UNCLASSIFIED

DTIC

Technical
Report

distributed by

Defense Technical Information Center
DEFENSE LOGISTICS AGENCY

Cameron Station • Alexandria, Virginia 22314

UNCLASSIFIED

NOTICE

We are pleased to supply this document in response to your request.

The acquisition of technical reports, notes, memorandums, etc., is an active, ongoing program at the Defense Technical Information Center (DTIC) that depends, in part, on the efforts and interests of users and contributors.

Therefore, if you know of the existence of any significant reports, etc., that are not in the DTIC collection, we would appreciate receiving copies or information related to their sources and availability.

The appropriate regulations are Department of Defense Directive 5100.36, Defense Scientific and Technical Information Program; Department of Defense Directive 5200.20, Distribution Statements on Technical Documents; Military Standard (MIL-STD) 847-A, Format Requirements for Scientific and Technical Reports Prepared by or for the Department of Defense; Department of Defense Regulation 5200.1-R, Information Security Program Regulation.

Our Acquisition Section, DTIC-DDA-1, will assist in resolving any questions you may have. Telephone numbers of that office are: (202) 274-6847, 274-6874 or Autovon 284-6847, 284-6874

June 1982

MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

SELECTION OF APPLICANTS FOR THE

AIR TRAFFIC CONTROLLER OCCUPATION

Donald B. Rock
John T. Dailey
Herbert Ozur*
James O. Boone
Evan W. Pickrel

Federal Aviation Administration
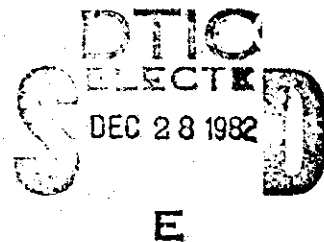Washington, D.C.

*Office of Personnel Management

July 1981

DTIC
ELECTE
DEC 28 1982

E

NOTICE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FAA-AM-82-11	AD.4122795	

4. Title and Subtitle	5. Report Date
	July 1981
SELECTION OF APPLICANTS FOR THE AIR TRAFFIC CONTROLLER OCCUPATION	6. Performing Organization Code
	8. Performing Organization Report No.

7. Author's

9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
FAA Office of Aviation Medicine 800 Independence Ave., SW Washington, D.C. 20591	
	11. Contract or Grant No.
	13. Type of Report and Period Covered

12. Sponsoring Agency Name and Address	
Federal Aviation Administration Office of Aviation Medicine 800 Independence Ave., SW Washington, D.C. 20591	
	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

This report covers personnel research efforts during the past nine years directed toward improving the selection of applicants to work in the Air Traffic Control occupation. The report summarizes the various research efforts and makes specific recommendations for changes in present OPM tests through which applicants qualify for employment consideration in the ATC occupation, and the rating and ranking procedures used to establish a register of eligible applicants for appointment to the occupation.

New tests developed for inclusion in a revised ATC test battery include the Multiplex Controller Aptitude Test (MCAT) and an Occupational Knowledge Test (OKT). Both tests are paper-pencil, machine scorable, multiple choice type tests. The MCAT presents simulations of air traffic flow in each item. The applicant must solve air traffic problems or a related cognitive skill problem in completing each item. The OKT measures important aspects of knowledge associated with prior experience that are related to successful performance as an air traffic control specialist.

The revised test battery recommended to OPM includes the Multiplex Controller Aptitude Test and the present OPM test, Abstract Reasoning and Letter Sequence. The Occupational Knowledge Test is used for granting additional earned credit in place of the present OPM Rating Guide. This new test battery meets all professional requirements for an effective test including fairness.

17. Key Words	18. Distribution Statement
Selection-Air Traffic Controllers Occupational Knowledge Aptitude Testing Personnel Research	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		

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

SELECTION OF APPLICANTS FOR THE
AIR TRAFFIC CONTROL OCCUPATION

CONTENTS

SELECTION OF APPLICANTS FOR THE AIR TRAFFIC CONTROL OCCUPATION

INTRODUCTION

This report has been prepared to consolidate various personnel research efforts during the past 9 years directed toward improving the selection of applicants to work in the air traffic control (ATC) occupation. The position classification standards for the Air Traffic Control Specialist (ATCS) Series (GS-2152) established by the Office of Personnel Management (OPM) define three major options within the occupation.

- **Flight Service Stations (FSS)** which provide preflight briefing and assistance and advisory services to pilots during flight;

- **Terminals** which provide control and separation of air traffic at airports, and,

- **Centers** which provide control and separation of enroute air traffic.

The Federal Aviation Administration (FAA) is the major employing agency for ATCSs with a workforce of about 26,000 employees in the GS-2152 series. They represent almost 50 percent of the total FAA full-time staffing. Selection and hiring of applicants for the occupation by FAA over the past 5 years has ranged between 1,700 and 2,000 new employees annually and can be expected to continue at about the same levels. The Department of Defense (DOD) employs a total of about 300 civilian air traffic personnel. DOD may select applicants from the same OPM register used by the FAA, although their use of the register is minimal. Positions as ATCSs are essentially nonexistent in the private sector.

This report summarizes the various research efforts and, based on the results, makes specific recommendations for changes in:

- Present OPM tests through which applicants qualify for employment consideration in the ATC occupations, and,

- The rating and ranking procedures used to establish a register of eligible applicants for appointment to the occupation.

The research documents and studies referenced in the text of this report are identified in the bibliography.

1

HISTORICAL DEVELOPMENT OF PRESENT SELECTION PROCEDURES

Selection of ATC personnel has been an area of concern for a number of agencies over a considerable time period. Continuing research since the early 1950's has been directed toward improving the processes by which applicants or candidates for civil and military ATCSs are evaluated, tested, and selected. Reference (1) provides a survey of selected research studies on ATC selection during the period 1952-1972.

In August 1960, the FAA Civil Aeromedical Institute (CAMI) began administering a heterogeneous battery of commercially available tests on an experimental basis to newly selected ATCSs at the FAA Academy in Oklahoma City. After the 9 week Academy ATC training course, the student's average academic test scores and average laboratory scores were summed to form a composite which was correlated (Pearson product - moment formula) with the composite of the experimental aptitude test scores. The coefficients ranged from .35 to .54. Based on these findings, it was concluded that aptitude tests could improve selection of ATC applicants who would be successful in the occupation.

The commercial tests used experimentally were considered to be more susceptible to compromise than those controlled by the Civil Service Commission (CSC). Consequently, those commercial tests that showed the most promise were used to identify CSC tests which were similar in factor content. Table I identifies the five tests and the zero-order correlations derived from these studies (2). These CSC tests and an additional Air Traffic Problems Test developed for FA' were then employed, beginning in 1961, in another series of testing sessic . at the FAA Academy. Subsequent regression analysis resulted in identifying five best predictors of training success using course grade and pass or fail status of ATC students.

TABLE 1 CSC TEST VALIDITIES

Testing Dates: September 1960-June 1962	(N = 183)	(N = 143 - pass) (N = 40 - fail)
CSC Tests[1]	Course Grade	Pass/Fail
CSC 51 Spatial Patterns	.37**	.27**
CSC 24 Computations	.28**	.16*
CSC 157 Abstract Reasoning	.28**	.18*
CSC 157 Letter Sequence	.55**	.45**
CSC 135 Oral Directions	.23**	.23**
CSC 540 Air Traffic Problems	.41**	.29**

*p $\leq$ .05; **p $\leq$ .01

[1]Test number assigned by CSC

2

Beginning in 1962, the test battery identified in Table I served as a major determinant in the selection of one-third of the applicants with little or no prior ATC related experience. Applicants with experience continued to be selected without a requirement for formal testing. CAMI continued to collect and analyze data on the test battery after its adoption by the CSC. In January 1964, the CSC battery was introduced as a means to qualify all applicants regardless of their previous experience, although air traffic related experience continued to be used as a ranking factor for placing applicants on the CSC register for selection.

In October 1968, further modifications were made to the ATC selection system aimed at relieving a shortage of ATC personnel in a rapidly expanding ATC system. A major change provided for hiring applicants with specialized ATC experience (particularly in military radar control) at a higher grade (GS-9/GS-11) without requiring that they take the CSC test battery.

These ATCS selection procedures remained essentially unchanged from 1968 until 1973 when a maximum entry age (31 years) for the Center and Terminal options was established by Public Law 92-297. At this time, CSC and FAA re-established the requirement that all ATC applicants take the CSC test battery and achieve a score of 210 or higher for employment eligibility. From 1973 to the present, the ATC selection procedures have remained essentially the same.

The methods and standards for establishing ATC employment eligibility have varied from time-to-time since 1962. However, the total selection process has typically involved two elements: the use of the CSC test battery to determine eligibility for employment; and the weighting and use of prior aviation-related experience either directly or indirectly in selecting from among those applicants eligible for employment. References (2), (3), (4), and (5) provide sources for more detailed information on the development of ATC selection procedures and (Chapter I of reference 6) presents a more extensive discussion of current selection processes.

3

PROBLEM AND RESEARCH OBJECTIVES

Attrition of ATCSs during the developmental training period has been an area of continuing concern for the FAA. Analysis of training losses over the past 9 years highlights the fact that attrition ranged between 25 and 40 percent during the 2 to 5 year training period required to reach the full performance level (FPL). The major part of this attrition resulted from the inability of trainees to acquire and demonstrate the skills and knowledge to progress satisfactorily through the training program. This has been the case indepen-dent of whether or not the training has been conducted on a completely decen-tralized basis or when selected phases of training have been conducted at the FAA Academy on a centralized basis.

Attrition rates of this magnitude involved investment losses of millions of dollars annually to the FAA in salaries, training, and travel costs. Of equal importance to maintaining an adequate staff of full performance controllers is the irretrievable loss of "time," since a new applicant hired as a replacement must start at the beginning of the training cycle.

This significant investment loss to the FAA has two major components. First, the screening and selection criteria used to determine which applicants to hire for the occupation; second, the length of time taken to identify trainees who do not demonstrate the ability to progress to the full performance level required by the occupation.

To establish more comprehensive and uniform training, and to provide earler identification of trainees who do not demonstrate the necessary skills and abilities, the FAA re-established a centralized ATC Initial Qualification Training Program at the FAA Academy for the terminal and center options and restructured the follow-on training conducted at the ATC facilities in January 1976. Selected phases of training leading to FPL are conducted on a pass/fail basis using defined job-related training outcomes and standards. A comparable pass/fail training program was established for the FSS option in September 1978.

Terminal, center, and FSS initial qualification training at the FAA Academy consists of approximately 15 weeks of instruction in both academic and labora-tory environments where students must demonstrate the ability to learn and apply the knowledge and skills taught at the Academy. Failure to pass the training is the basis for removal from the ATC occupation or the FAA. While the attrition rate fluctuates from class to class, the average loss rate during the initial qualification training over the past 3 years in terminal and center training has been 24 percent. Four percent of the students volun-tarily withdrew from training and 20 percent failed the course. This rate varies for each year. For 1977, the failure rate was 21 percent and with-drawal 5 percent; for 1978, failure 30 percent and withdrawal 8 percent; and in 1979, failures were 28 percent and withdrawals 7 percent.

4

The FAA Academy training has been effective in providing early identification of those trainees who do not demonstrate the necessary abilities to progress to the journeyman ATCS level. Between 1970 and 1976, when all ATC training was conducted on a decentralized basis at the individual ATC facilities, the average length of service for ATC trainees who separated from the occupation was between 18 and 24 months. With initial qualification training conducted at the FAA Academy, most attrition which can be expected to occur due to performance takes place during the first 5 months of employment. By itself, this change in FAA ATC training provides a cost avoidance on the order of $10.0 to $12.0 million a year.

There still remains, however, the concern with the overall attrition rate of ATC trainees. For example, salary and training costs for the initial ATC qualification training program represent about $10,000 for each employee hired. With a hiring rate of 1,800 new employees annually, a 24 percent attrition rate equates to an investment loss of about $43 million each year.

Consequently, the primary purpose of this personnel research is to develop improved selection procedures in order to hire those applicants who clearly demonstrate aptitute for ATC and whose potential for success is greatest.

References (8), (9), and (10) provide sources for more detailed information on ATC traininee attrition, costs, and training programs.

ATC JOB ANALYSIS

The Dictionary of Occupational Titles provides a general description of the functions and activities of ATCSs in each of the options. In addition, more detailed job/task analyses have been completed for a variety of purposes.

. **Position Classification.** In 1978, the CSC completed a revised Classification Standard for the GS-2152 ATC occupation which identifies knowledge, skills, and aptitudes required in the FSS, terminal and center options.

. **Performance Evaluation.** Between 1971 and 1975, the System Development Corporation under contract with the FAA completed a detailed job/task analysis of the terminal and center options, to identify the basic skills needed for successful performance of each task required of a controller, as a basis for developing objective performance standards and measures. The results are currently used by the FAA in conducting "over the shoulder" evaluations and extended training reviews for ATCSs under the ATC Technical Appraisal Program (TAP).

. **ATC Work Attributes.** As part of a broader research study on ATC selection, Education and Public Affairs (EPA), Inc., under contract with the FAA, completed an analysis of the work attributes associated with the ATC occupation and proposed methods for their measurement. Their report (7) identifies nine major demands of ATC work (e.g., cognitive requirements); over 30 attributes (e.g., judgment and reasoning ability) associated with job demands; and, describes one or more ATC tasks related to these attributes.

In addition to these more recent ATC job analyses, others going back to the 1950's have been conducted (1).

6

OVERVIEW OF RESEARCH STUDIES

Personnel research related to different facets of the ATC occupation has been
a continuing FAA program since the early 1960's. This report, however, will
focus on the results of studies directed to initial selection of applicants,
particularly the development of screening tests and the use of prior aviation
experience in the selection and placement process. Over the past 9 years,
studies have encompassed samples of four major population groups.

. FPL ATCSs;

. Developmental ATCSs with several years of ATC training who have not
  reached FPL status;

. New ATC appointees who were just entering the occupation; and,

. Applicants for ATC work.

While a number of interrelated studies have been completed, perhaps the best
way to provide an overview is to briefly identify them in a chronological
sequence although they were obtain conducted as parallel efforts.

. 1972 - FAA ATCSs. A major study (12) of some 14 experimental tests as
  candidates for improving selection of ATCS applicants was completed by
  EPA. The study was based on concurrent analysis of about 800 ATCSs
  covering all three options and included both journeyman and new
  appointees in the ATC occupation.

. 1977 - FAA ATCS. Based in part on longitudinal analysis of the infor-
  mation obtained in the 1972 study, EPA completed a comprehensive
  analysis of 11 experimental predictors involving over 3,000 employed
  and separated ATCSs. This study (13) also encompassed all three ATC
  options.

. 1976-1977 - Applicants for ATC Work. Since no information was avail-
  able on the characteristics of the population which applies for ATC
  work, CSC, at FAA's request, administered three of the experimental
  tests or predictor forms along with the present CSC test battery to
  about 7,500 applicants for ATC work when the exam was opened in the FAA
  Eastern and Southern Regions. The information obtained provided a
  variety of necessary data on the characteristics of individuals who
  express an interest in the ATC occupation by taking the qualifying CSC
  test. It also provided a basis for correcting for restriction in range
  in the statistical analysis of CSC tests and experimental predictors
  administered to those individuals hired for ATC work.

. 1976-1978 - New Appointees to the ATC Occupation. With the establish-
  ment of the ATC initial qualification training at the FAA Academy in
  1976, the Office of Aviation Medicine and CAMI initiated a continuing
  program of experimental test development, administration, and analyses
  of data on new hires in the center and terminal options. Experimental
  tests, including those used in the 1977 study by EPA and administered

to the applicant population by CSC, were given to new ATC hires at
the FAA Academy prior to the beginning of training. In addition, the
CSC test battery scores were obtained for each individual whenever
possible. About 3,000 new ATC hires were included in this study with
complete data on all variables for over 1,800 appointees.

Analysis of the data for this sample together with the informa-
tion obtained on the 7,000 applicants by CSC in 1976-77 provided the
basis for identifying those existing CSC or experimental tests which,
in combination, resulted in the best prediction of success in ATC
training. It should also be pointed out that this sample group (and
subsequent ATC training classes) provide the basis for longitudinal
analysis of tests and training scores with on-the-job training and
success as employees progress toward FPL status.

1978-1979 - New Appointees to the ATC Occupation. Based on analyses
completed through 1977, the FAA Office of Aviation Medicine developed
parallel forms for those experimental tests which provided the best
predictive capability. Beginning in June 1978, two parallel versions
of the tests were administered to each employee entering ATC training
between June 1978 and January 1979. About 1,100 trainees comprise
the sample for this study. The data provided the basis for item
analysis of the parallel test forms and together with other variables,
their correlation with training success or failure.

1978 - Applicants for ATC work. During September - October 1978, the
CSC reopened the ATC examination for applicants. At FAA's request,
CSC administered the parallel forms of the experimental tests used at
the FAA Academy along with the standard CSC ATC test battery. CSC
also obtained information on sex and ethnic groups, for these
applicants, which previously had been unobtainable. This applicant
population was comprised of about 7,000, including about 1,000 who
took the examination on a "walk-in" basis as part of the affirmative
action recruitment program. This information provided a basis not
only for item analysis of parallel test forms relative to the applicant
population, but an opportunity to evaluate the effect of changes
in the qualifying test battery on selection alternatives and criteria.

The general relationship between these studies is shown in Figure 1
They are summarized in the following sections with respect to:

- Objectives

- Sample Descriptions

- Predictors

- Criterion Measures

- Analytical Methodology

- Results

FIGURE 1.   GENERAL RELATIONSHIPS BETWEEN ATC SELECTION RESEARCH STUDIES
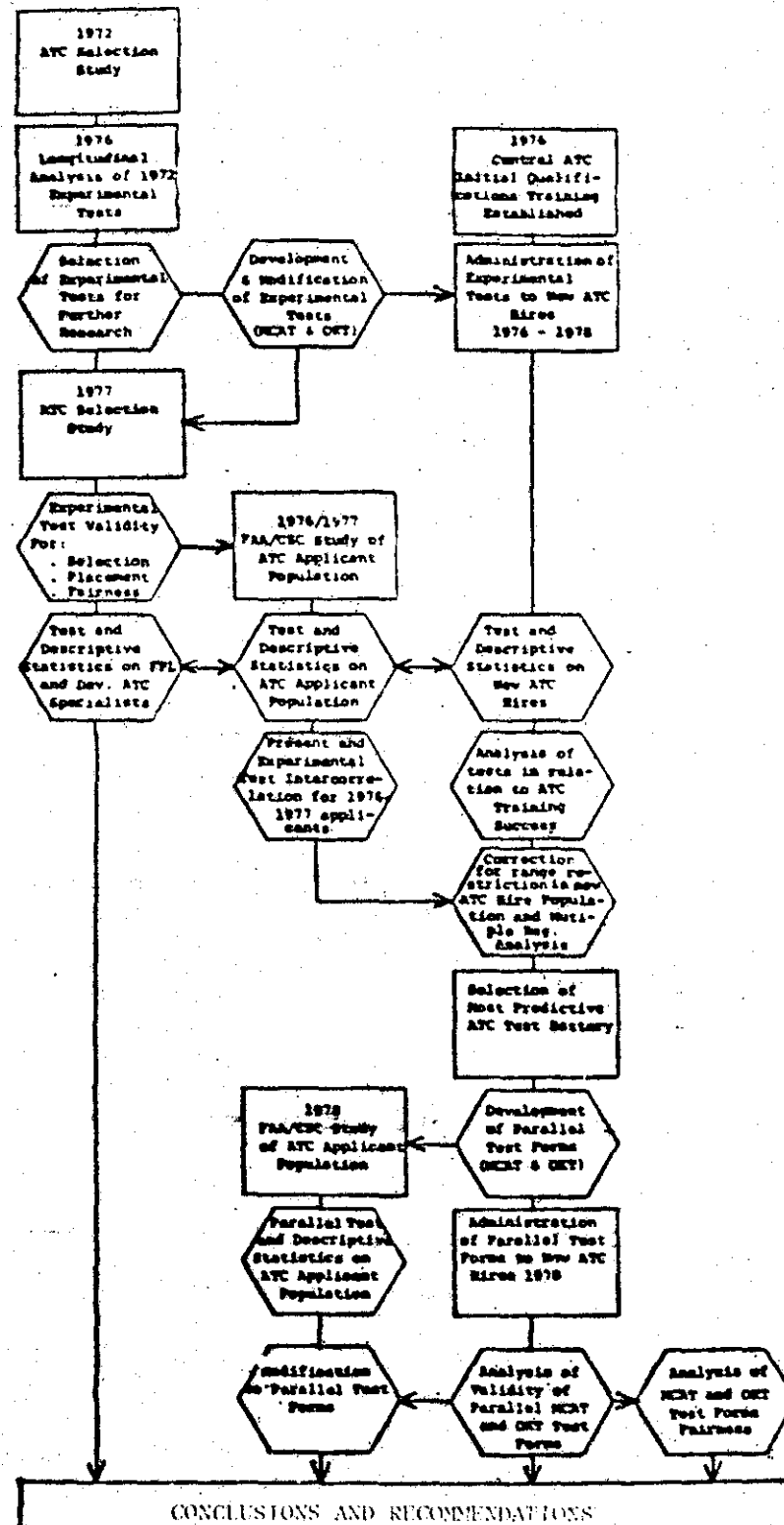
1972 STUDY - SELECTION OF AIR TRAFFIC CONTROL SPECIALISTS

Objectives. The focus of this research was on selection as related to job performance. It was based on concurrent validity analysis. The primary questions addressed were:

- To what extent is it possible to predict the quality of job performance of a journeyman ATCS from a battery of tests adminis-tered at the time of job application?

- To what extent can improvement be achieved by assigning applicants consistently to one of four selected ATC options (FSS, Center, IFR or VFR terminals) and within these, into high or low density (activity) facilities?

- To what extent do the measures selected to accomplish the above objectives affect black and white applicants with an equal degree of fairness?

Sample Description. The total sample was comprised of approximately 800 employees who were either journeymen ATC specialists (FPL) or new ATC appointees including an oversample of black ATC specialists. Sample selection provided for distribution between four types of ATC facilities (Centers, IFR and VFR Terminals, and Flight Service Stations). Samples for Centers and IFR terminals were further stratified between high and low activity facilities. Racial distribution, including the oversample, was 93 percent white and 7 percent black.

Using 15 major cities in the U. S., the sample was randomly selected (except for the oversample) from ATC facilities within a 100-mile radius of the 15 hub cities. Sampling was controlled to exclude ATC specialists over 36 years of age and to insure that FPL Specialists had no less than 3 years nor more than 10 years of ATC experience with FAA. Because of the small number of women in the ATC workforce, it was not possible to stratify the sample selection based on sex. Participation on the part of ATC specialists was voluntary. All subjects were administered a battery of paper-pencil tests or forms including the present CSC test battery.

In addition, a subsample of about 260 ATC Specialists who took the paper-pencil battery were administered a series of "psychomotor" tests in a separate testing session at FAA facilities at Oklahoma City. Of the journeymen, 34 percent took the psychomotor tests together with 32 percent of the new appointees. Selection of this subsample was controlled for region, type of ATC facilities and variance on the confidential supervisory performance evaluation which was used as the criterion measure. Racial distribution of the "psychomotor" subsample was 83 percent white and 17 percent black.

11

Predictors. In selecting predictors for this study, four objectives were considered:

- They should cover as broadly as possible, the range of job and worker attributes identified with the ATC occupation.

- Experimental tests should not substantially overlap areas already covered by the existing CSC test battery which was to be administered to the ATC Specialists.

- Experimental tests should be selected which, based on prior research, appeared to have potential validity for ATC selections.

- The two selected test batteries (paper-pencil and psychomotor) should not require more than 8 hours each to administer.

Within these objectives, the following were selected:

- Paper-Pencil Battery

  - Aptitude Tests

    Current CSC test battery (5 parts)
    Minimum Coins test
    Dial and Table Reading Test

  - Knowledge and Interest tests

    Dailey Technical and Scholastic Test
    ATC General Information Test

  - "Personality" Tests

    Concept-Adjective Test
    Closure Speed Test

  - Background Information

    Biographical Inventory

- Psychomotor Battery

    Directional Headings Test
    Hidden Patterns Test
    Press Test
    Controller Decisions Evaluation (CODE) Test
    Multiple Task Performance Test
    Compressed Speech Test

12

The Paper-Pencil Battery was administered to the total sample group.
Test administrators were FAA employees, trained by Education and Public
Affairs. The Psychomotor Battery was given to the subsample at FAA
facilities in Oklahoma City where equipment required for administration was
available. Detailed discussion on the source, nature of the test, method
of scoring, reliability and validity for the selected tests is provided
elsewhere (1). The contractor obtained and kept all test results
and information on individual ATC Specialists on a confidential basis.
No individual data were provided to FAA.

Criterion Measure. After examining a number of alternatives, a Confidential
Supervisory Evaluation form was used as the criterion against which experi-
mental tests were validated. These evaluations were obtained by the con-
tractor directly from the ATC Specialists' supervisor. They were not reviewed by
FAA nor were copies provided to the agency. The contractor retained all
individual evaluation data on a confidential basis. The evaluation covered
the broad areas of performance within a number of task behaviors for each
area.

Performance Area	Task Items
Knowledge	3
Perception	6
Comprehension	6
Memory	2
Communication	8
Judgement	4
Traffic Management Techniques	4
Performance Work Stress	4
Interpersonnel Skills	3
Other personnel skills	5

In addition an "Overall Performance" category consisting of four task
items and a summary evaluation were obtained. The summary evaluation
was a seven-point rating scale with "1" being the highest performance and
"7" the lowest.

Descriptive Information. Based on biographic responses from 304 FPL ATC
Specialists, the following background and educational information was
obtained:

- 98% were men; 2% women
- 96% had prior military service
- 72% claimed prior (military) experience as a controller
- 2% claimed experience as a pilot
- 48% took the CSC test battery for appointment; 10% more than
  once,before passing
- 97% completed high school
- 25% attended college; less than 1% completed college

13

Analytical Methodology. Data collected were subjected to a variety of statistical treatments. Among the principal methods were:

- Multiple regression analysis in which test measures were correlated with performance evaluation measures to determine how well they predicted job performance.

- Analysis of variance for testing differences between groups; minority and non-minority; journeymen and new appointees; and, those hired with previous aviation-related experience compared to those hired without such experience.

- Multiple discriminant function analysis, in which tests were evaluated in terms of their ability to maximize placement within option and activity levels.

Results. In summary, this study concluded that:

1. Capacity of the tests to predict job performance of journeymen ATC specialists produced mixed results. The CSC test battery was marginal in predicting job performance, as measured by supervisory evaluations. This result, however, was due primarily to the restriction in range of the CSC test predictors. This restriction resulted not only from the selection processes but also from the screening which took place during the developmental training period as the controllers included in the sample progressed to the journeyman (FPL) level. This study did not attempt to correct for the range restriction and consequently the predictor relationship of the CSC Tests to the job performance criteria was not determined. The psychomotor tests developed consistently significant correlations with supervisory job performance evaluations.

2. By combining paper-pencil and psychomotor tests, the analysis indicated that it was possible to improve assignments to the different ATC options based on tests results and, with additional tests, to high or low density center or IFR terminals.

3. If relevance to the job is the primary factor in determining test acceptability, the tests proposed from the research met that criterion. They also predicted job performance equally well for each racial group (black and white). However, blacks as a group consistently scored lower on the tests than did whites.

14

This study provided FAA with valuable information and insights on the problems associated with selection and placement of applicants for ATC work. However, action on the research results and recommendations was deferred due to a number of considerations including the complexity of the specialized equipment required for the psychomotor tests, logistical difficulties in test administration and complexity in test scoring and ranking of applicants.

Continuing concern with the rate of ATC trainee attrition and consideration of reestablishing centralized ATC training at the FAA Academy resulted in a review of FAA selection and screening policies in December 1974 (8) and completion of a cost analysis study of alternatives (9) in March 1975. Among the actions resulting from this review, FAA contracted with Education and Public Affairs, Inc., (EPA) for a follow-on analysis of ATC selection tests in June 1975. A summary of the final EPA report is provided elsewhere (10).

Objectives. The objectives of this research were directed to essentially the same concerns as the previous (1972) study completed by EPA: selection, placement, and fairness. However, the study design significantly expanded the sample size, structure, and representation; encompassed more criterion measures of ATC job success; and incorporated an evaluation of prior aviation-related experience and educational level as predictors of ATC job success.

Longitudinal Analysis of 1972 Experimental Tests. The first effort in this study was to analyze on a longitudinal basis the relationships between the experimental test results obtained during 1971 and criterion measures related to the ATC specialists who participated as of 1975. Validity coefficients of each test score with each criterion measure were calculated. A factor analysis of the tests was also undertaken to determine overlapping of underlying variables. The tests selected from these analyses were then subjected to multiple regression analysis to determine the minimum number of tests or test scores which would predict the maximum proportion of variance in each of the criterion measures. Table 2 identifies those 1971 experimental tests and predictors (X) which, based on their simple correlation with the various criterion measures established for the analysis, warranted further study by means of field validation.

16

Table 2 Predictor and Criterion Measures of ATC "Success" (1975) for Controllers Participating in 1971 ATC Research

	Separation	Progression + Attrition		Present Option		1971 Supv. Assess.	Sup/Staff Position
	DEV	DEV	FPL	DEV	FPL	FPL***	FPL***
CODE	X	X	X	X	X	X	-
Dial Reading	X	X	-	-	-	-	-
Dir. Heading	X	-	-	-	-	X	-
Air Traffic Prob.*	-	-	-	X	X	X	-
Arith. Reasoning**	-	-	-	-	X	X	-
ATC General Info.	X	-	-	-	-	-	-
Concept Adjective	X	X	-	X	-	X	-
Biographical Info.	X	X	-	X	X	X	X

*Air Traffic Problems CSC Test No. 540

**Arithmetic Reasoning was Part 5 of the Dailey Technical and Scholastic Test (TST)

***Supervisory assessment criterion data available for 1971 FPL ATCS only. No new 1971 appointees progressed to ATC supervisory or staff positions by 1975.

With some modifications these experimental tests used in 1971 became the "core" test battery for the 1977 research. A more detailed discussion of the methodology and results of this longitudinal study is provided elsewhere (14).

Sample Description - 1977 Study. A comprehensive sample design was constructed to define the ATC population to ensure a representative sample for three specific "year of hire" groups. These groups represented three ATC career "stages": (1) New Hires (1976); (2) Developmental ATC Specialists with 2 to 3 years ATC experience with FAA (1973 and 1974); and, (3) ATC Specialists with 2 to 6 years experience at the journeyman (FPL) level (1969 and 1970).

In addition to these primary ATCS samples, three additional ATCs samples were included:

- An oversample of currently employed women and minority ATCS's in the same three "year of hire" groups.

17

. ATC Specialists who participated in the 1972 research.

. A sample of ATC specialists who were hired during the three
  time periods sampled, but who had separated from ATC work
  before reaching FPL status.

Sample selection was constrained to exclude employees who were in ATC
staff or supervisory positions or over age 31 at the time they were
hired except in the FSS option.

Sample selection for the three primary groups was based on stratified
random sampling methods to provide a proportionally representative group
of the total constrained ATC universe for each of the four ATC "options"
(FSS, VFR, IFR, and ARTCC) with respect to both the initial and current
option of assignment.

Table 3 identifies the various ATC specialists samples by year group, the
total number desired and the samples actually obtained. Since participation was
voluntary, the number invited was expanded where possible to provide for
declination and nonresponses.

TABLE 3. ATC SPECIALIST BY YEAR HIRED, SAMPLE GROUPS AND ATC CAREER STATUS

Year Hired	ATC Sample Year Groups	ATC Career Status	Number Invited	Desired Sample	Obtained Sample	Percent of Desired Sample
1969-70	(1) Employed ATCS	FPL	1344	800	754	(94%)
	(2) ATCS Oversample	FPL	151	200	31	(6%)
	(3) Separated ATCS	DEV	---	---	362	--
1973-74	(1) Employed ATCS	DEV	1127	800	740	(93%)
	(2) ATCS Oversample	DEV	258	200	72	(36%)
	(3) Separated ATCS	DEV	---	---	166	---
1976	(1) Employed ATCS	New Hires (DEV)	610	610	590	(97%)
1971	(1) Employed ATCS	FPL	480	480	270	(56%)
	(2) Separated ATCS	FPL/DEV	---	---	74	---

18

The number of women and minorities who volunteered was not adequate for analysis by year group or ATC option. Consequently, the analysis on test fairness combined year and option groups to provide 235 women and 321 minorities. The oversample for women and minorities was used only for analysis of the fairness of the ATC "success" predictors. The final EPA report (Chapter VII of reference 14) provides a complete discussion on the sampling methodology.

Predictors. Two of the tests used in the 1972 study were modified for use in the 1977 research.

 . CODE (Controller Decision Evaluation). This test consisted of three film versions of a computer simulation of moving air traffic patterns appearing on a radar scope. Initially these were converted to slide projector presentation to eliminate the need for movie projector equipment and to simplify both the response recording and scoring. Group administration of the various adaptions of CODE during this test development phase clearly pointed up the practical problems of using film or slide projection equipment in test administration. Consequently, a paper-pencil version was developed which incorporated measures of abilities to identify potential conflicts of aircraft as well as the traditional kinds of aptitudes within an air traffic control context. The resulting test, the Multiplex Controller Aptitude Test (MCAT) was made available to EPA as a substitute for the CODE tests.

 . Arithmetic Reasoning. In the prior research, the arithmetic reasoning test was one part of the Dailey Technical and Scholastic Test (TST). Since this part could not be given as an isolated test, an Arithmetic Reasoning test developed by the Army Air Force, which was similar, was used in the 1977 research.

In addition to the predictors derived from the longitudinal analysis of the experimental tests administered in 1971, three other predictors were used.

 . Pre-employment Experience Questionnaire (PEQ)

    To obtain specific data from participating ATC specialists on various kinds of pre-FAA experience and education, a questionnaire was developed from the CSC Rating Guide elements used as a basis for granting additional credit in the ATC employment selection process.

ATC Occupational Knowledge Test (OKT)

This test was developed to be "job-knowledge specific." Consequently, it was not included in the 1977 research for the purpose of evaluating its use in screening ATC job applicants for employment eligibility. It was used to measure the "quality" of prior experience as a potentially improved basis for granting additional credit for experience in the selection process in place of the existing CSC Rating Guide.

Sixteen Personality Factor Questionnaire (16 PF)

The 16 PF Questionnaire is administered as part of the medical qualification process to all entering ATC Specialists. It is designed to measure important personality characteristics not otherwise measured and the 1977 research offered the opportunity to assess its utility for possible selection or placement purposes.

The experimental test battery used in the 1977 research may be divided roughly into two groups of instruments; (1) Cognitive Predictors, tests, which have predetermined right or wrong answers; and (2) Other Predictors where applicant responses ordinarily have no inherent value of "right" or "wrong" except as measurable against criteria or values external to the "test." The following identifies the predictors used within these two groups.

Cognitive Predictors

Multiplex Controller Aptitude Test (MCAT)
Directional Headings Test
Dial Reading Test
Arithmetic Reasoning Test
ATC General Information Test
Present CSC Test Battery (5 parts)
ATC Occupational Knowledge Test

Other Predictors

Pre-Employment Experience Questionnaire
Concept Adjective
Biographical Inventory
Sixteen Personality Factor Questionnaire (16 PF)

20

. ATC Specialists who participated in the 1972 research.

. A sample of ATC specialists who were hired during the three
  time periods sampled, but who had separated from ATC work
  before reaching FPL status.

Sample selection was constrained to exclude employees who were in ATC
staff or supervisory positions or over age 31 at the time they were
hired except in the FSS option.

Sample selection for the three primary groups was based on stratified
random sampling methods to provide a proportionally representative group
of the total constrained ATC universe for each of the four ATC "options"
(FSS, VFR, IFR, and ARTCC) with respect to both the initial and current
option of assignment.

Table 3 identifies the various ATC specialists samples by year group, the
total number desired and the samples actually obtained. Since participation was
voluntary, the number invited was expanded where possible to provide for
declination and nonresponses.

TABLE 3. ATC SPECIALIST BY YEAR HIRED, SAMPLE GROUPS AND ATC CAREER STATUS

Year Hired	ATC Sample Year Groups	ATC Career Status	Number Invited	Desired Sample	Obtained Sample	Percent of Desired Sample
1969-70	(1) Employed ATCS	FPL	1344	800	754	(94%)
	(2) ATCS Oversample	FPL	151	200	31	(6%)
	(3) Separated ATCS	DEV	---	---	362	--
1973-74	(1) Employed ATCS	DEV	1127	800	740	(93%)
	(2) ATCS Oversample	DEV	258	200	72	(36%)
	(3) Separated ATCS	DEV	---	---	166	---
1976	(1) Employed ATCS	New Hires (DEV)	610	610	590	(97%)
1971	(1) Employed ATCS	FPL	480	480	270	(56%)
	(2) Separated ATCS	FPL/DEV	---	---	74	---

18

The number of women and minorities who volunteered was not adequate for analysis by year group or ATC option. Consequently, the analysis on test fairness combined year and option groups to provide 235 women and 321 minorities. The oversample for women and minorities was used only for analysis of the fairness of the ATC "success" predictors. The final EPA report (Chapter VII of reference 15) provides a complete discussion on the sampling methodology.

Predictors. Two of the tests used in the 1972 study were modified for use in the 1977 research.

- CODE (Controller Decision Evaluation). This test consisted of three film versions of a computer simulation of moving air traffic patterns appearing on a radar scope. Initially these were converted to slide projector presentation to eliminate the need for movie projector equipment and to simplify both the response recording and scoring. Group administration of the various adaptions of CODE during this test development phase clearly pointed up the practical problems of using film or slide projection equipment in test administration. Consequently, a paper-pencil version was developed which incorporated measures of abilities to identify potential conflicts of aircraft as well as the traditional kinds of aptitudes within an air traffic control context. The resulting test, the Multiplex Controller Aptitude Test (MCAT) was made available to EPA as a substitute for the CODE tests.

- Arithmetic Reasoning. In the prior research, the arithmetic reasoning test was one part of the Dailey Technical and Scholastic Test (TST). Since this part could not be given as an isolated test, an Arithmetic Reasoning test developed by the Army Air Force, which was similar, was used in the 1977 research.

In addition to the predictors derived from the longitudinal analysis of the experimental tests administered in 1971, three other predictors were used.

- Pre-employment Experience Questionnaire (PEQ)

    To obtain specific data from participating ATC specialists on various kinds of pre-FAA experience and education, a questionnaire was developed from the CSC Rating Guide elements used as a basis for granting additional credit in the ATC employment selection process.

19

. **ATC Occupational Knowledge Test (OKT)**

This test was developed to be "job-knowledge specific." Conse-
quently, it was not included in the 1977 research for the purpose
of evaluating its use in screening ATC job applicants for employ-
ment eligibility. It was used to measure the "quality" of prior
experience as a potentially improved basis for granting additional
credit for experience in the selection process in place of the
existing CSC Rating Guide.

. **Sixteen Personality Factor Questionnaire (16 PF)**

The 16 PF Questionnaire is administered as part of the medical
qualification process to all entering ATC Specialists. It is de-
signed to measure important personality characteristics not other-
wise measured and the 1977 research offered the opportunity to assess
its utility for possible selection or placement purposes.

The experimental test battery used in the 1977 research may be divided
roughly into two groups of instruments; (1) Cognitive Predictors, tests,
which have predetermined right or wrong answers; and (2) Other Predictors
where applicant responses ordinarily have no inherent value of "right" or
"wrong" except as measurable against criteria or values external to the
"test." The following identifies the predictors used within these two groups.

. **Cognitive Predictors**

Multiplex Controller Aptitude Test (MCAT)
Directional Headings Test
Dial Reading Test
Arithmetic Reasoning Test
ATC General Information Test
Present CSC Test Battery (5 parts)
ATC Occupational Knowledge Test

. **Other Predictors**

Pre-Employment Experience Questionnaire
Concept Adjective
Biographical Inventory
Sixteen Personality Factor Questionnaire (16 PF)

20

ather than re-administer the present CSC test battery, the intent was to
obtain the CSC test scores for participating ATC specialists from existing
records. However, this did not prove to be feasible for a large number of
the employees in the sample. Consequently, analysis of the experimental
tests in relation to the existing CSC test battery as predictors of ATC
success was not possible in this study.

The "cognitive predictors" and "other predictors" were each analyzed
separately against the criterion measures used to define "success" in the
ATC occupation. They were also analyzed together in a supplemental report
(15) to determine the extent to which "other predictors" added to the
validity of the "cognitive predictors."

A discussion and description of each of the predictors used in the study
including relevant prior research, reliability and validity is provided
elsewhere (Chapter VI of reference 14).

For analytic purposes, the objective was to obtain pre-employment and
education information on all ATC specialists in the sample groups. The
experimental test battery was administered to approximately 50 percent
of the 1969-1970 and 1973-1974 sample and all new hires in the 1976 group.
Selection of employees to take the experimental test battery was controlled
to provide proportional distribution by ATC option within each year group.
The experimental tests were also administered to all employees included in
the ATCS oversample groups.

Predictor data were obtained in several ways. The Prior Experience Question-
naire (PEQ) was completed by the ATC specialists at the time they agreed to
participate. The experimental tests were administered to the new hires
at the FAA Academy on their first day of attendance by the EPA staff
assisted by FAA personnel. For those ATC specialists assigned to facilities,
the tests were given by FAA test administrators who were trained by EPA.
A full discussion of the data collection methodology is provided in the FAA
report (Chapter VIII of reference 14).

Criteria and Measures of ATC Success. To determine how valid experimental
tests, prior experience and education are as predictors of ATC success,
operational definitions (criteria) of success and ways to quantitatively
measure them had to be established. The following four criterion measures
were used in the study and were then combined into a single "aggregate"
measure of ATC success.

21

- Training Performance. This was measured by scores received on the ATC Laboratory Problems and the Controller Skills Test during initial ATC training at the FAA Academy. These scores were selected since they require students to demonstrate operational application of academic knowledge.

- On-the-job Performance. This was measured by confidential job/task assessments prepared by the employee's supervisor. It included 54 questions on ATC job tasks and four general questions on quality of job performance. Responses to a seven-point "overall" rating scale were selected as the measure for on-the-job performance.

- Progression. This was measured by the ATC "option" to which the ATCS was initially assigned when hired, compared to the "option" the ATCS was assigned to on January 1, 1976.

Four "options" were defined to represent general complexity levels of ATCS work (FSS, VFR, IFR, ARTCC) with ARTCC defined as the "highest" complexity level. Within this hierarchy, progression values were established as follows:

- "High"—assigned when an ATCS was in an option of a complexity level the same as or higher than the initial option to which assigned (i.e., VFR initial assignment; IFR current assignment on January 1, 1976).

- "Low"—asigned when an ATCS was in an option of a lower complexity level than the option to which initially assigned (i.e., ARTCC initial assignment; FSS current assignment on January 1, 1976).

