AM 68-14

A COMPARATIVE STUDY OF AIR TRAFFIC TRAINEE APTITUDE-TEST MEASURES INVOLVING NAVY, MARINE CORPS AND FAA CONTROLLERS

Bart B. Cobb, M.S.

Approved by

J. Robert Dille, M.D.
Chief, Civil Aeromedical Institute

Released by

P. V. Siegel, M.D.
Federal Air Surgeon

September 1968

Department of Transportation
FEDERAL AVIATION ADMINISTRATION
Office of Aviation Medicine

FOOTNOTE

This study is from the Civil Aeromedical Institute, Federal Aviation Administration, Oklahoma City, Oklahoma. The author is indebted to Lieutenant Commander F. J. Gallagher, Glynco NAS Air Controlman Training School Officer and Mr. Robert F. Jordan, FAA Air Traffic Representative, for arranging the study and for the collection of data at the Glynco facility.

Opinions or conclusions contained in this report are those of the author and do not necessarily reflect the views or endorsement of the Federal Aviation Administration.

Qualified requestors may obtain Aviation Medical Reports from Defense Documentation Center. The general public may purchase from Clearinghouse for Federal Scientific and Technical Information, U.S. Dept. of Commerce, Springfield, Va. 22151.

A COMPARATIVE STUDY OF AIR TRAFFIC TRAINEE APTITUDE-TEST MEASURES INVOLVING NAVY, MARINE CORPS AND FAA CONTROLLERS

I. Introduction.

Since 1964 the standards for the selection of personnel for training in air traffic control (ATC) work with the Federal Aviation Administration (FAA) have included an assessment of specific mental aptitudes, skills or abilities. A qualifying aptitude index, based upon a U.S. Civil Service Commission (CSC) battery of six subtests, constitutes the first of several eligibility requirements. This screening index is an outgrowth of extensive previous research[1,2,3,4] in which the six CSC test measures were identified from among 27 different variables as representing the best composite of aptitude scores for the prediction of training and job performance. Selection for controller training is also contingent upon an assessment of pre-employment experience, educational background, a medical examination, and an interview with management officials. Previous relevant experience, such as military air traffic control, is one of the more heavily-weighted selection factors. The qualification of a candidate having no previous ATC experience requires a much higher aptitude-screening index and either a college degree or an exceptional career background. The FAA consequently recruits most of its trainees from an applicant pool of former servicemen. Aside from other reasons, the implications arising from these recruiting practices are of sufficient import to warrant FAA interest and support in the development and effectiveness of ATC-selection-and-training programs of the United States Air Force, Army, Navy and Marine Corps.

The present investigation was undertaken on a cooperative basis with officials of the Glynco Naval Air Station (NAS), Georgia. The Glynco Training School has the responsibility of providing instruction and basic training for air traffic control personnel for the Navy and the Marine Corps. Plans for the study were first conceived in 1965 when officials of the Naval facility visited the FAA's Civil Aeromedical Institute (CAMI) for further indoctrination regarding the underlying research, development, and effectiveness of aptitude tests and other screening procedures used in the selection of applicants for FAA controller training. Naval officials subsequently discussed the possibility of being permitted the use of the operational battery of CSC aptitude tests for experimental administration and validation on ATC trainee groups at Glynco. Inasmuch as several policy reasons precluded this approach, CAMI scientists suggested an alternate approach involving the use of seven commercially-published test instruments which had been validated in previous research with FAA trainees.

Although only limited validation data were available for the seven tests as a group, each had been identified in previous phases of CAMI research[1,2] as a significant predictor of training-course grades for FAA-ATC students. Fortunately, one phase of this research had involved the experimental administration of both the complete CSC-ATCS-Aptitude Test Battery (before its adoption for operational screening) and four of the seven recommended tests to a sample of almost 800 FAA trainees. An analysis of these data revealed a correlation of .74 between the composite scores of the CSC battery and the four commercial tests. With course grade employed as the criterion, validity coefficients for the separate subtests of the CSC battery were found to range from .28 to .56 while those of the four non-CSC tests ranged from .35 to .54. A correlation of .52 was obtained between the training course grade and the overall score for all tests of the CSC composite. The corresponding validity coefficient for the four-test composite was .50.

On the basis of these and other supportive data, it was assumed that the seven commercially-published tests would constitute an appropriate

battery for experimental study at the Glynco training facility. The administration and validation of this battery at Glynco was the major objective of the present investigation. Other objectives included: (a) determination of possible differences between the aptitude levels of the Navy and Marine Corps trainees; (b) comparative evaluation of the experimentally-deprived aptitude measures versus the military-screening-and-classification (MSC) scores as predictors of training-course performance; (c) determination of the relationship between chronological age and aptitude performance; and (d) the establishment of projected estimates regarding the proportion of Navy and Marine students who (considered as potential applicants for FAA-ATC training) would be able to qualify on the CSC-ATC Aptitude Screening Battery. In order to facilitate a more direct comparison of Navy and Marine ATC trainees with those of the FAA, the design of the study was extended to include consideration of data for two groups of former FAA-ATC trainees.

II. Procedure.

In implementing the study, CAMI assumed responsibility for providing copies of the test booklets, the accomplishment of all analyses of data, and preparation of a report of findings. Glynco officials were responsible for the experimental administration and scoring of the tests and the collection of correlative data. Test scores, military - screening - and - classification (MSC) test data, chronological ages, and training-course performance information for over 950 Glynco students were subsequently forwarded to CAMI for analysis.

Samples. The two groups of FAA-ATC subjects selected for inclusion in the study represented the class inputs for the Aeronautical Center's Basic Terminal-Area-Traffic-Control (TATC) Training Course during the period September 1960 through June 1962. Even though it has since been discontinued (i.e., replaced by a program in which recruits receive intensified training and instruction within their facility of assignment), this former TATC training course was highly similar to that provided at the Glynco NAS. Both groups of FAA-TATC students had participated in a CAMI experimental aptitude-assessment program at the time of entry into training. As will be pointed out later however, the two groups had not been assessed with the same battery of aptitude tests.

The data represented in the various phases of the present investigation pertain to a total of 395 former students of the FAA-TATC Training Course and 959 Navy and Marine trainees of the Glynco NAS ATC Training Course. Two hundred and twelve FAA-ATC subjects who entered training during September 1960 through July 1961, designated as Sample 1, had been assessed with an aptitude test battery which included the seven commercially-published tests. The remaining 183 FAA-ATC cases, Sample 2, represent TATC classes of August 1961 through June 1962, and had been examined with both the complete CSC-ATC-Aptitude-Screening Battery (prior to its operational implementation) and a uniform set of four of the seven commercial instruments selected for use at Glynco. All but four of the 963 students who entered the Glynco Course from March 1966 through February 1967 were administered the uniform battery of seven commercial tests on an experimental basis at the time of entry into training. A group of 649, consisting of 435 Navy trainees and 214 Marines who entered training during the earlier two-thirds of the indicated time period, were designated as Sample 3. The remaining 310 Glynco cases, representing 207 Navy and 103 Marine trainees, were assigned to the fourth sample. In order to compare the Navy versus Marine students, Samples 3 and 4 were further divided into subgroups designated as "3a-Navy (N=435)," "3b-Marine (N=214)," "4a-Navy (N=207)," and "4b-Marine (N=103)."

Aptitude Tests. As shown in Table 1, the seven commercially-published aptitude tests constituted a uniform battery used in the assessment of the FAA subjects of Sample 1 and the Navy and Marine trainees of Samples 3 and 4. Three of the seven are subtests of the Psychological Corporation's well known Differential Aptitude Test (DAT), namely "DAT-Space Relations," "DAT-Numerical Ability," and "DAT-Abstract Reasoning." The remaining four, which represent subtests of the California Test Bureau's Test of Mental Maturity (CTMM, Advanced Form A edition), are referred to as "Analogies," "Inference," "Numerical Quantity — Coins" (which involves the mental manipulation of varying monetary amounts), and "Numerical Quantity-Arithmetic." The data available from previous research for the FAA subjects (Sample

TABLE I. Designation of Aptitude Test Scores Available For each Sample of The Study

	Sample 1 FAA N_t=212 N_p=193 N_f=17	Sample 2 FAA N_t=183 N_p=143 N_f=40	Sample 3 MILITARY N_t=649 N_p=581 N_f=64	Sample 4 MILITARY N_t=310 N_p=259 N_f=38
COMMERCIAL TESTS				
DAT Space Relations	X	X	X	X
DAT Numerical Ability	X	X	X	X
DAT Abstract Reasoning	X	X	X	X
CTMM Analogies	X	X	X	X
CTMM Inference	X		X	X
CTMM N. Q. Coins	X		X	X
CTMM N. Q. Arithmetic	X		X	X
CIVIL SERVICE COMMISSION TESTS				
CSC 51-Spatial Patterns		X		
CSC 24-Computations		X		
CSC 157-Abstract Reasoning		X		
CSC 157-Letter Sequence		X		
CSC 135-Oral Directions		X		
CSC 540-ATP I+II		X		
TYPES OF COMPOSITE SCORES DERIVED				
Commercial 7-test Composite	X		X	X
Commercial 4-test Composite	X	X	X	X
New 4-test Commercial Composite			X	X
CSC 6-test Composite		X		

2 of the present study) included scores on four of the seven commercial tests and the complete CSC-ATC-Aptitude-Screening-Test. The four commercial instruments which were administered to Sample 2 were the DAT subtests of Space Relations, Numerical Ability and Abstract Reasoning and the CTMM subtest Analogies. The six subscores of the CSC-ATC Test relate to the subtests of Spatial Patterns, Computations, Abstract Reasoning, Letter Sequence, Following Oral Directions, and Air Traffic Problems.

