better on the static posttraining test. In contrast, the dynamic test accuracy was 73 percent for those who scored below 70 on the posttraining examination.

Posttraining accuracy and average recognition range was not significantly correlated; the correlation was low ($r = 0.19$). However, when the range scores were adjusted to include a penalty for an incorrect answer, the correlation with posttraining accuracy increased ($r = 0.42$; $p < .05$).

GOAR versus Printed Training. The average accuracies for (a) 53 men trained with the GOAR Kit and (b) 47 men trained with the printed imagery were 81.3 percent and 82.0 percent, respectively. The difference was not statistically significant.

The mean recognition ranges for the correct responses for the GOAR and printed methods were 1,239 meters and 1,302 meters, respectively. This difference was not significant.

For the adjusted range scores, the averages were 1,002 meters (GOAR) and 969 meters (printed), respectively. This was not significant.

DISCUSSION

The field test results support the hypothesis that VAR proficiency as measured by static imagery tests does transfer to a dynamic aircraft recognition environment. However, it must be noted that accuracy in the dynamic test was greater than that obtained for the classroom test. Several differences existed between the two types of testing situations which may have contributed to the difference in accuracy.

a. Fewer aircraft were presented in the dynamic test. Of the 25 aircraft included in training, only 13 were available for the field test. That is, the discrimination burden on the observer in the field test was about 50 percent of that required in the classroom test. It would be expected that a reduction in the discrimination requirement (or load) would enhance discrimination accuracy.

b. The observers were given a list of the aircraft names for the field test but had to depend upon their memory of aircraft names in the classroom tests. It would be expected that providing a job aid which was limited to the specific set of responses required would enhance performance. Informal observation during the classroom tests indicated that persons often were "positive" what aircraft was being shown, but could not recall its name. In other words, discrimination occurred but the response was not available.

c. The aspect angles in the dynamic test were limited to near head-on views, whereas the classroom test had a greater variety of views, including the relatively more difficult rear views. It would be expected that the absence of the more difficult-to-discriminate views would produce higher accuracy.

d. In the field test, each aircraft was presented several times at the same aspect angle, whereas the classroom test only presented each view once. Also, although the research staff did not provide any information to the observers concerning correct answers during the field test, many men did converse with one another between trials concerning the aircraft's correct name on the preceding trial. It is almost certain that this peer feedback produced learning during the series of trials, thereby increasing total accuracy for the field test.

The average range of the aircraft at the time of recognition was consistent with results obtained in a full-scale test conducted in 1961 by the US Army Human Engineering Laboratories (*6*). However, the average range was substantially less than that obtained in a full-scale test conducted by HumRRO in 1965 (*7*). This disparity may be attributable to the lower accuracy scores obtained in the miniaturized test described here (82%). In the 1965 HumRRO test, recognition accuracy was about 94 percent.

SECTION 4

SATURATION AND RETENTION

PURPOSE

The objective of the third phase of the project was to estimate the percentages of forward area crewmen who could learn to recognize 20, 40, 60, and 80 aircraft. A secondary objective was to obtain data concerning the retention abilities of these crewmen.

METHOD

Personnel. The 4th Battalion, 1st Air Defense Artillery (4/1), designated about 100 men to participate in this phase of the research. Twenty to thirty men were provided by each of the four batteries. An effort was made by each battery to assign men who would be "stabilized" for the duration of the data collection (originally planned to be about 4 months). Written guidelines for conducting the training were provided by the contractor (see Appendix C).

Training Materials. The printed version of the VAR training kit was used by all four batteries. This kit consisted of the multiview and single-image cards described earlier. In addition, individual test booklets for each group of four to six aircraft were assembled and bound. After a soldier had completed training on a four- or five-aircraft group, he was issued a test booklet and an answer sheet. The training NCOs were requested to record the amount of time each person worked with the training materials. These learning times were to be recorded on the man's answer sheet.

Self-Paced Training. As implied by the description of the training materials, it was desired to establish an individualized, self-paced learning environment, which would include self-administered achievement tests that would be scored immediately by the training NCO.

The 80 aircraft were partitioned into 4 blocks of 20. Each block was further divided into 4 groups of aircraft. Each block contained a mixture of attack, utility, rotary-wing, and multi-engine aircraft. The blocks were not experimentally equated for learning difficulty and the aircraft in each group were selected on the basis of the researchers' judgments of similarity.

Training Standards. Each battery was issued the training and test materials for one of the four blocks of twenty aircraft. When four or five men had completed training on all aircraft in a block, the battery gave an end-of-block test using the printed test booklets. If a man achieved at least 70 percent correct on this printed test, he was scheduled for an end-of-block test administered by the research staff. This test used slides from the GOAR training kit and included views not employed in training in addition to some of the training views. At the beginning of this phase of the research, the HumRRO-administered, end-of-block tests consisted of 100 items for each block. It quickly became apparent when persons were successively tested on three or four blocks of twenty aircraft (300-400 views) that this lengthy test was fatiguing many men. As a result of test reliability studies being performed at the time on Phase I data, it was found that a number of items per block could be reduced to 20 without degrading reliability of measurement. As a result, the end-of-block tests were correspondingly reduced in length, so that the final test covering 80 aircraft only had 80 views. In addition, not all aircraft in a block were included in the end-of-block tests given by HumRRO; aircraft views were selected to provide statistically equivalent parallel forms (equal means, standard deviations, and item-total score correlations).

The training standard established for passing the HumRRO test was also 70 percent correct for the aircraft studied just prior to the test.[1] Although the end-of-block tests were cumulative in that they sequentially covered 20, 40, 60, and 80 aircraft, the men's retention scores for the test items for earlier learning blocks were not included in determining whether a person would proceed to a new block of aircraft. Individuals were allowed three attempts to meet the 70 percent standard for each block. If a person failed on all three attempts, he was "socially promoted;" *i.e.*, allowed to proceed in the program. Social promotions were employed to discourage soldiers from malingering. If individuals had been dropped from the program because they failed to meet the standards, it was felt that many men would pursue that course of action and the research results would underestimate the learning abilities of typical crewmen.

Independent Variables. Data was collected on two classes of variables which had been hypothesized to influence saturation and retention levels: (a) aptitude measures obtained from the Army Classification Battery: and (b) two measures of time spent in training: (1) elapsed time since training began and (2) total time spent with training materials as recorded on the achievement test answer sheets.

RESULTS AND DISCUSSION

Unit Variations. Table 10 presents a tabular summary of the overall results of the saturation study. This table indicates the number of men originally assigned to the project by each of the four batteries[2] (Column 2) and the number of these men who were still available for training upon termination of data collection on 1 May 1976 (Column 3). Column 4 indicates the frequency of failing an end-of-block test and Column 5 indicates the number of testing sessions conducted by the contractor. Column 6 presents the number of men who learned 80 aircraft. The last column (No. 7) presents the percentage of training commitments accomplished. The latter was computed as

[1]The 70 percent criterion was not arbitrarily established. It had been found in Phase I of the research that persons who achieved 70 percent or better on the end-of-training GOAR classroom test, had VAR proficiency of 90 percent or better in the miniaturized field test.

[2]Battery designations are coded to provide anonymity.

Table 10

General Results of Saturation Study

1 Battery	2 Number Assigned to Program	3 Number Remaining in Program	4 Number of Test Failures	5 Number of Block Testing Sessions Accomplished	6 Number Completing All 4 Blocks	7 Percentage of Total Training/Testing Program Completed
01	31	21	18	14	21	100%
02	20	16	9	13	16	100%
03	19	19	11	3	0	55%
04	30	27	30	5	3	52%
TOTAL	100	83	68	35	40	74%

equal to the ratio of the sum of products of the number of aircraft learned X number of persons achieving each level, divided by the product of the total aircraft to be learned (80) X the number of persons available for training:

$$e.g., \ 100 \times \frac{(5\text{x}20+4\text{x}40+3\text{x}60)}{12\text{x}80} = 45.8\%.$$

Two of the batteries completed the total training commitment, whereas the other batteries achieved a little over 50 percent of the training goals.

The general results reflect variations that existed among the four batteries with respect to several factors:

1. The batteries that achieved 100 percent of the training goals employed the desired self-paced, individualized approach to instruction. This is reflected in the number of contractor-administered tests that were given to each of these batteries. Whereas, Batteries 01 and 02 requested tests when four or five men had completed a block, the other batteries employed a more-or-less lock-step approach and did not request testing until all (Battery 03) or most (Battery 04) were judged to be prepared. Battery 03 definitely employed a lock-step training management approach. No person in that battery was tested on 80 aircraft and only 3 testing sessions were given over a 5-month time period.

