



**Federal Aviation
Administration**

DOT/FAA/AM-12/8
Office of Aerospace Medicine
Washington, DC 20591

Incremental Validity of Biographical Data in the Prediction of En Route Air Traffic Control Specialist Technical Skills

Dana Broach

Civil Aerospace Medical Institute
Federal Aviation Administration
Oklahoma City, OK 73125

July 2012

Final Report

NOTICE

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

This publication and all Office of Aerospace Medicine technical reports are available in full-text from the Civil Aerospace Medical Institute's publications Web site:
www.faa.gov/go/oamtechreports

Technical Report Documentation Page

1. Report No. DOT/FAA/AM-12/8	2. Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle Incremental Validity of Biographical Data in the Prediction of En Route Air Traffic Control Specialist Technical Skills		5. Report Date July 2012	
		6. Performing Organization Code	
7. Author(s) Broach, D		8. Performing Organization Report No.	
9. Performing Organization Name and Address FAA Civil Aerospace Medical Institute P.O. Box 25082 Oklahoma City, OK 73125		10. Work Unit No. (TRAIS)	
		11. Contract or Grant No.	
12. Sponsoring Agency name and Address Office of Aerospace Medicine Federal Aviation Administration 800 Independence Ave., S.W. Washington, DC 20591		13. Type of Report and Period Covered	
		14. Sponsoring Agency Code	
15. Supplemental Notes This work was completed under approved FAA Human Factors research task 1209AC082110. HRR523.AV9300, Evaluation of ATCS Biographical Data and Interview Selection Procedures.			
16. Abstract Previous research demonstrated that an empirically-keyed, response-option scored biographical data ("biodata") scale predicted supervisory ratings of air traffic control specialist (ATCS) job performance (Dean & Broach, 2011). This research focused on the validity of scores on the Controller Background Assessment Survey (CBAS) in predicting an objective, computerized measure of en route controller technical skills. Method. The analysis was conducted in two steps. First, computerized aptitude test battery ("AT-SAT") scores for 229 en route controllers were regressed on the Computer-Based Performance Measure (CBPM; Hanson, Borman, Mogilka, Manning, & Hedge, 1999). Second, biodata scores were entered into the equation. Results. AT-SAT scores accounted for 27% of variance in the criterion measure ($\beta=0.520$, adjusted $R^2=.271$, $p<.001$). Biodata accounted for an additional 2% of the variance in CBPM ($\beta=0.134$; adjusted $\Delta R^2=0.016$, $\Delta F=5.040$, $p<.05$). Discussion. The empirically-keyed, response-option scored biodata scale demonstrated incremental validity over the computerized aptitude test battery in predicting scores representing the core technical skills of en route controllers. Utility analysis suggested that even a small increment in validity was likely to have substantial organizational utility, given the high applicant volume and ATCS training costs. Further research to examine the relationship of CBAS scores to training outcomes at the FAA Academy and in field ATC facilities is recommended.			
17. Key Words Air Traffic Control Specialist, Personnel Selection, Biographical Data, Bootstrap		18. Distribution Statement Document is available to the public through the Internet: www.faa.gov/go/oamtechreports	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 13	22. Price

ACKNOWLEDGMENTS

Research reported in this paper was conducted under the Air Traffic Program Directive/Level of Effort Agreement between the Human Factors Division (ANG-C1), FAA Headquarters, and the Aerospace Human Factors Division (AAM-500) of the Civil Aerospace Medical Institute.

The opinions expressed are those of the author alone, and do not necessarily reflect those of the Federal Aviation Administration, the Department of Transportation, or the government of the United States of America.

INCREMENTAL VALIDITY OF BIOGRAPHICAL DATA IN THE PREDICTION OF EN ROUTE AIR TRAFFIC CONTROL SPECIALIST TECHNICAL SKILLS

The measurement of biographical data (or “biodata”) encompasses the notion of asking individuals to recall and report their typical, and sometimes, specific behaviors or experiences in a referent situation, generally from an earlier time in their lives (Mumford & Owens, 1987; Nickels, 1994). While approaches such as diaries have been used to collect biodata, the most common form is that of a scale, survey, inventory, or questionnaire. Such biodata instruments have demonstrated reasonable and useful reliability and validity in the prediction of job performance across a variety of occupations (see Stokes, Mumford, & Owens, 1994). Average cross-validities in the .3 to .4 range have been reported for biodata selection instruments in narrative and meta-analytic review (Asher, 1972; Hunter & Hunter, 1984; Reilly & Chao, 1982; Schmitt, Gooding, Noe, & Kirsh, 1984). Moreover, biodata scales can be constructed so as to have less adverse impact by race without significant loss in criterion-related validities (Dean, 1999).