Military - Screening - and - Classification - Test Scores. The Navy and the Marine Corps both employ MSC procedures which were originally designed to yield a mean of 100 for a military population. However, the procedures are different. The Navy MSC index, known as the "GCT+ARI," is based on the General Classification Test and a test of Arithmetical Reasoning. The Marine index, "⅓ (AR+VE+PA)," represents an average of performance scores on the tests of Arithmetical Reasoning, Verbal Ability, and Perceptual Ability. According to Glynco officials, an index of 110 is a recommended standard for the selection of both Navy and Marine ATC trainees. Findings in this study indicated that the Marines adhered to this policy while the Navy made 104 exceptions in selecting 642 trainees.

Criteria. The two criterion variables included in the present study were "Overall Training Course Grade Average" and "Pass-Fail Status." These comprehensive performance indices represented the only measures which were similar and directly comparable for the FAA and Glynco trainees.

Factor Weights for Derivation of Composite Aptitude Measures. Composite aptitude measures for all Ss were derived through application of sets of factor weights developed previously from performance data for several hundred FAA-ATC trainees. The FAA-ATC samples used here represent only a minor portion of the previous sample. The factor weight for each test was the inverse of the standard deviation of scores (i.e., 1/SD) obtained for the larger group. The omission of decimals to avoid fractional values had resulted in a two-digit weight for each variable. (The application of weights so developed is essentially equivalent to the equal

weighting or simple addition of test performance measures rendered in standard score form.) In the present study, the original two-digit factor weights were applied to derive two of the three composite-aptitude measures involving the commercially published instruments. These are referred to as the "Commercial 7-Test Composite" and the "Commercial 4-Test Composite." In contrast, the "New 4-Test Commercial Composite" and the "CSC 6-Test Composite" are based on single-digit factor weights. In each of the latter instances however, the weights represent a simple and proportionate reduction of the original set of two-digit weights to single-digit values. Those pertaining to the CSC tests are identical to the factor weights prescribed by the U.S. Civil Service Commission.

III. Results.

Numerous analyses were accomplished in this study and the results constitute a large body of data. The order in which the results are presented reflects a compromise between the sequence in which the statistical analyses were undertaken and the relevance of results across different analyses to one or more objectives of the study. For example, the empirical validities of the various aptitude tests, when used separately and in combination for prediction of Training Course Grade and Pass-Fail Status, are presented simultaneously in Table II. Even though coefficient

TABLE II. Empirical Validities of Aptitude-Test Variables

	Sample 1 FAA		Sample 2 FAA		Sample 3 MILITARY		Sample 4 MILITARY	
	Course Grade	P-F	Course Grade	P-F	Course Grade	P-F	Course Grade	P-F
	N_t	N_p-N_f	N_t	N_p-N_f	N_t	N_p-N_f	N_t	N_p-N_f
	r	r_{pb}	r	r_{pb}	r	r_{pb}	r	r_{pb}
COMMERCIAL TESTS	N=211	193-17	N=183	143-40	N=636	581-64	N=297	259-38
DAT Space Relations	.14*	.20**	.44**	.37**	.31**	.18**	.25**	.12*
DAT Numerical Ability	.33**	.19**	.36**	.24**	.38**	.22**	.40**	.27**
DAT Abstract Reasoning	.40**	.38**	.47**	.32**	.29**	.14**	.36**	.23**
CTMM Analogies	.12	.09	.27**	.25**	.15**	.11**	.08	.04
CTMM Inference	.27**	.25**			.24**	.21**	.13*	.12*
CTMM N. Q. Coins	.31**	.25**			.27**	.19**	.13*	.09
CTMM N. Q. Arithmetic	.38**	.16*			.33**	.19**	.31**	.24**
CIVIL SERVICE COMMISSION TESTS			N=183	143-40				
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-ATP I+II			.41**	.29**				
COMPOSITE SCORES	N=211	193-17	N=183	143-40	N=636	581-64	N=297	259-38
Commercial 7-test Composite	.39**	.31**			.42**	.26**	.39**	.26**
Commercial 4-test Composite	.32**	.29**	.52**	.40**	.39**	.23**	.40**	.25**
New 4-test Comm. Composite					.41**	.25**	.42**	.27**
CSC 6-test Composite			.54**	.39**				
MILITARY (MSC) SCORES								
GCT + ARI for Naval Ss					N=422	379-51	N=197	171-26
					.42**	.28**	.47**	.27**
1/3 (AR+VE+PA) for Marine Ss					N=184	173-12	N=74	62-12
					.18*	.01	.27*	.18

*p<.05
**p<.01

for all aptitude measures for all samples are shown, only those relating to Samples 1, 2 and 3 were determined in the initial stages of the study. Sample 4 was used, in a restricted sense, for cross-validation purposes.

Empirical Validities of the Commercially-Published Tests. The empirical validities of the commercially-published instruments appear in the upper portion of Table II. A comparison of the data for the first three samples reveals that the validity coefficients for some tests vary in magnitude from sample to sample. Yet, with the exception of CTMM Analogies for Sample 1, all test validities established with both FAA samples and the Glynco subjects of Sample 3 proved to be statistically significant. The validities obtained for Sample 1 are generally lower than those derived for Samples 2 and 3, and in most instances, are considerably lower than those previously established with more comprehensive groups of FAA trainees. In contrast, the test coefficients obtained for Sample 2 are substantial and within an expected range. With the exception of CTMM Analogies, this is also true for most of the test validities obtained with the Glynco trainees of Sample 3. Further comparison of the corresponding data across the different samples indicates that each test measure tended to correlate more highly with Course Grade than with Pass-Fail Status. This is not surprising inasmuch as the training-course-failure cases represent only 8.1 per cent of Sample 1, 21.9 per cent of Sample 2, and 9.9 per cent of Sample 3. In fact, the failure rates for Samples 1 and 3 are so low that the point-biserial coefficients could legitimately have been omitted.

Comparison of Empirical Validities of CSC and Non-CSC Tests for Sample 2. Within the context of the present study, performance scores on each of the six subtests of the CSC-ATC Aptitude Battery were available only for those FAA subjects designated as Sample 2. The experimental testing of these former FAA-TATC trainees had also included the administration of the three DAT subtests and CTMM Analogies. Validity data for the various aptitude measures for Sample 2 appear in Table II. Validities of the CSC subtests range from .23 to .55 for Course Grade and from .16 to .45 for the Pass-Fail criterion. All are statistically significant. However, the validities for each of the four non-CSC tests are in most instances even better than those derived for the CSC measures. Though not presented in any table of this report, intercorrelations of the various test measures were also computed for this sample. For example, a correlation of .72 was obtained between the scores of the DAT Space Relations Test and CSC Spatial Patterns. Such a correlation lends to the hypothesis that the similarly-named tests are also similar in factor content. In this same respect, DAT Numerical Ability was found to correlate .51 with CSC-Computations and a coefficient of .48 was obtained between DAT Abstract Reasoning and CSC Abstract Reasoning.

Various composite scores, involving different groups of test measures (and based on the application of previously established factor weights as described in preceding portions of this report) were derived for the subjects of the respective samples. Two such global scores were developed for every FAA subject in Sample 2. One of the two, involving the three DAT subtests and CTMM Analogies, is referred to as the "Commercial 4-Test Composite." The other is the "CSC 6-Test Composite". The validities of the two composites (see Table II) are almost equivalent. With Course Grade as the criterion, the validity of the Commercial 4-Test Composite is .52 while that of the CSC composite is .54. The corresponding validities with the Pass-Fail criterion are .40 and .39. The comparability of the validity data for the separate and combined subtests, the inter-group correlations for the similarly named but different subtests, and the correlation of .81 between the two composite scores of Sample 2 constitute convincing evidence of a high degree of equivalence in the two batteries. Inasmuch as Samples 1, 3, and 4 were not administered the CSC battery, this correlation of .81 (not presented in a table) provided the basis for an equation used in later phases of the study to estimate CSC-ATC Test performance from scores attained on the four commercial tests.

