An Evaluation of the Utility of At-Sat for the Placement of New Controllers by Option

Cristina L. Byrne

Dana Broach

AN EVALUATION OF THE UTILITY OF AT-SAT FOR THE PLACEMENT OF NEW CONTROLLERS BY OPTION

Cristina L. Byrne and Dana Broach
FAA Civil Aerospace Medical Institute
Oklahoma City, OK

In this study, we investigated the utility and fairness of using the Air Traffic Selection and Training (AT-SAT) test battery to place Air Traffic Control Specialist (ATCS) applicants into terminal or en route facilities. While results of statistical analyses indicated that AT-SAT could be considered a valid tool for use in placement, based on technical considerations only, it was concluded that it should not be used in that way due to lack of utility and potential for adverse impact. If the FAA were to use AT-SAT for placement, the risk of additional adverse impact and pay disparities should be evaluated against the marginal utility of placement in terms of changes in field training success rates.

The air traffic control specialist (ATCS, or controller) occupation is considered to be an intellectually challenging, important, and prestigious career field by the majority of recently hired developmental controllers (Cannon & Broach, 2011). The Federal Aviation Administration (FAA) projects hiring approximately 1,300 new controllers per year over the next five years to replace retiring controllers (FAA, 2014). Excluding rehires or others with previous ATCS experience, it is required that applicants receive a passing score on an aptitude test to be hired into the occupation (U.S. Office of Personnel Management [OPM], 2013). Currently, the computer-administered Air Traffic Selection and Training (AT-SAT) test battery is the aptitude test used by the FAA to assess applicants under the OPM occupational qualification standards.

The validity of AT-SAT as a predictor of ATCS job performance was demonstrated in two concurrent, criterion-related validation studies. The first study was reported in 2001 by Ramos, Heil, and Manning (2001a, b). Approximately 1,000 incumbent en route controllers took the proposed test battery and job performance data were collected. The correlation between scores on the test battery and a composite of the job performance measures collected was .51 without any corrections for range restriction or criterion unreliability. With correction for incidental range restriction, the correlation was .68 (Waugh, 2001). The American Institutes for Research (AIR®, 2012) conducted the second study, named the Concurrent Validation of AT-SAT for Tower Controller Hiring (CoVATCH). Incumbent air traffic control tower (ATCT) controllers (N = 302) took the current operational version of the AT-SAT test battery and job performance measures were collected (see Horgen et al., 2012). The correlation between a regression weighted composite of AT-SAT subtest scores and the composite of the two criterion measures was .42 without any statistical corrections (AIR®, 2012). These two studies independently demonstrated that AT-SAT is a valid predictor of ATCS job performance.

Before placement can be discussed, it is useful to understand the nature and structure of the organization within the FAA responsible for air traffic control operations and facilities. This organization, called the Air Traffic Organization, or ATO, can be divided into two major partitions, also referred to as options: Terminal Services and En Route/Oceanic Services. New hires can be placed into either the Terminal option or the En Route option. At en route centers, controllers handle high altitude air traffic between airports, work that is generally considered very complex and demanding. At TRACONs, controllers direct traffic within about 50 miles of an airport, usually during initial climb and final descent of the aircraft. This work can also be considered very demanding. The work at ATCTs involves directing air traffic on the runways and in the immediate vicinity of the airport, as well as issuing takeoff and landing clearances. This type of air traffic control (ATC) is generally considered somewhat less complex and less demanding than radar ATC, but that can vary greatly by location. Historically, the failure rate in on-the-job training for new controllers has been higher in en route facilities (Manning, 1998). Controller positions at en route centers generally have the highest pay grades in the occupation. Controller positions in towers generally have lower pay grades than en route positions. Thus, there are both organizational (success and failure rates in facility on-the-job training) and individual economic consequences attached to placement decisions. Moreover, because placement affects the terms and conditions of employment (especially starting pay), it is an employment decision as defined by the Uniform Guidelines on Employee Selection Procedures (29 C.F.R. § 1607.2B) (EEOC, 1978). Therefore, using AT-SAT scores for placement, as recommended by the Department of Transportation Inspector General (2010), must be validated.

