FAA Statement
on the
Barrier Analysis of the Air Traffic Control Specialist
Centralized Hiring Process

Administrator Michael Huerta has made an historic commitment to transform the Federal Aviation Administration (FAA) into a more diverse and inclusive workplace that reflects, understands, and relates to the diverse customers we serve. To meet this goal and satisfy the requirements of the Equal Employment Opportunity Commission MD-715, the Administrator tasked the Office of the Assistant Administrator for Civil Rights to conduct barrier analyses of the Air Traffic Control Specialist (ATCS) Centralized Hiring Process, Aviation Safety Inspectors, and Airway Transportation Systems Specialists.

The first study completed is on the ATCS series; therefore, the FAA is pleased to submit the reports entitled, "Barrier Analysis of the Air Traffic Control Specialist (ATCS) Centralized Hiring Process" and "Extension to Barrier Analysis of the Air Traffic Control Specialist Centralized Hiring Process." These reports reflect a collaborative effort undertaken by the FAA's Office of Civil Rights, Office of Human Resources, and the Air Traffic Organization. The primary purpose of these reports is to identify and analyze potential barriers to equal employment opportunities within the ATCS Centralized Hiring Process and to offer solutions to establish the foundation for improving the Process.

The reports reflect a detailed scope of work, approaches and methodologies, work plans, and analytical provisions including overall hiring conditions within the ATCS job series 2152. Our consultant, Outtz and Associates, was commissioned to conduct the barrier analysis, which began in April 2012, with the issuance of the final report in May 2013. The barrier analysis identified that four (4) of seven (7) decision points in the air traffic controller hiring process resulted in adverse impact to applicants from at least one demographic group. Subsequently, another independent consultant, APT Metrics, was hired to analyze the barrier decision points, specifically reflecting on the differential pass rates for protected group members. APT Metrics' report was finalized and issued in February 2013. These reports, in tandem, present recommendations and specific suggestions to improve the ATCS Centralized Hiring Process and to ensure that there will be no barriers to equal employment opportunity.

Significant progress is now underway. To date, progress includes the establishment of an Executive Steering Committee comprised of senior agency executives. The Steering Committee provides oversight for the new hiring process and has implemented multiple cross functional project teams to operationalize the recommendations identified in the report.

ATCS Centralized Hiring Process improvements being implemented to support the Fiscal Year 2014 hiring of air traffic controllers include (1) comprehensive outreach and recruitment, (2) improved automation enhancements to our application process, (3) revisions to the Air Traffic Selection Assessment Tools, and (4) standardization of human resource procedures in review of applications.

These efforts will result in important improvements in the ATCS Centralized Hiring Process, further demonstrating the FAA's commitment to equal employment opportunity for all.

Barrier Analysis of the Air Traffic Control Specialists (ATCS) Centralized Hiring Process

Dr. James L. Outtz | Dr. Paul J. Hanges

Outtz and Associates

May 8, 2013

Contents

TABLES

ACRONYMS

ATCS	Air Traffic Control Specialists
AFSM	Armed Forces Services Medal
AI	Adverse Impact Ratio
ATCT	Air Traffic Control Towers
AT-CTI	Air Traffic Collegiate Training Initiative
AT-SAT	Air Traffic Selection and Training
ATST	Air Traffic Scenarios Test
CAMI	Civil Aerospace Medical Institute
CBPM	Computer Based Performance Measure
CoVATCH	Concurrent Validation of AT-SAT for Tower Control Hiring
CPC	Certified Professional Controllers
CSP	Centralized Selection Panel
CTI	Collegiate Training Initiative
CTO	Control Tower Operators
EEO	Equal Employment Opportunity
EEOC	Equal Employment Opportunity Commission
EQ	Experience Questionnaire
EVHO	Enhanced Veterans Hiring Opportunity
FAA	Federal Aviation Administration
GPA	Grade Point Average
HR	Human Resources
HRMO	Human Resource Management Office
OPM	Office of Personnel Management
RMC	Retired Military Controller
RNO	Race and National Origin
SDI+	Self-Description Inventory+
UGESP	Uniform Guidelines on Employee Selection Procedures
VRA	Veteran’s Recruitment Appointment