Empirical Validities of "Commercial 7-Test Composite" and "Commercial 4-Test Composite" for Samples 1 and 3. The uniform battery of tests administered to the FAA trainees of Sample 1 and the Glynco trainees of Samples 3 and 4 included all seven commercial instruments but none of the CSC subtests. Two composite measures were computed at this stage of the study for every subject represented in Samples 1 or 3.

The first measure, based upon all seven tests, is referred to as the "Commercial 7-Test Composite." The second, based on the four commercial tests which had been administered (along with the CSC battery) to Sample 2, is referred to as the "Commercial 4-Test Composite." This composite is the same as previously described and computed for Sample 2. The validity coefficients for both composites appear in the lower portion of Table II. Validities of the seven-test composite for Course Grade are .39 and .42 for Samples 1 and 3, respectively. The corresponding validities of .32 and .39 for the four-test composite compare very favorably with those obtained for the complete seven-test battery. This is also true for the Pass-Fail validities, which (for either composite) are lower than those obtained with the Course Grade criterion. Validities of the two composites for Samples 1 and 3 are, without exception, lower than those obtained for Sample 2 with the four-test composite. Nevertheless, all of the validity coefficients are statistically significant.

Development of "New 4-Test Commercial Composite" For Maximized Validity. In view of the frequency of low and questionable validities for the CTMM Analogies Test, the specific combination of selected variables for the most valid composite was determined. Using intercorrelations and validity coefficients obtained for Sample 1, a multiple-regression analysis was undertaken. This analysis identified DAT Space Relations, DAT Numerical Ability, DAT Abstract Reasoning and CTMM Inference as the most effective predictor combination. A similar analysis on the Glynco trainees, Sample 3, indicated the same four measures as the most valid composite. For both samples, the multiple correlation ("R", or estimate of maximum possible validity) of the new four-test composite with Course Grade was found to be .45. With Pass-Fail employed as the criterion, an R of .39 was found for Sample 1 and an R of .31 for Sample 3. The new composite scores, derived through application of single-digit factor weights (described earlier), were computed for Sample 3. These correlated .41 with Course Grade and .25 with Pass-Fail Status. These coefficients of .41 and .25 approximate the indicated multiple correlations of .45 and .31 obtained by multiple regression analysis and thus serve to illustrate the appropriateness of the modified set of factor weights.

Validation of Tests and Composite Aptitude Measures with Sample 4. Complete records, consisting of performance scores for the entire group of seven commercial tests, training-course grades and pass-fail information, were available for 297 of the 310 Navy and Marine trainees. Validity coefficients of the separate subtests were obtained; three composite scores, the "Commercial 7-Test Composite," the original "Commercial 4-Test Composite," and the "New 4-Test Composite," were derived; and the validities of the three composite measures were determined. The validities of the separate and combined test measures are presented in Table II. The data indicate that CTMM Analogies has less predictive potential than any other test. With the exception of Number Quantity-Arithmetic, the validities of all the CTMM subtests are rather low. However, the correlations of the three DAT measures with Course Grade, which range from .25 to .40, are not appreciably different from those obtained with Sample 3. The validities of the three composite measures range from .39 to .42 for Course Grade and from .25 to .27 for pass-fail. Validities of the "New Commercial 4-Test Composite" are only slightly higher than those obtained for the other two composites.

Comparison of Empirical Validities of All Composites for All Samples. A review of all validity coefficients in Table II reveals no appreciable differences between the corresponding validities of the three composites for any sample. The validities of each composite are also remarkably similar across all samples. All the composite measures correlate about .40 with the training-course grades of the Glynco subjects, Samples 3 and 4, yet the data indicate that the three DAT tests and CTMM Inference constitute the best predictor composite.

Validities of the Military-Screening-and-Classification Test Scores. The empirical validities of the MSC Test scores for Glynco trainees, Samples 3 and 4, are shown in the lower portion of Table II. Since the MSC scores were prerequisites for trainee selection, it is important to recognize that these samples are probably quite restricted in range and that these validity coefficients may grossly underestimate the actual validities. (Each correlation presented is uncorrected for "restriction of range" effects.) The validities for the Marine MSC score range from .27 to .01 but probably have little meaning

TABLE III. Comparative Validities of Aptitude Tests and Composites for FAA, Navy, and Marine Corps ATC Trainees

	Sample 1 FAA		Sample 2 FAA		Sample 3a NAVY		Sample 3b MARINE		Sample 4a NAVY		Sample 4b MARINE	
	Course Grade	P-F	Course Grade	P-F	Course Grade	P-F	Course Grade	P-F	Course Grade	P-F	Course Grade	P-F
	N_t r	N_p-N_f r_{pb}	N_t r	N_p-N_f r_{pb}	N_t r	N_p-N_f r_{pb}	N_t r	N_p-N_f r_{pb}	N_t r	N_p-N_f r_{pb}	N_t r	N_p-N_f r_{pb}
COMMERCIAL TESTS	N=211	193-17	N=183	143-40	N=423	380-51	N=213	201-13	N=198	172-26	N=99	87-12
DAT Space Relations	.14*	.20**	.44**	.37**	.31**	.18**	.26**	.12	.23**	.10	.31**	.16
DAT Numerical Ability	.33**	.19**	.36**	.24**	.37**	.23**	.30**	.10	.41**	.27**	.43**	.29**
DAT Abstract Reasoning	.40**	.38**	.47**	.32**	.30**	.15**	.23**	.05	.34**	.23**	.42**	.21*
CTMM Analogies	.12	.09	.27**	.25**	.12*	.10*	.19**	.14*	.05	.02	.15	.08
CTMM Inference	.27**	.25**			.24**	.22**	.13	.10	.08	.10	.25*	.15
CTMM N.Q. Coins	.31**	.25**			.29**	.22**	.13	.03	.11	.05	.19	.17
CTMM N.Q. Arithmetic	.38**	.16*			.34**	.21**	.22**	.07	.33**	.20**	.33**	.33**
CIVIL SERVICE COMM. TESTS			N=183	143-40								
CSC 51- Spatial Patterns			.37**	.27**								
CSC 24- Computations			.28**	.16*								
CSC 157- Abstract Reas.			.28**	.18*								
CSC 157- Letter Sequence			.55**	.45**								
CSC 135- Oral Directions			.23**	.23**								
CSC 540- ATP I+II			.41**	.29**								
COMPOSITE SCORES	N=211	193-17	N=183	143-40	N=423	380-51	N=213	201-13	N=198	172-26	N=99	87-12
Commercial 7-Test Comp.	.39**	.31**			.42**	.28**	.35**	.14*	.37**	.23**	.48**	.32**
Commercial 4-Test Comp.	.32**	.29**	.52**	.40**	.38**	.23**	.35**	.15*	.38**	.23**	.50**	.28**
New 4-Test Commerc. Comp.					.41**	.27**	.33**	.14*	.39**	.26**	.51**	.30**
CSC 6-Test Composite			.54**	.39**								
MILITARY (MSC) TESTS					N=422	379-51	N=184	173-12	N=197	171-26	N=74	62-12
GCT + ARI					.42**	.28**			.47**	.27**		
1/3 (AR+VE+PA)							.18*	.01			.27*	.18

* $p < .05$
** $p < .01$

since later analyses revealed that the majority of the Marines possessed exceptionally high MSC scores. The Navy selected its trainees from a broader range of ability levels. The correlations of .42 and .28 obtained in Sample 3 for the Navy MSC score versus Course Grade and Pass-Fail Status and the corresponding coefficients of .47 and .27 obtained for Sample 4 should also be considered as possible underestimates of the true validities of this screening variable. Yet, despite the presumed restriction-of-range effects, the obtained, uncorrected validities are substantial and exceed those obtained for any of the experimental measures.

Differential Validities of Aptitude Measures for FAA, Navy, and Marine Corps Groups. Validities of the various aptitude measures were computed independently for the subgroups of Navy and Marine trainees of Samples 3 and 4. These data, together with those previously derived and shown for Samples 1 and 2, are presented in Table III. It is apparent that the validities for each of the seven subtests tend to fluctuate from sample to sample. There is little evidence that any particular subtest constitutes a unique or highly superior predictor of training performance for either the Navy or the Marine trainees. Also, no composite measure appears to be consistently more valid for one group than the others. Validities of the three composites vary from sample to sample but they are remarkably similar for any given sample. With Course Grade employed as the criterion, they range from .38 to .42 for Sample 3a (Navy); .37 to .39 for 4a (Navy); .33 to .35 for 3b (Marines), and .48 to .51 for Marine Sample 4b.

