FAA-AM-80-7

THE SELECTION OF AIR TRAFFIC CONTROL SPECIALISTS: I. HISTORY AND REVIEW OF CONTRIBUTIONS BY THE CIVIL AEROMEDICAL INSTITUTE

Edited by:
William E. Collins
James O. Boone
Allan D. VanDeventer
Civil Aeromedical Institute
Federal Aviation Administration
Oklahoma City, Oklahoma

JUNE 1980

Document is available to the public through the
National Technical Information Service,
Springfield, Virginia 22161

Prepared for
U.S. DEPARTMENT OF TRANSPORTATION
Federal Aviation Administration
Office of Aviation Medicine
Washington, D.C. 20591

NOTICE

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its content or use thereof.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FAA-AM-80-7		

4. Title and Subtitle	5. Report Date
THE SELECTION OF AIR TRAFFIC CONTROL SPECIALISTS: I. HISTORY AND REVIEW OF CONTRIBUTIONS BY THE CIVIL AEROMEDICAL INSTITUTE	JUNE 1980
	6. Performing Organization Code
7. Author(s) Edited by: William E. Collins, James O. Boone, and Allan D. VanDeventer	8. Performing Organization Report No.
9. Performing Organization Name and Address FAA Civil Aeromedical Institute P.O. Box 25082 Oklahoma City, Oklahoma 73125	10. Work Unit No. (TRAIS) 11. Contract or Grant No.
12. Sponsoring Agency Name and Address Office of Aviation Medicine Federal Aviation Administration 800 Independence Avenue, S.W. Washington, D.C. 20591	13. Type of Report and Period Covered 14. Sponsoring Agency Code

15. Supplementary Notes

Work was performed under Task AM-C-80-PSY-74.

16. Abstract

For two decades the FAA Civil Aeromedical Institute (CAMI) has engaged in active research programs exploring most aspects of the problems of effectively selecting air traffic control specialists (ATCSs) for the FAA. The results of those efforts have contributed directly to the establishment of revised ATCS selection standards by the Civil Service Commission (CSC). Early studies on the validity of aptitude tests for predicting successful completion of Academy training led to a decision to use such tests for part of the CSC screening standard. Later studies led to the establishment of a maximum age standard of 30 years for entry into ATCS training. In addition, CAMI researchers have continuously evaluated the validity of existing standards, have examined numerous variables and alternative aptitude measures, and have provided a number of data-based recommendations in an effort to upgrade the effectiveness of predicting success in ATCS training. This paper reviews that research with emphasis on aptitude screening measures, attrition, age, prior experience, education, sex, military ATCS training, and the Uniform Guidelines on Employee Selection.

17. Key Words	18. Distribution Statement
Air Traffic Control Specialists Personnel Selection	Document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	53	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

FOREWORD

This report was prepared at the request of the FAA Office of Aviation Medicine to serve as one of several sources documenting various aspects of research on the selection of air traffic control specialists (ATCSs). It is largely based on work accomplished by Bart B. Cobb (now deceased), whose published studies and unpublished mini-studies (memorandum reports both to FAA officials in personnel and training and to representatives of the Civil Service Commission (CSC--now the Office of Personnel Management)) spanned over a decade and a half; that work contributed significantly to FAA-CSC decisions regarding ATCSs and has been a prime source for contractors doing FAA-sponsored research on the air traffic control occupation. To the extent possible, much of the material was compiled directly from Cobb's writings, albeit reorganized and updated, with minimal editing; he would have preferred it that way. Similar treatment was accorded the work of W. Dean Chiles (also deceased) with regard to the section on the Multiple Task Performance Battery. Although the emphasis in this review is on Cobb's formidable contributions, that emphasis is not intended to obscure either the significant contributions of David K. Trites, with whom Cobb collaborated during 1960-65, or the more recent analytic work at the Civil Aeromedical Institute by Mary A. Lewis (now with PPG Industries, Inc.), James O. Boone, and Allan D. VanDeventer, the impact of which is still too fresh to place definitively in a historical perspective.

ACKNOWLEDGEMENT

We gratefully acknowledge the assistance of Jo Ann Steen, Lena Dobbins, and Marcia Grimm in the preparation of this manuscript.

TABLE OF CONTENTS

THE SELECTION OF AIR TRAFFIC CONTROL SPECIALISTS:

I. HISTORY AND REVIEW OF CONTRIBUTIONS BY THE CIVIL AEROMEDICAL INSTITUTE

History of Civil Aeromedical Institute Selection Research and Related Civil Service Commission Standards

Eligibility for Air Traffic Control Specialist (ATCS) training with the Federal Aviation Administration (FAA) has traditionally included consideration of an applicant's preemployment experience, educational background, the outcome of an interview with management officials, and the results of a medical examination. Previous relevant experience, particularly in military air traffic control, has always been heavily weighted in the selection process. Experience as a pilot and various types of previous work in communications and air surveillance have also been consistently viewed as important assets. In general, however, ATCS selection programs prior to 1962 involved no formal assessment of mental abilities or aptitudes (17).

Historically, research on ATCS selection is rooted in a 1950 contract by the Civil Aeronautics Administration (CAA--precursor to the FAA) for the development of aptitude tests that could be used to select ATCS trainees. The results of that contracted study (i) indicated that aptitude tests potentially could make an effective contribution in the selection process and (ii) provided the format for an Air Traffic Problems (ATP) test (50).

In 1956, Brokaw of the United States Air Force's Personnel Laboratory, in a joint effort with the CAA, administered a large number of aptitude tests to 197 ATCS trainees. His findings (8) indicated that a composite aptitude test score formed by adding together scores on tests of arithmetic reasoning, symbolic reasoning, code translation, and the ATP test could effectively predict instructors' ratings of training performance and supervisors' ratings of job performance approximately a year after training.

Then in 1959, a number of ATCS trainees were recruited on an experimental basis using the Federal Service Entrance Examination. For a variety of reasons, the results of that experiment were considered inconclusive (50).

A continuing research program on ATCS selection began in 1960 with the establishment of the FAA Civil Aeromedical Research Institute (now the Civil Aeromedical Institute--CAMI) in Oklahoma City, Oklahoma. The first study (46) was a followup of the 197 trainees tested by Brokaw in 1956. Job performance ratings were obtained by Trites (46) from supervisors of 143 of the former trainees; the vast majority of the latter were fully qualified ATCSs. To evaluate the effectiveness of the aptitude tests recommended by Brokaw, the 143 men were classified as either satisfactory or marginal in their job performance as determined from the ratings and comments of their supervisors. The distribution of the composite scores based on the four aptitude tests given each trainee in 1956 was divided into fourths and the percentages of satisfactory and marginal individuals in each fourth were computed (see Figure 1). Scores achieved on the aptitude tests given at the beginning of training were clearly related to job performance evaluations obtained 4 to 5 years after training.

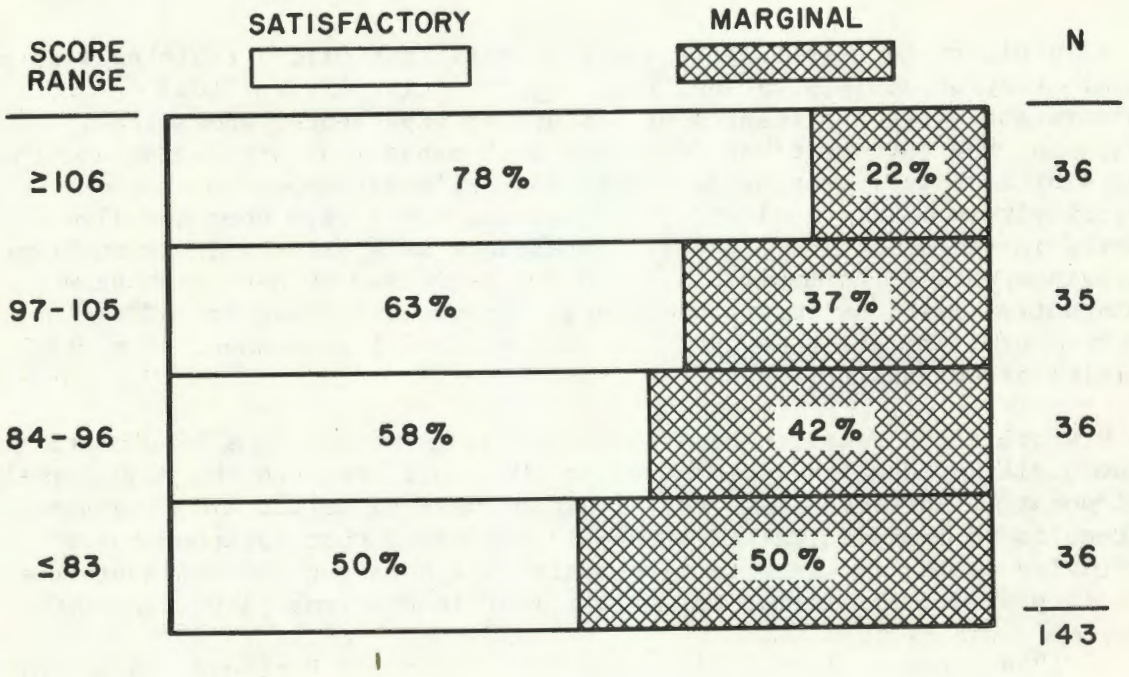


Figure 1. Percentages of marginal and satisfactory ATCSs in approximate fourths of composite aptitude test scores predicted from an early regression equation (from Trites and Cobb, 1964, ref. 50).

Concurrently with the collection of the facility performance data on the trainees tested in 1956, extensive experimental psychological testing of all En Route and Terminal trainees who entered the FAA Academy in Oklahoma City for basic ATCS training was begun in August 1960 (50). All of these trainees had been selected by medical and experience requirements and none by aptitude tests. The group of commercially available tests with which experimentation was started required 8 hours for administration. It

comprised those tests previously found most useful for prediction of ATCS training and job performance plus tests of a number of aptitude areas not previously examined. The tests yielded 44 different scores. Many of these tests were either commercially published instruments or aptitude assessment devices developed under contractual arrangement for the FAA. In fact, no Civil Service Commission (CSC--now known as the Office of Personnel Management, OPM) tests were included in the research prior to July 1961 (17). By that time a series of multiple-regression analyses, accomplished by Cobb and Trites in connection with followup studies of several hundred men, had identified a total of 8 tests (from a group of 27) from which a variety of summary measures having validity for prediction of ATCS trainee performance might be derived. Seven of the eight tests were commercially published instruments; the remaining one was the contractually developed ATP test.

While the seven published instruments of that experimental battery are referred to as "tests," they were actually parts, or subtests, taken from rather lengthy and comprehensive aptitude-measuring devices (17). Three of the seven were subtests of the Psychological Corporation's Differential Aptitude Test (DAT), viz, DAT Space Relations, DAT Numerical Ability, and DAT Abstract Reasoning. The first test measures the ability of an examinee to visualize objects and forms in two or three dimensions, the second is a test of arithmetical or computational skill, while the third provides a measure of nonverbal reasoning (specifically, determining, for each item, which of a series of choices (figures) properly carries out a principle of logical development exhibited by a sequence of figures). The remaining four were subtests of the California Test Bureau's Test of Mental Maturity (CTMM, Advanced Form A Edition)(45). In each item of CTMM Analogies, the examinee must recognize the relationship between a pair of drawings (objects) in order to identify, by analogy, one of four choices as being similarly related to a third. CTMM Inference involves the comprehension of statements that present premises underlying the derivation of logical conclusions. The subtest designated as CTMM Numerical Quantity-Arithmetic measures the ability to solve word-presented arithmetic problems, while CTMM Numerical Quantity-Coins involves the mental manipulation of interrelated amounts of money and numbers of coins.

No single group of trainees was ever administered all eight tests (the seven commercial tests plus Air Traffic Problems); one class was examined with seven, and several successive classes were administered either six or five, of the eight instruments (17). When an average of the academic and laboratory grades was computed for each Academy trainee and employed as the criterion, the validities of the composite measures, derived from the performance scores on five, six, or seven tests, were found to range from .35 to .54 (Pearson product-moment coefficients). More importantly, an analysis revealed that about 70 to 80 percent of the cases classified as training-course failures were, in most instances, represented in the lower half of the distribution of scores derived with each respective group of tests. Moreover, the Academy attrition rates were averaging over 30 percent despite the fact that most trainees were former military controllers.

The potential value of the tests for screening purposes was thus recognized, and, when similar results were obtained with additional samples, the FAA and the CSC agreed that aptitude-test measures should be employed on a tentative basis in the selection of some of the nonexperienced applicants (17). However, commercially published tests could not be used because such instruments were deemed more susceptible to compromise than those subjected to rigid CSC control procedures. Commission officials therefore examined their extensive file of CSC tests and selected several instruments that, in terms of factor content, appeared to approximate a number of the validated commercial tests. Since the ATP test had been developed specifically for the FAA and was still completely controlled, it was officially adopted as a CSC test. In addition, CAMI researchers were provided a number of other CSC tests, to administer and evaluate.

Commencing in August 1961, all incoming classes of Academy ATCS trainees were experimentally assessed with the entire group of tests extracted from the CSC files and with the ATP test. The restricted time available for each testing session precluded an examination of each class with the complete and previously validated battery of commercial tests. However, time beyond that required for the CSC tests was available to permit administration of a portion of the commercial battery. Followup studies of Academy trainees examined with the revised battery during the next 10 months revealed that composite scores based on five of the CSC tests and the ATP test could be used effectively to predict training outcomes (17). Composite scores of 190 and higher were attained by approximately 55 percent of all the examinees. Of these, about 70 percent successfully completed their training course and were certificated as ATCSs. In contrast, almost 75 percent of those with scores of 189 and lower failed to graduate and were eliminated from further FAA training (17). These results approximated those obtained in earlier analyses (with other groups) for the "commercial seven-test composite." Three of the tests involved in the six-test CSC composite were those that had been selected as "counterparts" of three commercial tests; these instruments provided measures of numerical, spatial, and nonverbal abstract-reasoning abilities. The new composite also included the ATP test (an extensive revision of the original test), an instrument known as Letter Sequence (a new test, not identified in the original research, which measured reasoning ability), and a test of following oral directions. The CSC battery is commonly referred to as "The CSC ATC Aptitude Screening Test" and its six elements as "subtests" (see Table 1).

Although CAMI was requested to continue its experimental testing program and obtain additional validation data (including, now, Flight Service Station (FSS) ATCS trainees for whom no previous experimental data were available), an early analysis, in which the aptitude test scores of the first 302 examinees were validated against the Academy training criteria, yielded findings (highly similar to those later obtained for the complete sample) that prompted CSC and FAA officials to authorize use of the battery, beginning in July 1962, in screening only those applicants unable to establish training eligibility in terms of the normally prescribed

Table 1. Description of CSC ATC Aptitude Screening Test Variables

(from Cobb, Lay, & Bourdet, 1971, ref. 18)

CSC Subtest	Weight	Description
CSC 24 Computations	1	A highly speeded test of arithmetic skill. The problems involve simple addition, subtraction, multiplication, and division. The aptitude factor is referred to as "numerical facility."
CSC 51 Spatial Patterns	2	A test consisting of two different types of spatial items. In one type, the task is to identify solid figures that can be made from unfolded patterns. In each of the other, three different views of an object are presented and the subject must select the correct object from one of four alternatives.
CSC 135 Following Oral Directions	1	In this test the subject must listen carefully to orally presented directions and information; then discriminate between relevant and irrelevant information in order to proceed toward the proper solution of a series of simple tasks.
CSC 157 Abstract Reasoning and Letter Sequence	2	In the "Abstract Reasoning" portion of the booklet, the task is to indicate which of a series of choices (figures) properly carries out a principle of logical development exhibited by a sequence of figures. In "Letter Sequence" the subject must indicate which of a series of letters properly carries out a principle of logical development exhibited by a sequence of letters.
CSC 540	1	A highly speeded test consisting of two parts of 30 items each. In each part, the subject is presented a flight data display for several aircraft and must determine whether certain changes in altitude may be directed without violating a specified time-separation rule.
CSC Composite or Total Weighted Scores		Sum of the weighted raw scores for each of the five CSC test booklets (i.e., six subtests).

qualification standards (i.e., qualifications with respect to aviation-related experience and/or education). Several thousand such applicants were operationally examined with the battery during the following 18 months and, although about half of them established training eligibility by achieving raw composite scores of 190 or higher (i.e., percentile scores of 70 or better), very few were selected (24). Candidates qualifying on the basis of previous ATC work and other aviation experience generally attained higher overall CSC eligibility ratings than those screened with the battery. Moreover, training quotas continued to decline and were usually met by selecting candidates having CSC percentile ratings no lower than 90. To attain a percentile rating of 90, an applicant having insufficient ATC-related experience to qualify for any credit points was required to achieve an exceptionally high aptitude test score of 257. Aptitude-screened applicants were, therefore, seldom able to effectively compete for the available training positions. In fact, most of the relatively few aptitude-screened candidates selected for training prior to January 1964 possessed at least some ATC-related experience that, although insufficient for exemption of the test screening requirement, warranted credit points to supplement ratings reflecting excellent levels of performance on the test battery.

Academy training performance records for the last of the 893 candidates examined with the CSC test battery for research purposes only were not available until October 1963. By that time, Cobb and Trites had completed several sets of analyses of Academy performance (see Figure 2) and had collected post-Academy information on training progress, experimental ratings of job performance, and other data for several hundred of the examinees who had successfully completed their basic training course some 12 to 18 months earlier.

A series of validation analyses completed shortly thereafter yielded findings of timely interest to officials seeking to improve ATCS selection (24). Perhaps the most important of the analyses was that which, in support of the preliminary findings obtained in the August 1961-May 1962 analyses, revealed that about two-thirds of the 271 Academy attritions among the 893 experimentally examined trainees scored no higher than 189 on the CSC battery, whereas two-thirds of the 622 graduates scored 190 or higher. In another analysis based on the entire sample, statistically significant ($p \leq .05$) correlations were obtained between the aptitude test variable and most of the Academy training-performance measures (i.e., grades). However, when dealing with data on the Academy graduates only, the aptitude-test scores (particularly those above 210) proved very unreliable for prediction either of Academy training performance or of promotions, ratings of job performance, and/or attrition-retention status during the first 12 to 18 months of facility training. Yet, the latter findings were of the general statistical type expected, because the deletion of data for the Academy attritions, the majority of whom were low-aptitude trainees, resulted in an attenuated distribution of test scores (24).

