Clearance number ASC 99-0211

DETERMINANTS OF US AIR FORCE ENLISTED AIR TRAFFIC CONTROLLER SUCCESS[1]

Dr. Thomas R. Carretta
Air Force Research Laboratory
Wright-Patterson Air Force Base, OH

ABSTRACT

This paper summarizes 3 recent US Air Force studies regarding enlisted air traffic controllers (ATCs). Results indicated the current selection composite had acceptable validity for predicting training performance. Cut-score analyses revealed that changing training qualification requirements for the ATC career field would reduce attrition, but would also have undesirable effects on eligibility. Results of a survey of enlisted ATCs indicated they were generally satisfied and motivated. In addition, they identified several abilities required for on-the-job performance that are not measured by current USAF selection procedures. Implications for enlisted ATC selection and training as well as future research directions are discussed.

INTRODUCTION

A review of 3 US Air Force (USAF) Class "A" mishaps in 1993 and 1994 implicated air traffic controller (ATC) loss of situational awareness as a contributing factor. As a result, the USAF began several initiatives to review ATC operations, including an examination of the manpower and personnel structure of the career field. One of the recommendations was to review current ATC selection procedures. The perception was that the current selection battery lacked measures of abilities related to ATC success, such as attention span, concurrent multiple task performance, decision making, and spatial reasoning. A related concern was that the screening system was deficient in identifying ATC trainees likely to succeed in the career field. It was noted that the attrition rate in apprentice-level training has increased since 1990, despite an overall reduction in the number of ATC trainees.

As a result of this review, several studies were undertaken to evaluate current ATC screening procedures and recommend potential enhancements. These efforts included (a) reviewing the literature on ATC selection research within the USAF, as well as other military services and the Federal Aviation Administration (FAA), (b) analyzing archival data on the relationship between selection test scores and performance in apprentice-level training, and (c) surveying incumbent USAF ATC personnel. This paper will focus on the results of the second and third efforts.

STUDY I: ASVAB UTILIZATION

Study I examined the validity of the Armed Services Vocational Aptitude Battery (ASVAB) for predicting apprentice-level ATC training performance. Analyses were done at the composite-level instead of the test-level, as the Air Force was interested in addressing the practical issue of examining the validity of the battery as it is currently used.

Method

Participants. Participants were 1,069 USAF enlisted personnel who entered ATC training in calendar years 1990-1995 and who were tested on the ASVAB. Most of the participants were male (71.1%) and White (81.2%). Education level was at least high school graduate or equivalent. Age at entry into the military ranged from 17 to 27 years. The graduation rate for apprentice-level training in the sample was 75.2% (804/1069). The most common reason for attrition was poor academic performance (n = 161). The other 104 eliminations occurred for a variety of reasons (e.g. fear of controlling, inadequate performance, self-elimination).

Measures. All USAF enlisted applicants are required to take the ASVAB prior to joining the military. The ASVAB is a 10 test, multiple aptitude battery. Its factor structure (Ree & Carretta, 1994) and reliability (Earles & Ree, 1992) have been studied, and it has been validated for training (Earles & Ree, 1992; Ree & Earles, 1991) and job performance (Ree & Earles, 1992; Ree, Earles, & Teachout, 1994).

---

[1] Previously published as Carretta, T. R. (1999). Determinants of US Air Force enlisted air traffic controller success. Proceedings of the Tenth International Symposium on Aviation Psychology, Columbus, OH, 581-586.

The tests are General Science (GS), Arithmetic Reasoning (AR), Word Knowledge (WK), Paragraph Comprehension (PC), Numerical Operations (NO), Coding Speed (CS), Auto and Shop Information (A/S), Mathematics Knowledge (MK), Mechanical Comprehension (MC), and Electronics Information (EI). The tests are not used separately, but rather are combined into composites. The Armed Forces Qualification Test (AFQT = AR + 2WK + 2PC + MK) score is used for entry into the US military regardless of job specialty, and the Air Force uses 4 Aptitude Indices for determining eligibility for specific jobs: Mechanical (M = GS + 2A/S + MC), Administrative (A = WK + PC + NO + CS), General (G = AR + WK + PC), and Electronic (E = GS + AR + MK + EI).

The minimum qualifying AFQT percentile score for USAF entrance is 40. In addition, entry into the ATC career field requires passing a flight physical and a Reading Aloud Test, vision correctable to 20/20, and a minimum score of 53 on the G composite (US Air Force Personnel Center, 1995).

Predictors were the ASVAB M, A, G, and E composites. The criteria included final school grade (FSG) during technical training (graduates only) and passing/failing (P/F) training (graduates and eliminees). FSG ranged from 70 to 99 and represented the average percent correct on several multiple-choice tests.

