

2005

The Air Traffic Selection and Training Battery: What It Is and Isn't (And How It Has Changed and Hasn't)

Raymond E. King

Andrew R. Dattel

Follow this and additional works at: https://corescholar.libraries.wright.edu/isap_2005



Part of the [Other Psychiatry and Psychology Commons](#)

Repository Citation

King, R. E., & Dattel, A. R. (2005). The Air Traffic Selection and Training Battery: What It Is and Isn't (And How It Has Changed and Hasn't). *2005 International Symposium on Aviation Psychology*, 388-392.
https://corescholar.libraries.wright.edu/isap_2005/57

This Article is brought to you for free and open access by the International Symposium on Aviation Psychology at CORE Scholar. It has been accepted for inclusion in International Symposium on Aviation Psychology - 2005 by an authorized administrator of CORE Scholar. For more information, please contact corescholar@www.libraries.wright.edu, library-corescholar@wright.edu.

THE AIR TRAFFIC SELECTION AND TRAINING BATTERY: WHAT IT IS AND ISN'T (AND HOW IT HAS CHANGED AND HASN'T)

Raymond E. King

Andrew R. Dattel

Civil Aerospace Medical Institute
Oklahoma City, OK

The Federal Aviation Administration (FAA) has developed a new selection procedure, the Air Traffic Selection and Training (AT-SAT) computerized test battery, to help select Air Traffic Control Specialists. AT-SAT is an aptitude test and not a test of air traffic control knowledge. Of the 264 applicants who have taken AT-SAT, 155 responded to a job announcement, while 109 previously passed the OPM (pre-employment) test and had to achieve a passing score on AT-SAT before they were admitted into training at the FAA Academy. Of the 155 job-announcement applicants, 131 (84.52%) achieved a passing score of 70 or greater (termed a “qualifying score”), while 24 applicants (15.48%) failed to achieve a minimum score of 70. Those who had been prescreened with the OPM test fared a bit better, with 104 (95.41%) achieving a qualifying score; five (4.59%) applicants failed. Current research efforts include equating a parallel form, rehosted on a Windows 2000 operating platform, with the assistance of research participants from the US Army, Navy, and Air Force. Another recent project was focused on reweighting the subtests and adjusting the overall constant to address issues of potential adverse impact, without compromising validity. A greater concern in this effort was to ensure that AT-SAT performance would predict job performance rather than just success or failure in training. Despite this reweighting effort and updating of the operating platform, the content of the battery remains unchanged. Future efforts will involve a longitudinal validation to compare performance on AT-SAT with success in training and on the job.

The development and validation of selection instruments for occupations where a sizeable number of applicants are needed to fill demanding positions play a critical role in reducing costs associated with attrition from training programs. Validation also ensures that those who are hired have (or are likely to develop) the necessary knowledge, skills, and abilities to perform successfully on the job. The duties of an Air Traffic Control Specialist (ATCS), specifically those providing *separation* services, which makes these employees individually responsible for more lives than the practitioners of any other occupation in the United States (Biggs, 1979). The FAA developed the AT-SAT battery to replace a two-stage selection process in which ATCS applicants completed an Office of Personnel Management (OPM) test battery and a nine-week screening program at the FAA Academy in Oklahoma City, OK. This previous selection process proved to be expensive and inefficient (Ramos, 2001). AT-SAT was developed based on the results of the Separation and Control Hiring Assessment (SACHA; Nickles, Bobko, Blair, Sands, & Tartak, 1995) job analysis of the duties of the ATCS options.¹

The SACHA job analysis reviewed the existing ATCS job analysis literature. An extensive assortment of documents was examined for terms suitable to the knowledge database, including FAA, military, and ATCS civilian courses. After reviewing and summarizing the existing job analysis information, the SACHA project staff visited sites to observe controllers from the two options and assignments. Subject-matter experts (SMEs) were also questioned about the qualities they considered necessary for effective job performance. The worker requirements determined necessary for the job of ATCS were then used to design a series of self-administering computerized tests to assess the ability of applicants to perform these tasks.