To use a test score for placement purposes, the relevant professional standards and principles require “evidence that scores are linked to different levels or likelihoods of success among jobs” (American Educational Research

Association, American Psychological Association, & National Council on Measurement in Education, 1999, p. 160). Relevant evidence might include a pattern of differential relationships between predictors and criteria by job type (Society for Industrial and Organizational Psychology [SIOP], 2003). In their analysis of the CoVATCH data, AIR® (2012) was not able to provide any evidence suggesting that AT-SAT could be used for placement by facility level which is based on complexity and traffic load. However, they did report evidence for validity by option in that the regression equation (i.e., the weight given to each subtest) for tower was not identical with the equation for en route as reported by Ramos et al. (2001a, b). Unfortunately, there were few differences in the recommended option placement when the two equations were used to hypothetically place individuals in their sample of 300 tower controllers. AIR® concluded that AT-SAT might be used for placement by option but further analyses were needed. Thus, the purpose of the current study was to extend the AIR® analysis by using AT-SAT validation data collected from both en route and terminal controllers.

Placement Rules

AT-SAT scores might be used for placement in many different ways. For example, persons with scores above some cut-off might be assigned to the en route option, while persons with scores below that cut-off might be assigned to terminal. Or scores might be categorized into ranges, with persons in the lowest range assigned to one option, persons in the highest range assigned to the other option, and persons with scores in the middle range assigned to either option, depending on agency needs. Thus, the first step in this analysis was to decide how AT-SAT scores would likely be used for placement. To use AT-SAT for placement decisions, AIR® suggested computing a score for each option, based on the option-specific regression equations. As no AT-SAT validation study has been conducted specifically for TRACON positions, the equation derived from the tower sample was used to represent the entire terminal option. The applicants would then be assigned to a score band within each option. For example, an applicant could be classified as Well-Qualified Terminal and Qualified En Route (or vice versa), Well-Qualified in both, or Qualified in both. Current use of AT-SAT defines Well-Qualified as a score of 85-100, Qualified as a score of 70-84.9, and Not Qualified as 69.9 and below.

The placement procedure suggested by AIR® is feasible but has three drawbacks. First, the overall ranking of an individual (which impacts hiring decisions) is confounded with their ranking within an option (which impacts placement decisions). This might make the initial selection of a candidate more complicated and less systematic with more judgment and consideration being required of decision makers for each individual case. Second, given the width of the categorical bands and the correlation found between the current en route score (used by AIR® as the basis for en route placement) and the tower score (r = .65, see Table 1), it would be expected that if the placement rules suggested by AIR® were used, a good portion of candidates would receive the same categorical ranking for both options (i.e., Well-Qualified or Qualified in both options). Third, the en route equation was reweighted to find an optimal balance between validity and the reduction of adverse impact, but the tower equation reported by AIR® was not weighted in a similar way. This suggests that the en route equation used by AIR® in their analysis would produce a different option score for reasons other than "true" subtest relationships to performance.

Table 1.
Correlations between Current, En Route, and Terminal AT-SAT Scores and 1st Facility Success.

	Current	En Route	Terminal
Current			
En Route	.880		
Terminal	.651	.793	
1st Facility Success	.120	.210	.176

Note. All correlations significant at $p < .01$, $n = 2{,}332$. Current, En Route, and Terminal AT-SAT scores are based on similar, but slightly different equations developed through two AT-SAT validation studies.

