Doc
FAA
AM
88
03

88/3

on Medicine
C. 20591

Studies Of Poststrike Air Traffic Control Specialist Trainees: II. Selection And Screening Programs

Carol A. Manning
Pamela S. Kegg
William E. Collins

RECEIVED
OCT 6 1988
FAA LIBRARY

Civil Aeromedical Institute
Federal Aviation Administration
Oklahoma City, OK 73125

July 1988

This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161.

U.S. Department of Transportation
Federal Aviation Administration

Doc
FAA
AM
88
03

Technical Report Documentation Page

1. Report No. DOT/FAA/AM-88/3	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle STUDIES OF POSTSTRIKE AIR TRAFFIC CONTROL SPECIALIST TRAINEES: II. SELECTION AND SCREENING PROGRAMS		5. Report Date JULY 1988
		6. Performing Organization Code
7. Author(s) Carol Manning, Pamela S. Kegg and William E. Collins		8. Performing Organization Report No.
9. Performing Organization Name and Address FAA Civil Aeromedical Institute P. O. Box 25082 Oklahoma City, OK 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

This work was performed under tasks AM-C-88-HRR-105 and AM-D-88-HRR-112.

16. Abstract

Specific contributions of aviation psychologists to the selection and Academy training of FAA air traffic control specialists are presented in an historical context. Research results which formed the basis for the written aptitude selection tests, The Occupational Knowledge Test (for assessing prior experience), and the pass/fail screens for the Academy's nonradar and radar programs are noted. Results of continuing validation research on all aspects of these selection/screening programs are presented in detail as well as the means by which the results are used to (i) predict training and performance outcomes, (ii) model the impact of program changes (iii) allow the introduction of changes (improvements) without compromising the validity of the programs, and (iv) assure compliance with the Uniform Guidelines on Employee Selection Procedures. The application of this line of research has produced considerable cost/ benefits to the agency.

17. Key Words Air Traffic Control Specialists, Selection, Training, Performance	18. Distribution Statement 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	23	

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

STUDIES OF POSTSTRIKE AIR TRAFFIC CONTROL SPECIALIST TRAINEES: II. SELECTION AND SCREENING PROGRAMS

The Air Traffic Control Specialist (ATCS) occupation in the Federal Aviation Administration (FAA) consists of three specialties, the enroute, terminal, and flight service station (FSS) options. Enroute and terminal specialists are responsible for ensuring the separation of aircraft either traveling between airports (enroute) or approaching or departing airports (terminal) through the issuance of clearances (instructions regarding allowable altitudes and directions of flight) to pilots. Flight Service Station specialists provide services to pilots such as briefing them on weather patterns, filing flight plans, and helping to locate lost aircraft. However, FSS specialists are not responsible for ensuring aircraft separation. Because the job of an FSS specialist involves different knowledges and skills from the job of enroute and terminal ATCSs and because their selection procedures differ, this report will discuss only selection and screening programs that apply to specialists assigned to the enroute and terminal options.

ATCS functions are very demanding and are of a critical nature. They require processing large amounts of information presented simultaneously through both visual and auditory sensory channels. Also required is visualization and manipulation of three-dimensional information which is, at best, presented in two dimensions. An ATCS must quickly and accurately evaluate incoming information, make precise and error-free decisions, and take action upon those decisions. Through application of established rules and procedures, as well as utilization of their own ability to prioritize and strategize events in a constantly changing environment, controllers are a crucial part of the National Airspace System (NAS).

The continued safety of the NAS requires that ATCSs be carefully selected and trained. Each candidate for the occupation is continually evaluated, from an initial aptitude selection test battery through grueling performance-based screening at the FAA Academy, and finally in on-the-job training, conducted at the assigned facility. Because of the safety-related, critical aspects of the job, identifying and screening for characteristics in individuals that will predict success in air traffic is especially important. In fact, research has demonstrated that not all individuals have aptitudes required to perform the duties required of an ATCS.

FAA aviation psychologists have been instrumental in providing a framework for the development of ATCS selection and training methods currently in use. Principles and methodologies from several specialized fields of psychology have been utilized to derive the research data upon which the current selection and training procedures are based. For example, the field of industrial/organizational psychology historically has developed approaches to selecting from a general pool of applicants individuals who are best suited for a job, through identification of unique qualities of both the job and of individuals who are particularly successful on that job. In many cases in the applied setting, establishing selection procedures is a process of continual development and refinement of methodologies to improve the

selection process and outcomes. The development of the FAA's procedures for selecting air traffic controllers presents an example of a selection process which has undergone a number of refinements over time. At each point that a new measure was added to the ATCS selection process, the change was based upon research findings and was modeled, evaluated, and monitored to establish its effectiveness. This report focuses on the application of psychological principles of personnel selection in hiring ATCSs. To provide a background on the profession, the following section reviews the career progression of an ATCS.

ATCS CAREER PROGRESSION

An individual interested in applying for an ATCS position must meet a number of evaluative criteria as shown in Table 1.

Table 1
ATCS INITIAL SELECTION REQUIREMENTS

* High school education or equivalent
* 3 years of general work experience (years in college may equate)
* 18 to 30 years old
* Qualification on Aptitude Tests (Office of Personnel Management Air Traffic Controller Test Battery)
* Medical Qualification (Class II airman's medical examination and minor psychiatric, the 16 Personality Factors Test)
* Security Clearance

Credit for demonstrated knowledge in air traffic control and for veteran's preference is added by the United States Office of Personnel Management (OPM) to the selection rating if the applicant's performance on the aptitude test battery results in a score of at least 70. Although those who successfully meet the qualifications in these areas are eligible for hiring, only those with the highest total ratings (typically 90 or above), are selected for hire.

The maximum age limitation was based on a series of investigations, beginning in 1961, which indicated that training attrition rates for trainees 31 years of age and older were generally two to three times higher than for younger trainees (Collins, Boone, and VanDeventer, 1981). Other experimental studies in which ratings of job performance were collected on full performance level (FPL) ATCSs from both their supervisors and their peers, revealed that mean performance ratings of controllers within every age category beyond 40 were significantly lower than those of the younger subgroups. Such findings

played a decisive role in obtaining Congressional legislation in 1972 which established an optional early retirement program for controllers and imposed an upper age limit of 30 in the selection of controller trainees. More recent analyses relating age and Academy performance since the 1981 ATCS strike (e.g., VanDeventer and Baxter, 1984) continue to demonstrate the negative relationship of age with Academy success.

ATCS candidates must successfully complete two additional aptitude screening processes prior to attaining the journeyman (or full performance) level: the performance-based FAA Academy nonradar and radar screens. Qualified candidates who accept appointment first proceed to the nonradar screening program at the FAA Academy in Oklahoma City. In this three month program, 60% of the final grade is based on student performance on a series of laboratory problems that simulate the control of air traffic under nonradar conditions. Candidates who are unsuccessful in the nonradar "Screen" are removed from the FAA.