EXECUTIVE SUMMARY

The Federal Aviation Administration (FAA) performs critical functions to ensure safe and efficient air travel in the National Airspace System (NAS). To accomplish this mission, the Agency relies heavily on trained service providers—the 35,000 controllers, technicians, engineers, and support personnel whose daily efforts keep aircraft moving safely through the nation's skies. These professionals are responsible for managing a very complex and highly automated network of interconnected systems to ensure the safety of the nation's commercial airliners. In particular, air traffic controllers, or the 2152 job series, carry out thousands of air traffic control activities daily and require significant training to prepare them for this job. The Federal Aviation Administration officials have unveiled a plan to hire 12,500 controllers during the next decade. The hiring surge was designed to offset the effects of an anticipated wave of retirements. The Federal Aviation Administration officials have projected that 11,000 air traffic controllers will leave the agency by 2014. The surge in retirements is linked to the 1981 Professional Air Traffic Controllers Association strike, when thousands of workers walked off the job and were replaced. These employees are now nearing retirement age.

Questions have been raised regarding the disproportionate underrepresentation of women and minorities among the air traffic control profession. This report provides the results of a barrier analysis for the Air Traffic Control Specialist (ATCS) centralized hiring process at the FAA.

The Equal Employment Opportunity Commission's (EEOC) Management Directive (MD) 715 outlines a comprehensive workforce analysis process to identify triggers for barrier analysis. MD 715 requires federal agencies to ensure the workplace is free of barriers that impede full opportunities to all persons in the workplace. The objective of the barrier analysis is the identification of specific barriers to the employment by race/national origin, sex, and disability. If the barriers identified are sufficiently detailed, effective actions must be taken by specific organizations to remove the barriers and improve the diversity of their workforces. The barriers must be specific, clear, and sufficiently detailed or else the barriers identified will be too general to yield appropriate actions to improve the diversity status of any particular class group. More importantly, agencies are required to file annual reports with EEOC that detail how well they are doing with identifying and addressing adverse effects from barriers to employment.

The FAA, Office of Civil Rights commissioned Outtz and Associates to conduct a barrier analysis of the ATCS position. The process was guided by MD 715 and the Uniform Guidelines on Employee Selection Procedures. Based upon the review of the ATCS Centralized Hiring

Process, seven critical decision points were identified for analysis as potential barriers to racial and gender employment opportunities:

1. Qualification determination of applications
2. AT-SAT testing phase
3. Preparation of the referral list of eligible and qualified applicant list
4. Centralized Selection Panel determination of selections to the interview process
5. Interview process
6. Medical clearance process
7. Security clearance process

Evidence of barriers was found for racial/ethnic minorities at the first four decision points. Evidence of barriers was found with regard to race/ethnicity and gender for the second decision point. No barriers were found for the last three decision points. Overall, women and minorities were underrepresented among those successfully completing the ATCS Centralized Hiring Process and being hired.

Recommendations to address the barriers include making improvements to the ATCS hiring process, developing a targeted marketing and outreach campaign to increase diversity within applicants pools, standardization of human resources (HR) evaluation criteria, removing inconsistencies in centralized selection process criteria, and revising how the AT-SAT is used in establishing best-qualified lists.

OVERVIEW

BACKGROUND

The FAA unveiled a plan for replacing three-quarters of the air traffic controller workforce by hiring 12,500 controllers during the next decade to offset the effects of an anticipated wave of retirements resulting in 11,000 controllers leaving the agency by 2014. The surge in retirements is linked to the 1981 Professional Air Traffic Controllers Association strike, when thousands of workers walked off the job. President Ronald Reagan fired the striking controllers and the FAA subsequently hired mandatory replacement workers. Those newly hired controllers are now nearing retirement age. In 2014 the hiring surge is estimated at or around 11,000 hires, and the FAA expects to have more than 16,000 air traffic controllers in the agency.

