JUN 13 1978

CONTROLLED

FAA-AM-77-25

ORIGIN

DEVELOPMENT OF NEW SELECTION TESTS FOR AIR TRAFFIC CONTROLLERS

John T. Dailey, Ph.D.
Evan W. Pickrel, Ph.D.
Federal Aviation Administration
Office of Aviation Medicine
800 Independence Avenue, S.W.
Washington, D.C. 20591

December 1977

Document is available to the public through the
National Technical Information Service,
Springfield, Virginia 22151

Prepared for
U.S. DEPARTMENT OF TRANSPORTATION
FEDERAL AVIATION ADMINISTRATION
Office of Aviation Medicine
Washington, D.C. 20591

12

Note

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

Technical Report Documentation Page

1. Report No. FAA-AM-77-25	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Development of New Selection Tests for Air Traffic Controllers		5. Report Date December 1977
		6. Performing Organization Code
		8. Performing Organization Report No.
7. Author(s) John T. Dailey, Ph.D., & Evan W. Pickrel, Ph.D.		
9. Performing Organization Name and Address Office of Aviation Medicine Federal Aviation Administration 800 Independence Avenue, S.W. Washington, D.C. 20591		10. Work Unit No. (TRAIS)
		11. Contract or Grant No.
12. Sponsoring Agency Name and Address Office of Aviation Medicine Federal Aviation Administration 800 Independence Avenue, S.W. Washington, D.C. 20591		13. Type of Report and Period Covered OAM Report
		14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

This report describes the development of a new Multiplex Controller Aptitude Test for initial screening of FAA Air Traffic Controller applicants. Its content includes the traditional types of aptitude test items used for today's screening. In addition it includes measurement of the ability to identify potential conflicts in air traffic, a skill that has been demonstrated experimentally to have a significant relation to success in the ATC speciality. Alternate forms of the test have been developed in a format that meets Civil Service test administration requirements. The test has been administered experimentally to groups whose abilities approximate those of the applicant population, and results indicate that it has satisfactory reliability characteristics. It has been administered experimentally to incoming students at the FAA ATC Academy and personnel on the job at operational facilities, and constantly produced higher correlations with ATC success than any other test used in the validation studies. The available data indicate that this new and customized instrument promises to be a significant improvement over the existing battery for screening FAA Air Traffic Controller applicants.

17. Key Words Selection Air Traffic Controllers		18. Distribution Statement Document is available to the public through the National Technical Information Service, Springfield, Virginia 22151	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages	22. Price

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

DEVELOPMENT OF NEW SELECTION TESTS FOR AIR TRAFFIC CONTROLLERS

This report describes creation of the Multiplex Controller Aptitude Test, a new measure to screen applicants for Federal Aviation Administration Air Traffic Controller positions. Its content includes items to measure the traditional types of aptitudes such as arithmetic reasoning and visualization that are included in today's Civil Service Commission tests. In addition it includes new job sample items from the controller activity, uses figures to show air traffic on a simulated radar screen, and requires identification of potential conflicts between aircraft in that simulated traffic. All test questions are presented in an air traffic control setting, providing a job-related appearance that gives the test high face validity not found in today's selection battery. The format for item sequencing also is a departure from another current aptitude battery design practice, that of clustering items into homogeneous subgroups. Instead, in these tests the items alternate from one type to another and spiral to increasing levels of difficulty, a mode of presentation found only in a few current tests such as the Standford-Binet. A result shown statistically is that non-confliction items included for measurement of aptitudes represented in today's selection battery show unexpectedly high commonality with the new air traffic items requiring detection of impending conflictions. The resulting measure shows promise of consistently producing higher validity coefficients than possible with the current test battery.