Those who pass the Academy screening program (now called "developmentals") are assigned to an air traffic control facility. Facilities are of differing levels of complexity and perform different types of air traffic services. Upon arrival at their facility, developmentals undergo phases of nonradar training, each conducted by the facility in a pass/fail mode, that emphasize the procedures appropriate for that facility's function and location. Developmentals eventually return to the Academy for a radar screening program, which assesses their aptitude to control air traffic in a radar environment. Additional field training, conducted for those who pass the radar screen, eventually results in a successful developmental's certification to perform various air traffic functions independently. This training program requires, on the average, 2.8 years for enroute developmentals to complete and between .5 and 1 year for terminal developmentals, depending on the type of facility to which they are assigned.

BACKGROUND OF ATCS SELECTION PROCEDURES

General Selection Guidelines. The principles governing the development and use of ATCS selection and testing procedures have their foundation in aviation psychology applications stemming back to World War I when psychologists at Kelly field devised tests that were successful in predicting performance in military flight training (Bond, Bryan, Signey, and Warren, 1968). World War II resulted in an expanded role for psychologists in all aspects of selection research for military needs, including flight training. The demonstrated utility of those efforts in identifying successful training candidates, reducing attrition, shortening training time, and producing cost-benefits led to a dramatic post-World War II expansion of personnel testing by industry, with little systematic evaluation or regulation. As a result, utilization of test scores for employment purposes sometimes resulted in the exclusion of various minority groups from employment. In 1978 the Uniform Guidelines on Employee Selection Procedures were adopted, which stated 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"

(Uniform Guidelines on Employee Selection Procedures, 1978). Consequently, subsequent research on selection procedures has been influenced both by the results of past research and by more recent legislation.

ATCS Research. Research on ATCS selection is rooted in a 1950 contract by the Civil Aeronautics Administration (CAA--precursor to the FAA) for the development of paper-and-pencil aptitude tests that could be used to select ATCS trainees. The results of the study indicated that some aptitude tests reflecting content areas identified from a job analysis potentially could make an effective contribution in the selection process. This study also provided the format for an Air Traffic Problems (ATP) test.

In 1956, the United States Air Force's Personnel Laboratory, in a joint effort with the CAA, administered thirty-seven aptitude tests to ATCS trainees (Brokaw, 1984). The findings indicated that a composite aptitude test score--created by adding scores on tests of arithmetic reasoning, symbolic reasoning, code translation, and the ATP test--effectively predicted instructors' ratings of training performance and supervisors' ratings of job performance approximately one year after training.

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; see Collins, et al, 1981). Under the direction of Trites and Cobb, CAMI research on Civil Service Commission (CSC) tests of similar content to the tests originally studied by Brokaw led the CSC for the first time to establish selection tests for screening ATCS trainees in 1962. The initial selection battery consisted of the following tests: Arithmetic Reasoning, Spatial Relations, Following Oral Directions, Abstract Reasoning, and Air Traffic Problems. During the period between 1962 and 1972, in addition to continuing validation research, aviation psychologists studied a number of factors relevant to ATCS selection and screening: attrition from the profession, age, prior experience, education, sex, and military ATC training. The research on age led directly to the establishment of an age 30 screening standard in 1973. Since that time, ATCS selection standards have remained essentially the same, although the aptitude test battery was changed in 1981.

In summary, this early line of ATCS selection research employed a number of commercial tests to identify an array of aptitudes having a demonstrated utility for identifying individuals with a high probability of success in the field of air traffic control. In addition to the commercial tests, aviation psychologists began developing tests (e.g., directional headings, air traffic problems) which had more direct relevance to this aviation career.

During the latter half of the 1970s, research focused on evaluation of the impact of the special entry "150 Program" for disadvantaged persons which was implemented in 1969, on studies relevant to the Uniform Guidelines on Employee Selection Procedures, on developing field performance ratings, on the development of a computerized, longitudinal ATCS database for continuing validation research, and on studies of optimal combinations of old and new aptitude screening measures to form a potential new aptitude screening battery (see Collins, et al, 1981). Thus, research began on the Multiplex Controller Aptitude Test, a major component of the current ATCS selection

battery initiated in 1981.

The development of a longitudinal ATCS database, begun in 1976, has been instrumental in allowing the accomplishment of research conducted during the 1980s. In the wake of the 1981 ATCS strike, the database proved to be invaluable for policy decisions directed at accelerating recovery from the loss of approximately 11,000 fired ATCSs (nearly two-thirds of the controller workforce at that time). Numerous alternate strategies were modeled and the potentially most effective policies were selected for implementation. For example, the use of the database allowed projections to be made regarding the number of students that must be trained, assessment of the impact of temporarily increasing the age limit for new hires, and identification of placement strategies which would decrease attrition. Demographic data were utilized to identify groups of individuals which might viably bypass the Academy and be successful in the field. Currently, besides supporting the investigation of Equal Employment Opportunity discrimination complaints and providing objective information used in making management-level decisions, the database is used to make statistical projections of the likely outcomes of program changes. One such program change, ("the common screen") a recent and major modification to the FAA Academy screening program, resulted from extensive data analysis and data-based decision-making. The factors leading to the change to "the common screen" are described later in this report.

ATCS SELECTION PROCEDURES: OPM SELECTION BATTERY, FAA ACADEMY NONRADAR AND RADAR SCREENING PROGRAMS

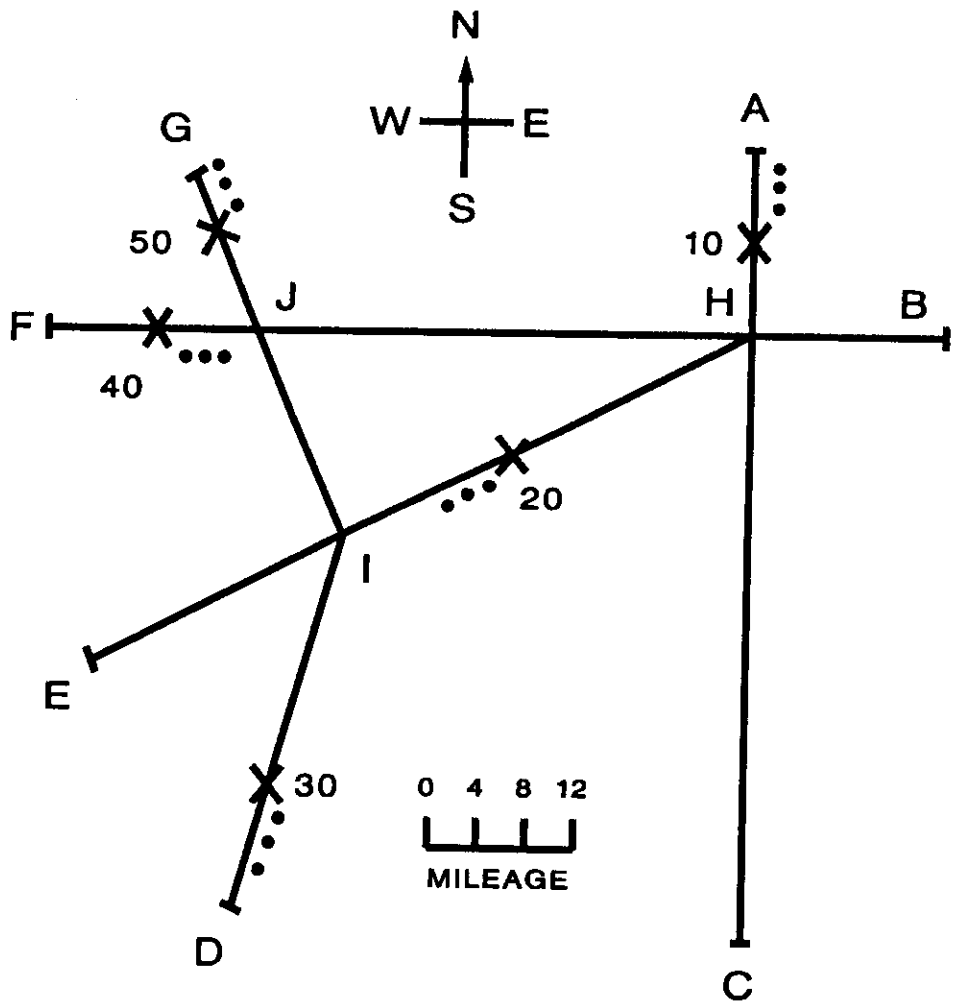
OPM SELECTION BATTERY.

