AD 663 093

Technical Report 68-1

A Classroom Method of Training Aircraft Recognition

by

Paul G. Whitmore, John A. Cox,
and Don J. Friel

HumRRO Division No. 5 (Air Defense)

AD ________

January 1968

Prepared for:

Office, Chief of
Research and Development
Department of the Army

Contract DA 44-188-ARO-2

HumRRO

The George Washington University
HUMAN RESOURCES RESEARCH OFFICE

40

A Classroom Method of Training Aircraft Recognition

by

Paul G. Whitmore, John A. Cox, and Don J. Friel

January 1968

Prepared for:

Office, Chief of Research and Development
Department of the Army
Contract DA 44-188-ARO-2 (DA Proj 2J024701A712 01)

HumRRO Division No. 5 (Air Defense)
Fort Bliss, Texas

The George Washington University
HUMAN RESOURCES RESEARCH OFFICE

Technical Report 68-1
Work Unit STAR
Sub-Unit I

The Human Resources Research Office is a nongovernmental agency of The George Washington University. The research reported in this *Technical Report* was conducted under contract with the Department of the Army (DA 44-188-ARO-2). HumRRO's mission for the Department of the Army is to conduct research in the fields of training, motivation, and leadership.

Published
January 1968
by
The George Washington University
HUMAN RESOURCES RESEARCH OFFICE
300 North Washington Street
Alexandria, Virginia 22314
Distributed under the authority of the
Chief of Research and Development
Department of the Army
Washington, D.C. 20310

DEPARTMENT OF THE ARMY
OFFICE OF THE CHIEF OF RESEARCH AND DEVELOPMENT
WASHINGTON, D.C. 20310

CRDBES

Feb. 14, 1968

SUBJECT: A Classroom Method of Training Aircraft Recognition

TO:

ADMINISTRATOR
DEFENSE DOCUMENTATION CENTER
ATTN: TCA (HEALY)
CAMERON STATION, BLDG.5
ALEXANDRIA, VA. 22314

1. This report describes the development of an experimental aircraft recognition training program for personnel manning forward area air defense weapons.

2. Presented with 16 aircraft on training slides, subjects of the course achieved 95% correct recognition after 16 training sessions. When presented with a different and degraded set of slide images, the subjects achieved 61% correct recognition, indicating that their recognition skill was generalized to an image condition other than the training slides. It was concluded that the prototype program was both effective and usable; several ways in which the program may be improved and refined are discussed in the report.

3. Modified portions of the prototype training method are being incorporated into Army Field Manual 44-30 (on visual aircraft recognition).

4. The findings of this report are expected to be of interest to those involved in training and operations for aircraft recognition, and, more generally, perceptual and research aspects of image recognition.

FOR THE CHIEF OF RESEARCH AND DEVELOPMENT:

C. J. CANELLA
Colonel, GS
Chief, Behavioral
Sciences Division

FOREWORD

During recent years, interest in visual aircraft recognition has been revived by the development of a variety of forward area air defense weapons systems. This report describes the techniques used in and the results obtained from the application of an initial prototype classroom method of training aircraft recognition.

This developmental effort was conducted by the Human Resources Research Office under Sub-Unit I of Work Unit STAR. Further work under Sub-Unit I will be concerned with developing a more efficient and flexible training concept. Subsequent Sub-Units will be concerned with aircraft recognition performance in the field, simulation of field conditions, and the selection of optimum recognition features for aircraft.

The prototype program described in this report was designed by Dr. John A. Cox. The record-keeping procedures were designed by SP 4 Don J. Friel. Dr. Paul G. Whitmore analyzed and interpreted the results and wrote the report.

STAR research, begun in 1965, is being conducted at HumRRO Division No. 5 (Air Defense), Fort Bliss, Texas. Dr. Robert D. Baldwin has been Director of Research during this period.

Military support has been provided by the U.S. Army Air Defense Human Research Unit and by the Air Defense Center. The Military Chief of the Human Research Unit at the time the study was conducted was MAJ A.D. Bell.

HumRRO research for the Department of the Army is conducted under Contract DA 44-188-ARO-2 and Army Project 2J024701A12 01, Training, Motivation, Leadership Research.

Meredith P. Crawford
Director
Human Resources Research Office

SUMMARY AND CONCLUSIONS

Military Problem

Visual recognition of aircraft is a critical skill for the effective use of virtually all existing and proposed forward area air defense weapons. Several informal contacts with personnel in cognizant Army agencies indicated that it was desired that trainees achieve a recognition accuracy level of from 90% to 99%. A previous pilot study showed that current Army guidelines and traditions for aircraft recognition training produced a performance accuracy of 20% on a degraded image test. Even when the inadequacies of the test were taken into account, the results strongly suggested that this approach to aircraft recognition training yielded unsatisfactory results.

Research Problems

The research effort had three objectives:

(1) To determine whether aircraft observers can be trained to a 95% level of recognition accuracy, and if so, to determine the average training time *per aircraft* required to reach such a level.

(2) To develop a prototype training approach for air defense units having a requirement for visual aircraft recognition.

(3) To gain direct experience with the conduct of aircraft recognition training as a basis from which to develop improved second-generation training methods and materials.

The traditional method of aircraft recognition training was considered unsatisfactory for the following reasons:

(1) Inordinate emphasis was placed on short-duration image exposures during training and testing.

(2) Inadequate emphasis was given to the necessity of learning to discriminate among various aircraft.

(3) The emphasis on selection and use of recognition features during training was inconsistent.

(4) The emphasis was on group rather than individual response in recognition practice.

(5) There was a lack of effective evaluation procedures during training.

(6) The level of student achievement obtained in the previous pilot study was far below a desirable level.

The 5-QQ-8 (SLARK #1) 35mm aircraft recognition slide kit currently available to the Army was considered unsatisfactory for the following reasons:

(1) A preponderant emphasis was placed on air-to-air rather than ground-to-air views of the aircraft.

(2) Distinctive background signatures on many of the slides enabled trainees to learn to identify the slide without necessarily looking at the aircraft image itself.

(3) Nationality insignia were present on many of the aircraft images.

(4) Different numbers of slides and different views were available for different aircraft.

(5) Image sizes varied inconsistently and, for the most part, were too large for training recognition at a maximum possible distance.

The Prototype Training Program

A sample of 16 U.S. and Soviet jet fighter/attack aircraft, representative of those which are currently most common, was selected as the content for the prototype training program. Slides judged to be more suitable for ground-to-air recognition than the 5-QQ-8 kit were available from a previous HumRRO pilot study.

Rationales

Examination of representative tactical situations suggested that there was no operational requirement for observers to recognize fleeting exposures (less than one second) of aircraft. Furthermore, there was no operational basis at this time for selecting a minimum exposure criterion within the one- to five-second range. The projectors which were available were equipped with exposure settings of 5, 8, and 15 seconds. Since the five-second exposure was the shortest, it was selected as the exposure to be used during testing.

Slide projections of aircraft images on generally available screens do not present the same perceptual information as do real aircraft in the natural world. It is not possible to match training images to real world images. Consequently, it was decided to train for small projected images, recognizing that such images could not be interpreted in terms of simulated target-to-observer distances in the natural world.

There were certain problems to be considered in determining whether observers should be trained to recognize aircraft on the basis of the whole image or on the basis of a fractionation of the image (image-analysis) into characteristics of the structural components of the aircraft, that is, recognition features. The whole-image approach attempts to prevent image analysis by using image exposures during training of 1/10 to 1/100 second. Whether this strategy actually prevents image analysis cannot be experimentally determined; brief exposures may only ensure that analysis proceeds more slowly and with less certainty over a greater number of exposures. Nor does the available experimental evidence support the opinion that such brief exposures during training produce higher recognition performance levels than do longer exposures. Because of the uncertainties involved in the whole-image approach, it was decided to base the prototype training program on teaching recognition features.

The Wings-Engine-Fuselage-Tail (WEFT) nomenclature system used during World War II was not adopted as a basis for specifying recognition features for the following reasons:

(1) It does not place emphasis on recognition features that discriminate among the aircraft at a distance.

(2) It allows and perhaps encourages trainees to select as recognition features characteristics that are common to all the aircraft in a program; that is, it does not necessarily lead to the selection of characteristics that discriminate among the aircraft.

(3) It uses esoteric aircraft terminology which is not familiar to most trainees.

Specific recognition features were selected judgmentally for each aircraft in the context of all the aircraft in the program by members of the research staff, most of whom had had previous experience observing aircraft in flight while participating in field studies conducted as part of other research efforts by HumRRO. These features were described in words that were considered to be familiar to the trainees.

The primary type of learning required for aircraft recognition training was identified as discrimination learning. A simultaneous paired comparison procedure was selected as the training procedure by which to accomplish discrimination learning. In order to improve the efficiency of discrimination learning, the aircraft were formed into groups on the basis of similarity and paired comparisons were made only among those aircraft within a single group.

In order to bridge the differences between the paired comparison conditions and the test conditions, single-image practice was considered necessary following discrimination learning within each aircraft group. It was also considered necessary to include periodic reviews of aircraft in previously learned groups. The practice and review activities were combined and made cumulative so as to include all previously learned aircraft as well as those currently being learned.

Eight-second exposures were deemed adequate during the early part of training. However, it was considered necessary to reduce the exposures to five seconds by the later parts of training in order to match the criterion requirement.

In order to have continual information regarding progress toward the 95% average achievement goal, it was decided to evaluate achievement on all the aircraft in the program at the end of each training session. It was further decided that, to prevent trainees from learning the test, five alternate forms rotated in sequence from session to session would be used. To evaluate the the generalization of learning, it was decided to use different views in testing than in training. And finally, to provide some common base by which to compare the program developed in this effort with the previous evaluation of the traditional approach, it was decided to administer the same type of end-of-training test as in that evaluation.

Class progression from one group of aircraft to the next was based on only those test images of aircraft in the group currently being taught. A class average of 80% was established as the criterion for progression.

Any training program which progresses from one portion to another only after the class as a whole attains some specified level of achievement is highly dependent upon the slowest learning trainees in the class. Consequently, provision was made for remedial training procedures whose application could be prescribed on the basis of the results obtained from the achievement tests.

General Structure of the Training Program

The prototype training program progressed through seven activities:

(1) Goal setting
(2) Aircraft familiarization
(3) Supplementary training
(4) Paired comparisons
(5) Recognition practice and review
(6) Achievement testing
(7) Remedial training

Activity 4 and Activity 5 were administered for each group of aircraft. Activity 6 was administered at the end of each 50-minute session. Activity 7 was administered whenever indicated by the results of a preceding achievement test. Activity 3 (supplementary training) was undertaken by each trainee at his own discretion, when and if he desired.

Results

(1) The 95% average achievement level was reached at the end of the 16th session, an average of one session *per aircraft*.

(2) The average attained by the class on a transfer test comprised of degraded images was 61%. The correlation of .82 ($p < .01$) between this test and training achievement tests indicated that the recognition skill acquired during training is not specific to the training slides; that is, it generalizes to some other image condition.

