603 416

PROBLEMS IN AIR TRAFFIC MANAGEMENT
IV. COMPARISON OF PRE-EMPLOYMENT, JOB-RELATED EXPERIENCE WITH APTITUDE TESTS AS PREDICTORS OF TRAINING AND JOB PERFORMANCE OF AIR TRAFFIC CONTROL SPECIALISTS

DAVID K. TRITES, Ph.D.
Chief, Selection Section

BART B. COBB, M.S.
Selection Section
Psychology Branch

12 p $1.00 hc
$0.50 mf

63-31

PROBLEMS IN AIR TRAFFIC MANAGEMENT
IV. COMPARISONS OF PRE-EMPLOYMENT, JOB-RELATED EXPERIENCE WITH APTITUDE TESTS AS PREDICTORS OF TRAINING AND JOB PERFORMANCE OF AIR TRAFFIC CONTROL SPECIALISTS*

David K. Trites, Ph.D.

Bart B. Cobb, M.S.

In addition to passing a second class aviation medical examination, individuals applying for training as an Air Traffic Control Specialist (ATCS) with the Federal Aviation Agency have, in the past, been required to have previous experience which was considered relevant to air traffic control work. This would seem to be a reasonable selection procedure, but other research[1][2][3] has demonstrated that aptitude tests effectively predict training and job performance of individuals selected in this manner. Thus, assuming the value of the medical requirements, the predictive worth of job-related experience can be questioned.

To evaluate the effectiveness of experience as a predictor, the experience backgrounds of several hundred ATCS trainees were examined. All of the trainees had been selected on the basis of the medical and experience requirements only, for training in either Enroute (Center) or Terminal (Tower) air traffic control procedures. These two training courses have been described in more detail in an earlier report.[2] In addition, approximately 94 percent of the Enroute trainees and 88 percent of the Terminal trainees had participated in an experimental program of aptitude testing. The present report summarizes the findings with respect to the prediction of training and job performance by various types of pre-employment, job-related experience, selected demographic variables and an aptitude test battery.

PROCEDURE

Samples — Trainees entering the ATCS Enroute or ATCS Terminal training courses at the FAA Academy, Oklahoma City, in September, 1960, through July, 1961, are to be referred to as the Enroute Sample (E) and the Terminal Sample (T), respectively. Of 503 Enroute Sample trainees, 471 participated in the experimental aptitude testing program; among the 503, 129 (25.6%) failed to complete the training course, and among the 471, 118 (25.1%) failed to complete the course. Of 242 Terminal Sample trainees, 212 participated in the testing program; among the 242, 19 (7.9%) failed to complete the course, and among the 212, 17 (8.0%) failed to complete the course. Minor variations in the number of cases entering into various calculations and distributions are due to missing data or to a few individuals who withdrew from training for reasons other than failure. These latter individuals, whose withdrawal was necessitated by reasons beyond their control, such as illness, death in the family, etc., were considered as neither successes nor failures in training and in most instances had no criterion data available.

For both samples, Table 1 contains means and standard deviations of age and four of the criterion variables to be described subsequently. The data revealed no apparent differences, or biases, between the total sample and that portion which was tested.

TABLE 1

Comparison of Certain Characteristics of the Total and the
Aptitude-Tested Parts of the Enroute and Terminal Samples

Variables	Enroute						Terminal					
	Tested			Total			Tested			Total		
	X	S.D.	N.	X	S.D.	N.	X	S.D.	N.	X	S.D.	N.
Age	28.5	6.3	470	28.5	6.4	501	27.8	5.9	212	27.9	5.9	241
Academic Gd.	83.6	7.4	470	83.6	7.3	502	87.0	5.3	211	87.1	5.3	241
Lab Gd.	78.1	13.0	467	78.0	13.0	499	83.3	5.2	207	83.3	5.4	237
A + L Gd.	80.8	9.3	468	80.8	9.2	499	85.0	5.6	211	85.0	5.6	241
Supervis. Rating	2.45	.54	301	2.45	.53	318	2.42	.67	180	2.43	.66	204

Experience and Demographic Variables — Source of these data was the "ATCS Registration Sheet" which each student completed immediately after arrival at the FAA Aeronautical Center. For those who were tested, the form was completed on the Saturday immediately preceding the start of their training course on the following Monday. For those not tested, the form was completed on the day the course began. Table 2 contains descriptions of the 12 specific types of experience recorded on the form, 3 variables representing a summation of the individual types of experience, the demographic variables of age and education, and the aptitude test variables. In coding the variables no attempt was made to achieve normalized distributions of the data.