Intercorrelations and Validities of Composites and MSC Score. The intercorrelations and validity coefficients of the three experimentally-derived composite aptitude measures and the MSC test scores of the Glynco trainees are presented in Table IV. The intercorrelations of the three composites range from .91 to .93 for the combined Navy groups and from .88 to .90 for

TABLE IV. Intercorrelations and Validities* of Composite Aptitude Measures for Navy and Marine Trainee Groups

Sample 3a+4a		4-test Comp.	New 4-test Comp.	MSC Test	Course Grade	Pass-Fail
Navy	Mean S.D.	N r	N r	N r	N r	N r_{pb}
Commercial 7-test Composite	2859.2 442.5	642 .92	642 .93	640 .71	621 .41	552-77 .26
Commercial 4-test Composite	1583.2 284.4		642 .91	640 .64	621 .38	552-77 .23
New 4-test Comm. Composite	296.9 44.7			640 .64	621 .41	552-77 .26
MSC Test GCT+ARI	119.4 10.1				619 .44	550-77 .28
Course Grade	79.3 8.3					551-70 .64
Sample 3b+4b						
Marine	Mean S.D.	N r	N r	N r	N r	N r_{pb}
Commercial 7-test Composite	3163.7 382.1	317 .90	317 .90	262 .51	312 .41	288-25 .23
Commercial 4-test Composite	1730.0 245.7		317 .88	262 .46	312 .41	288-25 .21
New 4-test Comm. Composite	325.5 38.0			262 .46	312 .41	288-25 .22
MSC Test 1/3(AR+VE+PA)	127.7 6.6				258 .21	235-24 .12
Course Grade	80.7 6.8					288-24 .52

*All intercorrelations and validities are significant at the .01 level or better except the Marine Corps MSC Test Score palidity coefficient of .12, which is significant at the .05 level.

the combined Marine groups. Validities of the three composites are, as previously discussed, virtually equivalent for the Navy and Marine groups. In other words, the predictive usefulness of either of the two four-test composites approximates that of the complete seven-test battery. However, comparison of the validities of the MSC test scores with those obtained for each of the experimentally-derived composites leads to some rather conflicting findings. For the combined Navy groups 3a and 4a, the validities of each composite are slightly exceeded by those of the Navy MSC score. For the combined Marine groups 3b and 4b however, the validities of the composites are, in every instance, appreciably higher than those of the Marine MSC score.

Inasmuch as the majority of the Marines were known to have been selected from relatively high MSC score ranges, this finding prompted a further statistical analysis to estimate the validity of each composite with the effect of the MSC score theoretically held constant. The resulting partial correlations are not shown but were quite moderate. For the combined Navy groups, the highest partial correlation was .11, for the "New 4-Test Commercial Composite." For the Marines, the partials ranged from .18 to .20. These results indicate that if the MSC qualification standards remain unchanged, the commercially-published tests involved in any of the three composites would produce only a moderate degree of improvement in the screening of Glynco candidates.

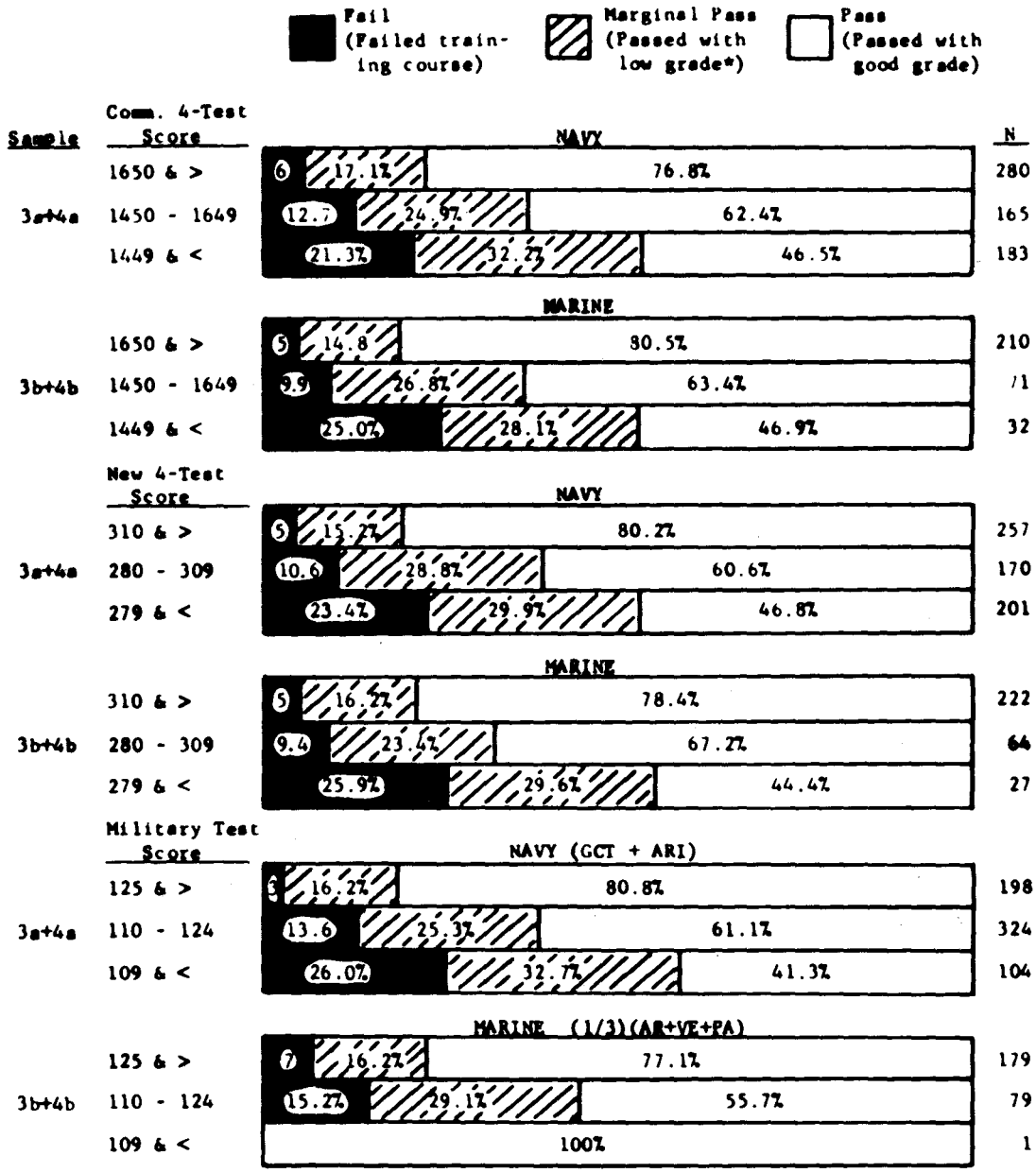

FIGURE 1. Percentages of Navy and Marine Corps ATC Trainees by Score-Range Groupings on the Commercial 4-Test Composite, New 4-Test Commercial Composite, and Military (MSC) Test who Failed, Marginally Passed, or Now-Marginally Passed the Glynco Training Course.

*Marginal Pass S's are those with training-course grades of 76 or lower; they comprise the approximate lower fifth of the distribution of grades.

Relative Screening Potential of MSC Scores Versus Two Four-Test Composites. The graphs in Figure 1 illustrate the relative effectiveness of the Commercial 4-Test Composite, the New 4-Test Commercial Composite, and the two MSC Scores as predictors of Pass-Fail Status and ATC Training Course Grades of the Navy and Marine trainees. For each graph, the aptitude measures have been arranged with reference to three score ranges. These groupings are not arbitrary. A score of 1650 on the Commercial 4-Test Composite is roughly equivalent to a minimum qualifying score of 210 on the CSC FAA-ATC Selection Test and a score of 1450 approximates a CSC-ATC test score of 190, which represented the FAA's qualifying standard during 1962 and 1963. Scores of 1449 and below thus correspond to CSC-ATC scores of 189 and less. Similarly, scores of 310 and 280 on the New 4-Test Commercial Composite are equivalent to 210 and 190 on the FAA test. The general, though not rigidly enforced, standard of an MSC score of 110 for selection as a Glynco trainee is reflected in the groupings established for the MSC score. However, grouping of MSC scores of 125 and higher to establish a third group was done on an arbitrary basis.

The first two graphs illustrate the potential of the Commercial 4-Test Composite to further improve the screening of candidates for the Glynco ATC training course. Data for the combined subsamples 3a and 4a reveal that 280 of the 628 Navy trainees attained scores of 1650 or higher. Only 6 per cent of these 280 high-scoring individuals failed the training course, 17.1 per cent passed with a "marginal" course grade (of 76 or lower), and the remaining 76.8 per cent passed with a grade of 77 or higher. The failure rate progresses to about 13 per cent for those having scores of 1450 to 1649 and to 21.3 per cent for those having relatively low scores of 1449 and less. The percentages representing marginal-pass cases increase in a similar manner, progressing to 24.9 for the intermediate score-range group and to 32.2 for subjects having scores of 1449 or less.