Procedures. Correlations were corrected for restriction in range (Lawley, 1943) using data from an historical database of USAF enlisted applicants. Next, the validity of a summed composite (M + A + G + E) and a regression-weighted composite of the 4 ASVAB scores was computed in both observed and corrected for range restriction form. The regression analyses used forced-entry. Finally, after correction for range restriction, correlations of the ASVAB composites and ATC pass/fail training score were corrected for dichotomization of the criterion. Significance tests were conducted for the correlations of the test scores with the training criteria and for regressions that used the observed (uncorrected) data. All statistical tests were done using a .01 Type I error rate. No significance tests can be done for the corrected data. The rational for examining the validity of the summed composite and the regression-weighted composite was to determine the maximum predictive utility that could be found by using the ASVAB composites in combination. The use of more than one ASVAB composite (M, A, G, E) in addition to the AFQT for job qualification is not unusual in the Air Force.

Results

Final school grade. All zero-order correlations of the ASVAB scores with FSG and the regression were significant at the .01 level. As shown in Table 1a, both before and after correction for range restriction, the G (.372 and .569) and E composites (.379 and .561) were the best predictors of FSG. The other 2 ASVAB composites were less valid (M, .293 and .428; A, .194 and .403). The 4 ASVAB composites were then summed and the resulting score correlated with FSG. The correlation was .394 for the observed data and .577 for the data corrected for range restriction. Finally, FSG was regressed on the 4 ASVAB composites. The multiple R was .411 for the observed data and .595 after range restriction correction. A comparison of the regression-weighted model with the individual ASVAB composites showed that considering all 4 ASVAB composites incremented prediction of FSG beyond using any individual ASVAB composite by itself (e.g., regression model vs. G alone: .411 vs. .372; $F(3, 799) = 8.92$, $p < .01$; regression model vs. E alone: .411 vs. .379; $F(3, 799) = 7.38$, $p < .01$). Despite this statistical significance, from a practical standpoint, after either G or E has been entered, the other ASVAB composites do little to increment the prediction of FSG (about .03 or .04 increment).

Passing/failing training. As with the FSG analyses, all of the statistical tests involving the ASVAB composites and passing/failing training were significant at the .01 level. On the composite level, E was the best predictor of passing/failing training, both before and after correction for range restriction (.353 and .454; see Table 1b). The M (.315 and .396) and G (.270 and .391) composites were similar in validity. The A composite had the lowest validity

Table 1.
Correlation Matrix: US Air Force ASVAB Composites and ATC Training Performance

a. Final School Grade (FSG; n = 804):

Score	M	A	G	E	FSG
M	1.000	0.218	0.597	0.754	0.428
A	0.056	1.000	0.622	0.488	0.403

b. Passing/Failing (P/F; n = 1,069):

Score	M	A	G	E	P/F
M	1.000	0.218	0.597	0.754	0.396
A	0.059	1.000	0.622	0.488	0.244

G	0.503	0.366	1.000	0.856	0.569	G	0.517	0.357	1.000	0.856	0.391
E	0.724	0.203	0.770	1.000	0.561	E	0.743	0.200	0.765	1.000	0.454
FSG	0.293	0.194	0.372	0.379	1.000	P/F	0.315	0.102	0.270	0.353	1.000

Note. Correlations below the diagonal are observed. Those above the diagonal were corrected for range restriction (Lawley, 1943). For FSG (n = 804), observed correlations greater than or equal to .085 are statistically significant at the .01 level (1-tailed test). For P/F (n = 1,069), observed correlations greater than or equal to .073 are statistically significant at the .01 level (1-tailed test).

(.102 and .244). As in the FSG analyses, the 4 ASVAB composites were summed and correlated with ATC P/F. The correlation was .358 for the observed data and .458 for the corrected data. The multiple R for the regressions of ATC P/F on the 4 composites were .365 and .465 for the observed and corrected correlations. It should be noted that in the P/F regression-weighted models, the G composite received a negative beta weight. This would be inappropriate and problematic in an operational selection system, because it would penalize applicants for good performance (i.e., high scores) on the tests making up the G composite.

Comparison of the regression-weighted model with the individual ASVAB composites showed little practical increment in predictive utility (e.g., regression model vs. E alone: .365 vs. .353; F (3, 1,064) = 3.28, ns). As observed with FSG, the E composite alone was nearly as predictive of ATC P/F as when used in combination with the other ASVAB composites. The fully corrected (range restriction and dichotomization) validities present a similar picture, with all values increasing as expected: M (.518), A (.519), G (.510), E (.593), summed composite (.599), and regression-weighted composite (.606).