The United States federal government has long had an interest in the development, validation, and use of biodata, reaching back to World War I, at least (see Farmer, 2002, for a review). The Federal Aviation Administration (FAA), in particular, has invested significant research effort in the development and validation of biodata instruments for the Air Traffic Control Specialist (ATCS, or air traffic controller) occupation. Early efforts focused on the validity of specific experience, such as having been a pilot or an air traffic controller in the military (Brokaw, 1957; Cobb & Nelson, 1974). The research program broadened following the 1981 strike by the Professional Air Traffic Controller Organization (PATCO) as the agency began to rebuild the ATCS workforce. Two biodata instruments were adapted by the FAA for research purposes: the Applicant Background Assessment (ABA); and the Biographical Questionnaire (BQ) (see Farmer, 2002, pp. 93-94 for more detailed history). The instruments were administered to most (but not all) newly hired controllers entering on duty for initial training at the FAA Academy between 1981 and 1992. Job-related outcomes such as performance in initial training at the FAA Academy and in on-the-job training (OJT) at the first assigned field facility were collected and matched with the biodata. The resulting datasets have been mined by multiple researchers to assess biodata predictive validity. Factors such as self-reported grades in high school mathematics were

found to be predictive of outcomes at the FAA Academy and in field OJT (Broach, 1992, 2008; Cobb, Young, & Rizutti, 1976; Collins, Manning, & Taylor, 1984; Collins, Nye, & Manning, 1992; Taylor, VanDeventer, Collins, & Boone, 1983).

Statistical analyses in the prior work were based on correlations, multiple regression, or discriminant analysis. While the FAA enjoyed large samples, capitalization on chance characteristics of a particular sample was a risk. The analyses used a traditional scaling method, in which the response options for an item were treated as interval-type data. For example, response options to items about grades in high school on various subjects were based on letter grades (“A+ to A-,” “B+ to B-,” etc.), where the “A” had a higher value than the “B” range and so forth. The biodata were, then, based on self-reports; as such, they would be vulnerable to misrepresentation by an applicant in operational use.

The veracity of biodata item responses, particularly in high-stakes selection processes, has long been a concern with empirically-keyed scales (Hogan, 2004; Lautenschlager, 1994). And the stakes are certainly high for the FAA’s ATCS selection process. For example, some applicants invest thousands of dollars in tuition and fees to attend 2- and 4-year colleges participating in the FAA’s Air Traffic Control Collegiate Training Initiative (ATC-CTI) with the hope of being hired by the FAA. They seek out and share information about the selection procedures in a number of on-line forums. The applicants are generally highly motivated to become controllers. The pay-off is the prestige of and compensation for the job. The stakes are equally high for the agency. Completion of all field training phases takes an average of 2 to 3 years for someone with no prior experience. False positives (training failures) waste training resources and can impact facility staffing, a critical concern for the agency as the “Post-Strike Generation” of controllers reaches mandatory retirement age. On one hand, given the high stakes, it is reasonable to expect that applicants will attempt to answer questions about life experiences, attitudes, and expectations in what is seen as an employer-desired direction. On the other hand, again given the high stakes, it is just as reasonable for the agency to counter such response sets.

The FAA began exploring an alternative framework for scoring biodata to mitigate both capitalization on chance characteristics and motivated response distortion in the

late 1990s. The alternative framework has two elements: *bootstrapping* (see Efron & Tibshirani, 1993); and *response-option scoring* (Kluger, Reilly, & Russell, 1991). Bootstrapping estimates the sampling distribution of a statistic (e.g., correlation between variables X and Y, or r_{xy}) by iteratively resampling cases from a set of observed data. Russell, Dean, and Broach (2000) demonstrated that the true population bivariate correlation between a predictor and criterion could be accurately estimated via bootstrapping with samples as small as 175-200 persons. The next study investigated response-option scoring. Response-option scoring assigns empirically derived weights to response options, based on the correlation between response option and criterion (Kluger, et al.) Bootstrapping was used to estimate the correlation between the response options to items in the ABA and BQ with a criterion measure. The criterion in that study was a composite of over-the-shoulder ratings used by peers and supervisors to assess typical on-the-job performance (Borman et al., 2001). The result was an 80-item empirically keyed, response-option scored biodata scale ("Controller Background Assessment Survey" or CBAS) (Dean & Broach, 2011). The purpose of this study is to build on the previous work through an investigation of the validity of the 80-item CBAS as a predictor of an objective, computer-based measure of controller technical skill and knowledge. The research question was "What is the incremental validity of CBAS in predicting performance on an objective, computerized measure of air traffic controller technical knowledge and skill?"

METHOD

Sample

Archival data for 229 controllers who participated in the concurrent, criterion-related validation of the AT-SAT test battery in the late 1990s were used in this study. The 229 controllers were a sub-set of 1,232 controllers who participated in that validation study. Full and complete data for the ABA, BQ, and the criterion measure were available for these 229 controllers. Demographic characteristics for the sample are summarized in Table 1. Demographic data on the other 1,003 controllers in the AT-SAT validation study and for FAA Academy first-time entrants between 1981 and 1992 are also summarized in Table 1. The CBAS sample was predominately male (83%) and white (87%). Over half of the sample had at least some college (59%), with more than an additional quarter (29%) reporting an undergraduate college degree. Most (76%) had no prior aviation-related experience as either a pilot or air traffic controller. The sample was similar to the other 1,003 controllers who participated in the AT-SAT validation study and to new hires entering the FAA Academy for the first time between 1981 and 1992 (Table 1).