Background

The current Civil Service Commission aptitude test battery for screening air traffic controller applicants has been operational since 1964. In 1970 the Federal Aviation Administration contracted with specialists in the field of personnel selection to search for ways to alleviate problems being experienced in selection and retention of air traffic controllers. They identified other available tests to increase the predictive validity of this battery, plus areas for the construction of new tests. One of these (Buckley, Note 1) made use of a technique developed to evaluate total man-machine system performance in control of air traffic by measuring how well a fully trained specialist could control simulated traffic presented on a film using the proposed designs. In an adaption of this technique for the testing of personnel skills, films of air traffic as seen on a radar scope are presented, and observers asked to identify the aircraft they predict will violate separation standards. This instrument added significantly to a composite for predicting on-the-job success (Milne, Note 2), but the equipment and space required for its administration generally are not available in Civil Service Commission testing situations. Therefore test development procedures were initiated to derive measures of this same skill in a format that meets Civil Service needs.

Test Development

The original Controller Decision Evaluation technique (Buckley, Note 1) required observation of simulated air traffic situations as they unfold in a film of a radar scope. The observer predicts and records potential conflicts as soon as he detects them, in an unstructured, free response mode. During experimental administration and study of three film versions for development of the Multiplex Controller Aptitude Test, it was found that identification of potential conflicts was easier, quicker and more accurate after observers adjusted to the scope and its targets, when only a few targets were in near-confliction, and when their rate of closure was slow. Items were more difficult when the observers were first exposed to the scope and its targets, when multiple targets were in near confliction and when the rate of closure among targets was rapid. False

positive conflict items did not discriminate and generally, for true conflicts, the greater the lead time available, the greater an item's discriminating power. Journeyman controllers identified potential conflicts sooner than developmental trainees, and identified almost all potentials that became real conflicts. Developmental controllers generally missed a number of real conflicts, were slower than journeymen in calling them out, and frequently did not attend to aircraft altitude separations. However, the range in test performance among developmentals was great, and some performed as well as journeymen. Some confliction items in these films were too easy and some had a negative relation to a developmental-journeyman criterion dichotomy.

The observer had much idle time during free response film versions of the test, as the traffic presented is light and aircraft widely separated. In this sense, the setting resembles a common air traffic situation, but an extended amount of testing time is required to present only a small number of confliction items. Better use of examinees' time was desired. An initial change during test development was to structure the items, using slides to present questions on the screen above the film. Mean response times for developmental and journeyman controllers to react to each conflict in the free response versions were determined. Journeyman mean times were selected to be determiners of aircraft positions when presenting confliction questions in the structured versions, for this maximized the discriminating power of each item. Two-choice conflict items, asking whether a pair of aircraft would conflict, were presented for thirty seconds when unexpected changes occurred among targets, such as when new aircraft entered the picture. Four-choice conflict items, asking which pair would conflict if a confliction occurs, were presented for forty-five seconds when traffic changes were slow. A "None of these" response was introduced to permit inclusion of non-conflictions that might be predicted to be conflictions, and this was used as the forth alternative in all four-choice confliction items. Items were assembled into the described test format, but they didn't utilize half the available testing time. This large amount of idle time between questions also provided much opportunity for observers to change their anwers to earlier questions regarding potential conflicts as the aircraft approached each other and the correct answers became more obvious. In the free response film version, this problem was controlled by a requirement for entries from a coded clock onto the computer-scored answer sheet whenever potential conflicts were reported. It was found with the paper and pencil version that presentation of a new item every 45 seconds would keep examinees so busy that they would have no opportunity to make such changes. More test items were needed to make this control procedure work.

The test films presented simulated traffic moving across a radar scope plus a table which includes detail data on each aircraft, identifying the target, its altitude, speed, and route on the scope. The scope uses lines to represent airways or highways in the sky, with alphabetical identifiers of starting, ending, and intersecting points on the airways. The top of the scope is North, and a mileage scale is provided at the bottom. Ample information is presented to prepare a variety of questions related to a controller activity. It appeared possible that most factors used in the current Civil Service Commission aptitude battery for controller selection could be measured within this simulated air traffic setting.