The current OPM selection battery was implemented in October 1981, about two months after the ATCS strike. The battery consists of three tests: the Multiplex Controller Aptitude Test (MCAT), the Abstract Reasoning Test which was included in the previous selection battery, and the Occupational Knowledge Test (OKT).

MCAT. The MCAT was developed to establish a test with higher predictive validity than tests included in the CSC selection battery. The MCAT was designed to approximate, in a simulated air traffic setting, activities required for the control of air traffic. Many of the MCAT items involve identifying situations that may result in conflicts between aircraft. Other items require computing time-distance functions, interpreting information, and spatial relations.

The original version of the MCAT, developed at the FAA's Technical Center, presented air traffic situations on film. Subjects observed radar-type situations as they developed, and answered questions about the changing relationships of aircraft at predetermined intervals. To reduce logistical problems, the test was next administered in a film and slide format, then on slides only, and finally in a paper-and-pencil format (Dailey and Pickrel, 1984).

The air traffic situations included in the paper and pencil version of the



Aircraft	Altitude	Speed	Route
10	5000	300	AHC
20	5500	300	EIHB
30	5000	450	DIJF
40	6000	450	BHJF
50	6000	300	GJBH

Figure 1. Sample information from Multiplex Controller Aptitude Test (MCAT).

MCAT provide an air route map showing permissible routes of flight through the sky. Aircraft locations are indicated on the air route map, as are locations where the routes of flight intersect. A table accompanying the map includes other relevant information, such as aircraft altitudes, speeds, and planned routes of flight (see Figure 1).

Extensive research suggested that MCAT scores were more highly correlated with ATCS Academy screening program performance than were many of the CSC tests (Rock, Dailey, Ozur, Boone, and Pickrel, 1981). Other analyses projected that the costs of training entrants who would eventually fail in the initial qualification screening program could be reduced by about three million dollars per year by replacing the operational CSC battery with a new selection battery consisting of the MCAT, the Abstract Reasoning test, and the Arithmetic Reasoning test, and using the Occupational Knowledge Test (see below) to award extra credit toward the earned rating. The Arithmetic Reasoning test was later removed from the battery due to its contribution toward unequal prediction of pass/fail status for certain minority groups. The resulting battery consisted of two versions of the MCAT (each assigned weights of 2 toward the computation of the composite score) and the Abstract Reasoning Test (assigned a weight of 1), and the OKT, used to award extra credit points.

OKT. The Occuational Knowledge Test (OKT) was developed to provide a more objective and reliable measure of ATCS job knowledge. Before 1981, extra points were awarded to an applicant's CSC rating based upon their claim of prior job-related experience. Items on the OKT, which cover seven knowledge areas related to air traffic control, were developed in conjunction with subject-matter experts, and were experimentally administered to incoming students and full performance level controllers prior to implementation.

Experimental administrations of the OKT to new hires suggested that OKT scores were more predictive of future success in the screening program than were ratings based on self-reported prior experience (Lewis, 1978). Addition of the OKT to the then-experimental selection test battery (including the MCAT) also increased the proportion of the variance in initial training scores accounted for by the selection test battery by 8% (a statistically significant amount). However, when the OKT was eventually added to the operational battery, the earned OKT score did not count toward qualification, but was used only to provide additional points to applicants who had already qualified on the basis of their performance on the other selection tests.

Results of implementation of current OPM selection battery.

As noted above, the current OPM selection battery, consisting of the MCAT, Abstract Reasoning Test, and OKT, was implemented in October 1981. VanDeventer's (1984a) comparison of OPM ratings with earned scores in the Academy pass/fail screening program suggested that the new OPM battery was more predictive of Academy success than the previous battery; a 7% increase was observed in the Academy pass rate for candidates who took the new battery as compared with those who took the former battery.

Analyses of the distributions of MCAT, Abstract Reasoning Test, and OKT scores were conducted after three operational test administrations (VanDeventer, 1984b). The analyses revealed that mean scores for each of the three component tests were slightly higher on the first operational administration than was estimated during experimental administrations of the test battery (the mean operational MCAT scores for versions 1 and 2 and the Abstract Reasoning score were 30.9, 37.1, and 30.2, respectively, as compared with the experimental scores of 30.6, 35.6, and 28.8). As a result, the distribution of test scores shifted upward and a higher percentage of applicants earned qualifying scores than had been projected.

The analyses also revealed that the mean MCAT scores increased as applicants repeated the test; in other words, the MCAT was learnable. For example, those repeating the MCAT once during the three operational administrations increased their scores on the first version of the MCAT by between 6.8 and 7.6 points on the second time they took the test; those repeating twice increased their scores by 11.6 points. Other test means in the battery did not increase as much as a function of test repetition (e.g., those repeating the Abstract Reasoning Test once increased their scores by between 2.3 and 2.8 points, while those repeating twice increased their scores by 4.6 points; changes in OKT scores ranged from a decrease by .3 to an increase of 2.3 points for those repeating once, and a decrease of 1.1 points for those repeating twice.) The results of analyses of earned scores across operational administrations suggested that earned ratings (based on scores from all three tests) increased between 6.1 and 7.0 points, on the average, for those who repeated once, and by 8.8 points for those repeating twice.

The increase in earned OPM ratings as a result of test repetition had a considerable impact on the subsequent screening process. Additional analyses revealed that the OPM rating earned during the initial testing session was more predictive of final disposition in the Academy pass/fail screen than was the rating earned after test repetition. While Academy entrants hired with OPM ratings of 95 and above had an Academy pass rate of 74.0%, those whose rating at first testing was between 90 and 94 had a pass rate of 72.3%; those whose rating at first testing was between 85 and 89 had a pass rate of 64.9%; those whose rating at first testing was between 80 and 84 had a pass rate of 58.9%. As mentioned above, an overall 7% increase in the Academy pass rate was observed after the implementation of the new battery. It is possible that an even higher increase in the pass rate might have occurred but for the inclusion of candidates who finally earned qualifying OPM scores by repeatedly taking the test.