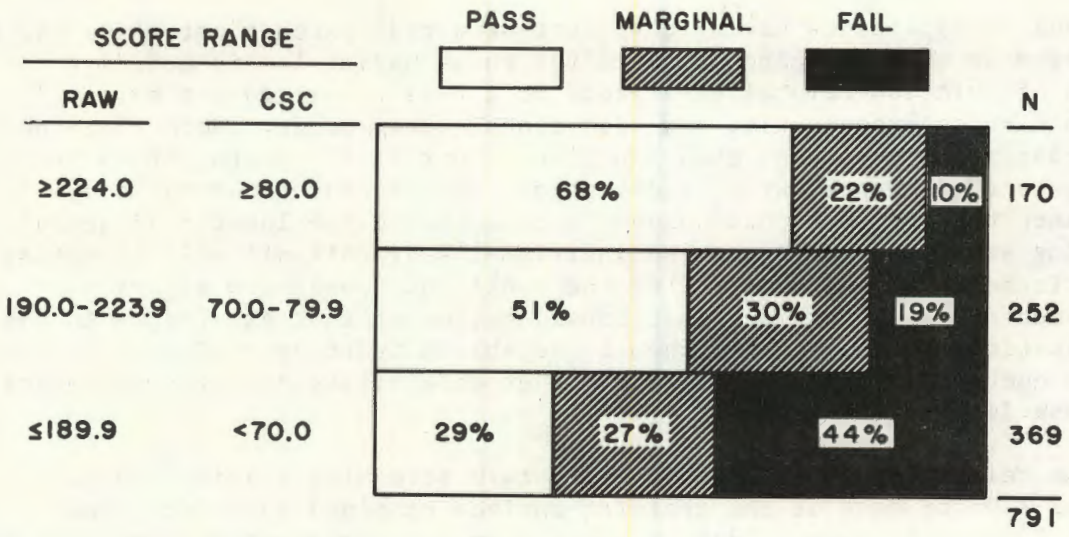


Figure 2. Percentages by three composite score ranges of passing, marginal, and failing trainees who took the CSC ATC Aptitude Screening Test experimentally during 1960-63: N=791 En Route, Terminal, and FSS trainees combined (from Trites and Cobb, 1964, ref. 50).

These results prompted a revision in the selection standards. Beginning in January 1964, the CSC battery was incorporated in the screening of all applicants regardless of their pre-FAA experience. Aside from other factors, eligibility required a composite CSC score of at least 210. Retention of a screening score of 190 was considered, but 210 was eventually adopted because it was contemplated by FAA officials that a further reduction in the number of ATCS trainee positions would be necessary pending an increase in congressional appropriations (24).

Revised Aptitude-Screening Procedures. Although the screening program instituted in January 1964 required that all applicants for ATCS training be examined with the CSC battery, they were screened in terms of three different aptitude-test-performance qualification standards (24). In accordance with procedures prescribed for each specific training option and entry pay grade being applied for, an individual's preemployment experience and/or level of education determined which of three tables was to be used in converting the applicant's composite raw score on the test battery to a percentile score. A percentile score of 70, considered a mandatory eligibility requirement, corresponded to a raw score of 210 on one conversion table, 225 on another, and 240 on the third. For example, candidacy for Tower or Center training at the GS-6 (General Schedule rating, i.e., pay grade) entry level required that ATC-rated applicants (usually former military controllers) score at least 210 on the battery, whereas 225 was considered minimally qualifying for instrument-rated, aircraft-pilot

personnel or applicants having navigator or air-dispatcher certificates, and 240 served as the screening standard for those having low-to-moderate amounts of aviation-related experience or a 4-year college degree, or certain diverse experiential and educational backgrounds. Moreover, the procedures stipulated that each applicant's percentile score, if 70 or higher, be supplemented with credit points awarded for those types of experience that had warranted use of either of the two lower test-score screening standards to derive the individual's overall eligibility rating (24). Inasmuch as the majority of the applicants possessed aviation experience of some sort, the dual consideration of that experience in the qualification process enabled them to establish training candidacy in far greater numbers, and generally with higher eligibility ratings, than those with nonaviation backgrounds.

The relatively high aptitude-test-score screening standards of 210, 225, and 240 for each of the training options remained in effect from January 1964 until August 1968, resulting in the screenout of more than half the applicants. However, no shortage of qualified candidates developed. Due to continuing budgetary limitations, training quotas remained unusually low throughout the entire 56-month period and consequently nearly all ATCS trainees were selected from among candidates having very high eligibility ratings--reflecting exceptional qualifications with respect to aptitudes, experience, and/or education.

Eligibility for ATCS training within the FAA has always been restricted to those applicants who, after having met other qualification requirements, satisfactorily pass a rigid medical examination. However, the procedures relating to medical certification of ATCS personnel were amended in 1966 to include consideration of personality attributes (22). Since that time, Cattell's Sixteen Personality Factor (16 PF) Questionnaire (9) has been used in the screening of all applicants. Physiologically qualified individuals for whom the 16 PF Questionnaire results indicate no significant emotional or mental problems are granted medical certification. Others, who usually represent but a small minority of the applicants, must submit to a psychiatric examination, the results of which may constitute grounds for ineligibility.

In August of 1968, a program for rapid expansion of the National Airspace System was initiated and a revised set of trainee-selection standards was adopted. In many respects, the new selection program was highly similar to that of the preceding 56 months. However, two new screening standards were implemented as a means of insuring an adequate supply of candidates (24). First, according to one of the new standards, applicants having highly specialized ATC experience (particularly in radar control) could be granted waivers of the aptitude-screening requirement and could also be appointed to training at pay grade GS-9 or higher rather than the normally prescribed entry grade GS-7 or lower. It was reasoned that such personnel would be able to complete developmental training more rapidly than others and thereby more quickly alleviate the shortage of full-performance-level controllers. Second, a score of 210 on the CSC test

battery was adopted as a common screening standard for most other applicants. The screening standard of 210 applied to (i) former military controllers unable to qualify under the "specialized experience" standard, (ii) pilots, navigators, air dispatchers, and others who would have confronted a test-score screening hurdle of 225 if they had applied in earlier years, (iii) 4-year college graduates with records of superior academic achievement, and (iv) applicants having master's degrees. A test score of 240 was prescribed for use in the screening of few applicants except those devoid of aviation-related experience, most of whom were college graduates with no evidence of superior academic achievement (24). The "specialized experience" standard remained in effect until April 1972. Throughout that time, however, less than one-fourth of the ATCS selectees entered as GS-9's or higher with waivers of the aptitude-screening requirement (22).

Only 710 (18.9 percent) of 3,751 trainees who arrived at the FAA Academy during November 1968 and the ensuing 17 months for basic ATCS training claimed to have entered the FAA as GS-9's or higher on the basis of highly specialized ATC experience (18). Some 446 of the 710 enrolled in the Academy's basic En Route course, while 264 entered the Terminal course. A study revealed that 16 percent of the 446 En Route trainees of GS-9 level and higher failed to successfully complete the En Route course and that the attrition rate of the 2,526 who enrolled in the same training course as GS-7's or lower was only slightly higher (i.e., 18.2 percent). Only 14 percent of the 264 Terminal trainees recruited with disregard of the aptitude-screening requirement failed the Terminal course, whereas a significantly higher ($p < .01$) elimination rate of 21.9 percent was obtained for the remaining 515 Terminal students (22). However, a later CAMI study (19) by Cobb, Mathews, and Nelson (in which December 1, 1971, served as a common date for determination of the attrition-retention status of every student who successfully completed either En Route or Terminal basic training at the Academy during 1969) indicated that the trainees selected under the "specialized experience" standard had slightly higher post-Academy attrition rates than those appointed to training with pay grades of GS-7 and lower. The difference between the post-Academy (i.e., facility-training) elimination rates of the two differentially selected Terminal subgroups was somewhat greater than that obtained between the En Route subsamples, but neither difference was statistically significant. Had the results not been confounded by aging effects, Cobb and his associates would have concluded that specialized ATC experience was of little or no value to most trainees after Academy graduation. However, almost 23 percent of the higher graded trainees of the combined En Route and Terminal options were 35 years of age or older, whereas slightly less than 14 percent of those appointed as GS-7's or lower were older than 34 (22).

Unpublished research by Cobb, involving several hundred ATCSs who had successfully completed either En Route or Terminal basic training at the Academy in 1969, revealed highly significant differences between the post-Academy attrition rates (as determined on December 1, 1971) for trainees aged "35 and older" vs. those "34 and younger." For Academy

graduates of GS-9 level and higher, the facility-training attrition rates were 42 percent and 17.5 percent for the older and younger subgroups, respectively (22). About 25 percent of the ATCSs of GS-7 level and lower who were over 35 years old were attrited after returning to their home facilities, whereas the post-Academy elimination rate of the younger ATCSs having similar pay grades was only 18 percent. Moreover, several earlier studies (15,20,47,50) had consistently shown chronological age to be inversely related (at highly significant levels) to scores on numerous aptitude tests, various indices of Academy training progress, and ratings of journeyman-level job performance.

As early as 1965, it was the view of some FAA officials that a special early-retirement program was needed for controllers and that the recruitment of ATCS trainees should be restricted to those qualified applicants who were relatively young (22). However, such proposed policies ran counter to the CSC regulations pertaining to all Federal service employees except those specifically exempted by congressional legislation. Research concerning age-related effects upon ATCS performance was intensified and, in 1972, the cumulative body of findings prompted congressional legislation authorizing the FAA and CSC to further develop and implement a proposed ATCS "Second-Career Program." The congressional bill, Public Law 92-297, became operationally effective on August 14, 1972. Since that time, ATCSs receive credit for 1.4 years of Federal service for each year of active control work; the normally prescribed minimum-age requirement of 55 does not apply to control personnel; early retirement is not mandatory but retention as an ATCS requires maintenance of job proficiency, and ATCSs are also offered training for other jobs (i.e., "Second-Career Training"). Moreover, as a means of countering aging and "burnout" effects, the FAA was permitted to establish a screening standard with respect to age.

Programs Designed to Aid Minorities and Women in ATCS Selection and Training. During the 1960's, Government-wide awareness of the social need to provide opportunities for certain economically and culturally disadvantaged groups prompted the FAA to implement a new source of recruiting for the ATCS occupation.

The new project was originally proposed in an FAA Organizational Bias Seminar, and was committed by the Administrator to the Secretary of Transportation in 1968. It was termed the "150 Program." The purpose was to provide for entry into the ATCS occupation at the GS-4 level rather than the previous GS-5 minimum accession level. The Predevelopmental program was designed primarily for persons of underprivileged backgrounds, and as a means of hiring more minorities in ATCS positions, which constitute about half of all FAA positions.

The program involved a training agreement with the CSC, whereby a number of positions were allotted on a 6-month recycling basis. Special 6-month qualifications training was given at the FAA Academy upon the trainee's entrance on duty. The courses were designed to provide the fully qualified GS-4 ATCS with a general background of aviation knowledge prior to entering the Academy.

On October 1, 1969, the CSC approved training agreements covering ATCSs for a 2-year duration. This allowed for accelerated promotion from the GS-4 entrance level to GS-5 after completion of 6 months of Academy training (30).

Since inception of the "150 Program," evaluation has taken place in several ways. The first 21 ATCS trainees who entered the Academy on February 9, 1970, were compared by the FAA Office of Personnel in biographical background with proposed course content, and based on this comparison course refinements were introduced. Those classes were graduated in July 1970. End-of-course critiques completed by trainees and instructors were evaluated, and further critiques were mailed and completed at field installations where graduates were subsequently assigned in February 1971. Several refinements in the program were made as a result of the initial evaluation process (29).

In early 1971 the FAA Office of Personnel conducted an informal evaluation that was based primarily on discussions with students, faculty members, and regional personnel staff. Refinements were made in recruiting efforts as a result of this study. In September 1971 a more formal evaluation of the "150 Program" was performed by the Office of Personnel (29). The general conclusion of the study was that the program "provided an additional rung in the ladder" for a high percentage of minorities who otherwise would not have become agency employees. A later study (1973) by the same office verified this general conclusion (30).

Boone (4) performed a path analytic study of the "150 Program" to determine if the program indeed had a direct impact on the trainee's ability to achieve success in Academy training. The two previous informal studies cited above (1971 and 1973) had shown that the "150 Program" resulted in a higher percentage of women and minorities in ATC work; however, no explicit evaluation had been performed to demonstrate that the increase in the number of women and minorities performing ATC work resulted from the instruction received in the program rather than from a mere increase in the number of women and minorities recruited and hired because of the "150 Program." The participants in the study consisted of all persons who came through the Predevelopmental program during calendar years 1974 through 1976. The effect of the program on Academy success was reviewed for nonminority men and women and for minority men (there were too few minority women to form a basis for analysis). The general conclusion of the study was that the Predevelopmental program overall aided the disadvantaged to achieve success in the FAA Academy. The general effect, however, was not true for the subgroups. The path models indicated that nonminority men and women (especially women) were aided by the program; however, minority men (predominantly Blacks) were not (4). The study suggested that recruitment and selection testing procedures could have produced this differential effect. A further study of attrition in Predevelopmental recruits was recommended.

Current Status of Selection Standards. The ATCS selection procedures remained essentially unchanged from August 1968 until April 1973. At that time, selection at the GS-9 level on the basis of "specialized experience" ceased (April 2, 1973). Simultaneous with the reinstatement of mandatory aptitude-screening procedures for all applicants, the selection procedure that automatically precluded the eligibility of any applicant more than 30 years old was implemented (24). With the exception of the two aspects just mentioned, the 1973 selection program was highly similar to that of 1968. Also, none of the current (1979) screening standards differ significantly from those of 1973. Nonetheless, recruiting demands have so drastically declined since 1968 that a current-day candidate has virtually no chance of being appointed to training unless his or her overall eligibility rating far exceeds that which would have warranted selection during the 1968-72 period. In addition, the Academy training program was modified in 1976 to include new pass/fail criteria based on a curriculum that assumed zero ATC knowledge on the part of the new trainees. CAMI was assigned the task of maintaining the longitudinal data base for all trainees beginning in 1976. That computerized base includes CSC (now OPM) selection test scores, biographical information, all Academy training scores, and subsequent field performance data from Academy graduates.

CAMI is also active in current efforts to update and improve ATCS selection procedures. These efforts are partly rooted in recommendations by a task force which was commissioned by the FAA in December 1974 to review the agency's selection policies. The task force, in a 1975 report (26), identified the following areas for further consideration: (i) the testing and screening of applicants for ATC work, (ii) the CSC rating guide used to grant additional points for certain types of related prior experience, and (iii) the evaluation of further recruitment and testing practices for cultural bias against women and racial minorities. As a result of that review, two major studies were performed with the goal of further improving ATCS selection through revised aptitude testing.

The first of these two studies, by Education and Public Affairs (EPA), a private company, was contracted by FAA (25,27,36) to determine the potential of an experimental test battery to predict ATCS success. An aggregate "success" criterion was used in the study, based on a composite of supervisory assessment and career progression. The experimental tests considered were the present ATC selection battery, an ATC General Information Test, ATC Occupational Knowledge Test, Multiplex Controller Aptitude Test, Directional Headings Test, Dial Reading Test, and an Arithmetic Reasoning Test. The EPA study (36) provided data-based predictions but was not able to demonstrate the relative value of the experimental tests vis-a-vis the present battery, since no information was available to EPA on the experimental tests from the total applicant group.

To evaluate the relative value of the tests that demonstrated potential in the EPA study, the CSC in 1977 administered two of the experimental instruments, the Multiplex Controller Aptitude Test (MCAT) and the Directional Headings Test (DHT), to approximately 7,000 ATCS applicants in

conjunction with the regular CSC ATC test battery. These data were to be used along with the analyses performed in the second major study which was conducted at CAMI by Boone and Lewis (5,6,7). The purpose of the CAMI effort was to determine which of the selected experimental tests, either independently or in combination with present CSC tests, were the best predictors of success at the FAA Academy. CAMI data came from trainees selected for ATCS work beginning May 1976 through April 1978. During their first day at the Academy, new trainees were tested by Boone and Lewis with experimental test batteries which included a Biographical Questionnaire, a Dial Reading Test (DRT), the MCAT, and the DHT. Only ATCSs who had a complete data set were included in the final sample of 1,828 trainees.

The DRT used in the CAMI study is Part I from the U.S. Air Force's Dial and Table Reading Test (32); the examinee is presented with seven dials for each set of six questions and is required to read the correct value on the correct dial in order to select the correct answer from among five given alternatives. The MCAT (28) presents pictures of simulated air traffic crossing a controller's display, and a table of altitudes, speeds, routes, and identifiers. The test is speeded to reflect changes in aircraft position with successive items. A primary task is to predict violations of aircraft separation standards. Other questions measure aptitudes such as table reading, spatial visualization, and arithmetic reasoning. Items alternate from one type to another with increasing levels of difficulty. The DHT (19) is a highly speeded test with items that present the examinee with one, two, or three pieces of information reflecting the cardinal points on a mariner's compass. In Part I of the test the examinee must determine very swiftly if the information conflicts or agrees. Each item is followed by five response categories: "North?," "East?," "West?," "South?," and "Conflict?," and the examinee must mark one of the response spaces under the correct category. Part II is similar except the examinee must mark the response space for the exact opposite of the heading in the item.