2. In contrast, Battery 04 tended to request contractor testing irrespective of the learning status of its personnel. Although 3 men from that battery did learn 80 aircraft, the largest number of test failures (30) also occurred in that battery.

3. It is not too surprising that units having common missions vary with respect to command emphasis and training approaches. However, it is very significant that such variations can have a powerful effect on the differential achievements of members of those units.

Learning Rate. Figures 4 - 7 present the learning (acquisition) curves for each battery. Each figure displays the cumulative percentage of battery personnel who achieved each subgoal (20, 40, 60, & 80 aircraft) as the training program progressed in time. For example in Figure 4, it can be seen that about 50 percent of the men had learned 60 aircraft by 10 March, about 17 calendar days after the first testing session was conducted. During that same time interval, about 95 percent of the men had learned 20 aircraft.

Figure 5 has a pattern similar to Figure 4, except that Battery 02 was, at first, somewhat slower in qualifying its men and then accelerated training to achieve the training goals. This acceleration is particularly evident in the acquisition curve for 80 aircraft. Essentially all 16 men learned the final block of aircraft in a 14-day time period.

Comparison of Figures 4 and 5 with Figures 6 and 7 reflects the apparent differences in command emphasis. The acquisition rates are very low for Battery 03, and as indicated, this battery terminated participation in the project without attempting the final block of training. Very low acquisition rates also characterized the fourth battery, but learning rates were just beginning to accelerate when it too had to drop out of the program.

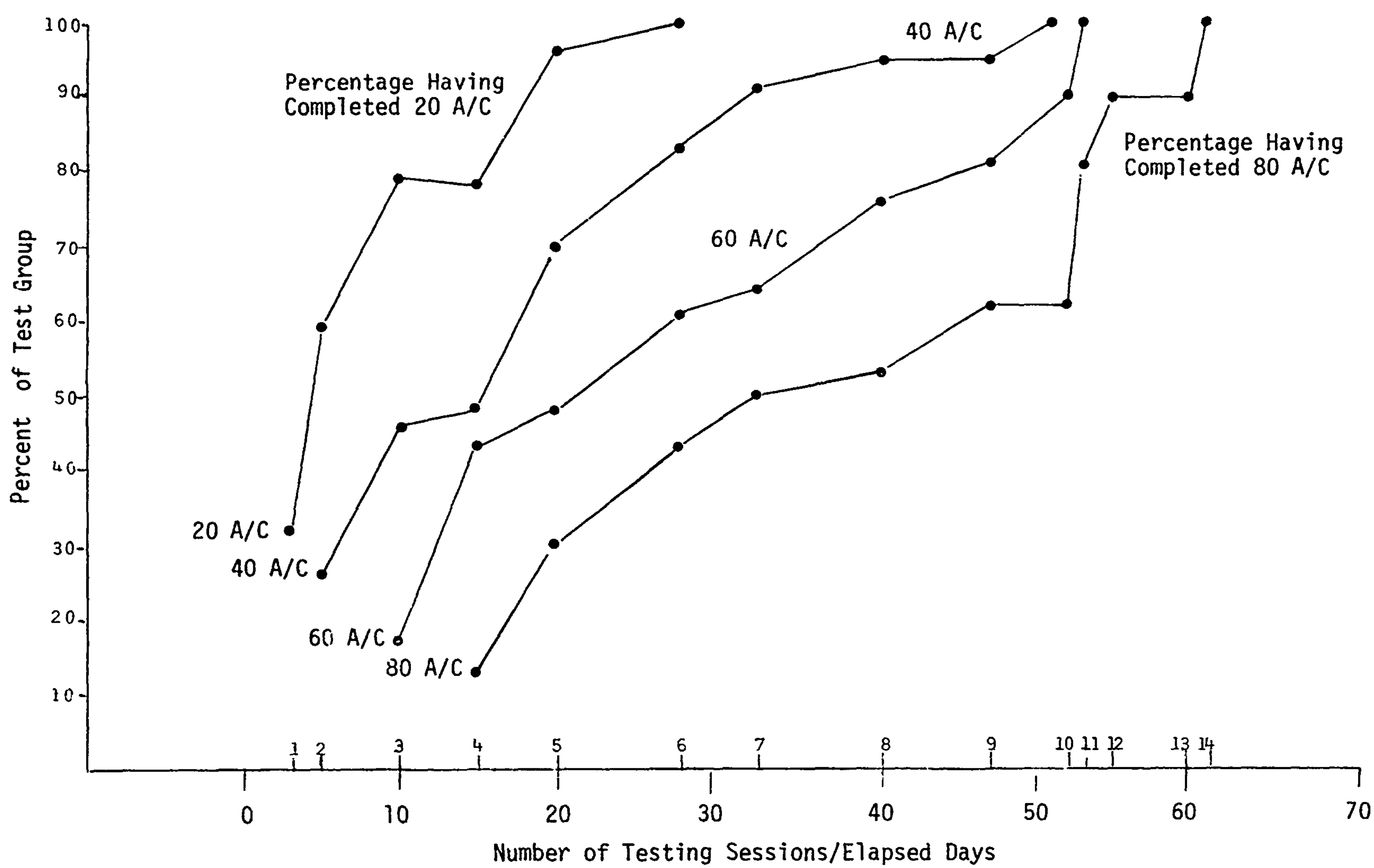


Figure 4. Acquisition Curves for Battery 01

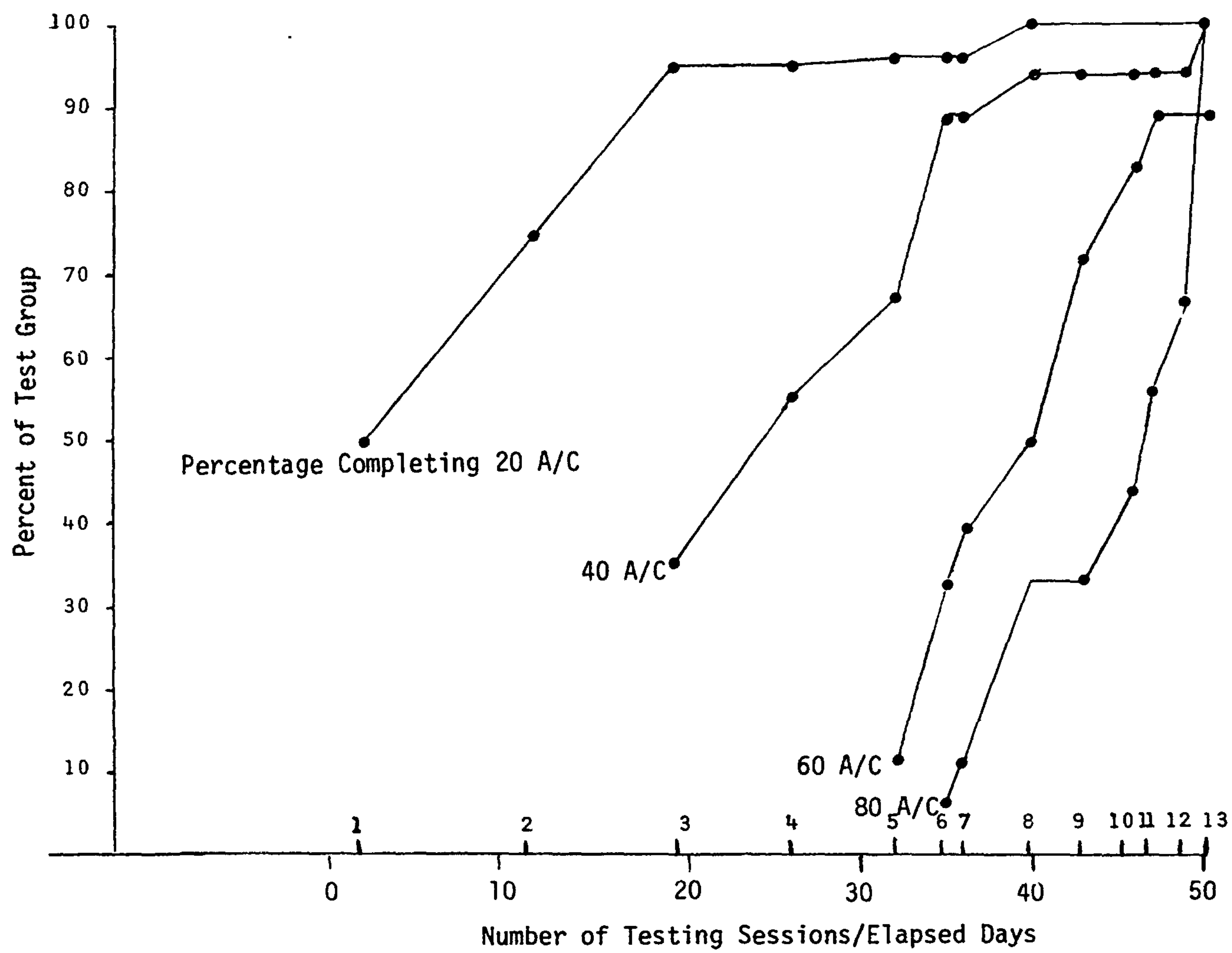


Figure 5. Acquisition Curves for Battery 02.