SCOPE

This report was commissioned in part due to the legal requirements contained in the EEOC MD 715, which requires federal agencies to ensure that the workplace is free of barriers that impede full opportunities to all persons in the workplace. Management Directive 715 provides policy guidance and standards for establishing and maintaining effective programs of equal employment opportunity under Section 717 of Title VII (Part A) and effective affirmative action programs under Section 501 of the Rehabilitation Act (PART B). Section II of these instructions contains operational guidance on how to identify barriers that tend to inhibit free and open workplace competition, and how to develop a meaningful plan to eliminate those barriers to equal employment opportunities.

METHODOLOGY

Before examining the effects of the ATCS Centralized Hiring Process, we first examined the diversity within the application sources used by the FAA as input into the hiring system. Prior to assessing the diversity of the applicant pools, the data had to be adjusted to the applicant level of analysis. That is, the data set originally represented applications, not applicants, in that a single applicant could apply multiple times throughout a fiscal year depending upon the number of posted announcements. The Race and National Origin (RNO) and gender analysis that we are reporting required an assessment of individuals (i.e., applicants) and not applications.

The methodology we used to conduct the barrier analysis also involved an in-depth, root-cause analysis on a multivariance analysis approach and the Uniform Guidelines on Employee Selection Procedures. The Federal government's need for a uniform set of principles on the

question of the use of tests and other selection procedures has long been recognized. The EEOC, the Civil Service Commission, the Department of Labor, and the Department of Justice jointly have adopted these uniform guidelines to meet that need, and to apply the same principles to the Federal government as are applied to other employers. These guidelines incorporate a single set of principles that are designed to assist employers, labor organizations, employment agencies, and licensing and certification boards to comply with requirements of federal law prohibiting employment practices that discriminate on grounds of race, color, religion, sex, and national origin. They are designed to provide a framework for determining the proper use of tests and other selection procedures. These guidelines do not require a user to conduct validity studies of selection procedures where no adverse impact results have been seen. However, all users are encouraged to use selection procedures that are valid, especially users operating under the Federal government's merit principles.

As such, for each year from 2006 to 2011, we analyzed the qualification decision rates made as a function of the RNO and gender subgroups that comprised more than two percent of the population of applicants. We determined from this analysis that Native American/American Indian/Alaska Native and Native Hawaiian/Other Pacific Islander subgroups would be excluded from all subsequent analyses in this report due to their less than two percent representation in the applicant population. The decision to exclude these two subgroups is consistent with the analysis recommendations of the Uniform Guidelines issued by the EEOC, Civil Service Commission, Department of Labor, and the Department of Justice in 1978.

We relied on several criteria to determine the presence of a barrier. Specifically, one criterion that we used was computing a statistic referred to as *effect size* or *d-ratio*. This is a common statistic used in the scientific literature and it is used to understand the magnitude of the difference in selection rates between two groups. The benefit of using *d-ratios* is that there are standards that have been proposed to help researchers and practitioners interpret the magnitude of the *d-ratio*. In particular, *d-ratios* are considered small if they are less than 0.30; they are considered moderate if they are in close to 0.50; and they are generally considered large if they are close to 0.80 or above (Cohen, 1988). Smaller *d-ratios* may be considered trivial whereas larger *d*-values are more problematic. This, however, is not always the case. There are conditions when even small *d-ratios* indicate significant, practical effects. As an example, this can occur in situations in which the overall selection ratio (i.e., number of vacancies divided by the number of applicants) is very low. Very low selection ratios occur when there are typically very few openings and many applicants. In this situation, even small differences in the selection rates for various applicant groups can have meaningful negative consequences and constitute a barrier. For this reason we used three criteria to identify barriers. Our criteria are based on the weighted average of statistical values across years. Our specific criteria are:

- a statistically significant difference between a minority or gender group and the majority group (in this instance Whites and males),

- an effect size of 0.20 or higher, or
- an adverse impact ratio below 0.80.

If a selection step or component of the hiring process meets at least two of the three criteria, we consider it a barrier. Examining Tables 3 through 6, for example, it can be seen that qualification decisions for African-Americans meet all three criteria and therefore constitute a barrier for that group.