Taking these drawbacks into account, we investigated an alternative approach to placement. The first step would be to categorize individuals using the current operational AT-SAT equation, which was weighted to mitigate adverse impact (Wise, Tsacoumis, Waugh, Putka, & Horn, 2001), into Well-Qualified, Qualified, and Not Qualified categories using the current cut scores as a basis for initial selection. Second, two additional composite scores would be computed based on (a) the original, unadjusted weights for en route (Ramos et al., 2001a, b), and (b) the tower equation developed by AIR® (2012) for terminal. For convenience, these will be referred to as the Current, En Route, and Terminal scores, respectively, throughout the rest of this paper. The applicant's hiring status would first be determined by using the Current score to determine the initial categorical rankings. Persons categorized as "Not Qualified" on the basis of their Current score would be removed from further consideration. Next, the En Route and Terminal scores would be computed for each person using the respective option-specific weights. Whichever score

was highest would serve as the placement recommendation. In the rare event of a tie, the applicant would be given a recommendation of "Either." The initial categorization (i.e., Qualified or Well-Qualified) based on the Current score would then be attached to this option recommendation.

Evaluation of Proposed Placement Approach

The following analyses were conducted to evaluate the proposed placement approach. First, logistic regression analyses were completed to verify the relationship of AT-SAT scores (computed using the three equations) to first facility training success, a criterion measure not used in the two previous concurrent, criterion-related validation studies. First facility training success refers to whether or not developmental controllers achieved certified professional controller (CPC) status at their first facility. Second, cross-tabulations were computed to examine the potential outcomes and utility of using AT-SAT for placement. Third, given that placement would constitute an employment decision encompassed by the Uniform Guidelines on Employee Selection Procedures, the potential for adverse impact was assessed against the 4/5ths rule (29 C.F.R. § 1607.4D) (EEOC, 1978).

The data used for these analyses were extracted from FAA AVIATOR, the Air Traffic National Training Database (NTD), the AT-SAT database, and the FAA Personnel and Payroll System (FPPS). Extracted information included AT-SAT test scores, race, gender, pay, and developmental training status. The sample used for the adverse impact analyses included anyone who had taken AT-SAT (N = 18,663) and who had race/gender information available (Race: N = 15,052; Gender: N = 14,115). The sample used for all other analyses included individuals who had AT-SAT data and a finalized first facility outcome (i.e., achieved CPC, failed, or transferred from first facility due to performance) by July 2012 (N = 2,332). In both samples, individuals had submitted an application for an ATCS position between 2007 and 2009.

Results

Logistic Regression

The results of the logistic regression analyses showed that the Current – $R^2 = .022$, χ^2 (1, 2332) = 32.71, $p \leq .001$, En Route – $R^2 = .064$, χ^2 (1, 2332) = 99.32, $p \leq .001$, and Terminal – $R^2 = .047$, χ^2 (1, 2332) = 71.88, $p \leq .001$ scores (based on the previously derived equations) were statistically significant predictors of first facility training success. The raw correlations between these scores and first facility training success, uncorrected for range restriction, were similar but not identical to each other and can be found in Table 1. These findings parallel the results obtained during both concurrent validation studies to assess the predictive of AT-SAT using ordinary least squares regression analyses and other types of job performance measures, as well as the results of a longitudinal validation of AT-SAT using first facility training success as the criterion (see Broach et al., 2013). Additionally, when logistic regression analyses were run separately for the En Route and Terminal samples (not restricting subtest weights based on the previously derived equations), the subtest scores were differentially correlated with first facility training success similar to the findings of AIR® that the subtest weights using a sample of tower controllers were not identical to those found in the original en route validation study. Taken together, this evidence demonstrates some degree of differential validity and prediction, both technical requirements as described by the Standards for Educational and Psychological Testing (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999), as well as the Principles for the Validation and Use of Employee Selection Procedures (SIOP, 2003).

Cross-Tabulation Analysis

The next step in the analysis was to cross-tabulate actual vs. hypothetical placement. "Actual Placement" was the official assignment of newly hired controllers to en route or terminal facilities without regard to their Current score on AT-SAT. "Hypothetical Placement" was the decision that would have been made using the En Route and Terminal scores derived from AT-SAT. Placement was "correct" when the actual placement matched the hypothetical placement based on En Route and Terminal scores; placement was "incorrect" when the actual placement did not match the hypothetical placement.