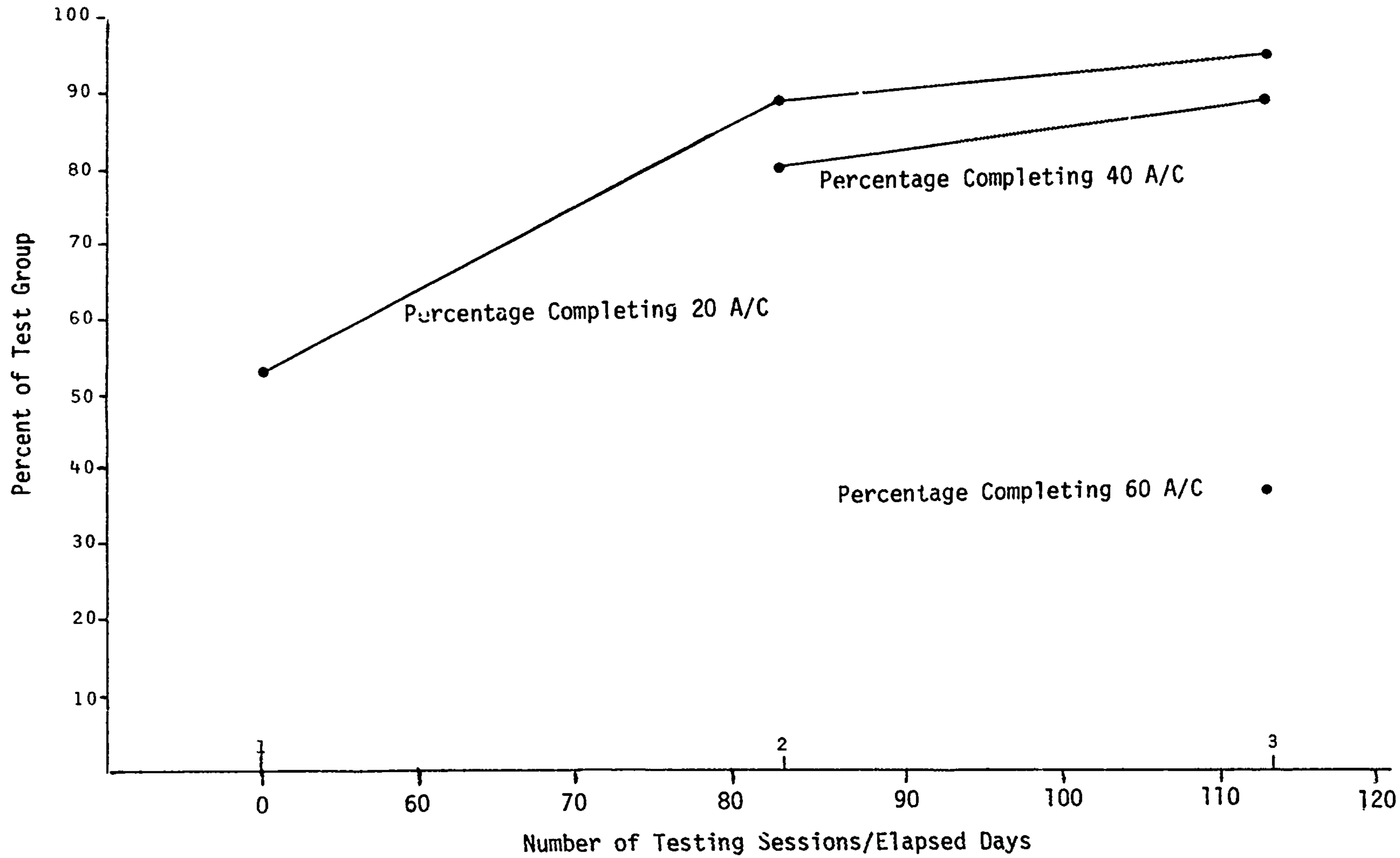


Figure 6. Acquisition Curves for Battery 03

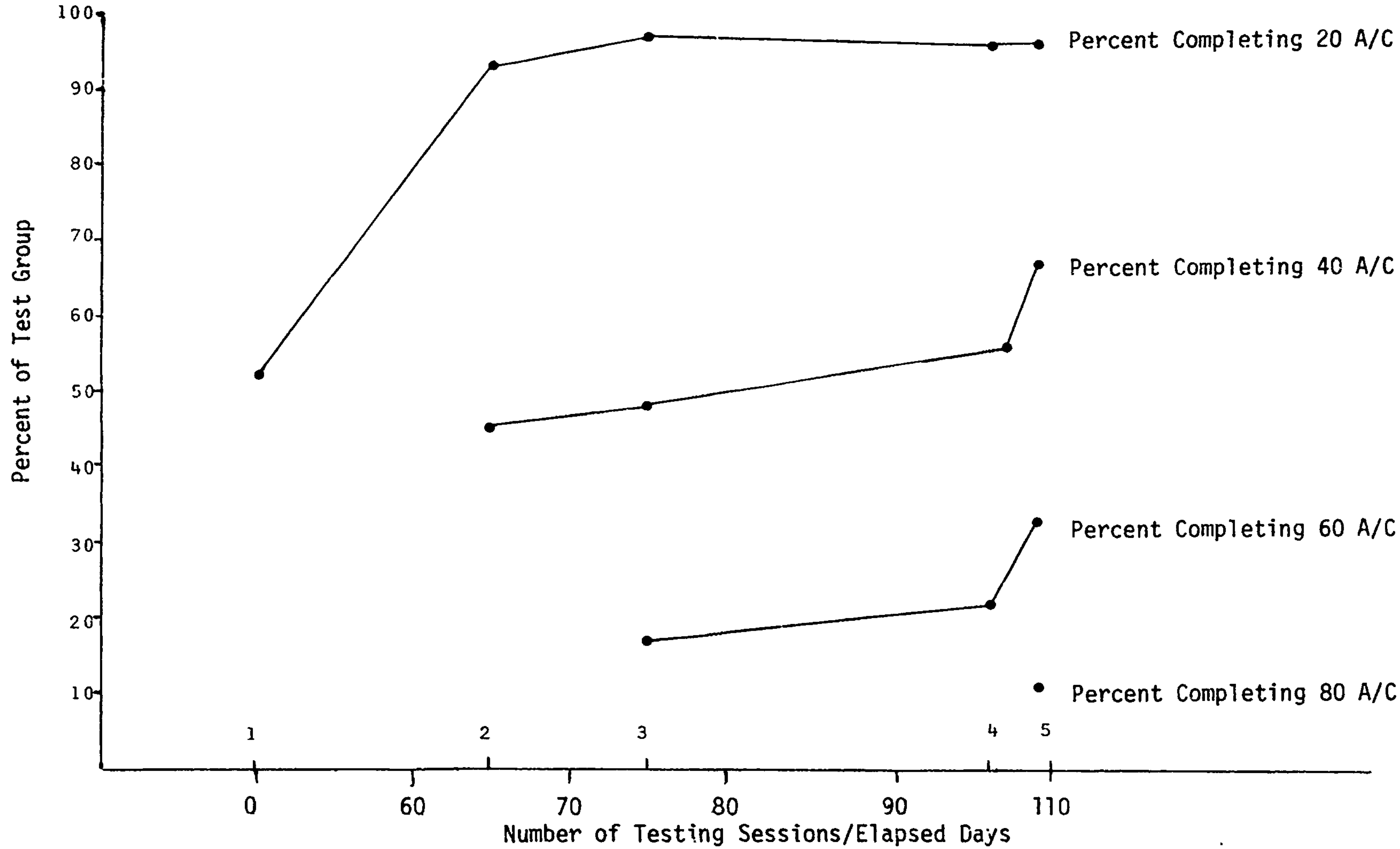


Figure 7. Acquisition Curves for Battery 04

One of the items of interest in the VAR research project was to estimate the number of training sessions required to learn the 80 aircraft. The best training management data was obtained from Battery 02 which held a total of 29 training sessions. These sessions also included the 13 contractor testing sessions. Each session involved about 11 students and two administrators. It was found that, on the average, 15 training/testing sessions were required to complete the 80 aircraft ($\sigma = 1.83$).