Measures

AT-SAT. AT-SAT is a computerized test battery designed to assess abilities and other personal characteristics required to perform critical and/or important ATCS job duties. The test battery was developed on the basis of a comprehensive job/task analysis (Nickels, Bobko, Blair, Sands, & Tartak, 1995). AT-SAT was validated in a concurrent, criterion-related validation study in 1997-1998 (Ramos, Heil, & Manning, 2001a, b); the FAA began using the test for ATCS selection in 2002 (King, Manning, & Drechsler, 2007). AT-SAT has eight subtests (Table 2). Scores from the tests are combined into a single overall score; a minimum of 70 is required to be considered eligible for consideration. The test takes about 8 hours to complete. The overall reliability of the composite total score (a weighted linear combination of 22 part scores) was estimated at .74 (Ramos et al., 2001b, p. 41).

Biodata. The development of the empirically-keyed, response-option scored 80-item scale CBAS is described in Dean and Broach (2011). Briefly, each response option for each item was assigned a statistical weight based on its point-biserial correlation with the composite of supervisory ratings of job performance from the AT-SAT validation study. If a given response was selected, it was assigned a value based on the bootstrapped estimate of the point-biserial correlation of that response with the criterion; otherwise a 0 weight was assigned. For example, the response options to a question about average grades earned in high school English classes were A (A- to A+), B (B- to B+), C (C- to C+), and "Less than a C average." The bootstrapped point-biserial correlations of each response option might be .21 for A, .39 for B, .07 for C, and -.12 for "Less than a C average." A participant marking response option A would earn an item score of .21, while a participant marking the "Less than a C average" would earn an item score of -.12. The item scores are then summed across the 80 items and normalized to a mean of 70, a standard deviation of 14, with a minimum of 0 and maximum of 100, to conform to traditional U.S. federal civil service scoring models. The scaled CBAS score distribution is shown in Figure 1; descriptive statistics are presented in Table 3. Scale reliability (Cronbach's α) was estimated as .74 (Dean & Broach, 2011).

Criterion. Two criterion measures were developed in the course of the AT-SAT concurrent validation study: a computer-based measure of situational judgment and a job performance rating. The computer-based performance measure (CBPM) served as the criterion in this incremental validity study. The CBPM was modeled on situational judgment tests. It was designed as an assessment of "the very important technical proficiency part of the controller job that involves separating aircraft" (Hanson et al., 1999, p. 204). Target performance constructs included procedural

Table 1. Demographics

Characteristic	CBAS (n=229)	AT-SAT (n=1,003)	FAA Academy (1981-1992) (n=25,277)
Sex ^a			
Male	189 (83%)	488 (84%)	21,001 (83%)
Female	40 (18%)	90 (16%)	4,276 (17%)
Race			
Minority	28 (13%)	74 (13%)	2,219 (9%)
Non-minority	195 (87%)	497 (87%)	22,529 (91%)
Age ^b			
At Entry-on-Duty	24.8 (2.9)	25.5 (3.4)	27.0 (3.1)
At AT-SAT validation	33 (2.8)	37.9 (6.2)	
Education			
HS/GED	23 (10%)	46 (8%)	2,493 (10%)
Some College	134 (59%)	315 (55%)	12,264 (49%)
Bachelor's Degree or higher	71 (31%)	207 (36%)	10,396 (41%)
Prior Experience			
None	169 (87%)	376 (75%)	16,475 (65%)
Pilot	7 (4%)	44 (9%)	2,490 (10%)
ATC	19 (9%)	84 (17%)	3,495 (16%)

Notes: ^aNumber in group (% of sample)
^bMean (Standard Deviation)

Table 2. Description of the AT-SAT tests

Test	Description
<i>Dials (DI)</i>	Scan and interpret readings from a cluster of analog instruments
<i>Applied Math (AM)</i>	Solve basic math problems as applied to distance, rate, and time
<i>Scan (SC)</i>	Scan dynamic digital displays to detect targets that regularly change
<i>Angles (AN)</i>	Determine the angle of intersecting lines
<i>Letter Factory (LF)</i>	Manage a “factory” with three production lines (with variable speeds) and products, package the products, provide supplies for packaging, and respond to inquiries and interruptions
<i>Air Traffic Scenarios Test (ATST)</i>	Control air traffic in an interactive, dynamic low-fidelity simulation of radar-based air traffic control
<i>Analogies (AY)</i>	Solve verbal and non-verbal analogies
<i>Experience Questionnaire (EQ)</i>	A questionnaire about life experiences relevant to air traffic control

Table 3. Descriptive statistics & correlations (n=229)

Variable ^a	Mean	SD	Min	Max	AT-SAT	CBAS	CBPM
AT-SAT	73.67	7.57	41.79	88.34	.76 ^b		
CBAS	69.98	13.96	22	100	.332***	.74	
CBPM	191.62	13.78	134.61	224.54	.520***	.292***	.63