The cross-tabulation was computed for those who were placed "correctly" or 'incorrectly" by their success in field training (Table 2), and the results were adjusted based on the typical proportion of hires assigned to each option (Table 3). The results of this analysis suggest that the utility of using AT-SAT for placement is marginal to slightly negative, depending on how the data are examined. The FAA could potentially see a 3% increase to 80% in the success rate of controllers "correctly" placed into the en route option as compared to the baseline success rate (without AT-SAT guided placement) of 77%. However, this gain could be offset by a 5% reduction to 74% in the

success rate of those "correctly" placed into the terminal option as compared to a baseline success rate (without AT-SAT guided placement) of 79% (see Table 3) for a net reduction in success rates across both options of 2%.

Table 2.
Cross-tabulation of Training Completion Rates at 1st Field Facility.

Actual Placement	Hypothetical Placement	Unsuccessful		Successful		
		N	%	N	%	Total
En Route**	En Route	111	20%	436	80%	**547**
En Route	Terminal	79	27%	218	73%	**297**
Terminal	En Route	153	17%	728	83%	**881**
Terminal**	Terminal	159	26%	448	74%	**607**
Total		**502**	**22%**	**1,830**	**78%**	**2,332**

Note. **Indicates "correct" placement – meaning that applicants were actually placed in the option that AT-SAT would have predicted had it been used for this purpose at the time of hire.

Table 3.
Training Success Rates at the 1st Field Facility with and without Placement.

	Success Rate without Placement	Success Rate with Placement
En Route (36% of positions)	77%	80%
Terminal (64% of positions)	79%	77%*
Across Options (weighted by number of positions)	78.28%	78.08%

Note. *Indicates rate adjusted for likelihood of filling 40% of terminal positions with applicants initially recommended for *En Route* placement

However, this loss must be reexamined and weighted within the context of the number of positions available in each option and the number of controllers being hypothetically placed in the en route option. The overall baseline success rate in terminal without placement is driven upwards by the higher success rate of individuals that would hypothetically have been placed in the en route option. Given the number of applicants that scored higher on the En Route equation, as compared with the number of positions typically available for en route controllers in recent years, it is estimated that approximately 40% of available terminal positions could be filled by individuals with en route recommendations. This would likely be the preferred policy given their apparent ability to succeed in either option.

Thus, to accurately estimate the overall success rate with AT-SAT guided placement for terminal, given the likely situation that 40% of the positions could be filled by applicants scoring higher on the En Route equation (who would likely have higher success rate – 83% vs. 74%), a weighted average was computed. The overall success rate, assuming placement of some applicants with En Route placement recommendations into the terminal option then becomes 77% [(83% success rate x 40% of the positions) + (74% success rate x 60% of the positions)] instead of 74% for terminal positions. This computation results in a success rate 2% lower than the current terminal success rate seen without using AT-SAT for placement.

In sum, if AT-SAT is used to guide placement by option, there is a potential increase in success rates for those placed in en route of 3% but a potential decrease for those placed in terminal of 2%, for an overall 1% increase in success rates. However, this estimate must also be considered within the context of the ratio of people hired into each option. Generally speaking, because more people are hired into the terminal option (accounting for approximately 64% of open positions yearly), the decrease in the terminal success rate must be weighed more heavily in the calculation of overall success rates computed with and without placement (Table 3). Taking the higher hiring rate in the terminal option, the net effect of using AT-SAT for placement would likely be a very slight reduction in the overall success rate across both options (Table 3).

Adverse Impact Analysis

As with other employment decisions, a placement decision carries with it the potential to impact an individual's ability to earn. Given the nature of this decision, the potential for adverse impact against members of protected groups must be considered. Using data from FPPS, it was determined that en route controllers earn on average approximately $20K more per year than terminal controllers. The difference in annual salaries was calculated using a snapshot of the FPPS data captured in July 2012. On average, receiving a recommendation for placement into the en route option would likely provide an individual with a greater opportunity to earn more over the course of employment and is, thus, considered the preferred option for calculating adverse impact.

Table 4.
Adverse Impact from Placement Decision.