Records were kept of the amount of time that each man spent training on each block of aircraft. Battery 02 spent, on the average, 113, 78, 81, and 80 minutes, respectively, on the first, second, third, and fourth blocks of aircraft. The total average time spent learning the 80 aircraft was 360 minutes ($\sigma = 115$) or about 5.14 minutes/aircraft. This does not include the amount of time spent taking the tests. As a comparison, Battery 01 spent 85, 78, 79, and 85 minutes or a total of 327 minutes learning the 80 aircraft (4.1 minutes/aircraft). These differences were not significant.

The difference between batteries in recognition accuracy was not affected by either differences in average aptitude (GT) or the amount of training time expended.[1] However, a significant difference between the number of days needed to accomplish training and recognition accuracy did occur ($p = .0005$), as indicated in Figure 8. This indicates the program established in Phase III trained unit personnel to approximately the same level of proficiency regardless of aptitude but training management variations produced differences in the number of days required to reach a given level of proficiency.

Retention. One of the secondary objectives of this phase of the research was to obtain estimates of the retention of VAR training. Since the contractor-administered tests were cumulative over successive aircraft blocks, it was possible to compare retention scores for VAR retests versus the initial test data. This evaluation was performed for all comparisons of VAR retest scores for all available data. Batteries 01 and 02 had three retests during a 10-week period. Batteries 03 and 04 were given two retests on the initial block during that interval. Table 11 presents the average scores for each battery for the initial administration and the subsequent retesting on the first block of 20 aircraft.

When averaged over all four batteries, accuracy declined between four and five percentage points from the original testing on Block 1 to each of the retests. The largest decline in accuracy occurred in Battery 03; a 14-point decrease from the initial average of 77.5 percent to an average of 63.5 percent on the second or last retest.

[1]Training times were not available for Batteries 03 or 04, making comparisons between training methodologies invalid.

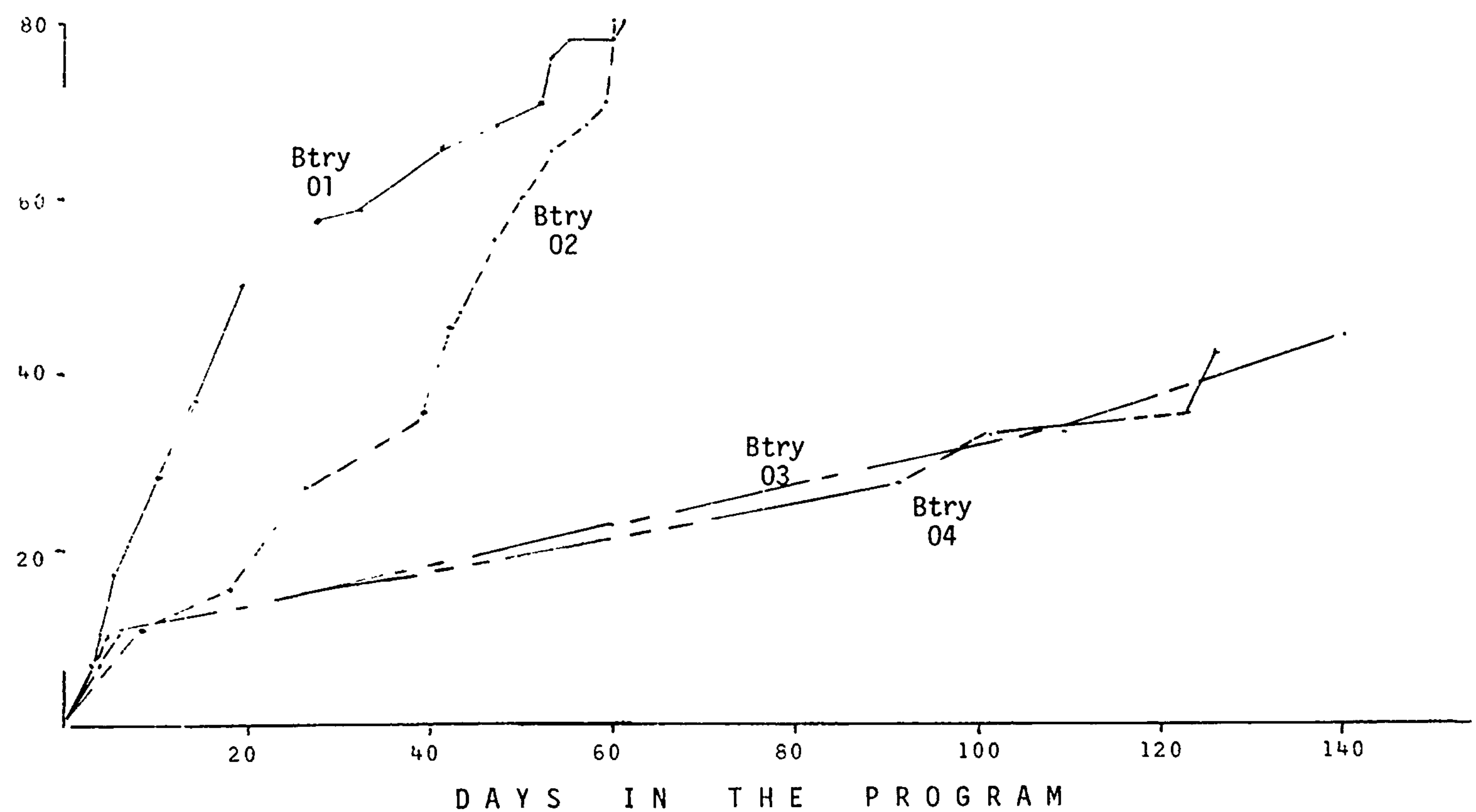


Figure 8. Program Completion in Terms of Mean Number of Aircraft Trained.

Table 11

Retention Data
for the
Initial Block of Aircraft

BATTERY		FIRST TESTING	SECOND TESTING	THIRD TESTING	FOURTH TESTING
01	$\overline{X}$%	87.4	88.4	86.2	82.6
	σ	10.8	14.1	12.8	14.3
	n	21.0	21.0	20.0	12.0
02	$\overline{X}$%	82.8	86.6	77.5	77.8
	σ	12.4	15.2	16.7	19.1
	n	16.0	16.0	16.0	16.0
03	$\overline{X}$%	77.5	63.1	63.5	85.7
	σ	11.3	23.2	21.7	11.0
	n	16.0	16.0	10.0	3.0
04	$\overline{X}$%	86.4	77.2	83.6	80.4
	σ	10.0	16.6	16.5	16.8
	n	22.0	22.0	11.0	31.0
OVERALL	$\overline{X}$	84.0	79.3	79.1	79.9
	σ	11.7	19.8	18.4	17.4
	n	75.0	75.0	56.0	28.0

Data Analysis. The retention scores were statistically analyzed in accordance with linear regression techniques. The initial analyses were designed to evaluate differences among the retention scores for the four batteries as a function of the following factors:

a. Variation in the average amount of training time for individuals in each battery;

b. Variation in the amount of elapsed time (calendar days) that each battery required to complete the program.

c. Differences among the average GT scores for the batteries.

Overall Losses. Initially, an analysis was performed for all available data (Blocks 1 through 3) to determine if the retention losses for all batteries were equal. This analysis indicated that there were significant differences in the losses experienced by the batteries ($F = 2.85$; $p = .02$). Supplementary analyses were performed in which the retention losses for Battery 01 were compared with those for each of the other batteries. The results were as follows: for 01 versus 02--$F = 1.18$, $p = .32$; for 01 versus 03--$F = 4.74$, $p = .001$; for 01 versus 04--$F = 0.35$, $p = .88$. These results indicated that Battery 03 had significantly greater losses than Battery 01 (and by inference, greater losses than occurred in either Battery 02 or 04).

Separate linear regression analyses were made for each of the first three blocks of aircraft. This analysis indicated that the retention losses for the first block were not statistically significant for Batteries 01, 02, but the loss for Battery 03 was significant ($F = 6.21$; $p = .01$).