Aptitude Test Variables — The aptitude test variables represented only a small part of a much larger test battery described by Cobb.[2] The five tests described in Table 2 are those which Cobb found most valid as a composite for prediction of Enroute training course criteria. The composite score was obtained by multiplying the scores on each of the five tests by weights which were proportional to the inverse of the standard deviation of each test and summing the weighted test scores. Selection of the five tests and determination of the weights given in Table 2, were based on Cobb's analysis of data for 124 ATCS trainees who form part of the present Enroute Sample. Thus, relationships between the composite score and training level criterion measures were somewhat inflated in that sample. However, since the composite score was subsequently validated and since the sample of 124 subjects comprises approximately one fourth of the present sample, the spurious inflation of relationships should not be too serious.

Training-Level Criterion Variables — Four criterion measures were developed from information entered by FAA Academy personnel on each trainee's final "Evaluation of Performance" record. These were: *Academic Grade Average* (Acad), *Laboratory Grade Average* (Lab), *Combined Academic plus Laboratory Grade Average* (A+L), and *Pass or Fail* (P-F) in the training course.

The Academic Grade Average was obtained by averaging the grades obtained on academic examinations of seven areas administered at an intermediate stage and at the conclusion of the training course. In the latter instance, the seven individual subtests constituted the Air Traffic Control Specialist Certification Examination (ATCSCE). Normally, the Academic Grade Average was based upon the 7 intermediate grades plus the 7 final grades; but a failing grade of 70 or less at the intermediate stage of training for any of the 7 academic areas usually entailed a "retake" of the examination for the specific areas involved. To be graduated from training no more than two failing grades could be obtained by a trainee on the first administration of the ATCS Certification Examination; and he had to successfully pass the "retake" examination for these areas. A trainee was automatically eliminated if he had three or more failing grades on the ATCSCE. Thus, theoretically an Academic Grade Average could be based on considerably more than 14 grades. In practice, the number infrequently exceeded 15 or 16.

The Laboratory Grade Average for Enroute Sample trainees was based on three final laboratory grades reflecting performance in Strip Writing, Control Procedures as an Assistant Controller, and Control Procedures as a Journeyman Controller. These grades reflected per-

TABLE 2

Description of Variables

Variable Name & Abbreviation*	Description and Coding
Experience Variables	**Coding for Variables 1 & 2**

	Amount of Experience	Codes
1. Pilot Experience (Pil)	No experience reported	1
2. Radio Operator: Air to Ground Communications (Radio:Air/Grnd)	Less than 1 year	2
	12 months through 23 months	3
	2 years through 4 years	4
	5 years through 6 years	5
	7 years through 8 years	6
	9 years through 10 years	7
	11 years through 15 years	8
	16 years or more	9

3. Ground Control
Intercept (GCI) — Long Range, high altitude interception technique used by military aircraft control and warning units

4. Station (Stat) — A unit primarily engaged in ground to air communications and pilot briefings

5. Ground to Air Communications (Grnd/Air)

6. Point to Point Communications (P to P) — Communcations from one fixed ground point to another fixed ground point

7. VFR Tower (VFR Tow) — A tower controlling air traffic under Visual Flight Rules (VFR)

8. Approach Control: Tower (App Con Tow) — A tower capable of controlling air traffic under Instrument Flight Rules (IFR) but without access to radar

9. Radar Approach Control: Tower (Rad App Con Tow) — A tower with access to radar as an aid in controlling air traffic

10. Center (Cent) — An air route traffic control center

11. Ground Controlled Approach (GCA) — A ground radar system used to assist aircraft during landing

12. Radar Approach Control Center (RAPCON) — USAF radar system used for approach control at Air Force airfields; similar system in use by the Navy is called a Radar Air Traffic Control Center (RATCC)

Coding for Variables 3 through 12	
Amount of Experience	Codes
No experience reported	1
Through 3 months	2
4 through 6 months	3
7 months through 1 year	4
13 months through 2 years	5
25 months through 3 years	6
37 months through 5 years	7
6 years through 10 years	8
11 years or more	9

13. Sum of Communications Experience (Σ Comm) — Sum of individual experience variables Nos. 4, 5, & 6

14. Sum of Air Traffic Experience (ΣAT) — Sum of individual experience variables Nos. 7 through 12

15. Sum of Relevant Experience (Σ Rel) — Sum of experience variables Nos. 3, 13, and 14

Demographic Variables

16. Age — Chronological age to nearest birthday on date of entry into ATCS training

17. Education (Educ)

Coding for Education	
Amount of Education	Codes
None reported	Blank
Less than High School Graduate	1
High School Graduate	2
Less than 1 year of college	3
1 year of college	4
2 years of college	5
3 years of college	6
4 years of college	7
5 years of college	8
6 or more years of college	9

TABLE 2 (continued)

