PORTATION

UNIVERSITY OF MINNESOTA

DEVELOPMENT OF AN ABILITY-BASED SELECTION INSTRUMENT FOR THE MARC-AIR TRAFFIC CONTROLLER TRAINING PROGRAM

Phillip Ackerman and
Ruth Kanfer
Psychology

Sponsored by:
Center for Transportation Studies

Development of a Ability-based Selection Instrument for the MARC-Air Traffic Controller Training Program

Final Report

to
Minnesota Technical College System
Minnesota Air Traffic Control Training Center
Flying Cloud Airport
10100 Flying Cloud Drive
Eden Prairie, MN 55347-4016

Submitted by

Phillip L. Ackerman
Ruth Kanfer

Department of Psychology
University of Minnesota
N218 Elliott Hall
75 E. River Rd.
Minneapolis, MN 55455

February, 1993

The financial support of oil overcharge funds distributed through the Minnesota Department of Administration is acknowledged, but the authors assume complete responsibility for the contents herein.

Executive Summary

This report reviews the completed task and deliverable components of the contract between the Minnesota Air Traffic Control Training Center (MATCTC) and the University of Minnesota, and describes the findings of our research concerning the implementation of a selection instrument for air traffic control (ATC) trainees. The report reviews the basic research plan, describes the results of test development and validation activities (using the University of Minnesota laboratory sample and the FAA student sample), and describes in detail the results of test development and validation for the MATCTC student samples. References are provided for previously submitted reports that detail subcomponents of the research program.

The final test battery delivered to MATCTC in December, 1992, had the following characteristics:

(1) Aptitude Assessment Battery (AAB) and self-report instrument (each of which provide a single summary score). Administration time of 115 minutes (1 hour, 55 minutes).

(2) Concurrent validity index for Aptitude Assessment Battery composite of r = .63, (with MATCTC Cumulative Grade Point Average) computed on N = 100 students who completed MATCTC program.

(3) Significant differences on the AAB composite between students identified as "academic drops" and whole sample ($\underline{M}_{\text{academic drop}}$ = 40.31, $\underline{M}_{\text{other drop}}$ 47.99, t = 2.95, $p < .01$). Significant differences between students identified as "academic drops" and those who dropped for unspecified or academically unrelated reasons ($\underline{M}_{\text{completing students}}$ 50.35, t = 4.65, $p < .001$).

(4) AAB composite correlated with OPM Test r = .45. [OPM test only correlated r = .46 with MATCTC Cumulative Grade Point Average] **The AAB is a significantly more valid predictor of MATCTC GPA than the OPM test (t = 2.15, $p < .02$).**

Overview

The broad goal of this project was to develop an Aptitude Assessment Battery (AAB) that satisfies the requirements of a valid predictor of air traffic controller success, in the context of the Minnesota Air Traffic Control Training Center (MATCTC) program. "Construct" validity was to be demonstrated with a laboratory study, in which scores on the AAB could be compared to extant off-the-shelf ability measures. "Criterion-related" validity was to be demonstrated both in the laboratory (with performance on a computerized air traffic control simulation program -- TRACON as the criterion) and in the field (with a concurrent job incumbent sample and success/failure in the Federal Aviation Administration "ATC Screen" selection/training program as the criterion). Additional validation indices were to be found with the field sample, by comparing the AAB with performance on the FAA's current selection instrument (designated as the "OPM Test"). Final validational checks for the specific application to MATCTC students were to be made with an aggregation of several classes of MATCTC students, with cumulative grade point average (GPA) as the criterion. As far as was feasible, the project also set out to evaluate gender and race differences in AAB performance and MATCTC GPA.

Test Development and Initial Validation

Initial design of the AAB followed review of the job and task analysis literature, and review of the predictor space. A battery of tests and a self-report measure were designed for tryout, along with reference measures (for construct validity). A two-stage process was used; the first stage of development and validation of the selection battery used a laboratory-based ATC simulation environment, and the second step was a cross-validation of the battery in a field-based study. Each of these steps will be discussed in turn.