The second graph of Figure 1 refers to the combined Marine Subsamples 3b and 4b. A total of 210, over 67 per cent of the 313 Marines, attained scores of 1650 or higher on the Commercial 4-Test Composite. About 44.6 per cent of the Navy trainees were represented in this score range. Although the score distribution is quite different, the percentages representing the fail, marginal pass, and non-marginal pass cases within each of the three score-range groupings are remarkably similar to those obtained with the combined Navy subsamples. The failure rate progresses from about 5 per cent for the upper group to approximately 10 per cent for the intermediate group and up to 25 per cent for the group of subjects having scores of 1449 and less. Percentages representing the marginal-pass cases reflect a progression from 14.8 per cent in the upper score-range category to over 28 per cent in the lowest.

Even though the results are not graphically depicted, a similar analysis was undertaken for the 393 FAA-TATC trainees of combined Samples 1 and 2. Only 117 trainees were found to have Commercial 4-Test Composite Scores of 1650 or better, 111 had intermediate scores of 1450 to 1649, and 165 (almost 42 per cent) had scores of 1449 or less. In contrast, only 29 per cent of the 628 Navy students and only slightly more than 10 per cent of the Marines made scores of 1449 and less. (Differences in aptitude levels of the groups will be discussed later in this report.) At this point however, it is more important to consider the screening potential of the Commercial 4-Test Composite. Of the 117 FAA Ss having high scores of 1650 and above, less than 6 per cent failed the TATC training course, about 15 per cent passed with marginal grades (i.e., in approximately the lower fourth of the distribution of passing grades), and almost 78 per cent were in the upper three-fourths of the combined classes. Almost 10 per cent of the 111 subjects with intermediate scores of 1450 to 1649 failed and about 16 per cent were categorized as marginal. Of the 165 having scores of 1449 and lower, 23 per cent failed, 30 per cent were marginal, and only 47 per cent passed with good grades. In summary, two-thirds of the 57 failures and over half of the 86 marginal cases could have been screened from the two FAA samples by using 1449 as a minimum qualifying score. The results would yield a graph highly comparable to that shown in Figure 1 for the combined Navy Subsamples 3a and 4a.

The second pair of graphs in Figure 1 illustrate the degree to which the New 4-Test Commercial Composite Score can be used to predict performance in the Glynco training course.

Over 200 of the 628 Navy trainees made scores of 279 or less on this composite compared to only 27 Marines. In each instance however, about 25 per cent of the subjects having such low scores represent training-course failures and almost 30 per cent are classed as marginal-pass cases. The failure rate and the proportion of marginal-pass cases are contrastingly lower for both the Navy and Marine trainees who attained scores of 310 or better. The two graphs appear to be almost congruent with the previous pair which pertained to the original 4-test composite.

One of the two graphs in the lower portion of Figure 1 illustrates the screening potential of the military screening-and-classification (MSC) score used by the Navy and the other reflects the potential of the Marine Corps MSC score. The two MSC indexes are based on different instruments, as described earlier, but are presumably equivalent and comparable. The data indicate that the majority of trainees in both groups were selected from relatively high MSC score ranges. In this respect however, the Marines constitute the more elite group. Only one of the 259 Marines had an MSC score of less than 110, compared to 104 of the 626 Navy students. The overall failure rates, which are not presented, are 9.3 per cent for Marine, 12.3 for Navy, and 14.8 for FAA trainees of combined Samples 1 and 2. The graph for the combined Subsamples 3a and 4a indicates that almost 81 per cent of the 198 Navy Ss with MSC scores of 125 or higher passed the Glynco course with good grades. About 61 per cent of the Navy trainees who have intermediate-range MSC scores completed the course with non-marginal grades compared to 41.3 per cent of those having MSC scores of 109 or less. The graph for the combined Subsamples 3b and 4b shows that approximately 77 per cent of the 179 Marines with MSC scores of 125 and higher were able to pass the course with good grades. In contrast, less than 56 per cent of the 79 who have Marine MSC scores within the range of 110 to 124 are categorized as non-marginal pass cases. A study of the differences between the Navy and Marine data indicates that if the number of applicants should vastly exceed the training demands, the Navy would be able to improve substantially its screening procedures by requiring an MSC score of 110 or higher of every ATC applicant.

The three sets of graphs in Figure 1 are all rather similar. Each of the three aptitude measures has considerable predictive potential. However, if the data were extrapolated to determine the relative screening potential of each measure, the results would be as follows: (1) Based on 1450 as a minimum qualifying score, the Commercial 4-Test Composite could have been used to screen 39 of the Navy's 77 failures and 8 of the 24 Marine Corps failures. However, this screening would have been accomplished at the sacrifice of 144 potential-pass Navy cases and 24 potential-pass Marine cases. (2) With the New 4-Test Commercial Composite, a cut between 279 and 280 would theoretically exclude 47 of the Navy's 77 failures and 11 of the 24 Marine failures while simultaneously eliminating 154 Navy pass cases and 16 potential-pass Marine cases. (3) The establishment of a screening cut between 110 and 109 would have eliminated 27 fails and 77 pass cases of the Navy and a single Marine who passed the training course. Based on the specified screening cuts, the New 4-Test Commercial Test Composite may appear to be somewhat superior to the other measures. On the other hand, if the applicant pool were rather limited and training demands were high, it would represent an impractical screening device.

Differences in Aptitude Levels of FAA, Navy and Marine Corps Trainees. The bar diagrams of Figure 2 illustrate the relative range and distribution of the Commercial 4-Test Composite Scores for the FAA, Navy, and Marine Corps ATC trainees. This composite, which involved DAT Space Relations, DAT Numerical Ability, DAT Abstract Reasoning and CTMM Analogies, is the only measure available for a direct comparison of the different samples. The horizontal line across the middle of the bar diagram shown for each sample represents the median-score level. Other points depicted in the figure correspond to the second, tenth, twenty-fifth, seventy-fifth, ninetieth, and ninety-eighth percentiles.

The median score for the 214 Marines in subsample 3b is higher than for all other groups. This median score of approximately 1775 was attained or exceeded by less than 20 per cent of the FAA subjects of Samples 1 and 2. At every point on the scale, the Marines of both Samples 3 and 4 appear to be superior to the FAA trainees and also substantially better than the Navy trainees. Based on a comparison of the median scores, it becomes quite evident that the FAA