OPM policy initially prevented the restriction of the number of testing repetitions allowed an applicant. However, changes to the testing process were made recently by OPM as a result of the problems mentioned above that were associated with increasing MCAT scores. These changes included implementing new MCAT versions determined to be more difficult than previous versions, reducing the amount of time allowed to take the MCAT, and limiting to one every 18 months the number of repetitions allowed for applicants who earn a qualifying score. At the same time, semiannual, scheduled administrations were eliminated; applicants are now allowed to take the tests at any time during the year.

ATCS NONRADAR SCREENING AT THE FAA ACADEMY.

Some history. As a result of concerns about standardization in ATCS training, the first centralized pass/fail program was implemented in March 1956 at the FAA Academy located at the (now) Mike Monroney Aeronautical Center in Oklahoma City (Boone, 1984). The pass/fail program in 1956 was eight weeks long and consisted of two parts: academic and laboratory. The method for assessing student performance in this program was based upon instructors' assessments. Few people failed the academic portion; however, about 30% failed the laboratory phase. This course served as the basic foundation for the pass/fail FAA Academy screening program in use today, although it has undergone a number of modifications over the years.

During various times over the next two decades, the program underwent periods of centralization and decentralization. Because the Civil Rights Act of 1964 and the resulting administrative guidelines focused attention on fair employment practices, a series of studies on the program was undertaken in the early 1970s. Until this time, the program had been conducted mostly as training, without much structured evaluation of outcomes. Some data were collected, but no systematic item analysis, reliability measurement, or validation studies were conducted on student performance measures. At the time, the student performance measures consisted predominantly of the expert judgments of full performance level controllers. Studies in the early 1970s questioned the reliability and validity of the training program (Boone, 1984).

In addition, congressional hearings held by the Committee on Government Operations in 1975 concluded that the FAA's basis for selecting ATCSs was inadequate for reasonable prediction of individuals with the potential for success in completing training, and that ATCS trainee attrition from the occupation (which typically occurred about two to three years into training) was unacceptably high. The Congressional committee recommended the development of a standardized, centralized, validated program designed to identify as early as possible and remove promptly from training those candidates demonstrating insufficient aptitude to become ATCSs. As a result of these recommendations, the FAA committed to developing a program designed to decrease the costs of attrition and improve the selection of ATCSs. The FAA's aviation psychologists collaborated to provide a scientific framework required for the design, implementation, and evaluation of this program.

The result of this collaboration was the implementation of two option-specific (enroute and terminal) pass/fail programs. The pass/fail nature of the process meant that an individual lost a job as an ATCS if he/she was unable to pass the screening phase. Since 1976, CAMI established and maintained a database containing all test and laboratory problem data from the two programs. In the continuing effort to increase the cost benefits of the selection process while maintaining a valid program the option-specific Academy programs were consolidated into a single initial screening phase, "the common screen."

A description of the 1976 Academy pass/fail program. The goal of the 1976 program was to assess the aptitude of an individual with no prior knowledge of the occupation by allowing him or her to learn a structured set of air

traffic control rules and principles, and then providing a series of laboratory job simulations in which the student has the opportunity to demonstrate the application of those principles. The student being evaluated performs the duties of an ATCS using standardized, timed scenarios encompassing the movement of aircraft through a specified airspace. Another student performs the roles of the aircraft pilots and other "controllers" participating in the scenarios.

Each option had a different program consisting of separate academic and laboratory phases. Both programs contained pass/fail segments covering (i) an introduction to aviation and air traffic control and (ii) procedures for nonradar control of air traffic in the type of airspace configuration relevant to the option. The terminal program also included a Control Tower Operator phase, which was not pass/fail. The nonradar screening phase for both options included academics, laboratory simulation problems, and a controller skills test (CST). The CST measured application of air traffic control principles to resolve air traffic situations in an objective paper-and-pencil format. The pass/fail composite Academy score for both options was calculated from the weighted sum of scores on the academics, laboratory problems, and the CST. A grade of 70 was required to pass.

Neither option-specific program remained static. Over time, through coordination between management, educators, and aviation psychologists, changes were made to increase the functionality of the programs. These changes included: increasing the number of laboratory problems from four to six; initiating the scoring of procedural errors; and altering the number of laboratory problems which counted toward the final grade from six to the best five of six.

Other minor changes occurred periodically. For example, new versions of academic tests and CSTs and new laboratory problems were introduced. Each change, while reducing the chance of program compromise, could potentially have had a negative impact on the validity of the program if it changed significantly the content or the measurement devices used in the program. Consequently, each change was overseen by CAMI researchers, who statistically modeled the impact of any proposed change on program validity and on student success rates. The results of those analyses determined whether or not the change would be effected. Another major change to the program, which will be discussed later in the report, occurred in October 1985, with the implementation of the "common screen".

The Academy screening programs, unlike any other ATCS training programs, have been validated and systematically evaluated since their implementation in 1976. Through the involvement of CAMI's aviation psychologists in both consulting with Academy program developers and analyzing the longitudinal database, Academy screening programs are continually monitored and proposed changes are very carefully modeled prior to implementation. Such activities ensure that the program remains a valid screen. Some of the results of the ongoing analyses, focusing on the nonradar screening, are presented below.

Table 2

ATCS SCREENING ASSESSMENT - ACADEMY NONRADAR LABORATORY PHASE

		JAN 1976 TO AUG 1981		AUG 1981 TO SEPT 1985		TOTAL	
ENROUTE	GRADUATES	2765	67%	4463	52%	7228	57%
	FAILURES	1042	25%	3041	36%	4083	32%
	WITHDRAWALS	331	8%	1032	12%	1363	11%
	ENTRANTS	4138		8536		12674	
TERMINAL	GRADUATES	3058	74%	3387	68%	6445	70%
	FAILURES	845	21%	1312	26%	2157	24%
	WITHDRAWALS	221	5%	298	6%	519	6%
	ENTRANTS	4124		4997		9121	
TOTAL	GRADUATES	5823	70%	7850	58%	13673	63%
	FAILURES	1887	23%	4353	32%	6240	28%
	WITHDRAWALS	552	7%	1330	10%	1882	9%
	ENTRANTS	8262		13533		21795	

Some pre- vs. post-strike comparisons. Table 2 compares performance of students who entered the FAA Academy before and after the 1981 ATCS strike (until October 1985, when the new "common screen" was implemented). Pass rates for the en route option were lower than for the terminal option both before and after the strike because the enroute program was designed to be more difficult than the terminal program. Furthermore, for both the enroute and terminal programs, pass rates prior to the strike were higher than pass rates after the strike.