Aptitude Test Variables

#18. DAT Space Relations (Sp) (Weight = 6)	Identify solid figures that can be made from an unfolded pattern (40 items; scored: Rights minus Wrongs)
#19. DAT Numerical Ability (Num) (Weight = 13)	Test of arithmetic or computational skill (40 items; scored: Rights minus ¼ Wrongs)
#20. DAT Abstract Reasoning (Abs) (Weight = 14)	Indicate which of a series of choices (figures) properly carries out a principle of logical development exhibited by a sequence of figures (50 items; scored: Rights minus ¼ Wrongs)
##21. CTMM Analogies (Analog) (Weight = 49)	Seven drawings of different objects are presented for each item. The first object has a definite relationship to the second which must be recognized in order to identify, by analogy, the drawing among the last four which is similarly related to the third drawing (15 items; scored: Rights)
**22. Air Traffic Problems (ATP) (Weight = 18)	Determine whether aircraft may be permitted to change altitude without violating a specified time-separation rule (30 items; scored: Rights minus Wrongs)
23. Composite Aptitude Test Score (Comp)	Sum of aptitude test variables Nos. 18 through 22, each weighted as indicated.

* Abbreviations are given in parenthesis following the names of the variables.

\# Part of the Differential Aptitude Test (DAT) Battery, Form A, 1947, published by the *Psychological Corporation*, New York, N. Y.

\#\# Part of the California Test of Mental Maturity (CTMM), Advanced, Form A, 1957 published by the *California Test Bureau*, Los Angeles, California.

** Originally developed by the *American Institute for Research*, Pittsburgh, Pa., under contract with the Civil Aeronautics Administration in 1950. Form used in the present research was an extensive revision of the original test. Revision prepared by the Selection Section, Psychology Branch, Civil Aeromedical Research Institute.

formance in simulated air traffic control work. An unsatisfactory grade of 70 or less for any one of the three areas normally resulted in elimination of the trainee from the course. The Laboratory Grade Average for the Terminal Sample trainees was based upon a single grade reflecting performance in terminal approach control operations. If the grade were less than 70, the trainee was eliminated from the course.

The Combined Academic plus Laboratory Grade Average was the simple arithmetic mean of the Academic and Laboratory Grade Averages of each trainee. Data from incomplete records of trainees who were eliminated or withdrawn from training were treated similarly; but the averages were based on fewer grades.

All trainees successfully completing the training course were considered as Pass cases; those unsuccessful, as Fail cases. Eight trainees in the Enroute Sample and two in the Terminal Sample were not considered as either Pass or Fail and did not enter into any calculations involving the Pass-Fail criterion.

Job-Performance Criterion Variable — Approximately 10 months after trainees had graduated from the FAA Academy, a letter describing the research project and data collection forms were sent to the Chiefs of the facilities to which the trainees had been assign-

ed. For each member of the sample at his facility, the Chief was asked to have four supervisors rate each subject using a 15 item checklist performance evaluation form. The form contained items related to work habits, ability, judgment, and reasoning, emotional stability, and relationships with others. The possible ratings were: Excellent, Very Good, Good, Fair, and Unsatisfactory. The *Average Supervisory Rating* (Super) was computed by assigning weights of 4 through 0 to ratings of Excellent through Unsatisfactory, respectively, summing all of the items rated by all supervisors, and dividing the sum by the total number of items rated. Unfortunately, it was not possible to obtain four supervisory ratings on all subjects, but the average number of forms completed per subject was 3.8. In an earlier study a corrected split-half reliability of .75 was obtained for Average Supervisory Ratings computed from a rating form containing 12 of the 15 items used in the present form.[3]

Statistical Methodology — To indicate the degree of relationship between each criterion variable and the experience, demographic, and aptitude test variables, product-moment or point-biserial correlation coefficients were computed for the Enroute and Terminal

Samples, separately. In every instance, correlations were computed using the maximum possible number of cases.

Since many of the variables had an appreciable relationship with age, the partial correlation technique was employed to eliminate, or nullify statistically, the influence of age. Both the original, or first order, and the partial correlations are reported; but in the following discussion reference usually is made only to the partial correlations.

Although the means and standard deviations of some of the variables revealed an extreme skewness in their distributions, conventional tests of statistical significance were applied to the first order correlations. In testing significance of the partial correlations the smallest N involved in the computation of a particular partial was used. This resulted in a conservative test of significance.

As a partial check on the appropriateness of the conventional test of the first order correlations, joint frequency distributions were compiled for the 12 individual experience variables vs. the Academic plus Laboratory Grade Average, and the 12 experience variables vs. the Pass-Fail criterion. To test the significance of the bivariate distributions all the experience variables were dichotomized as: no experience vs. some experience. Two by two contingency

tables were then established for the experience dichotomy vs. above and below the approximate median of the Academic plus Laboratory Grade Average, and the experience dichotomy vs. Pass or Fail in training. The significance of each contingency table was tested by chi square.

RESULTS

Means and standard deviations of all variables listed in Table 2, except age (see Table 1), are given in Table 3. It is readily apparent that many trainees had little, or none, of the individual types of experience. This difficulty was overcome to a large extent by pooling individual types of experience to form the three summary variables: Sum of Air Traffic Experience, Sum of Communications Experience, and Sum of Relevant Experience.