Job/Task Analysis

A substantial job analysis literature exists for FAA ATCs (e.g., Ammerman, et al., 1987). Our first step was to review and summarize these materials (Landon, 1991). Next, we conducted a series of critical-incidents elicitations with ATC Instructors (Landon, 1992), and conducted observations of ATC trainees, and full-performance-level controllers. A list of critical job components was derived to describe the criterion space.

Ability Battery Development

Development of a new ability battery was a multi-step, iterative process. A review of the selection literature was performed (e.g., see Sells, et al., 1984, for an extensive collection of research on ATC selection). Next, discussions with experts in ability theory, ability testing, and individual differences in learning/skill acquisition were conducted for a review of the predictor space. From these steps, we created tests for tryout in predicting individual differences in ATC performance. Abilities under consideration were as follows: Reasoning; Spatial Visualization; Numerical Abilities, Spatial Memory, Spatial Time Estimation.

Additional tests were selected or created to provide convergent and discriminant construct validity for these new tests, and for evaluation of incremental criterion validity. Based on Ackerman's (1988) theory of the cognitive determinants of individual differences in skilled performance, and the previous study of skill acquisition with the TRACON simulation (Ackerman, 1992), abilities associated with the first two stages of skill acquisition were included: General/Spatial Reasoning and Perceptual Speed, for the first and second stage of skill acquisition, respectively.

Two additional tests were adapted to provide reference to earlier studies of selection of ATCs, both of which involve speeded perceptual decision making and spatial memory: namely the Directional Headings Test (developed by the FAA - Civil Aeromedical Institute, Cobb & Mathews, 1972), and the Dial Reading Test (developed by the U.S. Army Air Force Aviation Psychology Research Program -- see Guilford and Lacey, 1947). Given that these two tests had been previously demonstrated to have high validity for predicting ATC training success, they provided an additional benchmark for showing conceptual overlap between the simulator ATC environment and the actual ATC environment.

Questionnaire Battery Development

The broader scope of the current research program (Kanfer & Ackerman, 1989) was concerned with the use of conative and self-regulatory training interventions, that may be ultimately used to increase the proportion of trainees who succeed in the high-pressure ATC training environment. Of particular interest here, was the potential use of a self-efficacy measure for prediction of training success. The particular measure was based on a body of research concerning self-regulatory determinants of task performance (e.g., for discussion see Kanfer, 1990, 1991; Kanfer & Ackerman, 1989). This measure concerns self-reported confidence in high-pressure studying, learning, and testing situations. The hypothesis was that self-reports of confidence in attentionally demanding situations would provide an incremental validity (to the ability measures) in predicting complex skill acquisition performance (for reviews, see Bandura, 1986; Kanfer, 1991).

Laboratory and field development work

The combined laboratory and field development work on the AAB has been extensively described in a document previously submitted to MATCTC, which is soon to be published in the *Journal of Applied Psychology* (Ackerman & Kanfer, in press). That document outlines the two studies that were conducted to converge on the development of the AAB for selection of air traffic controllers. The laboratory study (N = 112) used an air traffic control simulator (TRACON®) for initial development and evaluation of the selection battery. The field study (N = 206) of FAA air traffic controller trainees provided cross-validation data as a precursor to implementation. Implications for development of ability-based and self-efficacy-based selection measures for complex job performance are discussed, as are general issues for new selection research and application.

The overall results of those two studies were consistent with the Ackerman (1988) theory of the cognitive determinants of individual differences in skilled performance. That is, given the prior designation of the overall ATC task/job as requiring substantial controlled information processing resources, general and broad content (e.g., Spatial) abilities would be highly predictive of overall performance, especially in comparison to the perceptual speed abilities (which are more highly associated with practiced performance on consistent information processing tasks). The individual correlations between the ability tests and task/job criteria clearly indicated that the broad spatial and reasoning tests had the highest validity for the global performance measures, and the perceptual speed tests had much lower validity indices.

The measures designed to predict ATC performance represent significant progress in ability research. Although some traditionally-inspired measures were used (e.g., Verbal Analogy and Paper Folding), new tests that were derived from current theory of spatial abilities (Lohman, 1987), for example the Verbal Test of Spatial Ability, while having some roots in earlier psychometric work (e.g., Hoffman, Guilford, Hoepfner, & Doherty, 1968), and the Spatial Orientation test show that it is possible to break new ground in seeking applications of ability theory to selection.