A review of biographical data collected from Academy entrants (Collins, Manning, and Taylor, 1984) revealed a shift in the characteristics of the candidate population after the 1981 ATCS strike, which might be responsible for the difference between the prestrike and poststrike pass rates. As seen in Table 3, although only 30% of the entrants had no prior aviation-related experience before the strike, over two-thirds of the entrants had no aviation-related experience after the strike. Also, the proportion of students reporting ATC operations (military) experience after the strike was lower than the corresponding group before the strike because, to speed the strike recovery, most applicants with substantive operations experience entered through an alternative hiring program and bypassed the Academy. While 38% of entrants prior to the strike had a college degree, 43% of the post- strike entrants had a degree (however, other studies show no significant relationship between the amount of education attained and Academy performance, Collins, et al, 1981).

Table 3

PROFILE OF ACADEMY ATCS ENTRANTS

	PRE-STRIKE 1976-1981 (N=6059)	POST-STRIKE 1981-1988 (N=18,397)	
EXPERIENCE	30%	67%	NO RELATED EXPERIENCE
	20%	14%	AVIATION ONLY
	50%	19%	ATC OPERATIONS
AVERAGE AGE	26.5	26.3	RANGE 18-31, 18-35
EDUCATION	14%	10%	HIGH SCHOOL ONLY
	48%	47%	SOME COLLEGE
	38%	43%	COLLEGE DEGREE
SEX	83%	85%	MALES
	17%	15%	FEMALES
MINORITY STATUS	13%	8%	MINORITIES
	87%	92%	NON-MINORITIES

The next section discusses the methods by which the validity of the screening programs is assessed.

Field training status as a criterion measure of performance.

Basic considerations. One method for assessing the validity of the Academy program is to compare performance at the Academy with some measure of job performance. It is desirable to assess program validity for two reasons: first to comply with the United States' Civil Rights regulations as specified in the Uniform Guidelines for Employee Selection Procedures (1978); and second, to improve the cost effectiveness of the FAA's selection procedures. No selection procedure is infallible; with any procedure, some people are selected who do not complete training, while others are not selected who might have performed the job effectively. However, fewer of these kinds of errors are made in programs of higher validity.

To demonstrate the validity of selection procedures, it is desirable to compare performance on the various selection procedures (OPM selection test scores and Academy scores) with criterion measures of on-the-job performance, that is, how well the full performance level (FPL) ATCS controls air traffic. However, to date, no such measure has been developed for ATCSs. Instead, criterion measures of training performance are typically used. These measures include attrition from training, final disposition in the career field (particularly for individuals who failed at their initial facility and were subsequently reassigned to a facility of lower complexity), pass rates in each phase of training, supervisor ratings of training performance, time

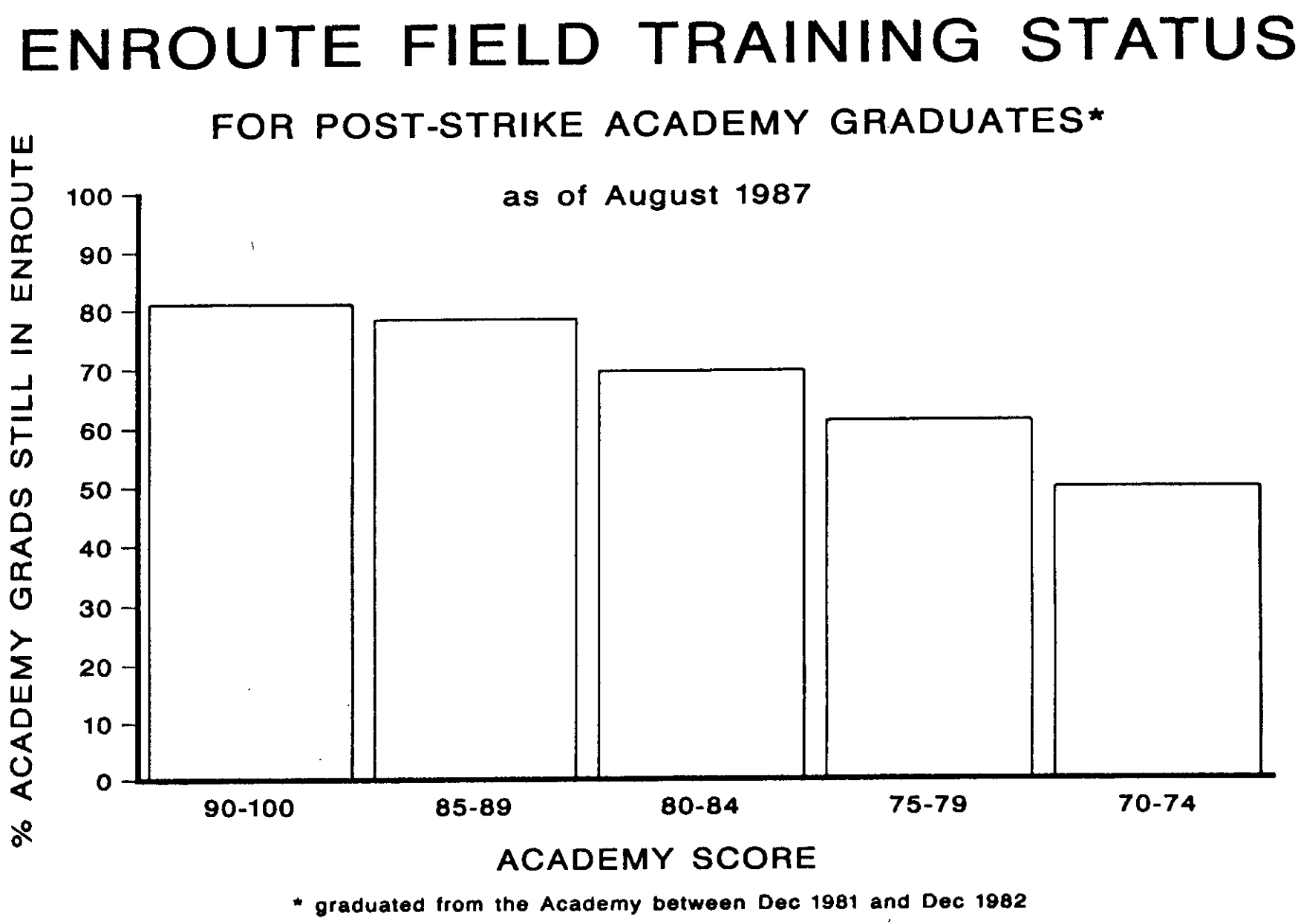


Figure 2. Relationship between Academy score and retention in enroute field training.

to complete training, and time to attain FPL status. Of course, these measures are influenced by the facility to which the developmental is assigned; some facilities are more complex than others, and the training they provide is more difficult than the training provided at less complex facilities. Some facilities perform different air traffic functions and consequently their ATCSs require different types of training. Thus, it may not be appropriate to average the criterion measures across different types of facilities.

Analysis. Several efforts have been made to examine the validity of Academy screening programs using the criterion measures described above. VanDeventer (1981) was the first to examine formally the criterion-related validity of the Academy screening program. Attrition information and supervisor ratings were obtained for a sample of successful Academy graduates about three years after their graduation. VanDeventer found that the correlations between the Academy composite score and the field supervisors' ratings were an acceptable .56 for those in the enroute option and .39 for the terminal option (corrected for restriction in the range of Academy scores).