Items were written utilizing this available information to measure such aptitudes as direction following, table reading, interpretation of data, spatial visualization and orientation, estimation of distances and relative target movements, and arithmetic. Some items included were of a very simple nature, and others were written in a multi-factor format to increase their level of difficulty. For example, the directions for this type of test required instruction on how to read the table, so initial table reading questions were included that were of the very easy, instructional type. Awareness of distances across the scope, reading the table to determine the speeds of aircraft, and mathematical computation to determine their rates of closure (all of these being related to horizontal separation) were required to solve a complex problem such as estimating the travel time in minutes between two aircraft at a given moment.

The relation between item types and total test homogeneity was determined, and this ratio was used to determine the number of items per type to include in the test. A result, for example,

was inclusion of twice as many target time-distance separation items as compass heading items in the test. The order of placement of conflict items had already been established by using the mean response time of journeyman controllers when targets were at certain location. New aptitude test items were placed in the remaining positions, alternating from one type to another and spiraling to increasing levels of difficulty as testing progressed. Three alternate forms of the test, MCAT 4, 6, and 7, were prepared.

Test Criterion Characteristics

Versions of the free response CODE test, the structured Multiplex Controller Aptitude Test, and other selection measures were administered to students at the US Navy Air Traffic Controller Training School, Memphis, during the week of August 18–22, 1975. Grades of these students were obtained as they progressed through classroom and laboratory training and passed or failed the course. Distribution statistics, intercorrelations and rotated factor loadings of the various tests with each other and with school grades were determined. Forty-eight variables were involved; five factors were extracted and rotated. Definitions of the rotated factors are presented below.

Rotated Factor I. Variables with highest loadings on this factor include CTO Block I Average, .76, ITEX Final Criterion Exam, .54, and Control Tower Training, Airport, .53. This factor might be defined as Control Tower Operator, the title given by the school to those classroom grades with highest loadings or portion of variance on this factor. Among potential screening measures, a new FAA air traffic controller ATC Occupational Knowledge Test I has the highest loading, .40.

Rotated Factor II. Variables with highest loadings on this factor include Spatial Test, .65,

Table 1

Intercorrelations of Selected Test Variables and Laboratory Grades

USN Air Traffic Control Training School

(109 Persons Tested August 18-22, 1975)

Laboratory Courses	Code	MCAT	OKT-I	GCT	Arithmetic Reasoning
Laboratory Flight Plans					
VFR	05	17	23	10	05
IFR	18	26	22	19	26
Stopover Composite	09	20	28	26	09
Performance Run	19	22	22	25	18
Laboratory Control Tower Training					
Basic	16	20	28	19	21
Intermediate	19	11	07	17	09
Advanced	26	17	27	19	08
Laboratory Radar Training					
Air Surveillance					
Week-1	-02	10	11	17	10
Week-2	-08	04	21	09	12
Precision Approach					
Week-3	08	21	13	14	08
Week-4	16	04	09	11	25
Course Average	22	27	43	34	23
Free Response Code		43	32	12	14
MCAT			42	20	23
ATC Occupational Knowledge				29	24
USN GCT					31

Mechanical Comprehension, .63, and Arithmetic Reasoning, .53. This factor might be defined as Mechanical-Spatial, since the spatial and mechanical comprehension selection tests have highest loadings on this factor. This factor seems to account for very little criterion variance.

Rotated Factor III. Variables with highest loadings on this factor include Course Average, .82, Basic Lab. Control Tower Training, .65, and Stopover Composite Flight Plan Lab, .60. This factor might be defined as Laboratory Performance, since laboratory activities have their highest loadings on this factor. Among potential screening measures, ATC Occupational Knowledge Test I and the Multiplex Controller Aptitude Test, with loadings of .22 and .21 respectively, have the highest loadings on this factor.