(3) Trainees consistently maintained the same relative position in the class from one achievement test to another ($W_c = .77$, $p < .005$).

(4) The silhouette sheets used in the training materials proved to be an effective medium for aircraft recognition training and greatly facilitated the efficiency of the training.

Conclusions

(1) The approach to aircraft recognition training taken in the prototype training program appears to be both effective and capable of being implemented.

(2) The recognition skill acquired during training generalizes to other image conditions.

(3) A substantial portion of the improvement obtained with the prototype program is probably due to increase in training time as compared with traditional treatments.

(4) There are a number of ways in which the prototype program can be improved.

(5) The feasibility of developing a program which uses printed rather than slide images and a coach-pupil rather than a group instructional method looks promising. Such a program would better meet the needs of units which are concerned with maintaining and updating previously acquired recognition skills and with training occasional replacements.

CONTENTS

A Classroom Method of Training Aircraft Recognition

BACKGROUND

Aircraft Recognition Training Practices in the U.S. Armed Forces

The basic programs for aircraft recognition training used by the U.S. Armed Forces during and since the Second World War have been quite similar. In general, they have begun with a verbal description of the structural characteristics of each aircraft. Typically, the WEFT nomenclature system (Wings-Engine-Fuselage-Tail) is used during this part of the training. The system specifies words and phrases to be used in describing the various types of wing, engine, fuselage, and tail structures which appear on aircraft. The extent to which the WEFT nomenclature is used in a program has tended to vary across schools and across instructors. In some instances, emphasis has been upon the utilization of the entire nomenclature system; in other applications, only some portion has been used. In some cases, the instructor has given a familiarization presentation of the WEFT description of each aircraft only once; in other instances, the instructor has required trainees virtually to memorize the descriptions.

Most Armed Forces programs have used 35mm black-and-white slides for presenting aircraft images during classroom training. Commonly, the slides show actual aircraft in natural environments. Since they represent views of opportunity of actual aircraft, the slides available for one aircraft may well show different views in different environments than those available for other aircraft. The aircraft images typically contain full nationality insignia. A few slides show unmarked aircraft models against the sky. Several additional sets of slides put out by the U.S. Naval Training Device Center during the last three or four years show aircraft models against a uniform background. These recent sets are consistent with respect to the view shown of each aircraft and include some paired comparison slides which show the same view of two different aircraft.

In both the older and more recent slide sets, the aircraft images constitute a substantial portion of the projected area of the slide. In the typical classroom set-up, the projector is placed so that the entire surface of the screen is illuminated. In such a set-up, it is not uncommon for most of the aircraft images to subtend angles of 12° to 45°, depending upon the maximum and minimum distance of trainees from the screen. For comparison, a typical jet fighter flying directly towards an observer 1000 meters away subtends an angle of less than 1°.

Following WEFT nomenclature training for several aircraft, the typical program has used single-image recognition practice. Slides are projected on the screen one at a time and the entire class responds vocally with the appropriate name for the aircraft represented in each slide. Depending upon the instructor and school, the duration of exposure of each slide may range from several seconds to 1/50 or 1/100 of a second. Most commonly, programs have used a progressively speeded exposure strategy, starting at one or two seconds during the early part of training and gradually reducing to 1/25 or 1/50 of a second by the end of training. Individual responding—either written or oral—has rarely been used. Most programs have used oral class response. Trainees have generally not been tested either during or at the end of training.

In July 1965, HumRRO conducted a pilot study in aircraft recognition training which included a program of the traditional type just described. The program began with a two-hour conference on the WEFT nomenclature system and its use in describing an aircraft's configuration. This session was followed by five 50-minute practice sessions in single-image recognition spaced over a three-day period. The slides used in the practice were presented with progressively decreasing exposures—from several seconds to 1/50 or 1/100 of a second—within each session.

Recognition of 16 representative jet fighter/attack aircraft was taught during the course of the training. They were drawn from the Army's SLARK #1 aircraft recognition kit, one of the older types of slide sets previously described. Three aircraft were introduced during each of the first four practice sessions, and four aircraft in the last session. The class responded orally throughout the sessions.

Following the last practice session, the class was given a 35mm color slide test containing 144 single-image slides, which were not part of the SLARK #1 kit. The test was composed of three views of each aircraft—head-on, tail-on, and side. One view for each aircraft represented a near distance (12-20cm projected image), one represented a medium distance (7-10cm), and one a far distance (4-5cm). The aircraft were viewed against natural sky and terrain backgrounds. Each test slide was exposed for five seconds with a five-second blank between slides. The trainees were instructed to write either the name or number designation of the aircraft in the proper numbered blank on their answer sheets. The class averaged 14% correct recognitions on this test.

The test had been designed to provide a basis for comparing groups trained by different methods with different kinds of slides. Because of the conditions under which the slides were made, many were faulty; for example, some of the aircraft images were partially obliterated by light spillover, many were too small for even experts to identify, and others were blurred. However, the test did provide a basis for comparison which gave no one group an advantage over any other. When the worst of the slides were removed, the average number of correct recognitions in the traditionally trained group rose to 20%. The test still constituted a degraded image test. Nevertheless, the results strongly suggest that the traditional approach to aircraft recognition training yields unsatisfactory results.

The Operational Situation

Aircraft recognition is a critical skill for the effective use of virtually all existing and proposed forward area air defense weapons. Such weapons include the infantry rifle, single- and multiple-mounted machine guns and cannons, and small, highly mobile missile systems such as Redeye and Chaparral. These various weapons are designed to meet different aspects of the forward area air threat. Thus, aircraft recognition requirements may be different for the different weapons. Since many of these weapons are still under development, manning and deployment doctrines are not yet available. Consequently, strict aircraft recognition requirements cannot be established for each separate weapon. However, some broad recognition guidelines can be set with regard to the degree of emphasis to be placed on the duration and size of the target image.

Under excellent visibility conditions, aircraft closing on an observer position can be positively recognized 50% of the time at a distance of approximately 3000 meters (1). An aircraft flying directly toward an observer at 400 knots would close at a rate of 205.6 meters per second. It would take 14.6 seconds to fly from a point 3000 meters away from the observer's position to a point directly over the observer's position. If the observer were manning a weapon which required a tail shot, he would have 14.6 seconds in which to view the aircraft before firing. If his weapon required a five-second preparation time, he would still have 9.6 seconds in which to view it before deciding whether to engage. If his weapon had a 2000-meter frontal engagement range, he would still have 4.9 seconds in which to view the aircraft before firing. Even if he had a weapon with twice this range, or if the aircraft closed at a faster rate, the critical restriction on recognition performance would not be the fleetingness of image duration, but rather the smallness of image size at the observer's eye.

Fleetingness of the target image can be critical only in situations in which the aircraft is revealed only during a break in high foliage or terrain. It is unlikely that a weapon could be brought to bear on an aircraft in such an instance, since some time would have to be consumed simply in acquiring the aircraft as it broke into view. Air defense weapons will undoubtedly be emplaced so as to provide a much better command of the surrounding airspace than is implied by situations in which targets are only fleetingly exposed.

OBJECTIVES

How much error in recognition accuracy can be tolerated in the forward area aircraft observer? Generating a precise answer to this question would require a complex cost-effectiveness analysis. A less precise answer can be based on military judgment. However, another question must be answered first: To what level of recognition accuracy can observers be trained and at what cost? This question in turn must be fractionated into two subordinate inquiries:

(1) To what level of recognition accuracy can observers be trained with respect to training images and at what cost?

(2) To what level of recognition accuracy can observers be trained with respect to actual aircraft in the natural world?

This study is concerned with the level of recognition accuracy to which observers can be trained.

Several informal contacts with personnel in cognizant Army agencies suggested that interest lay primarily at the 90% to 99% level of recognition accuracy. Ninety-five percent was selected as a representative training goal for research purposes. And since training cost is primarily a function of training time, the primary objective for this effort was formulated as follows: to determine whether aircraft observers can be trained to a 95% level of accuracy with respect to training images, and if so, to determine the average training time per aircraft required to reach such a level.

The typical kind of recognition program used during and since the Second World War was not considered satisfactory for the following reasons:

(1) Inordinate emphasis was placed on short-duration image exposures during training and testing.

(2) Inadequate emphasis was placed on discrimination learning.

(3) Inconsistent emphasis was placed on the selection and use of recognition features during training.

(4) Group rather than individual responding was used during training.

(5) There was a lack of effective evaluation procedures during training.

(6) The level of student achievement obtained in the 1965 pilot study was unsatisfactory.

Consequently, a second objective was formulated: to develop a clearly specified prototype classroom training approach for air defense units having a requirement for visual aircraft recognition.

Since the present effort constitutes but one of a series of continuing efforts dealing with aircraft recognition training and performance, a third objective was formulated: to gain direct experience with the conduct of aircraft recognition training as a basis from which to develop improved second-generation training methods and materials.

RATIONALES FOR THE TRAINING AND TESTING PROCEDURES

Exposure Duration and Size of the Criterion Images

In order to evaluate a trainee's ability to recognize aircraft, how long must he be allowed to view each aircraft image? The preceding descriptions of sample operational situations indicate that target exposures of less than one second, and perhaps as great as two seconds, are too short to allow for an effective engagement. But otherwise, there was little basis at the inception of this study for selecting some particular exposure duration from the one- to five-second range as constituting a minimum requirement. Since the projectors that were available were equipped with automatic exposure settings of 5, 8, and 15 seconds, the shortest—the five-second exposure—was selected as the criterion requirement.

How large (or small) should each aircraft image be in order to evaluate a trainee's ability to recognize aircraft? The preceding discussion of sample operational situations indicates concern with a number of air defense weapons characterized by a variety of engagement envelopes; that is, some of these weapons can engage aircraft at considerably greater ranges than can others. In general, however, the military advantage lies in identifying the aircraft at the greatest possible range (or smallest possible image size).

The problem is complicated by the fact that slide projections on generally available screens do not adequately simulate natural world images. Thus, a projected image which subtends the same visual angle for a given screen-to-observer distance, as does a natural world image at a given target-to-observer distance, presents less perceptual information to the observer than does the natural world image. There is a considerable loss of resolution on the screen (particularly beaded screens) so the projected image is blurred in comparison to the natural world image[1] (2). Increasing the size of the projected image will not necessarily lead to a match with respect to perceptual information. The difficulty with which some recognition features can be discriminated is primarily dependent upon image size, and the difficulty with which others can be discriminated is primarily dependent upon image sharpness and internal contrasts. Thus, an enlarged projected image may allow equal discrimination of the size and shape of an air intake but a much easier discrimination of wing position than does the smaller natural world image.

It is not yet possible to establish a direct correspondence between the characteristics of images as seen in the natural world. Consequently, it was decided to train for small projected images, recognizing that they could not be interpreted in terms of simulated target-to-observer distances in the natural world.