The combination of laboratory and field investigations contributed to both the research and application domains. From an applied perspective, we developed construct and criterion-related validity for the AAB in predicting air traffic control training success. Many of the new tests were further cross-validated in the field, providing a direct demonstration of the utility of the test battery. Moreover, the new test battery provided a greater validity than the current test used in the field, the OPM measure ($\underline{r}$ = .515 for the OPM measure vs. $\underline{r}$ = .558 for the new paper and pencil ability test battery), with shorter time demands (more than 150 min for the OPM measure vs. 90.5 min for the new paper and pencil battery, a reduction of 40% in testing time). In addition to the benefits of the current ability-theory approach, though, part of the increase in validity is probably due to the greater bandwidth of the test battery, in comparison to the current OPM composite. When the self-efficacy measure was added, the validity of the new battery increased to $\underline{r}$ = .591.

MATCTC Validation

Concurrent with the laboratory and FAA development work, data were collected from MATCTC students on the AAB. The data collection, analysis, and results of this phase of the project are described in the sections below.

Description of samples

Versions of the current AAB and Self-Report Questionnaire have been administered to

current students in MATCTC classes starting in December of 1991.[1] [Class #2 = 30 students, Class #3 = 32 students, Class #4 = 32 students, Class #5 = 32 students, and Class #6 = 22 students.] As of the end of the contract period, data had been obtained from a total of 148 students. Of this total, 31 students have failed to complete the program (16 designated as "Academic Drop" and 15 designated as "Other Drop"]. In addition, 17 students were currently in the MATCTC program, and as such have not yielded final cumulative GPA data. Thus, cumulative GPA data were available on 148 - 31 - 17 = 100 students.

Description of test evolution

The tryout of the AAB with MATCTC students (Classes #2 - 5) included the full array of paper and pencil tests, questionnaires administered to the FAA sample, most of the tests administered to the laboratory sample, and some of the computerized tests administered only in the laboratory study (namely, Time-Estimation, Spatial Memory, and Choice Reaction Time). Subsequent to Class #5, and in consideration of time, logistic, and test validity considerations, the test battery and questionnaire set were shortened to final form. (Specifically eliminated were the computerized measures; ability tests with low validity indices: Flight Checking, Letter/Number Substitution, Cancelling A's, and Verbal Analogy; and questionnaire items of low validity indices). The final battery can be administered in just under two hours, entirely with paper and pencil. The final battery includes the following measures (with administration times noted):

Tests in Administration Order	**Time**	**Cumulative Time**
1. Necessary Facts	11.5min	11.5min
2. Spatial Orientation	7.0	18.5
3. Math Knowledge	12.5	31.0
4. Spatial Analogy	12.0	43.0
5. Problem Solving	6.0	49.0
6. Paper Folding	14.0	63.0
7. Verbal Test of Spatial Ability	14.0	77.0
8. Dial Reading	20.5	97.5
9. Directional Headings	12.0	109.5
10. Questionnaire	5.0	114.5

[1] As a means toward getting an early assessment of the level of aptitude of MATCTC students, several off-the-shelf aptitude tests were administered to the first MATCTC graduating class (the AAB had not been created at that point). Given the lack of any overlapping tests, data from this class are not considered in this report.

In order to provide a common metric for the validation efforts reported below, *only the composite scores from the final test list are considered.* Therefore, even though AAB composites computed on early cohorts included additional tests, all composites have been recalculated based on the tests listed above. The AAB aptitude composite is a unit-weighted z-score sum of all of the 9 aptitude tests listed above. The self-confidence composite is a total of raw scores on the 18 items on the Questionnaire listed above. For ease of interpretation, composite scores reported to MATCTC are converted to T-Scores (mean of 50, standard deviation of 10), and are provided with corresponding percentile ranks.

Validation data

For the 100 students with cumulative GPA scores available, the AAB composite correlated $r = .63$ with GPA, which is both statistically significant and functionally important. In variance terms, the 2-hour AAB provides a single assessment of aptitude which accounts for 40% of the variance in cumulative GPA over a six-month period at MATCTC.