For the second block of aircraft there were no significant losses for any of the batteries; in fact the retest scores on Block 2 showed a significant gain for Battery 02 ($F = 6.68$; $p = .01$).

In contrast, for the third block of aircraft, Battery 02 showed a significant loss in accuracy ($F = 4.46$; $p = .04$). The losses for Batteries 01 and 04 were not significant (no retest data could be obtained for Battery 03 on Block 3).

Loss as a Function of Time. Records of the amount of time devoted to training were available for Batteries 01 and 02. Data concerning the total number of calendar days (elapsed time) required to complete (or suspend) the training program were available for all batteries. Linear regression analyses were performed to ascertain if each of these temporal factors was related to the retention losses that occurred (irrespective of the significance or lack of significance of the amount of loss).

When retention data for all batteries and all blocks were combined, it was found that the variation in retention losses among batteries was significantly and negatively correlated with both the amount of training time used and the amount of elapsed time. These results are interpreted to indicate that the persons who demonstrated larger retention losses required longer to complete the training in terms of elapsed time and spent more time in training; *i.e.*, they were poor and slow learners. This conclusion is supported by the additional finding that the amount of time spend in training was negatively correlated with GT scores ($r = -0.29$; $p < .05$).

SECTION 5

PREDICTING VAR PROFICIENCY

SCOPE

This part of the report discusses a number of factors concerned with the general problems of measuring and predicting VAR proficiency. Such factors as item difficulty, test reliability, and aspects of validity are discussed. In addition to the latter factors, this section presents analyses of VAR proficiency for familiar versus unfamiliar views. Analyses also are presented concerning response confusion; *i.e.*, analyses were made to determine which aircraft are most often confused. The section also discusses various methods of selecting men for VAR training or predicting those personnel who would provide the best VAR capabilities to a unit.

TEST RELIABILITY

It will be recalled that the initial phase of the research was concerned with a comparison of two methods of training visual recognition for a group of 25 aircraft. The proficiency test contained 250 items and was administered in 2 parts. The test had been designed so that each half of the test contained an equal number of views of each aircraft. Although different views were shown of each aircraft in each half, the number of familiar (*i.e.*, included in training) and unfamiliar views were equal for the two parts. Although the test parts were designed to be parallel, it was not possible to examine these characteristics before conducting the training programs.

The intercorrelation between the 132 pairs of scores obtained on the two parts of the posttraining test was computed. The correlation coefficient equalled .964, which when corrected for test length yielded an overall reliability of .999 for the 250-item test. The average (mean) and standard deviation (*SD*) for the two parts were 54.2 percent (*SD* = 25.2) and 57.8 percent (*SD* = 25.3). Slightly higher scores were obtained on the second part of the test, probably due to learning or increased familiarity. However, it may be concluded that a 250-item test wasn't required to evaluate a person's VAR proficiency.

REDUCING TEST LENGTH

The preceding analysis indicated that a 125-item test was just as useful as a 250-item test for evaluating VAR proficiency. Since the Army typically uses even shorter tests for evaluating this skill in field exercises, ARTEPs, and the soon-to-be-introduced SQTs, it was considered desirable to determine the effect that test length has on measurement error (*i.e.*, test reliability).

This information also was needed in order to minimize the amount (duration) of testing that had to be conducted in that portion of the research program which evaluated saturation and retention. Specifically, there was a need to determine if measurement errors were not excessive when only a 20 to 25-item test was used to evaluate learning of 20 - 25 aircraft. To create such tests, items in Phase I were ordered by level of difficulty. Two ten-percent samples (25 items) were selected from the 250 items and matched by level of difficulty. Linear regression between the 25-item versions and the full test (250 items) was 0.93 for one and 0.96 for the other. The reliability of the two tests was 0.90. These values were sufficiently high to justify using the short versions in Phase III.

PROFICIENCY FOR FAMILIAR VERSUS UNFAMILIAR VIEWS

In this research program, the observers were trained on 5 of a set of 10 views available for testing. In the first phase of the research, the men were tested on all 10 views, 5 of which had been included in the training programs. In the third phase of the program, which was concerned with saturation, some of the men were tested on all 10 views, but the majority were tested on a sample of these views.

Phase I Analysis. The scores (percentage correct) obtained on the familiar and unfamiliar views were compared by an analysis of variance method which also included the training method as an independent variable. The mean scores obtained by the total sample of 132 men are presented in Table 12. The difference between the average accuracy levels for the familiar and unfamiliar views was not significant ($p = .87$).

Table 12

VAR Accuracy for Familiar and Unfamiliar Views: GOAR versus Printed (Mean Percent Correct)

	GOAR	Printed	Average
Familiar	52.7%	59.7%	56.2%
Unfamiliar	52.5%	58.8%	55.7%
N	65	67	132

Phase III Analyses. The scores obtained by the observers in the saturation study were also analyzed. For a total of 39 observers, the average accuracy for the familiar views was 90 percent, whereas the accuracy for unfamiliar views was 94.6 percent. Although this difference was small numerically, statistically it was highly significant ($F = 44.35$; $p < .0001$).

The results of the analyses of Phase I and Phase III error rates were not in agreement. In Phase I, the error rates for familiar and unfamiliar views were equal. However, in Phase III, of a total of 872 errors made, 40 percent occurred on familiar views and 60 percent occurred for unfamiliar views, a difference of 20 percent. These different results are believed to be attributable to differences in VAR proficiency levels of the observers in the two studies. The average VAR proficiency of the 132 observers in Phase I was 59 percent, whereas,the average VAR score for the 19 observers in Phase III was 87 percent. The results suggest that when learning levels are relatively low, error rates are about equal on familiar and unfamiliar views. In contrast, when proficiency levels are relatively high the majority of the errors occur on unfamiliar views.

Accuracy as a Function of View. Anyone who has received or given VAR training is aware that recognition accuracy is more difficult for some aircraft views than others. In designing an optimum VAR training program, it would be desirable to place more emphasis on recognizing the more-difficult-to-discriminate views and less emphasis on the "easiest" views.

Data concerning relative accuracy in relation to view was obtained as part of this research program. Although such comparisons could be made for any of the tests given, it was believed by the research staff that the data obtained on the initial pretest given to 200 men would have the greatest validity. Since the man had not participated in the training methods comparison, presumably they would be equally familiar with all views, not just the subset of five included in the contractor-administered training programs. Table 13 presents the recognition accuracy for each of the 10 views of the 25 aircraft in the pretest.

Table 13

Recognition Accuracy for Ten Aircraft Views

View	Percent Correct	Rank Order*
0-0	55.4	9
0-15	62.8	4
0-35	63.6	1
0-90	61.4	5
45-15	60.9	6
90-0	58.4	7
170-0	51.8	10
190-15	56.0	8
315-35	63.5	2
340-15	63.4	3
OVERALL	59.7	

*1 = least difficult

In general, the least difficult images tend to be views which show the aircraft as incoming and over the shoulder. The most difficult views are head-on or flying away.

The Relative Difficulty of Various Aircraft. Data concerning recognition accuracy was analyzed for the group of 25 aircraft used to compare training methods.[1] Since this set of 25 constituted a sample of common NATO and Soviet aircraft, this data should be of use to designers and managers of VAR training programs. The recognition accuracy scores of 205 men on the pretest and posttest are presented in Table 14.

Table 14

Recognition Accuracy Scores for Twenty-Five Aircraft

Aircraft	Pretest Percent Correct (N = 205)	Posttest Percent Correct (N = 132)
A-7	59	59
Fiat	54	44
F-111	64	56
F-5	67	65
SU-11	37	30
A-4	44	44
F-4	72	71
Mirage	46	36
P1-B	44	31
Mig-21	41	39
SU-9	44	29
Mig-19	59	54
SU-7	52	44
OV-10	80	82
OV-1	63	70
U-21	70	58
Yak-28	67	64
UH-1	70	64
OH-58	55	56
OH-6	66	61
AH-1	85	86
MI-4	59	64
MI-6	53	54
MI-8	50	50
Chinook	92	93

[1]Although accuracy data was available for all 80 aircraft, analysis of all of this data was beyond the objectives of the contractual effort.

The highest recognition accuracy occurred for such aircraft as the Chinook, the Cobra (AH-1), the Bronco (OV-10), and the Phantom (F-4), which are often seen in the Fort Bliss area. In general, the lowest recognition accuracies occurred for the Soviet attack aircraft during both tests.