This paper focuses on the current status and future plans for the recently² implemented AT-SAT battery. AT-SAT is a computerized test battery comprised of eight subtests based on 22 individual scores that, when weighted (forming “part scores”) and combined, are totaled (with an overall constant added) for an overall score. AT-SAT comprises the following subtests: *Air Traffic Scenarios Test, ATST; Analogies, AY; Angles, AN; Applied Math, AM; Dials, DI; Experiences Questionnaire, EQ; Letter Factory,*

¹ There are three options in the 2152 occupational series: terminal, en route, and flight service station. Terminal controllers can be divided into two groups: tower cab and TRACON. AT-SAT is not used for the selection of flight service station personnel.

² AT-SAT was approved as the official ATCS selection test, for those applicants without previous air traffic control experience, on May 13, 2002, with June 2002 marking the first time the test was operationally used.

LF; and *Scan, SC*. AT-SAT is an aptitude test and not a test of air traffic control knowledge. The goal of AT-SAT is to gauge the likelihood of success in air traffic control training and, more importantly, subsequently on the job. Seven of the eight subtests assess aspects of cognitive ability, while one, *EQ*, assesses issues in the personal history/personality realm. Four (*ATST, AY, LF, SC*) of the subtests are dynamic; they are interactive and can only be administered via computer. The remaining four are static, similar to pencil-and-paper tests, but are administered via computer in AT-SAT.

Before operational use of AT-SAT was approved for hiring purposes, FAA employees who were members of minority groups raised concerns over potential adverse impact.³ Consequently, FAA management met with representatives from the groups to hear their concerns. The concern about the potential for adverse impact against African Americans seemed well founded, as only three out of every 100 black applicants were predicted to achieve a score of at least 70 (the minimum passing score – termed a “qualifying score”) on AT-SAT. The issue went beyond pass rates of minority applicants. By design, 38% of fully certified incumbent FAA controllers would not pass AT-SAT under the original scoring scheme. The original passing score of 70 had been calibrated so that only 62% of incumbent fully certified controllers would achieve an AT-SAT score equal to, or greater than, 70 in an effort to minimize FAA Academy failures and to compensate for the need for ATCSs to perform potentially more difficult duties in the future. The goal was to at least preserve and strive to improve the level of functioning in the workforce (Waugh, 2001).

In response, the FAA requested that scientists review the weights of AT-SAT subtests to reduce adverse impact. At the same time, there was an emphasis on maintaining the overall validity of the battery. Additionally and more importantly, management made the case that the cut score should be set at the point where most fully qualified incumbent FAA controllers would pass FAA’s entry-level aptitude test. Consequently, the AT-SAT subtests were re-weighted and the constant was adjusted. The content of the subtests themselves was not changed, rather the subtests were weighted differently. The

³ Adverse Impact – “A selection rate for any race, sex, or ethnic group which is less than 4/5 (80%) of the rate for the group with the highest rate” Uniform Guidelines on Employee Selection Procedures (1978), Sec 4D.

challenge was to retain adequate validity while reducing adverse impact. Test validity (job-relatedness) is determined by the strength of the correlation between the test score and job performance measures. After reweighting, the correlation between AT-SAT and job performance was reduced slightly, from .69 to .60. Compared with most validation coefficients, this is still a strong relationship with job performance. The relationship with job performance is especially important in this context as any remaining adverse impact can be justified by business necessity. In the end, however, it was found that potential adverse impact for women and Hispanics had been completely eliminated and had been greatly reduced for African Americans. Adverse impact will be continually assessed with job applicants. Finally, to further address the potential problem of adverse impact, the FAA decided to abandon a strict “top-down” approach to hiring and instead use a category ranking method. Under this scheme, successful examinees are divided into two groups: those scoring 85 and above (termed “well qualified”) and those scoring from 70 to 84.9 (termed “qualified”). Those in the well-qualified group will be offered employment before anyone in the qualified group. Within the ranges, veterans are hired before non-veterans, but selecting officials can consider other job-related factors, such as the ability to speak English well enough to be understood and self-reported interest in the job, dimensions that are not measured by AT-SAT.

A forthcoming study (Dattel & King, in preparation) applied the weights and additive constant developed to address potential adverse impact to the scores of 292 voluntary research participants who took the AT-SAT under the original scoring scheme. This rescoring increased the research participants’ overall scores by an average of 9.08 points, with the scores of Caucasians increasing by 8.84 points, African Americans by 9.82 points, and Hispanics by 11.03 points. Additionally, this rescoring increased the overall pass rate (scores equal to or greater than 70) in this sample from 36.3% to 68.2%. It is important to bear in mind that these test takers were not applicants and were, instead, voluntary research participants.