- Attrition. This was measured by whether or not ATC Specialists hired during the year groups sampled for this study were still employed in ATC work. Those still employed as ATC Specialists were assigned a "high" score; those separated were assigned a "low" score.

- Aggregate Criterion of ATC Success. This was constructed from the four individual criterion (training, on-the-job performance progression and attrition) and their measures and provides a five-point scale value for ATC "success."

The predictors of ATC "success" used in this study — experimental tests, prior experience, and education — were analyzed for validity against each of the four criterion measures of ATC"success" — training

22

scores, supervisory assessments, progression, and attrition. Final
conclusions and recommendations are based on the validity of the
predictors with the aggregate criterion of ATC success. A full
discussion of the development of each of the criterion measures is
provided in the EPA Report (Chapter V of reference 14).

Descriptive Information. Data on the education and experience background
of ATC specialists who participated in the 1977 study are provided in the
EPA Report (Chapter VII of reference 14). The most marked changes between
the three "year-of-hire" groups (and the 1972 sample) were in educational
level, military service and pilot experience. Table 4 compares the various
sample groups on selected variables.

TABLE 4.   EDUCATION AND EXPERIENCE LEVELS

Year Hired as ATCS

	Prior to 1971 (N=304)	1969 1970 (N=659)	1973 1974 (N=661)	1969-70 1973-74 Oversample (N=103)	1976 (N=592)
High School	72%	34%	24%	25%	16%
Some College	25%	53%	53%	59%	56%
College Degree(s)	1%	13%	23%	16%	28%
Military Service	96%	75%	74%	56%	71%
ATC Experience	72%	--	--	--	--
IFR	--	35%	39%	27%	32%
VFR	--	39%	38%	26%	37%
Pilot	2%	30%	34%	11%	33%

The sample of 304 ATCS hired prior to 1971 was derived from the 1972 study (11).
ATC Specialists could have been hired between 1960 and 1970.

Analytical Methodology. Before undertaking the validity analysis, a number
of preliminary analyses were made to:

. Substantiate that the sample of ATC specialists who did volunteer
  was not, except for the proportion of minorities, essentially
  different from those who did not volunteer.

. Calculate the means, standard deviations, reliability, and inter-
  correlations of experimental predictors and ATC success criterion
  measures.

23

. Determine that the order of MCAT test administration did not affect test results.

. Analyze training scores and convert them to standardized(Z) scores.

. Analyze pre-employment experience and establish empirical scale values for validation.

. Analyze predictors for inclusion in the validation study.

A full discussion of these preliminary analyses is provided in the EPA report (Chapter X of reference 14). The statistical results on distribution, means, standard deviations, and intercorrelations are provided in Tables IX.2 through IX.21 of that report (14). Results of preliminary statistical analyses for the non-cognitive predictors appear elsewhere (15).

As a result of these preliminary analyses, two of the cognitive tests were excluded from the validation analysis.

. Civil Service Tests. The inability to obtain the scores on the CSC test battery for a significant number of ATC specialists in the sample and the effect of restriction in range on those scores which were available, seriously interfered with proceeding with the validity analysis of the CSC tests in the EPA study.

. ATC General Information Tests. This test was excluded from validity analysis since it was only administered to the new hires attending the FAA Academy training; it was not given to the 1973-74 or 1969-70 ATC samples because of time constraints on field testing.

Consequently, the following predictors were included in the primary validation analysis:

Multiplex Controller Aptitude Test (MCAT)
Directional Heading Test
Dial Reading Test
Arithmetic Reasoning Test
ATC Occupational Knowledge Test (OKT)
Pre-employment Experience (PEQ)

The following non-cognitive predictors were included in the supplemental validation analysis:

24

Concept Adjective
Biographical Inventory
Sixteen Personality Factor Questionnaire (16PF)

In summary, the validation methodology for the cognitive predictors (and prior experience) consisted of the following steps:

1. Statistical analyses of each of the smallest homogeneous samples (i.e. each ATC option and all ATC options combined within each year group against each of the four individual ATC success criterion separately). This resulted in 118 separate analyses.

2. Predictors selected from Step 1 (based on validity coefficients and significant levels) for each ATC option and all ATC options combined across all year groups were then analyzed to identify the best overall set of predictors for each of the four individual criterion measures. This resulted in 17 separate analyses.

3. Results of each of the four individual criterion measures were then examined for each ATC option and all options combined to determine the best set of predictors across criterion measures. This resulted in 17 separate analyses.

4. The final set of predictors from Step 3 were then validated against the aggregate ATC success criterion leading to derivation of weighted test and experience scores. This involved five separate analyses.

5. The weighted test battery and experience scores were then validated against the four single criterion measures. This involved four separate analyses.

The specific analytical techniques applied and the rationale for each of the steps in the methodology used are fully discussed in Chapter X of reference 14. The results of the analyses and the validities obtained are provided in Chapters XI, XII, XIII, XIV and XV, of reference 14, the Appendices to the final EPA report (16).

The methodology previously discussed was modified for analyses of the non-cognitive predictors to compensate for the differences imposed by the use of "personality" as opposed to aptitude or knowledge tests.

The most significant change in methodology resulted from the need to empirically establish the appropriate directionality of the various non-cognitive predictor scales before undertaking validity analysis. The specific analytical techniques applied in evaluating the non-cognitive predictors are identified in reference 15.

The methodology to evaluate the fairness of the experimental test battery, prior experience and the Occupational Knowledge Test was directed to determining differential validity in accordance with the Uniform Guidelines on Employee Selection Procedures (17) which state:

> "When members of one racial, ethnic or sex group characteristically obtain lower scores on a selection procedure than members of another group, and the differences are not reflected in differences in measures of job performance, use of the selection procedure may unfairly deny opportunities to members of the group that obtains the lower scores."

The methodology for evaluating fairness aspects of the predictors used in this research is discussed in Chapter XVI of reference 14.

Results. The following briefly summarizes the results of this research study. More detailed discussion is provided in references 13, 14, and 15.

Selection

- Two of the experimental tests--Multiplex Controller Aptitude (MCAT) and Directional Headings--predicted the ATC success criteria established for the study at statistically significant levels of confidence for all ATC options combined and for the three primary year groups sampled (1969-70; 1973-74; and 1976). Two other experimental tests--Arithmetic Reasoning and Dial Reading--either did not predict the ATC success criteria or did not add appreciably to the prediction values obtained from MCAT and Directional Headings.

- Preemployment aviation-related experience and the ATC Occupational Knowledge Test, while not intended for use in determining initial appointment eligibility, predicted ATC success at statistically significant levels of confidence and increased the validity coefficients obtained with the experimental tests.

- Education beyond high school level prior to FAA employment did not predict ATC success in either a positive or negative direction. Essentially all of the controllers in the samples included in this study had at least a high school education.

- Other experimental instruments--16 Personality Factor Questionnaire, Concept Adjective and Biographical Information Questionnaire--did not add appreciably to the predictive capability of the experimental test battery, preemployment experience and the ATC Occupational Knowledge Test.

26

validities derived from the analysis of combined predictors against
Aggregate ATC Success Criterion by ATC option for all year groups
combined are provided in Table 5. Weights for the experimental tests
were derived from the multiple regression analysis. Separate weights
were developed for each test for each ATC option and for all ATC options
combined. The derived weights for all ATC options combined were MCAT
Conflict--37; MCAT Aptitude--21; Directional Headings--15; and, Dial
Reading--0.

TABLE 5. VALIDITIES OF COMBINED PREDICTORS AGAINST THE AGGREGATE ATC "SUCCESS" CRITERION (BY ATC OPTION -- ALL YEARS COMBINED)

ATC Option	Weighted Test Battery Scores		Weighted Test Battery Plus PEQ		Weighted Test Battery and PEQ Plus OKT	
	N	r	df	R	df	R
TSS	196	.23**	193	.26**	192	.26**
VFR	479	.26**	474	.44**	423	.45**
IFR	499	.26**	494	.39**	443	.43**
ARTCC	445	.30**	425	.32**	388	.37**
All Options	1309	.26**	1287	.32**	1205	.34**

** = $p \leq .01$

Academy trainees (1976) in the Terminal option were not identified by IFR
or VFR options. Consequently, they were included in both the IFR and
VFR option Ns, but only counted once in the total N's for all ATC
options.

Reference (13) Appendix A shows the validity coefficients derived for
the experimental test battery (unweighted), prior experience, ATC
Occupational Knowledge Test (OKT) and education level for each of the
four ATC success criteria (excluding the aggregate criterion) by ATC option.
Appendix B (13) shows the validity of selected experimental tests with the
aggregate ATC success criterion.

27

Placement

- The weighted experimental test battery did differentiate between the FSS, Terminal (IFR, VFR) and ARTCC options. Average scores for ATC specialists in FSS terminals and ARTCCs were different at statistically significant levels of confidence. The average FSS score was lowest and ARTCC highest. Table 6 provides the comparative mean scores for the various ATC options and year groups.

TABLE 6. MEAN WEIGHTED TEST BATTERY SCORES BY ATC OPTION

ATC Option	1969-70 & 1973-74 ATC HIRES			1976 NEW ATC HIRES (ACADEMY TRAINEES)		
	N	Mean	SD	N	Mean	SD
FSS	196	227.4	40.2	---	---	---
VFR	170	244.6	35.8 (TERM)	310	242.0	35.0
IFR	189	250.0	37.3			
ARTCC	182	264.3	36.6	263	247.9	38.4

Analysis of variance shows that the means of scores differ significantly at the 1 percent level of confidence between FSS, Terminal (VFR and IFR) and ARTCC for the 1969-70 and 1973-74 groups. The 1976 ATC group difference between terminal and ARTCC is significant at the 5 percent level

Test Fairness

- Women as a group scored lower on each of the predictors and on the aggregate criterion of ATC success. In each case, these differences were significant at the 1 percent level of confidence (Table 7).

TABLE 7 - MEANS, STANDARD DEVIATIONS AND t-TEST RESULTS FOR PREDICTORS AND THE AGGREGATE ATC "SUCCESS" CRITERION FOR MEN AND WOMEN

	Predictors									Criterion		
	Weighted Test Battery			PEQ			OKT			Aggregate ATC "Success"		
Sex	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
Men	1397	244.5	380	2736	3.7	2.9	1318	76.6	12.3	1254	3.1	1.4
Women	171	234.5	480	235	.8	1.6	158	64.6	16.3	158	2.6	1.4
	(t=3.14;p≤.01)			(t=15.02;p≤.01)			(t=11.21;p≤.01)			(t=4.56; p≤.01)		

28

Validities of the predictors against the aggregate ATC success
criterion are provided in Table 8. Except for prior experience,
these validities are significant at the 1 percent or 5 percent levels.

TABLE 8 - VALIDITIES OF THE WEIGHTED TEST
BATTERY, PEQ AND OKT AGAINST THE
AGGREGATE ATC "SUCCESS" CRITERION BY SEX
($*=p \leq .05$; $**=p \leq .01$)

Sex	Weighted Test Battery		Prior Experience (PEQ)		ATC Occupational Knowledge (OKT)	
	N	r	N	r	N	r
Men	1386	.23**	2721	.21**	1308	.22**
Women	165	.19**	229	-.04	154	.14*

It should be noted that few of the 229 women in the prior experience
sample had aviation-related experience. Consequently, the validity
results are due to lack of variance.

. Comparable statistical analyses were made between minority and non-
minority groups and between non-minority and blacks as defined by
OMB Circular A-46. The results with respect to the test battery,
prior experience and the ATC Occupational Knowledge Test are
provided in Tables 9, 10 and 11 respectively.

TABLE 9 - TEST SAMPLE

	PREDICTOR Weighted Test Battery			CRITERION Aggregate ATC "Success"		
	N	Mean	SD	N	Mean	SD
Non-Minorities	1323	247.7	371	1308	3.2	1.4
All Minorities	245	220.4	479	243	2.7	1.4
	(t=10.30:p .01)			(t=5.09:p .01)		
Non-Minorities	1323	247.7	371	1308	3.2	1.4
Blacks	145	207.4	407	144	2.4	1.3
	(t=12.30:p .01)			(t=6.31:p .01)		

TABLE 10 - PRIOR EXPERIENCE SAMPLE

	PREDICTOR Prior Aviation Experience (PEQ)			CRITERION Aggregate ATC "Success"		
	N	Mean	SD	N	Mean	SD
Non-Minorities	2115	3.9	2.9	2097	3.2	1.4
All Minorities	321	2.4	3.0	318	2.6	1.4
	(t=8.48:p .01)			(t=7.15:p .01)		
Non-Minorities	2115	3.9	2.9	2097	3.2	1.4
Blacks	194	1.7	2.8	193	2.4	1.3
	(t=10.27:p .01)			(t=8.39:p .01)		

TABLE 11 - ATC OCCUPATIONAL KNOWLEDGE SAMPLE

	PREDICTOR ATC Occupational Knowledge Test			CRITERION Aggregate ATC "Success"		
	N	Mean	SD	N	Mean	SD
Non-Minorities	1247	76.5	12.8	1235	3.2	1.4
All Minorities	229	69.1	14.7	227	2.7	1.4
	(t=7.85:p- .01)			(t=4.36:p .01)		
Non-Minorities	1247	76.5	12.8	1235	3.2	1.4
Blacks	134	66.9	15.2	135	2.5	1.3
	(t=8.12:p .01)			(t=8.12:p .01)		

Minorities as a group, and blacks as an element of the minority
group, scored lower on each of the predictors and on the aggre-
gate ATC success criterion. In each case, these differences were
significant at the 1 percent level of confidence.

Validity of each of the predictors with the aggregate success
criterion for minorities and non-minorities are provided in
Table 12.

Table 12   Correlations for the weighted test battery PEQ
           scale and OKT against the aggregate criterion
           by minority/non-minority status

	Weighted Test Battery		PEQ		OKT	
	N	r	N	r	N	r
Non-minority	1308	.203**	2097	.148**	1235	.201**
Minority	243	.219**	318	.247**	227	.272**

When blacks were analyzed separately from all other minorities,
the validity of prior experience and the Occupational Knowledge
Test were sustained (r = .259 and .256, both at the 1 percent
level of confidence). Validity of the weighted test battery was
.120 which did not reach the 5 percent level of confidence
(p = .075). However, the difference in the validity for non-minori-
ties and blacks was not statistically significant.

Conclusions

The results of this study supported the conclusion that several of the
experimental predictors were sufficiently valid and fair in predicting
ATC success that they should be further developed for operational use
in the selection of applicants for the air traffic control occupations.

However, the lack of sufficient CSC test scores for the sample popu-
lation and the restriction in range on those scores which were availa-
ble made it impractical to compare the validity of the current CSC test
battery with the validity of the experimental tests. In order to address
these questions, the Civil Service Commission, now the Office of
Personnel Management (OPM), and FAA jointly directed efforts to obtain
the necessary information from a group of about 11,500 applicants for
ATC work during the Fall of 1976 and Spring of 1977.

31

STUDY OF ATC JOB APPLICANTS 1976-1977

Objectives. Historically, there has been little or no information availa-
ble on the characteristics of individuals who apply for positions in the
Air Traffic Control occupation. Consequently, this study was developed
to collect and analyze data on ATC applicants with respect to sex, edu-
cation, prior aviation-related experience and the relationship between
these variables and scores on the CSC test battery and experimental tests.
One direct use of this information was to establish a basis for comparing
the validities of the CSC test and the experimental tests by correcting
for the restriction in range on CSC test scores for approximately 2,000
newly hired ATC trainees who attended the FAA Academy during 1976-1977 and
for whom CSC test scores were available.

Sample Description. During the period November 1976 through January 1977,
the Civil Service Commission opened the competitive ATC examination
in the FAA Eastern and Southern regions. These included the areas en-
compassed by Regions II, III and IV of the Standard Federal Regions.
Approximately 11,500 applicants took the ATC Civil Service test during
this period. Of this group, about 7,500 also completed the experimental
tests and the Prior Experience Questionnaire. While the sample from the
11,500 population could not be controlled, differences between the means,
standard deviations and other statistics on the CSC tests for the 4,000
who did not complete the experimental tests or provide prior experience
information were not statistically significant from those who did.

Predictors. In addition to the five-part CSC test battery, two experi-
mental tests, MCAT 706A and the Directional Headings Test, were administered
to the applicant group. These test forms were the same versions as used
in the 1977 study conducted by Education and Public Affairs (Chapter 1 of
reference 14). The same Prior Experience Questionnaire (15) was also given.

All tests were administered by CSC examiners as part of the normal com-
petitive testing procedures. Applicants were informed that the experi-
mental tests would have no bearing on their eligibility status and the
experimental predictors were administered after completion of the competi-
tive CSC test. The Directional Headings Test and the Prior Experience
Questionnaire had to be manually scored and converted to ADP records.
This was done by FAA under contract. Determinations on whether appli-
cants were men or women were coded on the Prior Experience Questionnaire
based on the applicants name. At this point, Federal regulations pro-
hibited obtaining any ethnic or minority group information from the
applicants.

Criteria. Since this study encompassed ATC applicants, it was not feasible to establish operational criterion measures of validity. However, a variety of statistical analyses were made based on men and women applicants and those who passed the CSC test in contrast to those who failed.

Analytical Methodology. The data obtained on this ATC applicant group were analyzed with respect to frequency distributions, means, standard deviations, analysis of variance and intercorrelation of predictors.

Results. Table 13 shows the means and standard deviations for the total applicant group, the pass group (those who scored above 209 on the CSC test), and the fail group by men and women applicants.

TABLE 13

SELECTION TEST PERFORMANCES, APPLICANT GROUPS

CSC TESTS

TEST SEX	TOTAL GROUP (N)	MEAN	SD	PASS GROUP (N)	MEAN	SD	FAIL GROUP (N)	MEAN	SD
CSC-24									
Total	(7412)	39.66	9.6	(3960)	44.80	6.8	(3722)	34.61	9.3
Men	(5720)	39.63	9.5	(2980)	44.51**	6.9	(2740)	34.32**	9.2
Women	(1607)	39.83	9.7	(663)	45.91	6.6	(944)	35.56	9.3
CSC-51									
Total	(7412)	26.65	6.9	(3960)	30.74	3.8	(3722)	22.60	6.4
Men	(5720)	27.40	6.4	(2980)	31.05	3.6	(2740)	23.43	6.4
Women	(1607)	23.98**	7.0	(663)	29.37**	4.4	(944)	20.20**	6.0
CSC-135*									
Total	(7412)	21.49	8.6	(3960)	27.00	5.1	(3722)	16.03	7.8
Men	(5720)	21.88	8.4	(2980)	26.97	5.1	(2740)	16.35	7.8
Women	(1607)	20.13**	9.0	(663)	27.11	5.0	(944)	15.23**	7.9
CSC-157*									
Total	(7412)	29.29	10.3	(3960)	36.65	6.0	(3722)	21.99	8.3
Men	(5720)	29.47	10.2	(2980)	36.40**	6.0	(2740)	21.93	8.2
Women	(1607)	28.62**	10.9	(663)	37.75	5.8	(944)	22.21	8.0
CSC-540									
Total	(7412)	28.98	13.1	(3960)	37.70	10.1	(3722)	20.34	9.6
Men	(5720)	29.78	12.8	(2980)	37.84	9.9	(2740)	21.03	9.5
Women	(1607)	26.11**	13.7	(663)	37.00	10.8	(944)	18.46**	9.8
CSC TOTAL									
Total	(7412)	202.03	52.3	(3960)	244.26	23.2	(3722)	160.16	37.3
Men	(5720)	205.04	50.9	(2980)	244.23	22.0	(2740)	162.42	36.8
Women	(1607)	191.20**	55.2	(663)	244.26	24.0	(944)	154.08**	37.8

EXPERIMENTAL TESTS

TEST SEX	TOTAL GROUP (N)	MEAN	SD	PASS GROUP (N)	MEAN	SD	FAIL GROUP (N)	MEAN	SD
MCAT (CSC-510)									
Total	(6822)	25.76	9.1	(3340)	31.65	7.2	(3481)	20.12	6.9
Men	(5241)	26.57	9.0	(2688)	32.14	7.0	(2552)	20.72	7.8
Women	(1498)	22.95**	8.7	(607)	29.53	7.2**	(891)	18.46**	6.5
DIR HEADING									
Total	(7073)	46.99	22.6	(3583)	59.93	17.3	(3489)	33.68	19.4
Men	(5462)	48.87	21.9	(2899)	60.63	16.6	(2562)	35.56	19.5
Women	(1526)	40.17**	23.3	(637)	56.58	19.4**	(889)	28.41**	18.2

*CSC 135 (following oral directions) and CSC 157 (Abstract Reasoning) tests are doubled weighted in computing the CSC test total score

**Significant at the 1 percent level of confidence.

From Table 13 it is evident that in total, women score somewhat below
men on most of the CSC tests and on the two experimental tests. These
differences, while statistically significant, are not statistically
important due to the large sample size. For example, further analysis
shows that sex accounts for only 1 percent of the variance in the CSC
total test score; 3 percent in the MCAT; and 2 percent in the Directional
Headings total test scores. In other words, most of the variance in these
three predictors is due to factors other than sex.

In terms of the "pass" group, the mean CSC test score for men and women
is essentially the same. Consequently, discounting veterans preference
and additional credit allowed for aviation-related experience, women who
pass the CSC test should have about the same opportunity to be selected
for the ATC occupation as men.

Frequency distribution for each of the five CSC tests and the two experi-
mental tests were also examined. Two of the CSC tests, CSC 135 and CSC 51,
show a marked negative skew (-1.30 and -1.80) respectively) and conse-
quently, provide little differentiation between the applicants.

FIGURE 2

CSC - FOLLOWING ORAL DIRECTIONS
135   (DOUBLE WEIGHTED)

CSC - SPATIAL RELATIONS
51

MEAN = 21.5
S.D. = 8.6

MEAN = 26.7
S.D = 6.7

34
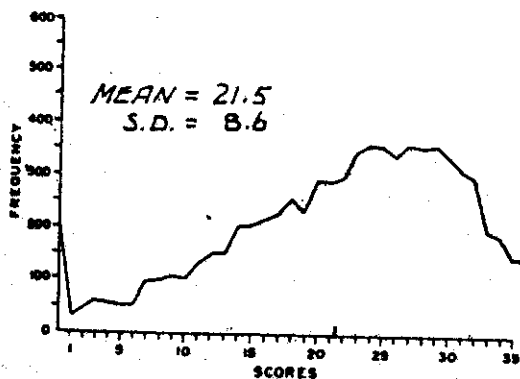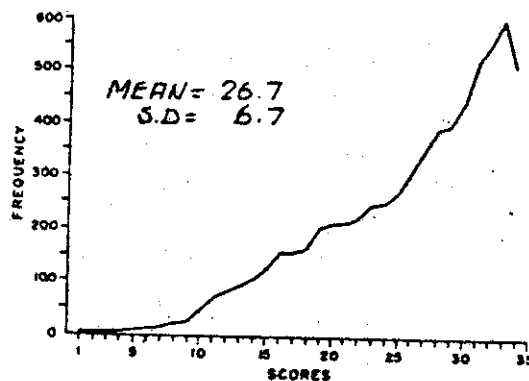
Frequency distribution graphs of all of the CSC and experimental tests are provided in reference 20.

Table 14 provides descriptive information on education and experience levels for men and women in the applicant group. Aviation-related experience data was obtained from the Prior Experience Questionnaire (PEQ). This identified the specific work or skill elements which, based on the present CSC Rating, allows granting extra credit points to those applicants who pass the CSC test.

TABLE 14

Descriptive Statistics-Applicant Group

EDUCATION LEVEL	TOTAL APPLICANTS	
Men with		
High school or less	1225	(21%)
Less than 3 years college	2352	(41%)
3 or more years college	2126	(37%)
Total	5703	(100%)
Women with:		
High school or less	355	(22%)
Less than 3 years college	618	(39%)
3 or more years college	630	(39%)
Total	1603	(100%)

EXPERIENCE LEVEL		
Men with		
No aviation-related experience	4393	(77%)
Aviation-related experience	1327	(23%)
Total	5720	(100%)
Women with:		
No aviation-related experience	1483	(92%)
Aviation-related experience	124	(8%)
Total	1607	(100%)

EDUCATION AND EXPERIENCE

With aviation-related experience	MEN	WOMEN
High school or less	21%	23%
Less than 3 years college	44%	39%
3 or more years college	35%	38%
No aviation-related experience		
High school or less	22%	22%
Less than 3 years college	40%	39%
3 or more years college	38%	39%

About 22 percent of the total applicants were women. However, as Table 14 shows, the distribution of education level for men and women is approximately the same. It also shows that approximately 78 percent of both the men and women have some education beyond high school.

35

With respect to aviation-related experience, 23 percent of the men as compared to only 8 percent of the women, identified experience which could result in granting extra credit points for selection eligibility. When education and experience are combined, the distribution of education level remains relatively consistent for both men and women and for those with or without aviation-related experience.

From this information it is evident that the level of education for men and women is essentially the same and does not serve to differentiate between the applicants based on their sex. The distribution of aviation-related experience between men and women does differentiate between the applicants with 23 percent of the men potentially being eligible for extra credit in contrast to only 8 percent of the women.

Next, the ATC applicant sample was analyzed in terms of passing or failing the CSC test in relation to sex, education and experience. Table 15 summarized the results of this analysis.

TABLE 15

ATC Applicant Sample

	TOTAL APPLICANTS	CSC TEST STATUS PASS	FAIL
Men	5720 (78%)	52%	48%
Women	1607 (22%)	41%	59%
Total	7327 (100%)	(50%)*	(50%)*

EDUCATION LEVEL

High School or less	1581 (22%)	32%	68%
Less than 3 years college	2973 (40%	49%	51%
3 or more years college	2757 (38%)	61%	39%
Total	7311 (100%)	(50%)*	(50%)*

EXPERIENCE LEVEL

No aviation-related experience	5876 (80%)	45%	55%
Aviation-related experience	1451 (20%)	68%	32%
Total	7327 (100%)	(50%)*	(50%)*

EDUCATION & EXPERIENCE

Aviation-related experience

High school or less	308 (21%)	52%	48%
Less than 3 years college	632 (44%)	67%	33%
3 or more years college	508 (35%)	79%	21%
Total	1448 (100%)	(68%)*	(32%)*

No Aviation experience with:

High school or less	1273 (22%)	27%	73%
Less than 3 years college	2341 (40%)	44%	56%
3 or more years college	2222 (38%)	57%	43%
Total	5836 (100%)	(45%)*	(55%)*

*Pass/Fail rates for the Total Group

36

In total, about 50 percent of the applicants passed the CSC test. However, proportionally more men (52 percent) passed than women (41 percent).

When specific education levels are examined, it is evident that those applicants with the most education have a significantly better chance of passing the test. Applicants with 3 or more years of college passed at about twice the rate of those who did not go beyond high school. The same relationship holds also for experience. Applicants with aviation-related experience also pass the CSC test at about twice the rate of those without experience.

When pass or fail rates are examined with education and experience combined, it is evident that: (1) level of education is essentially proportionally distributed between the experienced and non-experienced groups (Table 14); and (2) experience increased the pass rate in each of the education groups. Applicants with no college or experience pass at a 27 percent rate. Those with 3 or more years of college and aviation-related experience pass at a 79 percent rate.

TABLE 16

EXPERIENCE AND LEVEL OF EDUCATION

| | CSC TEST PASS RATE | | |
EDUCATION LEVEL	NO EXPERIENCE	WITH EXPERIENCE	% INCREASE
High School or less	27%	52%	+25%
Less than 3 years college	44%	67%	+23%
3 or more years college	57%	79%	+22%

Two-way analyses of variance were carried out for each CSC and experimental test. For this purpose the "experience group" was further classified into subgroups since the CSC Rating Guide provides different levels of extra credit based on specific types of experience. The "key" to the experience and education subgroups are as follows:

Aviation-Related Experience

0-1 No aviation-related experience
2-2 Comunication experience only
3-3 Non pilot with IFR or Air Defense Command (ADC) experience
4-4 Non pilot with VFR ATC experience
5-5 Pilot with IFR or ADC experience
6-6 Pilot   experience only
7-7 Pilot with VFR ATC experience

37

Education

0-1 High school or less
2-2 Less than 3 years of college
3-3 3 or more years of college

Table 17 provides the results of the analysis variance for the CSC total test score. The lowest mean score (173.10) is for the "No Experience and No College " Group; the highest mean score (241.00) is for the "Pilot with VFR Experience and 3 or more years of College" Group. The analysis of variance shows that experience (row) and education (column) each is statistically significant ($p \leq .01$) though there is no interaction between experience and education.

Further analysis, however, shows they are not statistically important for the total group since only 1 percent of the variance in CSC total score was accounted for by either variable.

TABLE 17

ATC TESTS VS. EDUCATION AND EXPERIENCE GROUPS

TOTAL CSC TEST SCORE (CSS TOTAL WT.)

N MATRIX OF CASES WITH NON-MISSION DATA ON ALL 3 VARIABLES

	ED GROUP GOES DOWN			EXP. GROUP GOES ACROSS				
	0-1	2-2	3-3	4-4	5-5	6-6	7-7	TOTAL N
0-1 -	1239	34	140	28	19	67	54	1581
2-2 -	2266	75	216	47	56	180	133	2973
3-3 -	2179	70	62	25	31	230	160	2757
SUM	5684	179	418	100	106	477	347	7311

MEANS OF VARIABLE   CSC TOTAL WT

	ED GROUP GOES DOWN			EXP. GROUP GOES ACROSS				
	0-1	2-2	3-3	4-4	5-5	6-6	7-7	MEAN
0-1 -	173.10	179.53	210.89	201.96	198.68	191.07	211.56	179.48
2-2 -	197.40	204.79	218.55	233.23	220.61	221.52	225.19	202.83
3-3 -	209.11	221.59	224.98	222.72	246.10	236.17	241.00	214.43
MEAN	196.59	206.56	216.94	221.85	224.13	224.31	230.36	202.16

SUMMARY OF ANALYSIS OF VARIANCE

Source						
Row	248889.27	2	124444.63	51.25	0.000
Col	152242.89	6	25373.82	10.45	0.000
R.BY.C	44320.17	12	3693.35	1.52	0.109
WITHIN CELL	17701966.50	7290	2428.25		

However, when the CSC total score was categorized by "pass" or "fail" groups, education and experience together accounted for nearly 10 percent of the variance in pass/fail status; so taken together, the effects of education and experience are important. This same relationship applies to the two experimental tests as well.

These data show: (1) that applicants with aviation-related experience have a higher probability of passing the CSC test (68 percent as compared to 45 percent for those without experience); and, (2) that few women have aviation-related experience. These two findings taken together, however, are not sufficient to explain the lower pass rate for women (41 percent) as compared to men (52 percent). Applying these probabilities to the experience and no experience groups, it would be expected that 712 women would have passed; only 663 did pass. For men, 2,879 men would be expected to pass; 2,980 actually did pass. Consequently, simply controlling for the fact that men have more aviation-related experience does not explain why women score lower on most of the tests.

Intercorrelation of each of the tests, education and experience are provided in Table 16.

<div align="center">

TABLE 18

</div>

ATC TEST CORRELATIONS -- TOTAL SAMPLE

POSITION LABEL	1 T 24.S CORE	2 T 51.S CORE	3 TS 40.S CORE	4 T1 57.5 CORE	5 T1 35.S CORE	6 C SC.T OTAL	7 CSC TOTA L.WT	8 TS 10.C	9 TS 10.A	10 TS 10.T OTAL	11 DH.A	12 DH.B	13 DH.T GTAL	14 ED.L EVEL	15 PEQ ACG	16 ED.G ROUP	17 XP.G ROUP	18 EXP GPX
T24 SCORE	100	33	56	80	51	75	69	44	52	53	46	47	49	27	6	26	10	11
T51 SCORE	33	100	46	59	54	69	75	42	60	57	54	54	57	11	18	12	23	19
T540 SCORE	56	46	100	99	58	85	80	56	63	6	55	56	59	16	13	16	14	5
T157 SCORE	50	59	59	100	63	83	89	49	63	62	53	56	57	22	13	22	17	14
T135 SCORE	51	54	58	63	100	81	79	52	66	65	54	58	59	20	15	21	16	7
CSC TOTAL	75	69	85	83	81	100	99	62	77	77	66	68	71	24	16	25	19	14
CSC TOTAL WT.	69	75	80	89	79	99	100	61	77	76	66	68	71	24	17	24	20	15
T510 C	44	42	56	49	52	62	61	100	69	89	54	55	57	14	17	14	16	5
T510 A	52	60	63	63	66	77	77	69	100	94	65	67	69	18	22	18	25	12
T510 TOTAL	53	57	65	62	66	77	76	89	94	100	55	67	70	18	22	18	24	10
DH.A	46	54	55	83	54	66	66	54	65	66	100	82	95	20	26	20	33	24
DH.B	47	54	56	56	58	68	68	55	67	67	82	100	95	20	22	20	27	21
DHTOTAL	49	57	59	57	59	71	71	57	69	70	95	95	100	21	26	21	31	25
ED. LEVEL	27	11	16	22	20	24	24	14	18	16	20	20	21	100	6	95	5	25
ED. LEVEL	6	18	13	13	15	16	17	17	22	22	26	22	26	6	100	6	67	25
PEQ ACG	26	12	14	22	21	25	24	14	16	16	20	20	21	95	6	100	4	37
ED.GROUP	10	23	14	17	16	19	20	18	25	24	33	27	31	5	87	4	100	100
EXP. GROUP	11	29	5	14	7	14	15	5	12	10	96	21	15	25	37	22	100	100
EXP. GPX																		

KEYS:

(6) CSC Total is the correlation of the unweighted CSC test scores
(7) CSC Total WT is the correlation of the weighted CSC test scores
(8) T510 C is the correlation of MCAT "aircraft conflict" scores
(9) T510 A is the correlation of MCAT "aptitude" scores
(14) Ed Level is the correlation of individual levels of education (01-13) See ref (14) c-1
(15) PEQ. ACG is the correlation of aviation-related experience scales derived empirically. See ref (12) Chapter 4 for the methodology for developing the scales
(16) Ed Group is the correlation of the three education groups; (1) high school or less; (2) less than 3 years of college; (3) 3 or more years of college.
(17) Exp Group is the correlation of the seven aviation-related experience groups
(18) Exp GP. X is the correlation of the group with aviation-related experience (excludingthe no experience applicants)

These correlation values were used to correct for restrictions in range on test scores of individuals hired as ATC trainees who attended the FAA Academy during 1976, 1977 and 1978.

Several multiple regression analyses were carried out using CSC test "pass" or "fail" as a criterion for the applicant group. Table 19 shows the results for the present CSC test battery, the order in which the individual tests entered the analysis and their contribution to the R^2 which measures the amount of variance accounted for by each test.

TABLE 19

ATC REGRESSION ANALYSIS TOTAL SAMPLE

FINAL SUMMARY OF REGRESSION ON DEPENDENT VARIABLE CSC.TOT.PF

MULTIPLE R	0.8058
MULT R SQUARED	0.6493
STD. ERROR OF EST.	0.2962
CONSTANT	0.0037

ANALYSIS OF VARIANCE	DF	SUM OF SQUARES	MEAN SQUARE	F RATIO	PROB. LEVEL
REGRESSION	5.	1203.100	240.620	2742.182	0.
RESIDUAL	7406.	848.846	0.088		

S T E P	NUM VAR NOW IN	MULT R	MULT R SQ.	CHANGE IN RSQ	VARIABLE ENTERED (* SHOWS DELETED)	B. RAW COEF- FICIENT	STAND ERROR OF B	BETA. STAND. COEF- FICIENT	FINAL F TO DELETE	F WHEN ENTERED OR DELETED	SIMPLE COR. WITH DEP.	PARTIAL COR. IN FINAL STEP
1	1	0.711	0.505	0.505	T157.SCORE	0.0148	0.0005	0.3047	894.793	7558.264	0.7106	0.3283
2	2	0.772	0.597	0.092	T540.SCORE	0.0096	0.0004	0.2530	896.450	1686.531	0.6815	0.2936
3	3	0.794	0.630	0.033	T51.SCORE	0.0152	0.0007	0.2034	519.011	663.970	0.6062	0.2966
4	4	0.802	0.644	0.014	T135.SCORE	0.0083	0.0006	0.1416	207.233	283.007	0.6390	0.1660
5	5	0.806	0.649	0.006	T24.SCORE	0.0050	0.0005	0.0967	121.331	121.331	0.6291	0.1270

Table 20 provides the results of the two experimental tests in predicting the CSC pass or fail status. The difference between the R^2 for these tests (45 percent) compared to the CSC tests (65 percent) indicate that they differ in some degree from the CSC tests in what they measure.

In the case of the experimental tests, test 510 A (MCAT Aptitude) which measures aptitude similar to the present CSC test entered the regression first and accounts for 40 percent of the total variance.

TABLE 20

ATC REGRESSION ANALYSIS TOTAL SAMPLE

FINAL SUMMARY OF REGRESSION ON DEPENDENT VARIABLE CSC.TOT.PF

MULTIPLE R	0.6675
MULT R SQUARED	0.4456
STD. ERROR OF EST.	0.3724
CONSTANT	0.6590

ANALYSIS OF VARIANCE	DF	SUM OF SQUARES	MEAN SQUARE	F RATIO	PROB. LEVEL
REGRESSION	3.	825.714	275.238	1984.871	0.
RESIDUAL	7406.	1027.232	0.138		

S T E P	NUM VAR NOW IN	MULT R	MULT R SQ.	CHANGE IN RSQ	VARIABLE ENTERED (* SHOWS DELETED)	B. RAW COEF- FICIENT	STAND ERROR OF B	BETA. STAND. COEF- FICIENT	FINAL F TO DELETE	F WHEN ENTERED OR DELETED	SIMPLE COR. WITH DEP.	PARTIAL COR. IN FINAL STEP
1	1	0.631	0.399	0.399	T510.A	0.0326	0.0012	0.3722	719.144	4913.448	0.6314	0.2975
2	2	0.662	0.439	0.040	SM.TOTAL	0.0057	0.0003	0.2570	444.547	528.509	0.5818	0.2375
3	3	0.668	0.446	0.007	T510.C	0.0141	0.0013	0.1170	92.136	92.136	0.5318	0.1106

Objective. This study was initiated by the FAA Civil Aeromedical Institute (CAMI) in July 1976 to evaluate the present CSC Test Battery and experimental predictors in relation to the success of newly hired ATC trainees in passing Initial Qualification training for the En Route and Terminal ATC options at the FAA Academy. Upon entry on duty, new ATC trainees are given a 2-week general orientation at their ATC facility (or regional headquarters). They then enter Initial Qualification training at the FAA Academy in Oklahoma City for the ATC option to which assigned. On their first day at the Academy, they are given a series of experimental tests and questionnaires. Their participation in the test program is voluntary. The experience has been that most all of the students volunteered. Extensive efforts were made to also obtain the individuals' test scores on the five-part CSC battery and their total earned ratings which include additional credit for veterans preference and aviation-related experience as provided by the CSC Rating Guide. This information was then correlated with their subsequent training status to evaluate the CSC and experimental tests which were most predictive of training success.

Sample Description. During the period July 1976 through April 1978, 3,008 students entered Terminal and En route ATC training at the FAA Academy. Selected characteristics of this total group are provided in Table 21.

TABLE 21

SELECTION CHARACTERISTICS-TOTAL GROUP

	TOTAL N	PASS N	PASS (%)	FAIL N	FAIL (%)	WITHDRAW N	WITHDRAW (%)
Men	2580	2034	(79%)	436	(17%)	110	(4%)
Women	428	272	(64%)	128	(30%)	28	(6%)
Total	3008	2306	(77%)	564	(19%)	138	(4%)
No Exp.	834	579	(69%)	206	(24%)	58	(7%)
Aviation-Related Exp.	1936	1579	(82%)	296	(15%)	61	(3%)
Unknown	229	148	(65%)	62	(27%)	19	(8%)
Total	3008	2306		564		138	

41

In order to evaluate the CSC and experimental predictors, it was necessary to have complete information on the students. The 138 students who withdrew did not have laboratory criterion scores and were eliminated from the sample. Similarly, the students for whom prior experience information was missing were also excluded. An additional 824 students were excluded because they did not have one or more test scores, either from the CSC test battery taken prior to their employment with FAA or were students who did not volunteer to take the experimental tests.

The final sample used for the analysis after these exclusions consisted of 1,827 students distributed as shown in Table 22.

TABLE 22

SELECTION CHARACTERISTICS-ANALYTICAL SAMPLE

	TOTAL N	PASS N	PASS (%)	FAIL N	FAIL (%)	% OF ORIGINAL POPULATION
Men	1587	1314	(83%)	273	(17%)	61.5%
Women	240	162	(68%)	78	(32%)	56.1%
Total	1827	1476	(81%)	351	(19%)	60.7%
No Exp.	558	409	(73%)	149	(27%)	66.2%
Aviation-Related Exp.	1269	1067	(84%)	202	(16%)	66.5%
	1827	1476		351		

The failure rate for this sample (19.2 percent) was essentially the same as for the total population (19.7 percent) after withdrawals were excluded.

Predictors. The tests used in this study were:

- Civil Service Test Battery

    (CSC 24; CSC 51; CSC 135; CSC 157; and, CSC 540)

- Directional Headings Test (DHT) Total Score

    DHT Part A
    DHT Part B

- Multiplex Controller Aptitude Test (MCAT) Total Score

    MCAT Aptitude (A)
    MCAT Conflicts (C)

- Dial Reading (DL RD)

- ATC Occupational Knowledge Test (OKT)

42

The Directional Heading Test was the same version used in the 1977 study conducted by EPA and which was administered to the 7,500 ATC applicants in 1976-1977.

The MCAT consisted of several different versions (606 A; 606 B; 706 A; 706 B; 607; 707) which varied in the number and mix of questions (aptitude/conflict) and in the length of time allowed for the test.

The OKT was form 101B which was essentially the same form utilized by EPA in their 1977 study. The Dial Reading Test was also the same test used by EPA.

Table 23 provides the means and standard deviations for the sample of 1,827 ATC trainees at the FAA Academy on the various predictors and compares them to results obtained for the 1976-1977 ATC applicants group as well as with the ATC sample group included in the 1977 ATC Selection Study conducted by EPA where comparable statistics were obtained.

TABLE 23

COMPARATIVE DESCRIPTIVE STATISTICS

	1976-77 ATC APPLICANTS				1977		1976-78 ATC	
	CSC TOTAL		CSC PASS		EPA STUDY		TRAINEES	
	(N=7412-6821)		(N=3690-3340)		(N=1323-1229)		(N-1827)	
TEST	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
CSC-24	39.66	9.6	44.80	6.8	NA		47.07	6.7
CSC-51	26.65	6.7	30.74	3.8	NA		31.90	3.3
CSC-540	28.98	13.1	37.70	10.1	NA		42.88	9.7
CSC-157	29.99	10.3	36.65	6.0	NA		38.17	6.4
CSC-135	21.50	8.6	27.00	5.1	NA		29.31	4.2
MCAT (A)	16.59	5.7	20.27	4.1	NC		23.14	4.5
MCAT (C)	9.17	4.2	11.38	3.9	NC		15.74	4.1
MCAT TOT	22.76	9.1	31.65	7.2	NC		38.87	7.6
DHT A	24.20	12.0	30.69	9.8	NA		32.68	9.0
DHT B	22.80	11.7	29.25	8.9	31.9	6.9	31.63	9.0
DHT TOT	47.00	22.6	59.93	17.3	NA		64.34	17.0
DL·RD	NA		NA		39.7	9.4	40.97	9.2
OKT	NA		NA		74.9	13.6	66.96	16.3

(Table 23 Continued)

NA – Statistics not available since the tests were either not
administered to these groups or could not be obtained from
available records. In the case of DHT, for the 1977 EPA
Study only Part B was scored and used in the analysis.