[1]Klaiber reports that the resolving power of a large sample of commercially available front and rear projection screens ranges from 4 lines/mm to 128 lines/mm. The resolving power of just the beaded front projection screens included in Klaiber's sample ranges from 4 lines/mm to 12 lines/mm.

Whole-Image vs. Image-Analysis Approaches to Training Design

Should potential aircraft observers be trained to recognize aircraft on the basis of the whole image or on the basis of a fractionation of the image into characteristics of the structural components of the aircraft? The previously mentioned WEFT nomenclature system is the major historical example of the latter approach. It provides the trainee with a language by which he can descriptively analyze aircraft images into characteristics of the wings, engine, fuselage, and tail. The opposing whole-image point of view was expostulated primarily by Renshaw (3), who stated: "For the attainment of the maximum skill in the visual perception of forms, wholes must be seen rather than a succession of discrete and disjointed parts."

From Renshaw's point of view, it is critical that the trainee be prevented from analyzing the image into structural components of the aircraft. The prevention of image analysis during training is difficult to accomplish and impossible to confirm. Even though the means for accomplishing such an analysis are not provided the trainee, he can make his own analysis using whatever descriptive terms he already knows or can invent. And he can accomplish such an analysis covertly so that the instructor is not aware of it.

In order to prevent—or at least discourage—the trainee from analyzing the image into parts, Renshaw suggests that the image be shown only in brief exposures. Although he does not clearly define what he means by "brief exposures," it has traditionally been interpreted to mean sub-second exposures, ranging from 1/10 to 1/100 of a second depending upon the agency or particular instructor responsible for the training. Using brief exposures does not, however, ensure that image analysis will not occur—more probably, it only ensures that analysis will proceed more slowly and with less certainty over a greater number of exposures. However, since image analysis can occur covertly, the point is virtually beyond experimental evaluation. In practical terms, the issue resolves itself into one of selecting the optimum image exposure duration for training. This problem will be discussed in a subsequent section.

In terms of the trainee's private experience, it seems likely that he would use an image-analysis approach during the early part of recognition learning and a whole-image approach during the later part of recognition learning. That is, when first confronted with the complex images of a given aircraft, he analyzes them into their simple components. As practice continues, this analysis becomes less and less necessary, until it largely ceases to occur as a conscious act. Testing this contention experimentally, however, would be difficult and costly, if possible at all.

Selection and Definition of Recognition Features

The image-analysis approach was selected as the basis for the development of the experimental program. In such an approach, salient structural characteristics of each aircraft are selected as recognition features: (a) The trainee may be required to learn a set of specified features, (b) he may be provided with a language to aid him in selecting his own features, or (c) he may be given no guidance whatsoever in the selection of features.

If he is given no guidance at all, he may select, as recognition features, a substantial number of unique details that are readily discriminable in the training images but are much more difficult to discriminate at long distances in natural world images. For instance, one of the aircraft might be distinguished in the training imagery by ridges on the upper surface of its wings. However, such ridges may not be sufficiently large to be seen at a great distance in the natural world. To reduce the size of the training image until the ridges can no longer be seen may also obliterate other salient characteristics which can be seen in the natural world images at a considerable distance. This problem can occur because of the differential effect of changes in the size of projected images on different kinds of perceptual information.

The WEFT nomenclature system does not provide adequate guidance for the selection of recognition features, since it is sufficiently exhaustive to include descriptive terms for both gross and detailed features (4).[1] It is not differentially sensitive to distance. Furthermore, it allows and perhaps encourages trainees to select many features which do not aid the discrimination of the aircraft since the feature may be common to all aircraft in the program. These difficulties can be minimized by providing the trainees with specified recognition features which have been preselected on the basis of their discriminability at great distances.

The visual recognition features which are used to define (or distinguish) a particular aircraft depend to a large extent upon the other aircraft that will be seen by the aircraft observer. The object is to select those features which will most effectively and efficiently discriminate among the aircraft to be taught in the program at the greatest possible distance (or smallest possible image size). It was decided to use a set of visual recognition features for each aircraft in the program that was selected judgmentally in the context of all the aircraft in the program. Members of the research staff served as judges. Most of them had previous experience observing actual aircraft in flight while participating in field studies conducted as part of other efforts by HumRRO. It would have been preferable to have selected visual recognition features on the basis of data obtained from a series of discrimination studies, but such an approach would have taken much too long.

And, finally, the selected visual recognition features were described in words that were considered to be familiar to the trainees. The use of esoteric aircraft terminology was not desired; it would only impose an added learning burden on the trainees. For instance, the visual recognition features of the F-4 (Phantom) selected for use in the experimental program were stated as follows:

(1) Wing tips slant up, tail slants down.
(2) Drooping short nose, swept up tail assembly.
(3) Air intakes cover side of body.
(4) Modified delta wing on bottom of body.
(5) Thick body.

For contrast, a WEFT description[2] of the same aircraft might read as follows:

Wings:
(1) Low mounted with extended chord.
(2) Marked (45°) leading edge sweep and slight trailing edge sweep.
(3) Moderate positive dihedral in outer panels.
(4) Squared-off wing tips.

Engines:
(1) Two jet engines mounted side-by-side.
(2) Large dual intakes mounted in sides of fuselage by forward mounted cockpit.
(3) Dual exhausts mounted low on fuselage under forward section of tail assembly.

[1]The *Aircraft Recognition Manual* cautions against the use of details, but does not implement this caution in all of its aircraft descriptions.

[2]This description is a composite made up from several Army and Navy sources. However, it is only slightly more extensive than any one of the source descriptions.

Fuselage:

(1) Forward slanting (drooping) pointed nose.
(2) Cockpit located well forward on fuselage.
(3) Stepped-up canopy.
(4) Fuselage thickens through engine area, then thins and tapers upward to a rounded tail cone.

Tail:

(1) Single tail assembly.
(2) Unusual empennage consists of massive fin and rudder.
(3) Horizontal stablizer mounted on fuselage with sharp (23°) negative dihedral (anhedral).

Discrimination Learning

The primary type of learning required for aircraft recognition training was identified as discrimination learning. A simultaneous paired comparison procedure was selected as the training procedure by which to accomplish discrimination learning. In such a procedure, two images representing the same view of two different aircraft are shown at the same time on the screen. In keeping with the image-analysis emphasis during early training, the trainee response during paired comparison training was designated as the identification of the preselected recognition features which differentiate the two aircraft one from another. Speed of response was not considered critical for this training activity. Instead, emphasis was considered more appropriately placed on obtaining from the trainees an exhaustive listing of the relevant recognition features which differentiated between the images in each pair. Exposure of each pair of images would be prolonged until such a listing was obtained.

Obviously, there are degrees of similarity among various pairs of aircraft. Extensive discrimination training is necessary only for those that are quite similar to each other. Consequently, the efficiency of discrimination learning could be improved by forming the aircraft into groups on the basis of similarity and restricting the paired comparison procedure to those aircraft within each group. Such a procedure would also reduce the number of new aircraft which trainees would be required to learn during any one portion of the training program.

A judgmental procedure was devised for forming the aircraft into groups. The stimulus materials consisted of cards, each of which contained silhouettes of the three plan-views (head-on, belly, and crossing) of one of the aircraft (see Figure 1). Personnel of HumRRO Division No. 5 (Air Defense) and the U.S. Army Air Defense Human Research Unit, Fort Bliss, Texas, served as judges. Ten judges were used, each of whom was instructed to sort the cards into piles of four on the basis of overall similarity. Subsequently, the groupings were formed by summing across judges. Aircraft which could not be clearly placed in one of two groups were submitted to a different and smaller set of judges for resolution of the conflict.

Single-Image Practice and Review

The criterion condition requires that the trainee name the aircraft represented by an image when the images are presented one at a time. Thus, it seemed necessary to include an opportunity for single-image recognition practice following the paired comparisons activity for each group of aircraft. This practice may be viewed as providing the trainee with an opportunity in which to shift from an image-analysis to a whole-image approach.

Plan-View Silhouettes Used for Obtaining Judgments of Similarity

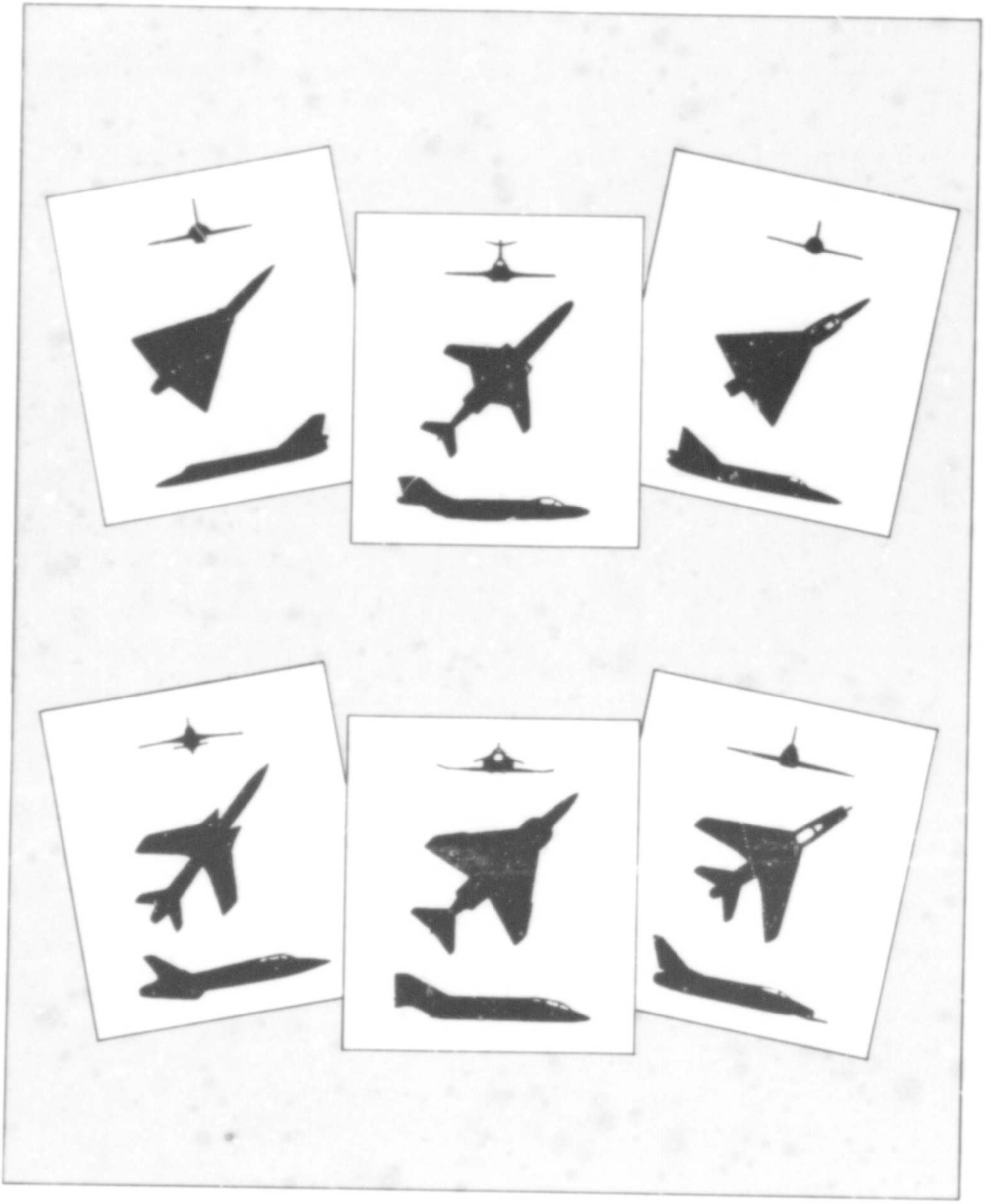


Figure 1

It also was considered necessary to include periodic reviews of aircraft in previously learned groups. The practice and review activities were combined and made cumulative so as to include all previously learned aircraft as well as those currently being learned.