It should be noted that there are several applicant categories whose members do not have to take and pass AT-SAT to be considered for employment. Military controllers and Department of Defense civilian controllers are included in this category as well as former PATCO controllers who are now eligible for rehire. These applicants still face a competitive process and are by no means

automatically hired; they are just exempt from having to take AT-SAT.

The report, *A Plan for the Future: The FAA's 10-Year Strategy for the Air Traffic Control Workforce* (http://www.faa.gov/newsroom/controller_staffing/WorkforcePlan.pdf), was submitted to the U.S. Congress in December 2004. This report provides a plan to mitigate pending controller retirements and contemplates strategies to achieve appropriate staffing levels. While previously military applicants with air traffic control experience were able to satisfy many of the FAA's hiring needs, there is a need to begin hiring more controllers. The availability of applicants with previous experience will quickly be exhausted. AT-SAT will thus become an instrument of increasing importance. How did the hiring need become so urgent? An overwhelming majority of the air traffic control workforce went on strike on August 3, 1981. During this time, President Ronald Reagan ordered the striking controllers to return to duty within 48 hours. When 10,438 (out of a workforce of approximately 15,000) striking controllers did not return to work in this timeframe, the president fired them. Facing a sudden shortage of controllers, the FAA hired 3,416 individuals in 1982 and another 1,720 in 1983. From 1982 through 1991, the FAA hired an average of 1,527 individuals per year. The majority of entrants met the 18 to 30 years-of-age entry requirement. This hiring wave created the potential for a large portion of the controller workforce to reach retirement age at roughly the same time. Based on recent projections, over the next 10 years, 73 percent of the agency's 15,000 controllers will become eligible to retire. Total losses over the next 10 years are expected to be nearly 11,000 (FAA, 2005).

The Current State of Affairs

To date, 264 applicants have taken AT-SAT as part of their job application process; 155 of these applicants responded to a job fair announcement (soliciting applicants for a specific position), while 109 had previously passed the OPM test (pre-employment test) and had to achieve a passing score on AT-SAT before they were admitted into training at the FAA Academy. Of the 155 job fair applicants, 131 (84.52%) achieved a score of passing score of 70 or greater, while 24 applicants (15.48%) failed to achieve a minimum score of 70. Those who had been prescreened with the OPM test fared a bit better, with 104 (95.41%) achieving a qualifying score; five (4.59%) applicants failed. AT-SAT was also taken by 727 research participants. These participants were students enrolled at the Academy but took the AT-

SAT voluntarily (their enrollment was obtained via voluntary consent and their continued employment was not contingent on their performance on AT-SAT). This group includes, but is not limited to, retired military personnel and graduates of collegiate training initiatives (CTI) who were previously hired with the OPM test.

Figure 1 presents overall AT-SAT results in a continuous, as opposed to a dichotomous (pass/fail), fashion. To aid in the comparison of results between groups (job announcement or "job fair" applicant, OPM applicant, research participants), all results have been transformed into the current weighting scheme. The groups are significantly different, ($F(2,930) = 38.440, p < .001$), with OPM applicants significantly outperforming Job Fair applicants and Research participants. Job Fair applicants significantly outperformed research participants.

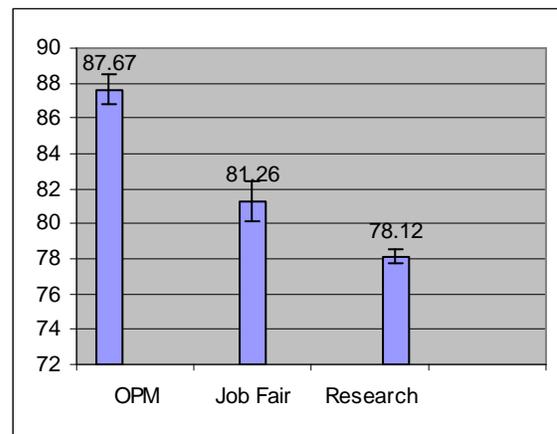


Figure 1. Overall AT-SAT Score by Participant Category (weighted under current scheme). Error bars indicate standard error.

The best way to appreciate results of AT-SAT is to consider the battery subtest by subtest (see Figure 2). As in Figure 1, all subtest scores have been transformed using the current weighting scheme. However, the means of the sub-test scores have been converted to standardized scores (z-scores⁴) for a more consistent presentation. There are significant group differences on six of the eight subtests (no significant differences were found between groups for ATST or AY), as delineated in Table 1.