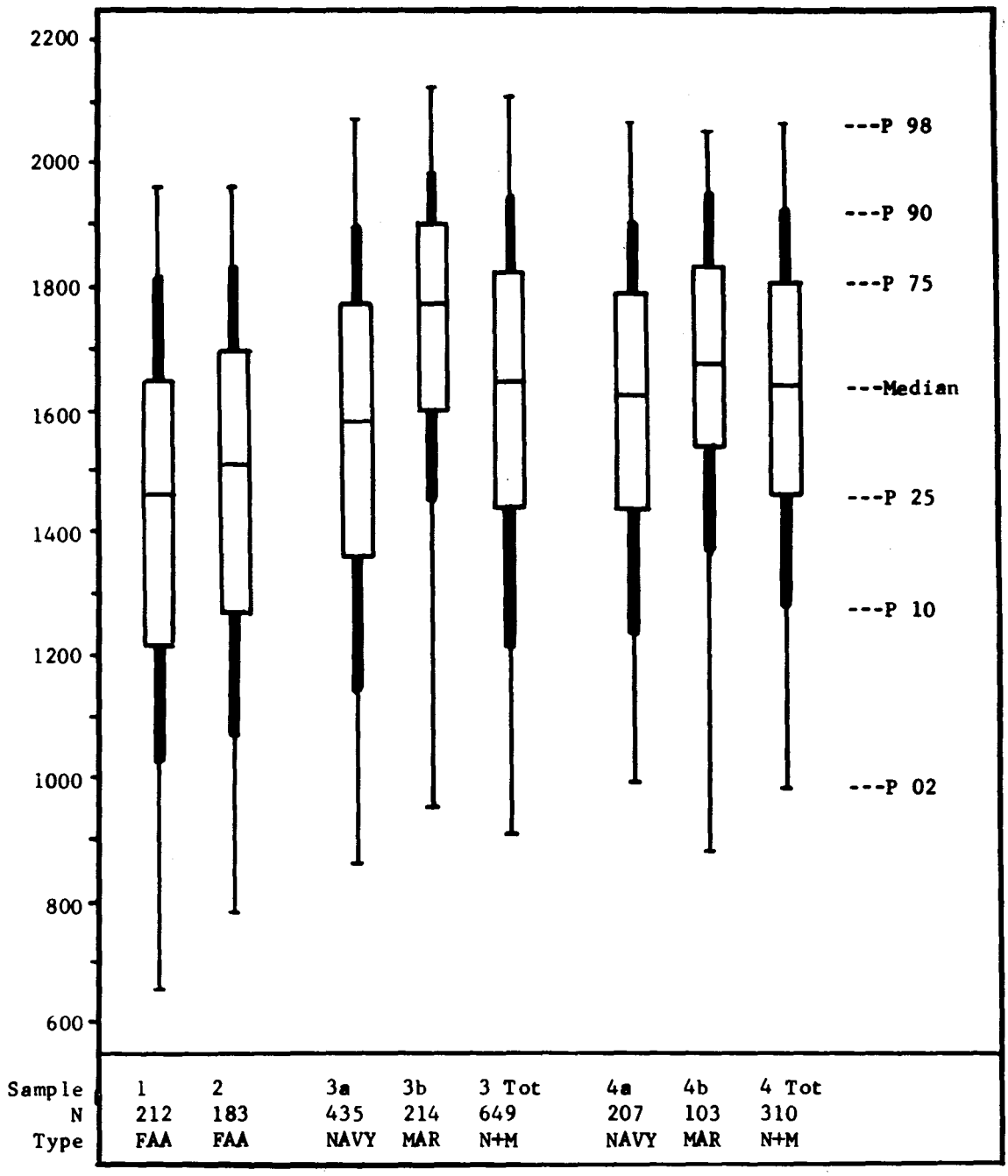

FIGURE 2. Relative Range and Distribution of Commercial 4-Test Composite Scores for FAA, Navy, and Marine Corps Trainees.

trainees of Samples 1 and 2 tend to have lower scores than the Navy students who comprise subsamples 3a and 4a and that the Marines generally have the highest aptitude scores. All differences between the means of the combined FAA samples, combined Navy samples and combined Marine samples were found to be statistically significant.

Comparison of FAA, Navy and Marine Ss in Terms of Estimated CSC Test Scores. The 183 FAA-ATC subjects designated as Sample 2 represent the only group in this study for whom scores were available on both the complete CSC-ATC-Aptitude-Screening-and-Selection Battery and any portion of the commercially-published

battery. Their experimental assessment at time of entry into training had included four of the seven commercial tests. As described earlier, these were the four tests used in derivation of the Commercial 4-Test Composite Score. The correlation between the Commercial 4-Test Composite scores and the CSC-ATC Test scores for Sample 2 was .81. This correlation thus permitted the development of an equation with which the former were used to estimate the CSC Test scores for subjects in Samples 1, 3, and 4.

Figure 3 illustrates the relative range and distribution of the estimated CSC-ATC Test scores for Samples 1, 3, and 4, and similar information regarding the existing CSC-ATC Test data of Sample 2. This figure is of the same

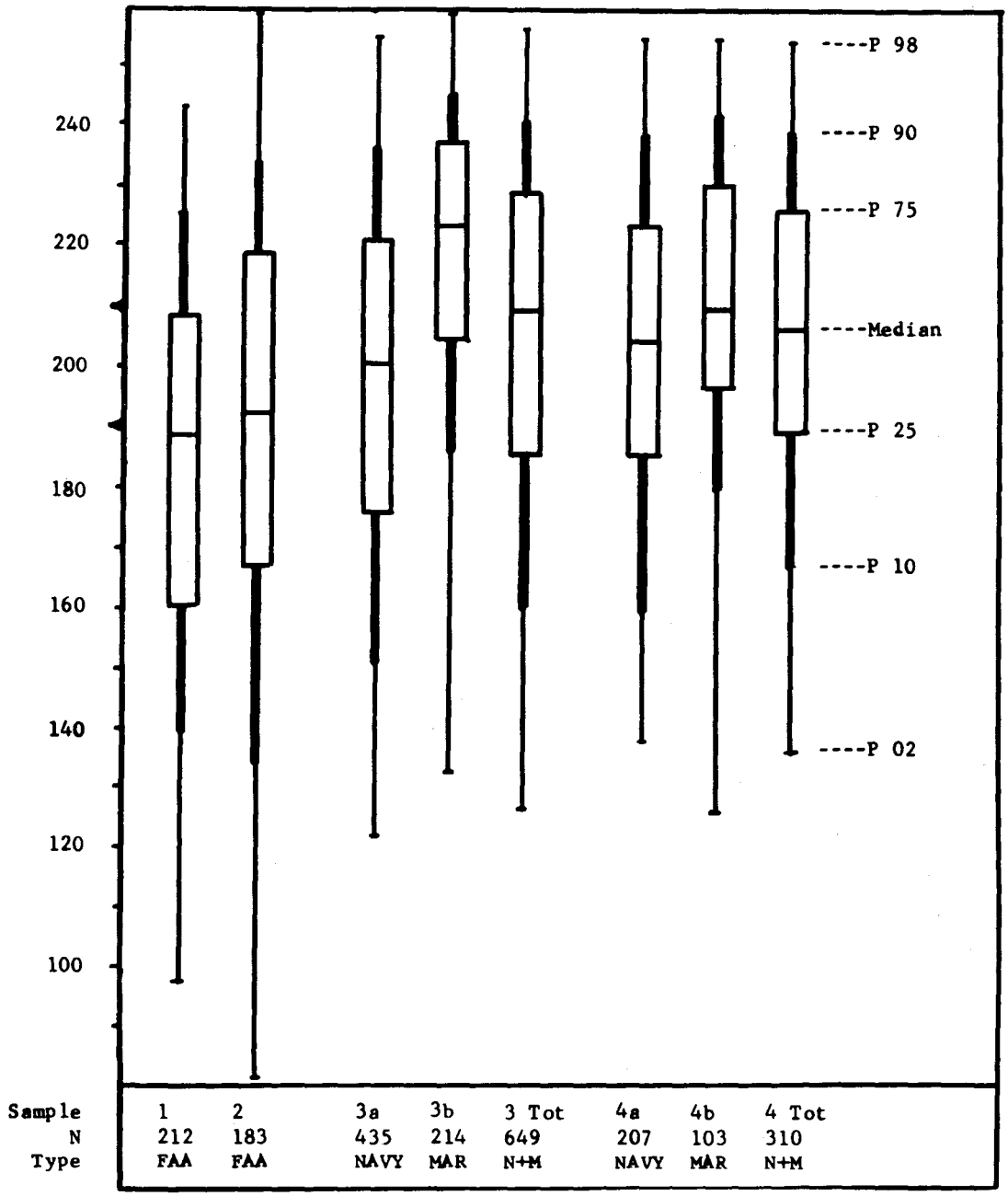

FIGURE 3. Relative Range and Distribution of CSC-ATC Aptitude Composite Scores for FAA, Navy, and Marine Corps Trainees. (Note: The CSC-ATC test scores were estimated through regression techniques for subjects of all samples except Sample 2.)

general type as Figure 2 and, inasmuch as the CSC scores were predicted from the composite scores depicted in Figure 2, the graphic patterns of the two are highly similar. However, Figure 3 permits an examination of the distributions relative to previous and current selection standards employed by the FAA. When the CSC Test first became operational in July 1962, a raw score of 190 represented the minimum qualifying score. However, the selection standards were revised in January 1964. Since that time, a CSC Test score of at least 210 has been required of every entrant into FAA ATC training. Figure 3 illustrates that a CSC-ATC Test score of 210 constitutes an exceptionally high score for the FAA subjects of Samples 1 and 2, whereas it is only slightly higher than the median of the estimated scores of the Naval trainees. In contrast,