Deciding on duration of exposure involved detailed consideration of previous research results. Renshaw (3) contends that brief exposures (1/10 to 1/100 sec.) in themselves facilitate recognition by both trained and untrained observers. Gibson (5) describes a study which is pertinent to Renshaw's contention. Three training and three testing conditions were used in the study. The training conditions were:

(1) Slow-Trained: one-second exposures used throughout the program.

(2) Intermediate-Trained: 1/5-second exposures during the very early part of training and 1/10-second throughout the rest of training.

(3) Fast-Trained: 1/5-second exposures during the very early part of training, gradually accelerated to 1/50-second during the rest of training.

Each training method was administered to a different group of about 170 trainees. Each group was made up of four classes. Each of four instructors taught one class in each group, so that differences in the quality of instructors were equally distributed throughout all three groups. The frequency of exposures and the amount of drill was the same for all classes in all groups. Gibson does not report how much total time was spent in training, but he implies that it was approximately equal for all groups. Furthermore, from his report it seems that training for each class was distributed over a two-week period. He does not report the number or type of aircraft taught.

Three exposure durations were established: one-second, 1/10-second, and 1/50-second, and one-third of each class was tested at each. The three sets of 20 slides used for the test were counterbalanced with exposure durations within and across training groups. The results are shown in Table 1. The means in this table are with respect to a maximum possible of 20. Reading each row from left to right across the columns, it is readily apparent that these data do not support Renshaw's hypothesis that brief exposures facilitate recognition. The data show that, in general, higher scores were made at the longer test exposures. There is one exception: The 1/50-second test condition for the Fast-Trained group produced a higher score than the 1/10-second condition for the same group. However, this difference is not statistically reliable, whereas all other differences within each row are. Furthermore, the differences between training methods within each test condition (column), with but one

Table 1

Proficiency for Various Combinations of Slide Exposure Durations During Training and Testing[a]

Group	*N*	Number of Correct Identifications[b]					
		1-sec. Test		1/10-sec. Test		1/50-sec. Test	
		Mean	SD	Mean	SD	Mean	SD
Slow-Trained	173	16.7	2.2	14.3	3.0	13.0	2.9
Intermediate-Trained	167	16.7	2.3	14.3	2.8	13.6	2.9
Fast-Trained	177	17.0	2.4	14.0	2.8	14.1	3.3

[a]Taken from Gibson (5).
[b]Maximum possible is 20.

exception, are not statistically reliable. The single exception is the statistically reliable difference between the Slow-Trained group and the Fast-Trained group under the 1/50-second test conditions. This difference would be pertinent only if it were operationally necessary to train men to recognize aircraft images presented in flash exposures of 1/50 of a second or less. As previously argued, this is not an operational requirement.

A recent study by Gavurin (6) also refutes Renshaw's contention. Six different training methods were tested, three of which used high-speed exposures. The subjects of the study were college students who were required to learn to a criterion of perfect performance. Those in the three high-speed exposure conditions took more trials to learn to criterion and made a greater number of errors. In comparing one of the high-speed exposure conditions with a matched prolonged exposure condition, the students in the high-speed condition took more than twice as many trials to learn to criterion and made almost three times as many errors during training as students in the prolonged exposure condition. The best training condition, both in terms of trials to learn to criterion and errors during training, was one that used prolonged exposures and allowed simultaneous comparisons of aircraft images.

In view of the above studies, it would appear that the only restriction that can be justified with regard to the duration of image exposures during training is that, sometime before the end of training, it match the duration of image exposures during testing. For the early part of the experimental training reported here, eight seconds seemed sufficient. The decision of when to shift from eight to five seconds seemed best made by the instructor.

Testing

As previously stated, one objective of this effort was to estimate the average training time required per aircraft to attain a level of 95% achievement. Continual information on progress toward this goal was deemed necessary. Consequently, it was decided to evaluate achievement on all the aircraft in the program at the end of each training session. Thus, during the early and middle stages of training, trainees would be tested on aircraft which they had not yet been taught. In an operational program, however, periodic achievement testing might well be restricted to those aircraft which have been or are currently being taught.

Periodic testing at the end of each training session provides both the trainee and the instructor with information concerning progress in training. Such information has several uses: (a) Progression from one group of aircraft to the next can be based on actual achievement rather than on some arbitrarily established time requirement; (b) it can act as an incentive for the trainees; and (c) it can be used as a basis for prescribing remedial training activities.

In order to prevent—or at least to discourage—trainees from learning the test, it was decided to use five alternate forms which would be rotated in order from session to session. Each form would consist of two views of each aircraft, with no view of any particular aircraft appearing in more than one form. In order to evaluate the generalization of learning, it was decided to use different views in testing than in training.

Progression from one group of aircraft to the next group was to be based on trainee response to only those test images of aircraft in the group currently being taught. The average progression criterion for the class was set at 80%.

In order to provide some common base by which to compare the program developed in this effort with the previous evaluation of the traditional approach, it would be necessary to administer the same type of end-of-training test as was previously used. This procedure was adopted. The worst slides in the test were replaced so as to bring the number back up to 144. However, many of the images in this test are still highly degraded in unrealistic ways. Although performance data obtained from the test can be used as a basis for comparative evaluation with respect to the two training approaches, it cannot be used as a basis for absolute evaluation with respect to field performance.

The degraded image test also provides an indication of generalization from the type of images used during training to some other kind of image condition. If a substantial correlation were obtained between performance in training and performance on this test, it would be reasonable to infer that the effects of training will generalize to other image conditions; that is, that they are not specific to the training images. If a low correlation were obtained, it would be reasonable to infer that the effects of training are specific to the training images.

Remedial Training

Any training program which progresses from one portion to another only after the class as a whole attains some specified level of achievement is highly dependent upon the progression of the slowest learning trainees in the class. Remedial training procedures are necessary in such an approach to reduce training delay. However, to prescribe effective remedial training procedures, it is first necessary to make a diagnosis, that is, to identify the slow learners and the nature of their difficulties. Such a diagnostic base is provided by the achievement tests administered at the end of each training session. To be effective, the tests must be scored and analyzed and remedial procedures prescribed before the next training session, either for the entire class or for selected members, depending upon the extent of the learning deficiency.

THE TRAINING PROGRAM

Selection of the Slides

Army Subject Schedule 44-2 (7), specified the use of the aircraft recognition slide (35mm) kit 5-QQ-8 (SLARK #1). Examination of this kit had suggested that there were characteristics of its slides that should be modified to make them more suitable for ground observer aircraft recognition training. The undesirable characteristics are:

(1) A preponderant emphasis on air-to-air rather than ground-to-air views of the aircraft.

(2) Distinctive background signatures on many of the slides which enable trainees to learn to identify the aircraft on the slide without necessarily looking at the aircraft image itself. The signatures include unusual backgrounds, unusual lighting, and unusual views.

(3) Nationality insignia and other markings on many of the aircraft images.

(4) A different number of views and lack of uniformity of views from aircraft to aircraft.

(5) Image sizes that vary inconsistently and are usually too large for training recognition at a maximum possible distance.

Slide images judged to be more suitable were available. Previous pilot work had included the development of experimental imagery for a large number of aircraft. Two sets of 35mm color slides were selected from this pool to develop recognition training for the following 16 jet fighter/attack aircraft:[1]

F-84	Thunderstreak	F-4	Phantom
F-86	Sabre	F-6	Skyray
F-100	Super Sabre	F-8	Crusader
F-101	Voodoo	Mig-15	Fagot
F-102	Delta Dagger	Mig-19	Farmer
F-104	Starfighter		Fishpot
F-105	Thunderchief		Fitter
F-106	Delta Dart	Yak-25	Flashlight

Both sets of slides were produced by photographing a model of each aircraft suspended in front of a backdrop of atmosphere-blue paper. Lighting conditions were constant for all photographs. The two sets differed in terms of the represented views of each aircraft and in terms of the camera-to-model distance used in making the slide.

Set I contained 10 views of each aircraft at one camera-to-model distance. Figure 2 presents a sample of these views. Set II contained five views of each aircraft at three different camera-to-model distances: far view (Set IIA), medium view (Set IIB), and near view (Set IIC). Figure 3 shows representative image sizes and Figure 4 shows representative image views from Set II.

There were not enough views for each of the 16 aircraft to provide a set of completely separate test slides. Consequently, many of the slides had to be used for both testing and training activities, with the restriction that no slide was to be used for both training and testing within any single training session. However, this restriction reduced the number of matched views of different aircraft available for paired comparisons, so it was sometimes necessary to use different views of two aircraft during paired comparisons.

Grouping of the Aircraft

The 16 aircraft were sorted into four groups on the basis of overall similarity, using the method previously described. Each group contained three to five aircraft. The final judgments yielded the following groupings:

GROUP #1	GROUP #2	GROUP #3	GROUP #4
F-102	F-84	F-100	F-101
F-106	F-86	F-8	F-104
F-4	Mig-15	Mig-19	F-105
F-6	Yak-25	Fitter	
Fishpot			

General Structure of the Training Program

The prototype training program progressed through seven activities:

(1) Goal setting
(2) Aircraft familiarization
(3) Supplementary training
(4) Paired comparisons
(5) Recognition practice and review
(6) Achievement testing
(7) Remedial training

[1]This set of aircraft constitutes an extensive sample of the most commonly deployed U.S. and Soviet jet fighter/attack aircraft during the period 1960-65.