In the Technical Report that follows, there is a complete description of the scope and methodology of the barrier analysis and our approach using the data that we received from the agency.

General Items to Note

Before discussing our specific findings, we note the following three general items. First, the percentage of applicants self-reporting RNO and gender status is quite high. This supports the premise that the samples reporting RNO and gender data are sufficiently large in relation to the total sample to indicate that there is little, if any, difference between the statistics from the self-report sample and what would be expected from the total sample. This is true even though more applicants self-reported RNO status than gender status. Second, the centralized hiring process was not fully implemented until 2007. The data for FY 2006 reflects the pre-2007 hiring process together with the ATCS Centralized Hiring Process currently under examination. Thus, even though we present results from FY 2006, we do not place as much weight on the FY 2006 results as we do on the results from FY 2007 to FY 2011. Finally, we computed weighted averages (e.g., Table 1) to provide an aggregate picture of the RNO and gender diversity composition in the various applicant sources across the fiscal years. This weighted average is computed using the formula described on page 34.

When we analyzed the variability of RNO and gender representation across sources of applicants, it became clear that the FAA has not used all applicant sources throughout the years. Indeed, announcements generating applications from the public source were only posted during FY 2007, FY 2008, and FY 2009. This is despite the extensive representation of African-Americans in this application pool (average 31.3%), as well as our finding that the General Public application source has the second highest pass-rate of minimum qualifications compared to the other sources of applicants. This finding (General Public source having the highest RNO diversity and second highest pass rate) is at odds with current FAA policies, such as not posting ATCS jobs to the General Public since 2009. This policy is consistent with general comments from trainers such as, “Please tell them not to send me any more public hires” (quoted in Barr, Brady, Koleszar, New, & Pounds, 2011; p. 16). The Barr et al. (2011) report interpreted this quote and other similar data, as indicating that the FAA “needs to review its hiring practices to

take advantage of the Air Traffic-Collegiate Training Initiative (AT-CTI) system it has created" (p. 16). As a result of our findings, we question these policies and conclusions. Given the limited number of applicants and especially the diminished diversity in the AT-CTI application source, there are serious diversity consequences for not fully using the General Public application source.

We also identified concerns about the RNO and gender diversity of the CTI schools. A recent report produced by the FAA titled "Air Traffic-Collegiate Training Initiative Diversity Data 2011-12" dated June 12, 2012, sought to determine the diversity of the potential applicant pool from the partner CTI colleges and universities. The authors of that report sought to establish a baseline that would be used to assess the effectiveness of future efforts to improve the diversity of the AT-CTI program. However, when the report was examined in detail, several issues became clear. First, there is a dropout of African-Americans from freshman year (16.9%) to the senior year (6.0%) in four-year CTI programs. This dropout rate of African-Americans is troubling. We recommend that the FAA undertake efforts to understand the cause for this selective dropout rate and take reasonable countermeasures to reduce it in the future. The FAA should investigate this problem by contacting four-year CTI schools and working with them to identify the possible causes of this problem.

Second, while the report "Air Traffic-Collegiate Training Initiative Diversity Data 2011-12" estimated the percentage of African-Americans in the CTI schools to be 12%, that percentage is an overestimate determined by including all students, regardless of their matriculation credits. For purposes of the FAA diversity assessment, the more important statistic is to examine the percentage of students from each ethnic subgroup who are in their senior year in their programs. Specifically, 6% of the CTI seniors from universities are African-American, which actually more closely corresponds to the 2011 CTI applicant source results that we report in Table 1F in Appendix A (5.4%) than the 12% African-American students reported in "Air Traffic-Collegiate Training Initiative Diversity Data 2011-12."

Finally, it would be helpful in future reports for the RNO and gender results to be reported at the school level as well as the individual level of analysis. A school-level analysis will identify diversity problems at the college or university level. This issue is not addressed in the analysis reported in "Air Traffic-Collegiate Training Initiative Diversity Data 2011-12."