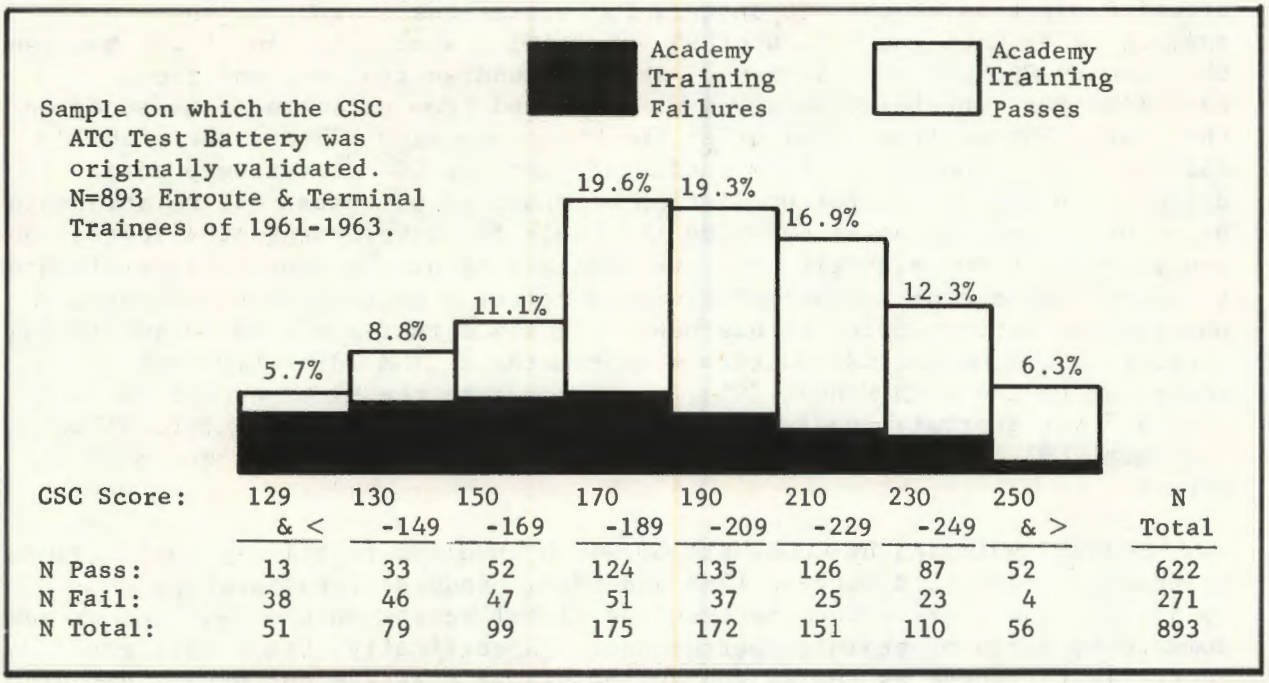
Scores on the DRT, MCAT, and DHT and on the tests composing the present CSC battery were correlated by Boone (5) with the averaged laboratory scores from Academy training. These correlations were then employed in an iterative stepwise regression (stepdown procedure). The tests that made a significant contribution in predicting Academy scores were then used to form a composite, and the multiple correlation was computed for the old test battery and the new test battery. Six regression models were explored; two of them demonstrated a statistically significant increase in the multiple correlation over the old test battery (5). Of the two models, one involved use of the MCAT with two current CSC tests (CSC 24, Computations; CSC 157, Abstract Reasoning), the other involved no current CSC tests but included the MCAT, DHT, and DRT. A report of the statistical findings was prepared in 1979 for OPM and FAA consideration (5). Final action on a new aptitude selection battery is currently pending.

Summary. Active research at CAMI on the problem of ATCS selection was undertaken soon after the establishment of the Institute in 1960. Under the

direction of Trites and Cobb, followup research immediately began on the predictive validity of the composite of aptitude tests originally studied by Brokaw. After further exploratory research on a varietv of aptitude measures, the validity of several such measures for predicting success in training was fairly well established. Further research at CAMI with CSC tests of similar content led the CSC for the first time to establish selection tests for screening ATCS trainees in July 1962. Continued validation research prompted a revision of CSC selection standards in January 1964, with all applicants being required to take the selection battery. However, prior experience and education continued to be the primary determinants of the selection decision. The high standards (and concomitant low selection rates) established in 1964 were revised in 1968 due to an expansion of agency hiring. During the period from 1962 through 1972, in addition to continuing validation research on the changing CSC criteria, Cobb and his associates at CAMI investigated a number of factors relevant to selection and screening of ATCSs. Those factors, which are discussed in more detail below, included research on attrition, age, prior experience, education, sex, and military ATC training. The CAMI research on age led directly to the establishment of the age 30 screening standard in 1973. At that same time, mandatory aptitude screening was reinstated (those with prior radar experience had been exempted in 1968). Since that time the selection standards have remained essentially the same. During the latter half of the 1970's, CAMI research focused on evaluation of the impact of the special entry "150 Program" for disadvantaged persons which was implemented in 1969, on studies of optimal combinations of old and new aptitude screening measures to form a potential new aptitude screening battery, on studies relevant to the Uniform Guidelines on Employee Selection (discussed below), on developing field performance ratings, and on the computerized, longitudinal ATCS data base for continuing validation research.

Some Followup Studies on Attrition

It had been ascertained in CAMI's primary validation study with the 893 pre-1964 ATCS trainees that Academy-basic-training attrition rates were: (i) 30.3 percent for the total group, (ii) only 16.4 percent for the 317 who attained composite raw scores of 210 and higher on the CSC test battery, and (iii) 38 percent for the group of 576 having scores of 209 and lower (see Figure 3). Thus, it had been anticipated that adoption of the CSC battery for screening purposes could result in a reduction of the attrition rate at the Academy from approximately 30 to about 16 percent. A reduction did occur but not to the extent expected (19). For example, 22 percent of the 2,822 men who presumably met or exceeded the aptitude screening requirement and who entered the Academy during November 1968 through March 1970 failed to successfully complete their training course. While the attrition rate of 22 percent was significantly lower than the 30.3 percent obtained for the total validation sample of 893 trainees, it was somewhat above the failure rate found for the 317 who had aptitude scores greater than 209. The disparity between the attrition rate predicted in 1964 and the rate obtained



CSC Score:	129 & <	130 -149	150 -169	170 -189	190 -209	210 -229	230 -249	250 & >	N Total
N Pass:	13	33	52	124	135	126	87	52	622
N Fail:	38	46	47	51	37	25	23	4	271
N Total:	51	79	99	175	172	151	110	56	893

Figure 3. Distributions of CSC ATC Aptitude Screening Test scores for the 893 experimentally examined entrants into Academy ATCS training during 1961-63 and for the 281 operationally examined trainees who entered during 1970 (from Cobb and Mathews, 1972, ref. 19).

in 1970 prompted Cobb and Mathews (19) to examine possible reasons for the obtained rate.

Two assumptions were made (19). First, an aptitude screening battery predicated on such a vast and diverse amount of research as had been accomplished by CAMI would have been expected to maintain considerable effectiveness over time as an initial screening device. Second, inasmuch as the CSC adheres to rigid control and accountability procedures, there was little likelihood that the aptitude measuring devices would have been subjected to compromise.

The most likely reason selected by Cobb and Mathews (19) for the disparity between the predicted and the obtained attrition rates was the impact of FAA ATCS selection practices on the data used to determine validity of the selection test battery--data based on such highly screened personnel were apt to be grossly attenuated. In addition to the medical, experimental, aptitudinal, and educational screening qualifications of ATCS applicants, several unpublished CAMI studies indicated that ATCS trainees were also well above average in intelligence (19). One investigation, which

preceded adoption of the CSC battery for operational purposes, was summarized by Cobb and Mathews (19) as yielding a correlation of .81 between the summary CSC aptitude scores of several hundred trainees and their so-called IQs (intelligence quotients) derived from performance measures on the CTMM. The aptitude scores for the group averaged 193 and the mean IQ was 110.5. Although the study indicated that the CSC scores were more useful than the latter for prediction of training outcomes, the relationship between the two variables provided the basis for development of a regression equation which subsequently permitted estimation of the general intelligence level of operationally examined recruits relative to that of the general population. For example, it has been determined that a minimally qualifying score (of 210) on the CSC battery approximates a CTMM IQ of 120, which, according to the CTMM Manual (45), corresponds to the 88th percentile for a "normal" (or general) population and about the 70th percentile for college freshmen (19). Thus, these trainees also represented a bright group of people.

Several analyses by Cobb, all of which involved relatively small groups of trainees recruited between 1968 and 1970, produced (predictably) relatively low correlations between the global scores on the CSC battery and summary measures of training performance. Specifically, the validity coefficients ranged as low as .04 and no higher than .26 for prediction of the pass/fail training criterion and from .03 to .39 for training grade averages (19). These and other similar findings, however, were expected on the basis of restriction of range and other factors noted below.

If it were assumed that ATC training performance evaluation standards remained unchanged from 1963 to 1970, a failure rate of somewhat less than 22 percent would have been anticipated for the later trainees. While little or no evidence was available to conclusively prove or disprove equivalency of the performance criteria, the issue was one frequently confronted in research aimed at progressive improvement of personnel selection procedures. It is commonly referred to as "the problem of creeping criteria," and, inasmuch as officials responsible for training traditionally strive to upgrade their programs, it is a problem that is normally and hopefully anticipated.

Cobb and Mathews (19) hypothesized that ATC training programs and performance evaluation standards had been upgraded; that the changes (the exact nature and extent of which would have been difficult to determine reliably) were partially attributable to training requirements associated with the increasing complexity of the air traffic management system and in part to the inherent tendency of training personnel to "customize" instructional material, instruction, and/or subjective performance evaluation standards relative to the general level and range of mental abilities represented in the incoming classes. Moreover, the vast majority of trainees recruited since 1963 were of high mental caliber and, thereby, comprised groups which would be characterized by substantial restriction of range of aptitudes (19).

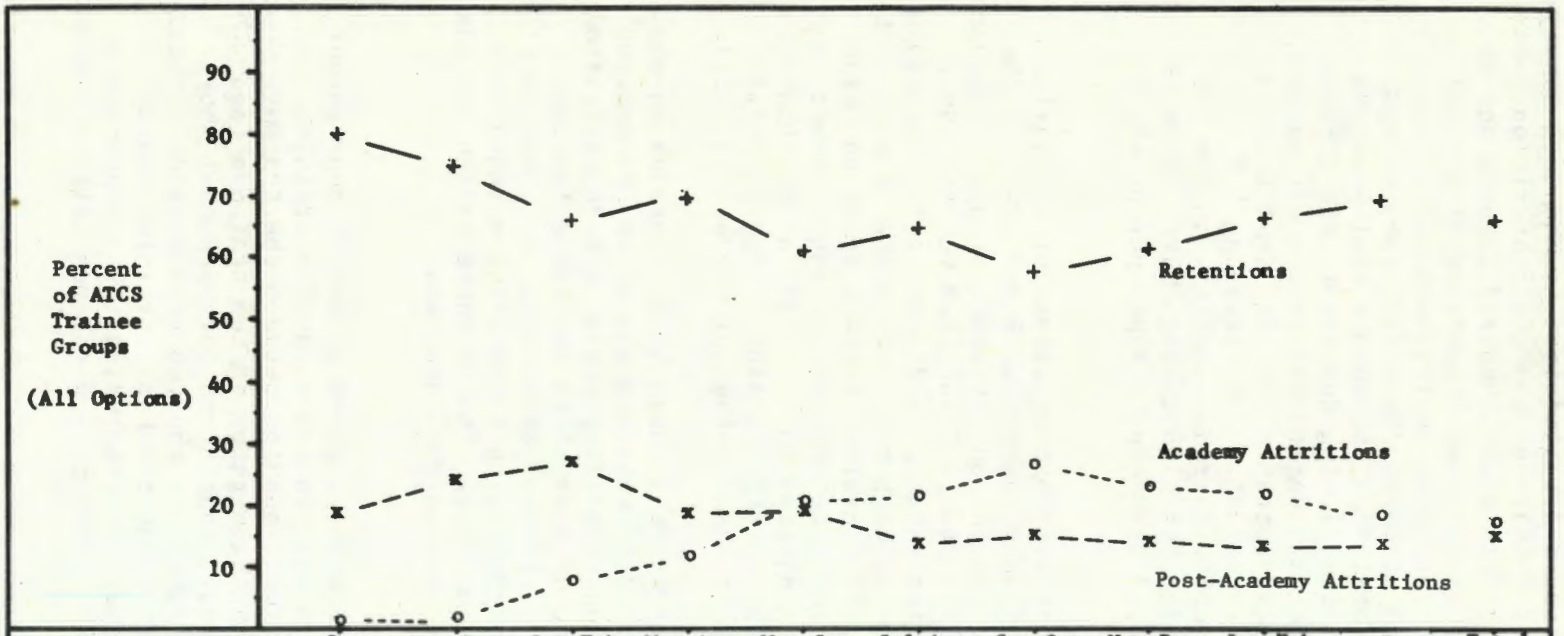
Cobb, Mathews, and Nelson (21) took a detailed look at the various ways in which the characteristics of the ATCS trainees and their attrition rates differed between 1960-63 and 1968-70. This study involved comparisons of Academy basic training elimination rates and post-Academy attrition and retention rates of personnel who were recruited during each of the two widely separated time periods for each of three different types of ATCS training. Data were obtained for a total of 6,367 former trainees. Exactly 2,000 of the 6,367 had entered the Academy during September 1960 through August 1963, before the CSC battery became operational in the screening of most applicants. The 2,000 included 733 entrants into Terminal training, 1,008 En Route trainees, and 259 FSS personnel. The remaining 4,367 trainees, the vast majority of whom were required to qualify on the CSC test, had entered the Academy during October 1968 through March 1970; of the 4,367, 935 were Terminal trainees, 3,159 were En Route personnel, and 273 were FSS personnel.

Percentages reflecting the Academy elimination rates for the earlier versus the later time periods, respectively, were: 20.9 and 19.3 for the Terminal personnel, 32.0 and 17.9 for the En Route trainees (this difference was statistically significant, $p < .01$), and 18.5 and 12.8 for the FSS personnel. The mean Academy elimination rate of 26.2 percent for the 2,000 total pre-1964 trainees was significantly higher ($p < .01$) than the 17.0-percent rate obtained for the 4,367 recruits of the more recent time period (21). Moreover, the 17.9-percent rate was only slightly higher than a rate of 16.4 percent that Cobb had projected on the basis of results obtained in an earlier study of the 893 pre-1964 trainees experimentally examined with the CSC battery (before it was adopted for operational use).

Followup procedures were employed whereby those Academy graduates who were still in ATC work on December 1, 1971, were designated as "retentions" while those eliminated after completion of Academy training were designated as "post-Academy attritions." The post-Academy attrition rates for the Terminal, En Route, and FSS entrants of 1960-63 were 16.0, 22.8, and 18.1 percent, respectively, whereas the corresponding rates for the recruits of 1968-70 were 10.1, 20.3, and 5.9 percent (21). The En Route option was the only one for for which the difference between the rates was not statistically significant.

Analyses pertaining to subgroups, each composed of one to four former incoming classes of the Academy's Terminal, En Route, or FSS training courses, yielded significant inverse relationships between the Academy elimination rates and post-Academy attrition rates of the trainees (see Figure 4). Such findings were remarkably consistent for personnel who entered each type of training during either of the two time periods. These findings also suggested the possibility that the basic training courses could have been utilized more effectively in precluding the advancement of many recruits to facility training from which they were eventually eliminated (21).

An investigative phase pertaining to 645 of the Academy Terminal graduates and 2,162 En Route basic training course graduates of January 1969



	Oct 1968	Nov-Dec 1968	Jan-Feb 1969	Mar-Apr 1969	May-Jun 1969	Jul-Aug 1969	Sep-Oct 1969	Nov-Dec 1969	Jan-Feb 1970	Mar 1970	Total
Retentions (Dec. '71)											
N Cases	120	326	102	200	538	356	166	208	588	233	2,837
Percent of Group	79.5	74.6	65.4	69.4	60.0	63.7	57.6	61.0	65.1	67.0	65.0
Post-Academy Attritions											
N Cases	29	107	42	54	171	78	45	52	121	51	750
Percent of Group	19.2	24.5	26.9	18.8	19.1	14.0	15.6	15.2	13.4	14.7	17.2
Academy Attritions											
N Cases	2	4	12	34	187	125	77	81	194	64	780
Percent of Group	1.3	0.9	7.7	11.8	20.9	22.4	26.7	23.8	21.5	18.4	17.9
Total N of Group	151	437	156	288	896	559	288	341	903	348	4,367

Figure 4. Percentages of entrants into the Academy's En Route, Terminal, and FSS courses during October 1968 through March 1970 who were Academy attritions, facility-training attritions, or were still in FAA ATC work in December 1976 (from Cobb, Mathews, and Nelson, 1972, ref. 21).

through March 1970 revealed that the post-Academy retention rate of the Terminal trainees was 87.8 percent and that the En Route retention rate was signficantly lower at 75.7 percent (21). A series of analyses in which no distinction was made with respect to the trainees' pay grades or entry-into-training qualifications indicated that the retention rates varied appreciably from facility to facility. However, the retention rates did not appear to be related, positively or negatively, to the traffic-density levels of the facilities. For example, the mean retention rate of assignees to the 12 Level-IV (or top-ranked) Terminal facilities was 89.2 percent, whereas the retention rates of trainees averaged 87.9 at all Level-I and Level-II facilities while the mean was 87.0 percent at the Level-III installations. Although significantly lower than those obtained for Terminal personnel, the mean retention rates of En Route personnel at the 13 Level-II Centers (75.4 percent) and 14 lower ranked Centers (75.8 percent) were highly comparable (21).

The study also revealed that personnel who qualified for entry into training at the GS-9 level and higher on the basis of pre-FAA specialized experience had significantly higher Academy graduation rates than the conventionally recruited trainees. However, the post-Academy attrition rates of the higher rated trainees averaged about twice that of the less experienced trainees at the Level-IV and Level-III Terminal facilities; also, their mean attrition rate at the 13 top-ranked Centers was higher, though not significantly so, than that of other trainees. It was only at the Level-I Centers and Terminal facilities of the two lowest levels that the mean attrition rates (or mean retention rates) reflected favorably on these more experienced trainees--and by only one percentage point (21).