Moreover, the AAB provides a superior predictive validity, when compared with the performance of the OPM test. The OPM test correlates $r = .46$ with MATCTC cumulative GPA, accounting for only 21% of the variance, nearly half as much as the AAB. [This difference is statistically significant ($t = 2.15$, $p < .02$).

When considering the validity indices above, additional information from students who did not attrit from the MATCTC program bolsters the statement of the utility of the AAB composite. From academic records and exit interviews, MATCTC staff designated students who attritted as either "Academic" or "Other." Although a designation of "other" may include students with academic problems, a designation of "academic" attrition was decisive. Although these individuals did not provide overall GPA's to allow inclusion for the correlational evidence, AAB composite scores among attritting and completing students provide additional information regarding the validity of the AAB. Specifically, if the AAB is a valid predictor of performance, "Academic" drops should have lower AAB scores than "Other" drops.

For the 16 students identified as "Academic" drops, performance on the AAB was as follows: $M = 40.31$, $sd = 6.18$. For the 15 students identified as "Other" drops, performance was as follows: $M = 47.99$, $sd = 8.22$. Although the nearly one standard deviation difference between the two groups shows an obvious distinction, a statistical test confirms the significance of the difference ($t = 2.95$, $p < .001$). A similar comparison to the completing students ($M = 50.35$, $sd = 8.63$), reveals analogous results ($t = 4.65$, $p < .001$), a difference in 1.16 *sd* units. Comparisons between "Other" drops and completing students revealed no significant differences (not surprising given the 0.27 *sd* unit difference between the two groups). In lay terms, the "Academic" drop students performed substantially more poorly on the AAB than the students who completed the program, or left

for other, unspecified reasons. *If the "Academic" students were allowed to continue in the MATCTC program, it is reasonable to expect that the overall correlation between the AAB and GPA would be increased.*

Gender/Race Comparisons

For overall validation purposes, gender and race differences are important considerations. We have endeavored to provide as much information as possible regarding these issues, however, given the relatively small sample sizes, it is impossible at this point to provide any decisive evidence regarding differences in regression estimates for the various subgroups to be evaluated. Below is the data collected to date regarding differences between gender and race groups.

Gender. A comparison across genders revealed no statistically significant (or meaningful) overall differences. For the AAB composite, the following results were obtained: $M_{\text{males}} = 50.43$, $sd = 8.40$; $M_{\text{females}} = 51.16$, $sd = 8.27$. The statistical test found that the extremely small advantage to females (less than 0.1 *sd* units) did not result in any significant differences ($t = .43$). Note, however, that females also had a small advantage in GPA $M_{\text{males}} = 3.31$, $sd = .36$; $M_{\text{females}} = 3.35$, $sd = .43$, again not reaching statistical significance ($t = .43$).

Race. Comparisons across racial groups were inconclusive, given the small number of students in each of the identified categories. For comparison purposes, mean AAB composite scores, GPA's, and total number of students in each category are presented below.

Group	AAB	GPA	Cases
White	51.36	3.36	89/86
African American	42.65	2.93	7
Hispanic American	63.26	3.75	2
Asian American	44.31	3.06	4
Native American	53.07	3.06	1

Again, while any comparison would be tentative at this point, it appears that AAB composite scores and GPA values show comparable differences among the various groups. Additional students will need to be tested before any conclusion about such differences can be made.

Self-Confidence Ratings

In our initial investigations (both in the laboratory and in the field), a measure of self-confidence (or self-efficacy) was a valid predictor of performance. We expected a similar result in the MATCTC sample. However, data collected to date failed to show a statistically significant relationship between the measure of self-confidence and cumulative GPA (r = .16, p = .12). Similarly, no significant differences were found between "Academic" drops and "Other" drops (M_{Academic} = 50.85, sd = 8.71; M_{Other} = 49.90, sd = 10.42). There are many potential explanations for this finding -- probably the most prominent one is the more nurturing environment provided by the MATCTC curriculum and staff vis à vis the laboratory and FAA Screen environments. However, we wish to withhold judgment on the utility of this measure until post-graduate data on MATCTC students regarding FAA developmental performance become available. We expect that the high-pressure environment on-the-job may yield significant validity coefficients for the self-confidence measure against training/field-attrition criteria.