Notes: ^aAT-SAT=AT-SAT Predictor Composite Score; CBAS=Score on Controller Background Assessment Survey; CBPM=Score on ATCS Computer-Based Performance Measure in AT-SAT validation study

^bScale reliability (Cronbach's α) on diagonal

***p<.001

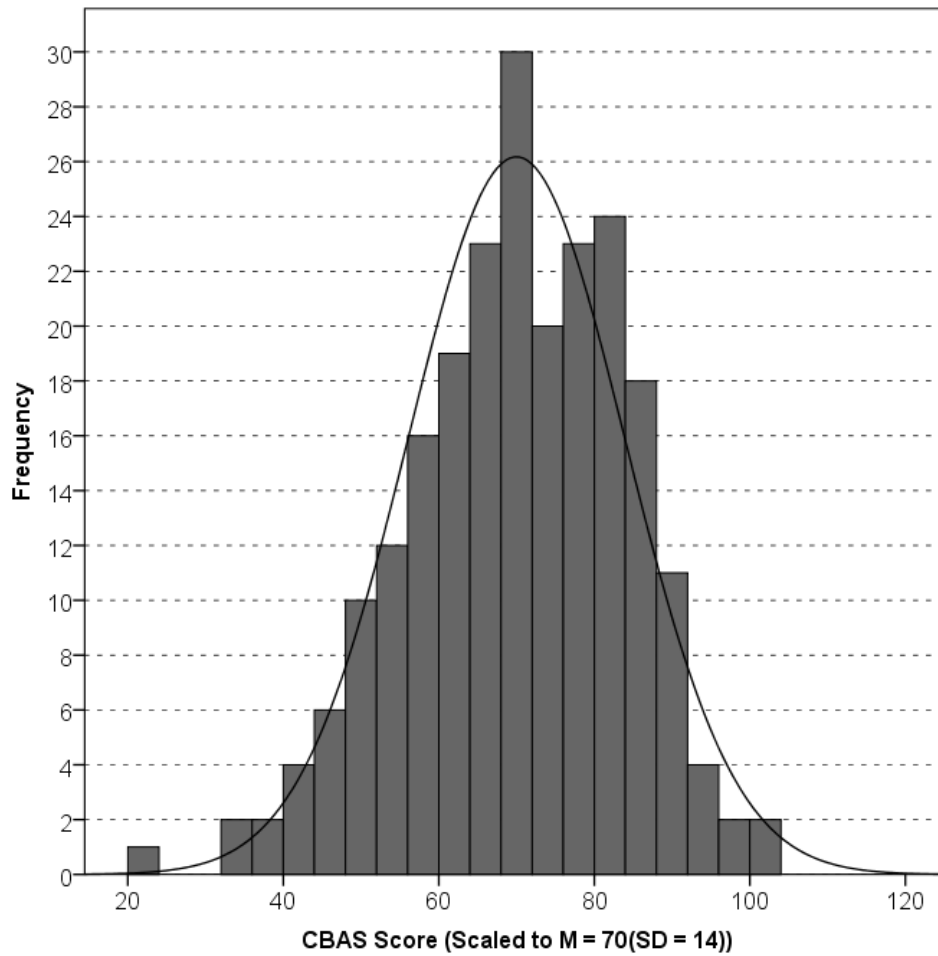
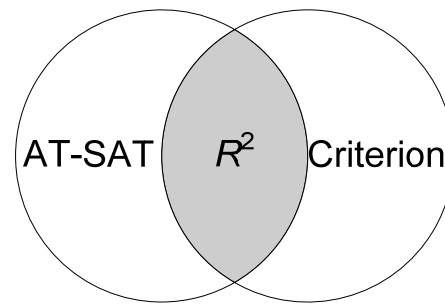


Figure 1. Distribution of CBAS Score ($n=229$)

knowledge about how to perform technical tasks, judgment and decision making, and conflict prediction (Hanson, et al., p. 204). Essentially, a realistic air traffic situation was presented to the participant on a simulated radar screen along with supporting information, such as flight progress strips for aircraft in the scenario, weather, and a map of the synthetic airspace. The controller was given 60 seconds to review the air traffic situation before aircraft began to move on screen, accompanied in some scenarios by pilot communications. The situation unfolded over a few minutes and then stopped. Test items and response options were then presented to the controller. For example, a scenario might involve two aircraft on intersecting flight

paths. The participant might be asked, “What control action should be taken to ensure separation between flights ABC123 and XYZ987?” Response options represented specific air traffic control actions such as (a) giving a speed control instruction to ABC123, (b) a change in heading for XYZ987, (c) an instruction to ABC123 to climb to a higher altitude, or (d) an instruction to XYZ987 to descend to a lower altitude. The participant had 25 seconds in which to select a response. Then a new scenario was presented on screen. There were 29 scenarios total in the CBPM, accompanied by 84 test items, with an internal consistency estimate of $\alpha=.63$ (Ramos et al., 2001b, Table 4-9, p. 88). For some items, such as conflict detection and