⁴ Z-scores range from a high of 1 to a low of -1 with 0 as the mean.

Sub Test	Means (SD)	Post hoc group differences ($\alpha=.05$)
<u>DI</u> $F(2,930)=13.24$, $p<.001$, $Mse=.048$	OPM=1.91 (.19) Job Fair=1.85 (.20) Research=1.79 (.22)	OPM > Job Fair OPM > Research Job Fair > Research
<u>AM</u> $F(2,930)=19.08$, $p<.001$, $Mse=34.901$	OPM=21.00 (5.34) Job Fair=16.18 (6.57) Research=19.40 (5.88)	OPM > Job Fair OPM > Research Research > Job Fair
<u>SC</u> $F(2,930)=53.77$, $p<.001$, $Mse=4.091$	OPM=9.77 (.77) Job Fair=9.33 (1.45) Research=7.90 (2.20)	OPM > Research Job Fair > Research
<u>AN</u> $F(2,930)=14.41$, $p<.001$, $Mse=.060$	OPM=1.67 (.20) Job Fair=1.53 (.26) Research=1.53 (.25)	OPM > Job Fair OPM > Research
<u>LF</u> $F(2,930)=80.86$, $p<.001$, $Mse=4.282$	OPM=6.30 (1.47) Job Fair=6.29 (1.50) Research=4.21 (2.21)	OPM > Research Job Fair > Research
<u>ATST</u> $F(2,930)=3.01$, $p=.050$, $Mse=.354$	OPM=2.08 (.52) Job Fair=1.98 (.59) Research=1.93 (.60)	
<u>AY</u> $F(2,930)=1.92$, $p=.147$, $Mse=1.881$	OPM=5.61 (1.21) Job Fair=5.26 (1.32) Research=5.48 (1.40)	
<u>EQ</u> $F(2,930)=105.02$, $p<.001$, $Mse=49.773$	OPM=33.50(4.52) Job Fair=32.72 (5.76) Research=25.01 (7.49)	OPM > Research Job Fair > Research

Table 1. Group Differences, Subtest by Subtest.

While the superior performance of the OPM applicants is not surprising, given that they were previously screened with the OPM test, one should view the weaker performance of the Research participants with a degree of caution. These participants, even though some were also pre-screened with the OPM test, may have been less motivated to do their very best as they had already been hired and were explicitly told that their performance on AT-SAT would not impact their employment with the FAA.

What's Current and What's Next?

Only one version of the AT-SAT battery was constructed during the initial development and validation effort, meaning that all persons who took AT-SAT received the same items and in the same order. Consequently, there was an increased likelihood that any improvement in the score of someone who retook the test was due to a practice effect (Heil, Detwiler, Agen, Williams, Agnew, & King, 2002). The use of one version (or "form") also suggests that the test may be more vulnerable to coaching since there is only one set of items that

must be trained. The result is a potentially incorrect hiring (false positive) decision, with an increased likelihood that such an applicant would not be ultimately successful. A score inflated as a result of coaching does not increase the individual's actual ability to perform air traffic control work. To guard against the empirically demonstrated compromising effects of practice and coaching (Heil et al., 2002) and to mitigate against the deleterious results of the security of AT-SAT being compromised, an alternate version has been developed. The first step in this process, the "pilot study," was to develop alternative items and test them on volunteer research participants to ensure that they were at the appropriate level of difficulty. The U.S. Air Force and Navy graciously supplied these participants from air traffic control schools at Keesler Air Force Base, MS, and Pensacola Naval Air Station, FL, respectively. The end result was two parallel forms. Current research involves equating these parallel forms, rehosted on a Windows 2000 operating platform, an upgrade from Windows 95. The U.S. Army at Fort Rucker, AL, has joined its sister services in collaborating in this effort by supplying participants who are either air traffic controllers or students in air traffic control training. For adequate statistical power, the goal is to collect data from a total of 1,500 participants across these military sites. Each military participant completes two four-hour test sessions in the course of a day. While these research participants may differ from actual job applicants, they are encouraged to do their very best. When completed, the parallel version of AT-SAT will be comprised of the same subtests, with similar items. The tests will be presented in a standardized fashion.