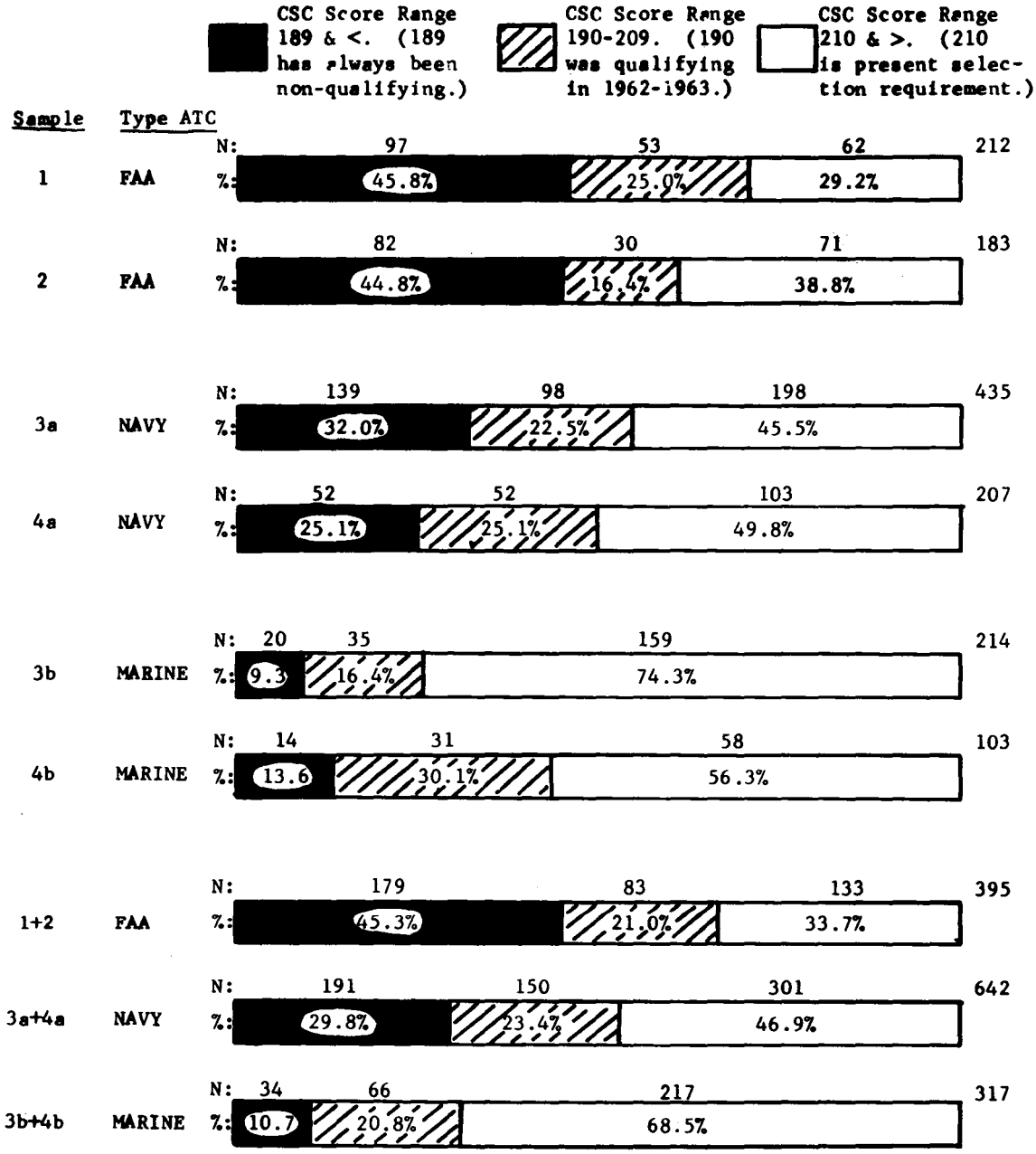

FIGURE 4. Percentages of ATC Trainees within each sample having actual or predicted CSC-Test Performance Scores Representing Low, Intermediate, or High ATC-Aptitude Ranges.

the majority of the Marines would be able to qualify. These findings are confirmed by the data presented in Figure 4. Over 45 per cent of the FAA subjects of combined Samples 1 and 2 have (actual or estimated) scores of 189 or less; 21 per cent have intermediate scores of 190 to 209, and only 33.7 per cent are in the qualifying range. For the combined Navy Subsamples 3a and 4a, the corresponding percentages are 29.8, 23.4, and 46.9. For the Marine Subsamples 3b and 4b, only 10.7 per cent are lower than 190; 20.8 per cent are in the intermediate range, and 68.5 per cent equal or exceed the requirement of 210.

CSC-ATC Test Score Versus MSC Score. Figure 5 reveals a high degree of correspondence between level of predicted CSC-ATC Test score and level of Military (MSC) Test score. (Though

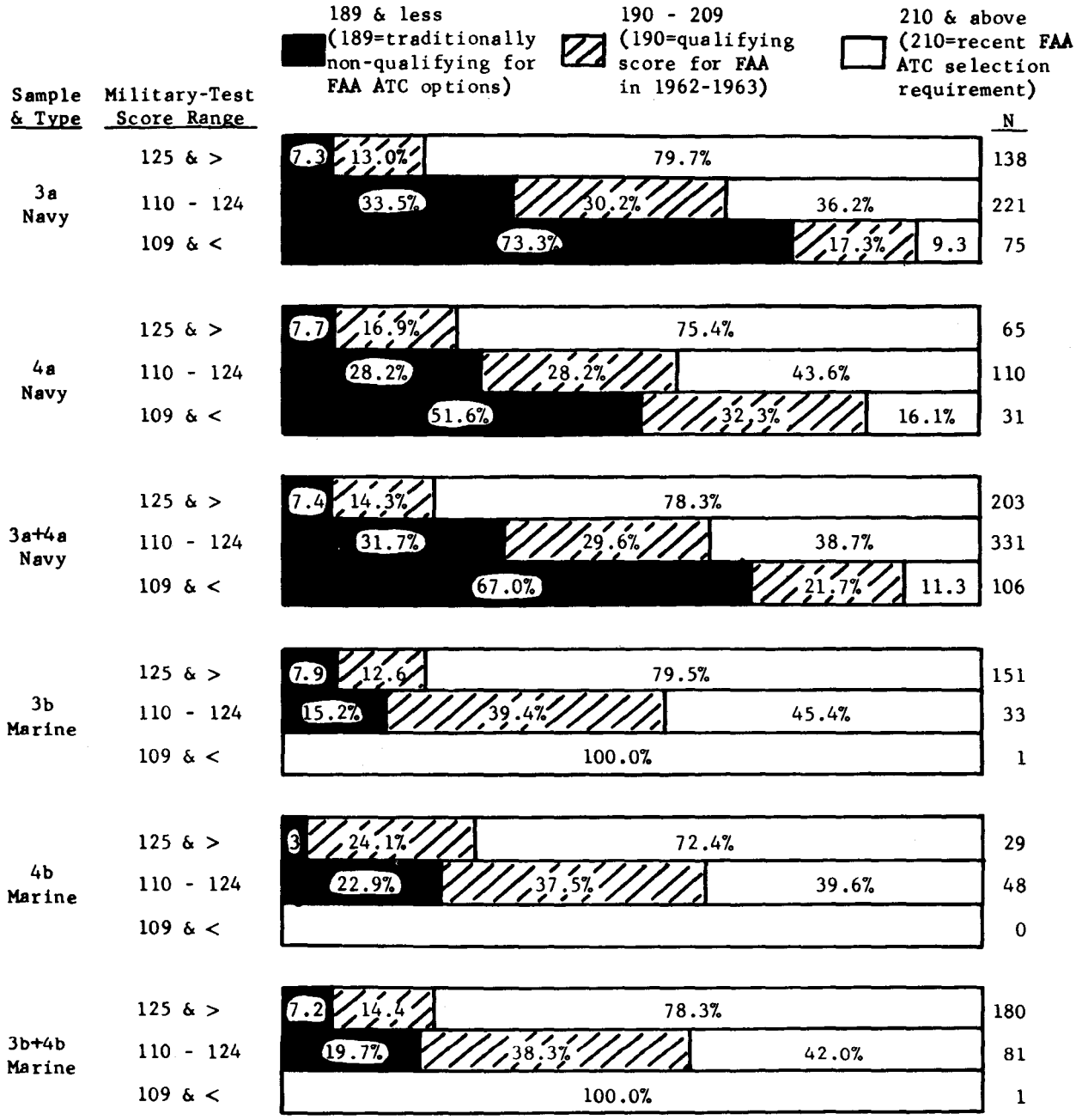

FIGURE 5. Percentages of Navy and Marine Corps ATC Trainees by Military-Screening-and-Classification Test Score Range Having Predicted CSC-ATC Test Scores of 189 or Less, 190 to 209, and 210 and above.

not shown, the correlation between the two was .64 for the Naval trainees and .46 for the Marines. The estimated Standard Error for the CSC score was 17.5; the Probable Error was 11.8.) The bar graphs pertaining to the separate and combined subsamples of both Navy and Marine Corps trainees indicate that the probabilities of attaining a qualifying score of 210 on the CSC Test range from about 72 to 80 per cent for those subjects having MSC scores of 125 or higher. Data for the various groups reflect a progressive decline in these probabilities from the high to the lower MSC-score-range categories. For example, 106 of the 640 Navy trainees of the combined subsamples 3a and 4a possessed MSC scores of 109 and less. Only 11.3 per cent of the estimated CSC scores for these 106 individuals are within the qualifying range of 210 and better, 21.7 per cent are within the nonqualifying range of 190 to 209, and 67 per cent are lower than the FAA's former selection standard of 190. With the exception of one case, all Marines included in this study had MSC scores of 110 or higher. One hundred eighty of these 261 Marines had an MSC score of 125 or better. Over 78 per cent of the CSC scores predicted for this elite subgroup are above 209. In contrast, nearly 60 per cent of the 81 Marines having MSC scores of 110 to 124 would probably not qualify on the CSC Test.