NC – Statistics not comparable. In the 1977 EPA Study, two forms
of MCAT were administered to the test sample (706 A, 606 A).
In the analysis all "aptitude" questions from both forms were
combined for a total aptitude score (mean = 42.8; SD-6.2) as
were all "conflict" questions (mean = 29.4; SD-5.6).

Criterion Measure. In this analysis, the ATC laboratory average score was
used as the criterion measure of ATC training success. Selection of this
criterion was based on several considerations. First, prior studies
demonstrated that these scores were the most reliable predictors of subse-
quent success as an Air Traffic Controller. Second, the laboratory training
phase consists of a series of operational air traffic control problems in
which students must demonstrate their ability to apply the academic know-
ledge and skills acquired in training. Finally, this phase is conducted on
a pass/fail basis. The laboratory score constitutes 65 percent of the
students total score in the laboratory phase of training. Almost all of
of the students who fail Initial Qualifications training do so in the
laboratory training phase. It should be pointed out that during the period
from July 1976 through April 1978 several changes were made in the laboratory
training phase which affected the pass/fail ratio of students. In July 1976
the weight given to the laboratory average was increased from 35 percent to 65
percent. In September 1976 the number of graded lab problems was increased
from four to six; and in May 1977 "procedural errors" were incorporated in
the grading for failure to handle aircraft in a timely manner during the
problem exercise. In order to accomodate these changes and the differences
between the Enroute and Terminal laboratory procedures, the laboratory
averages were converted to standard scores and data from the two ATC options
combined. The standardized laboratory average criterion is labeled "ZLAB."

The reliability of the laboratory average for Enroute training is .80
(2,223 students) and for the Terminal option .81 (1,982 students).

A detailed discussion of the ATC Initial Qualifications training program
and follow-on training at the ATC facility level for the Terminal and
Enroute options are provided in references 23, 24, and 25.

44

Analytical Methodology. The first step in the analysis was the calculation
of descriptive statistics. The remaining analyses presented several unique
problems.

Several different experimental forms of the MCAT were employed in the
testing at the FAA Academy and the order of administration varied for
each form. Consequently, it was necessary to standardize the scores for
these tests. Since MCAT 706 A was also administered to 6,821 ATC appli-
cants by OPM in 1976-1977, the scores on the various forms of MCAT
used by CAMI were standardized by linear conversions using the same metric
as MCAT 706 A. The order effect was handled by using the scores from MCAT
706 A given first. References 19 and 20 provide a more detailed dis-
cussion on the methodology used for these conversions.

The second problem involved in evaluating the CSC and experimental tests
was the restriction in range effect (21) since criterion (ZLAB) information
was available only on those individuals who were hired. This results in a
spurious low correlation of CSC test scores with the criterion. In order
to adjust the restricted correlations so they would reflect the relationship
between the tests and the criterion for the applicant group, the correla-
tions were corrected for their restriction in range. Usual methods for
this correction in the three variables case assume that unrestricted infor-
mation is available only on the variable used for selection or the third
incidental variable but not both. With the information derived from the
1976-1977 ATC applicant population, unrestricted data were available on
both variables. A modified procedure was developed to make use of all of
the available information in correcting for range restriction. A full
discussion of the procedure and the derivation of equations is provided in
reference 22.

The unrestricted correlation matrix from the total 1976-1977 ATC applicant
group used for correcting for range restriction is provided in Table 24.
The means, standard deviations and sample sizes (N's) were provided in
Table 23 except for the Earned Rating.

Table 24. Unrestricted Correlation Matrix From the CSC

	VARIABLE	1 T 24. SCORE	2 T 51. SCORE	3 T 540. SCORE	4 T 157. SCORE	5 T 135. SCORE	9 T 510. A	10 T 510. C	11 T 510. TOTAL	12 DH. A	13 DH. B	14 DH. TOTAL	19 EAR NED. RATE
1	T24.Score	100	33	56	50	51	52	44	53	46	47	49	66
2	T51.Score	33	100	46	59	54	60	42	57	54	54	57	75
3	T540.Score	56	46	100	59	58	63	56	65	55	56	59	78
4	T157.Score	50	59	59	100	63	63	49	62	53	56	57	85
5	T135.Score	51	54	58	63	100	66	52	65	54	58	59	77
9	T510.A	52	60	63	63	66	100	69	94	65	67	69	77
10	T510.C	44	42	56	49	52	69	100	89	54	55	57	60
11	T510.TOTAL	53	57	65	62	65	94	89	100	66	67	70	76
12	DH.A	46	54	55	53	54	65	54	66	100	82	95	68
13	DH.B	47	54	56	56	58	67	55	67	82	100	95	69
14	DH.TOTAL	49	57	59	57	59	69	57	70	95	95	100	72
19	EARNED RATE	66	75	78	85	77	77	60	76	68	69	72	100

45

The Earned Rating for the applicant group was derived by taking the raw weighted CSC total score, transmuting it to a metric scale of 0-100 and, where appropriate, adding veterans preference (5 points) and additional credit for experience (0, 5, 10, or 15 points based on the CSC Rating Guide criteria) to the scores of applicants who passed the CSC Test Battery. This Earned Rating is the basis on which OPM ranks and places eligible candidates on a register from which selections are made in accordance with OPM regulations.

Table 25 provides the restricted correlation matrix for the 1,827 ATC trainees at the ATC Academy. The means and standard deviations for this sample were previously provided in Table 23. It should be note that the stand-ardized (Z) scores were used for the MCAT tests in deriving the intercor-relation (mean = 0; SD = ± 1).

TABLE 25

RESTRICTED CORRELATION MATRIX USED IN REGRESSION ANALYSES

	CSC 24	CSC 540	CSC 157	DHT A	DHT B	DHT TOT	MCAT A	MCAT C	MCAT TOT	ZLAB	DL·RD
CSC 24	1.000	0.333	0.140	0.245	0.255	0.271	0.214	0.146	0.204	0.097	0.343
CSC 540	0.333	1.000	0.145	0.199	0.222	0.227	0.204	0.190	0.221	0.096	0.289
CSC 157	0.140	0.145	1.000	0.090	0.140	0.126	0.241	0.164	0.229	0.071	0.281
DHT A	0.245	0.199	0.090	1.000	0.803	0.950	0.258	0.268	0.296	0.207	0.338
DHT B	0.255	0.222	0.140	0.803	1.000	0.949	0.284	0.268	0.308	0.223	0.341
DHT TOT	0.271	0.227	0.126	0.950	0.949	1.000	0.282	0.282	0.316	0.227	0.356
Z MCAT A	0.214	0.204	0.241	0.258	0.284	0.282	1.000	0.596	0.903	0.246	0.481
Z MCAT B	0.146	0.190	0.164	0.268	0.268	0.282	0.596	1.000	0.882	0.250	0.445
Z MCAT TOT	0.204	0.221	0.229	0.296	0.308	0.316	0.903	0.882	1.000	0.277	0.518
Z LAB	0.097	0.096	0.071	0.207	0.223	0.227	0.246	0.250	0.277	1.000	0.272
DL·RD	0.343	0.289	0.281	0.338	0.341	0.356	0.481	0.445	0.518	0.272	1.000

The correlations of particular interest are those between the tests and the "ZLAB" criterion. These are zero order validity coefficients. The effects of restriction are immediately apparent in the low correlations between the CSC tests used in the actual selection and the ZLAB criterion. The two highest zero order coefficients are MCAT TOTAL (.277) and DL·RD .272, neither of which was restricted directly by selection of the trainees. Two CSC tests, 51—Spatial Patterns and 135—Following Oral Directions were omitted from the matrix. They were eliminated based on the negative skew previously discussed. Extreme selection, such as in the case of ATC trainees, results in a sharp reduction in the variance in the selected groups as was shown in Tables 13 and 23 which compares the means and standard deviations for the total ATC applicant group and the ATC appli-cant Pass Group, and the ATC trainee groups.

46

When correcting for restriction in range, the difference between the applicant group variance and the selected group variance is employed as a measure of the amount of curtailment that occurs due to selection. It was not determined if the skew resulted in a violation of the linearity assumption; however, the extreme disparity between the two variances for CSC 51 and CSC 135 resulted in a corrected correlation that was much higher than the other corrected correlations. In this analysis, if the correlations for CSC 51 and CSC 135 were used in the multiple regression analysis, none of the other tests either independently or in combination added anything significant to the multiple R beyond CSC 51 and CSC 135. These results were considered spurious and CSC 51 and CSC 135 were excluded from the following multiple regression analyses.

Table 26 provides the estimated unrestricted correlations (as well as the actual unrestricted correlation from the CSC sample). The correlations of primary interest are the correlations of the tests with the ZLAB criterion. After correcting for range restriction, the MCAT at .531, Dial Reading at .466, and Directional Heading at .461 have the highest zero order validity coefficients.

Table 26 Unrestricted and Corrected Correlations Used in Regression Analyses

	CSC 24	CSC 540	CSC 157	DHT 1	DHT 2	DHT TOT	MCAT A	MCAT C	MCAT TOT	ZLAB	DL·RD
CSC 24	1.000	.560	.500	.460	.470	.490	.520	.440	.530	.342	.515
CSC 540	.560	1.000	.590	.550	.560	.590	.630	.560	.650	.386	.501
CSC 157	.500	.590	1.000	.530	.560	.570	.630	.490	.620	.399	.502
DHT	.460	.550	.530	1.000	.820	.950	.650	.540	.660	.432	.498
DHT 2	.470	.560	.560	.820	1.000	.950	.670	.550	.670	.446	.510
DHT TOT	.490	.590	.570	.950	.950	1.000	.690	.570	.700	.461	.522
MCAT A	.520	.630	.630	.650	.670	.690	1.000	.690	.940	.503	.611
MCAT C	.440	.560	.490	.540	.550	.570	.690	1.000	.890	.475	.551
MCAT TOT	.530	.650	.620	.660	.670	.700	.940	.890	1.000	.531	.635
ZLAB	.342	.386	.399	.432	.446	.461	.503	.475	.531	1.000	.466
DL·RD	.515	.501	.502	.498	.510	.522	.611	.551	.635	.466	1.000

Results. The next step in the analyses was to employ the unrestricted and corrected correlations in a series of step-wise multiple regression procedures. Each model was a refinement of the previous model, with test scores regressed on ZLAB. The first model excluded total scores for DHT and MCAT, since these are the sum of their past scores that were included. Results are presented in Table 27.

Table 27

INITIAL REGRESSION MODEL

R = 0.5689   RSQ = 0.3236

V	BETA	B
CSC 24	0.0071	0.0011
CSC 540	-0.0066	-0.0007
CSC 157	0.0555	0.0043
DHT A	0.0513	0.0057
DHT B	0.0912	0.0101
MCAT A	0.1452	0.0322
MCAT C	0.1668	0.0407
DL 'RD	0.1856	0.0201
REG. CONST.	•	-3.0582

F-TEST :   TOTAL MODEL WITH PART SCORES
RSQ FULL   =   0.3236   Model 1
RSQ REDUCED =   0.0060   Model 0
DIFFERENCE =   0.3236
DFN = 7.   DFD = 1800.   F-RATIO = 123.020   P < 0.0001

47

In succeeding models, CSC 540 was eliminated because of its lack of contribution, and total scores were used for DHT and MCAT instead of part scores. The resulting multiple "R" was .5672. Then Dial Reading was eliminated because of its marginal value in the 1972 and 1977 FAA ATC Specialist studies, and DHT was eliminated because, in its current format (very brief and highly speeded) it is not suitable for operational use. Results are presented in Table 28. The initial model yielded a multiple "R" of .5689, the final model yielded a multiple "R" of .5407.

Table 28

FINAL REGRESSION MODEL

R = 0.5407   RSQ = 0.2924

V	BETA	B
CSC 24	0.0608	0.0090
CSC 157	0.0964	0.0074
MCAT T	0.4391	0.0570
REC. CONST.	-	03.2045

F-TEST   24, 157, MCAT TOTAL
RSQ FULL   =   0.2924   Model 5
RSQ REDUCED =   0.0000   Model 0
DIFFERENCE =   0.2924
DFN = 2.   DFD = 1800.   F-RATIO = 371.890   P < 0.0001

A factor analysis (principal axis analysis with varimax rotation) was performed to explore the characteristics of the the test scores. The results are provided in Table 29.

Table 29   Factor Analysis on Test Scores (Principal Axis Analysis - Varimax Rotation)

Variables	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
CSC 24	0.9731	-0.0550	-0.0261	-0.0181	0.2583	-0.0211
CSC 540	0.4396	0.3568	-0.1129	-0.5921	0.5849	-0.0564
CSC 157	0.3578	0.0785	-0.7302	0.0079	0.5883	-0.0350
DHT T	0.2638	-0.1068	0.0238	-0.0347	0.7255	0.0543
MCAT T	0.2715	-0.1042	-0.0066	-0.0305	0.7357	0.0522
ZLAB	0.2834	-0.1036	0.0143	-0.0415	0.7666	0.0563
DL·RD	0.3137	-0.0865	-0.0788	-0.0704	0.7740	0.0315
% variance accounted for	22.7188	2.5335	7.9083	5.1392	42.9871	0.2097

Two rather clear structures underlying the data with the orthogonal rotation appear. Factor 1 and Factor 5 account for 22.72 and 42.99 percent of the variance respectively. Factor 5 contains the largest loading for all tests and ZLAB with the exception of CSC 24. It is also notable that the division seems to occur on both factors between the CSC tests (24,540, 157) and the experimental tests (MCAT, DHT, DR·RD) and ZLAB.

48

The beta weights derived from the Final Regression Model (Table 28) were converted to raw score weights and then assigned unit weights. The following equation constitutes the composite score:

$$Y_c = 1(CSC\ 24) + 2(CSC\ 157) + 4(MCAT)$$

where Y_c = the composite test score for the battery. Using the unit weights, the multiple R derived was .5354 compared to .5407 with beta weights. The R^2 value derived was .2867 with unit weights compared to .2924 using beta weights.

A cross-validation study was performed to investigate the stability of the results. The sample was randomly separated into two groups, and the weights derived from the first sample were applied to the second sample to observe shrinkage in the multiple R. Distribution statistics and intercorrelations for the two groups are presented in Table 30.

TABLE 30 DISTRIBUTION BY RACE AND SEX AND
DESCRIPTIVE STATISTICS

SAMPLE 1

	Men	Women	Total
BLACK	47	17	64
HISPANIC	14	3	17
AM. INDIAN	0	1	1
ORIENTAL	6	1	7
ESKIMO	1	0	1
OTHER	730	94	824
TOTAL	798	116	914

	MEAN	S.D.
ZLAB	0.028	1.007
CSC 24	46.998	6.871
CSC 157	38.490	6.538
MCAT	35.608	7.451

CORRELATIONS

	ZLAB	CSC 24	CSC 157	MCAT
ZLAB	1.000	0.328	0.402	0.537
CSC 24		1.000	0.500	0.530
CSC 157			1.000	0.620
MCAT				1.000

SAMPLE 2

	Men	Women	Total
BLACK	45	16	61
HISPANIC	15	4	19
AM. INDIAN	0	0	0
ORIENTAL	7	1	8
ESKIMO	1	1	2
OTHER	723	101	824
TOTAL	791	123	914

	MEAN	S.D.
ZLAB	-0.020	0.990
CSC 24	47.026	6.853
CSC 157	38.252	6.244
MCAT	35.686	7.307

CORRELATIONS

	ZLAB	CSC 24	CSC 157	MCAT
ZLAB	1.000	0.326	0.396	0.527
CSC 24		1.000	0.500	0.530
CSC 157			1.000	0.620
MCAT				1.000

The regression equation for the first sample is presented in Table 31.

Table 31

CROSSVALIDATION CSC SELECTION STUDY

R = 0.5450   RSQ = 0.2970

VARIABLE	BETA	B
CSC 24	0.0354	0.0052
CSC 157	0.1023	0.0155
MCAT	0.4548	0.0615
REG. CONST.	-	3.0113

F-TEST		CROSSVALIDATION			
RSQ FULL	=	0.2879	Model 1		
RSQ REDUCED	=	0.0000	Model 0		
DIFFERENCE	=	0.2970			
DFN - 2.		DFD = 913.	F-RATIO - 192.847	$P < 0.0001$

Unit weights, derived from the beta values for Sample 1 were computed.
These unit weights (CSC 24 = 1; CSC 157 = 2; and, MCAT = 4) were then used
to compute the multiple Rs and R^2s in Sample 1 and Sample 2 with the
following results:

TABLE 32   CALCULATED MULTIPLE R VALUES

Sample	Multiple R	R^2
1	.5381	.2895
2	.5292	.2801

Full discussion of the analysis and results on the CSC and experimental
tests for appointees to the ATC occupation in 1976-1978 is provide in
reference [20] . The results from these analyses supported the conclusion
that CSC 24, CSC 157 and MCAT tests be used in combination as a selection
battery for screening applicant air traffic controllers. Some information
was provided that indicated a modified form of the Directional Headings
test suitable for operational use and a Dial Reading Test should be explored
in future studies.

In the analyses of tests in relation to ATC trainees hired during
1976-1978 discussed thus far, the ATC Occupational Knowledge Test (OKT) was
excluded from the various regression analyses. The OKT is a "job-knowledge
specific" test and is not intended for use as a "pass" or "fail" exam in
determining applicant eligibility for employment consideration. Rather its
use would be limited to measuring ATC related experience and knowledge as
a basis for granting extra point credit to those applicants who pass the
competitive test battery. Consequently, the ATC Occupational Knowledge
Test (OKT) was analyzed as a parallel effort.

Objective. The analysis was conducted to determine the relationship of OKT to:

- the present method of granting extra credit based solely on ATC related experience.

- Pass/fail status during Initial ATC Qualification training.

- An ATC selection test battery comprised of CSC tests 24, 157 and MCAT and the combined estimated validity of the test battery and OKT.

Sample Description. The sample for the analysis of OKT was the same group of 1827 ATC trainees (Table 22) that was used for regression analyses of the CSC and experimental tests previously discussed.

Predictor. The OKT (form 101B) is a 100-item multiple choice test covering air traffic control regulations, communication, flight service station work, navigation aids, weather, and radar. The score is the total number of correct responses. The mean and standard deviation for the 1827 ATC trainees was 66.96 and 16.3 respectively (Table 23). The Kuder-Richardson - 20 internal consistancy estimate of reliability yielded a reliability cofficient of .95 for a sample of ATC new hires.

Criterion. The same criterion measure, ATC laboratory average score (ZLAB), that was used in the multiple regression study was used in analyses of OKT.

Analytical Methodology. The 1827 ATC trainees were grouped into various ATC related experience categories based on their response to a biographical questionnaire administered on the first day of training. Four types of ATC related experience were identified.

1. Other experience (included air defense command, communications operator and prior ATC training without ATC operational experience).

2. Pilot experience

3. VFR or non-radar ATC experience

4. IFR or radar control ATC experience

These four types of experience (together with "No ATC related experience") were then used to classify each of the 1827 trainees into 16 experience subgroups which provided for combinations of experience as listed below.

Subgroup	Type of Experience	(N)	Group (N)
1	No ATC related experience	558 }	592
2	Other experience	34 }	
3	Pilot	297 }	389
4	Pilot + Other	92 }	
5	VFR only	99 }	
6	VFR + Other	27 }	191
7	VFR + Pilot	28 }	
8	VFR + Pilot + Other	37 }	
9	IFR only	86 }	
10	IFR + Other	29 }	144
11	IFR + Pilot	18 }	
12	IFR + Pilot + Other	11 }	
13	IFR + VFR	275 }	
14	IFR + VFR + Other	95 }	511
15	IFR + VFR + Pilot	57 }	
16	IFR + VFR + Pilot + Other	84 }	
Total		1827	

Because of the small number of trainees in some of the subgroups, they were combined into the following five major groups:

Group 1: No experience + Other experience only   (592)
Group 2: Pilot or Pilot + Other experience   (389)
Group 3: VFR + any additional experience except IFR   (191)
Group 4: IFR + any additional experience except VFR   (144)
Group 5: IFR + VFR + any additional experience   (511)

The test scores for OKT were correlated with experience and later correlated with the criterion ZLAB. Statistical analyses of each of the five experience groups in relation to OKT score ranges and pass/fail status was also conducted.

Results. The correlation of OKT scores with experience was .61. OKT correlated with the ZLAB criterion at .22 versus .11 for experience credit. This indicates that while OKT is closely related to experience, it is more predictive of success than the present method of giving credit based on experience in accordance with the OPM Rating Guide

The results of the statistical analysis of each of the five experience groups, their OKT scores and ATC training pass/fail status is provided in Table 33.

Table 33

Failure Rates by OKT Score Ranges and OKT Means by Experience Groups

Score	Other Total	Other % Fail	Pilot Total	Pilot % Fail	VFR Total	VFR % Fail	IFR Total	IFR % Fail	VFR + IFR Total	VFR + IFR % Fail	TOTAL Total	TOTAL % Fail
0-59	430	29.3%	67	26.9%	25	32.0%	11	27.3%	23	43.5%	556	29.7%
60-64	37	27.0%	47	23.4%	19	42.1%	9	33.3%	23	34.8%	135	29.6%
65-69	43	25.0%	55	9.1%	23	29.6%	13	23.1%	46	23.9%	179	20.1%
70-74	34	20.6%	77	13.0%	31	19.4%	13	10.7%	61	16.4%	226	16.4%
75-79	22	9.1%	67	10.4%	37	13.5%	34	11.8%	117	14.5%	277	12.6%
80+	25	12.0%	76	10.5%	53	13.2%	59	5.1%	241	7.1%	454	9.4%
TOTAL	592	26.9%	389	15.2%	191	21.5%	144	13.2%	511	14.3%	1827	19.2%

Means on OKT	51.49	69.94	72.25	76.28	78.0	66.96
std. dev.	14.8	10.9	11.4	9.5	9.7	16.3

While the mean OKT score for the total 1827 ATC trainees is 66.96 (SD = 16.3), examination of Table 33 shows marked differences for the five experience groups and, within each group by OKT score range as shown below.

Experience Group	N	Mean	SD	Percent of Failure
No + Other Experience	592	51.49	14.8	26.9%
VFR + Any other (except IFR)	191	72.25	11.4	21.5%
Pilot + Other Experience	389	69.94	10.9	15.2%
VFR & IFR + Any other	511	78.00	9.7	14.3%
IFR + Any other (except VFR)	144	76.28	9.5	13.2%

Of the total 1827 ATC trainees, 359 (19.6%) were in ATC related experience groups which presently earn extra credit under the CSC Rating Guide for this experience and who scored less than 70 on the OKT. The failure rate for the 359 was 26.2% as compared to 28.2% for the 511 trainees with no (or other) ATC related experience who also scored less than 70 on OKT.

As can be seen from Table 33, as OKT score ranges get higher, the proportion of failures generally continues to drop in each of the five experience groups. Rather than using the present method for granting extra credit for experience, the use of OKT scores would be significantly more predictive of ATC training success.

Thus far, OKT data reported has compared the OKT test scores to the present method of crediting ATC related experience and each of these in terms of their prediction of ATC training pass/fail status. For OKT to be truely useful, it must contribute to predicting ATC training success over and above what can be achieved by a new selection test battery.

53

As previously shown, a test battery comprised of CSC 24, CSC 157 and MCAT had a multiple R of .54 which accounted for 29% of the variance on ZLAB. If the correlations of OKT with ZLAB and the other test battery scores are corrected for restriction in range using Thorndike's formula 7, then the multiple regression coefficient including OKT can be estimated. The new Multiple R, calculated including OKT as well as the three tests previously included in the battery would be .60 and 36% of the ZLAB variance would be accounted for. This increase is significant ($F = 204.3$ $p < .001$), indicating that the addition of extra points using OKT scores would significantly improve prediction and that OKT should be used as a basis for determining extra credit. The analysis of OKT summarized here is discussed in more depth in reference [27] .

DEVELOPMENT OF THE MULTIPLEX CONTROLLER APTITUDE TEST

EXPERIMENTAL TEST FORMS. Development of the Multiplex Controller Aptitude Test was based, in part, on previous research efforts directed to developing performance indices for air traffic controllers during 1970-1971 (28). From this research, the initial versions of the Controller Decision Evaluation (CODE) film tests were developed and used in the 1972 study of Air Traffic Control Specialists conducted by Education and Public Affairs (12).

The CODE test used a motion picture film to present simulated air traffic in real time as they cross a controller's display scope, together with a mileage scale, lines to represent airways and intersections, and a data table to provide information on each aircraft's identity, route, speed and altitude. The test was administered in a free response mode, with the task being to predict violations of aircraft separation standards (conflicts) as early and accurately as possible.

Experimental administration of three motion picture films were carried out. Analysis revealed that most of what was measured in each 45 minute film could be derived from only 7 or 8 of the pairs of conflicting aircraft. In addition, there was considerable idle time during testing which offered the opportunity to measure more than simply conflict performance. Consequently, test development procedures were initiated to measure not only this skill but others as well.

An initial modification was to develop a structured test, using a slide projector to place conflict-type questions on the screen. However, when items were assembled into a film/slide version, half of the test time remained idle and this provided the opportunity for test subjects to change their answers as the test progressed and the correct answers became obvious. It was found that by presenting a new item every 45 seconds the examinees were kept sufficiently busy so that this situation could be controlled.

Ample information is available on the scope and table to ask a variety of questions. Items were written that utilized the available detailed information to measure such aptitudes as direction following, table reading, spatial visualization, and arithmetic reasoning regarding separation distances and times. Some items were very simple. Others were written in a multi-factor format to increase their difficulty. For example, initial instructional questions on how to read the table were very easy, but a complex problem such as estimating travel time in minutes between two aircraft (horizontal separation) required awareness of distances across the scope, reading the table to determine their speed, and mathematical computation to determine their rate of closure.

Items were administered experimentally, the relation between item types and total test homogeniety determined, and this ratio used to determine the number of items per type to include in the test. A result, for example, was inclusion

of twice as many time-distance separation items as compass heading items in the test. Conflict item locations had been established by mean journeyman response times when targets were at certain locations. Aptitude items were placed in the remaining positions, 45 seconds apart, alternating from one type to another and spiraling to increasing levels of difficulty as testing progressed. About 40 percent of the items were conflicts, and the other 60 percent were aptitudes. As films 4, 6 and 7 each provided a different pattern of air traffic, three different forms of the test, MCAT 4, 6 and 7 were prepared.

Versions of the CODE test, the film/slide version of MCAT, and other selection measures were administered to students of the Navy Air Traffic School, Memphis, and grades of these students obtained as they progressed through training. Analyses included intercorrelations and rotated factor loadings of the various tests with each other and with such school grades as laboratory flight plans, control tower, and radar control problems, as these grades provide the greatest range in criterion scores for predicting FAA student scores. The film/slide structured version had correlations with criteria that were as high or higher than those of the free response CODE. These results encouraged further development of MCAT.

A next step in test development was to create a "slides only" version, in essence taking pictures of the screen each time a new question appeared. Pacing of the slide presentation was the same as in the film plus slide version, so the amount of target movement, how far an aircraft moved from question to question, was unchanged. Free response, film/slide and all-slide versions were administered to students at the Air Force Controller School, Keesler, and grades of the students obtained as they progressed through training. The all-slide version generally showed higher correlations with the criteria than either the free response or film/slide versions.

During these test administrations, it became evident that the use of motion picture or slide projector equipment would not be feasible for operational test administration. The effect of equipment failure, room size, lighting, seating arrangement, and other factors could not be controlled at the many locations where the Office of Personnel Management administers the test to applicants.

A next step was to print the slide versions in paper-pencil format. Three forms were prepared and administered to several populations including: entering ATC students at the FAA Academy during all of 1976; ATC Specialists participating in the 1977 study conducted by Education and Public Affairs, Inc., under FAA contract; and, 7000 ATC applicants tested by the CSC in tha fall of 1976 and spring of 1977.

Each form was then lengthened to 55 items and administered to all students at the FAA Academy from 1977 until June 1978; students at various colleges and universities; and, students at the Army Air Traffic Controller School, Ft. Rucker. All three forms were administered to each of the Army students and differences in mean performances among forms were found that could effect form comparability. Also, student performances improved considerably on the second form they were administered, indicating that learning was still taking

56

place, but there was very little increase in performance on the third form. Differences in learning rate among students was noted, and this ability was recognized to be another potentially useful measure.

Development of various forms of MCAT has resulted in reliability correlations that vary with the restriction in range of the scores for the populations tested. Correlations between early forms were .31 to .50 for ATC Academy students who had greater restrictions in range of their scores; .60 to .66 for USAF ATC students; and, .87 to .90 for non-controller FAA employees who more nearly approximate the population for whom the test was designed. Validity of the MCAT test with the different criteria used for prediction of "ATC success" have been identified in the various studies previously discussed.

A more detailed description of the development of experimental forms of the MCAT test is found elsewhere (19).

The test is designed to measure the skills of the air traffic controller applicant within an ATC simulated setting. What is accomplished, in effect, is to teach persons a set of simplified ATC rules, and then test how well they can apply those rules in specific situations. Test subjects spend eight minutes studying the directions, and then start to answer the questions. Because the air traffic controller activity is complicated and rather difficult to learn, the test is designed so that initial items are very easy, and then progress in an order of increasing difficulty. The time limits are deliberately established to make the test a speeded test. While usual aptitude battery design practice is to cluster items into homogeneous subgroups, in this test the content of the items is alternated from one type to another and they spiral to increasing levels of difficulty. This mode of presentation is found in only a few tests such as the Stanford-Binet. A result shown statistically is that the non-confliction items show unexpectedly high commonality with those items requiring detection of impending conflicts. Apparently the detection of conflicts requires the same skills and operations measured by the different kinds of aptitude items presented.

Figure 3 provides the basic format of the MCAT. About 40 percent of the total 55 questions ask for identification of potential conflicts, or violations of separation standards. The remaining questions ask for such information as differences in the routes of pairs of aircraft, how far apart two aircraft may be at a given moment, their compass headings, and if they will come into conflict.

DEVELOPMENT OF PARALLEL MCAT TEST FORMS. For operational purposes, multiple forms of a test are necessary for retest purposes and control over compromise. Based on previous item analysis and experimental test form development, each of the three experimental test forms (407A, 607A and 707A) were split into two parts by placing the odd items in one part and the even items into a second part. These individual parts were then placed in various combinations to provide six available forms. Each form was a 55 item test and separated into two parts.

57

FIGURE 3

THE MULTIPLEX CONTROLLER APTITUDE TEST

	Flight Information		
Aircraft	Altitude	Speed	Route
10	5000	300	AHC
20	5500	300	EIHB
30	5000	450	DIJF
40	6000	450	BHJF
50	6000	300	GJHB
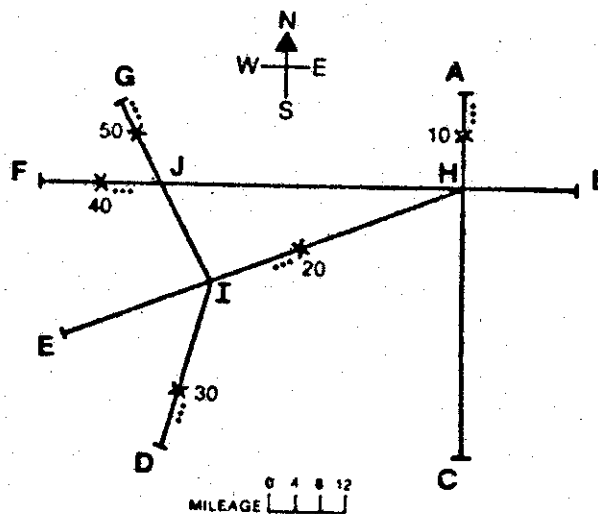
A typical example of the kind of information provided in the test problems is given above. The drawing shows the particular flight paths that aircraft must follow. Changes in routes can occur only at the intersection between two routes. Each x depicted on the routes represents an aircraft traveling in a particular direction indicated by the trailing dots.

To the side, a table containing critical flight information about each aircraft is provided. The number next to each x on the drawing allows the matching of the flight information to the correct aircraft. The flight information lists the altitude, speed in miles per hour, and route that each aircraft is flying.

Each part required a person to learn its unique mix and pattern of traffic and make predictions as to traffic behavior and potential conflictions. Part 1 was administered allowing 20 minutes for completion; 15 minutes was allowed for Part 2. The six forms were then administered to two sample populations; new ATC hires entering the FAA Academy for training after June 20, 1978, and to approximately 7800 applicants for ATC work who took the present ATC test administered by OPM in October-November 1978. Table 34 provides the descriptive statistics for each group for each form of the MCAT test.

TABLE 34

PERFORMANCES ON ALTERNATE FORMS OF MCAT

FORM NO.	ID	FAA ACADEMY STUDENTS 1978			OPM ATC APPLICANTS TESTED 1978		
MCAT	OPM	Mean	SD	N	Mean	SD	N
4o6e	120	33.8	7.45	459	29.3	9.15	1233
4e6o	130	37.4	7.35	458	29.3	9.70	2328
6o7e	140	40.4	7.45	434	33.1	10.40	1222
6e7o	150	37.6	6.80	319	31.5	8.20	1086
7o4e	160	33.8	7.65	135	29.6	8.70	1135
7e4o	170	39.8	7.25	356	32.2	10.34	882

The test is for use with OPM ATCS applicants, and those are the data of primary interest. The FAA Academy students data are included for future test development work, to estimate applicant scores on new forms of the test when experimental data are gathered on FAA Academy students. While all forms are quite comparable for the ATC applicant sample, item revisions will be made in forms 6o7e and 7e4o to increase their comparability.

Many different scoring methods have been explored during development of the test. The test could be scored for Rights and Wrongs, with separate scores for Conflicts and each of the Aptitudes, and with differential combining weights for all these subscores. The test could be scored for the rate of improvement from its early to later parts. Various multiple regression studies indicate that little is gained from maintaining separate scores for Conflicts and Aptitudes as was done in the 1977 study conducted by Education and Public Affairs, Inc. Consequently, performance on all items have been combined into a single score in subsequent analyses. The data also showed that little was gained through use of a combination of "Rights and Wrongs" instead of "Rights" only.

For a group of 617 ATC students at the FAA Academy tested between June and December, 1978, a total "Rights" score test-retest correlation of .60 between comparable two part 35 minute forms of MCAT was obtained. A planned doubling of the test length providing a total test time of 90 minutes should increase the reliability coefficient to .75 for this restricted group and provide an adequate level of reliability for the ATC applicant population.

The following summarizes the development of the MCAT test.

- CODE - film versions 4, 6 and 7. Scoring: 1971, 40 items; 1974, Stanine: 1975. corrected item scores - best items. Unstructured free response.

- Film Slide MCAT - July 1975. Time controlled for answering each item. MCAT 406 FS - 41 items, 606 FS - 43 items, 706 FS - 53 items. Structured multiple choice.

- All slide MCAT - October 1975. Time controlled for answering each item. MCAT 406 AS - 41 items, 606 AS - 43 items, 706 AS - 53 items. Structured multiple choice.

- Paper Pencil MCAT - January 1976. MCAT 406A - 41 items, 606A - 43 items (25 minutes test time) 706A - 53 items (30 minutes test time). Structured multiple choice.

- Paper Pencil MCAT - April 1976. Time controlled for answering each five minute group of items. Form 606B - 43 items, 706B - 53 items. Structured multiple choice.

- Paper Pencil MCAT - January 1977. Time controlled for working total test (35 minutes). 407A, 607A, 707A each 55 items. Structured multiple choice.

- Paper Pencil MCAT - June 1978. Parallel test forms. Two part tests with each part presenting a unique pattern of air traffic. Part 1 - 20 minutes, Part 2 - 15 minutes. Forms: 4o6e; 4e6o; 6e7o; 6o7e; 7e4o; 7o4e: each test 55 items. Structured multiple choice.

DEVELOPMENT OF THE AIR TRAFFIC CONTROLLER OCCUPATIONAL KNOWLEDGE TEST

The ATC Occupational Knowledge Test (OKT) is a subject element, paper-pencil, machine scorable, multiple choice type test. It measures important aspects of knowledge associated with prior experience that are related to successful performance as an air traffic controller specialist. The subject matter areas included are Air Traffic Control Regulations, Terminal Air Traffic Control, Enroute Air Traffic Control, Communications, Flight Service Stations, Navigational Aids, Weather, and Radar.

Experimental Test Development

The initial ATC Occupational Knowledge Test was designed to measure the acceptability of an applicant's claimed experience for qualifying at the GS-9 level or above. There was such a wide variety among applicant's claims of experience that they were difficult to evaluate, and the OKT would provide scores to verify the acceptability of claimed experience. The first step in test development was a two-week workshop in June 1970, in which eighteen journeyman air traffic control specialists from Terminals, Centers and Flight Service Stations participated. The workshop was to teach these people the specifications for the test and help them write items under the training and supervision of representatives of FAA Headquarters. This committee examined the exact types of qualifying experience that were given credit, and determined the specifications for the test to permit its application to verify claimed experience. Then over 300 items were written, and reviewed and edited by the Examination and Certification Section of the ATC Academy, Oklahoma City. Two test booklets were created, Form 15 containing 150 of the items, and Form 16 containing 160 items.

These forms were administered to two classes of students on the day they reported for training at the ATC Academy, and to samples of GS-7 and GS-9 Air Traffic Control Specialists in Flight Service Stations, Terminals, and Centers. The tests were found to correlate highly with past experience patterns, had acceptable validity coefficients with training success for Center, Terminal, and Flight Service Station training, and demonstrated ability to differentiate Air Traffic Controller from Flight Service Station job incumbents. The results of the experimental testing were analyzed at the individual item level, and these data used for creating multiple forms of the test.

Since planned initial application was a test for use as a partial basis for determining applicants' qualification for Terminal and Center job entrance above the GS-7 level, a first step was to eliminate items from consideration that had been found lacking by experts in the specialty area. The performance of virtuoso air traffic controllers was compared to performance of entering ATC Academy students on an item by item basis. Prime criteria for an item's selection were (1) its ability to differentiate between the two groups and (2) the item's difficulty level. The easiest 100 of the most discriminating items were selected for inclusion in Form 101 - Experimental and the next most difficult 100 items were assembled as Form 201.

A more detailed discussion of the initial development of the Occupational Knowledge Test is found elsewhere (30). Subsequent to 1970, FAA discontinued hiring ATC applicants at grades above GS-7 and work on developing the ATC Occupational Knowledge (OKT) was deferred until 1975, when further research was initiated to improve ATC selection.

In August, 1975, a battery of experimental selection tests, including Forms 101 and 201 of the ATC Occupational Knowledge Test, were administered to 109 students of the Navy Air Traffic School, Memphis. Student scores on Form 101 were distributed against their week in training when tested, and these showed increases in test performance levels as the students gained experience. Grades of these students were obtained as they progressed through training. Analyses included intercorrelations and rotated factor loadings of the various tests with each other and with school grades such as academic and laboratory grades and course average. Distribution statistics plus correlations with scores on selected variables are presented in Table 35. Mean scores on Form 201 are ten points lower than on Form 101, and the Form 201 total score correlations with other measures are similarly depressed.

TABLE 35
PERFORMANCE RELATIONS BETWEEN OKT AND SELECTED VARIABLES

Navy ATC Students August 1975	OKT 101		OKT 201	
	Mean	SD	Mean	SD
TOTAL GROUP (N-109)	61.6	8.9	51.5	6.7
CORRELATION WITH:	r		r	
Reading	-.02		-.01	
Education Level	.27		.10	
MCAT	.42		.22	
OKT 101	---		.16	
OKT 201	.16		---	
School Grades				
Academic Average	.48		.25	
Block II Composite	.28		.16	
Block III: Basic Lab	.28		.06	
Block III: Advanced Lab	.27		.13	
Course Average	.43		.25	

In the factor analysis, student performance on the ATC Occupational Knowledge Test (OKT) showed little or no communality with their ability to read and comprehend the material and very little relation with their mechanical-spatial aptitudes. OKT-101 had high loadings on the Air Traffic Controller Performance factor, the Control Tower Operator factor, and the Laboratory factor. OKT-201 had a similar factor structure but with less loading on each of the factors, perhaps as a function of restriction in variance, since it was a rather difficult test for this population.

62

A battery of experimental tests including ATC Occupational Knowledge Test
101 Experimental was also administered to forty-one students at the Air
Force Air Traffic Controller School, Keesler AFB, during October 1975.
Distribution statistics and correlations with selected school grades
are presented in Table 36. The performance of these students were quite
similar to those of the Navy Air Traffic Controller School sample.

TABLE 36
PERFORMANCE RELATIONS BETWEEN OKT AND SELECTED VARIABLES

USAF ATC STUDENTS October 1975	OKT 101		
		Mean	Standard Deviation
Total Group (N=41)		63.66	16.34
School Grades, Final School Correlations			r
Block I Basic FAA Certification			.42
Block II Basic Operations Specialist			.15
Block III Control Tower Operator			.32
Block IV Radar Controller (6CA/PAR)			.68

Based on the results of administering Form 101 and 201 of the Occupational
Knowledge Test to these groups, as well as to controllers at selected FAA
terminal and enroute facilities in December 1975, the best 100 items were
selected to create Form 101B of the OKT. This form was part of the experi-
mental test battery administered to FAA facility personnel and ATC Academy
students in the 1977 selection study conducted by Education and Public
Affairs, Inc (6). Distribution statistics and correlations of OKT-101B
with Supervisory Assessments obtained in the 1977 study for field personnel
are given in Table 37. Correlation with supervisory assessments for field
ATC personnel were among the highest found for variables in the experimental
test battery.

TABLE 37
DISTRIBUTION STATISTICS AND CORRELATIONS
ATC OCCUPATIONAL KNOWLEDGE TEST - 101B

(Facility ATC Specialists - 100 items, 97 keyed).

ATC Option and Year Hired	Correlation with Supervisory Assessment	N	Mean	SD
IFR 1969-70	.13	97	85.1	5.02
1973-74	.14	92	85.0	6.45
TOTAL	.13	189	85.0	5.75
VFR 1969-70	.13	69	82.0	9.71
1973-74	.29	100	80.6	8.39
TOTAL	.21	169	81.2	8.97
ARTCC 1969-70	.21	88	79.8	6.71
1973-74	.14	94	77.0	9.33
TOTAL	.17	182	78.4	8.26
FSS 1969-70	.01	99	75.0	9.65
1973-74	.34	97	71.5	13.04
TOTAL	.19	196	73.3	11.55
ALL OPTIONS 1969-70	.08	353	80.4	8.78
1973-74	.20	283	78.5	10.78
TOTAL	.15	736	79.4	9.91

63

Comparable data for a group of new ATC trainees entering the FAA Academy in 1976 are provided in Table 38.