STUDY 2: ALTERNATIVE CUT SCORE ANALYSIS

One way to improve the effectiveness of the G composite is to raise the minimum or "cut" score required for entrance into the ATC career field. Inspection of the graduation rates by individual G percentiles suggested that raising the minimum from 53 to 62 would produce a 5.4% decrease in the overall attrition rate from 25% to 20%. Alternatively, the E composite might be substituted for the G composite, with a cut score of 54 producing a reduction in attrition comparable to a cut score for G at 62 (4.9%).

Raising or changing a cut score can result in reduced attrition, but may have other less desirable consequences. For example, a cut score of the 90th percentile would clearly screen out a high number of applicants likely to fail, but would also "qualify" too few trainees to meet organizational needs. In addition, changes in cut score may have deleterious consequences on the rate of ethnic minorities and females that qualify for a job. Study 2 was conducted to address possible consequences of raising the G composite minimum or substituting an E composite minimum requirement of 54 for the present G minimum of 53.

Method

Participants. Participants were 216,207 USAF enlisted applicants who tested on ASVAB. There were 154,407 males and 61,800 females and 161,402 Whites, 37,478 African-Americans, 9,783 Hispanics, 902 Native Americans, and 4,467 Asians. The remaining 2,175 records lacked racial identity data.

Results

Raising the minimum G composite. Raising the minimum qualifying G score from 53 to 62 reduced the number of eligible males from 93,369 to 77,915, representing a reduction of almost 17%. For females, the number of eligible candidates fell from 31,592 to 23,709, a reduction of about 25%. Raising the minimum G score reduced the number of eligible White applicants from 107,585 to 87,614, a reduction of almost 19%. The number of eligible African-American candidates would fall from 12,232 to 8,087 (i.e., a reduction of about 34%).

Using a different composite. Changing the eligibility requirement for ATC training qualification to be an E composite of 54 *increased* the number of eligible male candidates from 93,369 to 100,193 (i.e., an *increase* of about 7%). However, if this procedure were used, the number of eligible female candidates would *decrease* from 31,392 to 24,334, a reduction of about 23%. If the minimum qualifying score were changed to be an E score of 54, the number of eligible White candidates would be reduced from 107,585 to 104,267, a reduction of only about 3%. The number of African-American candidates would fall from 12,223 to 11,640, a reduction of about 5%.

STUDY 3: ATC INCUMBENT SURVEY

Although the ASVAB composites were shown to be valid predictors of apprentice-level training, program managers for the enlisted ATC career field

were concerned that the ASVAB could not identify candidates likely to fail for non-academic reasons. They wanted to determine whether there were additional ability factors not covered by the ASVAB that could improve prediction of training performance. In response to program managers' concerns, a coordinated effort was undertaken to survey enlisted ATCs to identify the personnel characteristics and organizational factors that may influence training and job performance. It was intended that results of this effort be used to help design a preliminary selection system.

Method

Participants. The survey sample consisted of 181 incumbent enlisted ATCs. The majority of the participants were male (n = 155; 85.6%). The grade structure of the sample was: E-4 (n = 41; 22.7%), E-5 (n = 71; 39.2%), E-6 (n = 31; 17.1%), E-7 (n = 27; 14.9%), E-8 (n = 3; 1.7%), E-9 (n = 2; 1.1%), and missing (n = 6; 3.3%).

Measures. The survey (Siem & Carretta, in press) was designed to assess the importance of basic abilities, organizational aspects, and work environment factors thought to underlie ATC performance and to define key issues related to success in the enlisted ATC career field. Items that addressed organizational and personal concerns were developed based on interviews with trainers and program managers. The survey included 86 questions divided into 4 sections: Background Information (demographic information, job satisfaction), Motivation (preferences for different types of work environments, extent to which the ATC career field was rewarding), Situational (quality of life, acceptance of responsibility, decision making, attitudes toward temporary duty assignments), and ATC Abilities (importance of several abilities for successful ATC performance).

Procedures. Surveys were mailed to each duty location and supplied to participants by their supervisor. The survey was distributed to 200 incumbent ATCs at 19 bases. Completed surveys were placed in a sealed envelope and returned to Brooks AFB for analysis. The survey protocol had been reviewed and approved by the Air Force Occupational Measurement Squadron located at Randolph AFB, TX. Participation was voluntary and responses to survey questions were confidential. Informed consent was obtained from all participants prior to their participation.

Results

Because of the length of the survey, results presented below represent only a summary. Siem and Carretta (in press) provide a complete list of survey questions and detailed information on responses.