Other analyses have been performed to compare the validity coefficients VanDeventer (1981) obtained for prestrike Academy graduates with validity coefficients computed for poststrike Academy graduates. The analyses suggest that program validity has remained fairly consistent, in spite of the changes mentioned above in the characteristics of the population of poststrike Academy entrants. Furthermore, the instructor's assessment of performance on the laboratory problems, the most subjective measure of performance, was found, for all groups of Academy graduates, to have the highest correlation with retention in field training as compared to the remaining components of the Academy composite score (the rest of which are measured in a more objective manner.) Figure 2 shows a graphic representation of the relationship between the Academy score and field training retention for the enroute option. It is clear from examining the figure that only about half of those Academy graduates who earned the lowest passing Academy scores are successful in later field training. Of the different types of field training, the training provided at enroute and high-complexity terminal facilities is considered most difficult. If students earning lower Academy scores could be assigned to lower complexity facilities, it can be seen from examining the figure that the result would be a lower field training attrition rate at the higher complexity facilities, while other information suggests that such assignment would not significantly impact field training attrition at lower complexity facilities.

Considerations for implementing a more efficient Academy screening program (the "Common Screen.")

A change was made in October 1985 to consolidate the Academy's enroute and terminal option-specific screening programs into a single program, the "common screen." Ongoing analyses of the data in the CAMI longitudinal database, such as those described above, had begun to point to areas in which the program's effectiveness and cost benefits could be enhanced. Air traffic control personnel and aviation psychologists collaborated to review data and model various options. A proposal recommending implementation of the "common screen" was submitted to FAA management in 1984.

Some Background. The following factors influenced the decision to convert to a single "common screen": 1) Assigning students with higher aptitude to more complex facilities should result in a reduction in field training attrition at those facilities. 2) In planning the "common screen", phases of training not required to assess an individual's aptitude in the screening process could be resequenced so that only information required to complete the "common screen" would be presented during that phase. The remaining information would be presented later to those who passed the screening phase. As a result, it was expected that the FAA would spend less money training students who would eventually fail. 3) Training phases, which would be resequenced for presentation after screening phases, could be modified more flexibly in responding to changes in the dynamic ATCS occupation, because they would no longer fall strictly under the Uniform Guidelines on Employee Selection Procedures. (Changes to screening programs are often restricted until the change can be assessed for fairness and potential impact on program validity.) 4) The phases conducted at the Academy after the screen could be expanded to include some of the training previously conducted at field facilities. This would place less of a burden on field facilities to provide training, and result in more standardized training due to the use of a centralized training facility.

A Data-based Decision. A committee comprised of FAA administrative personnel involved in ATCS training was formed to examine issues surrounding the development and implementation of the "common screen" program. The committee decided that the new screen would be based on the option-specific enroute Academy program which, as described in the previous section, was found to have higher validity than the terminal program. As mentioned above, using a program of higher demonstrated validity means that fewer selection errors will be made.

The pass rate was also given close attention before implementing the new program. Statistical simulations of the effect of changing the weights on program components were conducted to identify weights that could be expected to produce the desired pass rate of 60%. Other simulations were conducted which suggested that changing the weights should have no significant impact on program validity. As a result of manipulating the weights, it was not necessary to change the content of the former program or the way labs were graded to achieve the desired pass rate.

Prior to the "common screen", applicants were assigned to specific options and facilities before they were formally hired, and, for the most part, no information about the aptitude of the applicant was used in making the assignment. When the "common screen" was implemented, procedures were changed so that facility assignments were made after students had passed the screen. This allowed the use of the Academy score, which had been shown to predict success in field training (see previous section), to make facility assignments. In principle, Academy graduates with higher scores would be placed in more complex facilities, and graduates with lower scores would be placed in less complex facilities.

Effectiveness of the "common screen."

In the first two years of the "common screen", the overall pass rate for the screen (60.2%) was very close to the projected 60% pass rate. Pass rates for individual groups of classes that enter the Academy at the same time may vary, but that is to be expected.

Adverse impact against minorities. Because a screening program in the United States is a type of selection procedure, due to its exclusion from further employment of individuals who fail, it must comply with the Uniform Guidelines on Employee Selection Proceedures (1978). Therefore, the screening programs must be evaluated periodically for their potential for adverse impact against protected groups. While adverse impact is not determined by statistical analysis alone, the Uniform Guidelines provide a "rule of thumb" for use in determining whether or not a selection procedure differentially impacts two or more groups in selecting one group over another. An "impact ratio" is calculated for each minority group by dividing each target group's pass rate by the pass rate of the group with the highest rate (typically nonminorities). The rule of thumb (also called the "4/5ths" rule) states that if the impact ratio for any minority group is greater than or equal to .80, the program is usually assumed to demonstrate no adverse impact. If the ratio is less than .80, it suggests that the program has adverse impact for a specific group.

TABLE 4
POST-STRIKE ACADEMY PASS FAIL RATES
BY MINORITY STATUS
THE SCREEN PROGRAM
January 1986 through January 1988 graduates

	ACADEMY PERFORMANCE						TOTAL	
	WITHDREW/NO SHOW		FAILED		PASSED			
	#	%	#	%	#	%	#	IMPACT RATIO*
MINORITY STATUS								
AM IND/ALASKAN	1	4.55%	7	31.8%	14	63.6%	22	**
ASIAN/PAC ISL	3	6.38%	18	38.3%	26	55.3%	47	**
BLACK	18	7.83%	95	41.3%	117	50.9%	230	.832
HISPANIC	12	8.57%	45	32.1%	83	59.3%	140	.969
WHITE	429	10.1%	1223	28.7%	2603	61.2%	4255	
UNKNOWN	19	11.2%	53	31.2%	98	57.6%	170	
TOTAL	482	9.91%	1441	29.6%	2941	60.5%	4864	

* No adverse impact demonstrated.
** The sample size is too small to calculate an adequate impact ratio.

Both the enroute and terminal Academy screening programs had some adverse impact against blacks in 1984-85. Neither program had adverse impact against females. Analyses conducted before implementation of the "common screen" projected that adverse impact against blacks would be reduced but not eliminated by using the weighting system recommended for the new program. Subsequent analysis (see table 4) conducted since the implementation of the "common screen" shows that the program has no adverse impact against blacks or any other minority group. The program demonstrates no adverse impact against females or Hispanics. The absence of adverse impact may be attributed partly to the change in the weights, partly to an increase in the percentage of minorities who participated in special emphasis programs (designed to increase the number of minorities and women in the ATCS occupation) prior to entering the "common screen" program, and partly to the success of such programs.

Effectiveness of Placement Process. A factor with considerable impact on the long-term success or failure of the "common screen" is the extent to which Academy performance is considered when assigning graduates to facilities. It may be recalled that one of the projected advantages of the "common screen" was to reduce field training attrition by assigning only those graduates with high Academy scores to the higher-complexity facilities and assigning graduates with lower scores to the lower complexity facilities.