Representative Image Views in Slide Set I

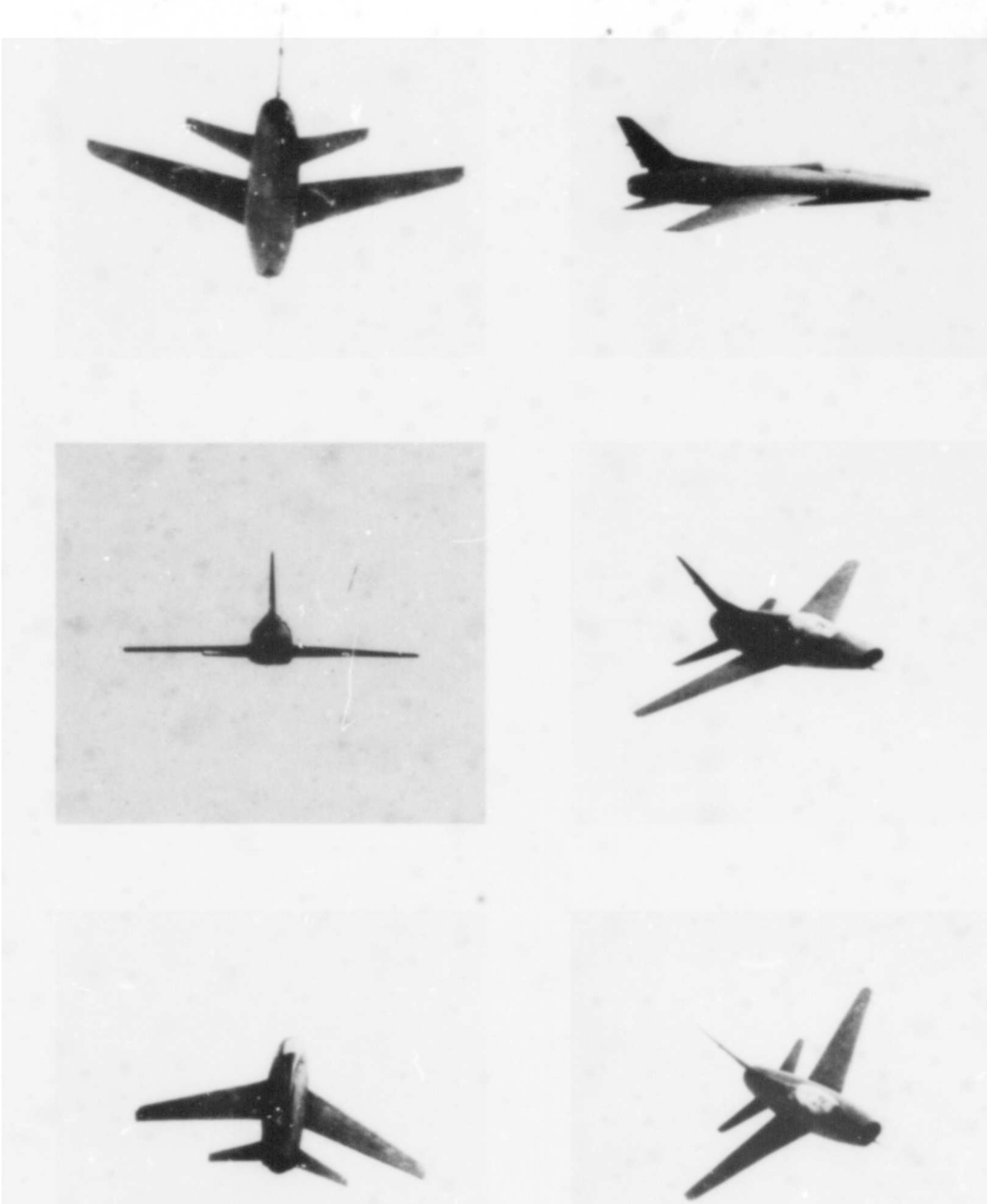
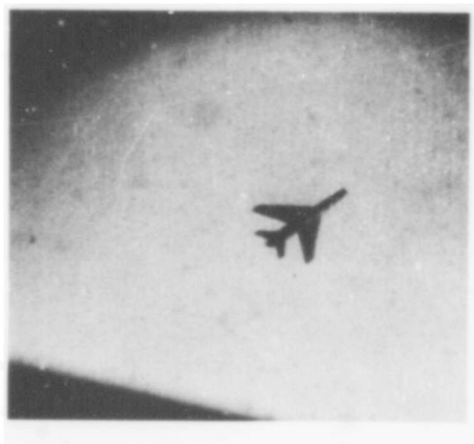


Figure 2

Representative Image Sizes in Slide Set II

Far

Medium



Near

Representative Image Views in Slide Set II

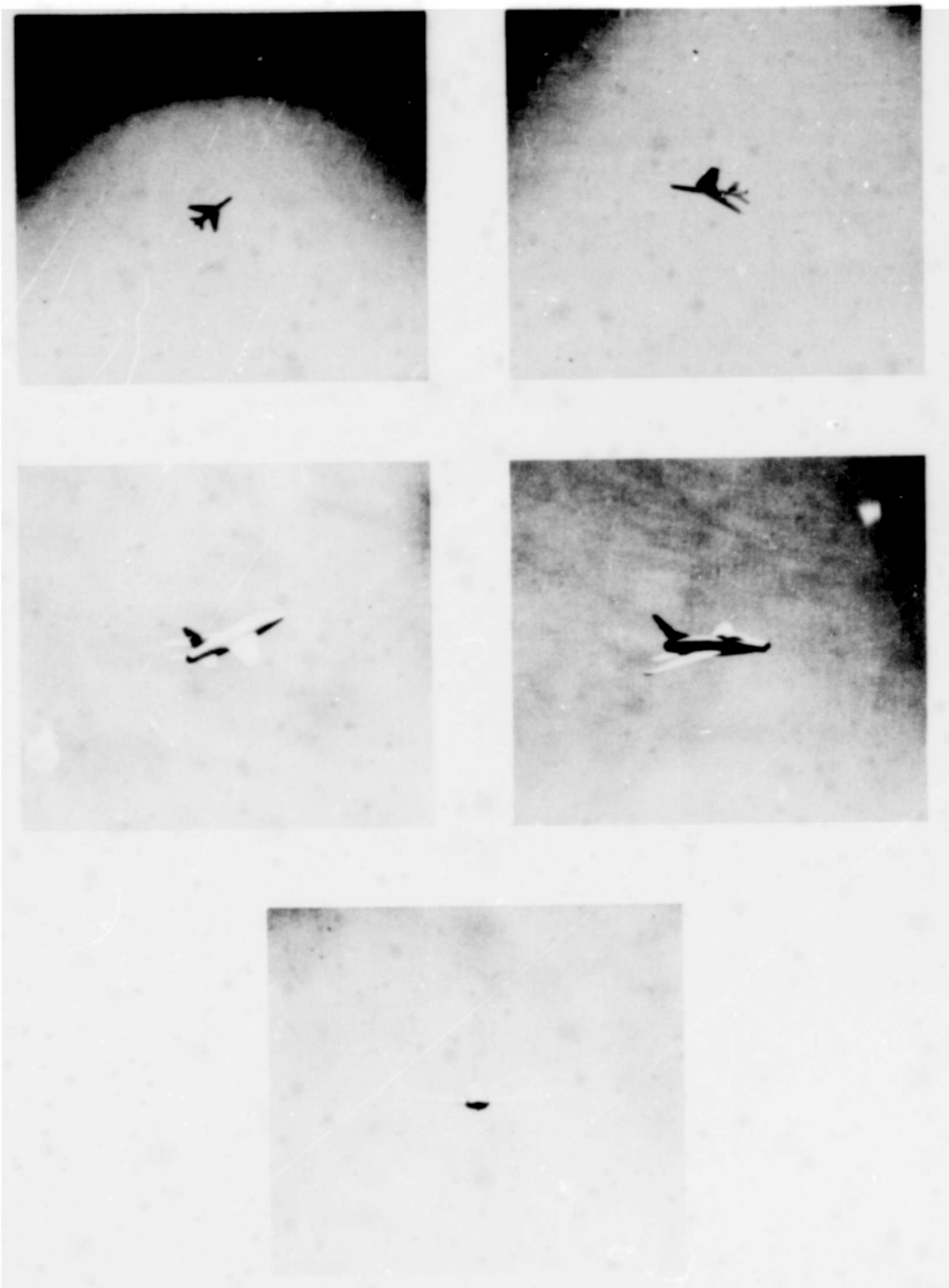


Figure 4

Activity 4 (paired comparisons) and Activity 5 (recognition practice and review) were administered for each of the four groups of aircraft. The progression from one group of aircraft to the next was made whenever the class achieved an average of 80% correct on the aircraft in the group currently being taught, as determined from the achievement test. Activity 6 (achievement testing) was administered at the end of each 50-minute session throughout the training program. Activity 7 (remedial training) was administered whenever indicated by the results of a preceding achievement test. Activity 3 (supplementary training) was undertaken by each trainee at his own discretion, when and if he desired. The remaining activities were administered only once during the program.

Functions and Means of Accomplishing Each Training Activity

(1) Goal setting. This activity had two functions:

(a) It provided both the instructor and the trainee with an indication of the trainee's pretraining aircraft recognition proficiency.

(b) It informed the trainee of the kind of performance that would be required of him by the end of the training program; that is, it established the performance objectives of training for the trainee.

It was accomplished by administering a short form of the proficiency test.

(2) Aircraft familiarization. This activity had two functions:

(a) It familiarized the trainee with the visual appearance of the recognition features of each aircraft.

(b) It informed the trainee of the aircraft and features that he would be required to learn to recognize during the training program; that is, it established the "enabling objectives"[1] (8) of training for the trainees.

It was accomplished by having the instructor display models and slides of each aircraft to the class while describing the recognition features and giving the name and number designation of each.

(3) Supplementary training. This activity had the function of providing trainees with material which would allow them to study during their own time, if they wanted to.

It was accomplished by providing each trainee with a sheet on which were printed silhouettes of the three plan-views and the name and number designations of each of the 16 aircraft.

(4) Paired comparisons. This activity had two functions:

(a) To lead the trainees into identifying and learning the recognition features for distinguishing among the aircraft in a group of similar aircraft.

(b) To lead the trainees into learning the name and/or number designations of the aircraft in a group of similar aircraft.

It was accomplished as follows:

(a) The name and number designations of each aircraft in a group were prominently displayed at the front of the classroom, and the instructor read them to the class.

(b) Two training slides showing the same view of two different aircraft in the group were simultaneously projected side-by-side

[1]Enabling objectives are defined as: ". . . the component actions, knowledges, skills, and so forth, the student must learn if he is to attain the terminal objectives. These bridge the gap between where the student is at the beginning of instruction and where he should be upon completion of instruction."

on the screen. The number designation of the aircraft appeared beneath its image.

(c) The instructor called on one trainee to describe the observable recognition feature differences between the two images.

(d) The trainee described the differences he saw. If he described actual but insignificant detail differences not listed on the card, the instructor admitted the accuracy of the differences and pointed out that they are not distinguishable at the distances at which identification must occur. The instructor's treatment of the trainee was one of helpful encouragement. If the trainee failed to designate the aircraft by name or number, the instructor asked him to do so.

(e) Another pair of slides was projected on the screen and the instructor called on another trainee to describe the observable recognition feature differences between the images. This procedure continued until the more important views of all aircraft in the group had been presented.

(f) This activity was administered only once for each group of aircraft.