Experience Variables vs. Criteria — First order and partial correlations of the five criterion variables with the experience variables and age are presented in Tables 4 and 5. The partial correlations indicated that many of the experience variables had either no relationship to training and job performance or had a negative relationship; i.e., the greater the amount of some specific types of experience the less well the trainee did. This was particularly true of GCI, and communications type experience.

TABLE 3

Means and Standard Deviations of the Experience Variables, Test Variables, and Education Based on the Maximum Possible Number of Trainees in the Enroute and Terminal Samples Separately

Variables	Mean		Std. Dev.		N	
	E	T	E	T	E	T
1. Pilot	2.38	1.92	2.64	2.16	503	242
2. Radio: Air/Grnd	1.38	1.28	1.23	1.13	503	242
3. GCI	1.39	1.23	1.42	1.07	503	242
4. Station	1.39	1.10	1.38	.68	503	242
5. Grnd/Air	2.12	1.72	2.22	1.88	503	242
6. P to P	1.39	1.23	1.34	1.08	503	242
7. VFR Tower	2.68	4.19	2.30	2.51	503	242
8. App Con Tow	1.80	2.53	1.68	2.16	503	242
9. Rad App Con Tow	1.14	1.14	.77	.72	503	242
10. Center	1.44	1.20	1.28	.78	503	242
11. GCA	2.21	2.26	2.14	2.25	503	242
12. RAPCON	1.75	1.46	1.68	1.35	503	242
13. Σ Comm	4.89	4.07	3.51	2.57	503	242
14. ΣAT	11.01	12.81	4.83	4.21	503	242
15. Σ Rel	17.28	18.10	5.06	4.46	501	242
17. Educ	2.90	2.79	1.59	1.49	489	241
18. Space	60.79	58.98	18.93	19.62	472	212
19. Numerical	22.17	21.68	8.13	7.52	474	212
20. Abstract	33.92	34.48	7.14	7.16	474	212
21 Analogies	6.80	6.71	1.99	2.03	475	212
22. ATP	13.46	13.34	5.84	6.06	475	212
23. Composite	1703.	1688.	363.	375.	471	212

TABLE 4

Enroute Sample: First-Order Correlations and Partial Correlations (Age Held Constant Statistically) between Experience Variables and Criterion Measures

Variables	A + L		P − F[1]		Acad		Lab		Super		Age
	r	Partial r	r	Partial r	r	Partial r	r	Partial r	r	Partial r	r
1. Pilot	—12**	07	—17**	02	01	14**	—18**	02	—17**	—07	64**
2. Radio: Air/Grnd	—23**	—19**	—15**	—11*	—20**	—18**	—20**	—16**	03	07	18**
3. GCI	—14**	—13**	—12**	—11*	—16**	—16**	—10*	—09*	—01	00	04
4. Station	—11*	—11*	—11*	—11*	—17**	—17**	—06	—06	—02	—01	02
5. Grnd/Air	—15**	—15**	—13**	—13**	—23**	—22**	—09*	—09*	—05	—05	03
6. P to P	—12**	—12**	—15**	—15**	—21**	—21**	—05	—04	—15**	—15**	03
7. VFR Tower	14**	09*	16**	11*	21**	18**	09*	03	11*	08	—21**
8. App Con Tow	17**	16**	16**	14**	19**	18**	13**	11*	03	01	—09
9. Rad App Con Tow	06	06	06	07	08	08	04	04	02	02	01
10. Center	14**	13**	08	07	08	07	15**	14**	09	08	—05
11. GCA	13**	10*	18**	16**	14**	13**	11*	09	09	08	—10*
12. RAPCON	16**	12**	15**	11*	14**	11*	15**	10*	05	02	—17**
13. Σ Comm	—18**	—18*	—18**	—18*	—29**	—29**	—10*	—10*	—10	—09	03
14. Σ AT	29**	24**	30**	25**	31**	29**	23**	18**	16**	12**	—24**
15. Σ Rel †	11	06	13**	08	05	02	12**	07	09	06	—19**
16. Age †	—27**		—28**		—15**		—30**		—18**		
Total N	499		495		502		499		318		501

[1]Point-biserial correlations; all other correlations are product moment. Decimal points omitted from all correlations.

*Significant at less than the .05 level.

**Significant at less than the .01 level.

†For Σ Rel and Age the Ns used in the correlations with Acad and Lab were 500 and 498, respectively. For the other correlations in these columns, the Ns were as indicated in the Total N row.