Rotated Factor IV. Variables with highest loadings on this factor include Multiplex Controller Aptitude Test, .58, Free Response Code, .49, ATC Occupational Knowledge Test I, .46, Navy General Classification Test, .40, ACT Occupational Knowledge Test II, .36, Charts & Publications, Base Ops., .36, and Precision Approach Radar Lab., .30. This factor might be defined as Air Traffic Controller Performance, since selection tests involving the controller activity and classroom and laboratory performance variables have significant loadings on this factor.

Rotated Factor V. Variables with highest loadings on this factor include Reading, .70, Radioman, .58, and Precision Approach Radar Lab., .52. This factor might be defined as Reading, since that variable has the highest loadings on this factor.

Selected Intercorrelations

Selected data are presented in Table 1 to highlight three of the experimental selection tests being studied here, their interrelations with each other and with laboratory Flight Plans, Control Tower, and Radar Training, as those grades provide the greatest range in criterion scores for differentiating performance among FAA student controllers. ATC Occupational Knowledge Test I, a measure of each person's past controller experience, showed highest relation to performance in laboratory classes. The thirty minute Multiplex Controller Aptitude Test generally had correlations with criterion variables that are as high or higher than those between the Free Response CODE test and those same criterion variables. These results encouraged further development of the Multiplex Test.

Table 2

Selection Test Variables and School Grades

Selected Intercorrelation Data

USAF Air Traffic Control Training School

(461 Persons Tested October 20-24, 1975)

Test and Course Variables	1	2	3	4	5	6	7	8	9	10	11	12
1. Final Score in Block- I		34	41	39	-32	08	35	15	20	19	42	-01
2. II			49	31	-33	07	21	15	21	30	15	09
3. III				45	-36	28	33	16	17	27	32	39
4. IV					-51	28	06	18	07	01	68	01
5. No. Counseling Sessions						-12	-25	02	-05	-04	11	13
6. CODE 7- (Free Response)							*	28	*	*	*	*
7. MCAT 606AS								*	61	59	*	*
8. 606FS									66	*	52	38
9. 706FS										57	41	20
10. 406FS											*	*
11. Occupational Knowledge Test- I												
12. II												

*No Cases

Media For Presenting Items

A next step in test development was to convert the slides plus film presenting moving targets to a "slide only" presentation, in essence capturing pictures of the scope with targets in the same position as when each question appeared on the screen. Pacing of slide presentation was the same as in the film plus slide version, so the amount of target movement from question to question was unchanged. This format was easier for test administrators since they no longer had to cope with film projection problems, and the stationary targets presented by slide seemed easier to read than the moving targets presented by film. Persons taking the experimental version in this slide format did grumble that the test seemed to take control over their time, as each item had to be completed quickly before it left the screen and a new item appeared. This task characteristic has some commonality with controller situations in the real world, as traffic movement tends to control their involvement and pace of work. Another merit of this version would be its case for transition into a paper and pencil format that would meet Civil Service Commission requirements for use in their very decentralized field testing situations. Films plus slide and "slide only" versions of the Multiplex Controller Aptitude Test plus other selection measures were administered to students at the USAF Air Traffic Controller Technical School, Keesler Air Force Base, Mississippi, during the week of October 20–24, 1975. Grades of these students were obtained as they progressed through classroom and laboratory training and passed or failed the course. Distribution statistics and intercorrelations of the various tests with school grades were determined. Twenty-one variables were included, and selected data are presented in Table 2. ATC Occupational Knowledge Test I, a measure of the individual's past controller experience, shows highest selection test relation to final scores in training. The slide version of Multiplex generally shows higher correlations with the criteria, Block I–IV Final Score, than CODE 7 (Free Response) and the combination film-slide versions of MCAT.