Currently, AT-SAT is not used for placement decisions; that is, scores are not used to assign successful applicants to centers or terminal facilities. There is growing interest in determining if AT-SAT can be effective in placing new hires into facilities. Other future research efforts include longitudinal validation: comparing performance on AT-SAT with success in training and on the job. The ultimate goal of research with AT-SAT is to ensure that those selected to enter the ATCS career field possess (or will develop) the necessary knowledge, skills, and abilities to ensure that air traffic moves in a safe and expeditious manner.

References

Biggs, D. (1979) *Pressure cooker*. New York: Norton.

Dattel, A. R. & King, R. E. (in preparation). Gauging the impact of attempts to reduce potential adverse impact.

FAA (2005). *FAA announces 10-year air traffic controller staffing plan*. Retrieved January 20, 2005 from http://www.faa.gov/newsroom/controloler_staffing/WorkforcePlan.pdf. Federal Aviation Administration.

Heil, M. C., Detwiler, C. A., Agen, R. A., Williams, C. A., Agnew, B. O., & King, R. E. (2002). *The effects of practice and coaching on the Air Traffic Selection and Training Battery* (DOT/FAA/AM-02/24). Washington, DC: FAA Office of Aerospace Medicine.

Nickels, B. J., Bobko, P., Blair, M. D., Sands, W. A., & Tartak, E. L. (1995). *Separation and control assessment (SACHA) final job analysis report*. Bethesda, MD: University Research Corporation.

Ramos, R. A. (2001). Air traffic selection and training (AT-SAT) project. In R. A. Ramos, M. C. Heil, & C. A. Manning (Eds.) *Documentation of validity for the AT-SAT computerized test battery, Volume 1* (DOT/FAA/AM-01/5) (pp. 1-5). Washington, DC: FAA Office of Aerospace Medicine.

Waugh, G. (2001). Predictor-criterion analyses. In R. A. Ramos, M. C. Heil, & C. A. Manning (Eds.) *Documentation of validity for the AT-SAT computerized test battery, Volume II* (DOT/FAA/AM-01/6) (pp.37-42). Washington, DC: FAA Office of Aerospace Medicine.

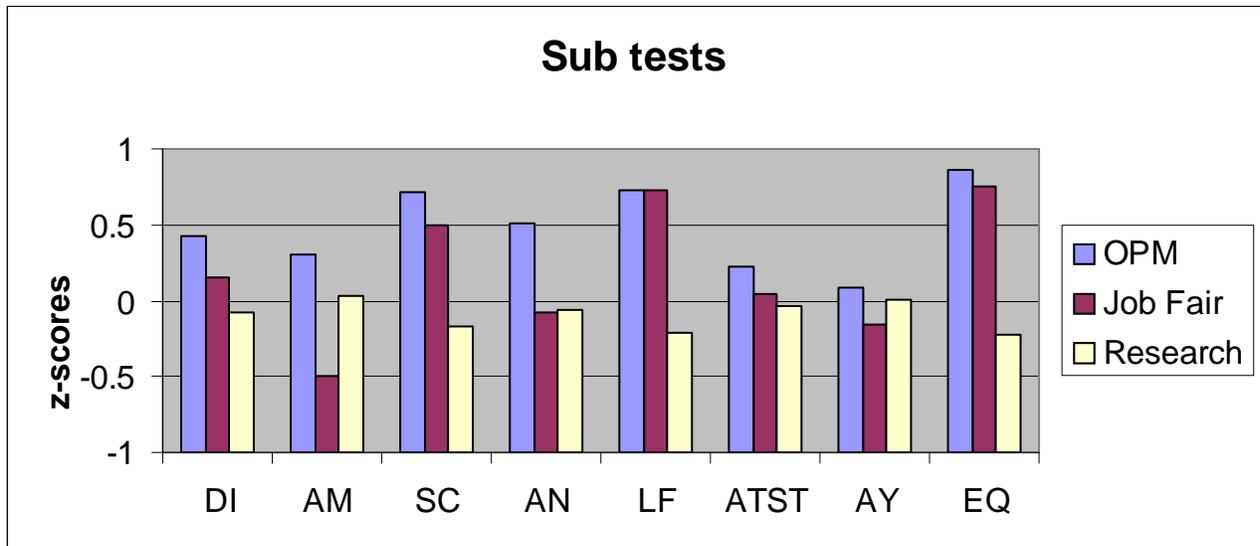


Figure 2. Group Differences, Subtest by Subtests