General Conclusions

As can be seen from the various results concerning the AAB, we consider the overall project to be one of great success. The AAB test battery is easy to administer (requiring under 2 hours of administration time), and is a valid predictor of MATCTC performance, accounting for 40% of the variance in cumulative GPA scores. Moreover, the AAB battery is superior in every comparison (e.g., administration time and validity) with the OPM test battery. In addition, the test battery is sufficiently general to well-predict acquisition of skills on a computerized simulation of ATC skills (TRACON), and of performance in the FAA Screen (see Ackerman & Kanfer, in press, for further details).

Follow-up

We see several sources of potential follow-up to the current project. Test validation, in general, is an ongoing process. As additional experience is obtained, with test administration and data collection, estimates of the validity of the test can be made with increasing precision. Furthermore, if changes are made to the curriculum, a reevaluation of the test battery should be in order. Finally, it will be ultimately desirable to obtain FAA records to evaluate the long-term validity of the test battery, when compared with on-the-job performance appraisals.

References

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. Journal of Experimental Psychology: General, 117, 288-318.

Ackerman, P. L. (1992). Predicting individual differences in complex skill acquisition: Dynamics of ability determinants. Journal of Applied Psychology, 77, 598-614.

Ackerman, P. L., & Kanfer, R. (In press, June, 1993). Integrating laboratory and field study for improving selection: Development of a battery for predicting air traffic controller success. To appear in Journal of Applied Psychology.

Ammerman, H. L., Becker, E. S., Jones, G. W., Tobey, W. K., & Phillips, M. K. (1987). FAA air traffic control operations concepts volume I: ATC background and analysis methodology. Washington D.C.: FAA report no. DOT/FAA/AP-87-01.

Bandura, A. (1986). Self-regulation of motivation and action through goal systems. In V. Hamilton, G H., Bower, & N. H. Fryda (Eds.), Cognition, motivation, and affect: A cognitive science view (pp. 37-61). Dordrecht, Holland: Martinus Nijhoff.

Cobb, B. B., & Mathews, J. J. (1972). A proposed new test for aptitude screening of air traffic controller applicants (FAA-AM-72-18). Washington, DC: U.S. Department of Transportation, Federal Aviation Administration.

Guilford, J. P., & Lacey, J. I. (Eds.). (1947). Army air forces aviation psychology program research reports: Printed classification tests. Report No. 5. Washington, DC: U. S. Government Printing Office.

Hoffman, K. I, Guilford, J. P., Hoepfner, R., & Doherty, W. J. (1968). A factor analysis of the figural-cognition and figural-evaluation abilities. Tech. Rep. No. 40. Los Angeles: University of Southern California, The Psychological Laboratory.

Kanfer, R. (1990). Motivation and individual differences in learning: An integration of developmental, differential, and cognitive perspectives. Learning and Individual Differences, 2, 221-239.

Kanfer, R. (1991). Motivation theory and industrial/organizational psychology. In M. D. Dunnette and L. Hough (Eds.), Handbook of industrial and organizational psychology: Theory in industrial and organizational psychology (Vol. 1, pp. 75-170). Palo Alto, CA: Consulting Psychologists Press.

Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. Journal of Applied Psychology Monograph, 74, 657-690.

Landon, T. A. (1991). Job Performance for the en route ATCS: A review with applications for ATCS selection. Unpublished report to the Minnesota Air Traffic Controller Training Center.

Landon, T. A. (1992). Critical performance examples for air traffic controllers. Unpublished report.

Lohman, D.F. (1987). Spatial abilities as traits, processes, and knowledge. In R.J. Sternberg (Ed.) Advances in the psychology of human intelligence (Vol. 4, pp. 181-248). Hillsdale, NJ: Lawrence Erlbaum Associates.

Sells, S. B., Dailey, J. T., & Pickrel, E. W. (Eds.). (1984). Selection of air traffic controllers (FAA-AM-84-2). Washington, DC: U.S. Department of Transportation, Federal Aviation Administration.