However, in concurring with the program, management reserved the right to assign students to facilities based on factors other than the Academy score. The priorities considered when making the placement decision were to be i) FAA hiring requirements (e.g., if terminal facilities had openings, then graduates would be assigned to those facilities, regardless of score); ii) Academy performance; and, iii) student preference.

This priority ranking for placement purposes was necessary to comply with congressional hiring mandates specifing that the controller work force be increased by approximately 1000 in FY-86 and 87. It was anticipated that most new graduates would be assigned to the enroute option; thus, the recommendation to assign only those with high scores to the most complex facilities (usually enroute and high level terminals) would not be feasible in all cases.

During the two-plus years following implementation of the new program, Academy performance has been only partially utilized in making placement decisions. The correlation observed between the Academy score and the complexity of facility to which the student is assigned was .28 overall ($N=2787$, $p < .001$). However, over time, the extent to which the score has been utilized appears to have improved with increased acceptance and enforcement of the practice. Utilization appears to be related to the extent to which those who make facility assignments are educated about the reason for using the scores in making placement decisions. One of the functions of CAMI psychologists has been to educate those making placement decisions on a continual basis about the purpose of the screen and placement programs.

Field training attrition. Another question that must be addressed is whether graduates of the "common screen" have lower field training attrition rates. It was projected that if graduates with high Academy scores were placed in

the more complex facilities, a lower field training attrition rate in those facilities would result. However, from the discussion above, it can be seen that the score is not always the basis for the placement decision. Consequently, the "common screen" may have no positive effect on field attrition rates. In fact, it was anticipated that because hiring requirements for the enroute option were so high in 1986, most graduates, even those who would have previously failed the enroute program, would be assigned to enroute. The result might be a higher field training attrition rate for the graduating class of 1986 than for previous years.

Because the average time to attain FPL status for enroute ATCSs is 2.8 years (although some developmentals require more than 4 years), it will not be possible to examine the relationship between the score earned in the "common screen" and field training status until comparable before-and-after screen groups have completed all their field training. However, a preliminary assessment of the effect of the "common screen" on field training attrition can be made by comparing attrition rates at equivalent times in training for groups of Academy graduates who entered the different versions of the screening program.

The results of this preliminary assessment show that about 88% of those who graduated from the screen in 1986 remained in enroute field training while about 87% of the 1985 graduates remained in enroute field training as of November 1986. During 1987, changes occurred in the way enroute field training is conducted that might have affected the loss rate for "common screen" graduates. Consequently, it cannot be determined whether the field training attrition rate was influenced by the "common screen;" on the other hand, the numbers do not suggest that the screen has increased field training attrition.

Utility of ATCS Academy screening. The 1976-based Academy screening programs were intended to identify, through a performance assessment in a simulated job environment, the majority of candidates who passed the OPM selection battery but did not have the potential to become full performance level ATCSs. The option-specific screening programs occured early enough in the training cycle to provide a more cost- effective identification of candidates having the highest potential for success before both the agency and the candidate invested several years in training.

A comparison of attrition rates before and after implementation of the Academy screen demonstrates the effectiveness of having an early screen. Of the individuals who entered the profession prior to the 1976 Academy program, 38% left the Agency. This loss often occurred two to three years into the person's career. After the Academy screening programs were implemented, the field loss rate was cut to 8%, because 30% of the individuals who would eventually become attritions were screened out within a few weeks. Observations since the 1981 ATCS strike reveal a 40% loss rate at the Academy; however the field loss rate has remained at about 10%. These data emphasize the continued utility of employing the initial screening program in spite of changes in population characteristics that occurred after the strike.

RADAR SCREENING AT THE FAA ACADEMY.

Research was conducted during the late 1970s and early 1980s to develop another screening program based upon the application of principles of air traffic control in a radar environment. The radar simulation was a sophisticated design which closely simulated the actual operating environment in an enroute or terminal radar facility with computer ability to synthesize and present a wide variety of air traffic situations.

Boone (1980) assisted in the initial planning for utilizing the Radar Training Facility (RTF) as a new screening program. He also conducted an extensive initial validation effort (1983) involving design, implementation, and formative evaluations to determine if performance measures designed for the option-specific radar screening programs were reliable, if the performance measures had a proper concurrent relationship with measures in the corresponding nonradar screening program, and if the difficulty level of the new program was appropriate.

Initial plans were to implement radar screening programs for both options in a pass/fail mode during the early 1980s. The programs were originally intended to follow each nonradar option-specific screening program in the training sequence. However, after the 1981 ATCS strike, the requirement for hiring new trainees increased until the RTF facility could not accommodate the required number of incoming students. Consequently, the RTF programs were rescheduled to occur in their present positions, before the beginning of facility-specific radar training (about one to two years after Academy graduation).

The radar programs never actually functioned as planned because ATCS developmentals who take an RTF course in its current location in the training program have a more extensive knowledge of the principles of radar air traffic control than do the inexperienced students for whom the courses were originally designed. Current hiring requirements continue to prevent moving this course to its originally intended position in the screening program. Consequently, present plans are to conduct each course in a training instead of a screening mode.

SUMMARY

This report reviewed the development of an approach used by the FAA to select ATCSs through the contributions of aviation psychologists. The foundations of this psychological research effort date back to the 1950s, when studies demonstrated that aptitude test scores could be successfully utilized to predict success in the ATCS occupation. Since that time, the selection process has evolved through various additions and refinements to include three screening components: the OPM selection test battery, the FAA Academy nonradar screen, and the FAA Academy radar screen.

Because the air traffic occupation is directly involved with the safety of the traveling public, it is critically important that the most qualified individuals are selected for training and ultimately for performing the job. The FAA's research efforts identified performance and aptitude measures which

could be used to predict training and occupational success. Efforts in program evaluation have relied upon the analysis of longitudinal data bases to improve the predictive validity and utility of current selection procedures. Thus, the foundations of the ATCS screening programs are research-based, ensuring that the programs remain job-related, valid predictors for selecting ATCSs.

Lessons learned. Many lessons have been learned over the years about the selection of ATCSs as a result of psychological research. Although the continual refinements in ATCS selection procedures reflect these lessons, several specific examples are worth noting. First, claims of previous experience in the occupation are not as effective in predicting later success in the FAA as are paper-and-pencil measures of job knowledge. The use of the OKT has improved considerably the selection of those with prior experience.

Second, the performance-based Academy screening program appears to be a better predictor of later success in training than any combination of the extensive number of written tests (administered either operationally or administratively.) While the Academy screening program is costly to administer due to the requirement to train candidates about procedures before testing their aptitude for application, the cost of administering the program is considerably less than the cost of increasing field attrition which would result from using a selection procedure of lower validity than the current program.

Third, the effectiveness of a newly implemented test or a program modification may be negatively affected by administrative policies or constraints not anticipated during experimental administration of tests or during the modeling of program outcomes. Two examples of this are the increase in MCAT scores for applicants who repeated the test, and the potentially negative impact on field training attrition rates of regional policies not requiring the use of Academy scores in making facility assignments.