	Hypothetical Placement			En Route Placement Rate	Adverse Impact Ratio[a]
	En Route	Terminal	Total		
By Ethnicity					
Asian	228	228	456	.50	.95
Black	713	2,324	3,037	.23	**.45**
Hawaiian-Pacific Island[b]	26	49	75	.35	**.66**
Hispanic-Latino	269	556	825	.33	**.62**
Native American-Alaskan Native	30	35	65	.46	.88
White	4,632	4,209	8,841	.52	
Multi-racial	462	569	1,031	.45	.86
No groups marked	357	358	715	.50	.95
Total	**6,717**	**8,328**	**15,045**		
By Sex					
Female	1,103	2,320	3,423	.32	**.66**
Male	5,350	5,686	11,036	.48	
Total	**6,453**	**8,006**	**14,459**		

Note. [a]Adverse impact ratio calculated with respect to whites for ethnicity and male for sex.
[b]Groups comprising less than 2% of the applicant pool are italicized. Bold ratios are less than what is acceptable under the 4/5^{ths} Rule (0.80).

Using the placement rules previously described, assigning controllers to an option using their AT-SAT scores could result in differential placement rates by race and sex into the terminal and en route options (Table 4). For example, just 23% of black candidates would be recommended for placement in en route, compared to 52% of white candidates (adverse impact ratio = .23/.52, or .45, where the threshold for adverse impact is defined as a ratio of .80 or less by the Uniform Guidelines on Employee Selection Procedures [EEOC, 1978]). The adverse impact ratio for Hispanic/Latino applicants was .62 and for females was .66. This adverse impact would result in addition to the adverse impact these protected groups already face in assignment to the Well-Qualified category ranking for initial selection considerations.

Summary

Looking at both of the concurrent validation studies and this current set of analyses together, there is sufficient evidence to suggest that the abilities required to perform air traffic control tasks do vary, to some limited degree, by option. The regression analyses (calculated repeatedly using different samples and at different times) have, in fact, derived different equations for the two options, which overlap but are not completely identical. This evidence can help provide the technical justification required, if the FAA were to pursue the use of AT-SAT for placement.

However, it is not clear that the variation by option is of a sufficient degree to justify differential placement given the minimal utility observed. Moreover, the utility of using AT-SAT to guide placement is minimal – and might be slightly counterproductive for the FAA. The cross-tabulations indicated that the success rate in en route would increase if AT-SAT is used for placement but would decrease in terminal. Taken across both options, field training success rates would not likely change in a meaningful way provided that the number of candidates typically hired for each option in recent years remains consistent. Finally, placement using AT-SAT could potentially have adverse impact on individuals in protected classes. That is, members of protected classes would be placed into higher paying en route facilities at less than 80% of the rate of majority members of each class (race and gender). Differential placement rates on the basis of AT-SAT scores could create troubling pay disparities by race and sex. If the FAA were to use AT-SAT for placement, the risk of additional adverse impact and pay disparities should be evaluated against the marginal utility of placement in terms of changes in field training success rates..

Acknowledgements

Research reported in this paper was conducted under the Air Traffic Program Directive /Level of Effort Agreement between the Federal Aviation Administration Headquarters and the Aerospace Human Factors Division of the Civil Aerospace Medical Institute sponsored by the Office of Aerospace Medicine and supported through the FAA NextGen Human Factors Division.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing* (4th ed.). Washington, DC: American Psychological Association.

American Institutes for Research. (2012). *Validate AT-SAT as a placement tool.* (Draft report prepared under FAA contract DTFAWA-09-A-80027 Appendix C). Oklahoma City, OK: Federal Aviation Administration Aerospace Human Factors Research Division (AAM-500).

Borman, W. C., Hedge, J. W., Hanson, M. A., Bruskiewicz, K. T., Mogilka, H. J., Manning, C., Bunch, L. B., & Horgen, K. E. (2001). Development of criterion measures of air traffic controller performance. In Ramos, R. A., Heil, M. C., & Manning, C. A. (2001). *Documentation of validity for AT-SAT computerized test battery, Volume II*. (Report No. DOT/FAA/AM-01/6). Washington, DC: Federal Aviation Administration Office of Aviation Medicine.