Although other types of occupations that make use of one or more sets of initial screening tests as well as formal centralized training are not directly comparable to air traffic control work (nor to each other), their attrition rates were examined to provide at least some perspective from which ATC Academy attritions of 1968-70 might be viewed (21). For example, published reports available at that time (circa 1970) about several occupational groups that were carefully prescreened prior to entry-into-training revealed training-attrition rates ranging from 22 to 43 percent. Included among these were attrition rates of 35.6 percent during United States training of Peace Corps volunteers (31); 22 percent of nursing students in their first year of training (33) (a different study (44) of nursing turnover in a teaching hospital showed that 38 percent of the nurses stayed on the job less than 1 year); 43 percent of Army officer candidates during 23 weeks of training (42); and 35 percent of a sample of naval aviation students (excluding those who failed to complete training due to medical or disciplinary problems) who, in addition to prescreening, were typically recent college graduates (2).

Since ATCSs make up about half of the FAA work force, some comparisons of their annual attrition rates with rates of non-ATCS personnel provided another occupational perspective. The same caveat noted above with regard to comparing different types of occupations (as well as different GS-grade

levels and other characteristics), of course, applies to these comparisons. However, based on FAA personnel records, Cobb, Mathews, and Nelson (21) calculated the average attrition rate for all FAA employees not in the air traffic control occupational specialty to have been approximately 13.5, 9.3, and 9.1 percent for calendar years 1969, 1970, and 1971, respectively (averaging 10.6 percent for the 3 years). The rates during the same time periods for all ATCSs were 5.7, 7.8, and 5.0 percent (averaging 6.3 percent). If the assumption is made that those who were attrited from the FAA Academy would have been later ATCS attritions anyway, then, following selection and Academy screening, the attrition rates of ATCSs would have dropped to approximately 3.3, 5.6, and 4.9 percent, respectively, for the 1969 through 1971 periods (a 3-year attrition rate averaging 4.6 percent).

These 3-year attrition rates were also examined by means of longitudinal comparisons between ATCSs and noncontroller FAA personnel (21). Data were made available by FAA's Office of Manpower regarding all full-time GS employees (i.e., excluding temporary or part-time help) hired during the 6-month period from October 1968 through March 1969. During that period, 949 ATCSs (including many who entered the Academy during April or later) were hired at the GS-5 through GS-9 levels; a total of 217 non-ATC employees were also hired at those same levels. The overall attrition rate (as of December 1971) for the 217 noncontrollers was 32.7 percent; for the 949 ATCSs, the rate was 24.6 percent (including Academy failures).

Summary. CAMI research on ATCS attrition demonstrated that the overall 1968-70 attrition rates compared favorably with attrition rates in other occupations utilizing highly competitive prescreening prior to entry-into-training as well as in non-ATCS FAA occupations. Other ATCS attrition research demonstrated that:

(i) The 1964 change in CSC entry standards significantly reduced attrition rates;
(ii) Attrition rates were negatively related to CSC composite scores;
(iii) "Special entry" ATCS trainees had significantly higher Academy graduation rates; however, they also had higher post-Academy attrition rates;
(iv) Attrition rates were not related to the air traffic density characteristics of a facility; and
(v) Attrition rates at the Academy were inversely related to post-Academy attrition rates.

Thus, initial selection screening and Academy training-screening appeared to be appropriately related to attrition rates by 1970, although overall attrition rates remained high and were costly to the agency. Additional attrition research, detailed in the next section, led to revised CSC selection standards aimed at reducing attrition.

Age as a Selection Factor

Perhaps the most important of all the CAMI studies on ATCS selection were those demonstrating the effects of age on training and job performance.

Training records for groups of ATCS personnel recruited since 1960 and a number of unpublished earlier CAMI studies consistently revealed a definite relationship between chronological age and probability of attrition.

To enter ATCS training at the time CAMI began its research program, an individual had to be at least 21 years old, but there was no upper age limit. It had long been felt by training and supervisory personnel even in 1960 that a marked negative relationship existed between age at entry into training and subsequent performance (50). That is, the older trainees did not perform as well as their younger classmates either in training or subsequently.

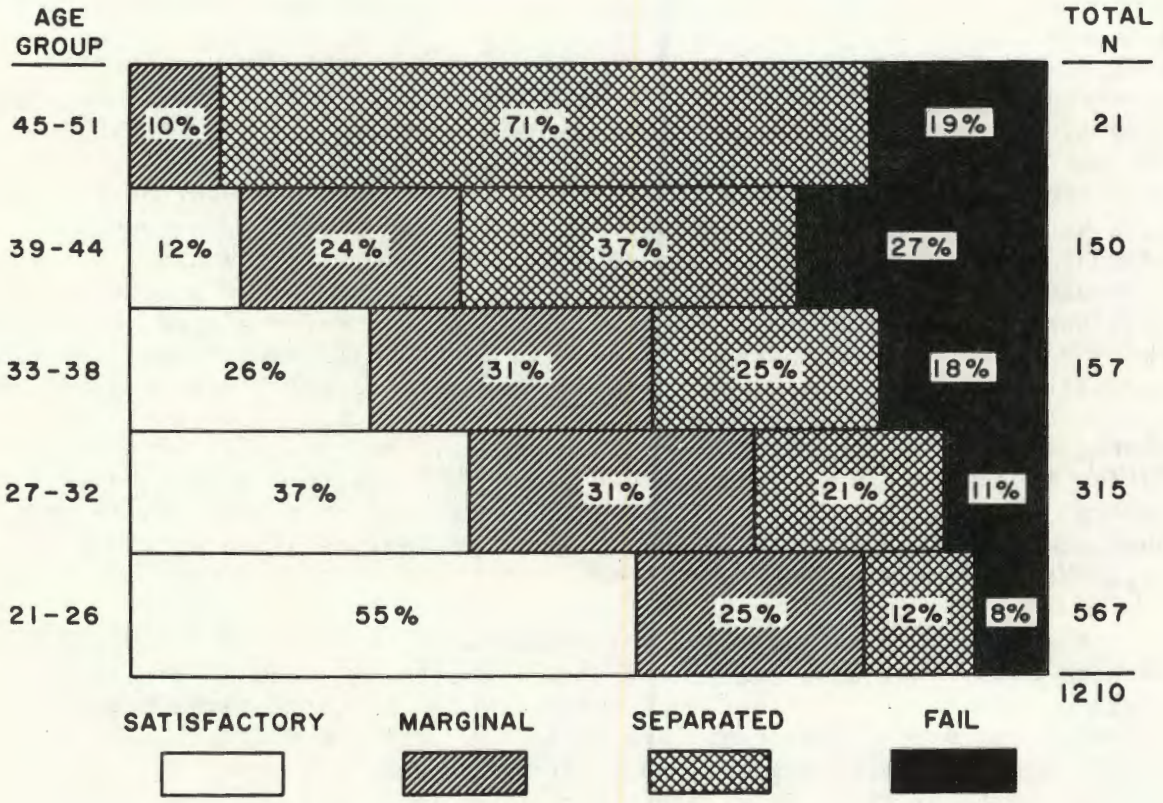


Figure 5. Percentages of ATCSs by age groups in four training and job performance categories: Satisfactory, marginal, separated, and failed (from Trites and Cobb, 1964, ref. 50).

In CAMI's first assessment of this relationship, the data given in Figure 5 were compiled by Trites and Cobb (50) from several different samples of ATCS trainees. In the figure, "Fails" are those not completing training, "Separated" are individuals completing training but no longer with the FAA, "Marginals" are individuals still with the FAA but about whom a job supervisor had some reservations concerning adequacy of job performance or

potential, and the "Satisfactory" individuals are those satisfactorily working at their ATCS specialty. The negative relationships between age at entry into training and job performance are obvious. The older the trainee, the less likely he or she was to complete training or remain with the FAA (50). In an unpublished 6-year followup study conducted by Cobb on 688 trainees (who comprised the successive classes of the Academy's basic training courses during August 1961 through March 1963), it was found that (i) 72 percent of 118 men of age 36 and over failed their initial training course whereas only 33 percent of 570 younger students failed, and (ii) older trainees who did pass the Academy also tended to experience greater difficulty than their younger colleagues in subsequent phases of on-the-job training (18). In a later published study that included these ATCSs and later trainees, entry age proved to be inversely related to both the aptitude-test variable and the criterion measures (15). Some 668 (74.8 percent) of the 893 examinees were no older than 30 and the remaining 225 were almost equally distributed among the age brackets 31-35, 36-40, and 41-and-older. Although only 148 of the 668 youngest trainees were Academy attritions, 63.5 percent (N=94) of the 148 scored 189 or lower on the CSC test battery, compared to 35.8 percent (N=186) of the 520 youngest graduates. Trainees of the successively older age brackets had progressively higher Academy attrition rates and lower mean scores on the test battery. Almost 78 percent (N=59) of the 76 trainees aged 41 and older (many of whom were military retirees with lengthy ATC experience) failed to successfully complete Academy training, and only 22 (37.3 percent) of the 59 attained test scores of 190 or higher, including 9 with scores of 210 and higher (24). The consensus that older controllers generally performed their duties less effectively than younger ATCSs was supported by the results of several studies (15,48,50), all of which indicated that performance, as evaluated by supervisors or peers, was apt to decline after age 40 without regard to tenure or experience in ATC work (18).

A series of published CAMI investigations (18,20,22,48,49,50), dating back to 1961, indicated that the training attrition rates of groups of trainees of age 31 and older were generally two to three times higher than those of the younger trainees (see Table 2). Other studies (15,23,48) in which experimental ratings of job performance were collected on journeyman-level ATCSs from both their supervisors and their peers revealed that the mean performance ratings of controllers within every age category beyond 40 were significantly lower than those of the younger subgroups (see Figure 6). Such findings and other unpublished analyses that Cobb provided at the request of the FAA and the CSC played a decisive role in the FAA's obtaining Congressional legislation in 1972 permitting the establishment of an optional early retirement program for controllers and also the imposition of an upper age limit of 30 in the recruitment of controller trainees (24). (Note: The maximum age limit of 30 does not apply to FSS personnel nor does the early retirement program.)

Summary. Implementation of the standard that precludes the training eligibility of any applicant older than 30 has been viewed by many within the FAA as representing the most significant step ever taken to improve the

Table 2. Comparison of Attrition (Attn) and Retention (Ret) Rates by CSC ATC Aptitude Screening Test Level, Type of Rated Pre-FAA Experience, and Dichotomized Age Grouping for 2,349 En Route and Terminal ATCSs Who Entered the Academy (Acad) During 1969. P-A Refers to Post-Academy Attrition (from Cobb & Nelson, 1974, ref. 22).

		CSC ATC Aptitude Test Score															
		Unknown				209 & Lower				210 & Higher				All Aptitude Levels			
Age	Pre-FAA Experience Category	N Tot	Acad Attr Rate %	P-A Attr Rate %	Ret Jan 1973 %	N Tot	Acad Attr Rate %	P-A Attr Rate %	Ret Jan 1973 %	N Tot	Acad Attr Rate %	P-A Attr Rate %	Ret Jan 1973 %	N Tot	Acad Attr Rate %	P-A Attr Rate %	Ret Jan 1973 %
31 & older	Non-ATC Pilot	71	57.7	21.1	21.1	1	100.0	0.0	0.0	115	27.8	33.0	39.1	187	39.6	28.3	32.1
	ATC & Pilot	16	25.0	18.8	56.3	3	33.3	66.7	0.0	9	0.0	22.2	77.8	28	17.9	25.0	57.1
	Non-Pilot ATC	125	35.2	26.4	38.4	40	27.5	25.0	47.5	48	10.4	25.0	64.6	213	28.2	25.8	46.0
	Other	84	64.3	10.7	25.0	5	60.0	20.0	20.0	148	27.0	26.4	46.6	237	40.9	20.7	38.4
	Total	296	48.3	20.3	31.4	49	32.7	26.5	40.8	320	24.1	28.4	47.5	665	35.5	24.7	39.8
30 & younger	Non-ATC Pilot	97	35.1	21.7	43.3	0	0.0	0.0	0.0	245	9.0	20.4	70.6	342	16.4	20.8	62.9
	ATC & Pilot	29	13.8	20.7	65.5	11	0.0	0.0	100.0	24	0.0	29.2	70.8	64	6.3	20.3	73.4
	Non-Pilot ATC	277	17.3	15.5	67.1	88	6.8	20.5	72.7	350	3.1	19.4	77.4	715	9.1	18.0	72.9
	Other	165	54.6	15.8	29.7	3	0.0	0.0	100.0	395	9.9	20.3	69.9	563	22.9	18.8	58.3
	Total	568	31.0	16.9	52.1	102	5.9	17.7	76.5	1014	7.1	20.2	72.7	1684	15.1	18.9	66.0

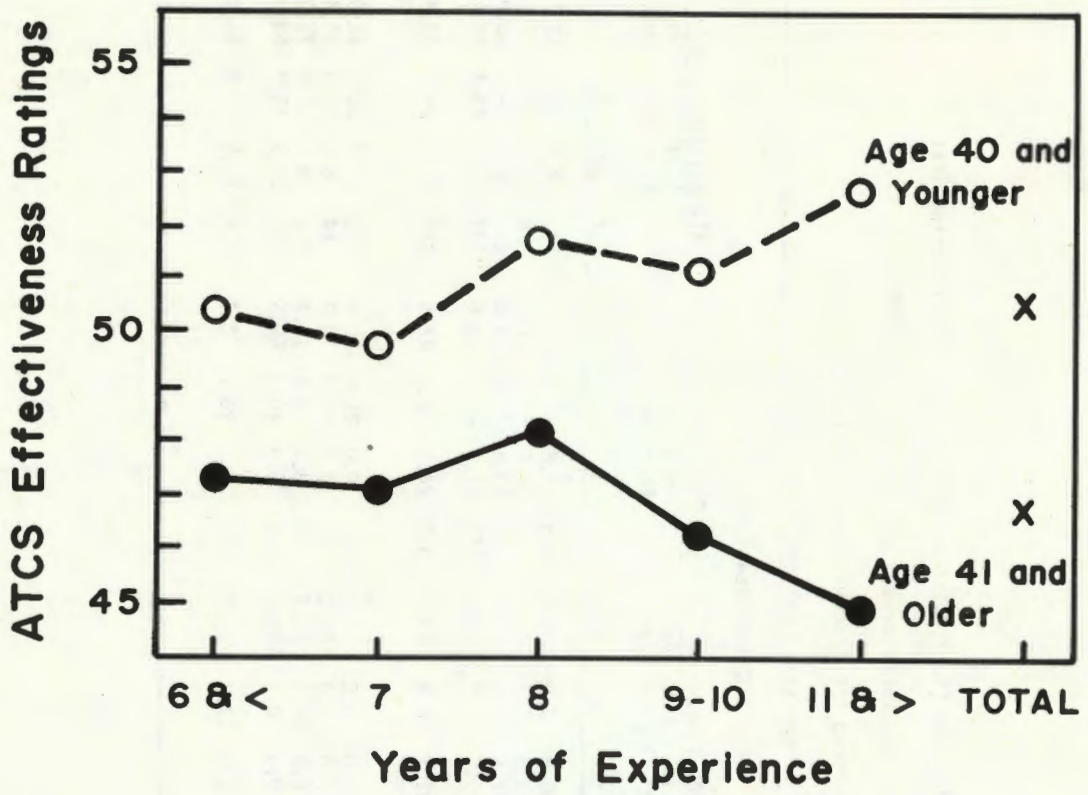


Figure 6. Means of ratings combined from supervisors and peers for "Older" vs. "Younger" radar ATCSs of different experience groups (adapted from Cobb, 1967, ref. 15).

selection process (24). This belief stems from consideration of findings obtained in a number of CAMI studies. Due to the standards prescribed for evaluating and weighting various types of aviation-related experience as a selection factor, relatively high proportions of the selectees who established their candidacy ratings during periods preceding imposition of the age limitation were more than 30 years old, including appreciable percentages of selectees in their 40's and some who were more than 50 (24). However, the followup studies of groups recruited during 1960 through 1971 always revealed significant inverse relationships between training-entry age and training-performance measures (15,23). The effects of age pervaded all experience subgroups, indicating the need for establishment of a standard to preclude the qualification and selection of older applicants, regardless of their experience backgrounds (22). Moreover, research (15,23) also demonstrated that the job proficiency of full-performance-level ATCSs, or journeyman-level controllers, generally tended to decline progressively after age 40. Such findings retrospectively attested to the validity of the FAA's current policy of recruiting only relatively young personnel for ATCS training.

Aviation-Related Experience as a Selection Factor

Throughout the history of the FAA and that of its predecessor organization, the CAA, ATCS selection programs have included standards predicated on the philosophy that almost any type of aviation-related experience should be of value for prediction of success in ATCS training and work. Inasmuch as previous experience in air traffic control (usually acquired in military service) has always been considered of paramount importance, standards have invariably prescribed that it be heavily weighted, directly or indirectly, as a selection factor. Other types of aviation experience traditionally regarded as important, but generally weighted more moderately than prior ATC work, include experience (military or civilian) as an aircraft pilot, a navigator, a communications expert, a radar surveillance specialist, and a flight dispatcher. Prior to implementation of mandatory aptitude-test screening procedures in 1964 (and exclusive of brief trial periods for procedures resulting in the selection of relatively few trainees), the eligibility ratings of medically qualified ATCS applicants were determined primarily on the basis of assessments of aviation-related experience and education (24).

Briefly stated, selection programs have always been formulated to result in the recruitment of as many controller trainees as possible from applicants who, in addition to other qualifications, possessed previous ATC experience. The appropriateness of this policy has been repeatedly confirmed by the results of CAMI followup studies of personnel who entered ATCS training during the decade ending in 1970 (24). Unfortunately, however, the pool of former military controllers has progressively diminished during the past 19 years, and the FAA has therefore recruited increasingly greater proportions of its ATCS trainees from candidates having other aviation backgrounds and also from those having no aviation experience of any type--but who qualified on the basis of aptitude-test measures and assessments of education.