TABLE 5

Terminal Sample: First-Order Correlations and Partial Correlations (Age Held Constant Statistically) between Experience Variables and Criterion Measures

Variables	A + L		P − F[1]		Acad		Lab		Super		Age
	r	Partial r	r	Partial r	r	Partial r	r	Partial r	r	Partial r	r
1. Pilot	—12	—02	—13*	—01	04	10	—32**	—18**	—17*	—14*	51**
2. Radio: Air/Grnd	—07	—04	03	07	—07	—06	—09	—05	—17*	—16*	15*
3. GCI	—06	—01	—04	01	—03	—01	—11	—04	01	03	21**
4. Station	—10	—10	02	03	—15*	—15*	—07	—07	01	02	03
5. Grnd/Air	—07	—08	03	02	—13*	—14*	—06	—08	—01	—02	—04
6. P to P	—08	—10	—02	—04	—09	—10	—10	—12	06	05	—04
7. VFR Tow	08	03	13*	09	—01	—03	26**	21**	02	00	—20**
8. App Con Tow	20**	20**	06	07	10	10	25**	27**	02	02	00
9. Rad App Con Tow	06	09	06	09	04	05	06	11	03	04	12
10. Center	05	04	08	05	07	06	01	—02	01	00	—09
11. GCA	04	06	01	03	01	02	03	06	—01	—01	09
12. RAPCON	11	11	03	03	09	08	11	11	06	05	—03
13. Σ Comm	—11	—12	02	01	—17**	—17**	—10	—11	02	02	—03
14. Σ AT	22**	22**	15*	14*	10	10	35**	35**	05	04	—07
15. Σ Rel	13*	13*	14*	14*	—01	—01	25**	26**	06	06	—03
16. Age †	—21**		—25**		—09		—34**		—09		
Total N	241		240		241		237		204		241

[1] Point-biserial correlations; all other correlations are product moment. Decimal points omitted from all correlations.

* Significant at less than the .05 level.

** Significant at less than the .01 level.

† Correlations between Age and the first four criteria were computed with an N which was one less than the Ns given in the Total N row.

On the other hand, those types of experience related more directly to air traffic control tended to have positive relationships with the criteria; i.e., the greater the amount of such air-traffic-control related experience the better the trainee tended to do. The data for the Sum of Air Traffic Experience (Variable 14) and the Sum of Communications Experience (Variable 13) illustrate these contrasting relationships. Each statistically significant correlation of Σ AT with one of the criteria had a positive sign, whereas each statistically significant correlation of Σ Comm with one of the criteria had a negative sign. Combining the Sum of Air Traffic, the Sum of Communications, and GCI Experience to form the total Sum of Relevant Experience was, therefore, inappropriate. As revealed by the attenuated relationships with the criterion measures, this particular combination of experience failed to improve the prediction of training or job performance.

Of all the individual experience variables, Pilot Experience correlated most highly with age resulting in relatively large changes in the correlations when the effect of age was statistically controlled. Before age was statistically controlled, seven of the ten correlations between Pilot Experience and the Criterion variables in Tables 4 and 5 were significant at less than the .05 level and were negative in sign. With age controlled, the correlations with Laboratory Grade and Average Supervisory Rating in the Terminal Sample remained significantly negative; but the non-significant and positive correlation of .01 between Pilot Experience and Academic Grade in the Enroute Sample was increased to a statistically significant positive value. Considering the contradictory findings in the two samples and inasmuch as 7 of the partial correlations were not statistically significant, it seems that Pilot Experience can not be considered as an effective predictor of training or job performance measures.

A comparison of Tables 4 and 5 revealed many more significant correlations in the Enroute Sample. Out of 75 experience vs. criteria correlations in each sample, 45 were significant for the Enroute and 15 were significant for the Terminal. The magnitudes of the correlations also tended to be greater in the Enroute Sample.

Noteworthy exceptions to this occurred in the case of the VFR Tower and Approach Control Tower Experience. Both of these individual types of experience were more highly related to the Laboratory Grade Average in the Terminal Sample than in the Enroute Sample. A similar difference in favor of the Terminal Sample was also found with the Sum of Air Traffic Experience. The results seemed logical since these types of experience might be expected to generalize more readily to Terminal air traffic control than to Enroute and be of particular importance in the laboratory phase of training.

With the job performance criterion, two of the experience variables in the Enroute Sample had significant correlations, but only one of these, the Sum of Air Traffic Experience, was positive and this at a relatively low level, .12. In the Terminal Sample, the only two significant correlations with the Average Supervisory Rating were negative in sign.

As far as the magnitudes of the correlations were concerned, only one of the experience vs. criteria coefficients exceeded a value of .3 and only 11 exceeded a value of .2. Considering that there were 60 significant correlations in the tables, 29 of which were negative, and that of the 11 correlations greater than .2, 3 were negative, it can be concluded that previous job-related experience was a poor predictor of subsequent training and job performance. In fact, as the large number of negative correlations indicate, some types of experience which have been considered as qualifying should not have been.

The distribution data and the significance of the chi squares for the 12 individual experience variables vs. criterion measures given in Table 6 are in close agreement with the correlations of Tables 4 and 5. In every instance where the observed and expected values differed, tetrachoric correlation coefficients would have the same algebraic sign as the corresponding correlations in Tables 4 and 5.