Broach, D., Byrne, C. L., Manning, C. A., Pierce, L. G., McCauley D., & Bleckley, M. K. (2013). *The validity of the Air Traffic Selection and Training (AT-SAT) test battery in operational use.* (Report No. DOT/FAA/AM-13/3). Washington, DC: Federal Aviation Administration Office of Aerospace Medicine.

Cannon, M. M. & Broach, D. (2011). *Studies of next generation air traffic control specialists: Why be an air traffic controller?* (Report No. DOT/FAA/AM-11/12). Washington, DC: Federal Aviation Administration Office of Aerospace Medicine.

Equal Employment Opportunity Commission, Civil Service Commission, Department of Labor, & Department of Justice. (1978). Uniform guidelines on employee selection procedures. *Federal Register, 43*(166), 38290-39315.

Federal Aviation Administration. (2014). *A plan for the future: 10-Year strategy for the air traffic control workforce 2014-2023.* Retrieved from http://www.faa.gov/air_traffic/publications/controller_staffing/media/CWP_2014.pdf

Horgen, K., Lentz, E. M., Borman, W. C., Lowe, S. E., Starkey, P. A., & Crutchfield, J. M. (2012, April). *Applications of simulation technology for a highly skilled job.* Paper presented at the 27th Annual Conference of the Society for Industrial and Organizational Psychology, San Diego, CA.

Manning, C. A. (1998). Air traffic controller field training programs, 1981-1992. In D. Broach (Ed.), *Recovery of the FAA air traffic control specialist workforce, 1981-1992.* (Report No. DOT/FAA/AM-98/23). Washington, DC: Federal Aviation Administration Office of Aviation Medicine.

Nickels, B. J., Bobko, P., Blair, M. D., Sands, W. A., & Tartak, E. L. (1995). *Separation and control hiring assessment (SACHA) final job analysis report* (Deliverable Item 007A under FAA contract DFTA01-91-C-00032). Washington, DC: Federal Aviation Administration, Office of Personnel.

Ramos, R. A., Heil, M. C., & Manning, C. A. (Eds.). (2001a). *Documentation of validity for the AT-SAT computerized test battery, Volume I.* (Report No. DOT/FAA/AM-01/5). Washington, DC: Federal Aviation Administration Office of Aviation Medicine.

Ramos, R.A., Heil, M.C., & Manning, C.A. (Eds.). (2001b). *Documentation of validity for the AT-SAT computerized test battery, Volume II.* (Report No. DOT/FAA/AM-01/6). Washington, DC: Federal Aviation Administration Office of Aviation Medicine.

Society for Industrial and Organizational Psychology. (2003). *Principles for the validation and use of employee selection procedures* (4th ed.). Bowling Green, OH.

U. S. Department of Transportation Office of the Inspector General. (2010). *Review of screening, placement, and initial training of newly hired air traffic controllers.* (Report. No. AV-2010-049). Retrieved from http://www.oig.dot.gov/audits?tid=71

U.S. Office of Personnel Management. (2013). General schedule qualification standards: Air traffic control series - 2152. *OPM Classification and Qualifications.* Retrieved from http://www.opm.gov/policy-data-oversight/classification-qualifications/general-schedule-qualification-standards/2100/air-traffic-control-series-2152/

Waugh, G. (2001). Predictor-criterion analyses. In Ramos, R. A., Heil, M. C., & Manning, C. A. (2001). *Documentation of validity for AT-SAT computerized test battery, Volume II*. (Report No. DOT/FAA/AM-01/6). Washington, DC: Federal Aviation Administration Office of Aviation Medicine.

Wise, L. L., Tsacoumis, S. T., Waugh, G. W., Putka, D. J., & Hom, I. (2001, December). *Revision of the AT-SAT.* (Report No. DTR-01-58). Alexandria, VA: Human Resources Research Organization (HumRRO).