Another barrier that we identified concerned the decision regarding the applicant sources that will be used to sample applications. The definition of the applicant pool affects the extent to which different RNO and gender groups are able to pass the minimum qualification criterion. This application source barrier precedes all other barriers in this process because it not only is the first step in determining the hiring process, but it also influences where to draw the applicants from outside the legal requirements that must be met and what sources are the best

sources for diversity in the applicant pools. The FAA data clearly shows which applicant pools have the greatest diversity. Despite this data, however, a decision was made by the FAA to exclude one of the most diverse gender and RNO applicant pools. This decision seems to be the result of attempts to manage the application workload, rather than a legitimate issue of business necessity. This decision had major consequences for the diversity of the eventual pool of potential ATCSs.

All the sources, except the General Public source, require the applicants have air traffic control training or experience. The basis for this requirement is that the content of this previous training or experience will be sufficiently similar to the actual job. This is what is called a *content validation approach* to establishing the business necessity of a practice or policy. Unfortunately, this policy is inconsistent with the fact that the FAA has a training school in place to train applicants. Further, this requirement is inconsistent with the job description, which states that the job requires a high school diploma, three years of general work experience, or a combination of college and general work experience. Thus, the policy requiring applicants to have the same skills that the agency provides in training violates the content validity provisions of the Uniform Guidelines on Employee Selection Procedures.

Some of our interviews raised questions regarding the consistency with which the three years of progressively responsible experience standard is being applied to evaluate the qualifications of the public announcement pool of applicants. We see this inconsistency as a definite barrier. The variability in the standards used to assess applicants allows for errors and biases to enter the hiring process. Psychological research has shown that the more inconsistency in an evaluation process, the greater the likelihood that bias can occur (Roch, Woehr, Mishra, & Kieszczynska, 2012).

Our analysis regarding the AT-SAT indicated that its use is problematic. First, using the original pass score of 70, thousands of people were passing the AT-SAT. This created an administrative burden on the centralized hiring process that the FAA attempted to resolve by creating a “highly qualified” band of applicants (i.e., 85 or above AT-SAT score). Prior work rescoring the AT-SAT was not focused on correcting the administrative burden caused by the numbers of applicants passing the test. Rather, this earlier rescoring of the AT-SAT addressed a serious problem in that only 62% of fully certified controllers would pass the test using its original weighting scheme. In other words, 38% of good, currently employed controllers would fail the AT-SAT. The reweighting directly corrected this problem and as a result, the pass rate using a cut score of 70 for certified controllers was 98.8%.

With regard to using the AT-SAT with actual applicants, the report describing the AT-SAT reweighting indicated that 67.5% of the applicants should pass the test using the 70 or better score. This report also indicated that only 19.2% of the applicants would score 85 or better on

this test. Unfortunately, when the actual pass rates were compared with these expected rates (see Tables 11 and 12), it is clear that something is wrong with the AT-SAT. The current pass-rate of the AT-SAT (70 pass score) is averaging around 95.3%, and the percentage of people scoring 85 or higher on the test is averaging 60.2%. More troubling, there is evidence that the percentage of people scoring 85 or higher on the AT-SAT in certain RNO classifications has been steadily increasing over the last three years at a higher rate than others (see Table 11).

Finally, the FAA practice of reviewing and considering all applicants in the highly qualified band before any applicants in the qualified band is highly questionable. This practice creates adverse impact and there is no evidence that an applicant who scores at or near 85 (highly qualified band) on the AT-SAT is substantially superior to an applicant who scores at or near 84 (qualified band).

These findings are even more troubling in the context that the criterion measure used in the corrected validation study seems to have overlooked that the FAA trains selected applicants for the position. In other words, training performance should have been considered as a criterion measure in that study. Our analysis indicates the policy of considering applicants in the highly qualified band before considering applicants in the qualified band is a root cause of the diminished RNO diversity when examining the consequences of using the AT-SAT.