It should be noted that the overall test average was lower for the posttest than the pretest in this comparison because only the two batteries having the lowest average pretest scores took the posttraining test. The two batteries having the highest average pretest scores did not participate in the training experiment.

Response Confusion. The posttraining test data for Phase I (GOAR *vs.* Printed) were analyzed to determine which of the 25 aircraft were most often confused with one another. That is, if a person gave an incorrect response, what was that erroneous response most likely to be? Table 15 presents the results of that analysis. The table lists each aircraft and indicates the error rates for the two or three most common confusions within the set of

Table 15

Response Confusions for Twenty-Five Aircraft

Correct Aircraft	Most Common Errors 1st	2d	3d
Mohawk	Bronco (6%)	UTE (5%)	--
Bronco	Mohawk (6%)	--	--
A-4	Mirage (7%)	F-4 (5%)	F-5 (5%)
F-4	F-5 (2%)	F-111 (2%)	SU-11 (2%)
F-5	F-4 (3%)	A-4 (2%)	F-111 (2%)
F-111	F-5 (4%)	F-4 (3%)	SU-11 (3%)
A-7	Fiat (4%)	SU-7 (3%)	Mig-19 (2%)
P1-B	A-7 (5%)	Fiat (5%)	SU-9 (4%)
Mirage	A-4 (8%)	Mig-21 (5%)	F-5 (4%)
Mig-19	SU-7 (7%)	Mig-21 (5%)	SU-9 (4%)
SU-9	Mig-21 (16%)	SU-7 (15%)	Mig-19 (8%)
Mig-21	SU-9 (6%)	Mig-19 (5%)	SU 11 (3%)
Yak-28	Mig-21 (2%)	P1-B (2%)	--
SU-11	Yak-28 (6%)	F-111 (5%)	F-5 (4%)
SU-7	SU-9 (10%)	Mig-19 (9%)	Mig-21 (4%)
Ute	Mohawk (9%)	Bronco (2%)	A-4 (2%)
Fiat	A-7 (16%)	Mig-19 (3%)	Mig-21 (2%)
OH-58	UH-1 (14%)	OH-6 (8%)	--
CH-47	--	--	--
AH-1	UH-1 (2%)	MI-6 (2%)	--
UH-1	OH-58 (12%)	OH-6 (3%)	--
OH-6	OH-58 (8%)	UH-1 (4%)	MI-8 (2%)
MI-4	MI-8 (10%)	MI-6 (5%)	AH-1 (2%)
MI-8	MI-4 (16%)	MI-6 (8%)	--
MI-6	MI-8 (10%)	AH-1 (7%)	MI-4 (6%)

25 aircraft. These error rates varied from 16 percent to 2 percent. The highest error rates tended to occur among the Soviet aircraft; *e.g.*, SU-7 was confused with SU-9 or MIG 19, and the Soviet helicopters were confused with one another.

A tally also was made of the frequency with which aircraft were confused with other aircraft from the same or different national origin. This tabulation is presented in Table 16.

Table 16

Frequencies of Confusions
for
Aircraft of Different National Origins

Aircraft Shown	National Origin of Error		
	US	NATO	Soviet
US	12	1	7
NATO	3	1	3
Soviet	3	1	9

From Table 16 it can be seen that US aircraft were confused with Soviet aircraft about half as frequently as they were confused with each other. Similarly, if a Soviet aircraft was incorrectly recognized it was most often confused with another Soviet aircraft.

ACB Versus VAR Proficiency. The results of this research indicate a moderate relationship between VAR proficiency and several of the aptitude area scores of the Army Classification Battery. In particular, Aptitude Area CO had the highest intercorrelation with VAR proficiency of all the area scores, perhaps due to the inclusion of the Pattern Analysis subtest in the CO measure.

These results suggest that VAR training is to be given to only some of the personnel in an air defense unit, then an examination of the CO scores of the available personnel would provide a basis for selecting the persons to receive such training. Such a procedure would tend to increase the overall cost-effectiveness of the VAR training given by units by reducing the training burden on instructors and relieving men who possess low VAR learning potential for assignment to other mission-oriented duties.

Prediction Using a Screening Test. In the initial phase of the research, a very high intercorrelation (0.88) was obtained between the scores obtained on pre- and posttraining tests. In addition, a moderately high correlation (0.67) was obtained between scores for the posttraining classroom test and the dynamic test. These findings raised the possibility of using the results from a short VAR screening test to select persons to participate in more extensive training. Two sets of data were available to examine the feasibility of using a pretraining measure to predict VAR performance.

The first set required scoring the first 25 items of a more lengthy test given early in the saturation study. Such data were available for 31 of the men who completed all four blocks of the training program. The intercorrelation between the score on the 25-item test and the average score obtained for the first administration of each of the four block tests was 0.32 ($t = 1.72; p > .05$). Although the mean score increased significantly ($t = 5.01$) from 67 percent on the 25-item test to 84 percent for the average score for 80 aircraft, the scores were not intercorrelated sufficiently to predict final proficiency with much accuracy.

The second set of data consisted of pretest scores obtained in April 1975 and the average of scores obtained for the blocks attempted in the 1976 saturation study. Such data were available for only 21 of the men tested. The intercorrelation between these pairs of scores was 0.34 which also was not statistically reliable ($t = 1.60; p > .05$) although the increment in average proficiency from 68.5 percent to 81.4 percent was ($t = 3.27; p < .01$).

Even though both analyses yielded nonsignificant correlations, it is believed that the results are inconclusive for two reasons. First, the average scores on the pretest and the short screening test are higher than the average pretest score obtained for 205 men in April 1975. Second, and more importantly, the average score on the saturation test was very high (87%) and the variability in scores was very low ($SD = 9.67$) Such a restricted variance greatly reduced the likelihood of obtaining a high intercorrelation.

Additional analyses were made to determine if performance on the 25-item screening test was correlated with the amount of training time required to complete the four blocks of instruction. Two analyses were performed: (1) the VAR score was correlated with the number of calendar days that were required to learn all four blocks, and (2) the VAR score was correlated with the total amount of time each person had possession of training materials.

1. *Elapsed time.* A correlation of -0.52 was obtained between the screening test scores and the number of days that persons were in training. This correlation was statistically reliable ($t = 3.29; p < .01$).

2. *Training time.* The correlation between VAR score and the total amount of time spent in training was -0.40 ($t = 2.22; p < .05$).

Each of these analyses indicated that VAR proficiency was negatively correlated with the duration of training required to complete the program. Persons with high scores on the VAR screening test involving 20 aircraft took less time per training session and required fewer training sessions than needed by persons having low scores on the screening test. It was of interest to observe that the amount of elapsed time varied from a low of 10 to a high of 57 days for the 31 men included in this analysis. Training time varied from 125 to 735 minutes. These results reflect the influence of individual differences in learning ability and provide a striking illustration of the training flexibility and savings that can be achieved through the use of individualized, self-paced instructional methods.

SECTION 6

OBSERVATIONS ON TRAINING MANAGEMENT AND COST/EFFECTIVENESS

This study provided a relatively rare opportunity for research personnel to monitor the conduct of training in a tactical (TOE) organization. Whereas the vast majority of training research has been conducted in well controlled laboratory-type environments, very few attempts have been made to conduct such research in an environment that is largely beyond the control of research personnel. The VAR research conducted with the support of the 4th Bn, 1st ADA, provided a test bed which can be characterized as possessing many factors which could not be controlled or influenced by the project staff.

In the initial phase of the research, which compared the relative effectiveness of two training methods, the training and testing had to be performed on "a-not-to-interfere" basis. This admonition meant that any and all other mission and training requirements had a priority that was equal to, if not greater than, the requirement to conduct VAR training. As a result of this restriction, it was found that more than half the men in the two training programs attended less than half of the training sessions. The results obtained for the posttraining test reflected the impact of this fact: proficiency as a result of training increased only 10 percent over the level that existed before training began. Very little additional VAR skill was acquired through the additional training and the primary determiner of posttraining skill was the individual's skill level before the new training began. It was not surprising that people who don't attend training learn very little. What is disturbing is that such programs are clearly not cost effective, either in terms of the learner's time or the cost of providing the instruction.

The institutional environment that initially existed in the research phase concerning saturation and retention also was characterized by considerable absenteeism. This condition was quickly remedied by the battalion commander when the absenteeism problem was brought to his attention. At that time, the battalion decided to designate for VAR training only a portion of the men from each battery and to require these controlled personnel to receive VAR training on a regular basis. The results of this change in battalion policy was reflected in the training achievements made by a majority of the controlled personnel. VAR accuracy on successive tests was very high and the vast majority of men still available for training at the end of the program had achieved the goal of learning 80 aircraft. This achievement level was actually beyond the expectations of the operational, doctrinal, and research staffs.