Unpublished studies in which biographical data were collected and analyzed for large samples of ATCSs recruited from the 1960's through the early 1970's generally showed that 40 percent or more were former military controllers and that 40 to 45 percent of the remaining selectees held aircraft-pilot ratings (22). The results of several early CAMI studies (18,20,49,50) suggested that various types of pre-FAA ATC experience were primarily beneficial to the ATCSs during the basic-training phase only, whereas all other experience, including aircraft-pilot experience, appeared to be of questionable value at any stage of the training. Yet, little or no information had been gleaned through the 1960's to indicate whether the pilot-experience standards should be abolished, drastically revised, or modified only slightly.

The validities of aviation-related experience (pilot, navigator, etc.) and also other kinds of experience for prediction of training progress had never been firmly established (22). While several CAMI studies, though

focusing on other objectives, had included a comparison of the training attrition rates for groups of nonrated trainees and those having pre-FAA ratings as pilots, ATCSs, or communications specialists, an issue seldom dealt with was the amount of each type of experience. Moreover, the available results had generally been difficult to interpret due to the presumed (but unassessed) interaction effects of numerous variables (22). The trainees established their eligibility for training on the basis of various factors and standards and, inasmuch as chronological age was not considered in the screening process until 1973, those with rated experience tended to be somewhat older than the nonrated trainees. Research (17,18,19,20) had consistently shown age to be inversely related to performance on the CSC ATC test battery and to measures of training performance. Yet, no studies prior to 1973 had been accomplished to determine the extent to which the aviation-related experience might vary as a function of the interaction effects of age and aptitude level and also education.

Most of the early CAMI studies referred to above involved comparison of Academy-training-performance criterion measures only for trainees having different types of experience. In each such investigation, the trainees who held pre-FAA ATC ratings were found to have significantly higher training-course grade averages and/or graduation rates than all other groups, whereas the mean differences between the pilot-rated trainees and the communications-rated and nonrated personnel were usually rather small and not statistically significant (22).

Such results, however, did not necessarily warrant the conclusion that pilot experience should be completely disregarded in the screening of ATCS applicants; the standards might be merely too liberal. Under the standards, a total of five points was (and still is) credited toward the overall eligibility rating of each pilot-rated applicant having 350 or more hours of logged flight time. A cursory review by Cobb of biographical-questionnaire response data for several hundred pilot-rated ATCS trainees recruited during 1969 revealed that about half of them possessed no more than a private pilot license and 350 to 500 hours of logged flying time, and less than 30 percent had 1,000 hours or more (22).

In preparing for a direct study of the issues during 1972, Cobb and Nelson (22) contacted a number of Academy instructors and other long-tenured ATCS personnel and solicited their opinions regarding the validity of pilot experience for selection purposes. Each was consulted on an individual basis and granted assurance of anonymity. According to Cobb and Nelson (22), few expressed satisfaction with the current standards, several stated that pilot experience should not be considered in the selection process unless the applicant held a commercial license with an instrument rating, and some felt that all pilot experience should be disregarded.

Of all the discussants, those who had entered civilian ATCS training shortly after World War II seemed to have the strongest opinions (22). Many with military service as Airport Tower Operators said that pilot experience

had always received unwarranted emphasis; they felt that the selection standards used for several years after 1945 tended to favor the pilot-rated applicants, and that the military controllers frequently represented a minority of the ATCS-trainee recruits because they generally experienced greater difficulty in establishing highly competitive eligibility ratings. However, most of the long-tenured ATCSs who had served as pilots during the war felt that their candidacy for the military pilot-training programs, and commissioned-officer status after graduation, had been based on exceptionally high qualification standards (e.g., mental abilities and/or education, etc.). Some alleged that the majority of the pilot-rated ATCSs recruited during those years possessed more than 1,000 hours of flying time, usually with a great deal of navigational training and experience, and that the validity of that experience was manifested in the relatively low attrition rates of the pilots for all phases of the ATCS training program. Several claimed that the then-current issue regarding the use of pilot experience as a selection factor would never have arisen "if the standards had remained sufficiently high to preclude the entry of so many 'leisure-time' or 'Sunday' pilots" (22). Unfortunately, Cobb and Nelson (22) were unable to locate any factual information concerning the backgrounds and training progress of ATCSs recruited during the immediate postwar years. However, it was commonly acknowledged that military pilots with extensive flying experience and navigational knowledge represented progressively smaller proportions of the pilot-rated personnel selected each year since about 1955.

Increased skepticism regarding the validity of pilot experience for selection purposes (22) followed publication of the report of the Air Traffic Controller Career Committee in January 1970 (1). The committee report stated that no evidence could be found indicating any type of pre-FAA experience other than ATC work to be useful for prediction of FAA ATCS training progress or subsequent job performance. The report (1) recommended "elimination of credit for pilot experience" in the selection process. The same recommendation was made in a July 1970 contract report to the FAA by Colmen (25). However, neither of the two investigative bodies cited any studies other than those by CAMI as a basis for their conclusions and recommendations concerning pilot experience, and those cited studies did not necessarily imply that all such experience should be completely disregarded (22). None of the studies had focused directly on the issue of pilot experience; none included determination of the attrition-rate probabilities for ATCSs relative to their flying time or types of ratings held (e.g., private license, commercial license, instrument rating, air transport rating, etc.); and the interaction effects of age, aptitude, and education on the validities of pilot experience and other types of experience had never been assessed and compared. Thus the comprehensive longitudinal study by Cobb and Nelson (22) was undertaken to obtain such information and thereby minimize much of the conjecture that would otherwise arise in the formulation of future standards relating to the evaluation of prior experience in the selection of ATCS trainees.

That longitudinal study of 4,092 ATCS trainees examined the validities of various types of aviation-related experience, separately and in

combination, for prediction of success in FAA ATC work. Success was defined as retention status within the ATC system several years after entry into training. Of the 4,092 ATCSs, 1,740 entered Academy basic training during September 1960 through August 1963, before the CSC ATC battery became operational in the screening of most applicants. The remaining 2,352 ATCSs, the majority of whom were selected from among aptitude-screened applicants, entered Academy training during 1969. Both groups entered prior to establishment of the current standard under which eligibility for training is restricted to personnel no older than age 30. The obtained results clearly demonstrated that success in FAA ATC work was far more contingent on entry age than on type of aviation-related experience, level of aptitude, or level of education. The findings led Cobb and Nelson (22) to suggest that ATCS applicants who meet the existing age and aptitude screening standards should not be awarded credit points toward their eligibility ratings for any type of experience other than ATC work (see Figure 7), and that even the latter should be conservatively assessed and weighted in the selection process, particularly with respect to military control experience that involved no instrument flight rule (IFR) operations. Cobb and Nelson (22) noted parenthetically that if such a procedural change were followed, an indirect result would likely be a relative improvement in the competitive ranking of women and minority candidates who, for various sociocultural reasons, probably do not obtain the types of pre-FAA experience for which credit is given in selecting ATCS candidates.

Candidates having only pilot experience when they entered ATCS training during either of the two widely separated time periods had unusually low retention rates, even lower than those of groups having no aviation-related experience of any type. Speculation that many of the pilot-rated ATCSs might have volitionally attrited in order to take jobs of higher pay or jobs more compatible with their flying interests, prompted a search of the standard personnel records of all FAA employees; only 14 of the 254 attrited pilot-rated ATCSs of 1969 were still in the FAA (in non-ATC jobs) at the beginning of 1973. Of the 14, 6 were aviation safety officers (5 of whom were GS-12's); 1 was a GS-12 in an educational training program, and the other 7 were in various jobs (e.g., electronics technicians, flight data aids, wage board employees, clerks, etc.) with lower pay grades (22). Unfortunately, data were unavailable for a similar followup study of the remaining 240 attrited pilots who had left the FAA.

Thus ATC experience that involved only visual flight rule (VFR) operations proved to be considerably less valid than ATC-IFR experience for prediction of success in either the En Route or Terminal option. Cobb and Nelson (22) therefore concluded that, when possible, all En Route and Terminal trainees should be selected from among the best qualified of the aptitude-screened candidates younger than 31 who possess pre-FAA IFR control experience.

In a later CAMI report by Lewis (35), the feasibility of using scores on the Occupational Knowledge Test (OKT), a test containing questions on general ATC information to determine the quality of an applicant's ATC

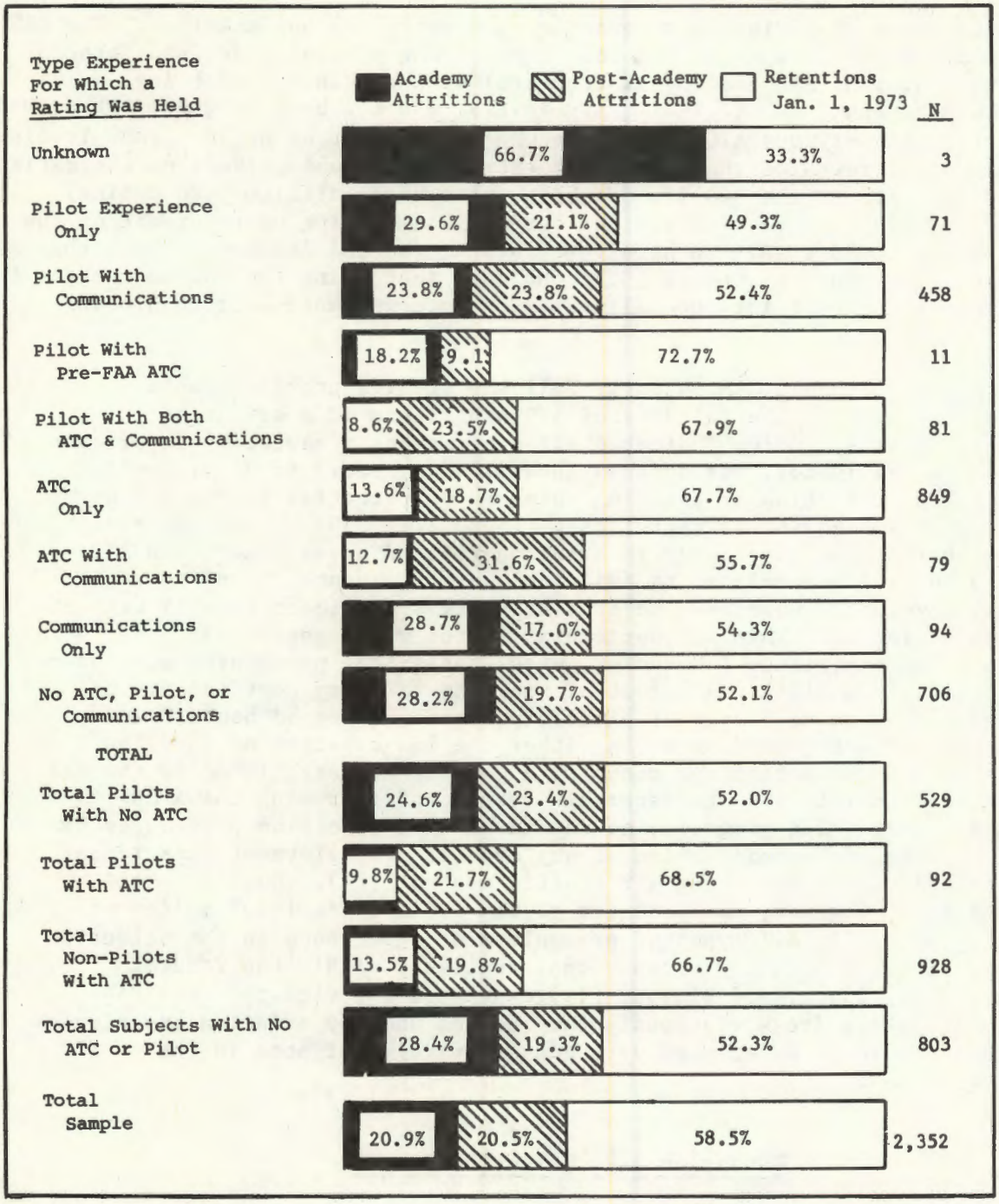


Figure 7. Attrition and retention rates by pre-FAA experience categories for 2,352 entrants into the Academy's En Route and Terminal basic training courses during 1969 (adapted from Cobb and Nelson, 1974, ref. 22).

experience prior to giving extra experience points (43) was examined. The OKT had been developed from a pool of test items (some specially devised, others drawn from existing FAA Academy instruments) with the intent of being "job-knowledge specific" (37,43). Its initial use had been intended for screening applicants with previous ATC experience who would enter at higher grade levels. The Lewis study revealed that assigning extra credit subjectively on the basis of verified experience (the quality of which was usually difficult to define) resulted in a failure rate of 7.6 percent while assigning extra credit on the basis of OKT scores would result in a failure rate at the FAA Academy of less than half that rate. Consequently, Lewis (35) concluded that using the OKT as a basis for assigning extra credit for specialized experience was more fair than using verified experience.

Summary. Although CAMI's early followup studies provided ample evidence attesting to the validity of ATC experience as a selection variable, the same studies indicated all other types of aviation experience (e.g., pilot, navigator, air defense surveillance, etc.) to be virtually worthless for prediction of training outcomes or retention in FAA ATC work several years subsequent to recruitment (15,20,22,49,50). More importantly, variables reflecting experience in aviation-related areas other than ATC (and also in fields unrelated to aviation) were frequently found to correlate inversely (sometimes at statistically significant levels) with the criterion variables. This was particularly true with respect to aircraft-pilot experience. Moreover, when statistical procedures were used to theoretically nullify the effect of age, the validity coefficients of variables representing "years of pilot experience," "logged hours," and "pilot ratings" were found to be in either the low negative or very low positive range. Such findings prompted a committee, established by the FAA Administrator in late 1974 to ascertain methods of improving the ATCS selection and training programs, to recommend that selection procedures be revised to preclude consideration of any type of preemployment experience except that directly involving air traffic control (24). Such a recommendation, however, runs counter to the CSC's traditional policy of evaluating virtually all types of preemployment experience in the selection of personnel for almost any occupational specialty within the Federal service. Cobb, Young, and Rizzuti (24) expressed the view that the FAA should nonetheless press vigorously for changes whereby aviation experience other than ATC would be at most very conservatively weighted in the selection process.

Education as a Selection Factor

The ATCS selection programs of all time periods have included the mandatory requirement that an applicant either hold a high school diploma or provide evidence (e.g., a General Education Development (GED) certificate) of an educational background deemed equivalent to that of a typical graduate. Education beyond the high school level (although not mandatory) has traditionally received significant weighting, directly or indirectly, as

a factor in the derivation of each applicant's overall eligibility rating. Selection programs have invariably included provisions whereby the "general experience" requirements, specified for applicants not having backgrounds in aviation, could be met wholly or in part by the substitution of college-level education. Regardless of experience, however, applicants with 4-year college degrees have usually been provided a variety of standards helpful in establishing training candidacy. As mentioned earlier, there was a period during which college graduates with records of superior academic achievement were screened in terms of an aptitude-test-score standard that was considerably below that designated for screening of their comparably experienced and otherwise equally qualified colleagues. Also, a similar policy prevailed at one time that pertained to all college graduates, irrespective of academic records. However, the greatest emphasis of education as a selection factor was during 1971 and 1972 when applicants with 4-year college degrees having at least 1 year of graduate work and 12 months of specialized ATC experience could be granted waivers of the aptitude-test-screening phase and also be appointed to training at grade GS-9 rather than GS-7 (24).

Throughout the history of the CAA and the FAA, education has received notable consideration as a factor in the selection of ATCS trainees. However, virtually all the selection procedures relating to education except those concerning records of superior academic achievement have pertained to levels of education, or years of education completed. The types of college studies (e.g., college courses, areas of major and minor study, etc.) pursued by the applicants had generally received little or no consideration (24). While relevant experience has usually been deemed more important than education as a selection factor, applicants lacking qualifying preentry experience, but meeting other prerequisites, have generally been allowed to substitute college education (as measured by years) for experience, in accordance with differential rates prescribed for the various types of experience.

Although the selection procedures pertaining to education have varied from time to time, they have consistently reflected an implied assumption that success probabilities in ATC training and work tend to vary in accordance with levels of education attained by personnel considered equally qualified in other respects. However, the belief that educational level is indeed appropriate and valid for selection purposes had no apparent research base. To the contrary, various phases of early CAMI research (all data based on trainees who entered ATCS training in the decade of the 1950's) on ATCS selection in which level of education was included as a variable for "peripheral study" in analyses relating to the validation of age, preemployment experience, and aptitude-test performance indicated that the training attrition rates of ATCS personnel tended to increase (rather than decrease) in accordance with the preentry levels of education (14,20,22,49). Several unpublished CAMI studies involving trainees recruited during various time periods had shown that college graduates generally had significantly higher attrition rates than selectees having either high school diplomas only or 1 year or less of college (24).

Such findings, however, did not necessarily imply the need for reformulation of selection procedures relating to education. In fact, the early findings could not be reliably interpreted because they were undoubtedly confounded by (unassessed) interaction effects of age, experience, aptitudes, and other factors considered in the selection process. Virtually all the early (pre-1975) CAMI validation data pertaining to education were derived in conjunction with research undertaken primarily for other purposes. Some of the CAMI studies in which level of education had been included as a variable for ancillary study involved determination of the coefficients of correlation for the educational variable versus Academy training-performance measures and other criteria (14,20,22,49,51). Certain phases of research also had included the determination and comparison of attrition rates of ATCSs categorized in terms of educational level (22). In none of these studies, however, were analyses conducted to assess the effects of other selection factors on the validity of the educational variables or the influence of the latter on the validities of the former. Moreover, selection procedures bearing on education had traditionally included consideration of level only, and, due in part to that fact and the ready availability of data not requiring transposition, CAMI investigators invariably refrained from dealing with other aspects of education when conducting research undertaken primarily for other purposes.

Thus, a major study (24) specifically focused on education was undertaken for three basic purposes: (i) to assess the validity of level of education as a selection factor, separately and in combination with other factors officially considered in the selection of ATCS trainees; (ii) to determine whether data and information relating to other aspects of education (e.g., recency, major courses of study, etc.) were sufficiently related to training attrition-retention status to warrant consideration in the establishment of eligibility ratings; and, if so, (iii) to derive a set of factor weights, reflecting their relative performance, for recommended use in future revisions of the ATCS selection standards.