Furthermore, current use of the AT-SAT is a barrier to RNO and gender diversity at the FAA. The following components of the AT-SAT are a barrier based on either race/ethnicity or gender:

- Dials Test
- Applied Math Test
- Angles Test
- The Letter Factory Test
  - The Letter Performance Scale
- Air Traffic Scenarios Test (ATST)
  - ATST Efficiency Score
  - ATST Safety Subtest
  - ATST Procedures Scale
- Analogies Test

In addition, the decision to develop and use the highly qualified band resulted in substantial reduction in minority and gender representation in the FAA from the application sources required to take this test. Unfortunately, the FAA did not appear to question the consequences of adopting this policy change regarding the AT-SAT pass-score decision. Given that (a) the original AT-SAT validity report indicated that the cut score of 70 was justified, (b) the original AT-SAT validity report warned about the adverse impact consequences of using the

AT-SAT if alternative diversity recruiting methods were not adopted, and (c) several subsequent reports indicated that the AT-SAT would minimize adverse impact if a cut score of 70 was maintained, the FAA was not justified from a diversity perspective to use the highly qualified band to reduce its administrative burden caused by too many applicants passing the test. The high passing rate was probably indicative of more underlying problems with the exam than inappropriate pass scores.

Finally, as revealed in our results, several of the scales of the AT-SAT show substantial problems with regard to RNO and gender diversity. Specifically, the Dials subtest, the Applied Math subtest, the Angles subtest, the AT-Scenarios Efficiency and Safety scales caused problems with both RNO and gender diversity. The Scan subtest, the Letter Factory Situation Awareness and Performance Scales, the AT-SAT Scenarios Procedure scale, and the Analogies Correct Scale caused problems with RNO diversity. Steps should be explored to reduce the severity of the adverse impact of this instrument. For example, the 2001 validity report assesses the incremental validity for all of the AT-SAT scales. It should be noted that the validity report shows that a number of the AT-SAT components that result in adverse impact do not exhibit incremental validity (Waugh, 2001a) (see Technical Report page, 135).

Based upon our review, we identified seven critical points at which barriers to racial and gender employment opportunities can be created. These seven decision points are as follows:

1. Qualification determination of applications
2. AT-SAT testing phase
3. Preparation of the referral list of eligible and qualified applicant list
4. Centralized Selection Panel determination of selections to the interview process
5. Interview process
6. Medical clearance process
7. Security clearance process

We considered each point separately and provide the root-cause analysis of the multivariance step analysis that led to the stated conclusions and recommendations regarding each of the identified barriers.

Point 1 – Qualification Determination: The first step in the centralized hiring process is the qualification determination. This step involves verifying whether the qualifications of the applicants match the minimum qualifications specified in the job announcements. The determination of whether someone matches the minimum qualifications for the position is either a system decision (i.e., the applicant is missing an objective qualification) or is a decision made by HR personnel (i.e., a subjective decision regarding whether the person has three years of progressive work experience or the requisite combination of experience and education).

Point 2 – AT-SAT testing phase: Following the qualification determination phase, applicants from the General Public and the CTI sources must complete the AT-SAT test. Applicants from other hiring sources do not take the AT-SAT test, but move to the next step of the hiring process, the preparation of the referral list of eligible and qualified applicants by state (see Figure 1 on page 28). The AT-SAT consists of eight subtests. Some of these subtests are composed of multiple scales. A total of 23 AT-SAT scales are standardized, weighted, combined, and transformed into the final AT-SAT composite score. Applicants who score 85 and above are considered highly qualified; those who score 70 to 84.9 are considered qualified.

First, we examined the RNO and gender composition of the AT-SAT applicants in the highly qualified band followed by the qualified band. Table 11 shows the percentage of applicants in the highly qualified band by RNO. As can be seen from this table, the percentage of African-Americans passing at the highly qualified level is substantially lower than the percentage of Whites. The same is true for Hispanics/Latinos compared to Whites. In other words, there is evidence that this highly qualified band is resulting in a substantial adverse impact for RNO minorities.