Command emphasis at the battalion level was placed on accomplishing VAR training. However, current policies concerning decentralization of training responsibility to the company/battery level existed and the results of this policy were reflected in the variability that occurred in the training accomplished by the various batteries.

In Battery 01, the commander and the training officers accepted the commitment to accomplish the training goals in accordance with the instructional guidance and policies prescribed by the research staff. Several training sessions were scheduled by the training officer and the controlled personnel were required to attend at least one of these sessions each week. A concerted effort was made to conduct reviews of the group of aircraft in each block. The training was individualized and self-paced. In some instances the learners paired-off and conducted their learning/practice in their quarters. Towards the end of the program, training was being conducted during after-duty hours to eliminate a recurrence of one unfortunate instance when an inspecting officer made a repremanding observation that formal classroom training was not being given in accordance with the official battery training schedule. The battery officers had been reprimanded for not using lock-step instruction. The achievements of Battery 01 were exceptional and the men completed the program goals ahead of schedule.

In Battery 02, the commander did not have confidence in the program objectives and had little regard for individualized instructional concepts. Although the commander's support was absent, the training officer accepted the requirement and sought to follow the instructional guidance. This battery got off to a slow start, in part due to a lack of command emphasis. The battery training officer, however, began to place a strong emphasis on VAR training, including scheduling training sessions after normal duty hours and on week ends. The personnel in this battery rapidly began achieving program goals and most of the men reached the final criterion.

In Battery 03, the commander and the training officer did not follow the instructional guidance. Lock-step instruction occurred and only three official VAR tests were scheduled over the four-month period. Reviews apparently were not scheduled and retention scores were the lowest of all four batteries. No useable time records were kept and no one achieved the program objectives.

In Battery 04, training was sporadic. Although three men completed the whole program, test failures were the highest of all batteries. Many retests had to be scheduled because the personnel had not learned the aircraft. A considerable amount of training and testing time was wasted as a result.

LITERATURE CITED

1. Baldwin, Robert D. *Capabilities of Ground Observers to Locate, Recognize, and Estimate Distance of Low-Flying Aircraft.* Alexandria, VA: Human Resources Research Organization. March 1973. HumRRO Technical Report 73-8.

2. Whitmore, Paul G., Rankin, William C., Baldwin, Robert D., and Garcia, Sandra. *Studies of Aircraft Recognition.* Alexandria, VA: Human Resources Research Organization, February 1972. HumRRO Technical Report 72-5.

3. Whitmore, Paul G., Cox, John A., and Friel, Donald J. *A Classroom Method of Training Aircraft Recognition.* Alexandria, VA: Human Resources Research Organization, January 1968. HumRRO Technical Report 68-1.

4. Miller, Elmo E. and Vicory, Arthur C. *Comparison and Evaluation of Printed Programs for Aircraft Recognition.* Alexandria, VA: Human Resources Research Organization, October 1971. HumRRO Technical Report 71-22.

5. Baldwin, Robert D., Frederickson, Edward W., and Hackerson, Edward C. *Aircraft Recognition Performance of Crew Chiefs With and Without Forward Observers.* Alexandria, VA: Human Resources Research Organization, August 1970. HumRRO Technical Report 70-12.

6. Wokoun, William. *Detection of Random Low Altitude Jet Aircraft by Ground Observers.* Aberdeen Proving Ground, MD: US Army Ordnance Human Engineering Laboratories, June 1960. Technical Memorandum 7-60.

7. Wright, A. Dean. *The Performance of Ground Observers in Detecting, Recognizing, and Estimating Range to Low-Altitude Aircraft.* Alexandria, VA: Human Resources Research Organization, December 1966. HumRRO Technical Report 66-19.

APPENDICES

Appendix | Page

A PROCEDURE USED FOR GOAR KIT TRAINING ... A-1

B PROCEDURE USED FOR PRINTED IMAGERY TRAINING ... B-1

C INSTRUCTOR GUIDELINES FOR PHASE III: SATURATION TRAINING ... C-1

APPENDIX A

PROCEDURE USED FOR GOAR KIT TRAINING

Members of X Battery, 4/1 ADA, were trained in aircraft recognition skills using slides from the GOAR Kit. Training was begun on 19 June 1975 and final review was conducted during the week of 25 August 1975. The training sessions were to be scheduled so that 1st Platoon was to train on Tuesday afternoons, 2d Platoon on Wednesday afternoons, and 3d Platoon on Thursday afternoons. This schedule was followed for the most part except for about four sessions in which two platoons had to be scheduled at the same time. These sessions resulted in classes that were too large for best training results. Classes were usually held in a conference room in X Battery which could be darkened for the slide training. Several classes had to be held in the dayroom which was not similarly equipped for training.

The procedure used in all GOAR slide training sessions followed the outline listed below.

1. Attendance Rosters. Attendance rosters were passed out at the beginning of each training session to obtain the names and platoons of those persons attending the session.

2. Three-View Familiarization Slides. Each training session trained four to five new aircraft The aircraft were introduced to the students by showing a three-view slide (front, side, & bottom views) of

each aircraft. The instructor stressed the name and number and country of origin of each aircraft. Distinguishing features were pointed out such as wing design and location, the number, shape, and location of intakes or exhausts, etc. Any features that could be related to an aircraft's purpose or name were commented upon. For example, the Soviet HIP helicopter has a bulge low on each side of the fuselage that could be called "hips."

3. Paired Comparison Familiarization. In addition to the three-view slides, five different views of each aircraft were used for all training and review sessions. These views were, in terms of their heading angle and climb angle: 0°-0°, 0°-35°, 45°-15°, 90°-0°, and 190°-15°. The paired comparison training involved using two slide projectors to project identical views of two similar aircraft on the screen at once. Students were called upon to identify each aircraft and to point out their distinguishing features and to comment upon how the aircraft differed in appearance. Then, the next two identical views of the same two aircraft were shown and another student called upon. This process continued until all views of each aircraft had been commented upon.

The aircraft paired were those considered by the instructor to be most similar. When a session involved five aircraft, the odd aircraft was paired with the one previously seen which seemed to be most difficult to learn. This portion of the session took about 20-30 minutes.

4. Single Image Practice. After the paired comparison training, the slides were re-ordered randomly according to a prepared list. As each

slide was shown, a student was called upon to identify the ai· raft. If he missed the identification, the instructor asked for responses from the class and then pointed the salient differences between the incorrect aircraft and the correct one. It was also possible to index to the slide named incorrectly by the student so that the class could see the differences between the views.

Single image practice was continued until all of the slides had been seen or until each student had responded to several slides. At this point, the students usually seemed to know the aircraft fairly well.

5. End-of-Session Test. The slides were re-ordered according to a prepared test list and placed in every other slot of a Carousel tray. The odd numbered slots contained cardboard squares such that the screen would be dark between test slides. A numbered slide was substituted for the cardboard square before every fifth slide so that the students could keep track of the response numbers. Each aircraft slide was presented for five seconds "on" followed by a five-second period in which the screen was dark before the next aircraft slide was presented. Therefore, each student had ten seconds to write down the name or number (or both) of each aircraft slide. All five training views of each aircraft were used in the test and no previously trained aircraft were included.

Grading

The students were asked to exchange papers after the test so that they could be graded. The test slides were shown again and class members usually called out each aircraft as it was presented. The wrong responses were marked and the papers handed back so that the owner could look them over. The papers were then handed to the instructor and the class dismissed. HumRRO personnel then went over each test to see that the grading was correct and the grades were recorded.

Review Sessions

Review sessions were scheduled at the mid point and at the end of training. The first review came after the first three training sessions and included the twelve aircraft trained in those sessions. A second review was held after training sessions 4, 5, and 6, and included the 13 aircraft trained. in those sessions. The third review was an overall review of all twenty-five aircraft.

The reviews were conducted much like the training sessions except that additional single image practice was substituted for the paired comparison training. This was done to keep session time down since 5 views of each of 12-13 aircraft were used in the first 2 reviews.

The final review used 154 views of 25 aircraft. All five views of the more difficult aircraft were used, but only four views of the easier aircraft were included. Each review session began with the three-view familiarization slides.

Sixty-item tests were given at the end of the first 2 reviews, but no test was given after the final review since the 250-item GOAR posttest was to be given during the following week.