Step 1: Regress AT-SAT on the criterion
 REGRESSION
 /VARIABLES = ...
 /STATISTICS = DEFAULTS CHA ...
 /DEPENDENT = CRITERION
 /ENTER = AT-SAT



Step 2: Enter CBAS to estimate ΔR^2
 /ENTER = BIODATA

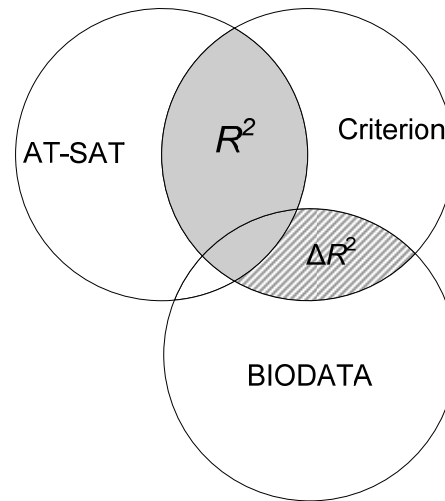


Figure 2. Incremental validity analysis (with SPSS syntax)

avoidance, there was a single correct answer; for others, the responses had been ranked by subject matter experts (all controllers) in terms of their effectiveness.

Procedure

A hierarchical regression analysis was conducted to assess the incremental validity of CBAS. In a hierarchical regression analysis, the variance in the dependent variable can be uniquely partitioned based on the order in which the (correlated) independent variables are entered (Cohen, Cohen, West, & Aiken, 2003, p. 158). In this instance, the first variable entered was the AT-SAT score. The standardized regression weight (β) for the

independent variable entered in the first step is equal to its zero-order correlation with the criterion. CBAS scores were entered in the second step to estimate the effect of the second independent variable, taking into account the first predictor and the correlation between the two predictors (see Cohen et al. p. 67). The critical question in incremental validity is the additional variance (change in the multiple correlation coefficient, or ΔR^2) in the criterion explained by the additional test score (Hunsley & Meyer, 2003; see Figure 2). No corrections were made for predictor incidental restriction in range or unreliability.¹ SPSS Version 19 was used for all statistical analyses.

¹Incidental (or indirect) restriction of range is the reduction in variance on the alternative predictor due to explicit selection of the sample on the current predictor. Unreliability is the degree of unsystematic variance in scores due to errors of measurement, as opposed to reliability (Ghiselli, Campbell, & Zedeck, 1981; Guion, 2011).

RESULTS

Descriptive statistics for and uncorrected correlations between the three measures are presented in Table 3. Results of the hierarchical regression are presented in Table 4. AT-SAT scores accounted for about 27% of variance in the CBPM criterion measure ($\beta=0.520$, adjusted $R^2=.271$, $p<.001$). This is consistent with results from the validation of AT-SAT ($r_{xy}=.52$, Ramos et al., 2001b, Table 5.5.1, p. 120). The biodata scale accounted for an additional 2% of the variance in CBPM scores, without corrections ($\beta=0.134$; $\Delta R^2=0.016$, $\Delta F=5.040$, $p<.05$).

DISCUSSION

The empirically-keyed, response-option scored biodata scale demonstrated incremental validity over the computerized air traffic controller aptitude test battery in predicting scores representing the technical knowledge and skills of en route controllers. In other words, after taking AT-SAT into account, CBAS accounted for just a bit more of the variance in the criterion measure. While an additional 2% seems small in an absolute sense, that increment can have substantial utility in high-stakes selection such as that for air traffic controllers. For example, after taking into account aptitude (as measured by the written aptitude test used between 1981 and 1992 (Broach, 1998)), personality explained an additional 6-9% of variance in performance in the FAA Academy Screen (Schroeder, Broach, & Young, 1993). The utility analysis, based on the Taylor-Russell tables (Taylor & Russell, 1939), demonstrated that incorporation of a "Big Five"

personality test into the aptitude testing in use at the time would have increased the FAA Academy Screen pass rate by about 3% (from 55% to about 58%). The avoided "lost" costs were estimated at about \$600,000 per year.

A similar logic was used to estimate the potential utility of adding biodata to the current controller selection process. Utility, in this analysis, is defined as the change (increase or decrease) in the proportion of new controllers successfully completing their training and the avoided (or additional) costs associated with attrition. The expected utility of the selection procedure depends on (a) the base rate of satisfactory performance, (b) the validity of the selection procedure, and (c) how it is implemented.