Background information. Examination of responses to the job satisfaction questions revealed that generally the enlisted ATCs had positive feelings about their job. Comparing their job to other enlisted specialties, the ATCs rated it as more interesting, providing a greater likelihood of using their training and talents, and providing a greater sense of accomplishment. Enlisted ATCs also stated that they were seldom made to feel uncomfortable in their job and usually were treated with respect. When asked about the likelihood of reenlistment, about 15% indicated they would retire (with at least 20 years service), 25% indicated that they would probably/definitely not reenlist, and 60% indicated that they probably/definitely would reenlist.

Motivation. Mean responses to the Motivation questions indicated a very positive attitude toward the ATC career field. Responses were made using a 7-point scale from (1) Strongly Disagree to (7) Strongly Agree. Means for 7 of the 12 questions were 6 or greater and indicated that the respondents liked the work environment and the high level of responsibility associated with their duties and that they found the job rewarding and exciting.

Situational. Overall, responses to these questions can best be described as neutral. Enlisted ATCs were neither extremely satisfied nor dissatisfied with the quality of life, temporary duty assignments, and technical instructors' concern toward students. The highest rated questions indicated they felt the ATC job carried a greater level of responsibility than other enlisted specialties ($\underline{M}$ = 6.0, SD = 1.3) and that mistakes were treated more severely for ATCs than other enlisted specialties ($\underline{M}$ = 6.0, SD = 1.2).

ATC abilities. Questions regarding the importance of various abilities for successful ATC performance were divided into 2 sections (i.e., 16 agree-disagree scales; 29 requirements scales). Responses to the agree-disagree questions used a 7-point scale that ranged from (1) Strongly Disagree to (7) Strongly Agree. Overall, the means for these questions were very high, indicating that respondents felt these abilities to be important for successful job performance. The mean value across all 16 agree-disagree questions was 5.975. Mean values for individual questions ranged from 4.1 to 6.7 and 12 questions had values of 6 or greater. The ability rated least important had to do with understanding basic geometry ($\underline{M}$ = 4.1, SD = 1.5). The abilities rated most important dealt with the ability to prioritize ($\underline{M}$

= 6.7, SD = 0.6), assimilate information and make correct decisions (M = 6.6, SD = 0.6), work well in stressful environments (M = 6.5, SD = 0.8), and anticipate what has not yet happened (M = 6.5, SD = 0.7).

For the requirements scales, respondents rated the importance of several abilities *relative* to their importance for other enlisted specialties. Scale values ranged from (1) Very Low to (7) Very High. Results were consistent with a view of the ATC job requiring high levels of cognitive capacity and information processing and the ability to work well under stress. The abilities rated least important had to do with exerting muscular strength (M = 2.8, SD = 1.4) and the psychomotor abilities of control precision (M = 3.3, SD = 1.5) and multi-limb coordination (M = 3.7, SD = 1.5). The most highly rated abilities were memorization and retention of new information (M = 6.1, SD = 0.9), spatial orientation/visualization (M = 6.1, SD = 1.0), the ability to work well in stressful environments (M = 6.1, SD = 0.9), the ability to shift between two or more sources of information (M = 6.0, SD = 0.9), and combine and organize information (M = 6.0, SD = 1.0).

DISCUSSION

Results indicated that current USAF selection procedures offer good prediction of enlisted ATC training performance. ASVAB validities were consistent with prior studies of enlisted ATC trainees (Stoker, Hunter, Batchelor, & Curran, 1987) and for a similar training specialty, weapons directors (Ree & Carretta, in press).

Alternative cut score analyses examined the impact on attrition rate for either raising the minimal G composite or for using the E composite instead of G. Results indicated that although raising the minimal G composite would reduce attrition by about 5%, the number of enlistees eligible for ATC training would decline by over 20%, making it difficult to recruit enough candidates for training. Using the E composite in lieu of G also would reduce attrition by about 5%, but would have less of an overall impact on reducing the number of eligible candidates. However, using the E composite would be unacceptable as it would produce adverse impact for female candidates.

Results of a survey of enlisted ATCs indicated a high level of job satisfaction and motivation. Respondents liked the work environment and the high level of responsibility associated with their duties and said they found the job rewarding and exciting. Further, in most respects enlisted ATCs felt their job to be comparable to other enlisted specialties. The most notable exceptions were that ATCs felt their job carried a greater level of responsibility than other enlisted specialties and that mistakes were treated more severely for ATCs than other enlisted specialties.