Table 3

Multiplex Controller Aptitude Test

Distribution Statistics for Various Populations

	FAA ATC Academy Classes				USAF ATC School Students				FAA Employees	
	January 1976		March 1976							
	Enroute	Terminal	Enroute	Terminal					College	High School
Tests	N=91	N=106	N=87	N=106	N=316	N=141	N=289	N=62	N=18	N=22
MCAT 606 (43 Items)										
FS Mean					28.81				30.73	17.45
SD					5.32				6.15	3.44
AS Mean						27.35				
SD						6.08				
A Mean	35.16	35.24								
SD	3.90	3.95								
B Mean			32.22	32.12						
SD			5.53	4.71						
MCAT 706 (53 Items)										
FS Mean							30.51			
SD							5.83			
AS Mean	39.96	37.11								
SD	4.40	8.68								
A Mean			40.20	39.15						
SD			7.48	7.41						
B Mean									36.23	17.23
SD									8.49	4.32
MCAT 406 (41 Items)										
FS Mean								20.47		
SD								4.87		
A Mean									30.50	13.89
SD									4.08	4.96

FS: Film-slide version, with time controlled for working each individual item.
AS: Slide version, with time controlled for working each individual item.
A: Paper-and-pencil version, with time controlled for working total test.
B: Paper-and-pencil version, with time controlled for working each five-minute group of items.

A next step in test development was to print the slide versions in paper and pencil format. The slide version and combination film-slide version permitted time control at the individual item level, allowing thirty seconds for response to two-choice items and forty-five seconds for response to four-choice items. One paper and pencil version allows uninterrupted work throughout testing time, with announcements when fifteen minutes have elapsed and when only five minutes remain to work on the test. A segmented paper and pencil version allows five minutes for response to each cluster of items presented, with a cluster containing as many as eleven items when they are all of the two-choice type. Three parallel forms of the written test were prepared, MCAT 406 including 41 items, 606 containing 43 items and 706 containing 53 items. Time limits are 25 minutes plus directions for MCAT 406 and 606 and 30 minutes plus directions for MCAT 706. Each form now has been extended to 55 items. MCAT 407 includes 23 conflict and 32 aptitude items, MCAT 607 contains 22 conflict and 33 aptitude items, and MCAT 707 contains 23 conflict and 32 aptitude items. Time limits are 35 minutes for each of these tests.

Reliability

The questions presented in Multiplex Controller Aptitude Test, Form 4, 6 and 7, remained unchanged for some length of time, though the media used for presenting items changed from film to slide to paper and pencil. This change in media for presenting items has some effect on an individual's test performance, but all forms are qualitatively similar. Correlations of these parallel forms provide a minimum estimate of the measure's test-retest reliability when administered to members of a similar population.

These forms have been administered to students of the USAF Air Traffic Controller Technical School in all phases of training and to a sample of forty non-controller high school and college level FAA employees, groups whose abilities approximate those of the general population of applicants that take the Civil Service Commission tests. They also have been administered to entering FAA Air Traffic Controller Academy students in the new centralized Air Traffic Controller training program, persons selected off the top of the Civil Service Register for controller applicants. Table 3 presents distribution statistics for these groups on the various forms of the test. The range in test scores is highly restricted for the ATC Academy group when compared to the range for the more general population.

Correlations between forms were lowest, .31 to .50, for the ATC Academy students who had a greater restriction in range of their scores. They range from .60 to .66 for the USAF ATC students and .87 to .90 for non-controller FAA employees whose abilities more nearly approximate that population for whom the test is designed. These coefficients indicate that the test will provide reliable measures of performance under well controlled conditions, but a combining of items from two forms would provide a more satisfactory measure to use in an operational testing situation.

Test Validity

For test validation purposes the criterion used was "success" in air traffic control work. Success as defined here for ATC applicants is hierarchical, including (1) satisfactory completion of the initial, formal training program, (2) satisfactory performance on the job, and (3) progression or upward mobility within the ATC system. Another element, (4) attrition, may be a measure of non-success. Those enrolled in the initial, formal training program are a highly select group, for they scored high enough on the Civil Service Commission tests to be hired, but group membership becomes even more selective as those who fail to learn or perform adequately are separated. This selection process continues as they progress up the ATCS career ladder and are evaluated for satisfactory performance on the job and advancement into the more demanding and higher paid positions within the ATC system. The range in scores on these selection tests will be greatest for the group when they first enter formal training, but become ever more restrictive with career progression. As this restriction in range may have a direct effect on the size of the validity coefficients, the validities should be highest when the group is in initial training and become lower for that portion of the group which progresses up the carrer ladder.