The FAA plans to hire between 800 and 1,100 new controllers per year between now and 2020 (FAA, 2011). About 20% will be lost in training (estimated from FAA, 2011, Figure 4.10, p. 35 and Figure 5.1, p. 37). The base rate of "satisfactory performance" (in terms of completing training), given selection on the basis of AT-SAT scores, then, is 80%. The proposed biodata instrument (CBAS) could be implemented either (a) before AT-SAT is administered, as part of the initial on-line application process, or (b) as part of AT-SAT. Validity is the zero-order correlation between the CBAS score and the job performance criterion in the first implementation scenario (.29 in Table 3, rounded to .3 for the utility analysis). If CBAS were added to the AT-SAT composite score (Scenario 2), validity is the CBAS standardized regression coefficient in the incremental validity analysis ($\beta_{CBAS}=.13$ [Table 4], rounded to .1). Finally, hiring rates from very lenient (90% selected) to very strict (just 10% selected) were considered. The Taylor-Russell (Taylor & Russell,

Table 3. Descriptive statistics & correlations ($n=229$)

Variable ^a	Mean	SD	Min	Max	AT-SAT	CBAS	CBPM
AT-SAT	73.67	7.57	41.79	88.34	.76 ^b		
CBAS	69.98	13.96	22	100	.332***	.74	
CBPM	191.62	13.78	134.61	224.54	.520***	.292***	.63

Notes: ^aAT-SAT=AT-SAT Predictor Composite Score; CBAS=Score on Controller Background Assessment Survey; CBPM=Score on ATCS Computer-Based Performance Measure in AT-SAT validation study

^bScale reliability (Cronbach's α) on diagonal

*** $p<.001$

Table 4. Hierarchical regression analyses

Step	Predictor ^a	β	t	R	R^2	Adj- R^2	ΔR^2	F Change
1	AT-SAT	.520	9.183***	.520	.271	.268	.271	84.333***
2	AT-SAT	.476	7.993***	.536	.287	.280	.016	5.040*
2	CBAS	.134	2.245*					

Notes: ^aAT-SAT=AT-SAT Predictor Composite Score as computed in Ramos, Heil, & Manning (2001a, 2001b); CBAS=Controller Background Assessment Survey score

* $p<.05$, *** $p<.001$

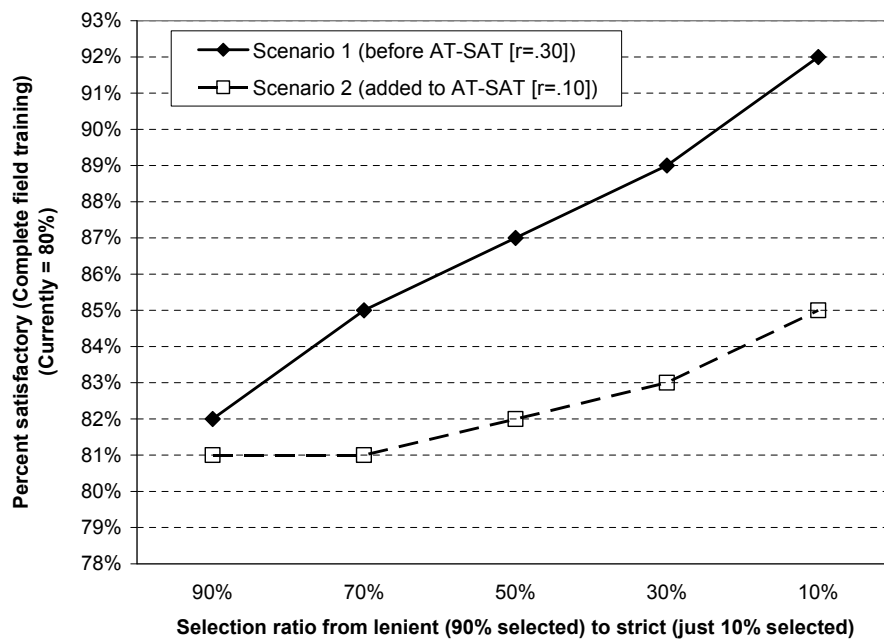


Figure 3. Gain in proportion of new hires successfully completing ATCS training as a function of selection rate (from lenient to strict) and CBAS validity for two implementation scenarios (before AT-SAT or as part of AT-SAT)

1939) tables were used to estimate the proportion of new controllers that would be successful in each of the two implementation scenarios.

The results for the utility analysis are illustrated in Figure 3 for the two implementation scenarios. If biodata were made part of the on-line application process (Scenario 1), the expected proportion of new hires completing field training could increase from 80% to 82% at a lenient selection rate (90% selected). At more stringent selection rates, the expected proportion successful increases to as much as 92% in the first implementation scenario. Scenario 2 (adding CBAS to AT-SAT) is more conservative. With a lenient selection rate (80-90% selected), the proportion of successes increases very modestly to just 81%. Even at the strict selection rate (just 10% selected), the proportion of successes increases to just 85% in the second implementation scenario. However these are statistical predictions given the assumptions of the Taylor-Russell approach. The likely gains in practice are more likely to be closer to those of the second implementation scenario.

But even very modest gains in the proportion successful in field training can have substantial financial implications in terms of avoided costs. FAA plans to hire 800-1,000 new controllers per year for the next several years in its annual controller workforce plan (FAA, 2011). In that same plan, FAA estimated the personnel compensation and benefits cost of each trainee at about \$93,000 per developmental (p. 47). Increasing the field training success rate by just 1% (8-10 controllers at \$93K each per year) equates to about \$744,000 per year in avoided direct

costs of failures, using the more conservative estimate assumptions (lower validity [.1], higher selection rate [.9]) of the second implementation scenario.