TABLE 38
DISTRIBUTION STATISTICS AND CORRELATION:
ATC OCCUPATIONAL KNOWLEDGE TEST - 101B

(Academy ATC Trainees - 100 items, 97 keyed)

Academy Classes 4-8	Correlation with Z Lab Score	N	MEAN	SD
Jan - June 1976				
Total Terminal and				
Enroute by Subgroups		803	67.3	16.42
Aviation experience	.40	349	76.6	10.88
No aviation experience	.27	452	60.2	16.23
Men	.31	700	68.7	16.06
Women	.11	103	57.5	15.62
Minority	.36	86	67.5	16.51
Non-Minority	.29	710	68.7	16.06
Terminal	.31	372	70.9	15.44
Enroute	.32	431	64.2	16.33

As would be expected, mean scores on OKT for this trainee group were lower and the standard deviations higher than for the more experienced facility ATC specialists hired in 1969-70 or 1973-74. Correlations with the non-radar laboratory training scores (Z scores) for the Academy were highest for those with prior aviation experience and for minorities. Correlations for the terminal and enroute option were also among the highest found for variables in the experimental test battery.

The ATC OKT 101B was also evaluated in relation to present methods for granting extra credit for aviation related experience. The ATC Occupational Knowledge Test 101B was administered to 784 ATC trainees who entered the FAA Academy's 16-week ATC training program between July and December, 1976. All trainees completed a pre-employment questionnaire. Based on responses to the questionnaire, the trainees were assigned to one of three experience groups in accordance with evaluations of claimed experience as made by the Office of Personnel Management rating procedures. It was found that while scores on the OKT were highly correlated with claimed experience (.64), OKT had a higher correlation with successful completion of the non-radar lab (.25) than did claimed experience (.12). It was determined that use of an OKT score of 75 or above to assign extra credit would result in a failure rate of 3.1 percent for those receiving credit, while use of the present experience rating would result in a failure rate of 7.6 percent for new hires who now receive extra credit for experience. The results held up for a cross-validation sample of 432 trainees who entered the Academy during 1977. The full results of this analysis of OKT is provided elsewhere

In order to obtain data on the ATC applicant groups within available testing time, a 60 item form of OKT (101-C) was developed from form 101B. This was administered to two groups of ATC applicants during September-November 1978, together with parallel forms of MCAT after they had completed the existing operational ATC test battery used by OPM. One group consisted of 5331 scheduled applicants. The second group was 669 "walk in" applicants

in testing sessions arranged to encourage women and minorities to apply for
ATC work. Mean scores for the scheduled applicants on the 60-item OKT (101C)
was 24.9 with a standard deviation of 11.80 and for the "walk in" group, 20.1
and 9.71 respectively. OKT Form 101C was also administered to groups of ATC
trainees at the ATC Academy during 1978. From these analyses it was con-
cluded that the ATC occupational Knowledge Test Form 101B has high reliability
(.91) and that the slight loss of reliability with reduction in test length
would not be of consequence.

Development of Parallel Test Forms

The purpose of the OKT is to define the domain of ATC Knowledge that is
demonstrated by applicants for the occupation as a basis for granting extra
credit to those who successfully pass the basic ATC selection tests admini-
stered by OPM. For operational usage, multiple forms are needed to meet
retest requirements and provide some capability for control over compromise.
An essential step in the development of alternate forms was to define the
types of items and number of items of each type to include in each succeeding
form, to structure the domain of knowledge to make it consistent with the
knowledge requirements of air traffic control. This was accomplished pri-
marily by determining what types of knowledge were retained after training
and early apprenticeships experience and were known by nearly all full per-
formance level controllers. Table 40 identifies the steps involved in deter-
mining the final domain of ATC occupational knowledge included in OKT and
provide a basis for developing future comparable forms.

As the Federal Aviation Regulations state that, to become certified as a
Terminal ATCS, the person must pass a written knowledge test on (1) Flight
Rules in FAR Part 91 (Air Traffic Control Rules), (2) Terminal Traffic
Control Procedures, (3) EnRoute Traffic Control Procedures, (4) Communica-
tions Operating Procedures, (5) Flight Assistance Service, (6) Air Navigation
and Aids to Navigation, and (7) Aviation Weather, the FAA Basic Certification
Test was designed to test knowledge in each of those subject matter areas.
The original form of the certification test is listed first in Table 38, and
the table shows that an equal percentage of items (14 percent) were allocated
to each subject matter area. For development of the OKT, radar questions were
added to the item pool (line 2), and the best items selected (lines 3-6) to
measure candiates' possession of the knowledge required. The percentage
allocation per subject matter area is presented in line 7 for version
101B of the OKT. Each of the seven subject matter areas is further defined
by subject sub-elements which are identified in Table 40. Table 41 shows
the classification of items in Form 101B by element and sub-element. Numbers
in the cells are item number. If the item number is underlined, that item
was found to be known by 90 percent or more of the 1975 sample of 50 full-
performance level Air Traffic Controllers.

65

TABLE 39

ATC OCCUPATIONAL KNOWLEDGE TEST - SUBJECT MATTER ELEMENTS
PERCENT OF ITEMS - EACH ELEMENT

Form of OK Test	ATC Rules	Airport Traffic Procedures	Inflight Traffic Control Procedures	Communications Operations Procedures	Flight Assistance Service Procedures	Air Navigation & Aids to Navigation	Aviation Weather	Number of Items
1. Original FAA Basic Certification Test All Items	14%	14%	14%	14%	14%	14%	14%	140
2. Form 16 All Items	12½	12½	25	12½	12½	12½	12½	160
3. Form 101 - Exp. All Items	20	12	28	12	2	19	6	99
4. Form 101 - Exp. Items known by 90+% of Group of Full Performance Level (N=50, 1975)	17½	9	32	12	3.5	21	5	57
5. Form 101 - Exp. Items known by 100% of Group of Full Performance Level ATC (N=16, 1970)	18	9	28	13	2	25	4	53
6. Form 101B All Items	21	12	28	12	2	19	6	100
7. New Operational Forms (Recommended)	18	10	29	13	4	21	5	100
Sub-element: 1	1	*	10	9	*	*	*	
2	*	2	3	2	*	5	*	
3	7	2	*	1	*	*	2	
4	8	1	15	1	2	11	*	
5	1	1	1	*	*	*	2	
6	*	3		1	*	0	0	
7	1	1		*	2		1	

TABLE 40

CATEGORIES (ELEMENTS AND SUB-ELEMENTS) OF
ATC OCCUPATIONAL KNOWLEDGE

1. AIR TRAFFIC RULES

01. Pertaining to Aircraft
02. Pertaining to Airway
03. Pertaining to Airspace
04. Pertaining to Flying Conditions (Ratings, Weather, Speed Maneuvers, Altitude, Flight Level)
05. Pertaining to Operations/Instructions
06. Pertaining to Administrative Messages (Reports, Phraseology)
07. Pertaining to Safety/Emergency

2. AIRPORT TRAFFIC PROCEDURES

01. Pertaining to Traffic Information
02. Pertaining to Traffic Patterns
03. Pertaining to Runway Operations
04. Pertaining to Flight Movement (Destination Changes, Arrivals, Departments)
05. Pertaining to Administrative Messages (Phraseology, Authorization, Official Reports, Documents)
06. Pertaining to Meterological Information (Ceiling, Visibility)
07. Pertaining to Safety (Messages, Warning Devices, Collisions)

3. INFLIGHT TRAFFIC CONTROL PROCEDURES

01. Pertaining to Separation (Vertical, Lateral, etc.)
02. Pertaining to Clearance (Routing)
03. Pertaining to Approach (Flight Level, Holding, Alt. Setting)
04. Pertaining to Radar (Contact, Ident, Interference, Traffic Information)
05. Pertaining to Flight Information and Data

4. COMMUNICATIONS OPERATING PROCEDURES

01. Pertaining to Flight Movement and Control Messages
02. Pertaining to Administrative Messages (Phraseology)
03. Pertaining to Meterology Information
04. Pertaining to Clearance
05. Pertaining to Safety
06. Pertaining to Service/Maintenance

5. FLIGHT ASSISTANCE SERVICE PROCEDURES

01. Pertaining to Search and Rescue
02. Pertaining to Overdue Aircraft
03. Pertaining to Emergency Communications
04. Pertaining to Meterological Information
05. Pertaining to Facilities
06. Pertaining to Procedures

6. AIR NAVIGATION AND AIDS TO NAVIGATION

01. Pertaining to Frequencies
02. Pertaining to Instrumentation
03. Pertaining to Position/Distance/Direction
04. Pertaining to Navigation (Time, etc.)
05. Pertaining to Navigational Aids

7. AVIATION WEATHER

01. Pertaining to Visibility/Ceiling
02. Pertaining to Weather Conditions (Temperature, Winds, Tornado, etc.)
03. Pertaining to Clouds
04. Pertaining to Administrative Messages (Phraseology)
05. Pertaining to Forecasts/Reports
06. Pertaining to Responsibilities/Procedures

TABLE 41

OKT 101B
ITEMS CODED BY ELEMENTS AND SUB-ELEMENTS
(Underlined Item Numbers Mean 90% + Right for FPL C&L -1975)

Sub Element	I ATC Rules	II Airport Traffic Procedures	III-Inflight Traffic Control Procedures	IV Communications Operating Procedures	V-Flight Assistance Service Procedures	VI-Air Navigation & Aide to Navigation	VII Aviation Weather
1	45,		#1 3, 4, 10 25, 70, 71, 72 74, 99	17, 18, 21, 23, 32 75		19, 43, 44 81, 84	
2		47, 64	6, 40	14, 16, 41, 77			50, 86, 89
3	5, 38 52, 54, 55, 68	48, 67		15,	78,	13, 27, 28, 29, 30, 33, 34, 79 80, 82.	
4	7, 9, 37 51, 53 56, 73, 76	8, 42, 62	12, 20, 22, 26, 36, 46, 90, 91 92, 93, 94, 95 96, 100				87, 88
5	39, 97	69	60, 98	24		11, 31, 49, 83	
6		59, 61, 66			65		85
7	2, 35 67, 63	58					
	21 items	12 items	28 items	12 items	2 items	19 items	6 items 100

The subject matter areas or elements and percentage allocations per area as determined above were used to describe the types and number of items per type to be included in new alternate forms of the test. Fourteen hundred multiple choice items were drafted and reviewed, and items selected from this pool to create eight 100-item parallel forms of the test. Twenty items were common to all forms. These forms were administered experimentally, item analysis accomplished, and the eighty best items in each form identified for keying. These eighty item forms are identified as ATC Occupational Knowledge Test 102A, 102B, 102C, 102E, 102F, 102G, and 102H. Use of a Rights Only scoring procedure has been practiced and is recommended.

The forms have been administered to incoming students on the first day of training at the ATC Academy since November 21, 1978. The mix of trainees attending the Academy varied during this period with regard to levels and types of aviation-related experience. In order to control these differences, each group of students was given two forms of the Occupational Knowledge test. Consequently, the comparison of results for alternate forms of OKT can only be made for those tests taken by the same group of trainees.

The descriptive statistics obtained for each trainee group which took two different forms of the test are provided in Table 42.

TABLE 42

ALTERNATE FORMS OF OKT

	(N=97)			(N=22)	
	Mean	SD		Mean	SD
Form A	41.95	14.05			
Form C	42.55	11.95			

	(N=67)			(N=22)	
	Mean	SD		Mean	SD
Form A	38.91	14.61		49.59	13.35
Form B	36.25	12.15		44.73	10.84

	(N=86)	
	Mean	SD
Form B	38.35	14.80
Form D	43.45	12.75

	(N=384-386)			(N=69-70)			(N=22)	
	Mean	SD		Mean	SD		Mean	SD
Form E	46.49	13.80		45.51	13.98		48.18	11.66
Form F	44.34	13.40		42.60	13.13		44.50	12.10

	(N=407-408	
	Mean	SD
Form G	46.28	13.95
Form H	48.38	13.95

The means and standard deviations for the two forms taken by each group can
be compared since the experience mix for that group was common across both
forms taken. However, the data cannot be compared for the same form taken by
different groups of trainees since the experience mix was not common.

The alternate forms do show some differences which require minor adjustments
by shifting items of a common type but ranging in difficulty level,from one form
to another.

Reliability estimates for the various forms have been derived and are presented
in Table 43. These reliabilities are considered satisfactory for operational
testing use and administration.

TABLE 43

RELIABILITY OF ATC OCCUPATIONAL KNOWLEDGE TEST

Testing Date	Student Groups	OKT 102 Taken First vs. Second		Correlation
3-6-79	Enroute N=66	E,F,G,H (80 item)	E,F,G,H (80 item)	.904
7-3-79	Enroute, Terminal FSS, N=89	E,F,G,H (80 item)	E,F,G,H (80 item)	.893
11-21-79	Enroute, Terminal N=183	A,B,C,D (20 core items)	A,B,C,D (20 core items)	.877

Objectives. The results of the analysis of the experimental tests for a group of 1827 new ATC trainees at the FAA Academy during the period 1976-1978 supported the conclusion that a test battery comprised of the Multiplex Controller Aptitude Test, Abstract Reasoning (CSC-157), and Arithmetic Reasoning (CSC-24) would provide an operational selection test battery for the ATC occupation with significantly better prediction of success in ATC training.

This conclusion, together with the development of parallel test forms of MCAT made it desirable to administer these tests to a new group of ATC applicants in order to compare the results of the proposed test battery with the existing CSC tests with respect to the composition of the two groups who "passed" each test battery and the effect on score distribution within each group. A change in policy also made it possible to obtain race, sex and ethnic data which had not been possible with the 1976-1977 applicant group.

Arrangements were made between FAA and the Office of Personnel Management to administer the parallel forms of the MCAT and a 60-item version of the ATC Occupational Knowledge Test when the present CSC tests were administered during September-November 1978 to ATC job applicants.

Sample Description. The present CSC tests, together with MCAT and OKT were administered to a total of 6,000 ATC job applicants. Some 5,331 applicants were scheduled for examination through established OPM procedures. In addition, FAA had arranged with OPM for a series of "walk-in" test sessions which did not require advanced scheduling by applicants with OPM. A total of 669 applicants took the test on a "walk-in" basis. These "walk-in" sessions were established as a means of encouraging more women and minorities to compete for positions in the ATC occupation. Table 44 provides the distribution by race, ethnic groups and sex for 5295 of the 5331 "scheduled" ATC job applicants. Table 45 presents these data for 664 of the 669 "walk-in" applicants.

TABLE 44

	Amer. Indian	Asian	Black	White	Hisp.	Other	TOTAL
Men	42 (72%)	38 (76%)	650 (58%)	2860 (76%)	213 (79%)	21 (70%)	3824 (72%)
Women	16 (28)	12 (24%)	463 (42%)	915 (24%)	56 (21%)	9 (30%)	1471 (28%)
TOTAL	58 (1%)	50 (1%)	1113 (21%)	3775 (71%)	269 (5%)	30 (1%)	5295 (100%)

TABLE 45

	Amer Indian	Asian	Black	White	Hisp.	Other	TOTAL
Men	2	3	145 (50%)	160 (55%)	44 (65%)	-	354 (53%)
Women	1	4	146 (50%)	132 (45%)	24 (24%)	3	310 (47%)
TOTAL	3	7	291 (44%)	292 (44%)	68 (10%)	3	664 (100%)

71

There are several points of interest regarding the racial and sex distribution of these two applicant groups. For the "Scheduled" applicant group, (5295) 92 percent were black or white men and women, with 71 percent of the group represented by white men and women. Generally men represented between 70-79 percent of each racial group except for blacks; only 58 percent were men while 42 percent were women; almost twice the ratio of other groups.

For the "Walk-In" applicants the distribution indicates that the objective of having more minorities and women compete for the ATC occupation was fairly successful. Some 47 percent of the total group were women in contrast to 28 percent for the scheduled applicants and 55 percent were minorities in comparison to 29 percent of the scheduled applicants.

Predictors. Tests used in this study were the existing CSC selection test battery and parallel forms of MCAT and OKT:

OPM Test Battery

- CSC-24   Arithmetic Reasoning
- CSC-51   Spatial Relations
- CSC-135   Following Oral Direction
- CSC-157   Abstract Reasoning
- CSC-540   Air Traffic Problems

MCAT	OPM ID	ODD/EVEN TEST COMBINATION
Form	120	(4o6e)
Form	130	(4e6o)
Form	140	(6o7e)
Form	150	(6e7o)
Form	160	(7o4e)
Form	170	(7e4o)

OKT - 101c   (60-item version)

As in previous experimental testing of applicant groups, the MCAT and OKT test were administered after completion of the regular CSC test battery. In addition to the tests, applicants completed forms provided by OPM on a voluntary basis to obtain racial, sex, and other information for separate analysis by OPM.

Criterion and Analytical Methods. Since this study encompassed ATC applicants, it was not feasible to establish operational criterion measures of validity. However, a number of statistical analyses were made based on pass/fail eligibility comparison, score distributions and mean and standard deviations on the various tests.

72

Results. Descriptive statistics for the "scheduled" and "walk-in" groups were developed for each test. It should be pointed out that of the 260 "Scheduled" applicants who identified themselves as "Hispanic", all but 74 also identified themselves with another racial group. For example, 139 were identified as "Hispanic-White". In this analysis, such cases were included in the racial group. Consequently, 104 applicants in the "other" category shown in Table 43 were comprised of 74 Hispanic applicants and 30 of some other racial groups not identified. Descriptive statistics for the "Scheduled" group of applicants are provided in Table 46 and for the "Walk-In" group, in Table 47. (In Table 44, the 24 applicants in the "other" category consist of 21 "Hispanic" and 3 otherwise not identified.)

TABLE 46 SCHEDULED ATC APPLICANTS

	AMERICAN INDIAN (N= 66)		ASIAN (N= 59)		BLACK (N= 1146)		WHITE (N= 3914)		OTHER (N= 104)		TOTAL (N= 5331)	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
CSC-24	40.1	9.2	41.9	7.9	35.4	9.6	42.4	7.8	36.8	8.6	40.7	8.9
(Men)	-	-	-	-	(34.3)	(9.7)	(42.2)	(7.9)	-	-	-	-
(Women)	-	-	-	-	(37.0)	(9.2)	(43.2)	(7.6)	-	-	-	-
CS-51	28.6	4.8	28.3	5.8	23.1	6.5	29.1	5.2	27.0	6.5	27.7	6.1
(Men)	-	-	-	-	(24.3)	(6.4)	(29.6)	(5.0)	-	-	-	-
(Women)	-	-	-	-	(21.3)	(6.2)	(27.7)	(5.6)	-	-	-	-
CSC-135	21.6	9.0	19.9	8.2	16.6	8.1	24.7	7.0	20.1	8.5	22.7	8.1
(Men)	-	-	-	-	(16.7)	(8.0)	(24.8)	(6.9)	-	-	-	-
(Women)	-	-	-	-	(16.5)	(8.3)	(24.2)	(7.4)	-	-	-	-
CSC-157	28.0	8.2	31.0	10.9	22.1	9.3	30.9	8.9	26.1	9.0	28.8	9.7
(Men)	-	-	-	-	(21.5)	(9.3)	(30.5)	(8.9)	-	-	-	-
(Women)	-	-	-	-	(22.9)	(9.1)	(32.1)	(8.8)				
CSC-540	29.4	12.7	30.5	12.6	22.7	11.9	33.1	11.7	25.6	12.5	30.6	12.5
(Men)	-	-	-	-	(23.4)	(11.9)	(33.6)	(11.6)	-	-	-	-
(Women)	-	-	-	-	(21.8)	(11.9)	(31.3)	(11.9)	-	-	-	-
NCAT	31.1	8.6	31.0	8.6	22.1	7.4	34.2	8.1	28.0	8.4	31.4	9.5
(Men)	-	-	-	-	(22.9)	(7.8)	(35.0)	(8.0)	-	-	-	-
(Women)	-	-	-	-	(21.1)	(6.6)	(31.8)	(8.2)	-	-	-	-
OKT (Men)	24.3	10.8	24.9	11.3	17.4	8.7	27.2	11.7	21.8	11.8	24.9	11.8
	-	-	-	-	(16.8)	(9.9)	(28.8)	(12.0)	-	-	-	-
(Women)	-	-	-	-	(15.5)	(6.5)	(22.0)	(8.9)	-	-	-	-

TABLE 47 WALK-IN ATC APPLICANTS

	AMERICAN INDIAN (N= 5)		ASIAN (N= 10)		BLACK (N= 297)		WHITE (N= 328)		OTHER (N= 24)		TOTAL (N= 669)	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
CSC-24	60.0	4.2	48.0	9.5	37.5	9.7	42.3	7.8	39.4	9.6	40.1	9.1
CSC-51	29.0	4.0	31.8	4.2	23.0	6.2	28.6	5.0	28.8	4.2	26.1	6.1
CSC-135	19.0	7.2	24.6	7.3	16.7	7.4	23.7	6.8	19.7	7.6	20.4	7.9
CSC-157	30.0	11.6	34.6	12.2	25.0	9.4	32.0	8.9	28.7	11.6	29.8	9.9
CSC-540	36.4	8.4	45.2	12.1	23.6	12.7	32.4	12.5	30.4	13.4	28.7	13.5
NCAT	28.4	3.9	34.9	7.4	22.7	7.5	32.6	7.8	27.9	9.7	28.0	9.1
OKT	19.2	6.8	26.4	5.5	16.2	6.5	23.7	10.9	19.5	8.9	20.1	9.7

For the "scheduled" applicant group, the mean scores for women were examined for both the "Black" and "White" groups to assess their relationships to the scores of men in the same racial group. Table 43 indicates that women score somewhat higher on CSC tests 24 and 157 than men in each racial group, but somewhat lower on CSC 51, 540, MCAT, and OKT, with the greatest difference between men and women being on the OKT. The most marked difference in test scores between racial groups is for the "Black" and "White" groups, with the mean test scores for "Blacks" generally about one standard deviation below the mean for "Whites". This holds for both men and women in each of the groups.

Relatively few differences between mean scores of the "Scheduled" and "Walk-In" groups are evident by racial groups. Within the "Black" group, walk-in applicants scored about 2 points higher on CSC 24 and 3 points higher on CSC-157. The "White" group walk-in applicants scored almost 2 points lower on MCAT and about 3 points lower on OKT. These lower scores may be a result of the higher percentage of women in the "White" walk-in group (45 percent) compared to the scheduled group (24 percent).

With the applicant group tested in 1976-1977, two of the CSC Tests (51 Spatial Relations and 135 Following Oral directions) showed marked negative skew. It was of interest to determine if this characteristic was replicated for these two tests with the 1978 applicant population. Figure 4 provides a frequency distribution for test scores for these two CSC tests. It is evident that again they provide little differentiation between the applicants.

FIGURE 4
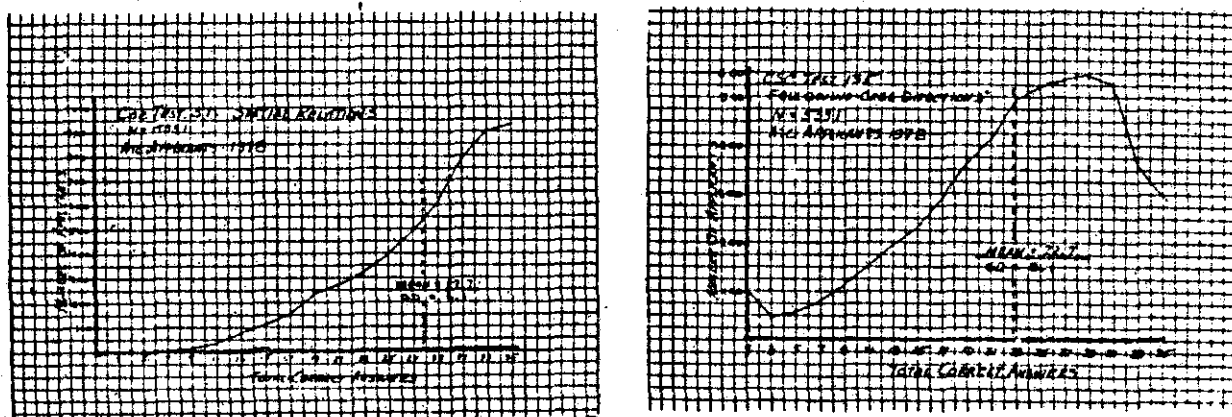
Veterans preference for the total group (scheduled and walk-in) was identified by sex and racial groups. Table 48 provides the distribution by sex for 5846 of the 6000 applicants.

TABLE 48

SEX DISTRIBUTION-1976-1977 APPLICANT GROUP

VET PREF	MEN	WOMEN	TOTAL
None	2406 (59%)	1640 (94%)	4046 (69%)
5 Pt.	1571 (38%)	97 } (6%)	1668 (29%)
10 + Pt.	123 (3%)	9 }	132 (2%)
TOTAL	4100 (100%)	1746 (100%)	5846 (100%)

Table 49 provides the distribution of Veterans Preference points by racial and ethnic group.

TABLE 49

VETERANS PREFERENCE DISTRIBUTIONS

VET. PREF.	AMERICAN INDIAN	ASIAN	BLACK	WHITE	HISP.	OTHER	TOTAL
None	42 (70%)	40 (70%)	951 (73%)	2724 (68%)	223 (68%)	27 (84%)	4037 (69%)
5PTS.	17 } (30%)	17 } (30%)	318 } (27%)	1210 } (32%)	99 } (32%)	5 } (16%)	1666 (29%)
10 + PTS.	1 }	- }	41 }	83 }	7 }	- }	132 (2%)
TOTAL	60	57	1340	4017	329	32	5835

Since Veterans Preference points are added to the test battery score for those who pass the test, it is evident that only 6 percent of the women could benefit, and thus be ranked higher on the OPM register for appointment eligibility. This in contrast to 41 percent of the men who could benefit from military service.

Eligibility for Veterans Preference points by racial and ethnic groups is about evenly distributed (30-32 percent) except for the "Black" group (27 percent) where the higher ratio of women applicants to men reduces the number who receive Veterans Preference points.

Table 50 compares the total ATC applicant group to the civilian labor force (CLF) by racial and ethnic group based on 1970 census data with amendments through 1977.

TABLE 50

ATC APPLICANTS—CIVILIAN LABOR FORCE

	Civilian Labor Force	% ATC Applicants
American Indian	.3%	1%
Asian	.8%	1%
Black	9.7%	24%
White	85.0%	68%
Hispanic	4.2%	6%

In terms of ATC applicants, blacks as a group are significantly over-represented in comparison to their percentage of the civilian labor force while whites as a group are underrepresented. As a group, women were also underrepresented; they comprised 28 percent of the ATC applicant group compared to 38 percent of the civilian labor force.

Analysis of current OPM Test and OKT

Next, the current OPM test battery scores were analyzed in terms of "pass/fail" score distribution by ethnic or racial group and by men and women. Because of the small number of Asian and American Indian applicants, analysis of minority groups has been directed to Blacks and Hispanics.

Table 51 provides the results of those who took the current OPM ATC test with respect to "pass/fail". The minimum passing score on the test battery is 70 which equates to the approximate mean score on the test where about half of the total applicants pass and half fail.

TABLE 51

PASS/FAIL—OPM ATC TEST

Group	Total Applicants	Pass OPM Test		Fail OPM Test	
		N	%	N	%
Men	4191	2236	(53%)	1955	(47%)
Women	1785	799	(45%)	986	(55%)
Total	5976	3035	(51%)	2941	(49%)

76

Compared to the 7500 applicants tested in 1976-1977, this group had about the same pass rate for men (53 percent compared to 52 percent) and a somewhat higher pass ratio for women (45 percent compared to 41 percent).

Table 52 shows the "pass/fail" rates by racial group.

TABLE 52

PASS/FAIL BY RACIAL GROUP

Group	Total Applicants	Pass N	%	Fail N	%
White	4067	2556	(63%)	1511	(37%)
Hispanic	339	128	(38%)	211	(62%)
Black	1407	264	(19%)	1143	(81%)
Total	5813	2948	(51%)	2865	(49%)

In view of the disproportionate number of minorities who did not pass the OPM test, an analysis of score distribution was made by racial group (Table 53).

TABLE 53

OPM TEST SCORE DISTRIBUTION BY RACIAL GROUP

Group	Fail OPM Test (N = 2865) Below 50 N	%	50-59 N	%	60-69 N	%	Pass OPM Test (N = 2948) 70-79 N	%	80-89 N	%	Above 89 N	%	Total
White	185	(4%)	406	(10%)	920	(23%)	1128	(28%)	975	(24%)	453	(11%)	4067
Hispanic	63	(18%)	56	(17%)	92	(27%)	70	(21%)	41	(12%)	17	(5%)	339
Black	457	(32%)	381	(27%)	305	(22%)	189	(13%)	63	(5%)	12	(1%)	1407

It is evident from this analysis that for the Black group in particular, the higher fail rate is comprised of a large number of applicants who did very poorly on the test battery. The 457 Blacks who scored below 50 on the test comprised 40 percent of the total number who failed to pass the test in contrast to about 30 percent for Hispanic and 12 percent for Whites.

77

Figure 5 provides a graph of the OPM ATC test score distribution for selected racial groups.

FIGURE 5
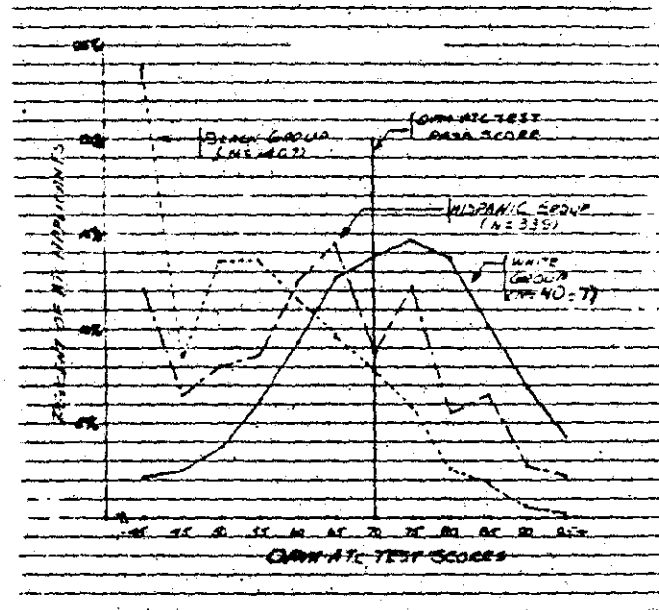
Table 54 compares the distribution of ATC applicants, ATC applicants who passed the OPM test and the Civilian Labor Force (CLF).

TABLE 54
CIVIL LABOR FORCE-ATC APPLICANTS-PASS GROUP

	Civ. Labor Force	% of ATC Applicants	% of Applicants in Pass Group
American Indian	.3%	1%	1.0%
Asian	.8%	1%	1.1%
Black	9.7%	24%	8.9%
White	85.0%	68%	84.8%
Hispanic	4.2%	6%	4.2%
	100.0%	100.0%	100.0%

Administration of the ATC Occupational Knowledge Test (OKT) provided the first opportunity to evaluate the results for a sizable sample of ATC applicants. The OKT used was version 101 C, a 60-item test. Consequently, the results of this test were reviewed for the total group, men, women, racial and ethnic groups and relationship of the results to those applicants with veterans preference.

Tables 43 and 44 provided means and standard deviations on the OKT Test results for the various groups. The "walk in" group of applicants had a lower mean score than the "scheduled" group; 20.1 in comparison to 24.9 and this difference was evident in all racial groups except for blacks where the mean score was 17.4 (scheduled) and 16.2 (walk in). This would indicate that there was relatively little difference in ATC related knowledge for blacks of either group but that other applicants in the "Walk In" group had less ATC related knowledge or experience that those in the "Scheduled" group. In view of the objective of the "Walk In" recruitment program, these differences would be expected.

Table 55 provides the descriptive statistics for men and women on OKT.

TABLE 55

DESCRIPTIVE STATISTICS, SEX VS. OKT

Group	(N)	Scheduled Mean	SD	(N)	Walk In Mean	SD
Men	(3835)	26.9	12.2	(356)	22.2	11.2
Women	(1473)	19.8	8.7	(312)	17.6	6.9

Table 56 provides the OKT score distribution for men and women and for racial groups. In deriving this distribution, the raw scores were converted to a scale of 0-100. In this analysis, the Hispanic group was distributed within the racial groups shown.

TABLE 56

Distribution of OKT Scores

Group	N	Below 15	15-24	25-34	35-44	45-54	55-64	65-74	75-79	Above 79
Men	(4191)	3%	11%	27%	20%	9%	10%	9%	8%	3%
Women	(1785)	5%	20%	43%	19%	5%	4%	3%	1%	*
White	(4242)	2%	9%	27%	22%	10%	10%	9%	7%	3%
Asian / American Indian	(140)	3%	14%	35%	18%	9%	9%	6%	5%	1%
Black	(1443)	8%	32%	39%	13%	2%	2%	2%	2%	*

* Less than 1%

79

Approximately 20 percent of the men scored above 64 as compared to 4 percent of the women; 19 percent of the white group; 12 percent of the Asian/American Indian; and, 4 percent of the black group scored above 64 respectively.

Previous studies with ATC trainees at the Academy as well as developmental and Journeymen ATC specialists provided a significant correlation between scores on the OKT test and success in ATC training and work. These studies also supported the conclusion that the OKT provided a better basis for granting extra credit to ATC applicants who passed the OPM ATC Test than the present rating guide used for this purpose. Consequently, for further analysis, this applicant group's OKT scores were given the following "extra points":

OKT Scores	Extra Credit
65-69	3 points
70-74	5 points
75-79	10 points
80+	15 points

These point values are similar to those presently provided by the OPM rating guide for specific types of ATC related experience which now provides 5, 10 or 15 points depending on the specific experience.

In most cases, the types of ATC related experience for which an applicant is presently granted extra credit is obtained through military service and training. For example, Air Traffic Control experience which is granted the most points (10-15) is obtained through military experience. This also tends to be the case, although to a lesser extent, for pilot and other types of credited experience.

In order to assess the relationship between military experience and OKT scores, an analysis of those applicants with veterans preference, their OKT scores, and the extent to which they could earn extra credit was completed. The results are provided in Table 57 below.

TABLE 57

		Veterans Preference				
		No		Yes		
Score	Credit	N	%	N	%	Total
65-69	3 pts	103	(45%)	126	(55%)	229
70-74	5 pts	76	(33%)	153	(67%)	229
75-79	10 pts	61	(24%)	190	(76%)	251
80+	15 pts	26	(10%)	233	(90%)	259
		266	(28%)	702	(72%)	968

Prior analysis of the relationship between OKT scores and "pass/fail" of
ATC training for 1827 students indicated that the largest drop in total
failure rate was between those trainees who scored between 60-64 (29.6
percent failure rate) and 65-69 (20.1 percent failure rate). Consequently,
for this analysis comparisons were based on providing 3 points for those
applicants who scored between 65 and 69. About 16 percent (968) of the
total 6000 applicants could have been eligible for additional point credit
based on OKT test scores. A total of 702 (72 percent) also could receive
points for veterans preference. The increasing number of applicants with
veterans preference who scored "highest" on the OKT indicate that most of
them would probably have had military air traffic control experience.

In contrast, those applicants who were not veterans constituted 28 percent
of the total group who could receive extra credit based on OKT scores.
While applicants in this group score "lower" on OKT, use of this test to
grant extra credit would broaden the opportunity for non-veterans who
acquired ATC related knowledge through other experience to also qualify for
extra credit.

Under current OPM rating and ranking procedures, an applicant must make a
passing score of at least 70 on the present ATC test battery before
additional credit is given for veterans preference or aviation related
experience. Of the total 968 applicants who could be eligible for extra
credit based on OKT scores, 272 (28 percent) scored below 70 on the present
OPM ATC test battery. Table 58 shows the distribution by "point groups".

TABLE 58
ELIGIBILITY, VETERANS PREFERENCE, AVIATION RELATED EXPERIENCE

Extra Points	Total	Failed OPM ATC Test		Total Eligible for Extra Credit	
15 pts	259	58	(22%)	201	(78%)
10 pts	251	74	(29%)	177	(71%)
5 pts	229	62	(27%)	167	(73%)
3 pts	229	78	(34%)	151	(66%)
	968	272	(28%)	696	(72%)

The 696 applicants who passed the OPM ATC test and would be eligible for
extra credit based on OKT scores, represented 23 percent of the 3083 appli-
cants who passed the test.

Next the data were examined in terms of the ATC test score distribution and
the effect of adding additional credit for veterans preference and aviation
related experience on the score distribution for those who passed. Figure 6
provides the distribution of raw OPM test scores for the total applicant
group.

81
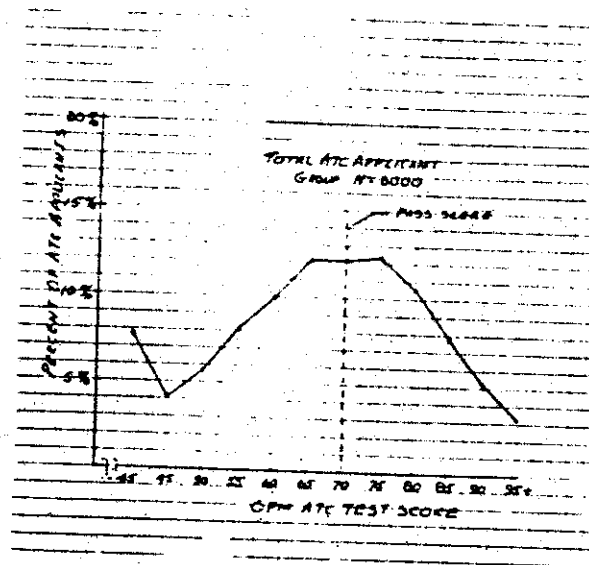
FIGURE 6

Figure 7 provides the score distribution for all ATC applicants who passed the OPM test after adding extra credit for veterans preference and OKT test scores.

FIGURE 7
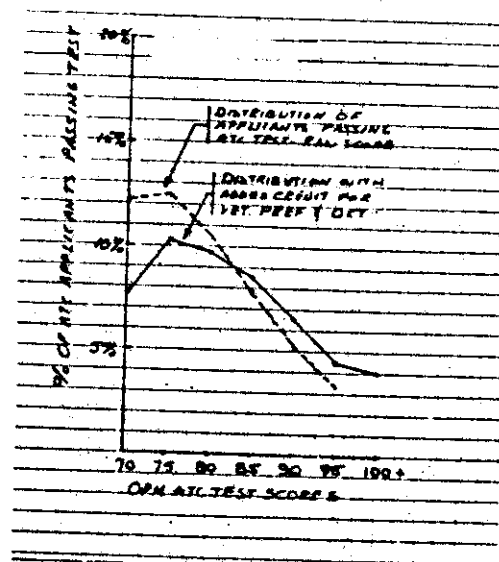
It is evident that granting extra credit for veterans preference and OKT significantly shifts the distribution to higher scores, with a total of 899 applicants scoring 90 or above (29 percent of those passing) compared to 500 (16 percent of those passing) before adding extra credit.

Figure 8 shows the distribution of OPM scores after adding extra credit for veterans preference and OKT test scores for men and women.

FIGURE 8
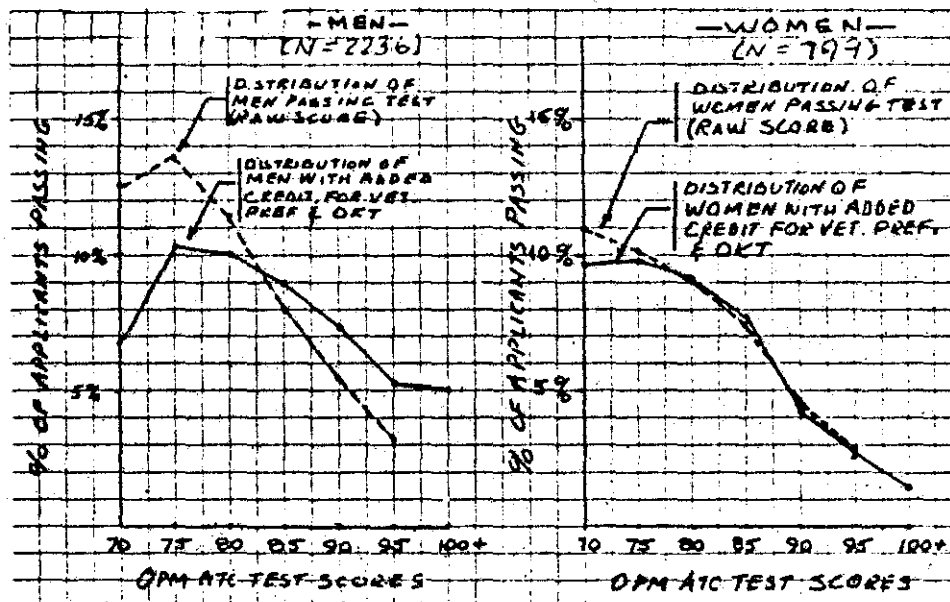
While a significant shift to higher scores is apparent for men, there is relatively little change for women. Of the two extra credit factors, veterans preference accounts for more of the shift than OKT credit.

For every woman who receives extra credit for being a veteran, 16 men also get veteran's preference credit. By comparison, for every woman who would get extra credit for aviation-related experience based on OKT scores, 9 men would also get extra credit. While both factors result in moving more men than women to higher total score ranges for appointment consideration, almost twice as many men are moved up based on veterans preference than on OKT scores.

The effect of granting extra credit on score distribution can also be examined in terms of the relative representation of men, women and minority groups in various score ranges with and without the extra credit.

Table 56 shows the percentage of men and women in each of four score range groups based on raw OPM test scores (before veterans preference or OKT extra credit is applied) and compares those to the percentage of men and women in each group after applying extra credit. It should be noted that by adding extra credit it would be possible to have a maximum total score which exceeds 100.

TABLE 59

OPM TEST SCORE GROUPS
(Pass Sample N=3035)

Group	70-79 Raw Score (N=1440)	70-79 With Credit (N=1076)	80-89 Raw Score (N=1106)	80-89 With Credit (N=1086)	90-100 Raw Score (N=489)	90-99 With Credit (N=644)	100+ With Credit (N=229)
Men	74%	67%	73%	73%	73%	79%	89%
Women	26%	33%	27%	27%	27%	21%	11%
	100%	100%	100%	100%	100%	100%	100%

Based on the raw OPM test scores, women comprise 26-27 percent of each of the score range groups as compared to men. Since women comprised 26 percent of the _total_ pass group, they are also proportionately represented (as compared to men) in each of the score range groups. However, when veterans preference and OKT credit is included, the percentage of women increased to 33 percent in the 70-79 score group; remains the same in the 80-89 group; and drops to 21 percent and 11 percent in the two "highest" scoring groups. Table 59 provides a quantative measure of the extent to which veterans preference and aviation related experience, as measured by OKT, differentially affect men and women for appointment consideration.