Table 6 does not contain data for the Pass-Fail criterion in the Terminal Sample. An examination of that sample revealed that in only two instances did more than 4 of the 19 failures indicate that they had had a particular type of

TABLE 6

Frequencies of Trainees Reporting Various Types of Experience and Chi Square Tests of the Experience—No Experience Dichotomy vs. Pass—Fail and Above—Below the Approximate Median of Academic plus Laboratory Grade Average

Variables	Enroute				Enroute: A + L Grade				Terminal: A + L Grade			
	Fail	Pass	Total	X^2	≤83	≥84	Total	X^2	≤85	≥86	Total	X^2
1. Pilot	45	80 (92) #	125	8.57*	80	45 (58)	125	7.11	25	21 (25)	46	NS*
2. Radio: Air/Grnd	23	27 (37)	50	11.47	40	11 (24)	51	13.97	13	4 (9)	17	7.00
3. GCI	14	22 (27)	36	NS	27	11 (18)	38	4.98	9	3 (7)	12	4.39
4. Station	17	23 (30)	40	6.10	25	15 (19)	40	NS	5	1 (3)	6	NS
5. Grnd/Air	39	71 (81)	110	6.47	71	39 (51)	110	6.67	22	11 (18)	33	6.81
6. P to P	21	22 (32)	43	12.67	34	10 (20)	44	10.78	9	3 (7)	12	4.39
7. VFR Tower	33	166 (147)	199	15.51	88	111 (92)	199	11.98	71	92 (89)	163	NS
8. App Con Tower	11	95 (78)	106	17.21	41	66 (50)	107	12.97	30	61 (49)	91	9.47
9. Rad App Con Tow	2	16 (13)	18	NS	9	9 (8)	18	NS	4	7 (6)	11	NS
10. Center	9	50 (44)	59	4.06	20	39 (27)	59	10.56	8	9 (9)	17	NS
11. GCA	17	130 (109)	147	22.79	61	87 (69)	148	13.20	33	39 (39)	72	NS
12. RAPCON	10	86 (71)	96	15.12	29	68 (45)	97	27.45	10	21 (17)	31	NS
** TOTAL N	129	366		495	268	231		499	110	131		241

Numbers in parentheses are expected values for Pass or Above Median trainees computed from a 2 x 2 contingency table. If the expected value is less than the observed value, the sign of the relationship is positive; if greater, negative.

* With 1 degree of freedom, a X^2 value of 3.84 is significant at the .05 level, 6.64 at the .01 level, and 10.83 at the .001 level. NS replaces X^2 values not reaching the .05 significance level. Yates correction was not applied.

** Total N is the maximum possible frequency which could occur as an entry in each column. From the Total N and the table entries, the 2 x 2 contingency tables can be reconstructed.

experience. Consequently, chi squares were computed for just these two variables — Pilot Experience and VFR Tower Experience. Neither was significant. Because of the small number of failures in the Terminal Sample and the extreme skewness of the individual experience variables, the relationships with the Pass-Fail criterion must be interpreted cautiously.

Age vs. Criteria — Detailed studies of the implications of training entry age for training and job performance have been reported elsewhere.[4,5] However, it should be noted in Tables 4 and 5 that the age-criteria correlations were all negative and that the largest coefficients occurred with the Laboratory Grade Average. This suggests that the older trainees had the most difficulty in acquiring performance skills as contrasted with academic knowledge. Since it is performance skill which is most critical in the control of air traffic, the finding is of some import.

Aptitude Tests and Education vs. Criteria.—Tables 7 and 8 contain the correlations of the aptitude tests and Education with the training and job performance measures. None of the Education vs. criteria correlations were statistically significant.

Out of 30 correlations in each sample between the tests and criteria, 27 were significant for the Enroute and 19 were significant for the Terminal. All significant correlations were positive in sign. With respect to the magnitude of the correlations in the Enroute Sample, 6 exceeded .4, 14 exceeded .3, and 22 exceeded .2. In the Terminal Sample, 3 exceeded .4, 9 exceeded .3, and 15 exceeded .2. A comparison of these magnitudes and levels of significance with those for the experience variables leaves no doubt as to the superiority of the individual aptitude tests and the composite test score for the prediction of the Academic Grade Average, the Combined Academic plus Laboratory Grade Average, and Pass or Fail in both Enroute and Terminal Samples and for the prediction of the Laboratory Grade Average in the Enroute Sample.

For the prediction of the Laboratory Grade Average in the Terminal Sample, experience was superior to the aptitude tests. This is most apparent in the correlation between Lab and ΣAT.

In the area of job performance, the aptitude tests were not very effective predictors of the Average Supervisory Rating. Although two of the individual tests and the composite test score were significantly and positively related to the Average Supervisory Rating in the Enroute Sample, none were significantly related in the Terminal Sample.