Point 3 – Referral List: Table 151 shows the descriptive statistics for the referrals for each racial/ethnic subgroup that comprised more than two percent of the population over the period from 2006 to 2011. Table 152 shows the results of the statistical significance tests comparing the referral rate of each minority racial/ethnic subgroup to the majority White subgroup. As shown in this table, the White subgroup is referred significantly more often than African-Americans, Hispanics/Latinos, and the Multiracial subgroups.

Table 153 shows the adverse impact ratios of the referral rates. The effect sizes of these subgroup differences are shown in Table 154. They are medium for African-Americans and Multiracial subgroups but small for Asians and Hispanics/Latinos. Given that (a) there is a statistically significant difference in referral rates for African-Americans and Whites (based on the average across years), (b) the adverse-pact ratio associated with this difference is 0.61 (Table 154), and (c) the effect size is 0.48, we conclude that the process for preparation of the referral list of eligible and qualified applicants is a barrier to African-Americans.

Point 4 – Centralized Selection Panel: Table 159 shows the descriptive statistics for the referrals to the interview for each racial/ethnic subgroup that comprised more than two percent of the population over the period from 2006 to 2011. Table 160 shows the results of the statistical significance tests comparing the referral rate of each minority racial/ethnic subgroup to the majority White subgroup. As shown in this table, the White subgroup is referred significantly more often than African-Americans are based on the average across years. In addition Table 161 shows that the adverse impact ratio for this difference is below 80%. Table

162 shows that effect size is above 0.20 based on the average across years. Therefore we conclude the Centralized Selection Panel process is a barrier for African-Americans.

Point 5 – Interview: Our review of the interview data indicated that there was no adverse impact. This was due to the fact that the pass rates on this decision point are over 85% for all RNO and gender subgroups.

Point 6 – Medical Clearance: Table 167 shows the analyses for the medical clearance. No adverse impact was found for either RNO or gender diversity.

Point 7 – Security Clearance: We conclude that Security Clearance: Conditional Suitability and Final Suitability are not barriers.

Based on our analyses, we found evidence of barriers to racial/ethnic minorities at the first four decision points in the ATCS hiring process. We also found evidence of a barrier with regard to gender at the second decision point (i.e., AT-SAT testing phase). No evidence of any barriers for the remaining decision points were found.

SUMMARY RECOMMENDATIONS

The Executive Recommendations that follow summarize our conclusions. Each of the decision points was reviewed by a root-cause analysis using a multivariance approach to identifying and eliminating barriers in the ATCS hiring process. As such, the barriers described in the analysis and our methodology are detailed in the relevant sections under the decision point. This is a summary of the recommendations in the Technical Report. Based upon our review, we make the following recommendations:

- A positive climate for race and national origin (RNO) and gender diversity needs to be established.
    - RNO and gender diversity need to be considered a high priority and any changes to the ATCS selection system need to be documented prior to adoption.
- The technical quality of the entire ATCS hiring process needs to be improved. By technical quality, we mean the aspects of the process that make its application consistent and accurate.
    - The process needs standardization, monitoring, and overall improvement on an on-going basis.

- A Steering group responsible for monitoring and improving the ATCS Hiring Process should be established. At a minimum this committee should include a representative from the Office of Human Resources (AHR) and a representative from the Office of Civil Rights (ACR). ACR should have oversight per policy order 1400.8A, Chapter2, Section 4e (5/16/12).
    - This committee could be an existing FAA unit/division or a group of persons drawn from various units throughout the FAA who are stakeholders in the ATCS selection process.

❖ There should be continued community outreach efforts to educate applicant populations about the ATCS occupational series.

❖ Qualification decision point:

- A standardized training program for HR specialists should be established. The training program should focus on how to properly evaluate applicant qualifications.
- The list of collegiate training initiative (CTI) schools needs to be reconsidered with regard to the degree to which it fosters a diverse applicant pool.
- CTI school differential effectiveness
    - Barr et al. (2011) reported that not all AT-CTI programs are equal. They recommended that the FAA track the success of the graduates from each program. The AT-CTI schools differ in terms of equipment (e.g., simulations time and equipment) and the length of training (e.g., 2-year versus 4-year education) given to the students. The Barr et al. (2011) report recommended categorizing the AT-CTI schools by levels based upon the strength of curriculum. We highly recommend analyzing the quality of the product from these schools as well as their diversity. RNO and gender diversity need to be explicitly considered when determining the sources for applicants in each upcoming recruitment year.
- CTI diversity data for 2011-12 indicated that a significant number of African-Americans drop out of CTI programs between the freshman and senior year. The FAA should investigate this problem by contacting four-year CTI schools, working with them to identify possible reasons for this phenomenon to correct the problem.