Table 60 compares the same data by minority groups.

TABLE 60

OPM TEST SCORE GROUPS
(Pass Sample N=2948

Group	70-79 Rating (N=1387)	70-79 Augmented Rating (N=1042)	80-89 Rating (N=1079)	80-89 Augmented Rating (N=1058)	90-100 Rating (N=482)	90-99 Augmented Rating (N=622)	100+ Augmented Rating (N=226)
White	81%	80%	90%	89%	94%	93%	93%
Hispanic	5%	5%	4%	4%	4%	4%	4%
Black	14%	15%	6%	7%	2%	3%	3%
	100%	100%	100%	100%	100%	100%	100%

In the case of Black and Hispanic minority groups, it is evident that granting veterans preference and OKT credit has very little impact on the proportionate representation within each OPM score group for those applicants who pass the OPM test. Overall, Table 60 indicates that veterans preference and credit for aviation experience affects each of the racial groups essentially the same with respect to their score distribution.

Analysis of Experimental ATC Test and OKT

Based on prior research conducted by the Civil Aeromedical Institute on
a group of 1827 ATC trainees at the FAA Academy, it was shown that an
experimental ATC test battery comprised of the Multiplex Controller
Aptitude Test (MCAT), the OPM Test 157 - Abstract Reasoning, and the
OPM Test 24 - Arithmetic Reasoning provided a high multiple correlation
(R=.54) with ATC training laboratory scores (19). It was also shown
that the use of an ATC Occupational Knowledge test added significantly
to the statistical correlation (26) (27).

Since parallel forms of MCAT and a 60-item version of OKT (101 C) were
administered together with the current OPM ATC test, it was possible to
construct an Experimental ATC Test Battery comprised of MCAT, Abstract
Reasoning, Arithmetic Reasoning, and OKT and compare these results to
the current OPM test battery for the same 6000 applicant group.

Consequently, the analyses previously discussed for the OPM ATC test were
also completed using this Experimental ATC test battery. Each of the
three tests (excluding OKT) were equally weighted in deriving a composite
score for the experimental ATC test battery.

The table numbers used in this section will be the same as those used in
the OPM test analysis followed by a suffix "E" where the same data is
presented for the Experimental ATC test.

Table 61 provides the results of using the Experimental ATC test with
respect to "pass/fail". A minimum raw score of 100.47 was used which also
equates to the mean score on the test battery where about half of the
applicants pass and half fail.

TABLE 61
PASS/FAIL - EXPERIMENTAL ATC TEST

Group	Total Applicants	Pass Exp. ATC Test N	(%)	Fail Exp. ATC Test N	(%)
Men	4191	2312	(55%)	1879	(45%)
Women	1785	852	(48%)	933	(52%)
Total	5976	3164	(53%)	2812	(47%)

With the Experimental ATC test battery (excluding OKT), a somewhat higher
percent of the total applicants passed compared to the OPM test (53 percent
compared to 51 percent).

With respect to sex groups, both men and women passed the Experimental ATC
test at a higher rate, (55 percent men compared to 53 percent on the OPM
test and 48 percent women compared to 45 percent with the OPM test). However,
the mean raw score on the Experimental ATC test was used as a "pass" score.
On the current OPM ATC test the mean raw score for this applicant group
was 207.1 compared to the established pass score of 210. Consequently, the
"higher" pass rate is primarily a result of the pass scores utilized.

85

Given this general comparison of the two tests, one key point then is
the extent to which the individual applicants who passed the two tests
in fact differ. Table 62 shows the number of applicants who: (1) passed
both tests; (2) failed both tests; (3) passed the OPM test but failed the
Experimental ATC test; or (4) failed the OPM test but passed the Experi-
mental ATC test.

TABLE 62
TEST RESULT DIFFERENCES - INDIVIDUAL APPLICANTS

Passed Both (1)	Failed OPM/Passed ATC (4)	
2727 (46%)	437 (7%)	Pass Exp. ATC 3164
Passed OPM/Failed ATC (3)	Failed Both (2)	
308 (5%)	2504 (42%)	Fail Exp. ATC 2812
Pass OPM 3035	Fail OPM 2941	TOTAL 5976

From Table 62 it is evident that using the two different tests affected the
pass/fail status of 745 (12 percent) of the total applicants. In terms of
the Experimental ATC test, 437 different applicants (14 percent of those who
passed) would be eligible for appointment consideration who would not be
eligible based on the current OPM test. While most of the shifts between
pass/fail on the two tests occur around the passing score of 70, some
applicants who scored as low as 45 on the OPM test, passed the ATC Experimental
test. Conversely, some applicants who scored as high as 85 on the OPM test
failed the ATC experimental test.

Table 63 shows the pass/fail rates on the Experimental ATC test by racial
groups.

TABLE 63

PASS/FAIL RATES - EXPERIMENTAL RACIAL GROUPS

Group	Total Applicants	Pass Exp. ATC Test N	(%)	Fail Exp. ATC Test N	(%)
White	4067	2678	(66%)	1389	(34%)
Hispanic	339	141	(42%)	198	(58%)
Black	1407	256	(18%)	1151	(82%)
Total	5813	3075	(53%)	2738	(47%)

For the Experimental ATC test, the percentage of the White group who passed was somewhat higher than on the OPM test (66 percent compared to 64 percent). The Hispanic group also increased in pass rate (40 percent compared to 38 percent) and the Black group pass rate decreased slightly (18 percent compared to 19 percent).

Again, because of the disproportionate number of minorities who did not pass the Experimental ATC test, analysis of score distribution is provided in Table 64.

TABLE 64
SCORE DISTRIBUTIONS—EXPERIMENTAL ATC TEST

Group	Below 50 N	(%)	50-59 N	(%)	60-69 N	(%)	70-79 N	(%)	80-89 N	(%)	89 N	(%)	Total
White	39	(1%)	329	(8%)	1021	(25%)	1758	(43%)	857	(21%)	63	(2%)	4067
Hispanic	23	(7%)	74	(21%)	101	(30%)	105	(31%)	33	(10%)	3	(1%)	339
Black	245	(17%)	465	(33%)	441	(32%)	227	(16%)	28	(2%)	1	(*)	1407

Fail Exp ATC Test (N=2738) → | ← Pass Exp. ATC Test (N=3075)

It is evident from this analysis that, for all racial groups, the number of applicants who did very poor on the Experimental ATC test (scores below 50) is much smaller than on the OPM test. For example, 32 percent of the Black applicants scored below 50 on the OPM test compared to 17 percent on the Experimental test. This also held true for applicants who scored very high on the OPM test where 11 percent of the White group scored above 89 in contrast to only 2 percent on the Experimental ATC test.

Figure 9 provides a graph of the Experimental ATC test score distribution for selective racial groups.

FIGURE 9
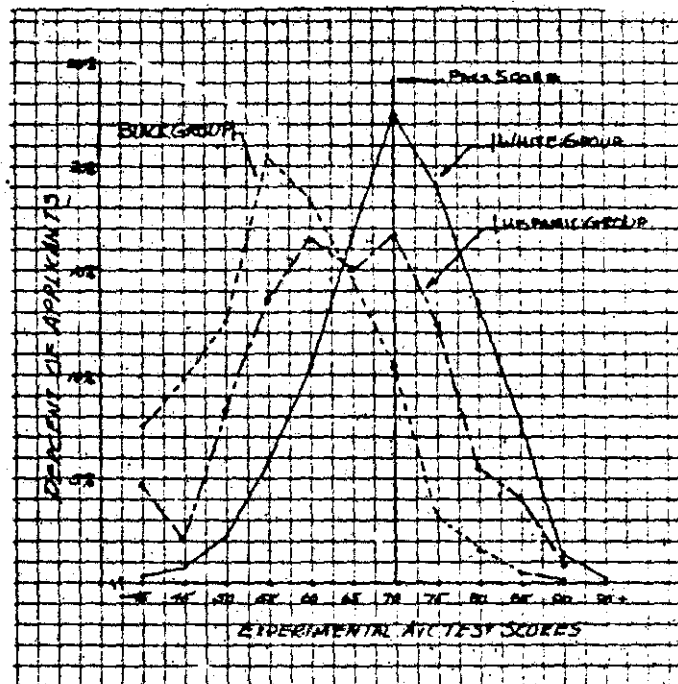
Table 65 shows a comparison of the distribution of applicants who passed
the Experimental ATC test with those who passed the OPM test, and the
distribution of ATC applicants with the Civilian Labor Force (CLF).

TABLE 65

EXPERIMENTAL ATC TEST – OPM TEST RESULTS

Group	Civilian Labor Force	% of ATC Applicants	% of OPM Test Pass Group	% of ATC Exp. Test Pass Group
Am. Ind.	.3%	1%	1.0%	1.2%
Asian	.8%	1%	1.1%	1.3%
Black	9.7%	24%	8.9%	7.6%
White	85.0%	68%	84.8%	85.5%
Hispanic	4.2%	6%	4.2%	4.4%
	100.0%	100.0%	100.0%	100.0%

In analyzing the relationship of the Experimental ATC Test and the Occupational Knowledge Test (OKT), the discussion and descriptive statistics for the total applicant group, as provided in Table 55, 56, and 57 also apply here.

However, some applicants who passed the OPM Test failed the Experimental ATC Test (and vice versa). Consequently, the number of applicants (and individual applicants) who could be eligible for extra credit based on OKT scores will differ for the two tests.

Table 66 provides the number and distribution by "point groups" of the total applicant group, those who failed the Experimental ATC Test, and those who would be eligible for extra credit based on their OKT scores.

TABLE 66
TOTAL APPLICANT POINT GROUPS

Extra Points	Total	Failed Exp. ATC Test		Total Eligible for Extra Credit	
15 pts.	259	60	(23%)	199	(77%)
10 pts.	251	70	(28%)	181	(72%)
5 pts.	229	59	(22%)	170	(78%)
3 pts.	229	85	(37%)	144	(63%)
	968	274	(28%)	694	(72%)

Of the total 968 applicants who could be eligible for extra OKT credit, a total of 206 failed both the OPM and the Experimental test. A total of 66 failed the OPM test but passed the Experimental ATC test. Conversely, 68 passed the OPM test but failed the Experimental ATC test. Consequently, the two tests affected 340 applicants or 35 percent of the applicants who scored above 64 on the OKT.

Figure 10 provides the distribution of raw Experimental test scores for the total applicant group in comparison to the OPM test score distribution.

89

FIGURE 10

TOTAL ATC APPLICANT GROUP
(N = 6000)
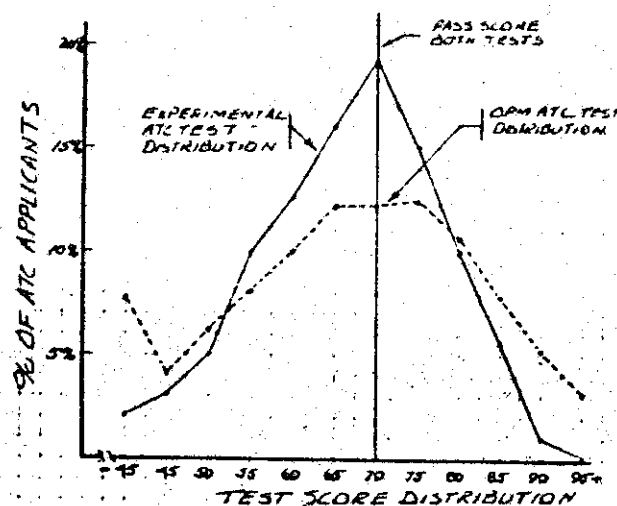
Those applicants who scored below 45 on the OPM ATC test tend to cluster in the 50-59 score range on the Experimental ATC test. Table 67 also provides the comparative distribution statistics for the "pass group" for each of the tests.

TABLE 67

PASS GROUP TEST SCORE DISTRIBUTION

	70-74		75-79		80-84		85-89		90-94		95+		
	N	%	N	%	N	%	N	%	N	%	N	%	TOTAL
OPM Test	697	(23%)	743	(24%)	635	(21%)	471	(16%)	310	(10%)	179	(6%)	3035
Exp. ATC Test	1253	(40%)	895	(28%)	606	(19%)	341	(11%)	63	(2%)	6	(*)	3164

The Experimental ATC test provides a considerably greater differentiation on test scores between the applicants who pass as compared to the OPM test with a much higher clustering of applicants in the 70-79 score range (67 percent compared to 48 percent) and a much smaller number of applicants with scores of 90 and above (2 percent compared to 16 percent).

90

The Experimental ATC test data were also examined in terms of test score distribution and the effect of additional credit for veterans preference and aviation-related experience on the score distribution for those who passed.

Figure 11 provides the score distribution for all ATC applicants who passed the Experimental ATC test after adding extra credit for veterans preference and OKT test scores above 64.

FIGURE 11

EXPERIMENTAL ATC TEST
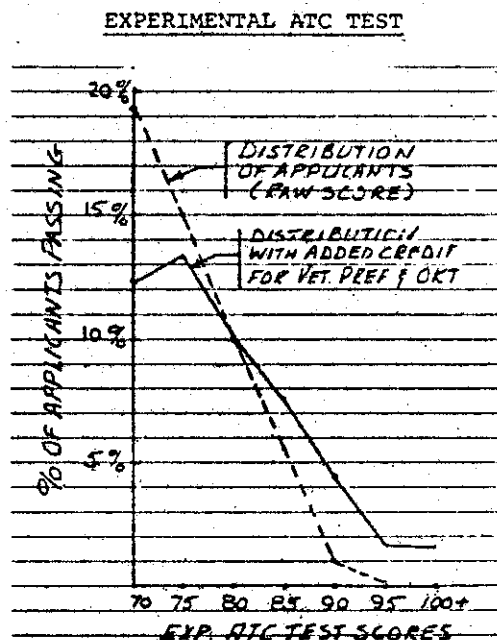
With the Experimental ATC test, a total of 485 applicants scored 90 or above (15 percent of those passing) with the addition of xtra credit for veterans preference and OKT scores compared to 69 (2 percent of those passing) before adding extra credit. This is in contrast to the OPM test where 899 (29 percent of those passing) scored 90 or above after adding extra credit.

Figure 12 shows the distribution of Experimental ATC test scores after adding veterans preference and OKT test score credit for men and women.

91

FIGURE 12

— MEN —
(N = 2312)

— WOMEN —
(N = 852)

DISTRIBUTION OF
MEN PASSING TEST
(RAW SCORE)

DISTRIBUTION OF
MEN WITH ADDED
CREDIT FOR VET.
PREF & OKT.

DISTRIBUTION OF
WOMEN PASSING
TEST (RAW SCORE)

DISTRIBUTION OF
WOMEN WITH ADDED
CREDIT FOR VET.
PREF & OKT.

% OF APPLICANTS PASSING

70 75 80 85 90 95 100+
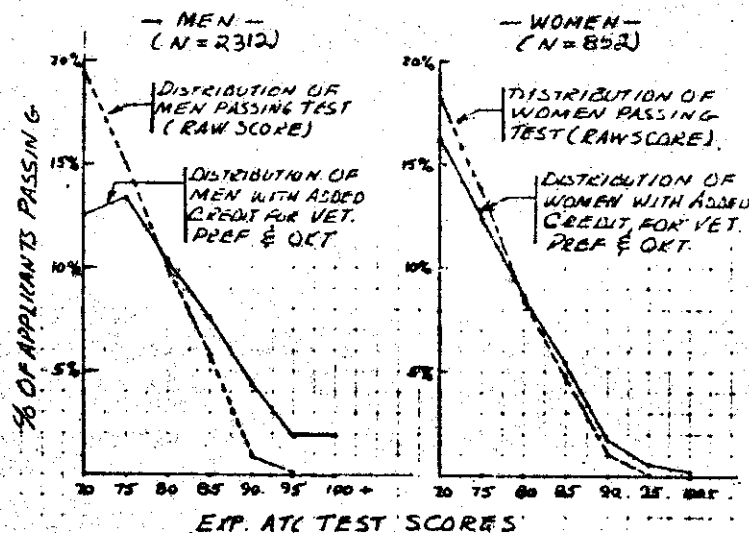
EXP. ATC TEST SCORES

Table 68 also shows the percentage of men and women in each of four score
range groups based on raw Experimental ATC test scores (before veteran
preference or OKT extra credit is applied) and compares those to the pe-
centage of men and women in each group after applying extra credit.

TABLE 68

EXPERIMENTAL ATC TEST SCORE GROUPS

| | 70-79 | | 80-89 | | 50-100 | 90-99 | 100+ |
	Raw Score (N=2118)	With Credit (N=1523)	Raw Score (N=947)	With Credit (N=1156)	Raw Score (N=69)	With Credit (N=374)	With Credit (N=111)
Men	72%	65%	75%	76%	71%	89%	95%
Women	28%	35%	25%	24%	29%	11%	5%
	100%	100%	100%	100%	100%	100%	100%

The overall effect of combining the Experimental ATC test score with
veterans preference and OKT scores is to increase the proportion of women
in the lower score group (70-79) and decrease their representation in the
higher score group (90+) in comparison to use of the current OPM test.

Table 69 compares the same information by minority groups.

TABLE 69

Exp. ATC Test Score Groups
(Pass Sample N=3075)

	70-79		80-89		90-100	90-99	100+
	Raw Score (N=2090)	With Credit (N=1475)	Raw Score (N=918)	With Credit (N=1127)	Raw Score (N=67)	With Credit (N=364)	With Credit (N=109)
White	84%	81%	93%	92%	94%	94%	93%
Hispanic	5%	6%	4%	4%	5%	3%	5%
Black	11%	13%	3%	4%	1%	3%	2%
	100%	100%	100%	100%	100%	100%	100%

Again, the addition of credit for veterans preference and OKT scores has a
relatively small impact by racial group except in the 70-79 score group
where the effect is to decrease the percentage of the white group by 3 percent
(84 to 81 percent) and increase the precentage of the black group by 2 percent
(11 to 13 percent).

The analyses of the Experimental ATC test battery previously discussed were
based on equal weighting of each of the three tests (CSC-24, CSC-157, and
MCAT) included in the experimental test battery.

Multiple regression analysis using the ATC laboratory problem scores as the
criterion for a sample of 1827 ATC trainees at the FAA Academy resulted in
a correlation (R) value of .5407. Converting the beta weights to unit
weights of:

    1 x CSC 24 Arithmetic
    2 x CSC 157 Abstract Reasoning
    4 x Multiple Controller Aptitude

resulted in a correlation value (R) of .5354. In order to assess the effect
of using the weighted test scores, the 1978 ATC applicant group test results
were further analyzed.

Table 70 provides the pass/fail results using the Weighted Experimental
ATC test (excluding OKT) for men and women. With the test weighted, the percentage
of men who passed increased from 53 percent (unweighted) to 56 percent. Women
decreased from 48 percent (unweighted) to 43 percent.

TABLE 70
PASS/FAIL, WEIGHTED EXPERIMENTAL ATC TEST, BY SEX

Group	Total Applicants	Pass Weighted Exp. ATC Test N	(%)	Fail Weighted Exp. ATC Test N	(%)
Men	4191	2348	(56%)	1843	(44%)
Women	1785	768	(43%)	1017	(57%)
Total	5976	3116	(52%)	2860	(48%)

Table 71 provides the pass/fail rates for selective racial groups with the Weighted Experimental ATC test (also excluding OKT). Again, the mean score on the composite weighted test was used as the "cut" score for the applicant group.

TABLE 71
PASS/FAIL, WEIGHTED EXPERIMENTAL ATC TEST BY RACE

Group	Total Applicants	Pass Weighted Exp. ATC Test N	(%)	Fail Weighted Exp. ATC Test N	(%)
White	4067	2698	(66%)	1369	(34%)
Hispanic	339	136	(41%)	203	(59%)
Black	1407	198	(14%)	1209	(86%)
Total	5813	3032	(52%)	2781	(48%)

The pass rate for the White group remained the same (66 percent) as compared to the Unweighted test. The Hispanic group dropped from 42 percent to 41 and, the Black group dropped from 18 percent to 14 percent.

Table 72 provides the weighted test score distribution by racial groups.

TABLE 72
WEIGHTED TEST SCORE BY RACIAL GROUP

	(N=2781) Fail Weighted Exp. ATC Test ⟶						(N=3032) ⟵ Pass Weighted Exp. ATC Test						
	Below 50		50-59		60-69		70-79		80-89		Above 89		
Group	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)	N	(%)	Total
White	20	(*)	273	(7%)	1076	(27%)	1645	(40%)	987	(24%)	66	(2%)	4067
Hispanic	22	(7%)	65	(19%)	100	(30%)	100	(30%)	32	(9%)	4	(1%)	339
Black	187	(13%)	522	(37%)	500	(36%)	167	(12%)	31	(2%)	0	(0%)	1407

*Less than 1 percent

By weighting the test, more of the White applicants shift to the 60-69 and 80-89 score ranges and more of the Black applicants shift to the 50-69 score range. As a total group, fewer applicants score below 50 with the weighted test as compared to the Unweighted test (20 percent versus 25 percent).

94

Figure 13   shows the score distribution for selected racial groups based on
the Weighted Experimental Test.

FIGURE 13
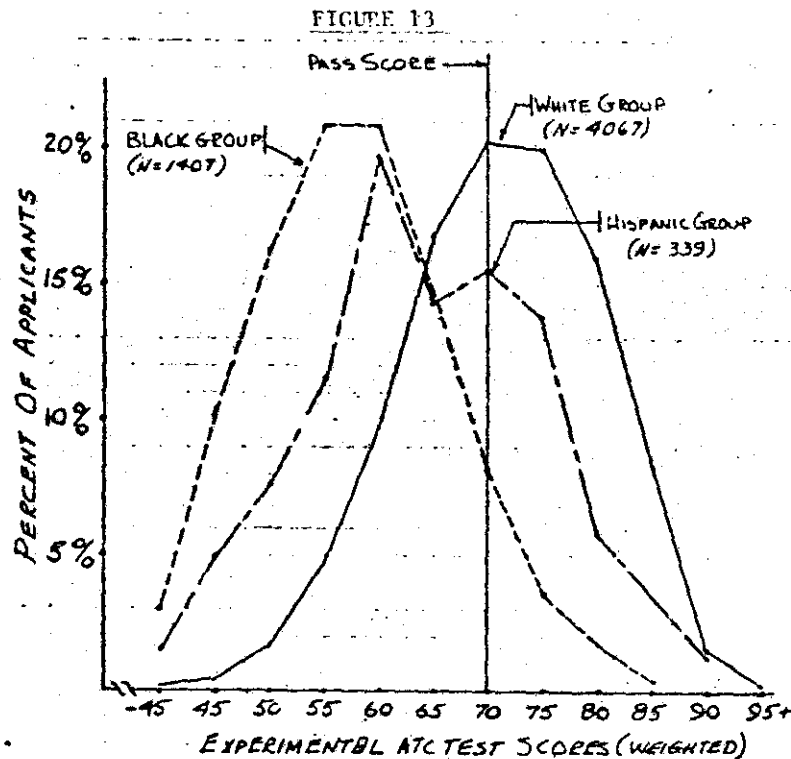
Comparing the Weighted Experimental ATC test battery to the present OPM test
in terms of pass/fail status of the applicants shows that: (1)   2606 passed
both tests; (2) 2431 failed both tests; (3) 510 passed the Weighted ATC test
but failed the OPM test; and, (4) 429 passed the OPM test but failed the ATC
Weighted Test battery. Use of these two different test affect the pass/fail
status of 939 (16 percent) of the applicants, With the Weighted Experimental
ATC test, 510 different applicants (16 percent of those who passed) would be
eligible for appointment consideration that would not be eligible based on
the current OPM test.

In comparing pass/fail results of the Unweighted and Weighted Experimental
ATC Tests, 8 percent of the total applicants were affected; 219 failed the
Unweighted Test but passed the Weighted Test and 267 passed the Unweighted
but failed the Weighted Test.

95

Table 73 compares the distribution of the ATC applicants and those who passed the OPM, Unweighted and Weighted ATC test with the Civilian Labor Force (CLF).

TABLE 73
COMPARATIVE DISTRIBUTIONS

Group	Civilian Labor Force	% of ATC Applicants	% of OPM Pass Group	% of Exp. ATC Test Pass Group	% of Weighted Exp. ATC Test Pass Group
American Indian	.3%	1%	1.0%	1.2%	1.1%
Asian	.8%	1%	1.1%	1.3%	1.0%
Black	9.7%	24%	8.9%	7.6%	6.4%
White	85.0%	68%	84.8%	85.5%	87.1%
Hispanic	4.2%	6%	4.2%	4.4%	4.4%
	100.0%	100%	100.0%	100.0%	100.0%

Figure 14 provides the distribution of weighted test scores for the total applicant group in comparison to the OPM test score distribution. In both cases, the test scores exclude extra credit for veterans preference or OKT.

FIGURE 14
TOTAL ATC APPLICANT GROUP
(N=6800)

96
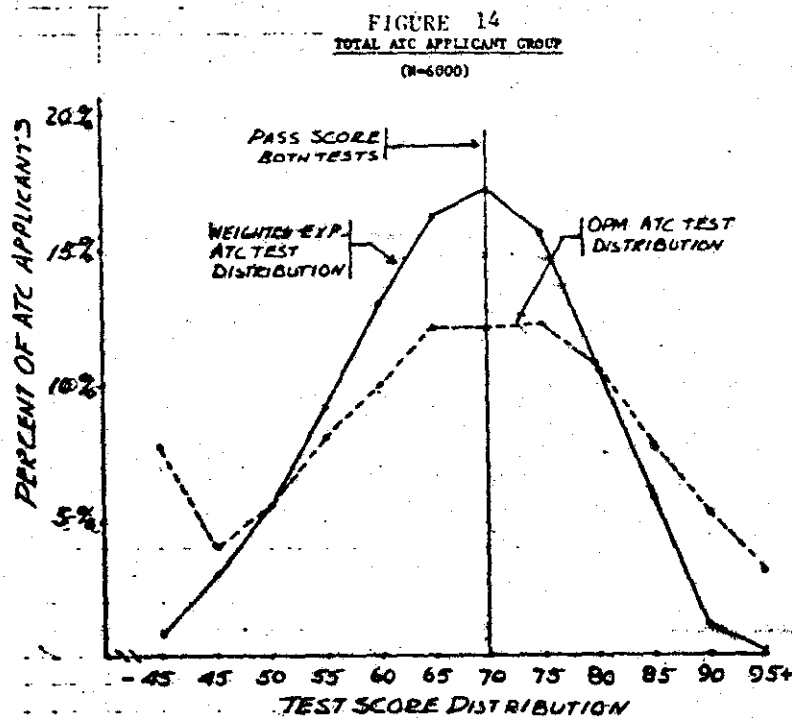
Table 74 shows the percentage of men and women in each of four score range groups based on Weighted ATC test scores before applying veterans preference or OKT credit and compares these to the percentage after applying extra credit. Table 75 provides the same information by racial group.

TABLE 74

EXPERIMENTAL WEIGHTED ATC TEST SCORE GROUPS

	70-79		80-89		90-100	90-99	100+
	Raw Score (N=1988)	With Credit (N=1426)	Raw Score (N=1076)	With Credit (N=1162)	Raw Score (N=72)	With Credit (N=402)	With Credit (N=126)
Men	74%	67%	77%	78%	88%	92%	96%
Women	26%	33%	23%	22%	12%	8%	4%
	100%	100%	100%	100%	100%	100%	100%

TABLE 75

EXPERIMENTAL WEIGHTED ATC TEST SCORE GROUPS

	70-79		80-89		90-100	90-99	100+
	Raw Score (N=1912)	With Credit (N=1381)	Raw Score (N=1050)	With Credit (N=1136)	Raw Score (N=70)	With Credit (N=392)	With Credit (N=123)
White	86%	85%	94%	93%	94%	93%	94%
Hispanic	5%	6%	3%	3%	6%	4%	5%
Black	9%	9%	3%	4%	0%	3%	1%
	100%	100%	100%	100%	100%	100%	100%

Table 76 provides a comparison of the total score distribution with extra credit and the pass rates for the present OPM ATC test battery and the Experimental ATC test battery, both unweighted (UW) and weighted (W). These distributions reflect the composition of the OPM competitive register from which selection of applicants would be made based on the "Earned Rating" for each of the test batteries with extra credit given for veterans preference and OKT scores. (See Tables 59, 74, and 56.)

TABLE 76

EXTRA CREDIT EFFECT ON PASS RATES

Test Battery Earned Rating	70-79 N	(%)	80-89 N	(%)	90-99 N	(%)	100+ N	(%)	Total Pass N	(%)	Total Applicants N
OPM Test	1076	(35%)	1086	(36%)	644	(21%)	229	(8%)	3035	(51%)	5976
Exp. ATC Test (UW)	1523	(48%)	1156	(36%)	374	(12%)	111	(4%)	3164	(53%)	5976
Exp. ATC Test (W)	1426	(46%)	1162	(37%)	402	(13%)	126	(4%)	3116	(52%)	5976

Table 77 provides the descriptive statistics for the applicant group on each of the three composite test batteries by men and women and racial groups.

TABLE 77
DESCRIPTIVE STATISTICS—COMPOSITE TEST BATTERIES

Group	OPM Test Mean	SD	Unweighted Exp. ATC Test Mean	SD	Weighted Exp. ATC Test Mean	SD
Men (N=4191)	210.0	46.0	101.9	22.6	227.3	55.5
Women (N=1785)	199.8	49.3	98.3	22.7	210.7	56.1
White (N=4067)	220.1	39.8	107.5	19.0	241.5	46.7
Hispanic (N=339)	191.7	49.7	93.9	23.4	206.5	56.1
Black (N=1407)	165.0	45.8	79.6	21.0	170.5	47.2
Asian (N=57)	210.7	48.3	103.8	21.7	235.8	50.8
Amer. Indian (N=61)	204.5	41.6	99.2	20.5	222.9	49.6
Total Group	207.1	47.2	100.5	22.7	222.3	56.2
"Pass Score"	210.0	-	100.47	-	222.27	-
Transmuted Score	70.0		70.0*		70.0*	

*Transmuted scores were derived as follows:

Unweighted Experimental ATC test battery

$$Ts = 10.56(\frac{Rs-Rs\bar{x}}{RsSD}) + 70 \quad \text{where:}$$

Rs = Raw score on unweighted ATC test
Rs$\bar{x}$ = Mean of unweighted ATC test (100.47)
RsSD = Standard deviation of unweighted test battery (22.67)

Weighted Experimental ATC test battery

$$Ts = 10.68(\frac{Rs-Rs\bar{x}}{RsSD}) + 70 \quad \text{where:}$$

Rs$\bar{x}$ = 222.27
RsSD = 56.22

Table 78 provides the intercorrelations for the three test batteries.

TABLE 78
INTERCORRELATIONS-COMPOSITE TEST BATTERIES

	Present OPM Test	Exp. ATC Test Unweighted	Exp. ATC Test Weighted
OPM	1.00	.92	.88
ATC (Unweighted)		1.00	.97
ATC (Weighted)			1.00

Conclusions. Given the high attrition rate of ATC trainees and the objective of developing improved methods of selecting applicants for the ATC occupation, attention is focused on the selection tests used in the hiring process. The results of using the Experimental ATC test battery which was derived from the research studies previously discussed are not unexpected.

The number of applicants who score high on the Experimental ATC test (85 and above) is much smaller than on the current OPM test. Based on the validity studies which have been conducted, those applicants who score high on the Experimental ATC test can be expected to have the highest probability of success in the developmental training for the ATC occupation.

Granting extra credit for military service is required by law. This does impact on the competitive appointment consideration for women but does not appear to differentially affect individual minority groups. Aviation-related experience and knowledge which is measured by the Occupational Knowledge test has been shown to be a positive predictor of success in ATC training in the research studies previously discussed. Granting extra credit for the demonstration of this knowledge improves the competitive position of selected applicants but this is consistent with the objective of developing improved selection procedures in order to hire those applicants who demonstrate aptitude for air traffic control work and whose potential for success is greatest.

99

Objective. Research studies discussed up to this point were directed to:

1) Identifying an improved ATC test battery;

2) Development of parallel test forms for those tests which
   consistently produced significant correlations with the
   various criterion measures of ATC success; and

3) Evaluation of experimental and OPM tests in relation to
   applicants for ATC work.

Results of these studies supported the conclusion that a new test battery
comprised of the Multiplex Controller Aptitude Test (MCAT), Abstract
Reasoning (CSC-157), and Arithmetic Reasoning (CSC-24) provided consistently
high correlation with criterion measures. It was also shown that the ATC
Occupational Knowledge Test provided a better measure of aviation-related
knowledge than the present Rating Guide used to evaluate experience. The
Occupational Knowledge Test (OKT) also provides significant correlations
with success in the ATC occupation which added to the validity obtained from
the test battery alone.

With the development of parallel forms of MCAT and OKT, the next step was to
validate the proposed test battery with a new group of ATC trainees attending
the FAA Initial Qualification ATC Terminal and En route training program. Since
this study has not been published elsewhere, this section will discuss the com-
plete methodology and findings.

Sample Description. The sample of trainees used for this validation analysis
consisted of 953 new trainees attending the FAA Academy during the period
June 1978 through December 1978. As in previous studies, tests were admin-
istered on a voluntary basis on the first day of attendance at the Academy.
The composition of the sample with respect to sex and racial groups is provided
in Table 79.

TABLE 79

SEX AND RACE - 1978 SAMPLE

Men	800 (85.0%)		Oriental	10 (1.1%)
Women	141 (15.0%)		Hispanic	23 (2.4%)
Unidentified	12	--	Black	81 (8.5%)
			Other	839 (88.0%)
Total	953		Total	953

100

Table 80 compares the composition of this sample with the 1978 ATC appli-
cant group who passed the OPM test and those who passed the OPM test with
total scores (earned ratings) above 89.

TABLE 80

1978 SAMPLE VERSUS 1978 ATC APPLICANTS

Group	1978 ATC Applicants		Passed OPM Test		Passed OPM Test 90+ Score		1978 ATC Trainee Sample	
Men	4191	(70%)	2236	(74%)	722	(82.6%)	800	(85%)
Women	1785	(30%)	799	(26%)	151	(17.2%)	141	(15%)
	5976	(100%)	3035	(100%)	873	(100%)	941	(100%)
Amer. Ind.	61	(1.0%)	31	(1.1%)	6	(.7%)	–	(--)
Asian	57	(1.0%)	35	(1.2%)	15	(1.7%)	10	(1.1%)
Hispanic	339	(5.7%)	128	(4.2%)	31	(3.6%)	23	(2.4%)
Black	1407	(23.7%)	264	(8.8%)	26	(3.0%)	81	(8.5%)
White	4067	(68.6%)	2556	(84.8%)	791	(91.0%)	839	(88.0%)
	5931	(100.0%)	3014	(100.0%)	869	(100.0%)	953	(100.0%)

The representation of minorities in the ATC trainee sample is somewhat
greater than would be expected if all trainees had been selected through
competitive appointment processes from the OPM register. Because of the
large number of applicants in relation to the relative small number of
vacancies, competitive selections of ATC trainees are normally made in
score ranges above 90. As shown in the analysis of the OPM test scores
for the 1978 ATC applicants, the White group comprised about 91 percent of
all applicants who scored 90 and above on the OPM test battery after in-
clusion of credit for Veterans Preference and aviation-related knowledge/
experience. However, in this group of over 900 ATC trainees, the White
group is 88 percent of the total.

Because of the difficulties in hiring women and minorities through com-
petitive appointment processes from OPM registers, FAA established an
alternate recruitment program in 1968 which provides for non-competitive
appointment of individuals who are already federal employees with status
in the career Civil Service system. These individuals are required to
pass the same ATC test battery with a score of 70 or more as the appli-
cants appointed from the OPM register. However, having passed the test,
they can be hired at the GS-5 level non-competitively. This alternative
recruitment program is identified as the ATC Predevelopmental Program.

101

Approximately 200 ATC trainees are hired each year through this alternate recruitment program. This represents about 10 percent of annual new ATC hires. These employees receive a year of predevelopmental training in air traffic control related subjects including 17 weeks of formal classroom instruction at the Predevelopmental Training Center at the University of Oklahoma followed by training assignments at FAA terminal, en route centers, and flight service station field facilities. Upon completion of predevelopmental training, these employees are promoted to GS-7 (the normal entry grade for ATC trainees) and assigned to the FAA Academy for Initial ATC Qualification Training.

In order to further determine the composition of the 1978 ATC trainee group, Table 81 identifies the number of predevelopmental employees included in the sample by sex and racial group.

TABLE 81

PREDEVELOPMENTAL EMPLOYEES BY SEX AND RACE

Group	Non-Predevelopmental (Competitive Appt.)			Predevelopmental (Non-Competitive Appt.)			Total	
	N	% of Total	% of Group	N	% of Total	% of Group	N	%
Men	739	(92.4%)	(88.9%)	61	(7.6%)	(55.5%)	800	(100%)
Women	92	(65.2%)	(11.1%)	49	(34.8%)	(44.5%)	141	(100%)
	831	(88.3%)	(100.0%)	110	(11.7%)	(100.0%)	941	(100%)
Asian	9	(90.0%)	(1.1%)	1	(10.0%)	(1.0%)	10	(100%)
Hispanic	11	(47.8%)	(1.3%)	12	(52.2%)	(10.9%)	23	(100%)
Black	26	(32.1%)	(3.1%)	55	(67.9%)	(50.0%)	81	(100%)
Other	797	(95.0%)	(94.5%)	42	(5.0%)	(38.1%)	839	(100%)
	843	(88.5%)	(100.0%)	110	(11.5%)	(100.0%)	953	(100%)

It is evident that the Predevelopmental recruitment program provides a major avenue for entry of women and minorities into the GS-7 Initial Qualification ATC training program. About 12 percent of this 1978 ATC trainee sample entered the ATC occupation through the Predevelopmental Program. Approximately 35 percent of the women, 52 percent of the Hispanics, and 68 percent of the Blacks in this group of over 900 ATC trainees came into the occupation through Predevelopmental recruitment efforts.

Table 82 compares the composition of this group of ATC trainees with the Civilian Labor Force (CLF) in relation to both the ATC trainees hired through competitive and non-competitive OPM procedures.

TABLE 82
CIVILIAN LABOR FORCE VS. ATC HIRES BY CLASS

Group	CLF	Competitive 1978 ATC Trainees	Total (Comp. & Non-Comp.) 1978 ATC Trainees
Men	62%	89%	85%
Women	38%	11%	15%
Amer. Indian	.3%	--	--
Asian	.8%	1.1%	1.1%
Hispanic	4.2%	1.3%	2.4%
Black	9.7%	3.1%	8.5%
Other	85.0%	94.5%	88.0%
	100.0%	100.0%	100.0%

Predictors. The predictors used in this analysis were:

- Multiplex Controller Aptitude Test (4o6e, 4e6o, 6e7o, 6o7e, 7340, and 7o4e)

- CSC Test 24 - Arithmetic Reasoning

- CSC Test 157 - Abstract Reasoning

- ATC Occupational Knowledge Test
  OKT 101B, 100 items; OKT 101C, 60 items, OKT 102
  (Forms A, B, C, D, E, F, G, H), 80 items keyed

In this study, two of the parallel forms of MCAT were administered to each student. MCAT 1 was the first form administered; MCAT 2 was the second.

Table 83 provides the intercorrelation of raw scores for each half or part test on each of two parallel forms of MCAT administered to 617 ATC students at the FAA Academy.

TABLE 83

INTERCORRELATIONS-MCAT PARALLEL FORMS

	1	2	3	4	5	6
First MCAT Administered (MCAT 1)						
1. First Half, Rights	—	.58	.89	.48	.42	.51
2. Second Half, Rights			.89	.48	.49	.56
3. Total Rights				.89	.89	.60
Second MCAT Administered (MCAT 2)						
4. First Half, Rights					.52	.85
5. Second Half, Rights						.89
6. Total Rights						—

103

A test-retest correlation of .60 between comparable two-part 35 minute
forms of MCAT was obtained. Test reliability was .75 for this restricted
group. Administering two forms of MCAT would provide an estimated reli-
ability of about .88 with an unrestricted population. Some 265 trainees
in the total sample (953) took forms 101B or 101C of the OKT. For pur-
poses of this analysis, these test scores were converted to a scale common
to forms 102 of the OKT.

The descriptive statistics on test scores for this sample of ATC trainees
is provided in Table 84.

TABLE 84

DESCRIPTIVE STATISTICS-MCAT 1978 SAMPLE

	MEN			WOMEN			TOTAL		
	(N)	Mean	SD	(N)	Mean	SD	(N)	Mean	SD
MCAT 1	(800)	37.4	6.9	(141)	34.1	7.8	(953)	36.9	7.1
MCAT 2	(800)	42.9	5.8	(141)	40.7	6.4	(953)	42.6	6.0
CSC 24	(515)	46.6	6.5	(67)	47.2	7.1	(592)	46.6	6.6
CSC 157	(515)	38.8	6.2	(67)	39.8	6.0	(592)	38.8	6.2
OKT	(800)	64.9	15.5	(141)	57.7	15.1	(953)	63.9	15.6

	ORIENTAL			HISPANIC			BLACK			WHITE		
Test	(N)	Mean	SD	(N)	Mean	SD	(N)	Mean	SD	(N)	Mean	SD
MCAT 1	(10)	36.2	6.7	(23)	37.1	6.6	(81)	29.8	7.9	(839)	37.6	6.7
MCAT 2	(10)	43.2	5.7	(23)	42.7	7.4	(81)	36.6	6.8	(839)	43.2	5.5
CSC 24	(1)	44.0	—	(7)	48.7	6.3	(39)	44.2	7.1	(545)	46.9	6.5
CSC 157	(1)	43.0	—	(7)	38.7	5.4	(39)	35.3	6.1	(545)	39.1	6.2
OKT	(10)	61.3	13.5	(23)	59.0	19.9	(81)	66.2	13.1	(839)	63.9	15.8

About 12 percent of the 953 ATC trainees entered the occupation through the
Predevelopmental Program. It was of interest to establish the extent to
which mean scores and standard deviations on each of the tests differed
between the Predevelopmental (non-competitive appointments) and non-prede-
velopmental (competitive appointments) groups. Table 85 provides the
descriptive statistics for each of these groups by sex. It is important
to recognize that the scores for the Predevelopmental trainees on MCAT
and OKT were obtained _after_ they had completed a year of training. Scores
on the CSC-24 and CSC-157 tests were obtained for both groups based on the
current ATC test battery administered prior to employment in the ATC
occupation.