Fourth, real world considerations may complicate the application of psychological principles to the development or modification of selection procedures. Air traffic control is a particularly high visibility occupation. Recommendations and pressures to make changes in the validated ATCS selection program can arise from Congress, the media, or the public. While such suggestions are sincere efforts to solve aviation problems, they may be based on a naive interpretation of the conditions affecting the air traffic control system. The challenge to aviation psychologists involves balancing the requirements imposed by forces outside the FAA with the internal, data-based actions necessary and sufficient to ensure that the selection/screening programs retain their validity.

From about 1990-2020, considerable changes will be made to the equipment used to depict the location of aircraft in the National Airspace System. Near-term changes should be fairly transparent to the ATCS users of the equipment (although these changes will have a major impact on the jobs of those who perform equipment maintenance.) However, longer-term changes may significantly impact the tasks performed by ATCSs, and the skills and aptitudes required to perform the tasks. We know that the new equipment will

perform problem detection and resolution functions, but we do not yet know the extent to which the solutions provided by the computer will be considered supporting data for the ATCS or will be mandatory solutions implemented automatically without human intervention. Depending on the extent to which the ATCS of the future will have a role in the decision making process regarding the control of air traffic, future ATCS job functions may be directed less toward making decisions to separate aircraft and more toward monitoring the operations of the equipment that makes the separation decisions and and coordinates them with other aircraft computers.

If the occupational requirements change, associated selection and training programs must also change. To anticipate changes in selection and training program requirements, an integrated psychological research program is currently being conducted in the following areas: 1) developing, refining, and analyzing job/task analyses describing future ATCS job requirements, 2) examining factors affecting performance on systems functionally equivalent to future ATCS automated equipment; 3) developing color vision tests for employees who operate new color-based ATCS equipment, 4) developing selection tests to identify candidates with aptitude to perform the job of ATCS given future job/task requirements.

Approaches to resolve anticipated changes for future ATCS selection and new training needs require aviation psychologists to work closely with FAA management to define requirements addressing those changes, formulate implementation plans, develop new programs, provide regular status reports regarding program implementation efforts, and conduct periodic program evaluations. As a part of this process, program evaluation measures must be developed and integrated into the design and implementation stages of program development to allow for the systematic collection, analysis and dissemination of data relevant to program outcomes for top FAA management's applications to decisions regarding program implementation and change.

The ongoing research and data collection efforts made by FAA aviation psychologists provide a data-based support system for assessment of ATCS selection and training needs and ensure that efforts will continue to refine and enhance the selection procedures. Since the mandate of the Committee on Government Operations in 1975, the FAA has succeeded in developing a standardized, validated, centralized selection and training program which provides early identification of candidates who demonstrate the aptitude to perform competently in the profession. The commitment of FAA management to employing data-based decision-making strategies has clearly contributed to the stability of the ATCS selection program since 1976. In addition, the contributions of research conducted by aviation psychologists have resulted in real dollar savings to the public, as well as in the selection of individuals who have the aptitude to perform as effective ATCSs.

REFERENCES

Bond, N.A., Bryan, G. L., Signey, J. W., and Warren, N. D. Aviation Psychology. Los Angeles: University of Southern California, 1968.

Boone, J. O. Radar Training Facility initial validation. FAA Office of Aviation Medicine Report No. FAA-AM-83-9, 1983.

Boone, J. O. The FAA air traffic controller training program, with emphasis on assessment of student performance. In Sells, S. B., Dailey, J. T., and Pickrel, E. W. (Eds.), Selection of air traffic controllers. FAA Office of Aviation Medicine Report No. FAA-AM-84-2, 1984.

Boone, J. O., Van Buskirk, L., and Steen, J. The Federal Aviation Administration's Radar Training Facility and employee selection and training. FAA Office of Aviation Medicine Report No. FAA-AM-80-15, 1980.

Brokaw, L. D. Early research on controller selection, 1941-1963. In Sells, S. B., Dailey, J. T., and Pickrel, E. W. (Eds.), Selection of air traffic controllers. FAA Office of Aviation Medicine Report No. FAA-AM-84-2, 1984.

Collins, W. E., Boone, J. O., and VanDeventer, A. D. (Eds.), The selection of air traffic control specialists: History and review of contributions by the Civil Aeromedical Institute, 1960-1980. Aviation Space and Environmental Medicine, 1981, 52, 217-240.

Collins, W. E., Manning, C. A., and Taylor, D. K. A comparison of prestrike and poststrike ATCs trainees: Biographic factors associated with Academy training success. In VanDeventer, A. D., Collins, W. E., Manning, C. A., Taylor, D. K., and Baxter, N. E. Studies of poststrike air traffic control specialist trainees: I. Age, biographical factors, and selection test performance related to Academy training success. FAA Office of Aviation Medicine Report No. FAA-AM-84-6, 1984.

Dailey, J. T. and Pickrel, E. W. Development of the Multiplex Controller Aptitude Test. In Sells, S. B., Dailey, J. T., and Pickrel, E. W. (Eds.), Selection of air traffic controllers. FAA Office of Aviation Medicine Report No. FAA-AM-84-2, 1984.

Lewis, M. Use of the Occupational Knowledge Test to assign extra credit in selection of air traffic controllers. FAA Office of Aviation Report No. FAA-AM-78-7, 1978.

Rock, D. B., Dailey, J. T., Ozur, H., Boone, J. O., and Pickrel, E. W. Selection of applicants for the air traffic controller occupation. FAA Office of Aviation Medicine Report No. FAA-AM-82-11, 1981.

Uniform guidelines on employee selection procedures. Federal Register, 1978, 43, 38290-38315.

VanDeventer, A. D. Field training performance of FAA Academy air traffic control graduates. Presented at the annual Scientific Meeting of the

Aerospace Medical Association, 1981.

VanDeventer, A. D., and Baxter, N. E. Age and performance in air traffic control specialist training. In VanDeventer, A. D., Collins, W. E., Manning, C. A., Taylor, D. K., and Baxter, N. E. Studies of poststrike air traffic control specialist trainees: I. Age, biographical factors, and selection test performance related to Academy training success. FAA Office of Aviation Medicine Report No. FAA-AM-84-6, 1984.

VanDeventer, A. D. A followup evaluation of the new aptitude testing procedures for selection of FAA air traffic control specialists. In VanDeventer, A. D., Collins, W. E., Manning, C. A., Taylor, D. K., and Baxter, N. E. Studies of poststrike air traffic control specialist trainees: I. Age, biographical factors, and selection test performance related to Academy training success. FAA Office of Aviation Medicine Report No. FAA-AM-84-6, 1984a.

VanDeventer, A. D. An evaluation of the initial three operational administrations (1981-1982) of the new Air Traffic Control Specialist Selection Procedure. Civil Aeromedical Institute Memorandum Report No. AAC-118-84-3, 1984b.

☆U.S. Government Printing Office : 1988 - 516-018/80421