Table 4

Correlations Between Selection Tests and Various Criteria

Criteria	Number of Cases	Selection Tests: Arithmetic Reasoning	MCAT Conflicts	MCAT Aptitudes	Directional Headings	Dial Reading	Selected Battery
Training Scores[a]							
Terminal Option	301-310	136*	321**	256**	191**	288**	338**
Enroute Option	257-263	267**	434**	402**	218**	321**	473**
Both	558-573	202**	370**	323**	202**	257**	395**
Supervisory Rating[b]							
Flight Service Station	195-245		254**	276**		204*	293**
Terminal VFR	169-190		227**	220**	221**		289**
IFR	188-188		154**	048			166
Enroute	181-241		075	125		026	128
All Options	733-833		156**	151**	123**		176**
ATC Progression[c]							
Flight Service Station	157-159						
Terminal VFR	191-193			231*		075	232**
IFR	179-179	148	118	133	120		178
Enroute	199-200	107	298**	285**	170	280**	357**
All Options	727-731	058	094*	117*		074	119*
Aggregate Criterion[d]							
Flight Service Station	196-249		220**	262**		197**	272**
Terminal VFR	479-518		254**	195**	179**	157**	275**
IFR	499-623		235**	178**	146**	134**	246**
Enroute	445-514		253**	204**	158**	166**	287**
All Options	1309-1603		238**	223**	174**	149**	265**

Note. Decimals Omitted *$p < .05$ **$p < .01$
Note. Data are from Mies, J.M., Colmen, J.G. and Domenech, O., May 1977.
a Table XI. 1
b Table XII. 2
c Table XIII. 2
d Table XV. 2

Success in Training. A battery of predictors including Multiplex Controller Aptitude Test 606A and 706AS, scored separately for Conflicts and Aptitudes, plus Directional Headings, Dial Reading and Arithmetic Reasoning tests because of the promise they demonstrated in an earlier study (Milne, Note 2), were administered to the January 1976 Class of students entering the ATC Academy centralized training course. Persistent mechanical problems in operation of film, slide and slide programmer equipment gave strong support to the Civil Service Commission's stated need for development of these measures in a paper and pencil format. Thus the March 1976 class of students were administered only paper and pencil versions MCAT 706A and 606B, plus the other selection measures. Satisfactory completion of the initial, formal training program for those enrolled in the Enroute option was based on scores received on four ATC laboratory problems plus scores on a Controller Skills Test. For those enrolled in the Terminal course, performance was based only on the ATC laboratory problems. These scores require students to demonstrate operational application of academic knowledge. Correlations between selection tests and training scores are presented in Table 4. The MCAT Conflict and Autitude scores constantly produced higher validities than any other test used. In fact, addition of the other tests increased the multiple correlation by only .017 for Terminal, .012 for Enroute, and .008 for both options combined.

Performance on The Job. Test validation in the operational situation, test correlation with supervisory ratings and career progression scores, was accomplished as part of a larger ATC selection research effort (Mies, Note 3). A comprehensive sampling design was constructed to define the ATC population, constraints were established, and stratified random sampling methods then applied to select the primary samples of specialists from within the total available ATC population. On-the-job performance was measured by confidential job-task assessments prepared by each employee's supervisor,

Table 5

Integer Weights Assigned to Tests

By ATC Option and All Options Combined

ATC Options	MCAT Conflicts	MCAT Aptitudes	Directional Headings	Dial Reading
FSS	18	40	(0)	7
VFR	55	10	21	(0)
IFR	49	8	13	(0)
ARTCC	33	34	7	5
All Options	37	21	15	(0)