TABLE 85

DESCRIPTIVE STATISTICS BY SEX - 1978 SAMPLE

	Non-Predevelopmental ATC Trainees						Predevelopmental ATC Trainees					
	Men			Women			Men			Women		
	N	Mean	(SD)	N	Mean	(SD)	N	Mean	(SD)	N	Mean	(SD)
MCAT 1	(739)	37.8	(6.6)	(92)	35.2	(7.4)	(61)	32.3	(9.1)	(49)	32.4	(7.9)
MCAT 2	(739)	42.3	(5.4)	(92)	41.5	(6.2)	(61)	38	(8.1)	(49)	39.5	(6.8)
CSC-24	(495)	46.7	(6.5)	(46)	48.1	(7.1)	(20)	44.3	(6.6)	(19)	45.0	(6.7)
CSC-157	(495)	38.9	(6.2)	(48)	40.8	(5.7)	(20)	35.6	(5.7)	(19)	37.3	(6.2)
OKT	(739)	64.4	(15.7)	(92)	52.0	(15.6)	(61)	72.5	(10.8)	(49)	66.1	(9.4)

On all tests except OKT, the mean score of the Predevelopmental group is lower than the non-predevelopmental group. On OKT it is evident that the year of predevelopmental training for employees who are hired non-competitively is quite successful in providing knowledge on air traffic rules, regulations, and procedures which the OKT was designed to measure. Both men and women Predevelopmentals score significantly higher than their non-predevelopmental counterparts as reflected by the mean scores on OKT for the two groups. In addition, the standard deviations on OKT for the Predevelopmental group was considerably lower, indicating a much more restricted range of test scores.

On the two present CSC tests (157 and 24), women in both groups score somewhat higher than men. This is consistent with the pattern derived from the 1976-1977 ATC applicant group who passed the OPM ATC test and the 1978 ATC applicant group.

On the MCAT, men and women in both groups increased their mean test scores on the second administration of this test. However, the difference in mean scores between the predevelopmental and the non-predevelopmental groups remained fairly constant for each of the tests. Men in the predevelopmental program scored between 4 and 5 points lower and predevelopmental women between 2 and 3 points lower on each of the tests as compared to non-predevelopmental trainees.

Table 86 provides the descriptive statistics for the predevelopmental and non-predevelopmental groups by race.

TABLE 86

Non-Predevelopmental ATC Trainees

	Oriental			Hispanic			Black			White			Total		
	(N)	Mean	(SD)	(N)	Mean	(SD)	(N)	Mean	(SD)	(N)	Mean	(SD)	(N)	Mean	(SD)
MCAT 1	(8)	36.4	(5.2)	(11)	41.7	(4.7)	(24)	30.8	(6.4)	(778)	37.7	(6.7)	(821)	37.5	(6.7)
MCAT 2	(8)	43.4	(5.7)	(11)	43.2	(5.3)	(24)	38.5	(5.6)	(778)	43.2	(5.4)	(821)	43.1	(5.5)
CSC 24	(1)	44.0	—	(4)	51.8	(6.2)	(15)	42.4	(7.1)	(532)	46.9	(6.5)	(552)	46.9	(6.5)
CSC 157	(1)	43.0	—	(4)	42.0	(3.6)	(15)	35.4	(6.3)	(532)	39.1	(6.2)	(552)	39.0	(6.2)
OKT	(8)	60.9	(15.4)	(11)	48.1	(20.5)	(24)	58.3	(15.1)	(778)	63.5	(16.0)	(821)	63.1	(16.1)

Predevelopmental ATC Trainees

	Oriental			Hispanic			Black			White			Total		
	(N)	Mean	(SD)	(N)	Mean	(SD)	(N)	Mean	(SD)	(N)	Mean	(SD)	(N)	Mean	(SD)
MCAT 1	(1)	34.0	—	(12)	32.9	(5.2)	(56)	29.5	(8.3)	(43)	36.4	(6.9)	(112)	32.6	(8.1)
MCAT 2	(1)	37.0	—	(12)	42.3	(9.2)	(56)	35.8	(7.1)	(43)	42.1	(6.0)	(112)	39.0	(7.5)
CSC 24	-	—	—	(3)	44.7	(4.5)	(24)	45.4	(7.0)	(13)	43.9	(6.8)	(40)	44.8	(6.6)
CSC 157	-	—	—	(3)	34.3	(4.5)	(24)	35.3	(6.1)	(13)	38.2	(6.5)	(40)	36.1	(6.1)
OKT	(1)	60.0	—	(12)	69.1	(13.6)	(56)	69.9	(10.4)	(43)	70.1	(10.3)	(112)	69.7	(10.7)

As a group, the Predevelopmental trainees scored lower on all tests as compared to the non-predevelopmental trainees except on OKT. The greatest difference in mean scores were associated with the MCAT test which also had greater score ranges for the Predevelopmental group. These differences were generally evident within each of the racial groups as well.

On the OKT, each racial group of Predevelopmentals generally scored signi-
ficantly higher with a more restricted range of scores than their non-pre-
developmental counterparts. However, it does not appear that the academic
ATC knowledge acquired during the one year of training carrys over, in
application, to the MCAT for the predevelopmental group.

On the MCAT, there is a consistent pattern of "learning" between the first
and second administration of the test for all trainees although the non-
predevelopmental trainees generally maintain a constant difference in mean
scores on both tests as compared to the predevelopmental trainees. This
pattern supports the use of a double length MCAT as a means of measuring
this "learning ability" on the part of applicants.

Criterion. In this analysis, whether or not the student passed the ATC
Academy training course or failed/withdrew was used as the criterion
measure against which the tests were validated. Of the students who with-
drew, those who did so for medical reasons or because of personal or family
emergencies, were excluded from the sample since their training was inter-
rupted through no cause of the student. In these cases, the student is
allowed to reenter training after the medical situation or emergency is
resolved.

Essentially all of the students who fail or withdraw do so in the laboratory
phase of training which requires the trainees to demonstrate and apply their
knowledge and skill to air traffic control problems conducted in a laboratory
environment. The structure of the grading of the laborato , phase of training
is provided in Figure 15. Estimated reliability for the six en route control
problems is .79 and .75 for the terminal control problems for students attending
the FAA Academy during the period June-December 1978.

The ATC laboratory training phase consists of six operational ATC problems
designed to replicate the non-radar air traffic control environment. Trainees
conduct identical control problems within a common air traffic control sector
using identical equipment and evaluation criteria for each ATC option. These
operational problems provide stimulus and decision-making conditions repre-
sentative of the job environment and require responses from the trainees in
terms of actions that are both possible and required to control air traffic
on-the-job with a minimum of restrictions. Students are given a series of
practice problems before each of the six graded problems as part of the
training and learning process. Each graded operational problem takes about
one hour and is scored by an instructor using established grading standards
for defined types of student errors or deficiencies. Additionally, each
student's performance is evaluated by the instructor on a rating scale of
40 to 100. Each problem is scored and evaluated by a different instructor.

After the operational problems, a Controller Skills Test (CST) is administered.
This test was designed to measure the application of knowledge and skills
taught during the first months of training. Three basic elements for evalua-
tion are distributed within the test: (1) Application of aircraft separation
standards—students must respond to control situations presented by flight
strips and charts; (2) responding to or forwarding information received

106

which pertains to coordination of aircraft movement or information with other, controllers; and (3) other ATC control items such as board management, timeliness of actions and phraseology. The CST is a 50-item multiple choice test with an administration time of one hour. The correlation between the average score on the six operational problems and the CST for this group of trainees was .50 for those in the terminal option (N=454) and a correlation of .54 for enroute students (N=473).

During the academic training of the ATC Laboratory Phase, a bloc test is administered for each subject. A comprehensive knowledge test 's also administered to measure the degree to which students have learned the academic portions of ATC subjects taught during laboratory training. This comprehensive phase test is administered prior to giving any operational laboratory control problems. Figure 9 provides a breakdown of the ATC laboratory training and the basis upon which a student's total score for this phase of training is derived. Students must complete this phase of training with a total score of 70 or more to pass the FAA Academy and continue on-the-job training at assigned ATC facilities.

Analytical Methodology. The first step in the analysis was the calculation of descriptive statistics on the ATC trainee sample using those students who had test and criterion data. It should be noted that the sample size was affected by the lack of complete information on all variables, particularly on CSC tests 24 and 157 where scores could not be retrieved from FAA or OPM records.

Multiple regression analysis was carried out using the experimental test battery with and without OKT to determine the multiple correlation of the test battery with "pass" or "fail" status of the trainees. Separate regression analyses were conducted for the total group and for subgroups: men, women and racial groups where subsample sizes warranted.

The selection utility of the test battery was examined to assess the impact on training costs; impact on the sample of ATC trainees in terms of those who would (or would not) have been hired if the test battery had been used for selection decisions; and, an analysis of the pass/fail status of those who would have been either eligible or ineligible for hiring based on the test.

107

FIGURE 15

COMPONENTS AND WEIGHTS USED IN COMPUTING THE
FAA ACADEMY ATC LABORATORY TRAINING TOTAL SCORE

	**Extra Credit 13.00%	Sixth Problem	Extra Credit Instructor Assessment Problem Errors	2.60% 3.90% 6.50%
	**Instructor Assessment 19.50%	Fifth Problem	Extra Credit Instructor Assessment Problem Errors	2.60% 3.90% 6.50%
Lav Average *65.0%		Fourth Problem	Extra Credit Instructor Assessment Problem Errors	2.60% 3.90% 6.50%
	Problem Errors 32.50%	Third Problem	Extra Credit Instructor Assessment Problem Errors	2.60% 3.90% 6.50%
		Second Problem	Extra Credit Instructor Assessment Problem Errors	1.30% 1.95% 3.25%
		First Problem	Extra Credit Instructor Assessment Problem Errors	1.30% 1.95% 3.25%

Controller Skills Test   25.00%

Comprehensive Phase Test 8.00%

Block Average 2.00%

*The lab average constitutes 65% of the total score for this phase of
ATC training.

**On each lab problem the instructor gives a performance rating for that
problem that is averaged with the student's problem performance.
Since the rating is not allowed to be below 40, essentially the student
is given a certain amount of extra credit in the computation of the
problem average. Each of the six problems is graded by a
different instructor.

108

Results. The following discusses the analysis and findings in terms of
the ATC sample used, the descriptive statistics, multiple regression, and
utility of the tests.

ATC Sample Used. As pointed out previously, the sample of 953 ATC trainees
was affected by the lack of complete information on all variables. As a
consequence, the total sample was not available for some of the analyses,
particularly the multiple regression. Table 87 provides a distribution of
the total ATC trainee sample (953) showing the number of trainees who had
data available for each of the variables.

TABLE 87

AVAILABLE DATA - 1978 ATC TRAINEE SAMPLE

	Variables	Status As Predev. NonPredev.	MCAT 1 MCAT 2 OKT	CSC 24 CSC 157	Pass/Fail Criterion
Sex	Men	788	800	515	790
	Women	133	141	67	137
	Total	921	941	582	927
Race	Black	80	81	39	79
	Hispanic	23	23	7	23
	Asian	9	10	1	9
	Other	821	839	545	821
	Total	933	953	592	932

Availability of test scores on CSC 24 and 157 had the most impact on the
ATC sample which could be used in some of the analyses. Table 88 provides
a comparison of the means and standard deviations for the group of trainees
with CSC 24 and 157 scores, the group without, and the scores on those tests
common to both groups.

In addition to common test data, OPM Transmuted scores and Earned Ratings
were available for many of these ATC trainees. The Transmuted OPM score
is the score the trainee made on the present OPM test battery before additional
credit for veteran's preference or aviation-related experience is applied.
The Earned Rating is the total OPM score including extra credit for veteran's
preference and aviation-related experience. Table 88 provides the means and
standard deviations for these data. The data for the pass/fail criterion is
also provided. These data were examined for statistically significant
differences.

TABLE 88
DESCRIPTIVE STATISTICS - CSC TEST SCORES

Common Tests and Data	Total ATC Sample			ATC Sample With CSC 24/157			ATC Sample Without CSC 24/157			Sig. Test Results	Est. Effect Size
	(N)	p or Mean	SD	(N)	p or Mean	SD	(N)	p or Mean	SD		
MCAT 1	(953)	36.88	(7.12)	(592)	37.07	(6.94)	(361)	36.56	(7.40)	t = 1.07	d = .07
MCAT 2	(953)	42.59	(5.96)	(592)	43.01	(5.87)	(361)	41.90	(6.05)**	t = 2.80	d = .19
OKT	(953)	63.92	(15.64)	(592)	63.23	(15.82)	(361)	65.04	(15.31)	t = -1.75	d = .12
OPM Score (T)	(704)	87.51	(7.69)	(591)	87.21	(7.62)	(113)	89.08	(7.80)**	t = -3.10	d = .24
OPM (ER)	(651)	94.93	(5.66)	(540)	95.07	(5.43)	(111)	94.22	(6.61)	t = 1.44	d = ..4
Pass/Fail (p)	(939)	.63	(.48)	(585)	.62	(.49)	(354)	.64	(.48)	z = - .62	h = .04

** (p < .01)

A t-test was performed to examine the significance of differences on the tests and OPM scores. A z-test was used to measure the difference between the proportions for pass/fail. The differences are not significant operationally.

A number of analyses were based on sex and race in addition to test scores. Consequently, test scores on MCAT 1, MCAT 2 and OKT were examined for statistically significant differences by race and sex between the ATC trainees with CSC 24 and 157 scores and those without scores on the latter two tests. Table 89 provides the results. The "Hispanics" and "Asian" groups were not used because of the small sample numbers.

As shown in Table 89, group differences were not significant for the "Women" or "Black" groups. Statistically significant differences were found for the "Men" and "Other" groups where the sample sizes were large. However, the estimated effect size in Table 89 are all below the Cohen's criterion (31) for small.

TABLE 89
MCAT ANALYSES BY SEX AND RACE

Group	Test	TOTAL ATC SAMPLE (N)	Mean	SD	ATC SAMPLE WITH CSC 24/157 (N)	Mean	SD	ATC SAMPLE WITHOUT CSC 24/157 (N)	Mean	SD	SIG. TEST RESULTS	EST. EFFECT SIZE
Men	MCAT 1	(800)	37.4	(6.9)	(515)	37.6	(6.6)	(285)	37.0	(7.3)*	t = 2.31	d = .09
	MCAT 2	(800)	42.9	(5.8)	(515)	43.3	(5.7)	(285)	42.2	(6.0)**	t = 6.11	d = .19
	MCAT TOT	(800)	80.3	(11.25)	(515)	80.9	(10.9)	(285)	79.2	(11.8)*	t = 2.46	d = .15
	OKT	(800)	64.9	(15.5)	(515)	64.1	(15.8)	(285)	66.5	(14.8)	t = -1.85	d = .16
Women	MCAT 1	(141)	34.1	(7.8)	(67)	33.3	(8.0)	(74)	34.8	(7.7)	t = - .81	d = .19
	MCAT 2	(141)	40.7	(6.4)	(67)	40.7	(6.6)	(74)	40.6	(6.3)	t = .08	d = .02
	MCAT TOT	(141)	74.8	(13.1)	(67)	74.0	(13.3)	(74)	75.4	(13.0)	t = - .27	d = .11
	OKT	(141)	57.7	(15.1)	(67)	56.3	(14.3)	(74)	58.9	(15.8)	t = - .38	d = .17
Black	MCAT 1	(81)	29.8	(7.9)	(39)	29.7	(7.7)	(42)	29.8	(8.2)	t = - .03	d = .1
	MCAT 2	(81)	36.6	(6.8)	(39)	36.3	(7.5)	(42)	36.9	(6.1)	t = - .26	d = .09
	MCAT TOT	(81)	66.4	(13.2)	(39)	66.0	(13.9)	(42)	66.7	(12.8)	t = - .01	d = .05
	OKT	(81)	66.2	(13.1)	(39)	65.6	(13.4)	(42)	66.8	(13.0)	t = - .14	d = .09
Hispanic	MCAT 1	(23)	37.1	(6.6)	(7)	33.4	(6.7)	(16)	38.8	(6.1)		
	MCAT 2	(23)	42.7	(7.4)	(7)	40.1	(5.7)	(16)	43.8	(8.0)		
	MCAT TOT	(23)	79.8	(12.2)	(7)	73.5	(9.5)	(16)	82.6	(12.5)		
	OKT	(23)	59.0	(20.0)	(7)	60.4	(24.7)	(16)	58.4	(18.4)		
Asian	MCAT 1	(10)	36.2	(4.7)	(1)	31.0	(—)	(9)	36.8	(4.6)		
	MCAT 2	(10)	43.2	(5.7)	(1)	42.0	(—)	(9)	43.3	(6.0)		
	MCAT TOT	(10)	79.4	(9.5)	(1)	73.0	(—)	(9)	80.1	(9.8)		
	OKT	(10)	61.3	(13.7)	(1)	86.0	(—)	(9)	58.6	(11.2)		
Other	MCAT 1	(839)	37.6	(6.7)	(545)	37.7	(6.6)	(294)	37.4	(6.9)	t = 1.25	d = .04
	MCAT 2	(839)	43.2	(5.5)	(545)	43.5	(5.4)	(294)	42.5	(5.6)**	t = 6.25	d = .18
	MCAT TOT	(839)	80.7	(10.8)	(545)	81.2	(10.5)	(294)	79.9	(11.1)*	t = 2.17	d = .12
	OKT	(839)	63.9	(15.8)	(545)	63.1	(15.8)	(294)	65.4	(15.5)	t = -1.78	d = .15

* p(≤ .05)
** p(≤ .01)

The question as to whether the two groups can be considered equivalent cannot be answered definitely. However, the analysis supports the assumption that they are, considering:

- No group differences were found on four of the measures (Table 88).

- No group differences were found for the "Women" or "Black" groups.

- Where group differences were found, the estimated effect sizes were small.

- Where group differences were statistically significant, they were in a positive direction for MCAT, but a negative direction for OKT. They do not consistently favor one group or the other.

Table 90 shows the distribution of the total 953 ATC sample in comparison to the ATC trainees (592) who had scores on all test variables by Prede-velopmental and non-predevelopmental groups.

111

TABLE 90
DISTRIBUTIONS—INITIAL AND REDUCED SAMPLES

| | Total ATC Sample | | | | Predevelopmental | | | | Non-Predevelopmental | | | |
| | Initial Sample | | Reduced Sample | | Initial Sample | | Reduced Sample | | Initial Sample | | Reduced Sample | |
Group	N	%	N	%	N	%	N	%	N	%	N	%
Men	800	(85.0%)	515	(88.5%)	61	(55.5%)	20	(51.3%)	739	(88.9%)	495	(91.2%)
Women	141	(15.0%)	67	(11.5%)	49	(54.5%)	19	(48.7%)	92	(11.1%)	48	(0.8%)
Total	941	(100.0%)	582	(100.0%)	110	(100.0%)	39	(100.0%)	831	(100.0%)	543	(100.0%)
Black	81	(8.5%)	39	(6.6%)	55	(50.0%)	24	(60.0%)	26	(3.1%)	15	(2.7%)
Hispanic	23	(2.4%)	7	(1.2%)	12	(10.9%)	3	(7.5%)	11	(1.3%)	4	(.7%)
Asian	10	(1.1%)	1	(.2%)	1	(1.0%)	-	(--)	9	(1.1%)	1	(.2%)
Other	839	(88.0%)	545	(92.0%)	42	(38.1%)	13	(32.5%)	797	(94.5%)	532	(96.4%)
Total	953	(100.0%)	592	(100.0%)	110	(100.0%)	40	(100.0%)	843	(100.0%)	552	(100.0%)

Descriptive Statistics. Of the total 953 ATC trainees, 582 men and women and 592 trainees based on race had complete data on all test variables. The next step was to compute a total weighted test battery score and descriptive statistics using MCAT 1, MCAT 2, CSC 157 and CSC 24 for the total group; those who passed the ATC training and those who failed or withdrew. In addition, the number of students, means and standard deviations for the Predevelopmental trainees (non-competitive appointments) and non-predevelopmental trainees (competitive appointments) are shown in Table 91.

At this point, OKT scores were excluded, the tests used were weighted as follows:

$$2(\text{MCAT } 1) + 2(\text{MCAT } 2) + 1(\text{CSC } 157) + 1(\text{CSC } 24)$$

Weights were derived from the regression analysis which is discussed later.

TABLE 91
DESCRIPTIVE STATISTICS—TWO DEVELOPMENTAL CATEGORIES

GROUP	PASS (N)	MEAN	(SD)	FAIL (N)	MEAN	(SD)	TOTAL (N)	MEAN	(SD)
Men	(329)	253.0	(22.5)	(186)	234.2	(28.9)	(515)	247.2	(26.9)
Women	(36)	246.6	(32.9)	(32)	223.2	(32.4)	(67)	235.1	(29.9)
Black	(12)	246.0	(15.9)	(26)	198.5	(23.1)	(39)	213.5	(28.2)
Hispanic	(4)	244.6	(27.3)	(3)	221.0	(13.5)	(7)	234.6	(24.2)
Asian	(1)	233.0	(—)	(—)	—	(—)	(1)	233.0	(—)
Other	(345)	254.7	(28.7)	(194)	237.0	(27.6)	(545)	248.7	(25.9)
Total Group	(—)	—	(—)	(—)	—	(—)	(592)	243.6	(27.6)

NON-PREDEVELOPMENTAL ATC TRAINEES

GROUP	PASS (N)	MEAN	(SD)	FAIL (N)	MEAN	(SD)	TOTAL (N)	MEAN	(SD)
Men	(313)	255.3	(22.4)	(176)	236.7	(27.4)	(495)	248.3	(25.8)
Women	(36)	248.8	(31.6)	(21)	233.3	(30.5)	(68)	241.6	(27.1)
Black	(5)	236.2	(15.4)	(9)	203.6	(18.7)	(15)	214.3	(22.9)
Hispanic	(2)	267.5	(10.6)	(2)	228.6	(6.2)	(6)	247.8	(23.7)
Asian	(1)	293.0	(—)	(—)	—	(—)	(1)	233.0	(—)
Other	(356)	254.9	(22.6)	(188)	237.6	(27.4)	(537)	248.3	(25.9)
Total Group	(—)	—	(—)	(—)	—	(—)	(552)	247.6	(26.1)

PREDEVELOPMENTAL ATC TRAINEES

GROUP	PASS (N)	MEAN	(SD)	FAIL (N)	MEAN	(SD)	TOTAL (N)	MEAN	(SD)
Men	(7)	242.4	(26.8)	(13)	200.6	(27.2)	(20)	215.1	(33.0)
Women	(8)	250.5	(23.9)	(11)	204.0	(27.0)	(19)	218.9	(31.3)
Black	(7)	262.7	(16.9)	(17)	196.2	(25.4)	(24)	209.7	(31.5)
Hispanic	(3)	221.0	(7.1)	(1)	207.0	(—)	(3)	217.0	(10.0)
Asian	(—)	—	(—)	(—)	—	(—)	(—)	—	(—)
Other	(7)	260.7	(30.9)	(6)	217.0	(29.3)	(13)	234.5	(33.0)
Total Group	(—)	—	(—)	(—)	—	(—)	(40)	218.3	(32.5)

112

It should be pointed out that the total of passed and failed trainees in some cases do not equal the number shown for the total group. For example, the 320 men who passed and the 189 who failed is 509 compared to 515 shown for the total group. This is a result of having 6 trainees in the total group for whom information by sex was not available.

The means and standard deviations for this same group of ATC trainees were derived, including OKT. The total weighted test battery score was established as follows:

2(MCAT 1) + 2(MCAT 2) + 1(CSC 157) + 1(CSC 24) + 1(OKT)

Since the OKT is not intended for use in determining if applicants pass the ATC test but rather as an alternative to the present OPM Rating Guide for establishing extra credit for aviation-related knowledge, it was given a weight of "1" in deriving the weighted total. The results are provided in Table 92.

TABLE 92
EFFECT OF OKT WEIGHTS ON VARIOUS SAMPLES

TOTAL ATC TRAINEES

GROUP	PASS (N)	MEAN	SD	FAIL (N)	MEAN	SD	TOTAL (N)	MEAN	SD
Men	(320)	321.7	(25.6)	(189)	294.0	(29.8)	(515)	311.3	(30.3)
Women	(34)	307.3	(28.5)	(32)	275.2	(33.3)	(67)	291.4	(34.5)
Black	(12)	310.2	(18.0)	(26)	261.8	(28.7)	(39)	277.1	(33.8)
Hispanic	(4)	312.0	(34.3)	(3)	272.3	(11.0)	(7)	295.0	(32.8)
Asian	(1)	319.0	(--)	(-)	--	(--)	(1)	319.0	(--)
Other	(345)	320.5	(26.3)	(194)	295.2	(29.4)	(545)	311.3	(30.0)
Total Group	(-)	--	(--)	(-)	--	(--)	(592)	309.0	(31.4)

NON-PREDEVELOPMENTAL ATC TRAINEES

GROUP	PASS (N)	MEAN	SD	FAIL (N)	MEAN	SD	TOTAL (N)	MEAN	SD
Men	(313)	321.6	(25.6)	(176)	295.9	(28.6)	(495)	312.2	(29.4)
Women	(26)	308.9	(29.0)	(21)	278.4	(34.4)	(48)	294.7	(34.6)
Black	(5)	302.0	(9.2)	(9)	259.9	(18.2)	(15)	275.0	(24.8)
Hispanic	(2)	321.0	(55.2)	(2)	272.0	(15.6)	(4)	296.5	(43.5)
Asian	(1)	319.0	(--)	(-)	--	(--)	(1)	319.0	(--)
Other	(388)	320.6	(26.0)	(188)	295.6	(29.4)	(532)	311.5	(29.8)
Total Group	(-)	--	(--)	(-)	--	(--)	(552)	310.5	(30.3)

PREDEVELOPMENTAL ATC TRAINEES

GROUP	PASS (N)	MEAN	SD	FAIL (N)	MEAN	SD	TOTAL (N)	MEAN	SD
Men	(7)	324.4	(27.8)	(13)	268.7	(34.5)	(20)	286.2	(41.7)
Women	(8)	302.3	(28.1)	(11)	269.0	(31.6)	(19)	283.0	(33.9)
Black	(7)	316.0	(21.3)	(17)	262.9	(33.4)	(24)	278.4	(38.9)
Hispanic	(2)	303.0	(12.7)	(1)	273.0	(--)	(3)	293.0	(19.5)
Asian	(-)	--	(--)	(-)	--	(--)	(-)	--	(--)
Other	(7)	313.4	(39.6)	(6)	285.2	(29.0)	(13)	301.5	(37.1)
Total Group	(-)	--	(--)	(-)	--	(--)	(40)	287.0	(38.1)

The difference in test performance between those who pass and those who fail are consistent, for the total group and for the subgroups. For example, the mean score (Table 92) for the pass group is consistently above 300 while the mean scores for those who failed is consistently below 300. Generally there is a separation of about one standard deviation between the mean scores of those who pass and fail for each of the subgroups. For the total ATC sample, the smallest difference in mean scores is between the "Other" fail group (295.2) and the "Women" pass group (307.3).

This pattern generally holds for the weighted test battery which excludes OKT (Table 91). The most evident exceptions are the mean scores for the "Men" and "Other" non-predevelopmental "fail" groups compared to the "Black" pass group. However, the number in the Black group is small (5) and the mean difference slight, especially when compared to difference in the standard deviations.

Intercorrelations and Multiple Regression Analysis. The next step was to relate test results on parallel forms of MCAT and OKT together with CSC 24 and CSC 157 to the ATC Academy pass or fail/withdraw criterion.

Table 93 presents the intercorrelation of these tests for the total ATC sample and for each subgroup by sex and race. The correlation of particular interest are those between the tests and the criterion. These are zero order validity coefficients.

114

TABLE 93
INTERCORRELATIONS-POTENTIAL TEST BATTERY VARIABLES

TOTAL ATC TRAINEE SAMPLE

(Minimum N=585 -- Criterion and CSC 24)

Test	MCAT 1 (N=953)	MCAT 2 (N=953)	OKT (N=953)	CSC-24 (N=592)	CSC-157 (N=592)	Criterion (N=585)
MCAT 1	-	.60	.08	.29	.21	.35
MCAT 2		-	.12	.20	.20	.37
OKT			-	-.24	-.14	.25
CSC 24				-	.10	.10
CSC 157					-	.18
Criterion						-

MEN ATC TRAINEE SAMPLE

(Minimum N=509 -- Criterion and CSC 24)

Test	MCAT 1 (N=800)	MCAT 2 (N=800)	OKT (N=800)	CSC 24 (N=515)	CSC 157 (N=515)	Criterion (N=509)
MCAT 1	-	.57	.05	.32	.22	.34
MCAT 2		-	.08	.25	.21	.37
OKT			-	-.23	-.15	.21
CSC 24				-	.10	.12
CSC 157					-	.16
Criterion						-

WOMEN ATC TRAINEE SAMPLE

(Minimum N=66 -- Criterion and CSC 24)

Test	MCAT 1 (N=141)	MCAT 2 (N=141)	OKT (N=141)	CSC 24 (N=67)	CSC 157 (N=6?)	Criterion (N=66)
MCAT 1	-	.69	.11	.12	.?2	.37
MCAT 2		-	.17	-.07	.19	.38
OKT			-	-.25	.07	.31
CSC 24				-	.13	.08
CSC 157					-	.40
Criterion						-

OTHER (WHITE) ATC TRAINEE SAMPLE

(Minimum N=539 -- Criterion and CSC 24)

Test	MCAT 1 (N=839)	MCAT 2 (N=839)	OKT (N=839)	CSC 24 (N=545)	CSC 157 (N=545)	Criterion (N=539)
MCAT 1	-	.55	.10	.28	.21	.29
MCAT 2		-	.13	.19	.20	.33
OKT			-	-.24	-.13	.23
CSC 24				-	.10	.07
CSC 157					-	.17
Criterion						-

BLACK ATC TRAINEE SAMPLE

(Minimum N=38 -- Criterion and CSC 24)

Test	MCAT 1 (N=81)	MCAT 2 (N=81)	OKT (N=81)	CSC 24 (N=39)	CSC 157 (N=39)	Criterion (N=38)
MCAT 1	-	.63	.33	.15	-.29	.56
MCAT 2		-	.22	.06	-.24	.45
OKT			-	-.10	-.12	.24
CSC 24				-	-.07	.35
CSC 157					-	.07
Criterion						-

It should be pointed out that these validity coefficients are for a restricted population. Correction for restriction in range was not a concern in this analysis since the objective was to compute regression equations and determine the relative contribution of these variables for predicting course pass/fail. However, Table 94 provides a comparison of these validity coefficients with the restricted and unrestricted coefficients obtained for the prior ATC sample of 1,827 trainees who attended the FAA Academy during the period January 1976 – April 1978.

TABLE 94
VALIDITIES – POTENTIAL TEST BATTERY VARIABLES

| | ATC Trainees* Jan. 1976 – April 1978 | | ATC Trainees** June 1978 – Dec. 1978 |
	Restricted r	Corrected r	Restricted r
MCAT Exp. (Z scores)	.28	.53	-
MCAT 1 (Raw scores)	-	-	.35
MCAT 2 (Raw scores)	-	-	.37
CSC 24	.10	.34	.10
CSC 157	.07	.40	.18
OKT	.22	-	.25

* Criterion = ATC Lab. Average Score (Z scores. See Tables 25 and 26)
** Criterion = Academy Training Pass/Fail (Lab. Training)

The restricted correlations were next used in multiple regression analysis. The first regression analysis employed MCAT 1, MCAT 2, CSC 24 and CSC 157 as predictor variables. OKT was excluded since it is not intended as a screening test for applicants. Each of the four tests were included in the regression analysis in order to ascertain its contribution to the multiple correlation coefficient (R). The results are shown in Table 95.

TABLE 95
REGRESSION ANALYSIS – FOUR PREDICTOR VARIABLES

Test	TOTAL ATC SAMPLE (N=585) R	R^2	Beta	MEN (N=509) R	R^2	Beta	WOMEN (N=66) R	R^2	Beta
MCAT 1	.348	.121	.1832	.337	.114	.1800	.375	.140	.1430
MCAT 2	.405	.164	.2470	.400	.160	.2554	.412	.170	.2259
CSC 24	.405	.164	-.0109	.400	.160	-.0121	.419	.175	.0355
CSC 157	.416	.173	.0997	.406	.165	.0691	.523	.273	.3240

Test	OTHER (WHITE) (N=538) R	R^2	Beta	BLACK (N=38) R	R^2	Beta
MCAT 1	.292	.085	.1538	.565	.319	.4734
MCAT 2	.355	.126	.2294	.579	.335	.2016
CSC 24	.355	.126	-.0270	.640	.409	.2836
CSC 157	.366	.135	.0982	.689	.475	.2692

116

Of the four tests, CSC 24 and CSC 157 contributed to the multiple correlation coefficient for the "Women" or "Black" groups but made little or no contribution for the total group as shown below:

Contribution to R

Group	CSC 157	CSC 24
Men	.006	--
Women	.104	.007
Black	.049	.061
Total	.011	--

Other regression analyses were conducted for the total group which excluded CSC 24; the multiple correlation and beta weights obtained were essentially the same as those derived when CSC 24 was included.

In terms of the total group, the findings for CSC 24 are consistent with results for other ATC trainee samples and ATC applicant groups. For the sample of 1,827 ATC students (1976-1978) previously presented, the CSC 24 contributed least to the multiple correlation and had the smallest beta weights (Tables 27, 28, and 31). For the ATC applicant group (1976-1977), CSC 24 contributed least to passing the current OPM ATC test (Table 19).

However, since CSC 24 and CSC 157 do contribute to the prediction of the criterion for some of the groups, the decision was to include them in the test battery for further analysis. Consequently, the weighted test battery used in developing descriptive statistics and for examination of the tests consisted of:

2(MCAT 1) + 2(MCAT 2) + 1(CSC 157) + 1(CSC 24)

Table 96 provides the results of the multiple regression analysis for the four tests plus OKT. In this case, the order in which each test entered the regression analysis was determined by its contribution to the multiple correlation coefficient. As shown for each group, the tests entered in a different order, especially in the case of "Women" and "Blacks."

TABLE 96
REGRESSION ANALYSIS—PREDICTOR VARIABLES AND OKT

TOTAL ATC SAMPLE (N=585)

Test	R	R^2	Beta
MCAT 2	.374	.140	.2134
OKT	.429	.184	.2478
MCAT 1	.455	.207	.1547
CSC 157	.474	.225	.1398
CSC 24	.477	.228	.0588

MEN (N=509)

Test	R	R^2	Beta
MCAT 2	.369	.136	.2264
OKT	.421	.177	.2341
MCAT 1	.448	.200	.1539
CSC 157	.460	.211	.1109
CSC 24	.462	.214	.0530

WOMEN (N=66)

Test	R	R^2	Beta
CSC 157	.403	.162	.2997
OKT	.531	.282	.3352
MCAT 1	.592	.351	.1294
MCAT 2	.603	.363	.1873
CSC 24	.613	.376	.1199

OTHER (WHITE) (N=539)

Test	R	R^2	Beta
MCAT 2	.328	.108	.1924
OKT	.398	.158	.2583
CSC 157	.424	.180	.1397
MCAT 1	.438	.192	.1193
CSC 24	.440	.194	.0476

BLACK (N=38)

Test	R	R^2	Beta
MGAT 1	.565	.319	.3573
CSC 24	.625	.391	.3360
OKT	.698	.488	.3415
CSC 157	.744	.553	.2791
MCAT 2	.7	.576	.1974

117

Table 97 shows, in summary, the multiple correlation coefficient (R) values derived by group and the variance (R^2) accounted for in the pass or fail/withdraw criterion based on the restricted zero-order correlation coefficients (r).

TABLE 97
SUMMARY VALUES - REGRESSION ANALYSES

Group	Test Battery		Test Battery And OKT	
	R	R^2	R	R^2
Total Sample	.416	.173	.477	.228
Other	.368	.135	.440	.194
Men	.406	.165	.462	.214
Women	.523	.273	.613	.376
Black	.689	.475	.759	.579

Test Utility. The purpose of this analysis was to examine the practical utility of the new ATC test battery, relate the findings to the information obtained for the 1978 ATC Applicant Group and estimate the potential FAA training investment costs which could be saved by using the new ATC test for selection of applicants.

As previously noted, the weighted test scores for this sample of 585 ATC trainees was based on:

$$WT_1 = 2(MCAT-1) + 2(MCAT-2) + 1(CSC\ 157) + 1(CSC\ 24)$$

However, for the 1978 ATC Applicant Group, only one parallel form of MCAT was administered rather than the two which were given to this group of ATC trainees. In order to be able to relate the two groups, the weighted ATC test scores for the 585 ATC trainees were recomputed using only the first MCAT administered and then weighting each of the tests on the same basis as was used in the analysis of the 1978 ATC Applicant Group. The weighting used was:

$$WT_2 = 4(MCAT-1) + 2(CSC\ 157) + 1(CSC\ 24)$$

The total raw weighted scores were then transmuted for each trainee based on:

$$T_s = 10.68(\frac{RWS - 222.27}{56.22}) + 70$$

as was shown with Table 77. Again, a transmuted score of 70 or above was equated to "passing" the new ATC test for this ATC trainee group.

118

In order to assess the difference between these two weighting factors (WT_1 and WT_2), correlation coefficients were computed between the two weighted scores on the ATC test battery and between these and the scores on the present OPM test battery which were also available for the ATC trainees. Table 98 provides the descriptive statistics and correlations.

TABLE 98

DESCRIPTIVE STATISTICS AND INTERCORRELATIONS, ATC-OPM TESTS

Test	N	Mean	SD
OPM Score (Raw)	591	262.0	23.31
ATC Test (WT_1 Raw)	592	245.6	27.57
ATC Test (WT_2 Raw)	592	272.6	35.1

	OPM Score	ATC Score WT_1	ATC Score WT_2
OPM Score	-	.62	.65
ATC Score - WT_1		-	.93
ATC Score - WT_2			-

Since the weighting used to examine the utility of the new ATC test battery (WT_2) has a similar higher correlation coefficient with the OPM test scores (.65 compared to .62), the results should tend to be similar.

The mean transmuted score on the present OPM test battery for this group of ATC trainees was 87.2 with a standard deviation of 7.6. Their mean transmuted score on the new ATC test battery (WT_2) was 79.6 with a standard deviation of 6.7. This lower mean score is consistent with the findings derived from the 1978 ATC Applicant Group which demonstrated that applicants who passed both the present OPM test and the new ATC test, generally score lower on the ATC test (see Figure 13).

For this sample of 585 ATC trainees, Table 99 compares the score distribution for the present OPM test battery and the new ATC test battery for the total sample and for the Predevelopmental (non-competitive applicants) and Non-predevelopmental (competitive applicants) groups. These are transmuted test scores without Veterans Preference or aviation-related knowledge/experience credit. The 585 ATC trainees for whom complete test and criterion data were available represent 61 percent of the 953 attending the Academy during June through December 1978. The 40 Predevelopmental trainees represents 36 percent of the non-competitive hires; the 545 Non-predevelopmental trainees represent 65 percent of the competitive hires.

119

TABLE 99

SCORE DISTRIBUTIONS-585 TRAINEES

TOTAL ATC TRAINEES

Score Range	Present OPM Test N	Present OPM Test %	New ATC Test (WT2) N	New ATC Test (WT2) %
95+	121	20.7%	2	.3%
90-94	125	21.4%	27	4.6%
85-89	133	22.6%	98	16.8%
80-84	110	18.8%	161	27.6%
75-79	63	10.8%	147	25.1%
70-74	33	5.7%	102	17.4%
"Pass" Test	585	100.0%	537	91.8%
65-69	-	-	39	
60-64	-	-	8	8.2%
55-59	-	-	1	
"Fail" Test	0	-	48	
	585	100.0%	585	100.0%
Mean	87.2		79.6	
SD	7.6		6.7	

Score Range	NONPRED. TRAINEES (Competitive Hires) Present OPM Test N	Present OPM Test %	New ATC Test (WT2) N	New ATC Test (WT2) %	PREDEV. TRAINEES (Noncompetitive Hires) Present OPM Test N	Present OPM Test %	New ATC Test (WT2) N	New ATC Test (WT2) %
95+	119	21.8%	2	.4%	2	4.6	0	
90-94	122	22.4%	27	5.0%	3	7.0	0	
85-89	129	23.7%	97	17.8%	4	9.3	1	2.5%
80-84	101	18.5%	157	28.8%	9	23.3	4	10.0%
75-79	48	8.8%	136	24.9%	15	37.2	11	27.5%
70-74	26	4.8%	90	16.5%	7	18.6	12	30.0%
"Pass" Tests	545	100.0%	509	93.4%	40	100.0%	28	70.0%
65-69	-	-	31		-	-	8	
60-64	-	-	5	6.6%	-	-	3	30.0%
55-59	-	-	0		-	-	1	
"Fail" Test	0	-	36		0	-	12	
	545	100.0%	545	100.0%	40	100.0%	40	100.0%

Two significant points are evident from the data in Table 99.

First, 48 (8.2 percent) of the 585 ATC trainees, all of whom passed the present OPM test for employment eligibility, would not have passed the new ATC test and therefore would not have been eligible for appointment as an air traffic controller trainee. When the total group is further identified by competitive hires from established OPM certificates (Non-predevelopmental) and non-competitive hires (Predevelopmental) who were hired through alternate recruitment procedures, other differences are apparent: (1) 30 percent of the Predevelopmental Group would not have been eligible for appointment in contrast to 6.6 percent of the Non-predevelopmental hires (this 6.6 percent compares to 7.2 percent of the 1978 ATC Applicant Group who also passed the OPM test but did not pass the new weighted (WT2) ATC test; and (2) 23 percent of the Non-predevelopmental trainees scored 85 or above on the new ATC test in contrast to 2.5 percent of the Predevelopmental group. These data provide some further quantative insight into the descriptive statistics on the mean score differences between these two groups which were identified in Tables 85, 86, 91, and 92. It should be pointed out that the trainees in this sample could not score lower than 70 on the OKT and thus had to have higher scores.

Second, the score distribution of these ATC trainees on the new ATC test is markedly lower than on the present OPM ATC test. About 42 percent of the total trainee group scored 90 or above on the present OPM test in contrast to only 5 percent who scored this high on the new ATC test. In the Non-predevelopmental group, 44 percent scored 90 or more on the OPM test compared to 5.4 percent on the new ATC test. Of the Predevelopmental trainees, none scored above 89 on the new test but almost 12 percent did on the OPM test.

These findings are consistent with the results obtained for the 1978 ATC applicants where passing scores on the experimental ATC test were significantly lower than on the OPM test.

Given this marked difference in score distribution between the OPM and the new ATC test, the next step was to relate the scores on the present OPM and new ATC test to the pass or fail/withdraw criterion. Test scores and criterion data were available for 695 trainees for the OPM test and 585 trainees for the new ATC test (WT2). The results are provided in Table 100. Again, these scores do not include extra credit for veterans preference or aviation-related experience or knowledge.

TABLE 100

DISTRIBUTION COMPARISONS–OPM AND NEW ATC TESTS

| | ATC TRAINEES OPM TEST | | | | ATC TRAINEES NEW ATC TEST (WT2) | | | |
| | Total | | ATC Trng Fail/WD | Fail/WD Rate | Total | | ATC Trng Fail/WD | Fail/WD Rate |
Range	N	(%)	N	%	N	(%)	N	%
95+	145	(20.0%)	33	22.8%	2	(.3%)	0	
90-94	148	(21.3%)	47	31.8%	27	(4.6%)	5	18.5%
85-89	156	(22.5%)	62	39.7%	98	(16.8%)	20	20.4%
80-84	131	(18.8%)	59	45.0%	164	(27.6%)	46	28.0%
75-79	76	(10.9%	38	50.0%	147	(25.1%)	62	42.2%
70-74	39	(5.6%)	22	56.4%	102	(17.4%)	52	51.0%
Below 70	0	-	0	--	48	(8.2%)	38	79.2%
	695	100.0%	261	37.6%	585	100.0%	223	38.1%

Table 100 clearly identifies the difference between the present OPM test and the new ATC test with regard to the percentage of trainees who successfully complete the ATC training in each of the ATC test score ranges. For example, with scores between 90-94 almost 32 percent of the OPM test group fail or withdraw in comparison to only about 19 percent of the group with test scores in this range on the new ATC test. Of particular interest is the fact that of the 48 ATC trainees who scored below 70 on the new ATC test, 79 percent did not successfully complete the training.