Net utility to the agency would be determined by the actual benefits (increase in completions and avoided costs attributable to better selection) and costs for implementation and evaluation of biodata as part of the controller selection process. Further research is needed to evaluate the validity and utility of CBAS by linking it with actual training outcomes in the FAA Academy and field facilities, to better quantify costs and benefits. The FAA hired 6,484 new controllers in 2007-2010 according to the FAA in the annual controller workforce plans submitted to the U.S. Congress (FAA, 2011). In those same reports, FAA reported that the average training time was about 2 years for controllers assigned to terminal facilities and about 3 years for en route controllers. A substantial proportion of the controllers hired in 2007 and 2008 should have completed (or not) training at the first assigned field facility. Such data could provide an empirical basis for determining if and how CBAS might be operationally implemented and the likely benefits to the FAA.

A final consideration is the degree to which the response-option scoring key developed on the Post-Strike Generation can be applied to response data from the Next Generation of controllers. Comparison of response data from both generations, as well as the validity analyses suggested above, will provide an empirical basis for assessing the stability and validity of the response-option scoring key.

REFERENCES

- Asher, J.J. (1972). The biographical item: Can it be improved? *Personnel Psychology*, 25, 251–269.
- Borman, W.C., Hedge, J.W., Hanson, M.A., Bruskiwicz, K.T., Mogilka, H., Manning, C., Bunch, L.B., & Horgen, K.E. (2001). Development of criterion measures of air traffic controller performance. In R.A. Ramos, M.A. Heil, & C.A. Manning (Eds.) *Documentation of validity for the AT-SAT computerized test battery: Volume II* (pp. 1–12) (FAA Report No. DOT/FAA/AM-01/6). Washington, DC: Federal Aviation Administration, Office of Aviation Medicine.
- Broach, D. (1992, June). *Non-cognitive predictors of performance in radar-based air traffic control training*. Paper presented at the 4th Annual Convention of the American Psychological Society, San Diego, CA.
- Broach, D. (1998). Air traffic control specialist aptitude testing, 1981-1992. In D. Broach (Ed.), *Recovery of the air traffic control specialist workforce, 1981-1992* (pp. 7-16) (FAA Report No. DOT/FAA/AM-98/23). Washington, DC: Federal Aviation Administration, Office of Aviation Medicine.
- Broach, D. (2008). *Overview of CAMI Life Experiences Questionnaire (Version 1.0)*. Unpublished manuscript, FAA Civil Aerospace Medical Institute Aerospace Human Factors Research Division (AAM-500).
- Brokaw, L.D. (1957). *Selection measures for air traffic control training* (USAF Report No. PL-TM-57-14). Lackland AFB, TX: United States Air Force Personnel & Training Research Center Personnel Laboratory.
- Cobb, B.B. & Nelson, P.L. (1974). *Aircraft-pilot and other pre-employment experience as factors in the selection of air traffic controller trainees* (FAA Report No. DOT/FAA/AM-74/8). Washington, DC: Federal Aviation Administration, Office of Aviation Medicine.
- Cobb, B.B., Young, C.L., & Rizutti, B.L. (1976). *Education as a factor in the selection of air traffic controller trainees* (FAA Report No. DOT/FAA/AM-76/16). Washington, DC: Federal Aviation Administration, Office of Aviation Medicine.
- Cohen, J., Cohen, P., West, S.G., & Aiken, L.S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences (3rd Ed.)*. Hillsdale, NJ: Erlbaum.
- Collins, W.E., Manning, C.A., & Taylor, D.K. (1984). A comparison of prestrike and poststrike ATCS: Biographic factors associated with Academy training success. In A.D. VanDeventer, W.E. Collins, C.A. Manning, D.K. Taylor, & N.E. Baxter (Eds.), *Studies of poststrike air traffic control specialist trainees: I. Age, biographical factors, and selection test performance related to Academy training success* (pp. 7–14) (FAA Report No. DOT/FAA/AM- 84/6). Washington, DC: Federal Aviation Administration, Office of Aviation Medicine.
- Collins, W.E., Nye, L.G., & Manning, C.A. (1992). Post-strike air traffic control trainees: Biodemographic predictors of success in selection and screening. *International Journal of Aviation Psychology*, 2, 213–223.
- Dean, M.A. (1999, June). *A response-option examination of biodata adverse impact and criterion validity*. Presented at the 1999 International Personnel Management Association Assessment Council meeting, St. Petersburg, FL.
- Dean M.A. & Broach, D. (In press). *Development, validation, and fairness of a biographical data questionnaire for the Air Traffic Control Specialist (ATCS) occupation*. Manuscript submitted for publication as an Office of Aerospace Medicine technical report.
- Efron, B. & Tibshirani R.J. (1993). *An Introduction to the bootstrap*. New York: Chapman & Hall.
- Farmer, W.L. (2002). *Characteristics of biodata keys as a function of scaling method, sample size, and criterion*. Dissertation Abstracts International: Section B: The Sciences and Engineering, 63(3-B).
- Federal Aviation Administration. (2011, April). *A plan for the future: 10-Year strategy for the air traffic control workforce 2011–2020*. Washington, DC: Federal Aviation Administration.
- Ghiselli, E.E., Campbell, J.P., & Zedeck, S. (1981). *Measurement theory for the behavioral sciences*. San Francisco, CA: W. H. Freeman & Co.
- Guion, R.M. (2011). *Assessment, measurement, and prediction for personnel decisions* (2nd Ed.). New York, NY: Routledge Academic.
- Hanson, M.A., Borman, W.C., Mogilka, H.J., Manning, C., & Hedge, J.W. (1999). Computerized assessment of skill for a highly technical job. In F. Drasgow & J.B. Olson-Buchanan (Eds.), *Innovations in computerized assessment* (pp.197–220). Mahwah, NJ: Erlbaum.