TABLE 7

Enroute Sample: First-Order Correlations and Partial Correlations (Age Held Constant Statistically) of Test Variables and Education with the Criterion Measures

Variables	A + L		P — F[1]		Acad		Lab		Super		Age
	r	Partial r	r	Partial r	r	Partial r	r	Partial r	r	Partial r	r
17. Educ	-03	05	-08	00	03	07	-06	03	-08	-01	28**
18. Space	34**	31**	30**	26**	37**	35**	28**	24**	04	00	-18**
19. Numerical	36**	39**	26**	30**	43**	45**	27**	30**	09	11	08
20. Abstract	45**	43**	36**	33**	44**	43**	39**	36**	12*	09	-17**
21. Analogies	28**	25**	18**	15**	30**	28**	23**	20**	13*	11*	-14**
22. ATP	37**	34**	28**	25**	32**	30**	35**	31**	15*	12*	-17**
23. Composite	52**	49**	39**	37**	53**	52**	44**	41**	17**	14*	-17**
Max.	487		483		488		486		310		489
Range of N Min.	468		464		470		467		301		470

[1] Point-biserial correlations; all other correlations are product-moment. Decimal points omitted from all correlations.
* Significant at less than the .05 level.
** Significant at less than the .01 level.

TABLE 8

Terminal Sample: First-Order Correlations and Partial Correlations (Age Held Constant Statistically) of Test Variables and Education with the Criterion Measures

Variables	A + L		P — F[1]		Acad		Lab		Super		Age
	r	Partial r	r	Partial r	r	Partial r	r	Partial r	r	Partial r	r
17. Educ	00	04	-03	01	05	06	-08	-03	01	00	16*
18. Space	14*	12	20**	17*	26**	25**	06	02	01	-01	-11
19. Numerical	33**	34**	19**	20**	45**	45**	14*	15*	10	10	01
20. Abstract	40**	37**	38**	35**	45**	44**	30**	24**	-03	-06	-25**
21. Analogies	12	10	09	08	22**	22**	02	-01	01	00	-08
22. ATP	37**	34**	33**	29**	38**	37**	31**	25**	07	05	-22**
23. Composite	38**	35**	33**	30**	49**	48**	23**	18**	05	03	-18**
Max.	240		239		240		236		203		240
Range of N Min.	211		210		211		207		180		212

[1] Point-biserial correlations; all other correlations are product-moment. Decimal points omitted from all correlations.
* Significant at less than the .05 level.
** Significant at less than the .01 level.

DISCUSSION

Previously reported research and the present investigation have confirmed the value of aptitude tests and knowledge of chronological age for the prediction of training performance of Enroute and Terminal types of Air Traffic Control Specialists. In contrast to this, pre-employment, job-related experience, which has been used as a selection standard, must be fractionated into particular kinds of experience to obtain a clear evaluation of its worth as a predictor of training performance.

More specifically, communications, GCI, and Pilot experience had little value as predictors of training performance. Indeed, the more communications experience a trainee had, the less well he tended to perform. On the positive side, a trainee with experience more directly related to air traffic control work, such as Tower, Center, GCA, or RAPCON experience, tended to perform better in training.

Significant and positive predictions of the job performance measure by both aptitude tests and experience variables were obtained only in the Enroute Sample, with the aptitude tests being superior; but neither the tests nor experience can be considered highly effective predictors. The reasons for the lack of predictability of the Average Supervisory Ratings, particularly in the Terminal Sample, are not clear, but several possibilities exist.

First, with respect to the aptitude tests, it should be recalled that the present test battery was selected on the basis of prediction of Enroute training performance and was validated independently on additional samples of Enroute and Terminal trainees. Even though the tests do an effective job of predicting Terminal training performance, it is likely that more discriminating tests can be found. The lack of a significant relationship in the Terminal Sample between the Laboratory Grade Average and the

Space and Analogies tests lends some support to this likelihood. If tests were found which improved the prediction of Terminal training performance, it may be inferred that predictable elements of training performance exist which are different for Enroute and Terminal trainees. The different, but predictable, elements might also exist to some degree in the job situation. If so, the prediction of job performance measures by a test battery developed through analysis of Terminal training performance measures would be improved. Research is currently underway to investigate this possibility.

Another factor is related to the preceding. Enroute trainees are assigned to less than thirty Air Route Traffic Control Centers; whereas Terminal trainees can be assigned to one of several hundred Terminals. The diversity of activity between Terminals is greater than between Centers. This being so, it may be that for Enroute trainees there is more similarity between what is done on the job and what is taught in training. A greater overlap between behavioral elements required in the training course and on the job could account for the superior prediction of Enroute job performance by tests selected so as to maximize the prediction of training performance.