Survey respondents identified several abilities needed for successful on-the-job performance that are not measured by the ASVAB. These included memorization and retention of new information, spatial orientation/visualization, the ability to work well in stressful environments, the ability to shift between 2 or more sources of information, and the ability to combine and organize information. These survey results are consistent with a recent ATC job analysis reported by the RAF (Bailey, 1997). The most important abilities in the RAF analysis were spatial (i.e., reasoning/visualization), attentional capacity (i.e., ability to process and store information in real time; deal with multiple tasks involving auditory/visual information; concentrate over long periods; note and remember changes over short/long periods), and work rate (i.e., solve simple problems quickly and accurately).

Based on the results of the ability requirements survey, USAF ATC program managers felt that a screening device that measures these abilities may help reduce attrition in training. They also expressed an interest that the screening device resemble the tasks ATCs perform on the job (i.e., have face validity). As a result, the USAF has begun a study to evaluate the utility of a "job sample" test for enlisted ATC selection. In this test, which was developed by the FAA (Broach & Brecht-Clark, 1993), participants must control aircraft, adjusting their speed, altitude, and direction in order to send them to their proper destination (airport or transfer gate). Although data collection has begun, this project is expected to take a few more months to complete. Analyses will focus on the predictive utility of the job sample test and whether or not it adds to the predictiveness of the ASVAB.

In the mid-1980's, the USAF conducted a similar validation study (Stoker et al., 1987. In that study, the validity and incremental validity of several experimental tests for predicting enlisted ATC training outcome were examined in the presence of the ASVAB composites. The experimental tests included paper-and-pencil versions of an ATC job sample test (Multiplex Controller Aptitude Test or MCAT; Dailey & Pickrel, 1984) and 4 perceptual and spatial tests (Object Completion, Rotated Blocks, Perceptual Abilities, Electrical Maze). Regression analyses revealed that the MCAT and Rotated Blocks tests incremented the validity of the ASVAB

composites when predicting ATC pass/fail outcome. Despite Stoker et al.'s recommendations, neither the MCAT nor the Rotated Blocks tests were operationally implemented to augment the ASVAB for enlisted ATC candidate selection. Based on the results of Stoker et al., we are optimistic that the experimental computer-based ATC job sample test will demonstrate validity and incremental validity for enlisted ATC training.

REFERENCES

Bailey, M. (1997). *Air traffic controller task analysis* (Psychologist Report 7/97). Cranwell, UK: Royal Air Force, Directorate of Recruiting, Selection, and Initial Officer Training.

Broach, D., & Brecht-Clark, J. (1993). Validation of the Federal Aviation Administration air traffic control specialist pre-training screen. *Air Traffic Control Quarterly, 1*, 115-133.

Dailey, J. T., & Pickrel, E. W. (1984). Development of the Multiplex Controller Aptitude Test. In S. B. Sells, J. T. Dailey, & E. W. Pickrel (Eds.), *Selection of air traffic controllers* (FAA-AM-84-2), (pp. 281-297). Washington, DC: FAA Office of Aviation Medicine.

Department of the Air Force (1996). *Plan of instruction: Air traffic control operations apprentice and combat control apprentice,* POI E3ABR1C131 000. Keesler AFB, MS: Keesler Training Center.

Earles, J. A., & Ree, M. J. (1992). Training validity. *Educational and Psychological Measurement, 54,* 721-725.

Lawley, D. N. (1943). A note on Karl Pearson's selection formulae. *Proceedings of the Royal Society of Edinburgh, Section A, 62, Part 1,* 28-30.

Ree, M. J., & Carretta, T. R. (1994). Factor analysis of ASVAB: Confirming a Vernon-like structure. *Educational and Psychological Measurement, 54,* 457-461.

Ree, M. J., & Carretta, T. R. (in press). Lack of ability is not always the problem. *Journal of Business and Psychology*.

Ree, M. J., & Earles, J. A. (1991). Predicting training success: Not much more than *g*. *Personnel Psychology, 44,* 321-332.

Ree, M. J., & Earles, J. A. (1992). Intelligence is the best predictor of job performance. *Current Directions in Psychological Science, 1,* 86-89.

Ree, M. J., Earles, J. A., & Teachout, M. S. (1994). Predicting job performance: Not much more than *g*. *Journal of Applied Psychology, 79,* 518-524.

Stoker, P., Hunter, D. R., Batchelor, C. L., & Curran, L. T. (1987). *Air traffic controller trainee selection* (AFHRL-TP-87-19). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

U. S. Air Force Personnel Center (1995). *United States Air Force manual 36-2108: AFSC entry requirements*, attachment 41 (Oct. 1994). Randolph AFB, TX: Air Force Personnel Center.