- Hogan, J.B. (2004). Empirical keying of background data measures. In G.S. Stokes, M.D. Mumford, & W.A. Owens (Eds.), *Biodata handbook: Theory, research, and use of biographical information in selection and performance prediction* (pp. 69-107). Palo Alto, CA: CPP Books.
- Hunsley, J. & Meyer, G.J. (2003). The incremental validity of psychological testing and assessment: Conceptual, methodological, and statistical issues. *Psychological Assessment*, 15, 446-455.
- Hunter, J.E. & Hunter, R.F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, 96, 72-98.
- King, R.E., Manning, C.A., & Drechsler, G.K. (2007). *Operational use of the Air Traffic Selection and Training battery* (FAA Report No. DOT/FAA/AM-07/14). Washington, DC: Federal Aviation Administration, Office of Aerospace Medicine.
- Kluger, A.N., Reilly, R.R., & Russell, C.J. (1991). Faking biodata tests: Are option-keyed instruments more resistant? *Journal of Applied Psychology*, 76, 889-896.
- Lautenschlager, G.J. (1994). Accuracy and faking of background data. In G.S. Stokes, M.D. Mumford, & W.A. Owens, (Eds.), *Biodata handbook: Theory, research and use of biographical information in selection and performance prediction* (pp. 391-419). Palo Alto, CA: Consulting Psychologists Press.
- Mumford, M.D. & Owens W.A. (1987). Methodological review: Principles, procedures, and findings in the application of background data measures. *Applied Psychological Measurement*, 11, 1-31.
- Nickels, B.J. (1994). The nature of biodata. In G.S. Stokes, M.D. Mumford, & W.A. Owens, (Eds.), *Biodata handbook: Theory, research and use of biographical information in selection and performance prediction* (pp. 1-16). Palo Alto, CA: Consulting Psychologists Press.
- 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 Management.
- Ramos, R.A., Heil, M.A., & Manning, C.A. (2001a). *Documentation of validity for the AT-SAT computerized test battery: Volume I* (FAA Report No. DOT/FAA/AM-01/5). Washington, DC: Federal Aviation Administration, Office of Aviation Medicine.
- Ramos, R.A., Heil, M.A., & Manning, C.A. (2001b). *Documentation of validity for the AT-SAT computerized test battery, Volume II* (FAA Report No. DOT/FAA/AM-01/6). Washington, DC: Federal Aviation Administration Office of Aviation Medicine.
- Reilly, R.R. & Chao, G.T. (1982). Validity and fairness of some alternative employee selection procedures. *Personnel Psychology*, 35, 1-62.
- Russell, C.J., Dean, M.A., & Broach, D. (2000). *Guidelines for bootstrapping validity coefficients in ATCS selection* (FAA Report No. DOT/FAA/AM-00/15). Washington, DC: Federal Aviation Administration, Office of Aviation Medicine.
- Schmitt, N., Gooding, R.Z., Noe, R.A., & Kirsch, M. (1984). Meta-analyses of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology*, 37, 407-422.
- Schroeder, D.J., Broach, D., & Young, W.C. (1993). *Contribution of personality to the prediction of success in initial Air Traffic Control Specialist training* (FAA Report No. DOT/FAA/AM-93/4). Washington, DC: Federal Aviation Administration Office Aviation Medicine.
- Stokes, G.S., Mumford, M.D., & Owens, W.A. (1994). *Biodata handbook: Theory, research and use of biographical information in selection and performance prediction*. Palo Alto, CA: Consulting Psychologists Press.
- Taylor, H.C. & Russell, J.T. (1939). The relationship of validity coefficients to the practical effectiveness of tests in selection: Discussion and tables. *Journal of Applied Psychology*, 23, 565-578.
- Taylor, D.K., VanDeventer, A.D., Collins, W.E., & Boone, J.O. (1983). Some biographical factors associated with success of air traffic control specialist trainees at the FAA Academy during 1980. In A.D. VanDeventer, D.K. Taylor, W.E. Collins, & J.O. Boone (Eds.), *Three studies of biographical factors associated with success in air traffic control specialist screening/training at the FAA Academy* (pp. 6-11) (FAA Report No. DOT/FAA/AM-83/6). Washington, DC: Federal Aviation Administration Office of Aviation Medicine.