(5) Recognition practice and review. This activity had two functions:

(a) It provided trainees with practice in naming the images of the aircraft in the group currently under study.

(b) It provided trainees with review and practice in naming images of previously learned aircraft.

It was accomplished as follows:

(a) An image of one of the aircraft currently under study or previously learned was projected on the screen for several (e.g., 5 to 8) seconds.

(b) The screen was blank for several seconds, during which time the trainees wrote their designations of the aircraft image on a prepared answer sheet.

(c) The aircraft image was again projected on the screen along with its name or number designation. The trainee checked his answer for correctness, and the cycle was repeated. This procedure continued for all but the last 10 minutes of the 50-minute training session. The images of the various aircraft were presented in an intermixed order.

(d) Another aircraft image was projected on the screen and the cycle was repeated. This procedure continued for all but the last 10 minutes of the 50-minute training session. The images of the various aircraft were presented in an intermixed order.

(6) Achievement testing. This activity had several functions:

(a) It provided a means of identifying those trainees who were having difficulty learning to recognize the aircraft.

(b) It identified those aircraft which the class, or a substantial portion of the class, was having difficulty learning to recognize.

(c) It established intermediate goals for the trainees and kept them informed of their progress toward these goals.

(d) The average class results determined when the class as a whole was ready to progress from one group of aircraft to the next.

It was accomplished as follows:

(a) One of five forms of the test was administered at the end of each training session, beginning with the second session.

(b) Each form of the test was composed of two training slides for each aircraft in the entire training program. The training slides used in a given session did not include the test slides scheduled for use at the close of the same session.

(c) Each slide was projected for several seconds, with a blank interval of several seconds between slides.

(d) Trainees wrote their designations on prepared answer sheets during the blank periods.

(e) Answer sheets were scored and analyzed and the scores were posted at the next session, so that the trainees were informed of their progress and the instructor could take advantage of the analysis in planning the i[illegible]ructional activities for the next session.

(7) Remedial training. This acti[illegible] had the function of correcting those difficulties, if any, identified by the achievement testing.

It was accomplished by using either of two methods:

(a) Single-image method. An image of an aircraft which was causing difficulty was projected on the screen. The instructor asked a trainee to identify the image by aircraft name or number and to describe the recognition features that were observable in the view shown. He prompted the trainee, if necessary, by pointing to salient structures or by asking the trainee leading questions. Then he asked the trainee why the image could not be of some very similar aircraft, but without projecting an image of the second aircraft. Other trainees were called upon to help the first student if he was unable to give an appropriate response.

(b) Paired comparison method. Images of the same view of two aircraft which were confusing the trainees were projected on the screen. The instructor asked one trainee to identify the two images by aircraft name or number and to describe the recognition feature differences that were observable in the view shown. He prompted the trainee, if necessary. Then he had another view of the same two aircraft projected on the screen, and again he asked the same trainee to identify the two images by aircraft name or number and to describe the recognition feature differences that were observable in the view shown. This procedure might be repeated for several more views. Other trainees might be called upon to help the first student if he was unable to give an appropriate response.

The single-image method was designed for those instances in which the aircraft causing difficulty was being confused with more than just one or two other aircraft. The paired comparison method was designed for those instances in which only two or three aircraft were being consistently confused with each other.[1]

CONDITIONS FOR THE EXPERIMENTAL TRYOUT

The experimental aircraft recognition training program was administered to 27 trainees during July 1966. Twenty trainees were provided by the 1st Battalion, 33d Artillery (Nike Hercules), 6th Artillery Group, and seven by

[1]Other remedial training techniques were used during the experimental tryout to supplement the primary ones described in the text—for instance, having trainees draw a picture of an aircraft or pick th[illegible]rrect model from a display of models.

the Center Service Group, Fort Bliss. Three of these men missed one-third or more of the training sessions and were dropped from the final evaluation of the program, thus leaving 24 who completed the program.

Training was conducted in two separate one-hour (actually 50-min.) sessions per day and was scheduled to continue until the class reached an average achievement level of 95%. This level was reached by the end of the 16th session. Because of administrative and scheduling difficulties, the degraded image test was administered to only 18 of the trainees four days after the completion of training.

A beaded, front projection screen was used during training and testing. This type of screen typically possesses poor resolution characteristics. The slides were shown on 35mm projectors that had a tray capacity of 80 slides. Two projectors were used simultaneously during paired comparison training; they could be set to cycle automatically every five or eight seconds. The slide trays used for both the training and test images were loaded before each session began.

The instruction was conducted by a civilian member of the research staff who had many years of college teaching experience. He was assisted by an Army private who had completed college, but had no teaching experience.

Five alternate forms of the achievement test were used during the experimental tryout to prevent trainees from learning the test. The forms were presented in a fixed order during the training. Identical training and achievement test slides were not used in the same session. One form of the test was administered during the first session before training began as a means of establishing an untrained baseline. It is referred to as the pretest.

RESULTS

The average recognition level the trainees achieved on the pretest and on the tests given at the end of each 50-minute training session is shown in Figure 5, which covers all 16 aircraft. The data in Figure 6 are the same as that presented in Figure 5, but separated by aircraft groups. The vertical bars indicate the session at which each group of aircraft was introduced into the classroom training. Learning of aircraft in a group before that group is introduced into the classroom training can be accounted for only in terms of exposure to test images and use of the silhouette sheets. Since no direct and immediate feedback was available for test exposures, use of the silhouette sheets would appear to be the more powerful factor. Achievement on Group 2 aircraft rose to 33% before Group 2 aircraft were taken up in the classroom; achievement on Group 4 aircraft rose to 79% before Group 4 aircraft were taken up in the classroom. These results suggest that the silhouette sheets were an effective medium for aircraft recognition training and greatly facilitated the efficiency of the training.

Continued improvement in achievement beyond 80%, on aircraft in a group following progression to a subsequent group, can be accounted for in terms of both the silhouette sheets and the cumulative review incorporated into the single image practice sessions. It was the effect of these procedures that allowed the use of an 80% progression criterion during training and yet resulted in the attainment of the 95% terminal criterion. The 80% criterion on Group 4 was attained in the 12th session. Consequently, the 13th, 14th, 15th, and 16th sessions were concerned solely with bringing class achievement across all aircraft up to a 95% average.

Progression of Overall Achievement During Experimental Tryout

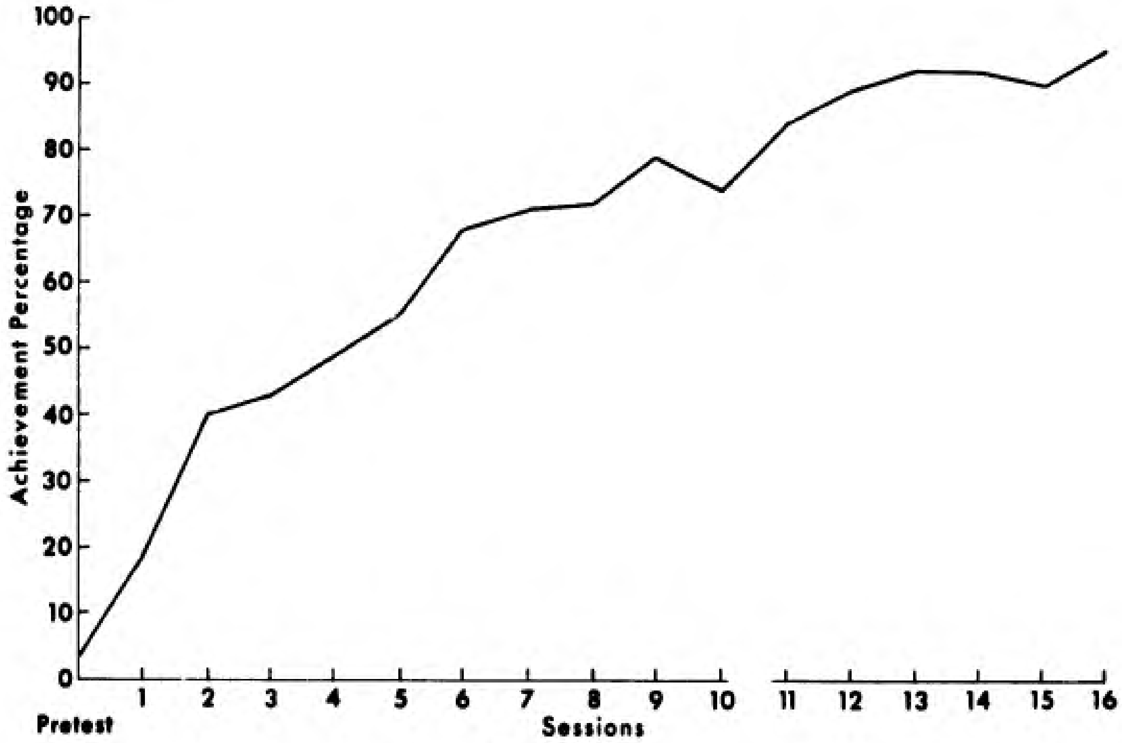


Figure 5

Progression of Achievement by Aircraft Groups During Experimental Tryout

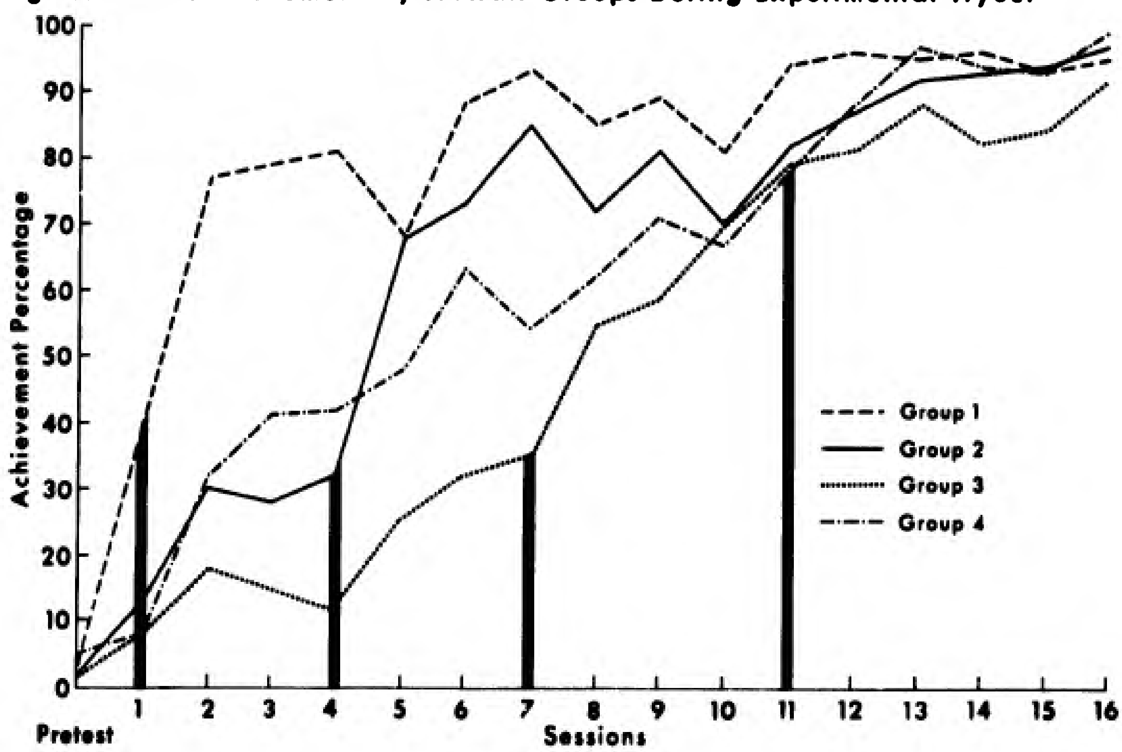


Figure 6

The 95% achievement level for all 16 aircraft was attained at the end of the 16th session. Thus, in projecting training time requirements for a program covering a given number of aircraft, it can be expected that as many 50-minute sessions will be required as the number of aircraft to be covered by the program.