Table 101 provides the score range distribution and failures/withdrawal rate for the 585 trainees in relation to the new ATC test identified by Non-predevelopmental and Predevelopmental ATC trainees.

TABLE 101

FAILURE/WITHDRAWAL RATE DISTRIBUTIONS BY DEVELOPMENTAL CATEGORY

New ATC Test Score Range	NONPRED. ATC TRAINEES				PREDEV. ATC TRAINEES			
	Total N	Fail/WD N	%	Fail/WD Rate %	Total N	Fail/WD K	%	Fail/WD Rate %
95+	2	0	0%	0%	0	0	-	0%
90-94	27	5	2.5%	18.5%	0	0	-	0%
85-89	97	20	10.1%	20.6%	1	0	-	0%
80-84	157	46	23.1%	29.3%	4	0	-	0%
75-79	136	57	28.6%	41.9%	11	5	20.8%	45.5%
70-74	90	43	21.6%	47.8%	12	9	37.5%	75.0%
Below 70	36	28	14.1%	77.8%	12	10	41.7%	83.3%
Total	545	199	100.0%	36.5%	40	24	100.0%	60.0%

These data focus attention on the 48 trainees who did not pass the new ATC test battery and who would not have been eligible for appointment if selections had been based on the new test. Of the 48, a total of 38 either failed or withdrew from the training; a loss rate of 79 percent. For Non-predevelopmental trainees the loss rate for this group was 78 percent and for Predevelopmentals it was 83 percent.

This has a direct and important impact on both recruitment for the air traffic control occupation and the cost of training for FAA. Had these 48 trainees not been hired, the overall loss rate would be reduced from 38.1 percent to 34.4 percent and by Non-predevelopmental and Predevelopmental groups, from 36.5 percent to 33.6 percent and from 60 percent to 50 percent respectively.

For each Predevelopmental trainee, FAA invests about $28,000 in salary and training costs and about $10,000 in each Non-predevelopmental student by the time their FAA Academy ATC training program is completed. For each student who fails or withdraws from training, this investment is lost and a new trainee has to be hired.

If the 38 trainees had not been hired based on their failure to obtain a minimum score of 70 on the new ATC test, FAA's loss in training investment costs would have been reduced by $560,000. Extending this for 2,000 new trainees entering Academy training on an annual basis would result in a cost avoidance of about $1,350,000 each year.

This projection of the cost avoidance utility of the new ATC test battery addresses only the benefit obtained if ATC applicants who scored below a minimum of 70 on the new ATC test battery were not hired. The other significant area where benefits could be obtained is to determine the difference in the group of ATC applicants who would be hired under competitive selection procedures with the current OPM test and the new ATC test with allowance of extra credit for veterans preference and aviation-related experience or knowledge.

122

In order to examine these differences, the OPM Earned Rating which is derived from the ATC trainee's score on the present OPM ATC test battery plus additional credit for veterans preference (0 to 10 points) and credit for aviation-related experience provided by the OPM Rating Guide (0-15 points), was available for 621 of the 843 Non-predevelopmental ATC trainees in this sample. The 110 Pre-developmental ATC trainees who were appointed non-competitively were not included in this analysis since extra credit for veterans preference or aviation-related experience is not a factor in their appointment eligibility; they need only to pass the OPM test battery with a score of 70 or above. The "equivalent" of the OPM Earned Rating using the new ATC test battery (WT2) and the ATC Occupational Knowledge Test (OKT) was derived for the Non-predevelopmental (competitive hires) ATC trainees. The 36 trainees who did not achieve a minimum score of 70 on the new ATC test were excluded from the analysis since they would not have been eligible for appointment. The "Earned Rating" for these trainees was based on their score on the new ATC test plus additional credit provided for their OKT scores as follows:

OKT Scores	Additional Points
80+	15
75-79	10
70-74	5
65-69	3

The difference between the OPM transmuted score and the OPM Earned Rating for each trainee was the additional credit given for veterans preference and aviation experience based on the OPM Rating Guide. Since the veterans preference points were not separately identified in the total of extra credit given in the OPM Earned Rating, it was not possible to specifically identify those ATC trainees who were 5-point or 10-point veterans.

In order to derive an "equivalent" Earned Rating with the new ATC test which would reflect extra credit for veterans preference and aviation-related knowledge based on OKT scores, the following procedures were used:

	OPM EARNED RATING	(Mean = 95.1; N=540)
MINUS	OPM TRANSMUTED SCORE	(Mean = 87.2; N-591)
=	TOTAL OPM EXTRA CREDIT (V_1)	(AVGV_1 = 7.9 points)

IF:	THEN:
V_1 is less than 5	VET points = 0 OKT credit = 0, 3, 5, 10, 15
V_1 = 5	VET points = 5 OKT credit = 0, 3, 5, 10, 15
V_1 = 10, 15, or 20	VET points = 5 OKT credit = 0, 3, 5, 10, 15
V_1 is more than 20	VET points = 10 OKT credit = 0, 3, 5, 10, 15

The net effect of this procedure would be a somewhat higher distribution of "Earned Rating" scores for the new ATC test than would be the case if specific veterans credit had been available for each ATC trainee.

Given these conditions, an estimated total Earned Rating could be computed based on the new ATC test for 486 of the ATC trainees appointed competitively (Non-predevelopmental) from OPM registers. The 48 students who scored below 70 on the new ATC Test were excluded from the sample. Table 102 compares the distribution of these Earned Ratings to the Earned Ratings for 621 trainees based on the present OPM ATC test battery and Rating Guide. The 486 trainees for whom estimated Earned Ratings could be computed using the new ATC test is a subsample of the 621 trainees who had OPM Earned Ratings available.

TABLE 102

NON-PREDEVELOPMENTAL ATC TRAINEES
(Competitive Hires)

Earned Rating Score Range	OPM Earned Rating N	(%)	Est. New ATC Earned Rating N	(%)
100+	205	(33.0%)	54	(11.1%)
90-99	314	(50.6%)	143	(29.4%)
80-89	99	(15.9%)	229	(47.1%)
70-79	3	(.5%)	60	(12.6%)
Total	621	(100.0%)	486	(100.0%)

About 84 percent of the 621 ATC trainees had OPM Earned Ratings of 90 or more. Since competitive selection from OPM registers requires hiring those applicants with the highest Earned Ratings first, this distribution of OPM Earned Ratings would be expected. However, the Earned Rating distribution based on the new ATC test for the 486 trainees who are part of the sample of 621, is quite different.

Only 41 percent of the trainees have Earned Ratings of 90 or above and almost 13 percent are below 80 as compared to less than 1 percent for the OPM Earned Rating. Consequently, many of the ATC trainees with Earned Ratings below 90 on the new ATC exam would not have been hired competitively since other applicants with higher ratings would have been selected in their place.

In terms of practical utility of the new ATC test battery and OKT, these differences in Earned Rating score distribution are of significance only if the pass or fail/withdraw rates also differ between the OPM and new ATC Earned Rating. These differences are provided in Table 103.

TABLE 103

NON-PREDEVELOPMENTAL ATC TRAINEES
(Competitive Hires)

Score Range	OPM Earned Rating Total N	Fail/WD Rate N	%	Est. New ATC Earned Rating Total N	Fail/WD Rate N	%	Diff. %
100+	205	54	26.3%	54	6	11.1%	-15.2%
90-99	314	122	38.9%	143	33	23.1%	-15.8%
80-89	99	47	47.5%	229	86	37.6%	- 9.9%
70-79	3	2	66.6%	60	38	66.3%	- .3%
Total	621	225	36.2%	486	163	33.5%	

124

It is clear from the data in Table 103 that the probability of successfully completing the ATC training program is greater in each of the Earned Rating score ranges above 79 for the new ATC test battery.

It is evident that an OPM register based on the new ATC Earned Rating would result in hiring a somewhat different group of ATC trainees than was in fact hired with the present OPM register. It is possible to relate the data in Tables 102 and 103 to the results obtained for the 1978 ATC Applicant Group on both the current OPM test and the new ATC test. This information would provide a basis for assessing the differences in selection of groups of applicants and the potential effect of these selections on fail/withdraw rates and on training investment losses.

The previous analysis of the 1978 ATC Applicant Group identified (Table 76) the Earned Rating score distribution of applicants who passed both the present OPM test and the new weighted ATC test battery (WT_2). If this group of 621 ATC trainees had been hired from the group of 1978 ATC applicants based on Earned Ratings obtained on the present OPM exam, the relationship would be as shown in Table 104. For the 1978 ATC Applicant Group, the number in each Earned Rating category in effect represents the OPM register from which applicant selections would be made.

TABLE 104

PROJECTED EARNED RATINGS - 1978 ATC TRAINEES

OPM Earned Rating Score Range	OPM Earned Rating Qualified 1978 ATC Applicants		OPM Earned Rating 1978 ATC Trainees (Competitive Hires)		"Selection" Ratio
	N	(%)	N	(%)	(%)
100+	229	(7.5%)	205	(33.0%)	(89.5%)
90-99	644	(21.2%)	314	(50.6%)	(48.8%)
80-89	1086	(35.8%)	99	(15.9%)	(9.1%)
70-79	1076	(35.5%)	3	(.5%)	(.3%)
Total	3035	(100.0%)	621	(100.0%)	(20.5%)

The registers maintained by OPM for ATC applicants are by separate employing jurisdictions. Consequently, each of the eleven FAA regions selects from its own certificate of applicants based on the geographic preference identified by the applicant and the Earned Rating. As a result, one region (Southern, for example) may be able to fill all of their ATC recruitment needs from applicants who have Earned Ratings above 100. Other regions (Alaska or Great Lakes, for example) may recruit a number of applicants with Earned Ratings in the 80's or even 70's. However, Table 104 shows that essentially all of the 621 ATC trainees scored 80 or above and, with a mean OPM Earned Rating of 94.9 for the total group, selections were made in most cases from the "top" of the registers with scores of 90 or more. The selection ratio of 89.5 percent between the 205 1978 trainees and the 229 1978 applicants who would have been available with Earned Ratings of 100 or more, is somewhat

125

higher than typically experienced by FAA regions. Generally, of those
applicants offered ATC positions, about 80 percent are selected; the other
20 percent decline or are dropped from consideration because of medical
or security reasons.

The next step was to project the composition for a group of 621 ATC trainees
by their Earned Rating if they had been selected from a competitive register
derived from the new weighted ATC test exam including extra credit for veterans
preference and OKT scores.

The number of 1978 ATC applicants in each Earned Rating score group derived
from the new ATC exam was used as the "selection register" as shown in
Table 105. This is the same distribution as previously provided in Table 76.
A maximum selection ratio of 80 percent in each Earned Rating group was
used as more representative of actual selection experience. Since fewer
applicants have Earned Ratings of 90 or more based on the new ATC exam, more
selections would be made from among applicants with Earned Ratings in the
80-99 range. Based on these factors, the projected distribution of the 621
ATC trainees based on the new ATC exam would be as indicated in Table 105.

TABLE 105

PROJECTED EARNED RATING DISTRIBUTIONS, 1978 ATC APPLICANTS

New ATC Earned Rating Score Range	Qualified 1978 ATC Applicants N	(%)	Est. 1978 ATC Trainees N	(%)	Selection Ratio (%)
100+	126	(4.0%)	100	(16.1%)	(80%)
90-99	402	(12.9%)	322	(51.9%)	(80%)
80-89	1162	(37.3%)	192	(30.9%)	(17%)
70-79	1426	(45.8%)	7	(1.1%)	(-)
Total	3116	(100.0%)	621	(100.0%)	(20%)

With the OPM Earned Rating distribution for the 621 ATC trainees and a pro-
jected Earned Rating distribution based on the new ATC exam, the fail/withdraw
rates provided in Table 88 can be applied to each group to estimate the difference
in loss rates between them as provided in Table 106.

TABLE 106

DIFFERENCES IN LOSS RATE ESTIMATES BY TEST EXAMINATION

OPM Earned Rating Scores	PRESENT OPM ATC EXAM 1978 ATC Trainees N	(%)	Fail/WD Rate N	(%)	New ATC Earned Rating Scores	NEW ATC EXAM Est 1978 ATC Trainees N	(%)	Est. Fail/WD Rate N	(%)
100+	205	(33.0%)	54	(26.3%)	100+	100	(16.1%)	11	(11.1%)
90-99	314	(50.6%)	122	(38.9%)	90-99	322	(51.9%)	74	(23.1%)
80-89	99	(15.9%)	47	(47.5%)	80-89	192	(30.9%)	72	(37.6%)
70-79	3	(.5%)	2	(66.6%)	70-79	7	(1.1%)	5	(66.6%)
Total	621	(100.0%)	225	(36.2%)	Total	621	(100.0%)	162	(26.1%)

126

Table[106] demonstrates the potential utility of the new ATC exam in terms of an overall reduction of the fail/withdraw rate of about 10 percent (from 36.2 to 26.1) for this sample of 621 ATC trainees.

With respect to FAA's training investment loss, projecting this reduction on an annual basis for 1,800 ATC applicants hired through competitive selection procedures would result in a cost avoidance of about $1,825,000 since significantly more of the trainees should successfully complete the ATC training program. By combining this with the $1,350,000 reduction in investment loss resulting from not hiring those applicants who pass the present OPM ATC test but fail the new ATC test, the potential cost avoidance which could be obtained is in the order of $3,000,000 a year.

UNIFORM GUIDELINES ON EMPLOYEE SELECTION REQUIREMENTS

The Uniform Guidelines (17) jointly established by the Equal Opportunity Commission (EEOC), Department of Labor (DOL), Department of Justice (DOJ), and the Office of Personnel Management (OPM), provides four basic requirements for selection procedures in relation to equal employment opportunity. Basically, these requirements are:

- to determine if a selection procedure has an adverse impact on employment opportunities of minorities or women

- if there is an adverse impact, the need to validate the selection procedure

- when empirical data demonstrates validity of the selection procedure (that is, it is predictive of or significantly correlated with important elements of job performance), the need to examine the fairness of the selection procedure

- consideration and investigation of alternative selection procedures. Where two or more selection procedures which are substantially equally valid for a given purpose, the procedure which has the lesser adverse impact should be used.

The various research studies which have been discussed encompassed a wide range of selection considerations and devices both of a cognitive and non-cognitive structure. The analytic methods applied combined new devices with present OPM selection devices and evaluated their interaction with various criteria of performance to identify those tests and procedures which provided the highest relationship between selection procedure scores and performance criteria which are functionally relevant to the air traffic control occupation. Based on this, the requirements of the Uniform Guidelines and other elements of EEO programs will be addressed in relation to the new ATC exam and selection procedures derived from the research.

Adverse Impact. For enforcement proceedings, the Uniform Guidelines state:

"A selection rate for any race, sex or ethnic group which is less than four-fifths (4/5) (or eighty percent) of the rate for the group with the highest rate will generally be regarded by the Federal enforcement agencies as evidence of adverse impact . ."

This definition was adopted as a practical "rule of thumb" and not a legal definition.

In the selection procedures for hiring applicants for the air traffic control occupation addressed by this research, adverse impact can occur at two points in the selection process. The first point is establishing basic eligibility for employment consideration by successfully passing the test battery. The second point is the Earned Rating score (including credit for veterans preference

128

and aviation-related experience or knowledge) which determines the competitive ranking of applicants who pass the written test.

The information base against which measures of adverse impact can be identified consist of the group of 5,976 applicants who took the present OPM and experimental ATC test battery in 1978. Ethnic, race or sex data has not been obtained by OPM for other groups of applicants.

For this group of 1978 ATC applicants, the relative selection (pass) rates by race and sex on the present OPM ATC test battery is provided in Table 107. Comparable data for the new ATC test (WT$_2$) is provided in Table 108. Both tables exclude any extra credit for veterans preference or aviation knowledge since the purpose is to evaluate the relative rates of basic eligibility status for the applicants. These selection rates identify the proportion of each sex and race group who passed the selection test battery with a score of 70 or more. Based on these rates, the 80 percent "rule of thumb" was applied to each group in relation to the majority group selected to derive the "adverse impact ratio" which would occur at this point in the selection process. When this ratio is less than 80 percent, the Uniform Guidelines state that this will generally be regarded as evidence of adverse impact. The Guidelines also point out that where the ratio is less than 80 percent, it may not constitute adverse impact where: (1) the differences are based on small numbers and are not statistically significant; or (2) where special recruiting or other programs cause the pool of minority or female candidates to be atypical of the normal group of applicants. Since these 1978 applicants include a "walk-in" group resulting from special recruiting efforts for minorities and women, the adverse impact analysis was done for the total group (including 664 "walk-in" applicants) and the "Scheduled Group" which would be representative of the normal pool of applicants.

TABLE 107

Adverse Impact Analysis
Present OPM Test Battery

Group	Total 1978 ATC Applicants	% Who Passed	Adverse Impact Ratio	Scheduled 1978 ATC Applicants	% Who Passed	Adverse Impact Ratio
	----------Total Group----------			-------Scheduled Group----------		
Men	4191	53%	---	3835	54%	---
Women	1785	45%	85%	1473	46%	85%
	5976			5308		
White	4067	64%	---	3775	63%	---
Hispanic	339	38%	(59%)	271	38%	(60%)
Black	1407	19%	(30%)	1116	18%	(29%)
Asian	57	61%	95%	50	56%	89%
Amer. Ind.	61	51%	80%	58	50%	(79%)
	5931			5300		

TABLE 108

Adverse Impact Analysis
New ATC Test Battery (WT$_2$)

Group	Total 1978 ATC Applicants	% Who Passed	Adverse Impact Ratio	Scheduled 1978 ATC Applicants	% Who Passed	Adverse Impact Ratio
	----------Total Group----------			---------Scheduled Group--------		
Men	4191	54%	---	3835	57%	---
Women	1785	43%	(77%)	1473	45%	(78%)
	5976			5308		
White	4067	66%	---	3775	67%	---
Hispanic	339	41%	(62%)	271	41%	(61%)
Black	1407	14%	(21%)	1116	14%	(21%)
Asian	57	61%	92%	50	56%	84%
Amer. Ind.	61	53%	80%	58	53%	(79%)
	5931			5300		

129

The data on the preceeding two tables show that including the "walk-in" applicant group has relatively little effect on the ratios derived. Generally, there is only a difference of one percent between the Total Group and the Scheduled Group except for Asians and here the number is quite small.

The present OPM ATC Test has an adverse impact on both the Hispanic and Black Groups. In the light of past FAA experience in hiring minorities from competitive OPM registers, this result is not unexpected but Table 107 provides a quantative measure of the degree of adverse impact for these two groups. Table 108 shows that the new ATC test also has some adverse impact on women applicants in addition to the Hispanic and Black Groups and that the degree of adverse impact on the Black Group is somewhat higher.

For American Indians, the addition of one or two more applicants to the "pass group" would result in a selection ratio of 80 percent or more on either the OPM or ATC test. Consequently, evidence of adverse impact based on this sample is not clear.

The analysis was also completed for those who passed the OPM and Weighted ATC test to determine if there was adverse impact in each of three score range groups. Candidates who pass the test battery are ranked on competitive OPM registers based on their total score and those with the highest scores selected first. This "Earned Rating", which includes additional points for veterans preference and OKT credit, was used for the analysis. The results are provided in Tables 109 (OPM Test) and 110 (Weighted ATC test).

TABLE 109

ADVERSE IMPACT ANALYSIS
PRESENT OPM TEST BATTERY
WITH VETERANS PREFERANCE AND EXPERIENCE CREDIT

	Total 1978 ATC Applicants Who Passed OPM Test	OPM Earned Rating ----70 - 79----			OPM Earned Rating ----80 - 89----			OPM Earned Rating ----90 + ----		
		N	Ratio	Adverse Impact	N	Ratio	Adverse Impact	N	Ratio	Adverse Impact
Men	2236	718	32%	---	796	36%	---	722	32%	---
Women	799	358	45%	No	290	36%	No	151	19%	Yes
	3035	1076			1086			873		
White	2536	831	33%	---	934	36%	---	791	31%	---
Hispanic	128	50	39%	No	47	37%	No	31	24%	?
Black	264	161	61%	No	77	29%	?	26	10%	Yes
Asian	35	12	34%	No	8	23%	Yes	15	43%	No
Amer. Ind.	31	10	32%	No	15	48%	No	6	20%	?
	3014	1064			1081			869		

130

TABLE 110

ADVERSE IMPACT ANALYSIS
NEW ATC TEST BATTERY (WT.)
WITH VETERANS PREFERENCE AND EXPERIENCE CREDIT

	Total 1978 ATC Applicants Who Passed New ATC Test	ATC Earned Rating 70 - 79			ATC Earned Rating 80 - 89			ATC Earned Rating 90 +		
		N	Ratio	Adverse Impact	N	Ratio	Adverse Impact	N	Ratio	Adverse Impact
Men	·2348	951	40%	---	905	39%	---	492	21%	---
Women	768	475	61%	No	257	34%	No	36	5%	Yes
	3116	1426			1162			528		
White	2698	1171	43%	---	1047	39%	---	480	18%	---
Hispanic	136	76	56%	No	39	29%	Yes	21	15%	?
Black	198	134	68%	No	50	25%	Yes	14	7%	Yes
Asian	35	16	46%	No	13	37%	No	6	17%	No
Amer. Ind.	32	19	59%	No	9	28%	?	4	13%	?
	3099	1416			1158			525		

While there is no evidence of adverse impact in the score ranges of 70-89 for women, both the OPM test and the ATC test have adverse impact in the score range of 90 or more. As identified earlier (Figure 8), a significant factor is the extra credit granted for veterans preference and, to a lesser extent, credit for aviation-related knowledge as measured by the Occupational Knowledge Test (OKT). It is also evident that the Weighted ATC test results in a greater adverse impact on women in the 90 and above score range than the present OPM test. This is primarily a result of the fact that raw test scores for all applicants are lower on the ATC test than the OPM test. However, relative to men, women score lower on the Weighted ATC test. As shown in Tables 59 and 74, women comprise 27 percent of the group with raw scores of 90 or more on the OPM test but only 12 percent of this group on the Weighted ATC test.

In comparing racial groups on the two tests, Tables 109 and 110 show a "?" in some of the score ranges. In these cases, the number of applicants is small and the addition or subtraction of one or two applicants in the score range would result in either meeting or not meeting the "80 percent rule of thumb". Consequently, evidence of adverse impact is not clear based on this sample of 1978 ATC applicants.

It is clear that there is an adverse impact on the Black Group on both the OPM and the ATC test in the score range of 90 or more. In the score range of 80-89 on the Weighted ATC test, and possibly on the OPM test, there is also an adverse impact. Again, this is primarily a result of the fact that as a group, Blacks who passed the test, score lower than the majority group.

In the case of the Hispanic Group, there is possibly adverse impact in the score range of 90 or more on the OPM and ATC test but not in the range of 80-89 for the OPM test.

131

In terms of competitive employment opportunities, based on the present OPM ATC test, between 80 and 85 percent of the appointments are generally made from among applicants with Earned Ratings of 90 or more. Consequently, employment prospects of applicants with scores in the 80-89 range would be characterized as "poor" to "fair". Based on the Weighted ATC test, however, about 40 percent of the applicants hired would have scores of 90 or more. Consequently, employment prospects of applicants with scores in the 80-89 range on the ATC test would be "fair" to "good" (Table 102).

When the score range of 80 or more on the Weighted ATC test was examined, there is an adverse impact for the women and Black Groups. For the Hispanic and American Indian Groups the numbers are small and the addition of one or two more applicants in the score range would meet the "80 percent" criterion. The results are provided in Table 111.

TABLE 111
ADVERSE IMPACT ANALYSIS
NEW ATC TEST BATTERY – SEX AND RACE

	Total 1978 ATC Applicants Who Passed New ATC Test	ATC Earned Rating 80 or More N	Ratio	Adverse Impact
Men	2348	1397	60%	---
Women	768	293	38%	Yes
	3116	1690		
White	2698	1527	57%	---
Hispanic	136	60	44%	?
Black	198	64	32%	Yes
Asian	35	19	54%	No
American Indian	32	13	41%	?
	3099	1683		

This analysis shows the Weighted ATC test has an adverse impact on women and some minority groups with respect to both initial eligibility (in passing the test) and by using the total scores, including extra credit for ranking success-ful candidates. In this situation, the Uniform Guidelines require evidence of the validity of test, the basis for establishing the cutoff pass score and the use of scores for rank ordering of candidates.

Test Validity, Pass Score and Candidate Ranking. The validity of the new ATC test battery as well as individual test components of the battery has been demonstrated in relation to a number of ATC performance criterion measures. In summary:

- 1977 ATC Study. For a sample of approximately 2,100 full performance, developmental and newly hired ATC specialists, MCAT, Directional Headings, Occupational Knowledge (OKT) and the Prior Experience Questionnaire (PEQ) provided significant

correlations (p $\leq$ .01) with the Aggregate ATC Success criterion for each of four ATC options (FSS, VFR, IFR and ARTCC) and for all options combined (see Table 5). In addition, mean test scores for each of the four ATC options were different at statistically significant levels of confidence (p $\leq$ .01) or (p $\leq$ .05) with the FSS option average score lowest and the ARTCC score highest (see Table 6).

- 1976-1978 ATC Study. For a sample of 1,827 ATC trainees, experimental forms of MCAT together with Abstract Reasoning (OPM 157) and Arithmetic Reasoning (OPM 24) provided a multiple correlation (corrected for range restriction) of .54 with the ATC laboratory problem average score criterion (see Table 28). Cross validation, using the weighted test scores derived from the multiple regression provided a multiple correlation value (R) of .5381 for one sample and .5292 for the other (see table 32). When test scores on the Occupational Knowledge Test (OKT) were combined with the experimental test battery (MCAT, OPM 157 and OPM 24), an estimated multiple correlation value of .60 was obtained with the increase significant at the .01 level of confidence.

- 1978 ATC Study. For a new sample of 585 ATC trainees, parallel forms of MCAT together with the present OPM tests 157 and 24 provided a multiple correlation value (R) of .42 for the total sample. The R values by sex and race ranged from .37 to .69 (see Table 95). The criteria used was the pass or fail/withdraw status of the trainee at the completion of ATC Laboratory training. Validity coefficients (r values) used in the multiple regression were not corrected for range restriction (see Table 94). When test scores on parallel forms of OKT were combined with the weighted test battery, the multiple correlation value for the total sample was .48 with R values by sex and race ranging from .44 to .76 (see Table 96. The increase in the variance (R^2) accounted for by the weighted test battery with OKT was significant at the .01 level of confidence for the total sample and for each race group. For women the increase was significant at the .05 level (see Table 97).

The utility of the new ATC test battery (as compared to the present OPM battery) was examined with respect to appointment eligibility (pass scores) and ranking of successful candidates based on total scores including credit for veterans preference and ATC-related knowledge.

To do this, the information available for the 1978 ATC trainee sample was related to the 1978 ATC applicant group. A common test battery, consisting of parallel forms of MCAT, OPM 157 and OPM 24 was administered to both groups. For the OKT, a 60-item test (101-C) was administered to the 1978 ATC applicant group and parallel forms of OKT 102 (80 item) administered to most of the 1978 ATC trainee sample. The common test battery was weighted:

4 x MCAT Total Right Scores
2 x OPM 157 Total Right Scores
1 x OPM 24 Total Right Scores

This weighting differs somewhat from the weights derived for the test battery which included two parallel forms of MCAT for the 1978 ATC trainee group (see page 114). However, the correlation coefficient between the two weighted ATC tests was .93 (see Table 98).

In the analysis of the new ATC test battery, the passing score was set at the mean raw score for the 1978 ATC applicant group (222.27). This was then equated to a transmuted passing score of 70 (see Table 77). With this base, equivalent scores on the same weighted ATC test battery were computed for 585 of the 1978 ATC trainees. A total of 48 ATC trainees (8.2 percent) failed to achieve a score of 70 on the new ATC test. Of these 48 trainees, 79 percent (38) failed the ATC Laboratory training or withdrew (see Table 100). Some 39 of the 48 trainees had scores in the 65-69 range and 9 scored below 65 (see Table 99). The score range of 70-74 on the ATC test included 102 of the 585 trainees (17 percent). Of this group, 51 percent (52) failed or withdrew (see Table 100). Consequently, use of the mean test score resulted in a 28 percent reduction in the fail/withdraw rates between the ATC trainees who scored in the 70-74 range in contrast to those who scored below 70 on the ATC test. The difference between a fail/withdraw rate of 51 percent in the 70-74 score range and a rate of 79 percent below scores of 70 clearly supports the "cutoff" score developed in the analysis.

The use of scores on the ATC test, with and without extra credit for veterans preference and ATC-related knowledge, as a basis for ranking successful ATC candidates was also examined. Since extra credit is a selection consideration only for those candidates who are appointed from competitive OPM registers, the analysis included only those 1978 ATC trainees who were hired competitively (Non-predevelopmental) and who scored 70 or above on the ATC test. On this basis, 509 of the 585 trainees were included in the analysis.

For the ATC test battery without extra credit, Table 101 (the Non-predevelopmental ATC Group) shows that the fail/withdraw rate in each score range progressively decreases as the test scores increase. In the 95-100 range there were no failures or withdrawals (only 2 trainees were in this range). In the 70-74 range, 47.8 percent failed or withdrew. The largest increase in the fail/withdraw rate is between the 75-79 range (41.9 percent) and the 80-85 range (29.3 percent); a difference of 12.6 percent in the fail/withdraw rate. Table 110 combines the ATC test score ranges and provides the failure/withdrawal rates for this sample of 509 trainees.

TABLE 112

1978 ATC TRAINEES
ATC TEST SCORES
(WITHOUT EXTRA CREDIT)

Score Range	N	Fail/Withdraw	Fail/Withdraw Rate
90 +	29	5	17.2%
80-89	254	66	26.0%
70-79	226	100	44.2%
Total	509	171	33.6%

134

These data clearly show that ranking ATC candidates, based on their test scores and selecting those with the higher scores first, will result in hiring applicants with the highest probability of successfully completing the ATC training program.

Ranking candidates based on total "Earned Rating" scores which include test scores plus extra credit for veterans preference and ATC-related knowledge based on OKT scores was examined next. From the sample of 509 ATC trainees who were hired competitively, the equivalent of the OPM Earned Rating for the ATC test could be computed for 486 trainees. Table 113 provides the fail/withdraw rates by Earned Rating score ranges and compares these to the rates based on the ATC test score ranges without extra credit.

TABLE 113

1978 ATC TRAINEES
(Non-Predevelopmental--Competitive Hires)

Score Range	ATC Test (Without Extra Credit)			ATC Test (With Extra Credit)		
	N	Fail/Withdraw	Fail/Withdraw Rate	N	Fail/Withdraw	Fail/Withdraw Rate
90 +	29	5	17.2%	197	39	19.8%
80-89	254	66	26.0%	229	86	37.6%
70-79	226	100	44.2%	60	38	63.3%
Total	509	171	33.6%	486	163	33.5%

It is evident from Table 113 that ranking candidates for selection consideration on the basis of the Earned Rating (which includes extra credit) also results in selecting those applicants first with the highest probability of successfully completing the ATC training. It is also evident that selection based on the Earned Rating results in a higher fail/withdraw rate in each score range than if candidates were hired solely on the basis of their test battery scores without granting extra credit.

In deriving the Earned Rating score for this sample of 486 ATC trainees, extra credit for veterans preference was included since it is granted by law. Extra credit based on OKT test scores was included based on the positive and significant correlations with ATC performance criterion used in the various research studies.

Consequently, it was hypothesized that extra credit for being a veteran, which has not been examined for validity with the criterion measures used in the research, does not predict (or negatively predicts) success in the ATC Laboratory training. In order to examine this, scores for the 509 ATC trainees were computed which excluded any veterans credit and included only the weighted ATC test battery score and extra credit for OKT scores.

135

Table 114 provides the score distributions for the group of 509 1978 ATC trainees before and after the application of extra credit for air traffic control related knowledge based on OKT points.

TABLE 114

1978 ATC TRAINEES - EFFECT OF EXTRA CREDIT

	1978 ATC Trainees (Distribution Without OKT Credit)			1978 ATC Trainees (Distribution with OKT Credit)	
	N	%		N	%
90+	29	5.7%	90+	139	27.3%
80-89	254	49.9%	80-89	245	48.1%
70-79	226	44.4%	70-79	125	24.6%
Total	509	100.0%	Total	509	100.0%

It is significant to note the addition of OKT points resulted in moving 110 of the ATC trainees to a score range of 90 or more--an increase of about 22 percent. It also resulted in reducing the number of trainees in the lowest scoring group (70-79) by 91--from 226 to 125--a reduction of about 20 percent.

Based on the validity data previously established for the OKT, it would be expected that those trainees who moved to the higher score ranges as a result of OKT extra credit would have lower fail/withdraw rates and that those who remained in the lowest score range would have higher fail/withdraw rates. The comparison of the fail/withdraw rates in the three score ranges for the 509 ATC trainees before and after applying OKT extra credit points is provided in Table 115.

TABLE 115

1978 ATC TRAINEES FAIL/WITHDRAW RATES: EFFECTS OF OKT CREDIT

	1978 ATC Trainees (Without OKT Credit)			1978 ATC Trainees (With OKT Credit)		
	N	Fail/ Withdraw	F/WD Rate	N	Fail/ Withdraw	F/WD Rate
90+	29	5	17.2%	139	19	13.7%
80-89	254	66	26.0%	245	78	31.8%
70-79	226	100	44.2%	125	74	59.2%
Total	509	171	33.6%	509	171	33.6%

Since the fail/withdraw rate for the total group of 509 trainees is a constant (33.6%), the rates can only vary between the different score ranges as shown in Table 115. In order to assess the effect of OKT extra credit on the total fail/withdraw rate, one can assume a selection made from among the group of 509 trainees. For example, if one were to hire 300 trainees

from the 509 in rank order (highest score range first) with an estimated
selection ratio of 80 percent in each score range, the fail/withdraw rate
for the 300 hired before application of OKT extra credit would be 30 percent.
The fail/withdraw rate for the 300 selected after application of OKT extra
credit would be 25 percent. If only 100 of the 509 were selected, the
fail/withdraw rate for the group without OKT extra credit would be 24 percent
in contrast to 14 percent for the group with OKT credit. These data support
the conclusion that ranking applicants for selection consideration based
on test scores and OKT credit will result in hiring applicants with the
highest probability of successfully completing ATC training.

Table 116 shows the fail/withdraw rates for the 509 trainees based on (1) the
ATC test alone; (2) the ATC test plus OKT credit; and, (3) the ATC test plus
credit for OKT and Veterans Preference points.

TABLE 116

1978 ATC TRAINEES FAIL/WITHDRAW RATES - EFFECT OF CREDITS

	ATC Test (No Extra Credit)				ATC Test (Plus OKT Extra Credit Only)				ATC Test (Plus OKT and Vet. Pref. Credit)			
Score Range	N	%	Fail Withdraw	F/WD Rate	N	%	Fail Withdraw	F/WD Rate	N	%	Fail Withdraw	F/WD Rate
90 +	29	(6%)	5	(17.2%)	139	(27%)	19	(13.7%)	197	(41%)	39	(19.8%)
80-89	254	(50%)	66	(26.0%)	245	(48%)	78	(31.8%)	229	(47%)	86	(37.6%)
70-79	226	(44%)	100	(44.2%)	125	(.25%)	74	(59.2%)	60	(12%)	38	(63.3%)
Total	509	(100%)	171	(33.6%)	509	(100%)	171	(33.6%)	486	(100%)	163	(33.5%)

The use of OKT credit resulted in moving 110 trainees into the 90 and above
score range compared to the ATC test alone. The number of fail/withdraws
increased by 14 for a rate of 13 percent for this group. When veterans
preference credit is also applied for 486 of the 509 trainees, an additional
58 trainees moved to the 90 and above score range and the number of fail/
withdraws increased by 20; a fail/withdraw rate of 34 percent for this group of
58 students. From the data presented in Table 116, granting extra credit for
veterans preference increased the fail/withdrawal rates in the higher score
ranges and therefore has a negative relationship to potential success in
ATC training.

Test Fairness. When empirical evidence of validity for a selection proce-
dure such as presented for the new ATC test has an adverse impact on a race,
sex or ethnic group, the Uniform Guidelines require examination of test
fairness where technically feasible.

In establishing the Uniform Guidelines on Employee Selection, it was recog-
nized there is serious debate on the question of test fairness. There are
several competing definitions of test or selection fairness, each of which
incorporates a different set of social values. Reference (32) provides an
analysis of three incompatible ethical positions in regard to fair and un-
biased use of tests together with differing statistical definitions of "test
fairness" and their relationship to specific ethical positions. These three

137

ethical definitions are characterized as:

- Unqualified Individualism

- Qualified Individualism

- The Quota Ethic

The General Accounting Office, in examini·g Federal selection tests and examinations (33) stated with respect to these three definitions:

> "The published literature on test validity indicates that most tests
> are either fair to minority groups or slightly biased in their
> favor by the second definition, (Qualified Individualism) which
> according to OPM, is the only concept of fairness consistent with
> merit system principles. The literature also indicates that by the
> first definition, (Unqualified Individualism) tests are slightly
> biased against minority groups, and if one subscribes to the last
> definition (Quota Ethic) tests have always been biased against
> minorities."

The "Qualified Individualism" definition of fairness adopted by the Office of Personnel Management (OPM) holds that tests are biased when those with equal chances of success on the job have unequal chances of being selected for the job. This definition relies solely on valid measures of aptitude, achievement and experience; maximizes productivity and provides equal opportunity consistent with merit principles. The statistical model upon which the Uniform Guidelines appear to be based, was stated by Cleary (33). Basically, this model requires the regression line which predicts the criterion (e.g. performance) from "scores" be the same for all cultural groups and, absent this, separate regression lines should be used as a basis for selection decisions. Statistical tests developed by Gulliksen and Wilkes (34) provide a procedure which meets the Cleary model. Basically, three independent and sequential tests are involved:

(1) population variances (standard error of estimates) are equal
(non-significant differences), then,

(2) slopes of the population regression lines are equal (non-significant differences), then

(3) intercepts of the population regression lines are equal
(non-significant differences).

If each of the three statistical tests is met, the regression lines are the same and the "fairness" definition of "Qualified Individualism" is satisfied in that "scores" used for selection predict the criterion equally for the various groups considered. On the other hand, if a statistically significant difference is found at any point in the sequential analysis, the test is considered to treat the groups unequally.

138

In examining the new ATC test battery in accordance with the "Qualified Individualism" model of fairness, test scores and criterion data obtained for the sample of 953 ATC trainees attending the FAA Academy during June 1978 through December 1978 were used in the analysis. The number of Hispanics (23 total--7 with OPM test data) was too small for analysis. There were no American Indians in this sample. There was no evidence of adverse impact on selection of Asians with the new ATC test battery. Consequently, the analysis for test fairness was completed with respect to men and women or for White and Black groups.

Sample sizes varied depending on the availability of data on the two OPM tests (OPM 157 and 24) included in the new ATC test battery. Test and criterion data were available for the following ATC trainees (Table 117).

TABLE 117

DISTRIBUTIONS BY SELECTED SAMPLES

Test	Men (N=800) Sample N	Women (N=141) Sample N	White (N=839) Sample N	Black (N=81) Sample N
MCAT 1	790	137	828	79
MCAT TOTAL	790	137	828	79
OPM 157	509	66	539	38
OPM 24	509	66	539	38

In conducting the analysis of test fairness, the tests were weighted in accordance with the values derived for the 1978 ATC trainees. These weights are:

$$\left. \begin{array}{l} 2 \times \text{MCAT 1 (Total Right Scores)} \\ 2 \times \text{MCAT 2 (Total Right Scores)} \end{array} \right\} = \text{MCAT Total} \times 2$$

1 x OPM 157 (Total Right Scores)
1 x OPM 24 (Total Right Scores)

The test for significance for all analyses of fairness was the .05 level of confidence.

Since almost all of the 1978 ATC trainee sample had test scores available on the two parallel forms of MCAT, an analysis of fairness of the MCAT by itself was completed. Table 118 shows the results for the first form of the test administered (MCAT 1) and the total score for both forms (MCAT 1 plus MCAT 2) by race and sex. The test means and standard deviations are based on raw test scores. Since MCAT was the only test used in this analysis, the raw scores were not weighted. The criterion value is based on a mean of "0" with a standard deviation of "1" and reflects, in effect, the percent of trainees who passed the Laboratory training; the validity coefficient (r) is the correlation value of the test with the test with the pass or fail/withdraw criterion.

139

TABLE 118

ANALYSIS OF TEST FAIRNESS

TEST	MEN						WOMEN						TEST OF SIGNIFICANCE		
	N	MEAN	SD	MEAN	SD	r	N	MEAN	SD	MEAN	SD	r	STD ERROR	SLOPE	INTER-CEPT
MCAT 1	790	37.40	6.90	.632	.483	.34	137	34.20	7.90	.577	.496	.370	NS	NS	NS
MCAT TOT	790	80.15	11.27	.632	.483	.40	137	74.91	13.20	.577	.496	.41	NS	NS	NS

TEST	WHITE						BLACK						TEST OF SIGNIFICANCE		
	N	MEAN	SD	MEAN	SD	r	N	MEAN	SD	MEAN	SD	r	STD ERROR	SLOPE	INTER-CEPT
MCAT 1	828	37.61	6.71	.647	.478	.29	79	29.82	7.92	.354	.481	.56	NS	NS	SIG
MCAT TOT	828	80.81	10.76	.647	.478	.35	79	66.35	13.33	.354	.481	.57	NS	NS	NS

140

In terms of sample sizes used in this analysis, the 137 women comprise 14.8 percent of the combined sample of men and women (927) compared to 15.7 percent of the 7,894 ATC trainees who were women hired from January 1976 through October 1980. The 79 Black trainees represent 8.7 percent of the combined sample of the Whi e and Black group (907) compared to 8.3 percent of the 7488 White and Black ATC trainees hired during this same period (see Table 123).

For men and women, the test of fairness shows no significant differences (NS). For the White and Black groups, the difference in the intercept of the regression line, was statistically significant (SIG) for the first MCAT test administered but differences in the standard error and the slope of the regression line were not significant (NS). When the total scores for both MCAT forms administered were analyzed, there were no significant differences between the two groups. This aralysis indicates that the total MCAT scores predict the ATC laboratory pass or fail/withdraw criterion equally for men and women and for the Black and White groups.