Note. Data are from Mies, J. M., Colmen, J. G. and Domenech, O., May 1977, Table XV. 3

All data were collected at ATC facilities. Correlations between selection tests and supervisory ratings are presented in Table 4. Correlation between MCAT Conflict and Aptitude scores and supervisory ratings are significant ($p. \leq .01$) for the FSS and Terminal VFR options and for all options combined. The correlation between MCAT Conflict scores and Terminal IFR supervisory ratings also was significant ($p. \leq .01$). The MCAT Conflict and Aptitude variables were prominent predictors in most regression equations which predicted job performance "success," although neither predicted at significant levels of confidence for the Enroute option. Directional Headings appeared valid only for the VFR option and for all options combined. Dial Reading entered for FSS only.

ATC Progression. ATC Progression was measured by comparison of the ATC option to which the specialist was initially assigned when hired to the option assigned on January 1, 1976. A progression from FSS to Terminal VFR and IFR and Enroute options was used to represent levels of increasing complexity. A progression of "high" was assigned within this hierarchy when a specialist was in an option of a complexity level the same as or higher than the initial option assigned. A progression value of "low" was assigned when a specialist was in an option of a lower complexity level than the option to which initially assigned. Correlations between selection tests and progression scores are presented in Table 4. Again the MCAT Conflicts and Aptitudes were prominent predictors.

Aggregate Criterion. An aggregate criterion of ATC "success" was constructed from combinations of the four individual criteria (training, on-the-job performance, progression, and attrition) and provided a five point scale value for ATC success. The Arithmetic Reasoning Test was valid only against the progression criterion, and was dropped from the battery. The Directional Headings Test was dropped for FSS, since it failed to enter the regression for supervisory assessment. Correlations between selection tests and the aggregate criterion are presented in Table 4. All correlations are significant ($p. \leq .01$) for each of the four options and for all options combined. The Dial Reading Test did not enter the regression for the VFR option and entered last with a negative "b" weight for IFR and for all options combined. The integer weights assigned to tests are summarized in Table 5. It is evident that MCAT (Conflict and Aptitude segments) is the major factor in the validities (R) derived from the multiple regression analysis.

Summary

A new test has been developed for initial screening of FAA Air Traffic Controller applicants. Its content includes the traditional types of aptitude test items found in today's Civil

Service screening battery. In addition, it includes a measure of the ability to identify potential conflicts in air traffic, a skill that has been demonstrated experimentally to have significant relation to success in the FAA ATC speciality. All test questions are preesnted in an air traffic control setting, which gives them a job-related appearance not found in today's selection battery. Alternate forms of the test have been developed in a paper-and-pencil format to meet Civil Service needs in their decentralized testing program. The test has been administered experimentally to groups whose abilities approximate those of the applicant population, and correlations between alternate forms indicate that it has satisfactory reliability characteristics. It has been administered experimentally to students entering the new FAA ATC Academy and personnel on the job at operational facilities, and constantly produced higher correlations with ATC success than any other test used in the validation studies. The available data indicate that this new and customized instrument, when used in combination with other selected measures, promises to be a significant improvement over the existing battery for screening FAA Air Traffic Controller applicants.

Reference Notes

1. Buckley, E. P. & Beebe, T. *The development of a motion picture measurement for aptitude for air traffic control.* FAA-RD-71-106. Atlantic City, New Jersey: National Aviation Facilities Experimental Center, January 1972.
2. Milne, A. M. & Colmen, J. G. *Selection of air traffic controllers for Federal Aviation Administration.* Final Report, Contract DOT-FA70WA-2371, January 1972.
3. Mies, J. M., Colmen, J. G. & Domenech, O. *Predicting success of applicants for positions as air traffic control specialists in the Air Traffic Service.* Final Report, Contract DOT-FA72WA-3646, May 1977.