Series of analyses were conducted on data pertaining to 2,352 ATCS recruits who entered the Academy basic training phase in 1969 (1,858 En Route and 494 Terminal trainees). The study (24) revealed all educational variables, both before and after consideration of effects associated with age and other selection factors, to be negligibly and/or inversely related to success in ATCS training as defined by Academy graduation status and/or retention in the ATC system 3 to 4 years following recruitment (see Figure 8).

The 181 college graduates among the 2,352 recruits had a higher Academy attrition rate (30.9 percent) than any other subgroup, and 208 whose highest level of education was reflected by a GED certificate for high school had the next highest Academy attrition rate (24.5 percent), whereas the lowest rate (18.0 percent) pertained to 876 having high school diplomas only, followed closely by the rate of those having less than 1 full year of college (18.3 percent). Post-Academy (i.e., facility training) attrition rates ranged from 16.3 to 26.5 percent and followed a somewhat similar rank-order pattern (24).

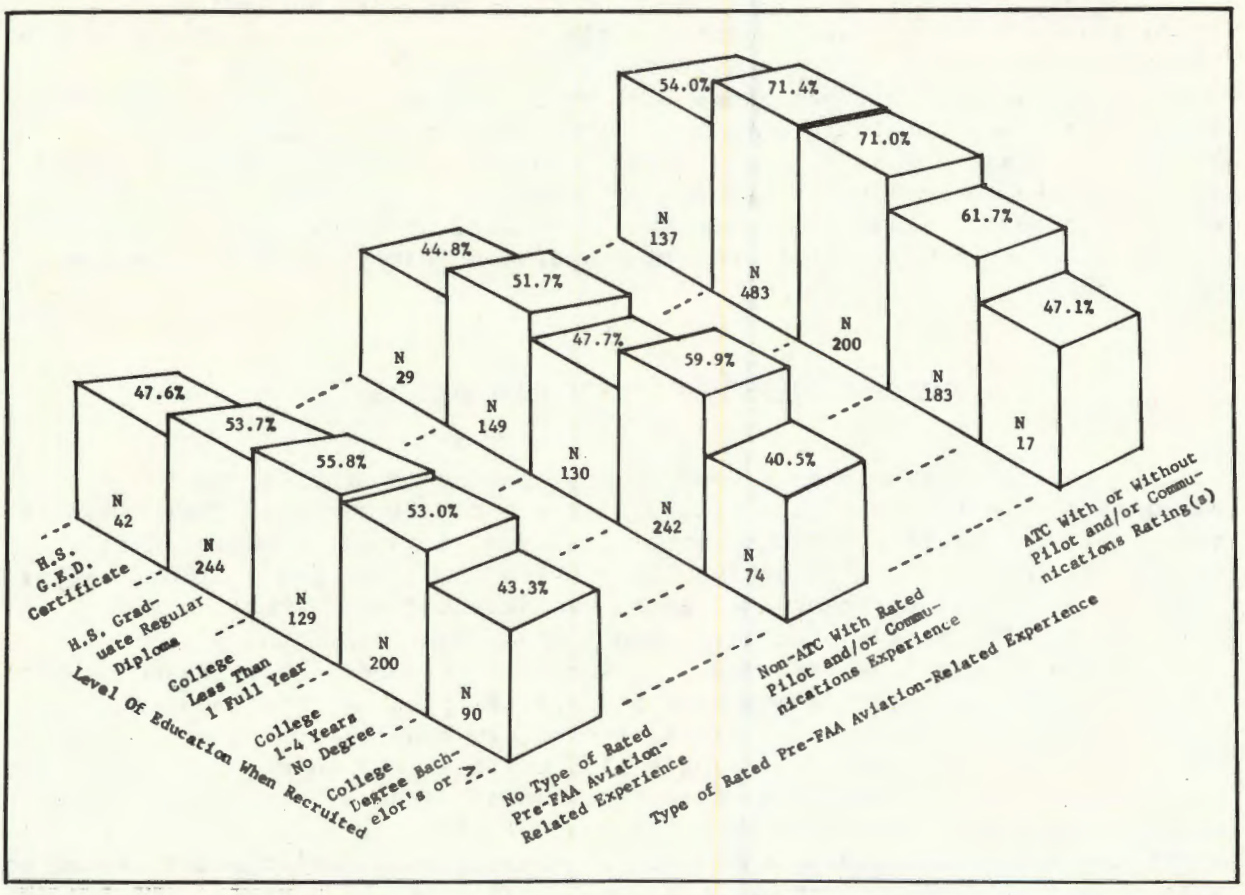


Figure 8. Retention rates by level of education and type of rated pre-FAA aviation-related experience for 2,349 selectees who entered Academy basic training in En Route or Terminal ATC procedures during 1969. Rates reflect proportions of ATCSs in each subgroup who were still in FAA ATC work as of January 1973 (from Cobb, Young, and Rizzuti, 1976, ref. 24).

Major courses of study listed by some 925 of 1,265 ATCS trainees who attended college were found to have little potential for prediction of training outcomes (24). The overall retention rate as of January 1, 1973, for the 1,265 former college students was 56.7 percent. When categorized on the basis of major studies, only those (N=141) majoring in the social sciences were found to have a retention rate differing significantly from that of the combined categories. Only 41.8 percent of the 141 social science majors were still in FAA ATC work at the beginning of 1973. Moreover, 53 recruits for whom major courses of college study were judged as being more directly related to aviation than were the majors of all others were found to have a retention rate of 56.6 percent, only one-tenth of a point lower than that of the entire group of former college students.

Summary. Thus, CAMI research through the mid-1970's found no educational variable that evidenced a significant interaction effect on the validities of other selection factors. In addition, all types of aviation-related experience except ATC were found to be grossly unreliable for prediction of training outcomes. These and other findings led Cobb, Young, and Rizzuti (24) to suggest that candidacy for ATCS training should be restricted to aptitude-screened applicants no older than 30 and that it would also be advisable to discontinue the award of credit points toward eligibility for all types of education and/or preentry experience except ATC experience.

Sex as a Factor in Performance and Attrition

While it had long been a governmental policy to provide equal employment opportunity in the Federal service to all persons, regardless of race, color, religion, national origin, or sex, a major emphasis on this policy was contained in Executive Order 11478 issued in August 1969, wherein the head of each department and agency was directed to "establish and maintain an affirmative program of equal employment opportunity" to eliminate all remaining vestiges of discriminatory practices. Based on the FAA's interest and participation in the overall program to eliminate sex as a discriminating factor in the selection of new Federal employees, a comparative study of female and male ATCS trainees was undertaken (20). The study was also prompted by the fact that women had never represented more than a very small proportion of all personnel directly involved in air traffic management. Since experience is heavily weighted in establishing an applicant's eligibility rating, the "best qualified" by normal standards are usually male veterans with experience in military ATC work or as pilots (13,14,46,49,50). It was therefore contemplated by Cobb in the early 1970's that comparative analyses of the various types of data available for groups of female and male trainees would yield findings of value to the FAA in its formulation of future recruitment, selection, and training programs.

As part of his comprehensive research program (18,19), Cobb and his associates had administered various aptitude tests, on an experimental basis, to all but three of the successive incoming classes of the Academy's basic training courses in Terminal, En Route, and FSS work during a 17-month period (November 20, 1968, through March 27, 1970). CAMI was provided a report on each trainee at time of entry into the Academy, reflecting types and amounts of preemployment experience, facility assignment, the overall eligibility rating (used for selection purposes), pay grade, and other information; the report on each trainee who took the operational CSC ATC test battery also included all part scores and the composite measure of performance.

Although participation in the experimental testing program was not mandatory, few declined to take part (20). Only 410 (10.7 percent) of the 3,843 trainees were not examined, and the majority of the 410 were in

classes for which testing sessions were canceled in order to avoid holiday travel conflicts. Generally, each group was examined on the day of arrival at the Academy. No more than 3 hours could be allocated for each testing session, and adherence to training schedules precluded more than one session per class or group. Inasmuch as the overall project specified that the potential of many different types of tests be explored as predictors of training performance, the groups of tests selected for experimental administration were intentionally varied. At least 1 week, and sometimes 2, intervened between incoming groups. Each group was administered only five to seven instruments, but periodic modification of the battery ultimately resulted in the collection of response data on samples of different sizes for 36 different tests or subtests; these yielded 41 performance measures (20).

The study compared age, education, pre-FAA experience, aptitudes, training-course performance measures, and post-Academy attrition rates of the 83 women who entered basic ATC training at the FAA Academy during the 17-month period with those of various samples of the 3,760 males who entered training during the same period. The study (20) revealed no significant differences between the means of the female and male trainees with respect to age and educational level. When samples of the 83 women were compared with samples of male trainees in terms of performance on 36 different apptitude tests, only four mean differences, all of which favored the women, proved statistically significant. Only 45.8 percent of the 83 women had pre-FAA ATC-related experience, while such experience was possessed by 63.9 percent of a sample of 798 men; the difference was statistically significant ($p < .01$). The means of the training course grade averages of the two groups differed by only three-tenths of one point, and there was no significant difference between the Academy attrition rate of 20.5 percent for the women and 23.2 percent for the 798 men. However, the groups differed markedly with respect to post-Academy attrition rates; 48.5 percent of the 66 women who completed Academy basic training were no longer in the air traffic management system as of December 1971, whereas only 22.5 percent of the 613 men (within the sample of 798) who graduated from the Academy were subsequently attrited (20).

These attrition rates, which showed (i) no sex differences in the proportion of trainees who completed FAA Academy training, but (ii) a percentage of women who subsequently left ATC work that was over twice that of male graduates, led to several studies of sex differences in job attitudes and attrition rates (38,39,40). In one of these, the 56 women who entered the Academy between December 1968 and November 1970, and who were no longer with the FAA as of June 1972, were matched on several variables with male attritions who had entered the Academy at the same time. Reasons for attrition were obtained from job-exit forms, telephone interviews, and a questionnaire. The majority of trainee attritions could be accounted for by three reasons: training difficulty or failure, family, and other employment. The category that accounted for most of the difference in attrition rates between the sexes was family-related (marriage, children) attrition cited by about one-third of the women, a finding that agrees with results of other studies of different occupational groups (39).

Summary. CAMI studies on sex as a factor in Academy performance have shown, for the relatively small proportion of women entering ATCS training around 1970, no differences from men in age, educational level, general aptitude factors, Academy training performance, and Academy attrition rates. However, significantly fewer women had prior ATC-related experience, and a significantly higher proportion of women attrited before reaching the full performance level. The difference between the sexes in reasons for attrition was mostly due to family-related reasons for women.

Research on Military ATCS Trainees

CAMI has had a long-standing interest in the selection and training of military ATCSs based in the fact that an applicant pool of former military controllers has always represented a prime source for the selection of FAA ATCS trainees. Aside from other reasons, these recruiting practices have prompted the FAA to maintain some awareness of the effectiveness of the controller selection-and-training programs in the U.S. Air Force, Army, Navy, and Marine Corps (17).

Plans for the first of the two CAMI investigations regarding military ATCS trainees were initially conceived in 1965, when representatives of the Glynco Naval Air Station, Georgia, visited CAMI for indoctrination regarding the underlying research, development, and effectiveness of aptitude testing and other procedures used in the screening and selection of applicants for FAA ATCS training. Naval officials subsequently discussed the possibility of being permitted the use of the operational battery of CSC tests for experimental administration and validation at Glynco. Although a number of policy reasons precluded this approach, Cobb suggested an alternate plan, subsequently accepted, involving the seven commercially published tests that had been validated in previous research with FAA trainees (17). After completion of the Glynco study (16), officials of Keesler Air Force Base, Mississippi, asked that a parallel study be conducted on samples of Air Force and Army ATC trainees.

The selection of tests for validation at the two training facilities was predicated on the findings that CAMI had obtained in previous research with FAA ATCS trainees. As mentioned earlier, a total of eight non-CSC instruments (seven commercially published tests and the Air Traffic Problems test), had been identified as yielding a number of different and valid composite measures for prediction of FAA ATCS performance criteria. Each composite score validated at a statistically significant level for prediction of the training measures. Inasmuch as the Air Traffic Problems test was subsequently designated as a CSC test, existing policies precluded its release for experimental use with military trainees. It was assumed that the remaining seven tests would constitute an appropriate battery for experimental studies involving military ATCS trainees (17).

Although two of the seven commercially published tests which were experimentally administered to the military ATCS training schools

consistently failed to correlate at statistically significant levels with the training-performance measures of every group (CTMM Analogies and CTMM Coins), composite scores based on the entire seven-test battery correlated with the academic plus laboratory grades of every group at somewhat higher levels than any of the measures based on combinations of test scores. Yet, in every instance, the validity of the "commercial seven-test composite" score was closely approximated by that of a composite measure reflecting performance on DAT Space Relations, DAT Numerical Ability, DAT Abstract Reasoning, and CTMM Inference (17).

Composite scores of both types correlated with the training-course performance measures significantly better than did the Air Force's General Aptitude Index (GI) and the Marine Corps' Military Screening and Classification (MSC) test score, and at about the same level as the aptitude screening measures used by the Army and Navy. Despite the high degree of comparability between the validities of each experimental composite and the MSC scores of the Army and Navy trainees, the results of multiple correlational analyses indicated that the experimental measures for these trainees, as well as for the airmen and marines, could have been used in combination with the MSC scores to attain some degree of improvement in the respective ATCS selection programs (17). However, the benefits stemming from such a procedure would have been considerably less for the Army and Navy than for the Air Force and Marine Corps. Moreover, while these studies were undertaken to determine the effectiveness with which various experimentally derived aptitude test measures might be used to improve military ATCS trainee selection procedures, the most important finding was that the screening potential of existing MSC aptitude-test scores could be more fully exploited to achieve that end (17). For reasons unknown, but perhaps due to recruiting demands, 16.5 percent of the Navy trainees, 14.8 percent of those in the Air Force, about 5 percent of the Army selectees, and one Marine were assigned to military ATCS training even though their MSC scores were below the normally prescribed minimum (which was 110 on the different test batteries employed by the Army, Navy, and Marine Corps, and 65 in terms of the Air Force GI). About 42 percent of the 138 men selected from these lower MSC-score categories either failed or marginally passed their training courses compared with 18.2 percent of the other 1,085 military ATCS trainees involved in the study. Moreover, other findings suggested that the recommended aptitude-screening standards could be raised by several points to increase their effectiveness. For example, MSC scores of less than 115 reflected the aptitude levels of 234 Army, Navy, and Marine trainees, and 51 airmen possessed a GI score of less than 70. Over 40 percent of these 285 students either failed or marginally passed training, whereas only 15 percent of the remaining 938 did so. The initial screenout, or disqualification, of every trainee having either an MSC score of less than 115 or a GI score under 70 would have rendered the commercial composite measures virtually useless for secondary screening purposes (17).

Thus, on the basis of these results, Cobb (17) concluded that the Army, Navy, Air Force, and Marine Corps had the capability, if desired or required, of upgrading their screening of ATCS personnel without new or

additional tests, by establishing higher minimum MSC requirements. Any changes in the military requirements would be of direct interest to the FAA since the agency has and will probably continue to select significant proportions of its ATCS trainees from among those applicants who have pre-FAA ATC experience and who also are able to qualify on the CSC ATC Aptitude Screening Test battery. At the present time, 210 is the minimum passing score on the battery. Such a screening standard (roughly equivalent to an MSC score of 125 and a GI score between 75 and 80) reflects a high degree of selectivity, such that numerous unpublished studies by Cobb had demonstrated that a downward revision in the aptitude-test screening standard of 10 to 15 score points (but no more than 20) was feasible and should be considered if the recruitment of personnel for FAA training ever reached a critical stage (17). On the basis of the results obtained in this study and those of previous investigations, Cobb (17) predicted that approximately one-half of the military trainees selected under these military ATCS selection standards would have been able to meet then-current (1971) FAA aptitude screening requirements; only about one-third of the FAA's applicant population prior to 1964 would have been able to do so. Moreover, if military cutoff scores were raised, for any reason, from then-present (1971) levels to MSC scores of 115 and to GI scores of 70, approximately two-thirds of the military controllers who qualified at those levels would qualify on the FAA's aptitude screening battery (17).

Summary. CAMI research on military ATCS selection procedures produced two major findings: (i) The same commercially available aptitude tests used in the development of the FAA aptitude screening tests were valid for predicting training success for military trainees and, if added to existing military tests, would improve prediction; and (ii) the effectiveness of military selection could be improved simply by increasing minimum cutoff scores on existing military screening tests. As a byproduct of this research, Cobb was able to outline the relationship between military selection/training procedures and the screening and entry of ex-military personnel into the FAA ATCS program.

Research on Alternative Aptitude Screening Measures

The early research in which six CSC tests were initially validated also yielded appreciable validities for a number of other CSC tests and several commercially published, copyrighted instruments. The latter group included several tests that apparently involved a speeded perceptual-discrimination factor and some that pertained to either coding skills, comprehension, rapid integration or processing of information, or memory (19). Although none of these tests appeared in the final group that the multiple regression analyses indicated most useful for selection purposes, the results led Cobb and Mathews (19) to suspect that the instruments were measuring important ATC-related aptitudes, albeit inadequately. Such findings, supplemented by those obtained for a number of different spatial tests, were interpreted as illustrating the need for development of tests "tailored" to fit the ATCS selection purpose more precisely.

The Directional Headings Test. While none of the various spatial tests included in CAMI's numerous experimental batteries had failed to validate at statistically significant levels, the validities of each had generally been much more variable from sample to sample than for nonspatial tests. Parenthetically, this also applies to CSC Booklet 51 of the operational screening battery. Cobb had concluded early on that a "visual imagery" type of speeded spatial ability was a major determinant of ATCS performance but was unable to locate an instrument fully appropriate in this respect. The development of a special test, the complexity of which would have required a major research effort, was not undertaken during the 1960's due to other research commitments having higher priorities (19).