The next step was to develop different test battery combinations with the MCAT test scores and OPM 157, OPM 24, and OKT test scores and analyze these various test batteries for significant differences between men and women and the White and Black groups. The intercorrelations between the various test combinations used and their correlation with the pass or fail/withdraw criterion are provided for each group in Table 93. Since a number of trainees did not have test scores on OPM 157 (and OPM 24), the sample sizes were reduced as shown in Table 119. The correlations shown are based on weighted test scores where appropriate using raw test battery scores (RS) as well as transmuted scores (TS). The correlations which combine a test battery with OKT use both the raw OKT score (RS) as well as OKT extra points (PTS) derived from the raw OKT score (i.e., 0, 3, 5, 10, 15 points).

The descriptive statistics for the samples used in the fairness analysis are also provided in comparison to the total sample for each group as was given in Table 84.

141

TABLE 119

TEST FAIRNESS ANALYSES OF TEST BATTERY OPTIONS

	MEN			(Table 84)			WOMEN			(Table 84)		
	N	MEAN	SD	N	MEAN	SD	N	MEAN	SD	N	MEAN	SD
MCAT 1	509	37.6	6.6	800	37.4	6.9	66	33.4	8.1	141	34.1	7.8
MCAT 2	509	43.3	5.7	800	42.9	5.8	66	40.7	6.7	141	40.7	6.4
MCAT TOT	509	80.9	11.0	800	--	--	66	74.1	13.4	141	--	--
OPM 157	509	38.8	6.2	515	38.8	6.2	66	39.9	6.0	67	39.8	6.0
OPM 24	509	46.7	6.5	515	46.6	6.5	66	47.3	7.1	67	47.2	7.1
OKT (RS)	509	64.1	15.8	800	64.9	15.5	66	56.5	14.4	141	57.7	15.1

MEN (N=509)

	(1) MCAT TOT.	(2) OPM 157	(3) OPM 24	(4) OKT (RS)	(5) MCAT TOT. 157 (RS)	(6) MCAT TOT. 157 (TS)	(7) MCAT TOT. 157 OKT (RS)	(8) (TS) MCAT TOT. 157 OKT (PTS)	(9) MCAT TOT. 157 24 (RS)	(10) MCAT TOT. 157 24 OKT (RS)	(11) MCAT TOT. 157 24 (TS)	(12) (TS) MCAT TOT. 157 24 OKT (PTS)	(13) Pass Fail UD
(1) MCAT TOT.	---	.24	.33	.03	.97	.97	.83	.82	.95	.96	.95	.82	.36
(2) OPM 157		---	.09	-.16	.48	.48	.32	.32	.45	.32	.45	.30	.16
(3) OPM 24			---	-.23	.32	.32	.14	.15	.53	.35	.53	.33	.12
(4) OKT (RS)				---	-.02	-.02	.54	.44	-.07	.46	-.07	.41	.21
(5) MCAT TOT + 157(RS)					---	1.00	.83	.83	.97	.86	.97	.82	.38
(6) MCAT TOT + 157(TS)						---	.83	.83	.97	.86	.97	.82	.38
(7) MCAT TOT + 157 + OKT(RS)							---	.94	.78	.98	.78	.92	.44
(8) MCAT TOT + 1.7(TS) + OKT(PTS)								---	.78	.92	.78	.98	.44
(9) MCAT TOT + 157 + 24(RS)									---	.85	1.00	.82	.37
(10) MCAT TOT + 157 + 24 + OKT(RS)										---	.85	.94	.44
(11) MCAT TOT + 157 + 24(TS)											---	.82	.37
(12) MCAT TOT + 157 + 24(TS) + OKT(PTS)												---	.44
(13) Criterion (Lab Pass or F/UD)													---

142

TABLE 119 (Continued)

---------------------------------------WOMEN---------------------------------------
(N=66)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	MCAT TOT.	OPM 157	OPM 24	OKT (RS)	MCAT TOT. 157 (RS)	MCAT TOT. 157 (TS)	MCAT TOT. 157 OKT (RS)	(TS) MCAT TOT. 157 OKT (PTS)	MCAT TOT. 157 24 (RS)	MCAT TOT. 157 24 OKT (RS)	MCAT TOT. 157 24 OKT (TS)	(TS) MCAT TOT. 157 24 OKT (PTS)	Pass Fail WD
(1) MCAT TOT.	---	.23	.04	.17	.98	.96	.89	.92	.95	.89	.95	.90	.33
(2) OPM 157		---	.12	.06	.42	.42	.38	.34	.43	.40	.43	.35	.40
(3) OPM 24			---	-.25	.06	.06	-.06	.01	.29	.15	.29	.20	.08
(4) OKT (RS)				---	.17	.17	.56	.41	.11	.51	.11	.37	.31
(5) MCAT TOT + 157(RS)					---	1.00	.91	.93	.97	.91	.97	.91	.39
(6) MCAT TOT + 157(TS)						---	.91	.93	.97	.91	.97	.91	.39
(7) MCAT TOT + 157 + OKT(RS)							---	.95	.86	.98	.86	.92	.46
(8) MCAT TOT + 157(TS) + OKT(PTS)								---	.89	.95	.69	.98	.41
(9) MCAT TOT + 157 + 24(RS)									---	.91	1.00	.92	.39
(10) MCAT TOT + 157 + 24 + OKT(RS)										---	.91	.95	.47
(11) MCAT TOT + 157 + 24(TS)											---	.92	.39
(12) MCAT TOT + 157 + 24(TS) + OKT(PTS)												---	.42
(13) Criterion (Lab Pass or F/WD)													---

	----------WHITE----------			(Table 84			----------BLACK----------			(Table 84		
	N	MEAN	SD	N	MEAN	SD	N	MEAN	SD	N	MEAN	SD
NCAT 1	539	37.7	6.6	839	37.6	6.7	38	29.7	7.8	81	29.8	7.9
NCAT 2	539	43.6	5.5	839	43.2	5.5	38	36.1	7.5	81	36.6	6.8
NCAT TOT	539	81.2	10.6	839	--	--	38	65.8	14.0	81	--	--
OPM 157	539	39.1	6.2	545	39.1	6.2	38	35.4	6.1	39	35.3	6.1
OPM 24	539	46.8	6.6	545	46.9	6.5	38	44.7	6.6	39	44.2	7.1
OKT (RS)	539	63.1	15.9	839	63.9	15.8	38	65.5	13.6	81	66.2	13.1

TABLE 119 (Continued)

------------------------------------WHITE------------------------------------

(N=539)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8) (TS)	(9)	(10)	(11)	(12) (TS)	(13)
	MCAT TOT.	OPM 157	OPM 24	OKT (RS)	MCAT TOT. 157 (RS)	MCAT TOT. 157 (TS)	MCAT TOT. 157 OKT (RS)	MCAT TOT. 157 OKT (PTS)	MCAT TOT. 157 24 (RS)	MCAT TOT. 157 24 OKT (RS)	MCAT TOT. 157 24 (TS)	MCAT TOT. 157 24 OKT (PTS)	Pass Fail WD
(1) MCAT TOT.	---	.23	.28	.07	.97	.97	.83	.82	.94	.85	.94	.82	.33
(2) OPM 157		---	.13	-.14	.48	.48	.31	.32	.45	.32	.45	.31	.17
(3) OPM 24			---	-.24	.28	.28	.09	.10	.50	.31	.50	.30	.07
(4) OKT (RS)				---	.03	.03	.58	.48	-.03	.50	-.03	.44	.23
(5) MCAT TOT + 157(RS)					---	1.00	.83	.82	.97	.86	.97	.82	.34
(6) MCAT TOT + 157(TS)						---	.83	.82	.97	.86	.97	.82	.34
(7) MCAT TOT + 157 + OKT(RS)							---	.94	.77	.98	.77	.92	.41
(8) MCAT TOT + 157(TS) + OKT(PTS)								---	.77	.92	.77	.98	.40
(9) MCAT TOT + 157 + 24(RS)									---	.85	1.00	.81	.33
(10) MCAT TOT + 157 + 24 + OKT(RS)										---	.85	.94	.41
(11) MCAT TOT + 157 + 24(TS)											---	.81	.35
(12) MCAT TOT + 157 + 24(TS) + OKT(PTS)												---	.40
(13) Criterion (Lab Pass or F/WD)													---

--BLACK--

(N=38)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8) (TS)	(9)	(10)	(11)	(12) (TS)	(13)
	MCAT TOT.	OPM 157	OPM 24	OKT (RS)	MCAT TOT. 157 (RS)	MCAT TOT. 157 (TS)	MCAT TOT. 157 OKT (RS)	MCAT TOT. 157 OKT (PTS)	MCAT TOT. 157 24 (RS)	MCAT TOT. 157 24 OKT (RS)	MCAT TOT. 157 24 (TS)	MCAT TOT. 157 24 OKT (PTS)	Pass Fail WD
(1) MCAT TOT.	---	-.28	.17	.27	.98	.98	.90	.92	.96	.91	.96	.91	.60
(2) OPM 157		---	-.11	-.11	-.06	-.06	-.10	-.18	-.09	-.12	-.09	-.19	.07
(3) OPM 24			---	-.09	.15	.15	.08	.09	.37	.27	.37	.25	.35
(4) OKT (RS)				---	.25	.25	.62	.49	.22	.58	.22	.47	.24
(5) MCAT TOT + 157(RS)					---	1.00	.92	.92	.97	.91	.97	.90	.64
(6) MCAT TOT + 157(TS)						---	.92	.92	.97	.91	.97	.90	.64
(7) MCAT TOT + 157 + OKT(RS)							---	.95	.88	.98	.88	.93	.62
(8) MCAT TOT + 157(TS) + OKT(PTS)								---	.89	.94	.89	.99	.61
(9) MCAT TOT + 157 + 24(RS)									---	.92	1.00	.91	.68
(10) MCAT TOT + 157 + 24 + OKT(RS)										---	.92	.95	.67
(11) MCAT TOT + 157 + 24(TS)											---	.91	.68
(12) MCAT TOT + 157 + 24(TS) + OKT(PTS)												---	.64
(13) Criterion (Lab Pass or F/WD)													---

144

The first test battery analyzed for fairness consisted of MCAT Total and OPM 157 raw scores. In this analysis, MCAT Total was weighted by "2" and OPM 157 weighted by "1" in accordance with the weights previously derived. Then the OKT raw scores (weighted "1") were combined with the MCAT and OPM 157 raw scores and this combination separately evaluated. The results for each sex and race group are provided in Table 120. It should be noted that at this point, a "passing" score for the combined MCAT and OPM 157 test scores was not established. Consequently, the addition of the OKT raw score was independent of whether or not the ATC trainee "passed" the MCAT/OPM 157 test battery and also independent of extra credit points (0, 3, 5, 10, 15) which would be derived from the raw OKT scores.

The combination of MCAT and OPM 157 raw test scores show that this test battery predicts the ATC laboratory pass or fail/withdraw criteria equally for men and women and for Black and White groups. The same results were obtained when MCAT, OPM 157 and OKT raw test scores were analyzed separately.

Use of raw scores in analyzing the test battery comprised of MCAT, OPM 157 and OKT does not take into account the proposed use of OKT in the selection process. Since OKT is a job specific test, it is not intended for use in the initial "pass/fail" determination of appointment eligibility for ATC applicants. Its use would be limited to determining if those applicants who passed the test battery (in this case MCAT and OPM 157) would also be given extra credit for air traffic control related knowledge as measured by OKT. Further, the amount of extra credit (0, 3, 5, 10, 15 points) would be based on the raw OKT score with no credit for raw scores below 65 and maximum credit (15 points) for raw OKT scores of 80 or above.

In order to evaluate the fairness of a test battery in combination with OKT in relation to its intended use, it is necessary to (1) establish a "passing" score on the test battery for an ATC applicant group, (2) transmute the raw test battery score in relation to the passing score, and (3) add the appropriate extra credit points based on OKT raw score for those who passed the test battery.

145

TABLE 120

ANALYSIS OF TEST FAIRNESS

TEST	MEN — TEST			MEN — CRITERION			WOMEN — TEST			WOMEN — CRITERION			TEST OF SIGNIFICANCE		
	N	MEAN	SD	MEAN	SD	r	N	MEAN	SD	MEAN	SD	r	STD ERROR	SLOPE	INTER-CEPT
MCAT TOT OPM 152 (RS)	509	200.62	24.19	.629	.484	.38	66	188.02	28.80	.515	.504	.39	NS	NS	NS
MCAT TOT OPM 157 OKT (RS)	509	264.71	28.68	.629	.484	.44	66	244.47	34.33	.515	.504	.46	NS	NS	NS

TEST	WHITE — TEST			WHITE — CRITERION			BLACK — TEST			BLACK — CRITERION			TEST OF SIGNIFICANCE		
	N	MEAN	SD	MEAN	SD	r	N	MEAN	SD	MEAN	SD	r	STD ERROR	SLOPE	INTER-CEPT
MCAT TOT OPM 157 (RS)	539	201.54	23.37	.640	.480	.34	38	166.93	26.88	.316	.471	.64	NS	NS	NS
MCAT TOT OPM 157 OKT (RS)	539	264.62	28.62	.640	.480	.41	38	232.41	33.03	.316	.471	.62	NS	NS	NS

(RS) Raw Scores

In order to transmute the total raw score for the ATC trainee sample, it was necessary to estimate the mean score which would have been obtained by an ATC applicant group on the ATC test battery and equate this mean score to a passing score of 70. Fortunately, the tests used in the analysis of the ATC trainee sample were identical to the tests taken by the 5,931 ATC applicants in 1978 except for the second form of MCAT which was not administered to the ATC applicants. Consequently, given the mean scores on each test which were available for both the ATC trainees and the ATC applicants, it is possible to estimate a mean for the ATC applicants on the second MCAT form. By combining this estimated mean for "MCAT 2" with the actual mean scores obtained by the ATC applicants on the other tests and weighting them, an estimated mean score of 161.2 for the 1978 ATC applicant group on a test battery comprised of MCAT Total and OPM 157 was derived as shown below:

		1978 ATC Applicants				1978 ATC Trainees	
	MEAN	x	WT.	= TOTAL	MEAN	x	WT.	= TOTAL
MCAT 1	30.6	2	61.2	36.9	2	73.8
*(MCAT 2	35.6	2	71.2)	42.6	2	85.2
OPM 157	28.8	1	28.8	38.8	1	38.8
	Est. Mean Raw Score 161.2						197.8

(1) ATC Applicant Mean – MCAT 1 $\dfrac{30.6}{36.9}$ = .829
     ATC Trainee Mean – MCAT 1

(2) ATC Trainee Mean – MCAT 2   42.6
     ATC Trainee Mean – MCAT 1   36.9
      Difference   5.7 (6.0)

*(3) Est. MCAT 2 Mean for Applicant Group = 35.6
      (6.0 x .829 = 5.0 + 30.6 = 35.6)

Since the correlation between the two differently weighted ATC tests for the ATC trainee sample was .93 (Table 98), an estimated mean score for the ATC applicant group of 159.5 was also derived using the ratio of the mean scores for the 1978 applicant and ATC trainee groups for the common weighted tests as shown on the next page.

147

(1) Actual Applicant Group Mean (N = 6000)
$\quad$ (4 x MCAT 1 + 2 x 157) $\qquad$ = $\qquad$ 181.6 (Table 43)

(2) Actual Mean ATC Trainee Group (N = 592)
$\quad$ (4 x MCAT 1 + 2 x 157) $\qquad$ = $\qquad$ 225.2 (Table 69)

(3) Actual Mean ATC Trainee Group (N = 592)
$\quad$ (2 x MCAT 1 + 2 x MCAT 2 + 1 x 157) $\qquad$ = $\qquad$ 197.8 (Table 69)

(4) $\dfrac{197.8}{225.2} \times \dfrac{X}{181.6}$ = 159.5 Est. MEAN for Applicant Group on ATC Test Weighted
$\qquad\qquad\qquad$ (2 x MCAT 1 + 2 x MCAT 2 + 1 x 157)

Given these two estimated mean scores for an applicant group on the ATC test battery comprised of MCAT and OPM 157, a mean value of 160 was used as a "passing" score and equated to a transmuted score of 70. The total raw weighted scores for each ATC trainee on MCAT and OPM 157 were then transmuted by:

$$T_s = (\frac{(RS - 160) \times 30}{MAX\ RS - 160}) + 70$$

$\quad$ RS = Total weighted raw score
$\quad$ MAX RS = 255.0
$\quad$ MEAN RS = 160.0

An analysis of fairness was completed using the obtained transmuted test battery scores for the ATC trainees. A separate analysis was completed using the transmuted scores plus extra credit points (0, 3, 5, 10, 15) derived from OKT raw scores for those ATC trainees who "pass" the test battery with scores of 70 or more. The results are provided in Table 121.

148

TABLE 121
ANALYSIS OF TEST FAIRNESS

TEST		N	TEST MEAN	SD	CRITERION MEAN	SD	r	N	TEST MEAN	SD	CRITERION MEAN	SD	r	TEST OF SIGNIFICANCE STD ERROR	SLOPE	INTER-CEPT
			MEN						WOMEN							
MCAT TOTAL OPM 157	(TS)	509	82.83	7.64	.629	.484	.38	66	78.85	9.10	.515	.504	.39	NS	NS	NS
MCAT TOTAL OPM 157 OKT	(TS) (PTS)	509	87.50	9.92	.629	.484	.44	66	80.88	10.75	.515	.504	.41	NS	NS	NS

TEST		N	TEST MEAN	SD	CRITERION MEAN	SD	r	N	TEST MEAN	SD	CRITERION MEAN	SD	r	TEST OF SIGNIFICANCE STD ERROR	SLOPE	INTER-CEPT
			WHITE						BLACK							
MCAT TOTAL OPM 157	(TS)	539	83.12	7.38	.640	.480	.34	38	72.19	8.49	.316	.471	.64	NS	NS	NS
MCAT TOTAL OPM 157 OKT	(TS) (PTS)	539	87.48	9.68	.640	.480	.49	38	76.14	11.90	.316	.471	.61	NS	NS	NS

(TS) = Transmuted Test Battery Score
(PTS) = OKT Points

The combination of transmuted test battery scores for MCAT and OPM 157 plus extra credit points for air traffic control related knowledge based on OKT scores replicates the operational use of the selection tests. The analysis of fairness provided in Table 121 shows that this test battery including OKT, predicts the ATC laboratory pass or fail/withdraw criterion equally for men and women and for Black and White groups.

The final step in the fairness analysis was to examine the entire test battery comprised of MCAT, OPM 157 and OPM 24 with and without extra credit for air traffic control related knowledge based on OKT scores. The results are provided in Table 123. In order to add OKT extra credit points, it was again necessary to transmute the total raw score for MCAT, OPM 157 and OPM 24 and to establish a passing raw score which was equated to a transmuted score of 70 as shown below:

	1978 ATC Applicants				1978 ATC Trainees			
TEST	MEAN	x	WT	= TOTAL	MEAN	x	WT	= TOTAL
MCAT 1	30.6	2		61.2	36.9	2		73.8
*(MCAT 2	35.6	2		71.2)	42.6	2		85.2
OPM 157	28.8	1		28.8	38.8	1		38.8
OPM 24	40.6	1		40.6	46.6	1		46.6
Est. Mean Raw Score				201.8				244.4

*See page 143 for derivation of estimated mean score for MCAT 2.

Again, an estimated mean raw score for the applicant group was derived using a ratio method as follows:

(1) Actual Applicant Mean (N = 6000)
(4 x MCAT 1 + 2 x 157 + 1 x 24)   = 222.3 (Table 62)

(2) Actual Mean ATC Trainee Group (N = 592)
(4 x MCAT 1 + 2 x 157 + 1 x 24)   = 272.6 (Table 83)

(3) Actual Mean ATC Trainee Group (N = 592)
(2 x MCAT 1 + 2 x MCAT 2 + 1 x 157 + 1 x 24)   = 245.6 (Table 83)

(4) $\frac{245.6}{272.6}$ $\frac{X}{222.3}$ =   200.3 Est. Mean for Applicant Group on ATC Test Weighted
(2 x MCAT 1 + 2 x MCAT 2 + 1 x 157 + 1 x 24)

Given these two estimated mean scores for an applicant group (201.8 and 200.3) on the ATC test battery comprised of MCAT, OPM 157 and OPM 24, a mean value of 200 was used as a "passing" score and equated to a transmuted score of 70.

The total raw weighted scores for each ATC trainee on MCAT, OPM 157 and OPM 24 were then transmuted by:

$$T_s = \left(\frac{(RS - 200) \times 30}{MAX\ RS - 200} \right) + 70$$

R_s = Total weighted raw score

MAX RS = 306

MEAN RS = 200

Table 122 shows the results of the fairness analysis on this ATC test battery (for both raw and transmuted scores) and for the ATC test battery transmuted scores in combination with extra credit points based on OKT for those trainees who scored 70 or more on the test battery.

As shown in Table 122, there were no significant differences on the ATC test battery between men and women. However, the addition of OPM 24 test scores (Arithmetic Reasoning) to the scores of MCAT and OPM 157 resulted in the estimated standard error (population variances) between the Black and White groups being statistically significant. The statistical tests developed by Gulliksen and Wilkes (34) require that if a significant difference is found at any of the three steps in the analysis, no further tests are made since the regression lines used for predictions are, by definition, unequal.

151

TABLE 122

ANALYSIS OF TEST FAIRNESS

TEST		MEN — TEST N	MEAN	SD	MEN — CRITERION MEAN	SD	r	WOMEN — TEST N	MEAN	SD	WOMEN — CRITERION MEAN	SD	r	TEST OF SIGNIFICANCE STD ERROR	SLOPE	INTER-CEPT
MCAT TOTAL OPM 157 OPM 24	(RS)	509	247.29	26.95	.629	.484	.37	66	235.27	30.08	.515	.504	.39	NS	NS	NS
MCAT TOTAL OPM 157 OPM 24	(TS)	509	83.38	7.63	.629	.484	.37	66	79.98	8.51	.515	.504	.39	NS	NS	NS
MCAT TOTAL OPM 157 OPM 24 (TS) OKT (PTS)		509	88.04	9.72	.629	.484	.44	66	82.01	10.09	.515	.504	.42	NS	NS	NS

TEST		WHITE — TEST N	MEAN	SD	WHITE — CRITERION MEAN	SD	r	BLACK — TEST N	MEAN	SD	BLACK — CRITERION MEAN	SD	r	TEST OF SIGNIFICANCE STD ERROR	SLOPE	INTER-CEPT
MCAT TOTAL OPM 157 OPM 24	(RS)	539	248.32	25.97	.640	.480	.33	38	211.62	28.61	.316	.471	.68	SIG	--	--
MCAT TOTAL OPM 157 OPM 24	(TS)	539	83.67	7.35	.640	.480	.33	38	73.29	8.10	.316	.471	.68	SIG	--	--
MCAT TOTAL OPM 157 OPM 24 (TS) OKT (PTS)		539	88.03	9.42	.640	.480	.40	38	77.16	11.45	.316	.471	.64	NS	NS	NS

The fact that the regression lines for the White and Black groups are statistically unequal, by itself, does not identify the practical implications for a selection procedure. Under the Cleary model of test fairness, if tests are "biased" (unequal), an alternative solution is to use the separate regression lines for selection decisions. This solution, however, explicitly introduces race, sex, or ethnic group (in this case race) as a predictor in the selection process. This is incompatible with the ethical position of "Qualified Individualism" and law (32).

In these circumstances, a single regression line derived by combining the groups typically has been used as the basis for selection procedures. The effect of this is to "overpredict" the criterion (e.g., performance) for that group which scored lower on the test and had a lower mean on the criterion measure (32). In the analysis of this ATC test battery, use of a single regression line obtained by combining the White and Black groups (or the total sample) would "bias" the selection procedure in favor of the Black group. This result conforms with the general findings in the published literature referred to by the General Accounting Office in their review of Federal selection tests (33).

Table 122 also shows that when extra credit points based on OKT are added to the transmuted scores for this ATC test battery, there were no significant differences between men and women or the Black and White groups. While this indicates that the combination of the test battery and OKT points predicts the ATC laboratory pass or fail/withdraw criterion equally for all groups, the test battery itself would be used to establish initial appointment eligibility for applicants. The fact that addition of OKT points results in equal treatment for those who pass the test does not adequately address the unequal treatment resulting from the use of OPM 24 under the Uniform Guidelines. Given that OPM 24 results in unequal prediction of the pass or fail/withdraw criterion between the White and Black groups, it should not be used as part of the ATC selection battery. From a practical viewpoint, the exclusion of OPM 24 has essentially no impact on the multiple correlation or the predictive value for the total group or for men and women since its contribution to the prediction of the performance criterion is very small (see Tables 95 and 96).

Alternative Selection Procedures. During the course of these research studies, a number of alternative instruments of a cognitive and non-cognitive structure were examined. The 1972, 1977, and 1978 studies of full-performance, developmental, and trainee air traffic controllers examined the relationship between non-cognitive instruments (including biographical data and personality tests) and performance criterion. Use of biographical data as an alternative selection procedure was examined in some depth.

In May 1979, the FAA proposed to the Office of Personnel Management (OPM) a demonstration pr. ect under the Civil Service Reform Act that addressed recruitment and selection of women and minorities in the air traffic control occupation (35). This proposal involved a 5-year period during which approximately half of the new ATC hires (750) in the FAA Southern region would be

selected in nonregister order as a control group and the other half (750) selected in the regular manner from the OPM register to form a noncontrol group. Candidates in the Southern region would be given a detailed multiple choice biographical questionnaire and "profiled" on relevant life experience dimensions and grouped into life profile clusters. The life profiles would then be used as an alternate means for ranking and selection of candidates for the control group. Success in ATC training would then be evaluated for both the control and noncontrol groups as a means validating the alternative selection procedure (36).

After review of the proposal and the biographical questionnaire, OPM pointed out a number of concerns on "Job Relatedness", Privacy, "Subjectivity", and "Public Relations" for most of the questions included in the biographical form (36). On December 4, 1979, the Deputy Director, OPM, suggested the use of numerical scores based on "life profiles" derived from empirical data established from validity studies. It was also pointed out that the "job relatedness" issue presented a difficult problem because of practical and legal requirements (37).

The biographical questionnaire proposed for this project has been administered to 545 of the ATC trainees attending the FAA Academy during the period 1976-1977. Information on sex, race, OPM selection scores, ATC training laboratory composite scores and training pass/fail status was also available for this sample of 545 students. These data were provided to the Institute for Behavioral Research, University of Georgia, for analysis. Eight factors were identified through factor analysis:

- o Academic factor
- o Social factor
- o Child Relationship factor
- o Initiative factor
- o IFR/VFR Experience factor
- o Parental Permissiveness factor
- o Physical/Sports factor
- o Socio/Economic factor

Based on these, the 545 trainees were clustered into six subgroups with similar life experiences. Because of the weak relationship between success or failure in the laboratory training and the subgroup memberships, the eight life experience factors were used to predict the laboratory score using regression and stepwise regression analysis. The results showed that the factors were not good predictors of the laboratory composite score which determined the pass/fail status of ATC trainees. Generally the results showed that ATC experience and tests involving mathematics and physical sciences would yield better predictions of success in AT training. The report (39) suggested that using a sample of working controllers to establish similar life experience and group membership might prove more meaningful. While this had been planned as part of the project methodology (36), the results of the University of Georgia study, the OPM concerns regarding validity of biographical data as a selection basis and cost consideration resulted in terminating this approach toward alternative selection procedures.

154

As pointed out in conjunction with the analysis of the sample of 953 ATC trainees attending the FAA Academy in 1978 (page 101), the FAA recognized the difficulty of selecting women and minorities from competitive OPM registers in 1968. As a consequence, the Predevelopmental ATC program was established. In 1974, Executive Order 11813 provided for noncompetitive conversion of Cooperative Education students and this authority was incorporated in the recruitment and selection program for women and minorities in air traffic and other FAA occupations.

These alternative selection programs have been and will continue to be a major vehicle to address the adverse impact on women and minorities as well as the Federal Equal Opportunity Recruitment Program (FEORP) requirements established by the 1978 Civil Service Reform Act.

Table 123 shows the total number of ATC trainees hired from January 1976 through October 1980; those hired from OPM register (competitive) and those hired through the Predevelopmental and Co-op program (non-competitive) by race and sex.

TABLE 123

ATC TRAINEES HIRES
January 1976 - October 1980

Group	Total Hires	% of Total	Competitive Hires	% of Group	Non-Competitive Hires	% of Group
Men	6653	(84.3%)	6218	(93.5%)	435	(6.5%)
Women	1241	(15.7%)	837	(67.4%)	404	(32.6%)
Total	7894	(100.0%)	7055	(89.4%)	839	(10.6%)
American Indian	21	(.3%)	18	(85.7%)	3	(14.3%)
Asian	66	(.9%)	61	(92.4%)	5	(7.6%)
Hispanic	220	(2.8%)	134	(60.9%)	86	(39.1%)
Black	618	(7.9%)	300	(48.5%)	318	(51.5%)
White	6870	(88.1%)	6442	(93.8%)	428	(6.2%)
Total	7795	(100.0%)	6955	(89.2%)	840	(10.8%)

155

CONCLUSIONS AND RECOMMENDATIONS

CONCLUSIONS: The various research studies previously discussed support the following conclusions.

- The experimental ATC test battery comprised of MCAT, OPM 157, and OPM 24 is a valid and statistically significant tool for use in screening applicants for the ATC occupation. The need for screening applicants is particularly important for this occupation because of: (1) the high cost of training to the taxpayer; (2) the large number of applicants in relation to the relatively few vacancies which are filled each year; and (3) the fact that there is essentially no "self screening" on the part of applicants since there are no educational or specialized experience requirements for appointment eligibility at the entry grade levels (GS-5 or GS-7).

- The use of the Occupational Knowledge Test (OKT) in conjunction with the experimental ATC test battery increases the predictive validity of the selection procedure. It provides a better method of establishing the applicant's knowledge relevant to the ATC occupation as compared to the Rating Guide presently used by the Office of Personnel Management. Further, the use of OKT allows applicants who have acquired this knowledge outside of the specific work experiences given credit under the present Rating Guide to earn extra credit for competitive selection consideration.

- The use of the experimental ATC test battery and OKT scores for ranking competitive ATC applicants for appointment consideration and selecting those with the highest scores first is a valid use of the applicants' test scores since it significantly increases the probability of success in the ATC occupation as measured by passing the Initial ATC Qualification Training Program. The addition of veterans preference points to an applicant's score on the ATC test battery is not a valid predictor of success in the ATC occupation.

- The statistical analyses encompassed a large number of different experimental tests including the five OPM tests presently used for scoring ATC applicants. The resulting experimental ATC test battery (MCAT, OPM 157, and OPM 24) was derived from multiple regression analysis. The tests were examined with respect to both unweighted and weighted test score values. Only the OPM 157 test of the present five OPM tests contributed to the multiple correlation (R) predicting the ATC training (performance) criteria. Weighting the tests increased the validity of the ATC test battery.

156

- Analysis of the utility of the ATC test battery indicates a significant potential for reducing overall cost of ATC training. The new test is effective in identifying those applicants who pass the present OPM ATC test but have high fail/withdraw rates in the Initial ATC Qualification training program. Operational use of the new ATC test battery would require more frequent administration (perhaps three times a year) in order to obtain more applicants with Earned Rating scores above 85 as compared to the present OPM test.

- Establishing the passing score on the new ATC test battery at the approximate mean score for an ATC applicant group is supported by the analysis of fail/withdraw rates by score range groups. Setting a higher passing score (for example a transmuted score of 80) would significantly increase the adverse impact on women and minorities as well as make it extremely difficult to recruit women and minorities through either competitive or non-competitive selection procedures.

- The present OPM ATC test battery and the experimental ATC test battery both have an adverse impact on some minority groups, particularly Blacks. When extra credit for veterans preference and ATC related knowledge or experience is granted to those applicants who pass the test battery, there is also an adverse impact on women. The experimental ATC test battery has a somewhat greater adverse impact on selection of Women and Blacks than does the present OPM ATC test battery. The veterans preference credit is a significant factor in the adverse impact on Women. While the statistical evidence does not support the validity of veterans preference in relation to performance criteria, it is required by law.

- Given the evidence of adverse impact on women and some minority groups, especially Blacks, analyses of test fairness were completed on the various components of the experimental test battery and the OKT. These analyses show that the requirements of the Uniform Guidelines on Employee Selection are met for the ATC test battery comprised of MCAT and OPM 157 and for MCAT, OPM 157 and OKT extra credit points. This was true for both men and women and the White and Black groups.

- The Multiplex Controller Aptitude test (MCAT) is the major component in the validities obtained with the new ATC test. The analysis of fairness for MCAT, by itself, included a sample of men and women as well as White and Black trainees which was proportional to the total population of ATC trainees attending the FAA Academy during the period of January 1976 through October 1980.

157

● The addition of OPM 24 test scores to the MCAT and OPM 157 test
battery revealed a significant difference in the population
variance between the White and Black groups indicating that an
ATC test battery comprised of MCAT, OPM 157 and OPM 24 did not
meet the fairness requirements of the Uniform Guidelines for the
White and Black groups. The effect of these differences would be
to bias the test battery somewhat in favor of the Black group if
a common regression line for the two groups were used. However,
there were no significant differences between men and women on
this ATC test battery.

RECOMMENDATIONS: Based on the conclusion derived from the research studies,
the following recommendations for the use of selection tests and procedures
to qualify, rate and rank applicants for the Air Traffic Control occupation
are supported.

● The ATC Test Battery used for qualifying applicants for placement
on competitive OPM registers or for qualifying applicants for
non-competitive appointment consideration should consist of:

(1) The two forms of the Multiplex Controller Aptitude Test
(MCAT) with the total correct scores weighted by a
value of 2.

(2) The present OPM Test, Abstract Reasoning and Letter
Sequence scored (R-1/4W) weighted by a value of 1.

(3) Each form of the MCAT has 55 questions segmented into
two parts; 20 minutes is allowed for Part A (27 or 28
questions); and, 15 minutes allowed for Part B (27 or
28 questions). Eight minutes is also allowed for
reading test directions and practice problems. There-
fore, a total of 43 minutes is required for administra-
tion of each MCAT form and the total test time required
for both MCAT forms is 1 hour and 26 minutes. The
Abstract Reasoning and Letter Sequence test is a two-
part test; each part has 25 questions. Test time for
Part A is 15 minutes; Part B is 20 minutes. There are
9 sample items with 5 minutes allowed for these practice
questions.

(4) Total testing time for the recommended ATC Test Battery,
including practice and test familiarization time, is
2 hours and 6 minutes.

● The composite weighted raw score on the test battery should be trans-
muted by a linear transformation of the distribution so that the
mean raw score of the applicant group is equated to a passing score
of 70.

158

- All applicants, either competitive or non-competitive, who score 70 or greater should be considered as eligible for appointment consideration in the Air Traffic Control occupation.

- The Occupational Knowledge Test (OKT) should be administered to all competitive and non-competitive applicants together with the recommended ATC Test Battery. The results of the OKT test should be used for granting additional earned credit in place of the present OPM Rating Guide.

  (1) The OKT is an 80-item test and is scored for correct answers only. Test time is 50 minutes. There are no practice items. Two minutes are allowed for reading test directions.

  (2) The total test administration time, including OKT, is 2 hours and 58 minutes.

  (3) Additional earned credit should be granted only to those applicants who achieved a score of 70 or greater on the qualifying ATC Test Battery (MCAT and OPM 157).

  (4) Additional earned credit points should be granted to qualifying applicants based on OKT scores as follows:

OKT Scores				
Right OKT Scores 0-80		Transmuted Scale 0-100		Additional Points Earned
52-55	=	65-69	=	3
56-59	=	70-74	=	5
60-63	=	75-79	=	10
64+	=	80+	=	15

  (5) For competitive applicants, additional points based on OKT scores should be added to the scores of applicants who achieve 70 or more on the ATC Test Battery. The total score, together with any veterans preference points, should be used to rank candidates for selection consideration on the OPM register.

  (6) For non-competitive applicants, the OKT scores together with the ATC Test Battery should be used as a basis for offering GS-7 appointments in the ATC occupation rather than GS-5 predevelopmental appointments to those individuals whose high scores indicate they have the aptitude and knowledge to enter the GS-7 Initial ATC Qualification training.

- Qualified applicants should be ranked for competitive appointment consideration based on the sum of their scores on the ATC Test Battery, the OKT and Veterans Preference credit with those applicants having the highest total score (Earned Rating) given first consideration for selection.

159

- The existing OPM register of qualified applicants (approximately 4,500) should be retained. Applicants who qualify on the basis of the new ATC Test Battery should be interspersed in their appropriate rank order with those applicants who are presently on the register.

- The six parallel forms of MCAT and the eight parallel forms of OKT which have been developed and used in conjunction with the validation studies should be placed into operational use for ATC applicant screening and selection.

- The FAA, Office of Aviation Medicine, should be delegated the responsibility from OPM for continued development of additional parallel forms of MCAT and OKT to insure against compromise of operational test forms.

160

BIBLIOGRAPHY

(1) Brokaw, L. Historic Overview of Research and Development of Air Traffic
     Controller Selection Selection Proceeding, Military Testing Association,
     October 15-19, 1979.

(2) Cobb, B. B. Air Traffic Aptitude Test Measures of Military and FAA Controller
     Controller Trainees. Federal Aviation Administration, Office of Aviation
     Medicine, Report No. AM-71-40-1971.

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

(4) Cobb, B.B., C. L. Young, and B. L. Rizzuti. Education as a Factor in the Selection
     of Air Traffic Controller Trainees. Federal Aviation Administration, Office
     of Aviation Medicine, Report No. AM-76-6, 1976.

(5) Colmen, J. G. Review and Evaluation of Present System for Selection of Air
     Traffic Controllers, Washington, D.C. Report for Phase I, Task I.
     Federal Aviation Administration Contract to DOT-FA-70WA-2371, Education and
     Public Affairs Inc., July 1970.

(6) Mies, J., J.G. Colmen and O. Domeneck. Predicting Success of Applicants
     for Positions as Air Traffic Control Specialists in the Air Traffic Service.
     Education and Public Affairs Inc., Contract DOT FA-75WA-3646, May 8, 1977.

(7) Worker Attributes for Air Traffic Control Jobs and Their Measurement.
     Education and Public Affairs Inc., 1970.

(8) Air Traffic Controller Selection and Retention Task Force Report. Federal
     Aviation Administration, December 1974.

(9) Air Traffic Controller Selection and Retention Cost Analysis. Federal
     Aviation Administration, March 1975.

(10) Training of U.S. Air Traffic Controllers. Institute for Defense Analysis.
     (IDA Report R206), January 1975.

(11) Air Route Traffic Control Center, Descriptive and Flow Diagrams of Control
     Functions. TM-5329/000/00, System Development Corporation, 30 June 1974.

(12) Milne, A. M. and J. G. Colmen, Selection of Air Traffic Controllers for
     Federal Aviation Administration. Education and Public Affairs Inc.,

(13) Summary of Education and Public Affairs Inc. Report on: Predicting Success
     of Applicants for Positions as Air Traffic Control Specialists in the Air
     Traffic Control Specialists in the Air Traffic Service. Federal Aviation
     Administration. August 1977.

(14) Mies, J. and J. G. Colmen. Final Report of Task I, Development of Recommen-
     dations for ATCS Selection. January 1976.

(15) Colmen, J. Validity of the Cattell 16 Personality Factor Questionnaire and Other "Non-Cognitive" Tests for Selection and Placement of Air Traffic Control Specialists. Education and Public Affairs Inc., Federal Aviation Administration Contract DOT FA-75WA-3646, May 8, 1977.

(16) Mies, J., J. G. Colmen and O. Domeneck. Appendices--Predicting Success of Applicants for Positions as Air Traffic Control Specialists in the Air Traffic Service. Education and Public Affairs Inc., Contract DOT FA-75WA-36, May 8, 1977.

(17) Uniform Guidelines on Employee Selection Procedures, FPM Letter 335-12, Appendix A, Office of Personnel Management.

(18) Statistical Analyses of ATC Applicants, Office of Personnel Management, (Unpublished, available from OPM and FAA).

(19) Dailey, J. G., and E. J. Pickrel. Development of New Selection Tests for Air Traffic Controllers. FAA Office of Aviation Medicine Report No. AM-77-25; 1977.

(20) 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, September 1979.

(21) Thorndike, R. L. Personnel Selection. Wiley and Sons; New York, 1945.

(22) Boone, J. O. and M. A. Lewis. The Development of a New Procedure to Make Maximum Use of Available Information When Correction Correlations for Restriction Range Due to Selection. FAA Office of Aviation Medicine Report No. AM-78-36; September, 1978.

(23) National En Route and Terminal Air Traffic Training Program. FAA Order 3220.18, January 15, 1976.

(24) Instructional Program Guide (IPG) Terminal. TP-12-9-1, May 1976.

(25) Instructional Program Guide (IPG) En Route. EP-12-(-1A, September 1977.

(26) Lewis, M.A. Use of the Occupational Knowledge Test to Assign Extra Credit in Selection of Air Traffic Controllers. FAA, Office of Aviation Medicine,

(27) Lewis, M.A. Implementation of an Occupational Knowledge Test to Assign Extra Credit in Selection of Air Traffic Controllers. FAA, Office of Aviation Medicine, Unpublished Report.

(28) Buckley, E. P., W. F. O'Connor and Beebe, T. A Comparative Analysis of Individual and System Performance Indices for the Air Traffic Control System. FAA, NAFEC, Atlantic City, New Jersey, 1970.

(29) Buckley, E. P., and Beebe, T. The Development of a Motion Picture Measurement Instrument for Aptitude for Air Traffic Control. FAA, NAFEC, November 1970.

(30) Dailey, J. T. Development of an Occupational Knowledge Test for ATCS. FAA Office of Aviation Medicine Report (Unpublished), April 1971.

(31) Cohen, J. Statistical Power Analysis for Behavioral Science. New York: Academy Press, 1969.

(32) Hunter, J. E. and Schmidt, F. L., Fairness of Selection Tests: A Critical Analysis. Personnel Research and Development Center, U.S. Civil Service Commission (PS 76-5), September, 1976.

(33) Federal Employment Examinations: Do They Achieve Equal Opportunity and Merit Principle Goals?, General Accounting Office, FPCD-79-46, May 15, 1979.

(34) Gulliksen, H., and Wilks, S. S., Regression Tests for Several Samples. Psychometrika, 1950, p. 91-114.

(35) Letter to Assistant Director, Policy Analysis, Office of Personnel Management, from Associate Administrator for Administration, Federal Aviation Administration, May 1, 1979.

(36) Research and Development Project on the Selection of Air Traffic Control Candidates, Federal Aviation Administration Southern Region, June 1979.

(37) Letter to Associate Administrator for Administration, Federal Aviation Administration, from Director, Personnel Research and Development Center, June 28, 1979.

(38) Letter from Deputy Director, Office of Personnel Management to Associate Administrator for Administration, FAA, December 4, 1979.

(39) Gauger, K., Analysis of the Federal Aviation Administration's Life Experience Questionnaire, Institute for Behavioral Research, University of Georgia, March 1980.