As the number of aircraft covered in a program increases, two counter-acting kinds of effects can be expected to occur. First, the average amount of time required <u>per aircraft</u> may <u>decrease</u> as a result of trainees having learned how to learn to recognize aircraft; that is, learning to recognize several aircraft may make it easier to learn to recognize several more. The trainees may have more confidence in their ability to learn. They may also have become more sensitive to structural differences among aircraft and have learned to recognize certain structures in a variety of views. Since these structures may also be critical to as yet unlearned aircraft, the trainees will have provided a basis for positive transfer. They will have learned a language for specifying structural differences and will have become familiar with terms and systems for designating aircraft. Second, the amount of time required <u>per aircraft</u> may <u>increase</u> because of the increased opportunities for confusion between aircraft and because of the greater memory burden.

The first group of effects are facilitating, the second are interfering. The nature of the facilitating effects suggests that they would peak out relatively early in an extensive program. The interfering effects, however, would continue to accrue throughout the program. Practice and review activities would become lengthier as more and more aircraft had to be covered. A more efficient way of handling a large number of aircraft might be to divide them among two or more programs based on aircraft similarity. Practice and review activities within each program would be cumulative within but not across programs. A practice and review program (with remedial training as required) cumulative across programs might be administered following the second and all subsequent programs. However, this suggestion is purely speculative, since no data bearing on these issues are available.

The individual scores made on the achievement tests given at the end of each session are shown in Figure 7. These are the same scores from which Figure 6 was derived. Figure 7 graphically illustrates the laggards on each test. Lagging would presumably have been even greater had remedial training procedures not been used throughout the course of the program to provide additional or corrective training to trainees who were falling behind.

Lagging might have been further reduced by splitting the class into a slow and a fast track. Before such a strategy could be undertaken, it would be necessary to determine the consistency with which trainees maintain the same relative position in the class from one achievement test to another. If consistency were low, then such a strategy would not be feasible.

The coefficient of concordance, a measure of consistency in performance across sessions, was .77 ($\underline{p} < .005$).[1] Since this coefficient can range from 0.00 to 1.00, the .77 value indicates a relatively high degree of consistency—sufficiently high to justify exploring the feasibility of splitting the class into a slow and a fast track.

On the basis of the average scores achieved on the recognition test at the end of the fourth session, the trainees were divided into fast and slow learners

[1]To apply this statistic, it was first necessary to fill in achievement test scores for trainees who were absent from some sessions. The scores were obtained by averaging the scores on the preceding and succeeding sessions. In this manner, 22 of 384 scores were obtained. The scores from each session were then ranked from highest to lowest, and the coefficient of concordance was computed from the total aggregate of these ranks.

Dispersion of Raw Achievement Scores During Experimental Tryout

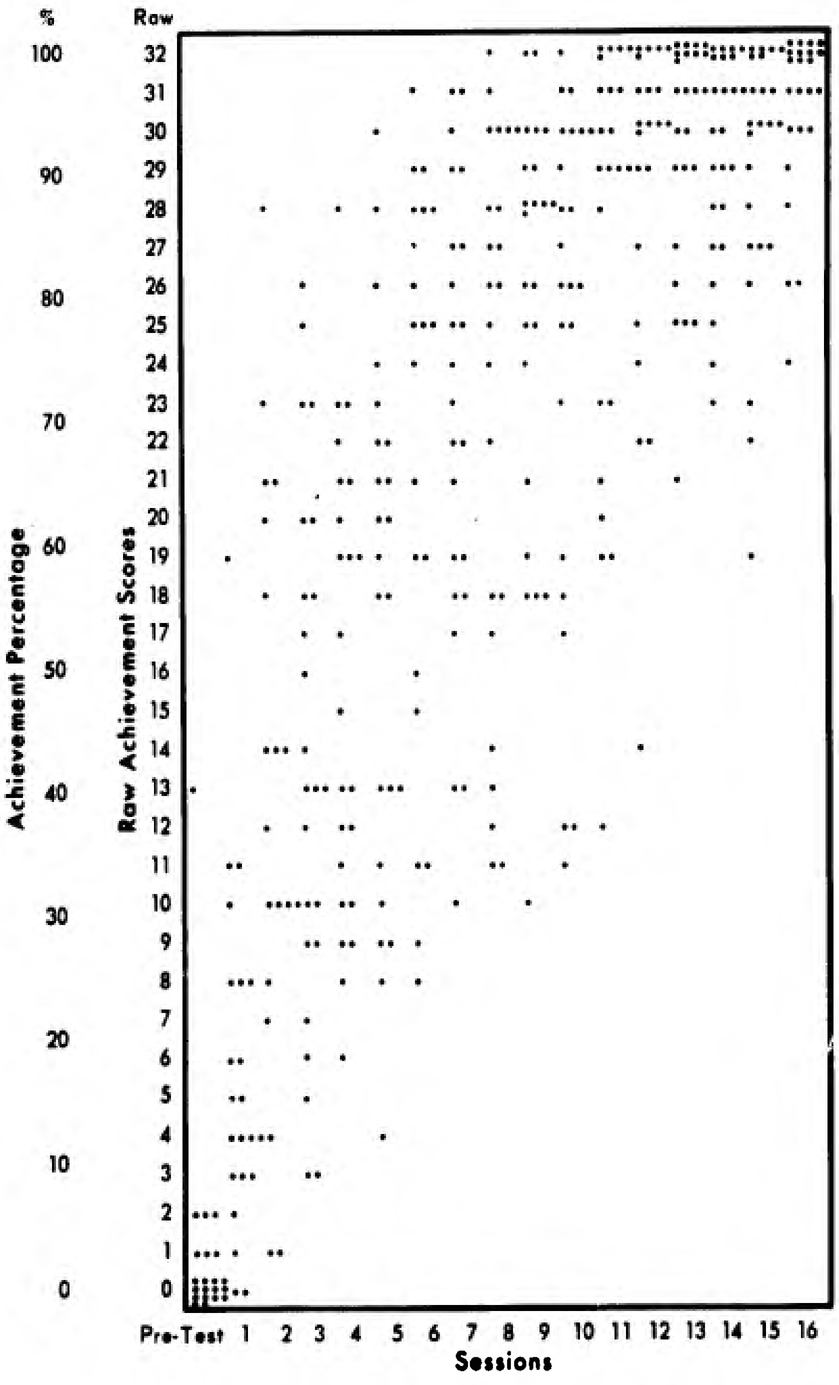


Figure 7

and their subsequent scores were analyzed on this basis. The average level of achievement of these two groups on the pretest and on the tests given at the end of each session is shown in Figure 8. The fast learners attained 95% achievement by the end of the 11th session. The slow learners had attained only 71% achievement by the end of the same session and only 91% by the end of the program (16th session).

Progress of Fast and Slow Learners During Training

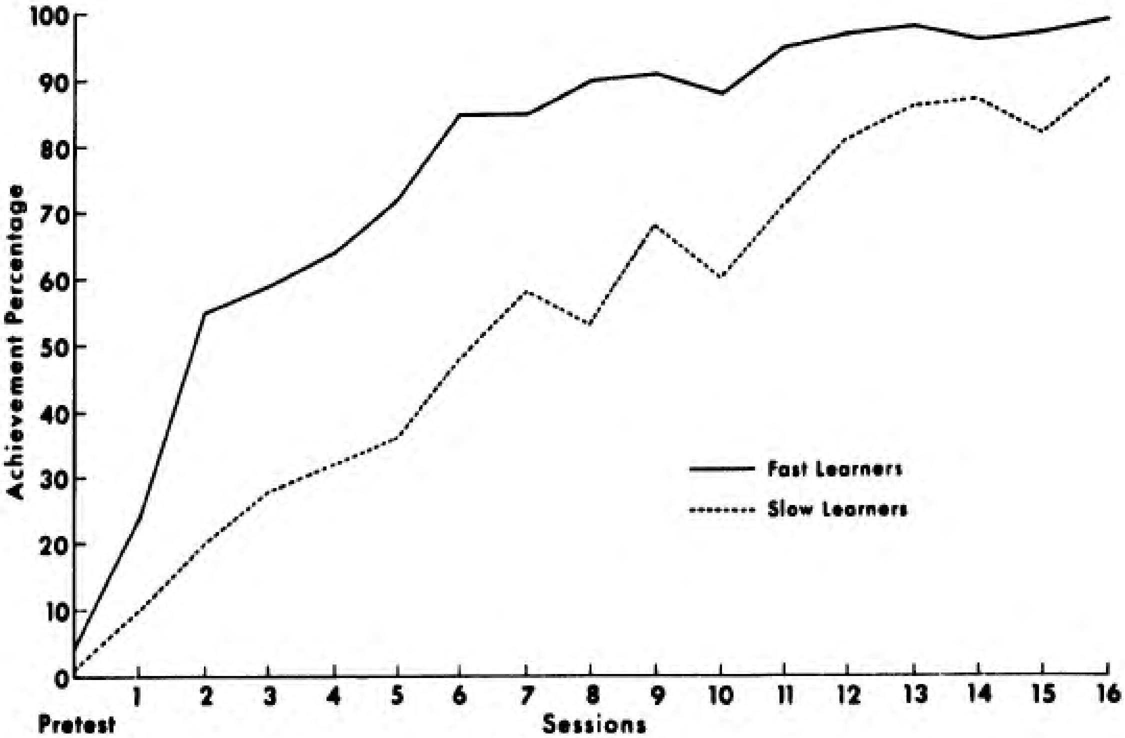


Figure 8

Consistency was such that only two trainees changed position from the first four sessions combined to all 16 sessions combined. Furthermore, the three trainees who were dropped from the program because of excessive absences invariably placed in the lowest fourth of the class during each of those sessions for which they were present.

Eighteen of the 24 trainees who completed the program were administered the degraded image test. The number correct on this test ranged from 63 to 106 with an average of 87.4 or 61% of the 144 slides in the test. This achievement is three times greater than the performance level obtained when the traditional program was evaluated in the 1965 pilot study.