However, early in 1970, the Directional Headings Test (DHT) was under "in house" development and experimental validation by Cobb on the basis that the aptitude screening process might be further improved through more adequate assessment of the applicants' speeded perceptual-discrimination and coding (or decoding) skills. In conceiving the type of task ultimately embodied in Part I of the DHT, initial efforts focused on the use of diagrams simulating a navigational compass. However, the pictorial approach was abandoned on the hypothesis that if the data pertinent to the solution of each problem were presented in a different format, such as described below, the test would be more difficult and therefore provide better opportunity for individual differentiation (19).

The DHT was a three-part test (in the final form, Part III was dropped) in which the examinee is allowed a total of only 90 seconds for the solution of 60 items, or problems, of each part (19). In each problem, the examinee is presented one, two, or three "bits" of information relating to the cardinal points on a mariner's compass. For example, the letter "N," the symbol "^," and the notation "360" each denote "North." Similarly, the letter "E," the symbol ">," and "090," either separately or in combination, denote "East." Other letters, symbols, and degrees correspond to "South" and "West." In Part I of the test, the examinee must rapidly interpret and collate the bits of presented information in order to ascertain whether the data are of a conflicting nature and, if not conflicting, then determine the directional heading to which they correspond; in Part II, the examinee must determine the exact opposite of the heading; and, in Part III, the examinee must ascertain the exact opposite of the directional heading while being subjected to aural distraction. In each item of Form A, the bits of information are followed by one of five questions: "North?," "East?," "South?," "West?," or "Conflict?." Two response categories are offered, "Yes" or "No." Approximately one-fourth of the items present "conflicting" data. For example, if the letter "S" were presented in an item of Part I with the symbol ">" and/or "270" and if the question were "Conflict?," a correct answer would be indicated by marking the space under "Yes." On the other hand, the same conflicting data presented with "West?" (or "North?," "East?," or "South?") would warrant a "No" answer.

The first preliminary version of the test consisted of only one part, with the task being essentially the same for all 60 items (19). Small

groups of Academy ATCS trainees were examined with the instrument and, although no validation analyses were possible at that time, considerable range and variance were found in the distribution of scores. Moreover, during test sessions, some individuals in every group appeared to become rather confused, or frustrated, when confronted with those items in which "conflicting" bits of information were presented and also hesitant in marking either "Yes" or "No" to many of those items where the questioned direction (e.g., "North?") failed to conform to the heading indicated by the data provided. When the CSC ATC battery was being validated some 8 years earlier, reactions of much the same type had been noted for many of the trainees during their assessment with CSC 135, Following Oral Directions. Inasmuch as CSC 135 proved to be one of the most valid components of the screening battery, Cobb hypothesized that many of the trainees who had performed poorly on the instrument might have done so because they had a "low frustration tolerance level." Although the latter phrase and the hypothesis to which it alludes were rather loosely formulated, they grossly reflected the reasoning that led to the development of Parts II and III of the DHT (19).

As mentioned earlier, DHT Part II is similar to Part I except that in each item the examinee is asked whether a specified direction represents the direct opposite of the heading (19). Part III is like Part II but the examinee is subjected to possible distraction by the aural presentation of the randomly arranged words, "West," "South," "East," "North," and "Conflict." The resulting three-part instrument, in which the examinees were offered no response categories other than "Yes" or "No," was designated as CAMI Directional Headings Test--Form A.

An alternate version of the test, Form B, was eventually developed; it offered five response categories (19). In each item of the latter, the "bits" of information are presented in the same manner as in Form A but the "question element" does not appear and the examinee merely indicates the answer he or she deems correct by marking a response space under one of the five columns labeled, "East," "South," "West," "North," and "Conflict."

Form A of the DHT was administered by Cobb and Mathews (19) on an experimental basis to a total of 586 students as they enrolled at the FAA Academy from January 20, 1970, through March 13, 1970, in either the basic En Route or Terminal training course. Form B was administered on a similar basis to 199 trainees, all of whom entered the En Route course on March 27, 1970. The vast majority of the 785 trainees examined with either Form A or Form B were new hirees and none possessed an FAA ATCS certificate; most had been selected for training on the basis of competitive ratings from among candidates who met exceptionally high qualifying standards in terms of operational aptitude test screening scores and/or evaluations of pre-FAA experience. Despite these prescreening effects, the DHT scores correlated .41 with an overall measure of training performance. Moreover, over 44 percent of the 15 examinees who failed the training course scored no higher than 29 on the DHT, whereas over 85 percent of the graduates scored 30 or higher. Reliability of the instrument, as determined by correlating the

scores based on even-numbered items versus scores based on odd-numbered items, was .93. The DHT was made available to EPA, the private company that was performing an ATCS selection study for the FAA as mentioned earlier. In their final report (41) based on samples of journeyman controllers with 3-10 years of experience, EPA researchers noted the effectiveness of the DHT in predicting job success by citing it as one of the best of the large numbers of aptitude tests they evaluated. Thus the promising DHT findings obtained by Cobb and Mathews (19) with ATCS trainees were independently confirmed with a different group of journeyman controllers.

Based on the above findings the DHT has been included in current efforts to develop a new ATCS screening battery (5). The test has shown good predictability, but in its present highly speeded form is not amenable to ordinary OPM national testing procedures (5). Efforts, perhaps, should be directed toward altering the DHT to a usable form or incorporating it in another instrument.

The Multiple Task Performance Battery. While the CSC ATC test battery does an efficient job of identifying those individuals who possess the "elemental" skills and knowledge necessary to becoming a satisfactory controller, it does not measure a particular kind of ability that may be said to be a defining characteristic of a good controller--the ability to perform several different tasks simultaneously (11). Thus, undoubtedly a number of trainees are eliminated from the ATCS training program, not because these trainees lack information in the academic sense, but because of an apparent deficiency in their skill at the concurrent, time-shared performance of a variety of tasks, viz, they cannot put the information together and perform satisfactorily. If this is true (and it is a widely held belief among FAA Academy instructors), then a technique that provides an objective, reliable index of the ability of an applicant to time-share multiple tasks of relevance to the job of an ATCS would be of value as an adjunct to the existing selection devices (11).

During the period from May 1970 to January 1971, exploratory studies were carried out by Chiles and his coworkers to examine the potential usefulness of a performance measurement device as a predictor of the future performance of ATCSs. Five groups totaling 229 ATCS trainees were tested on the CAMI Multiple Task Performance Battery (MTPB), and the predictor scores that were derived from the performance battery were correlated against a criterion provided by FAA Academy instructors (11).

The MTPB was originally designed for the Air Force as a device for use in research on complex performance of the sort demanded of aircrew personnel (10). The elements of the MTPB were selected to provide objective measures of "psychological or behavioral functions" of relevance to Air Force operations. The functions measured by the MTPB include monitoring, information processing, mental arithmetic, visual discrimination, and interindividual interaction in the execution of procedures (see Figure 9). These functions would appear to be relevant not only to aircrew activities but to complex jobs in general and to the job of the ATCS in particular. Moreover, the tasks, as routinely used over a number of years, have been

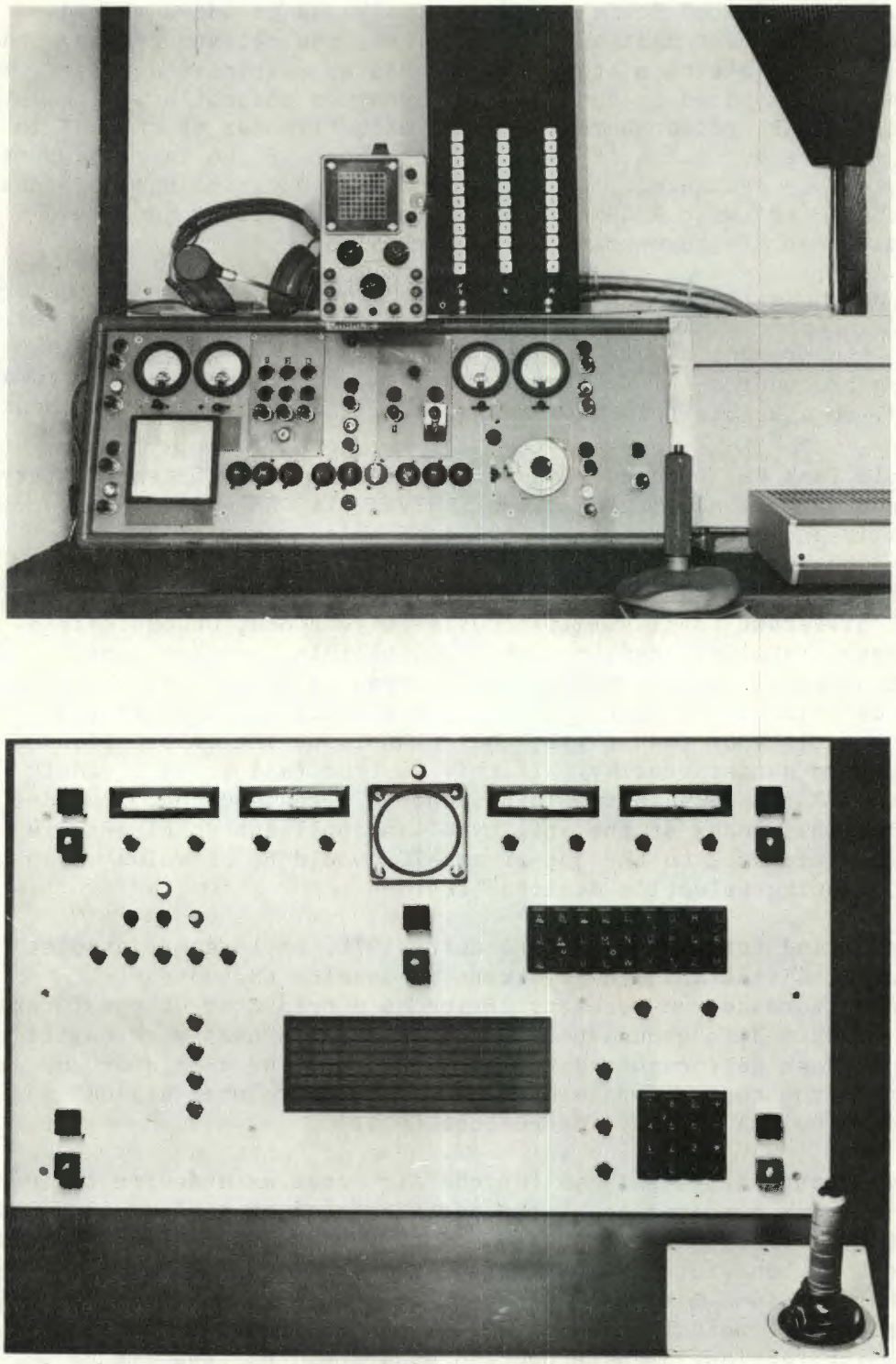


Figure 9. The CAMI Multiple Task Performance Battery. The upper photograph (A) depicts the console used in the studies discussed in this report, and the lower photograph (B) shows the current version of the test.

structured to impose varying levels of demand on the individual with respect to the requirements for time-sharing. Good performers on the MTPB are those individuals who not only possess ability on the individual tasks but, also, are readily able to shift their focus of attention from one kind of activity to another without disruption of the ongoing process (11).

The rationale for investigating the selection potential of the MTPB derived from the complexity of the ATC task situation with respect to the requirements for time-sharing. There was no particular reason to expect any of the individual tasks to be predictive of ATCS abilities even though FAA Academy instructors regarded the tasks as being generally of relevance to the job of the controller (11). The nature of the MTPB is such that an individual who fixates on the performance of one task will almost inevitably do poorly on one or more of the time-shared tasks. This is especially evident early in learning when skill is being acquired at the detailed, refined performance of the individual tasks. During this first hour or so of learning, the task ensemble should be sensitive to differences in the ability of individuals to shift attention rapidly from one task to a second task and to quickly adopt the proper "set" for the performance of that second task. Thus, the intent of a set of MTPB studies conducted with ATCS trainees was to provide an answer to an intermediate question, viz, the determination of whether the approach embodied in the MTPB showed sufficient promise as a selection device to warrant a large-scale investigation to develop the supporting data that would be required to demonstrate its efficiency as a predictor of ATCS skills in terms of cost-effectiveness criteria (11).

Five studies that involved a total of 229 ATCS trainees were undertaken by Chiles, Jennings, and West (11) to explore the possibility of using a performance measurement system such as the MTPB as a means for assessing the potential of trainees to become fully rated ATCSs. The primary criterion used in evaluating the predictive efficiency of the MTPB was the mean of three ratings for each trainee--one rating from each of three instructors per trainee. Specifically, the instructors were asked to express their opinion of each trainee's potential for becoming a fully rated controller by assigning a number from 1 to 100 to the trainee, with 70 being "minimally acceptable potential." These ratings were very easy to obtain, and they appeared to possess satisfactory reliability (estimated to be 0.84) (10). Previous research (8,46) had also shown that instructor ratings of this general sort were of substantial predictive value with regard to the on-the-job performance of ATCSs. In addition, 40 instructors were tested on the MTPB and then asked to complete a seven-item questionnaire designed to elicit their opinions as to the relation of the MTPB as an index of trainee potential (11).

The first study involved 19 En Route trainees; it served as a preliminary to the other studies in checking out training and testing procedures; it also permitted an examination of the appropriateness of the task parameters used. The second study involved 60 En Route trainees; it yielded a good validity coefficient against the criterion--product moment r

.54. The third study also involved En Route trainees (a total of 31), but a different set of task parameters was employed. The validity coefficient in this study was .53. The coefficients found in both of these studies were lower for data from the second and third hours of testing than for the first. The fourth study, which involved 30 Terminal trainees and the same task parameters as the third study, failed to replicate the validities found in the second and third studies; the correlation here was near zero. The validity coefficient found with 89 Terminal trainees in the fifth study, using still a different set of task parameters, was not as large as those of the second and third studies, viz, .24; however, when the MTPB performance index was based on measures judged to be more appropriate in relation to the changed task parameters, the correlation was a comparable .46.

The opinions of the instructors, as reflected in their responses on the questionnaire, supported the conclusion that the MTPB exhibited substantial content validity as a measurement system of relevance to the job of the controller (11).

It was concluded that the task parameters used in the first study (and retained in the second study) were closer to optimum than were any of the modifications used in the third, fourth, and fifth studies. The method of computing the composite index of performance (via equal weight to all tasks in terms of the variance of the composite index) was judged to be the best technique for developing an objective measure for use as a summary assessment of performance capability of the MTPB task ensemble. It was also concluded that 1 hour of testing yielded satisfactory reliability of the composite performance index (11).

Later, the official FAA personnel roster (dated January 1, 1973) was screened by Chiles and West (12) to determine whether each trainee who participated in the earlier study was still listed as having the job code of an ATCS. Thus, the elapsed time from initial testing to the followup ranged from 2 to 2 1/2 years. Since about 90 of the trainees were tested well along in their training program, their tenures in the ATCS program ranged up to more than 3 years on the followup date. Three main questions were of interest: (i) How well did the MTPB measures obtained in the earlier study predict the retention of trainees as controllers? (ii) How well did the criterion (instructors' ratings) used in the earlier study predict retention? (iii) How well did scores on the CSC ATC Aptitude Screening Test predict retention of those trainees for whom scores were available?

It is most important to note that, in general, these trainees represented a very select population (12). At the time of testing, two of the five groups (Studies II and III) had successfully completed the FAA Academy training program and had thus undergone secondary screening. In addition, 28 of the 61 ATCSs in Study II and 60 of the 89 trainees in Study V had been selected originally on the basis of having achieved satisfactory scores on the CSC ATC screening battery (the number of trainees who entered the program on the basis of CSC scores was not determined for the groups in Studies III and IV). Thus, it would be expected that the predictive

validity of all the various measures used would be substantially attenuated and perhaps even nullified because of the range restriction produced by the processes previously at work in eliminating most of the trainees with marginal aptitudes (12).

Point-biserial correlations between the previously obtained performance measures and the retention/termination status of the trainees were computed; correlations were also computed between the criterion used in the earlier study (instructor ratings of trainee potential) and the retention criterion, and between the CSC scores (where available) and the retention criterion. For only one group of trainees was the correlation between the retention criterion and the performance measure significant; this was for the subset of trainees (Study II) who had not been selected on the basis of their CSC scores. Of the 14 correlations between the instructor ratings and the retention criterion, 12 were significant, $p < 0.05$ or better, but none of the correlations between the CSC test scores and the retention criterion was significant (12).

A final determination of the potential of the complex performance battery approach to ATCS selection would require further validating research. Direct support for arguing that the MTPB has predictive power with respect to the potential of ATCS trainees derives from the results with the trainees of Study II. However, indirect support is also provided in that (i) the instructor ratings served as a good predictor of the retention criterion and (ii) the MTPB was a good predictor of the instructor ratings in the earlier study. Should FAA selection policies for ATCSs require an unequivocal assessment of the predictive efficiency of performance tests, the results of these studies provide a basis for optimism regarding the expected outcome of such research (12).

Summary. To date two major efforts have been initiated by CAMI researchers to develop alternative aptitude screening measures. These efforts involved the Directional Headings Test (DHT) and the Multiple Task Performance Battery (MTPB) and were based on early CAMI aptitude test research that suggested that measures developed specifically to screen ATCS trainees might profitably emphasize (i) speeded visual-spatial abilities and (ii) simultaneous multiple task performance requiring the ability to shift attention and time-share several tasks effectively. While both approaches showed promise, research has continued only on the DHT (5).

Research Related to the Uniform Guidelines on Employee Selection

The recent development of a set of uniform guidelines on employee selection procedures, rooted in the 1964 Civil Rights Act Title VII, prompted a different type of research focus at CAMI beginning in 1976, and coincided with both the start of a new pass/fail ATCS training program at the Academy and an accelerated pace of FAA-CSC activity to develop a new ATCS selection battery. One problem that confronted Congress in adopting

Title VII was the use of tests to select candidates for employment without proper evidence that the tests were predictors of job success. Congress attempted to produce a reasonable compromise that would eliminate discrimination and permit the use of tests as practical tools for employee selection. In Title VII, Congress allowed the use of "any professionally developed ability test provided that such test, its administration or action upon the results is not designed, intended or used to discriminate. . ." (52).