Actual or Estimated CSC Test Performance Versus Chronological Age. Previous research[4] based on a group of 306 FAA-TATC trainees, including the 183 subjects designated as Sample 2 in this study, revealed a negative correlation of .33 between CSC Test Score and Chronological Age. The mean CSC score for trainees over age 35 was significantly lower than for the younger groups. Similar results for the 395 FAA trainees of combined Samples 1 and 2 have emerged in the present study. The correlation between age and CSC score is −.25. Fifty-six of the 395 FAA Ss were 36 years of age or older when tested. The mean CSC score for this group is 170.4. The mean for those aged 35 and less is 194.1. The difference between the two groups is not as great as found in previous studies but it is statistically significant. (It should be noted that these results and those in Table V are based on analyses involving no distinction between predicted CSC Test scores and actual performance.)

Table V. Means of Actual or Predicted CSC-ATC Test Scores by Age Group for FAA, Navy, and Marine Corps ATC-Trainee Samples.

Chronological Age	Sample 1+2 FAA		Sample 3a+4a NAVY		Sample 3b+4b MARINE	
	N	M	N	M	N	M
45 & >	1	217.00				
43 - 44	9	149.50				
41 - 42	12	149.71				
39 - 40	16	184.34				
37 - 38	15	173.33				
35 - 36	6	210.75	1	164.50		
33 - 34	15	173.00				
31 - 32	21	195.69	3	187.83		
29 - 30	40	191.50	1	254.50		
27 - 28	36	191.51	7	187.36		
25 - 26	73	195.84	14	217.36	3	181.17
23 - 24	123	196.37	62	204.34	12	213.67
21 - 22	28	195.21	158	205.83	46	218.41
19 - 20			331	202.26	198	222.23
17 - 18			65	204.65	44	213.82
Total	395	190.75	642	203.71	303	219.68

Table V presents the means of the (actual and/or predicted) CSC Test scores, by age group, for the FAA, Navy, and Marine Corps trainees. In comparing means of CSC scores for subjects in successive two-year age intervals, considerable variability is evident within each array. Yet, the data for the FAA trainees follow a trend which reflects the previously-established differences between performance levels of the older and younger trainees. There is also some degree of correspondence between age and level of predicted CSC Test score for the Navy subjects of combined Subsamples 3a and 4a. Most of these trainees are relatively young; only 12 are over age 26. One of the 12, in the 35-to-36 year bracket, has an estimated score of 164.5. Ten of the 12, in various age brackets, were found to have an average score of approximately 187. The overall average for these 12 older individuals is 191.1, compared to about 204 for the 630 younger Navy trainees. The difference, however, is not statistically significant. All of the Marines are younger than 27 and no association between level of estimated CSC Test performance and age level was found for this group.

IV. Discussion.

Although most of the Navy and all of the Marine trainees included in this study were selected from relatively high MSC score ranges, the results indicate that further improvement in the screening and selection of applicants for Glynco ATC training could be attained by requiring a score of 110, or possibly 115, on the MSC Test and through secondary aptitude testing and screening procedures. A score of 110 on the MSC Test as a selection requirement would result in the exclusion of many potential fail and marginal-pass cases. However, if an MSC score of 110 constituted the qualification standard, the three DAT tests of Space Relations, Numerical Ability and Abstract Reasoning and the CTMM Inference test (or tests of a highly similar type) could be used to effectively screen additional numbers of individuals who would either fail or only marginally pass the training course. On the other hand, if an MSC selection score of 115 were chosen, supplementary aptitude testing and seconding screening procedures would be of only limited value. However, the feasibility of implementing any new method depends upon many factors. For example, the ratio of the number of applicants to the number of trainee positions to be filled should be a prime determinant of the specific qualification levels. A determination would also be in order regarding the extent to which priority selection of ATC personnel from high aptitude levels might conflict with demands for high-aptitude personnel for other crucial specialty areas.

The finding of insignificant relationships between chronological age and the experimentally-derived aptitude measures of the Navy and Marine Corps trainees is not surprising in view of the fact that only 12 of the Navy subjects (and none of the Marines) were over 26 years old. The mean performance level for the 12 older Navy trainees was admittedly lower than for the younger group of students but not significantly so. However, the age at time of testing ranged from 21 to 46 for the former FAA-TATC trainees. The older FAA subjects tended to have appreciably lower aptitude scores. Differences between the means of aptitude measures for dichotomized groups aged "35 and younger" versus "36 and older" was found to be statistically significant. This finding is in general agreement with results obtained in previous research with larger groups of TATC trainees and Center (or air-route-traffic-control) trainees. Observations during testing sessions at CAMI indicated that older examinees generally experience more difficulty with tests involving perceptual speed, rapid assimilation and processing of information, and the retention of instructions pertinent to the more novel tasks. Older individuals appear to be less "flexible" and more "persevering" than their younger colleagues in seeking solutions to the more difficult problems of many tests. Some authorities may theorize that recency of education and exposure to testing situations is partially responsible while others will contend that the older individuals are apt to have lower levels of motivation. Yet, it remains largely a matter of speculation as to why the older FAA trainees tend to have lower composite aptitude scores than the younger trainees.

The FAA-ATC trainees included in this study were recruited during 1961 and 1962 when an assessment of types and amounts of previous job-related experience (particularly military ATC work) constituted the primary basis for selection. Qualification involved no formal assessment of aptitudes. However, this was also a period during which research was being conducted to establish the degree to which the selection process might be improved through the use of aptitude tests. The FAA-ATC trainees included in this study were participants in that research and were administered an extensive and heterogeneous battery of aptitude tests on an experimental basis at the time they entered training. The results obtained in this study indicate that if the CSC-ATC test had been in operational use, almost 65 per cent of these trainees would never have been selected. Inasmuch as the majority of these men were selected on the basis of their military ATC experience, there must have been inadequacies in the procedures by which they were selected and assigned to military ATC training. Unfortunately, the existing research records are incomplete regarding the respective military organization in which each FAA trainee attained his entry-qualifying experience and no comparative study in this respect has yet been made.

The present study permitted a comparison of the average performance levels of FAA, Navy, and Marine Corps trainees on several groups of aptitude tests. In all instances, the FAA samples

had significantly lower mean scores than the Navy sample and the Marine sample was superior to both the FAA and Naval groups. With one exception, all Marines possessed MSC scores of 110 and higher. The average Marine MSC score was 127.7. Even though more than 100 of the Navy trainees had scores of 109 or less, their average was 119.4. Both the Marine MSC score and the Navy MSC index correlated significantly with each of the experimentally-derived composite measures. The latter were very highly intercorrelated. Finally, the CSC-ATC Test scores, which were available for one of the two samples of FAA trainees, correlated .81 with the Commercial 4-Test Composite scores. These relationships tend partially to explain the consistency with which the Marines were found to be the superior group. In brief, the Marines in this study were selected from exceptionally high aptitude groups, the Navy trainees were selected from mid-range levels, and the FAA subjects, who were required to meet no aptitude standards, tended to represent much lower levels.

It is of interest to the FAA that both the Navy and Marine Corps trainees of the Glynco ATC School represent rather select groups insofar as aptitudes are concerned. At some future date, many of these men will apply for FAA employment. During the interim they will have gained valuable military air traffic control experience. When they do apply, a far greater percentage of them should be able to qualify on the FAA-CSC-ATC Screening Test than the 25 to 35 per cent of the applicant population of recent years.

REFERENCES

1. Cobb, Bart B.: Problems in Air Traffic Management: V. Identification and Potential of Aptitude Test Measures for Selection of Tower Air Traffic Controller Trainees. *J. Aerospace Medicine*, 35, 1019–1027, 1964. Also published as: Civil Aeromedical Institute Report AM 65–19, FAA, Oklahoma City, Oklahoma, 1965.
2. Cobb, Bart B.: Problems in Air Traffic Management: II. Prediction of Success in Air Traffic Controller School. *J. Aerospace Medicine*, 33, 702–713, 1962. Also published as: Civil Aeromedical Institute Report 62–2, FAA, Oklahoma City, Oklahoma, 1962.
3. Trites, D. K.: Problems in Air Traffic Management: I. Longitudinal Prediction of Effectiveness of Air Traffic Controllers. *J. Aerospace Medicine*, 32, 1112–1120, 1964. Also published as: Civil Aeromedical Institute Report 61–1, FAA, Oklahoma City, Oklahoma, 1961.
4. Trites, D. K., and Cobb, Bart B.: Age, Aptitude, and Experience as Predictors of Performance of Air Traffic Controllers. Unnumbered Report, Civil Aeromedical Institute, FAA, Oklahoma City, Oklahoma, 1964.