The correlation between scores on the degraded image test and the average of the 16 achievement test scores for each trainee is .82 ($p < .01$). This is a substantial correlation and indicates that the effects of training are not specific to the training images; that is, the effects may be expected to generalize to other image conditions.

DISCUSSION

The substantial correlation ($r = .82$, $p < .01$) between training achievement and performance on the degraded image test indicates that the recognition skill produced by means of the prototype program is not specific to the slides used during training; that is, the skill will generalize to some other image condition.

The prototype program produced a performance level on the degraded image test three times as high as that obtained during the previous evaluation of the traditional approach. It also required 2.3 times as much training time as did the traditional program (7).[1] Consequently, it may well be that the improved performance was obtained largely because of increased training time. It should be noted, however, that it is not known whether increasing training time with the traditional approach would have resulted in a proportional increase in performance.

There was still a somewhat unsatisfactory amount of lagging on the achievement test given in the final session; four trainees had unsatisfactorily low scores ranging from 75% to 88%. If an individual rather than a class average achievement criterion had been applied, these trainees would probably have either been failed or been given additional training.

The strategy of splitting the class into two separate tracks has already been mentioned. Had such a strategy been employed and the class split at the end of the fourth session, the fast-learning track would have attained the 95% average achievement criterion within 11 training sessions. Furthermore, remedial instruction from the fifth session on would have been specific to the fast-learning track and could reasonably have accelerated their learning beyond the rate actually obtained from them. The time saved on the fast learners could then have been given to the slow learners to bring them up to the 95% average achievement criterion. Such an approach would consume the same number of trainee hours, but would require a greater number of instructor hours. Instructor hours could be held constant by splitting two concurrent classes and combining the slow tracks of each and the fast tracks of each to form two reconstituted classes. The primary advantages of such a split-track strategy is that it distributes training time in direct accord with learning rate and thus reduces attrition if an individual rather than a class average achievement criterion is used.

If a split-track strategy is used, then it is probably advisable to use a higher progression criterion for the slow track than the 80% used in the experimental tryout. Setting a higher criterion—say 90%—should reduce confusions with previously learned aircraft groups and thus reduce the need for remedial training and the amount of time spent in subsequent practice and review. The 80% criterion, however, appears to be quite adequate for the fast track.

The prototype program developed in this study represents a state-of-the-art effort. Although it produced the achievement level sought, the contribution of each of its separate components toward the attainment of that level has not been determined. A determination of those contributions might well lead to modifications which would result in improved training efficiency.

The relative contributions of the paired comparison and single-imag practice procedures is not known. There is a strong possibility that the major contribution to learning is made by the paired comparison procedure. If that theory should prove correct, it would be advantageous to place primary emphasis on the paired comparison procedure. This might be done in either or both of two ways:

(1) Paired comparison training might be repeated several times for each group of aircraft. It could also be made cumulative for review purposes.

(2) In a group procedure, only one trainee actively responds at any given time during paired comparison training. Trainee response could be increased by means of a coach-pupil procedure utilizing images printed on cards.

[1]ASubjScd 44-2 specified that three new aircraft be introduced during each nominal one-hour session. The prototype program averaged one aircraft per hour.

In addition, the reduction or elimination of single-image practice and review would lead to a substantial reduction in the number of slides required for training. Training slides might be further reduced by selecting for training both (a) those views which provide the greatest generalization of aircraft recognition to other views and (b) those operationally critical views which provide very little generalization to other views (for instance, flat head-on).

The prototype program differs from the traditional approach with respect to recognition features primarily in that the former provides the trainees with a specific set of preselected features for each aircraft in the program. However, neither approach actually tests the trainees on their learning of recognition features. The prototype program intermixes learning the list of recognition features for each aircraft with learning to discriminate among similar aircraft. Learning a list of features can be accomplished by presenting images of each aircraft separately and having the trainee respond with each recognition feature that can be seen in a given view and with the name of the aircraft. Discrimination learning, on the other hand, requires that the same view of at least two different aircraft be presented at a time and that the trainee respond by naming the recognition feature differences between the two and by naming the two aircraft.

These two different learning requirements could be separated by requiring the trainees to learn the list of recognition features for each aircraft to a specified and tested criterion before progressing to discrimination learning via paired comparisons. The primary effect of such an approach may be an improvement in retention—particularly savings in relearning time—rather than a reduction in initial training time. It should be noted that retention has not been determined for either the traditional approach or the prototype program.

Control of recognition feature learning could be increased in the prototype program, without making it a separate learning requirement, simply by converting the silhouette sheets into flashcards. The front of each flashcard would contain silhouettes of the three plan-views of an aircraft and the back would contain a listing of the recognition features and the name and number designation of the aircraft. This method has been followed in two applications of the prototype program since the conduct of the training reported here. Flashcards have been incorporated into the prototype program in two ways:

(1) Trainees have used them as references during the paired comparison, recognition practice and review, and remedial training activities.

(2) Trainees have used them for study on their own time.

The flashcards might be further improved by using good quality photographic reproductions of those aircraft views which provide the greatest generalization to other views in place of the three plan-view silhouettes. Again, operationally critical views which provide little generalization should also be included.

In another recent application, silhouette flashcards were used as the sole training medium in a coach-pupil method.[1] The objective was to determine the degree of learning that could be achieved within a restricted (approximately 30 hours) and relatively uncontrolled period. The test consisted of six slides for each of 17 aircraft, making a total of 102 slides. Trainees recorded their study time on prepared cards which were given to them at the beginning of the study. Study time per trainee averaged two hours and 40 minutes, ranging from one hour to five hours and 45 minutes. The average number of correct

[1]The technique was used in a test of a reduced-scale training program for use of small arms in an Air Defense role. The test was performed by E.W. Frederickson, R.D. Baldwin, and Robert J. Foskett.

recognitions on the test was 21.3 (21%), ranging from 1 to 56. For a similar time period and with approximately the same number of aircraft, trainees in the prototype classroom program described in this report averaged about 42%. Thus, the classroom program was about twice as efficient as the coach-pupil program.

Flashcards could be used in a flexible coach-pupil program as a means of meeting the aircraft (and armor) recognition training required by ATP 44-15 (U.S. Army Air Defense Artillery Battalions), 6 November 1959[1] for all enlisted personnel during Advanced Individual Training (4 hours) and Basic Unit Training (4 hours). An aircraft recognition test following the flashcard training might then be administered to all potential ground-aircraft observers. The final selection for ground-aircraft observers could be made on the basis of performance on the test. Those who were selected could then be given the advanced classroom training required during Basic Unit Training. Such a system would automatically select personnel with the optimum combination of interest and learning capability available in the unit.[2]

A classroom training method using projected slide imagery is not convenient for many types of Army units. Many units are concerned with maintaining and updating previously acquired recognition skills and with training occasional replacements. They need training materials that can be used with very small groups of (or perhaps single) trainees on a highly flexible training schedule. Their needs could be met by using printed rather than projected imagery—for instance, an organized set of carefully developed flashcards for each critical learning requirement that can be used either individually or by coach-pupil teams. Such an approach would have to be combined with a structured training management system consisting of a series of achievement tests and a record-keeping system to determine and control trainee progress. These procedures, along with photographic reproductions of optimally selected aricraft views, should considerably improve the efficiency of the coach-pupil method. In fact, such a system might well be developed to supplement or even replace the classroom approach for initial recognition training.

[1]This document is currently being revised, but it is understood that this requirement will also appear in the revision.

[2]Personnel should first be screened on the basis of their far visual acuity.

LITERATURE CITED

1. Wright, A. Dean. *The Performance of Ground Observers in Detecting, Recognizing, and Estimating Range to Low-Altitude Aircraft,* HumRRO Technical Report 66-19, December 1966.
2. Klaiber, Robert J. *Physical and Optical Properties of Projection Screens,* Technical Report NAVTRADEVCEN IH-63, U.S. Naval Training Device Center, Orlando, Florida, December 1966.
3. Renshaw, Samuel. "The Visual Perception and Reproduction of Forms by Tachistoscopic Methods," *J. of Psychol.* vol. 20, 1945, pp. 217-232.
4. Department of Defense. *Aircraft Recognition Manual,* Department of the Army FM 30-30, Department of the Navy NavWeps 00-80T-75, Department of the Air Force AFM 50-40, Marine Corps NavMC 2522, Department of Defense, Washington, June 1962.
5. Gibson, James J. (ed.). *Motion Picture Testing and Research,* Army Air Forces Aviation Psychology Program Research Report No. 7, U.S. Government Printing Office, Washington, 1947.
6. Gavurin, Edward I. *An Evaluation of Various Tachistoscopic and WEFT Techniques in Aircraft Recognition,* Technical Report NAVTRADEVCEN IH-40, U.S. Naval Training Device Center, Port Washington, New York, November 1965.
7. Department of the Army. *Visual Aircraft Recognition,* Army Subject Schedule No. 44-2, Department of the Army, Washington, September 1964.
8. Ammerman, Harry L., and Melching, William H. *The Derivation, Analysis, and Classification of Instructional Objectives,* HumRRO Technical Report 66-4, May 1966.

Unclassified

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*

Human Resources Research Office
The George Washington University
Alexandria, Virginia 22314

2a. REPORT SECURITY CLASSIFICATION

Unclassified

2b. GROUP

3. REPORT TITLE

A CLASSROOM METHOD OF TRAINING AIRCRAFT RECOGNITION

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

Technical Report

5. AUTHOR(S) *(First name, middle initial, last name)*

Paul G. Whitmore, John A. Cox, and Don J. Friel

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
January 1968	36	8

8a. CONTRACT OR GRANT NO.

DA 44-188-ARO-2

b. PROJECT NO.

2J024701A712 01

c.

d.

9a. ORIGINATOR'S REPORT NUMBER(S)

Technical Report 68-1

9b. OTHER REPORT NO.(S) *(Any other numbers that may be assigned this report)*

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY

Office, Chief of Research and Development
Department of the Army
Washington, D.C. 20310

13. ABSTRACT

A prototype classroom training program was developed to train observers to recognize 16 jet fighter/attack aircraft to a criterion performance level of 95% correct recognition at five-second exposures. Previously developed experimental 35mm color slides were used for training. The training method placed emphasis on recognition feature learning, discrimination learning by means of similarity groupings of aircraft and simultaneous paired comparisons, cumulative practice and review, periodic testing, and remedial training. The 95% level was reached during the 16th 50-minute session, an average of one aircraft per session. On a transfer test using degraded images the class averaged 61%—three times higher than a traditionally trained class in a previous pilot study. Most of this gain, however, may be due to increased training time. There was a substantial correlation between the transfer test and achievement, indicating that the recognition skill acquired during training would transfer to some other image condition. There are suggestions for improvement of the prototype program.

Unclassified
Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Aircraft Recognition Training						
Training Methods						
Perceptual Learning						

Unclassified
Security Classification