Problems Encountered During Training

The biggest problem encountered during the two-month training period was due to the fact that many students were unavailable for at least half of the training. Leave and details caused unavoidable absenteeism which resulted in many men attending less than half of the sessions. Possibly several weeks of review should have been scheduled at the end-of-training to try to expose men to aircraft that they had missed.

Motivation among the students to learn the aircraft was fairly good. A few people would come to a session, sign the roster, and then depart and a few made no effort to learn, but for the most part, the men were willing to try to learn.

The training session test scores benefited from some cheating, but due to the small size of the room, it was difficult to do much about this. Some cheating also occurred when men did not exchange papers and graded their own. They could then cross out a wrong answer and write in a correct one and the paper would look as if they had changed their minds during the test. These factors did not affect the final end-of-test results since the GOAR posttest and the field test were designed to make it difficult to cheat.

ATTENDANCE - VAR TRAINING

X Battery, 4/1 ADA
(GOAR Kit - Slides)

Date	Session	Platoon	No. Present
18 Jun 75	Group 1 A/C	2	18
19 Jun 75	Group 1 A/C	1	13
19 Jun 75	Group 1 A/c	3	19
24 Jun 75	Group 2 A/C	1	17
25 Jun 75	Group 2 A/C	2	21
26 Jun 75	Group 2 A/C	3	15
1 Jul 75	Group 3 A/C	2	10
1 Jul 75	Group 3 A/C	3	9
2 Jul 75	Group 3 A/C	1	12
15 Jul 75	Rev. - Gps. 1,2,3	1	8
16 Jul 75	Rev. - Gps. 1,2,3	2	14
16 Jul 75	Rev. - Gps. 1,2,3	3	12
22 Jul 75	Group 4 A/C	1	13
23 Jul 75	Group 4 A/C	2	22
23 Jul 75	Group 4 A/C	3	17
29 Jul 75	Group 5 A/C	1	24
30 Jul 75	Group 5 A/C	2	15
31 Jul 75	Group 5 A/C	3	24
12 Aug 75	Group 6 A/C	1	20
13 Aug 75	Group 6 A/C	2	17
14 Aug 75	Group 6 A/C	3	20
19 Aug 75	Rev. - Gps. 4,5,6	1	13
20 Aug 75	Rev. - Gps. 4,5,6	2	17
25 Aug 75	Rev. - Gps. 4,5,6	3	24
26 Aug 75	Overall Rev. (No test)	1	13
27 Aug 75	Overall Rev. (No test)	2	8
28 Aug 75	Overall Rev. (No test)	3	0

Two platoons were scheduled for same session. Sometimes this resulted in a very large class if the turn-out was good for each platoon.

APPENDIX B

PROCEDURE USED FOR PRINTED IMAGERY TRAINING

Perspective

In contrast to the structured classroom approach used in standard (GOAR) training, printed imagery training was designed to be self-paced (the individual could take as much or as little time on each aircraft as he desired) and flexible (the individual could work with the materials in any manner he desired). However, individuals were not allowed to remove materials from the classroom--this precaution was necessary in order to reduce the possibility of contamination to the GOAR-trained battery.

Environment

A battery classroom and its adjoining dayroom were utilized. The classroom seated approximately fifty students in desks and was used to present material and conduct training.

The dayroom, separated from the classroom by a partition, was used to test the proficiency of students after training. Approximately 10 individuals could be tested at any one time.

Students

Line personnel from Y Battery participated in the training. Not all personnel were present at each training session since leave, special activities, AWOL, etc., prevented a noticeable percentage from attending. Moreover, participants drifted in and out of the sessions as they were given more critical assignments.

Materials - The Printed Imagery

Printed imagery consisted of two sets of cards--8x10 inch and 2x3 inch.

The 8x10 inch cards were printed with 5 views of each aircraft. On the front of each card were the five views of each aircraft, the aircraft name and designation, and cues distinguishing it from other aircraft. On the reverse of the card were the same five views without identifiers.

Two-by-three inch flashcards were also used. On the front of the card was one of the five views of each aircraft; on the reverse, the aircraft name and designation.

Schedule

Twenty-five aircraft were presented in nine training sessions. Each training session consisted of two- to three-hour periods in which four to five new aircraft were presented. The fourth and eight sessions were used for reviewing aircraft from the first through third and fifth through seventh sessions, respectively. The ninth session consisted of a review of all twenty-five aircraft.

Procedure

At the beginning of each session, the instructor introduced himself and explained the purpose of the training. He then proceeded to demonstrate the printed imagery material and suggested ways it might be used--the 8x10 cards could be arranged so that the same view of different aircraft would be juxtapositioned; the "buddy" system could be used allowing the two individuals to present the cards to each other; or, in any manner desired.

The printed cards were then passed out and the students proceeded to use them at their own discretion. If a student wished to study an aircraft not presented in the session, he was allowed to do so even though he had not been trained on the aircraft.

When a student felt he had mastered the aircraft presented, he turned in his material and proceeded to the dayroom where the proficiency test was administered. In order to provide individual feedback, to answer questions, or assist the student in VAR, the student remained while the test was scored.

Effort was made to conduct the sessions as routine battery training. After presenting the materials and establishing procedures, battery officers and NCOs were allowed to conduct training as they saw fit. During the training cycle, various NCOs and their training officers coached students in identifying the aircraft. This coaching even extended to critiquing the individual's performance on the test given in the dayroom. In one instance, an NCO grading a student's examination made him return to the training area, spend more time on the training materials, and retake the test.

APPENDIX C

INSTRUCTOR GUIDELINES FOR PHASE III: SATURATION TRAINING

1. The VAR training program covers a total of 80 aircraft. For training purposes, the 80 aircraft have been divided into 4 block of 20 planes--Blocks W, X, Y, and Z.

Each block has been further divided into 4 groups of 4 to 6 aircraft.

2. The training aids consist of packets of printed images. There are 15 packets for each aircraft group. The VAR 35mm slides are not used for learning. (The slides, however, are used for periodic performance tests which will be given under the supervision of HumRRO civilians.)

3. Each training packet contains 2 kinds of aids.

a. Large cards which show 5 views of a specific aircraft plus a description of prominent recognition features. The aircraft's name, number, and national origin are also shown. The reverse side displays only the 5 views.

b. In addition to the multi-view cards described above, each packet contains a deck of flash cards. The deck consists of the 5 views of each aircraft in the group. The reverse side of the card presents the name, number, and national origin.

4. These training packets were set-up to be used in a coach-pupil or "buddy" system. Pairs of men are issued a packet and they work together during

learning. They study the five-view cards and learn features. These cards are also used for testing each other using the sets of unidentified views on the back side. Next, each pair of men works with the flashcard deck, in which the views are scrambled and must be identified in turn. The men alternate in the instructor-learner roles. This "buddy" training continues until both men believe they have mastered the material.

5. The men then are issued a test booklet (there are either 6 or 10 test booklets for each group of planes in each 20-aircraft block). Each man is issued one test booklet and one VAR answer sheet. The men must take the test independently. NO HELPING ONE ANOTHER. The test views are numbered and <u>either</u> the name or the ID number is written on the answer sheet. Do not require both name and ID number.

6. When completed, the VAR answer sheet and test booklet are returned to the instructor. He scores the answer sheet by comparing answers specified on the SCORING KEY for that block and group of aircraft.

7. Tell the learner his score and retain (hold) answer sheets for collection by a HumRRO civilian. Store completed answer sheets in the instructor's test folder for each aircraft group.

8. Record training time! It is <u>essential</u> that the amount of time spend in learning be recorded. The VAR answer sheet has blocks for recording time IN and time OUT. Record the clock time IN when the training packet is issued. Record time OUT when the man requests the test booklet. Record to nearest five minutes of clock time.

9. The training method used is very flexible. Aircraft are learned in groups of 4 - 6. After a man has gone through Groups 1 and 2 of Block W (W-1 & W-2) he reviews all the aircraft in both groups and then is given a slide test over both groups. After finishing the slide test, he is eligible to begin learning (and testing) on the next 2 groups in that block. Following a review (and testing) over these second groups of aircraft (W-3 & W-4, for example), the men are given a slide test which covers the whole block of 20 aircraft.

In this system men can be training at any time, may leave it temporarily, and/or return to it and pick up where they left off. In other words, it is individualized and self-paced. For example, suppose a man has gone through three groups and then goes on leave. When he returns, he picks up where he left off--he does not miss training on specific aircraft as is the case with "lock-step" classroom methods using 35mm slides.