Employers used this statement to support using any test developed by a professional, as long as there was no intent to exclude minorities. In 1966, the Equal Employment Opportunity Commission (EEOC) adopted guidelines that stated that the employer's intent was irrelevant if exclusion of minorities was the net effect of using certain tests. The EEOC guidelines declared a test unlawful for use in employee selection unless it predicted performance on the job. In 1965 the Department of Labor adopted the EEOC guidelines. In subsequent years the EEOC and Department of Labor expanded the guidelines to include all selection procedures as well as tests.

In 1971, the Supreme Court ruled in the "Griggs v. Duke Power Co." case that employer practices that systematically eliminated minorities were illegal unless the practices could be justified by "business necessity." Congress confirmed this view in a 1972 amendment to Title VII. During the early 1970's, various government agencies (the Department of Justice and the CSC) developed other guidelines. By 1976 two separate sets of guidelines were in existence. In 1977, the Carter administration authorized an effort to unify the two guidelines. In 1978, the EEOC, CSC, Department of Labor, and Department of Justice adopted the Uniform Guidelines on Employee Selection Procedures (52).

The 1978 Uniform Guidelines (52) state that "employer policies or practices which have an adverse impact on employee opportunities of any race, sex, or ethnic group are illegal under Title VII and Executive Order unless justified by business necessity." The guidelines go further in defining adverse impact as a selection policy where the selection proportion for any race, sex, or ethnic group is less than 80 percent of any other race, sex, or ethnic group. Business necessity is defined as a properly performed validation study. The validation procedures must be a statistical study. Expert or professional opinion or expert/professional development of tests according to the guidelines, is not acceptable in lieu of a proper statistical study (52).

While the 1978 guidelines answered several important questions about discriminatory selection practices, they raised other questions specifically related to ATCS selection procedures, which would have to be taken into account in current efforts to devise a revised battery of the ATC aptitude tests (5). A pertinent feature of these efforts is the fact that data on new tests are derived for the most part from new ATCS trainees, a population that has already passed screening hurdles. In response to one of these questions, Boone and Lewis (7) developed a new procedure to make maximum use

of available information when correcting validity coefficients for restriction in range due to selection. This new procedure was to be employed as a more accurate method of correcting spuriously low correlations between selection tests given to persons already selected for ATCS training (see 5) and ATCS job success. By means of Monte Carlo methods, the procedure was compared to the well-known Gulliksen and Thorndike procedures and was found in general to be a more accurate correction than Gulliksen's or Thorndike's methods, especially when restriction is extreme (7).

To assess the impact of another condition of the Uniform Guidelines, viz, test fairness to racial, ethnic, and sex groups, Lewis (34) used Monte Carlo data containing characteristics similar to those in current ATCS selection data to compare three well-known models of fairness; (i) Thorndike's Constant Ratio Model, (ii) Darlington's Conditional Probability Model, and (iii) Einhorn and Bass' Equal Probability Model. The models were compared for their robustness to sample size differences, different predictor and criterion correlations, and different selection and success ratios. Essentially, the study demonstrated that the models were equally fair under certain specified conditions and the application of the models depended on the goals and aims of the agency involved (34). Lewis emphasized the modification of minority recruitment practices as an effective means of complying with the Uniform Guidelines without necessitating the development of new selection devices.

In another study stimulated by the Uniform Guidelines (52) and the current efforts to develop a new ATC aptitude screening battery (5), the effect of recruitment procedures on correcting validity coefficients for restriction in range was examined by Boone and Lewis (5). The study demonstrated that recruitment procedures can result in widely varying unrestricted variances in applicant groups. Several examples were given. The effects of these variations in applicant group variance on the correction for restriction in range were reviewed with Monte Carlo methods. Changing the selection ratio from 1.5 to 3.0 resulted in a change of the estimated unrestricted validity coefficient from .605 to .755. Recommendations were offered in the study to help minimize this problem. Boone (3), in a report responding to an expressed need from the FAA Office of Aviation Medicine, delineated a mathematical procedure for eliminating outlier data (experimental aptitude test scores) which have a small probability of belonging to the remaining group of data under study. Mathematical derivations for the procedures were also presented. Basically, the process assumes a multivariate normal distribution and uses the Maholanobis D^2 distance function to determine the probability that a particular data vector belongs to the remaining group of data. Examples of how to apply the procedure to experimental test batteries for ATCS selection were given (3).

Summary. During the 1970's, new administrative and legal standards (that resulted from the adoption of the Uniform Guidelines on Employee Selection and from related court decisions) for selection programs coincided with a new ATCS pass/fail training program at the Academy and an accelerated

effort to develop a new ATCS selection battery. These legal changes created special statistical and methodological problems for CAMI researchers that were basic to addressing issues of the validity and fairness of the ATCS selection system. In the last half of the decade, CAMI researchers (i) developed new statistical procedures for correcting validity coefficients; (ii) studied the effect of recruitment practices on validity studies and examined methods for minimizing such effects; (iii) compared several statistical models of fairness to determine which is most appropriate under what conditions; and (iv) developed a statistical procedure for eliminating outlier data in selection test research. The products of these efforts are currently being used in continuing validation research at CAMI.

REFERENCES

1. Air Traffic Controller Career Committee: The Career of the Air Traffic Controller--A Course of Action. Executive Summary of the Report of the Air Traffic Controller Career Committee. Federal Aviation Administration, Washington, D.C., January 1970.

2. Bale, R. M., and R. K. Ambler: Application of College and Flight Background Questionnaires as Supplementary Non-Cognitive Measures for Use in the Selection of Student Naval Aviators, AEROSPACE MEDICINE, 42:1178-1181, 1971.

3. Boone, J. O.: A Statistical Procedure for Eliminating Extreme, Deviant Scores From the Longitudinal Air Traffic Control Data Base. In J. O. Boone and M. A. Lewis: The Selection of Air Traffic Control Specialists: Two Studies Demonstrating Methods to Insure an Accurate Validity Coefficient for Selection Devices. FAA Office of Aviation Medicine Report No. AM-79-14, 1979. (Also published in PSYCHOLOGICAL REPORTS, in press, 1980.)

4. Boone, J. O.: The Relationship of Predevelopmental "150" Training With Noncompetitively Selected Air Traffic Control Trainees to FAA Academy Success. FAA Office of Aviation Medicine Report No. AM-78-10, 1978. (Also published in AVIATION, SPACE, AND ENVIRONMENTAL MEDICINE, 49:1203-1211, 1978.)

5. Boone, J. O.: Toward the Development of a New Selection Battery for Air Traffic Control Specialists. FAA Office of Aviation Medicine Report No. AM-79-21, 1979.

6. Boone, J. O. and M. A. Lewis: An Example of the Effects of Recruitment Procedures on Correcting the Validity Coefficient for Restriction in Range. In J. O. Boone and M. A. Lewis: The Selection of Air Traffic Control Specialists: Two Studies Demonstrating Methods to Insure an Accurate Validity Coefficient for Selection Devices. FAA Office of Aviation Medicine Report No. AM-79-14, 1979.

7. Boone, J. O. and M. A. Lewis: The Development of the ATC Selection Battery: A New Procedure to Make Maximum Use of Available Information When Correcting Correlations for Restriction in Range Due to Selection. FAA Office of Aviation Medicine Report No. AM-78-36, 1978. (Also published in PROCEEDINGS OF THE 20TH ANNUAL CONFERENCE OF THE MILITARY TESTING ASSOCIATION, 1:906-918, 1978.)

8. Brokaw, L. D.: School and Job Validation of Selection Measures for Air Traffic Control Training. Lackland AFB, Personnel Laboratory, Wright Air Development Center Report No. WADC-TN-59-39 (ASTIA Document No. AD 214 884), 1959.

9. Cattell, R. B., and H. W. Eber: The Sixteen Personality Factor Questionnaire. Champaign, Illinois: Institute for Personality and Ability Testing, 1962.

10. Chiles, W. D., E. A. Alluisi, and O. S. Adams: Work Schedules and Performance During Confinement, HUMAN FACTORS, 10:143-196, 1968.

11. Chiles, W. D., A. E. Jennings, and G. West: Multiple Task Performance as a Predictor of the Potential of Air Traffic Controller Trainees. FAA Office of Aviation Medicine Report No. AM-72-5, 1972.

12. Chiles, W. D., and G. West: Multiple Task Performance as a Predictor of the Potential of Air Traffic Controller Trainees: A Followup Study. FAA Office of Aviation Medicine Report No. AM-74-10, 1974.

13. Cobb, B. B.: Problems in Air Traffic Management: II. Prediction of Success in Air Traffic Controller School. FAA Civil Aeromedical Research Institute Report No. 62-2, 1962. (Also published in AEROSPACE MEDICINE, 33:702-713, 1962.)

14. Cobb, B. B.: Problems in Air Traffic Management: V. Identification and Potential of Aptitude Test Measures for Selection of Tower Air Traffic Controller Trainees. FAA Office of Aviation Medicine Report No. AM-65-19, 1965. (Also published in AEROSPACE MEDICINE, 35:1019-1027, 1964.)

15. Cobb, B. B.: The Relationships Between Chronological Age, Length of Experience, and Job Performance Ratings of Air Route Traffic Control Specialists. FAA Office of Aviation Medicine Report No. AM-67-1, 1967. (Also published in AEROSPACE MEDICINE, 39:119-124, 1968.)

16. Cobb, B. B.: A Comparative Study of Air Traffic Trainee Aptitude-Test Measures Involving Navy, Marine Corps and FAA Controllers. FAA Office of Aviation Medicine Report No. AM-68-14, 1968.

17. Cobb, B. B.: Air Traffic Aptitude Test Measures of Military and FAA Controller Trainees. FAA Office of Aviation Medicine Report No. AM-71-40, 1971.

18. Cobb, B. B., C. D. Lay, and N. M. Bourdet: The Relationship Between Chronological Age and Aptitude Test Measures of Advanced-Level Air Traffic Control Trainees. FAA Office of Aviation Medicine Report No. AM-71-36, 1971.

19. Cobb, B. B., and J. J. Mathews: Proposed New Test for Aptitude Screening of Air Traffic Controller Applicants. FAA Office of Aviation Medicine Report No. AM-72-18, 1972. (Also published in AEROSPACE MEDICINE, 44:184-189, 1973.)

20. Cobb, B. B., J. J. Mathews, and C. D. Lay: A Comparative Study of Female and Male Air Traffic Controller Trainees. FAA Office of Aviation Medicine Report No. AM-72-22, 1972.

21. Cobb, B. B., J. J. Mathews, and P. L. Nelson: Attrition-Retention Rates of Air Traffic Control Trainees Recruited During 1960-1963 and 1968-1970. FAA Office of Aviation Medicine Report No. AM-72-33, 1972.

22. Cobb, B. B., and P. L. Nelson: Aircraft-Pilot and other Pre-Employment Experience as Factors in the Selection of Air Traffic Controller Trainees. FAA Office of Aviation Medicine Report No. AM-74-8, 1974.

23. Cobb, B. B., P. L. Nelson, and J. J. Mathews: The Relationships of Age and ATC Experience to Job Performance Ratings of Terminal Area Traffic Controllers. FAA Office of Aviation Medicine Report No. AM-73-7, 1973. (Also published in AEROSPACE MEDICINE, 45:56-60, 1974.)

24. Cobb, B. B., C. L. Young, and B. L. Rizzuti: Education as a Factor in the Selection of Air Traffic Controller Trainees. FAA Office of Aviation Medicine Report No. AM-76-6, 1976.

25. Colmen, J. G.: Review and Evaluation of Present Systems for Selection of Air Traffic Controllers. Report of Phase 1, Task 1. Education and Public Affairs, FAA Contract No. DOT-FA70WA-2371, Washington, D.C., July 1970.

26. Colmen, J. G., and J. Marshall-Mies: Selection of Air Traffic Controllers. Report of Task II, Design of Field Validation. Education and Public Affairs, Washington, D.C., January 1975.

27. Colmen, J. G., and J. Marshall-Mies: Selection of Air Traffic Controllers. Report of Task III, Design of Field Validation. Education and Public Affairs, Washington, D.C., January 1975.

28. Dailey, J. T., and E. W. Pickrel: Development of New Selection Tests for Air Traffic Controllers. FAA Office of Aviation Medicine Report No. AM-77-25, 1977.

29. FAA Office of Personnel: ATCs Intake Study, 1971.

30. FAA Office of Personnel: Evaluation of the "150" Program, 1973.

31. Gordon, L. V.: Clinical, Psychometric, and Work-Sample Approaches in the Prediction of Success in Peace Corps Training, JOURNAL OF APPLIED PSYCHOLOGY, 51:111-119, 1967.

32. Guilford, J. P., and J. I. Lacey, Ed.: Printed Classification Tests. Report 5, Army Air Forces Aviation Psychology Program Research Reports. Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1947.

33. Katzell, M. E.: Expectations and Dropouts in Schools of Nursing, JOURNAL OF APPLIED PSYCHOLOGY, 52:154-157, 1968.

34. Lewis, M. A.: A Comparison of Three Models for Determining Test Fairness. FAA Office of Aviation Medicine Report No. AM-79-3, 1979.

35. Lewis, M. A.: Use of the Occupational Knowledge Test to Assign Extra Credit in Selection of Air Traffic Controllers. FAA Office of Aviation Medicine Report No. AM-78-7, 1978. (Also published in AVIATION, SPACE, AND ENVIRONMENTAL MEDICINE, 49:1155-1159, 1978.)

36. Marshall-Mies, J., and J. G. Colmen: Development of Recommendations for ATCS Selection Tests. Report of Task 1. Education and Public Affairs, Washington, D.C., 1976.

37. Marshall-Mies, J., J. G. Colmen, and O. Domenech: Predicting Success of Applicants for Positions as Air Traffic Control Specialists in the Air Traffic Service. Education and Public Affairs, FAA Contract DOT-FA75WA-3646, Washington, D.C., 1977.

38. Mathews, J. J., B. B. Cobb, and W. E. Collins: Attitudes on En Route Air Traffic Control Training and Work: A Comparison of Recruits Initially Trained at the FAA Academy and Recruits Initially Trained at Assigned Centers. FAA Office of Aviation Medicine Report No. AM-75-3, 1975.

39. Mathews, J. J., W. E. Collins, and B. B. Cobb: A Sex Comparison for Reasons of Attrition of Non-Journeyman FAA Air Traffic Controllers. FAA Office of Aviation Medicine Report No. AM-74-2, 1974. (Also published in PERSONNEL PSYCHOLOGY, 27:535-541, 1974.)

40. Mathews, J. J., W. E. Collins, and B. B. Cobb: Job-Related Attitudes of Non-Journeyman FAA Air Traffic Controllers and Former Controllers: A Sex Comparison. FAA Office of Aviation Medicine Report No. AM-74-7, 1974.

41. Milne, A. M., and J. G. Colmen: Selection of Air Traffic Controllers for Federal Aviation Administration. Education and Public Affairs, Washington, D.C., 1972.

42. Peterson, P. B., and G. L. Lippitt: Comparison of Behavioral Styles Between Entering and Graduating Students in Officer Candidate School, JOURNAL OF APPLIED PSYCHOLOGY, 52:66-70, 1968.

43. Pickrel, E. W.: Development of Occupational Knowledge Tests for En Route, Terminal, and Flight Service Station Air Traffic Control Specialists. Unpublished document, FAA Office of Aviation Medicine, 1977. (Cited in reference 37.)

44. Saleh, S. D., R. J. Lee, and E. P. Brien: Why Nurses Leave Their Jobs--An Analysis of Female Turnover, PERSONNEL ADMINISTRATION, 28:25-28, 1965.

45. Sullivan, E. T., W. W. Clark, and E. W. Tiegs: Manual for California Test of Mental Maturity, Advanced Form. Los Angeles, California, California Test Bureau, 1957.

46. Trites, D. K.: Problems in Air Traffic Management: I. Longitudinal Prediction of Effectiveness of Air Traffic Controllers. FAA Civil Aeromedical Research Institute Report No. 61-1, 1961. (Also published in AEROSPACE MEDICINE, 32:1112-1118, 1961.)

47. Trites, D. K.: Problems in Air Traffic Management: VI. Interaction of Training-Entry Age With Intellectual and Personality Characteristics of Air Traffic Control Specialists. FAA Office of Aviation Medicine Report No. AM-65-21, 1965. (Also published in AEROSPACE MEDICINE, 35:1184-1194, 1964.)

48. Trites, D. K., and B. B. Cobb: Problems in Air Traffic Management: III. Implications of Age for Training and Job Performance of Air Traffic Controllers. FAA Civil Aeromedical Research Institute Report No. 62-3, 1962. (Also published in AEROSPACE MEDICINE, 35:336-340, 1964.)

49. Trites, D. K., and B. B. Cobb: Problems in Air Traffic Management: IV. Comparison of Pre-Employment, Job-Related Experience With Aptitude Tests as Predictors of Training and Job Performance of Air Traffic Control Specialists. FAA Civil Aeromedical Research Institute Report No. 63-31, 1963. (Also published in AEROSPACE MEDICINE, 35:428-436, 1964.)

50. Trites, D. K., and B. B. Cobb: CARI Research on Air Traffic Control Specialists: Age, Aptitude, and Experience as Predictors of Performance. FAA Civil Aeromedical Research Institute Unnumbered Report, 1964.

51. Trites, D. K., M. C. Miller, and B. B. Cobb: Problems in Air Traffic Management: VII. Job and Training Performance of Air Traffic Control Specialists--Measurement, Structure, and Prediction. FAA Office of Aviation Medicine Report No. AM-65-22, 1965. (Also published in AEROSPACE MEDICINE, 36:1131-1138, 1965.)

52. Uniform Guidelines on Employee Selection Procedures: Federal Register, Volume 43, No. 166:38290, August 25, 1978.