A third factor, having relevance to this problem, concerns the ratings by job supervisors. If indeed, there is more similarity between activities in Centers than in Terminals, then the job supervisors in different Centers would be making their evaluations on a more nearly comparable basis than would their counterparts in Terminals. The greater similarity of the behavior underlying ratings by Enroute supervisors in different Centers should operate to increase inter-center rater reliability; i.e., comparable behavior would be rated in the same way in different Centers. With an increase in the reliability of the ratings, there would be a corresponding increase in the probability of finding significant relationships between the ratings and the tests and experience variables.

Finally, in studies similar to this one, it is usually found that training performance measures can be predicted more accurately than job performance measures. This is partially accounted for by the greater behavioral homo-geneity of trainees who complete the training course. Since the course itself acts as a screening device, the poorer students, and it is assumed the poorer job prospects, are eliminated. With the elimination of the more inept trainees, the range of behavior evaluated by job supervisors becomes restricted. This, in turn, tends to lower correlations between the predictors and job performance measures.

Overall, the findings of the present study support the use of aptitude tests in preference to undifferentiated, pre-employment, job-related experience for the prediction of training and job performance of Enroute and the training performance of Terminal Air Traffic Control Specialists. If only one selection method could be used, aptitude tests should be the method of choice.

However, since there were positive relationships between some specific types of experience and training and job performance, additional weight in the selection process should be given for such experience. It is recommended that for selection purposes, Tower, Center, GCA, and RAPCON experience be considered as positive indicators for ATCS training in the Terminal and Enroute specialties, and that GCI, Pilot, Station, and Communications experience should be given no weight in the selection procedure.

Selection by a combination of aptitude tests and experience requirements would have an additional advantage. Individuals having experience of the kinds found to be useful as predictors, are somewhat familiar with the work actually performed as an Air Traffic Control Specialist. Applications for ATCS training from such individuals can be assumed to indicate a continuing interest in the occupation. These individuals also would bring to the training situation an aviation-related vocabulary which would reduce the amount of training time spent on terminology familiarization.

It is well to remember, however, that the definitions of various types of experience are far from standardized and that their relationships to specific jobs may change markedly with changes in the job elements underlying the experience. In contrast, aptitude measures are not affected by this limitation and should remain better predictors for longer periods of time.

The preceding recommendations with respect to experience variables apply only to the Enroute and Terminal specialties. Studies currently underway on Flight Service Station (FSS) trainees, a third type of Air Traffic Control Specialist, may reveal a quite different picture of the predictive value of the various types of experience. In a Flight Service Station, communications activity forms a large share of the work. Consequently, it would not be surprising to find that for FSS trainees communications experience was a positive predictor of training and job performance.

With continued research a number of the preceding questions may be answered; and it may be possible to achieve even better prediction of ATCS training and job performance by identifying additional, important aptitude areas, improving the measurement of the present areas, or by developing differential predictor weights for the different ATC specialties.

SUMMARY

A study of over 700 Enroute and Terminal Air Traffic Control Specialist trainees revealed that different kinds of pre-employment, job-related experience had differential value for the prediction of training performance. In general, experience most directly related to air traffic control work was a positive predictor; experience related to communications and piloting was negative. It was also shown that for Enroute trainees only, a composite variable (ΣAT) representing the sum of tower, GCA, RAPCON/RATCC, and center experience had a statistically significant, but small, relationship with ratings of job performance. In contrast, aptitude tests were superior to the experience variables for the prediction of all training course performance measures of both types of trainees, with the exception of Laboratory performance of Terminal trainees. In the latter instance, Σ AT was superior to the tests. For job performance, it was found that only ratings of Enroute trainees could be predicted by the tests and that although the tests were superior to experience as predictors, the relationships were small. As determined in other studies, it was again shown that age at entry into training was negatively related to training and job performance and that the negative relationships were greatest with grades in simulated air traffic control work in the training school laboratories. The findings of the study led to the recommendation that selection of individuals for Air Traffic Control Specialist training be based upon an aptitude test battery *and* pre-employment, job-related experience in tower, GCA, RAPCON/RATCC, and center work.

REFERENCES

1. Brokaw, L. D.: School and job validation of selection measures for air traffic control training. WADC-TN-59-39, 6570th Personnel Research Laboratory, AMD, AFSC, Lackland AFB, Texas, 1959.
2. Cobb, B. B.: Problems in air traffic management: II. Prediction of success in air traffic controller school. *Aerospace Med.*, 33:702, 1962.
3. Trites, D. K.: Problems in air traffic management: I. Longitudinal prediction of effectiveness of air traffic controllers. *Aerospace Med.*, 32:1112, 1961.
4. Trites, D. K. and Cobb, B. B.: Problems in air traffic management: III. Implications of age for training and job performance of air traffic controllers. Rept. No. 62-3, Civil Aeromed. Res. Institute, FAA, Oklahoma City, Okla., 1962.
5. Trites, D. K. and Cobb, B. B.: Problems in air traffic management: III. Implications of training-entry age for training and job performance of air traffic control specialists. *Aerospace Med.* (in press.)