❖ AT-SAT decision point

- The agency uses this cognitively loaded test while minimizing the use of noncognitive tests to select applicants, despite the fact that applicants are trained in required competencies, and many noncognitive competencies are important to job success. There is little, if any, evidence that use of the AT-SAT with today's

applicants predicts subsequent job performance, therefore less weight should be given to it. The original validation report recommended a cut score of 70. The Agency's use of highly qualified band creates additional RNO and gender diversity problems; therefore one band of 70 and higher should be used. There are alternative ways of combining multiple predictors to handle high applicant volume within ATCS selection (e.g., multiple hurdles, unit weighting of components, pareto-optimal weighting of components).

- We recommend using a version of the multiple hurdle approach in which the components with the least adverse impact are used first in the hiring process to identify applicants with maximum potential. We recommend that stringent but defensible pass scores be set for these front loaded components. Then the components that have the most adverse impact are used in the latter stages of the hiring process. Given that a large portion of the applicant pool will be eliminated by the first hurdle(s), the pass scores for subsequent components can be more lenient. Research has demonstrated that this approach maximizes diversity while minimizing reductions to criterion-related validity.

- Our overall recommendation is to revise the way AT-SAT is used (near-term), and eventually replace it (long-term).
  - Update the items for those components that can be shown to be valid.
  - Revise the test so that it assesses cognitive ability in the context of how that ability is used on the job while considering RNO and gender diversity.
  - Further, search for additional predictors to add to the selection protocol.
  - Apply a multihurdle approach to reduce adverse impact. In this approach, the experience components of the AT-SAT would be used to first identify the applicants with the most promise. This is accomplished by setting a strict pass score for this component of the AT-SAT. The cognitive AT-SAT components would then be used as the second hurdle to identify the applicants who will advance to the next step, the Centralized Selection Panel.
  - Immediately complete any AT-SAT validation projects that are in progress and use this information to improve the selection process.

❖ Centralized Selection Panel (CSP) decision point

- Remove inconsistencies in the CSP process by taking the following steps:
  - Either prohibit calls to applicants' references or mandate that calls be made to all references.

- Ensure that everyone understands the goals of the Centralized Selection Panel. Some Centralized Selection Panel members assume that applicants have already been screened and therefore think that their job is to identify a location for them.

- The use of location preferences of applicants is problematic. For example, some of the CSP members adhere strictly to the preferences listed by the applicants whereas others are less consistent. Further, applicants may be willing to take a job in a location other than the two listed in their application. Unfortunately these applicants are lost due to this artificial limitation.

  - A standardized training process for new CSP members should be developed.
  - There should be greater control of the information that the CSP receives.
    - Exploratory items on the application should be eliminated so that CSP participants do not see them.
    - Applicants' self-entry of AT-SAT and other scores on their application should be prevented.
  - A policy should be established and/or enforced regarding conflicts of interest in CSP recruitment and screening (e.g., forwarding an applicant to the next stage of process because a CSP member knows the applicant personally).

- Interview decision point

  - Make the interview process more consistent across interviewers by developing a specific set of questions or categories of questions to be asked. Guidelines should be developed as to the manner in which applicant responses are to be evaluated.
  - Alternative questions should be developed for use after a given set of questions has been used three times.
  - Rigorous, frame of reference training should be developed and implemented for interviewers. This training should and include the interview guidelines described above.
  - Standards should be developed regarding who can serve as an interviewer. These standards should address the level or type of training that must be completed and any experience deemed to be necessary. Interviewers should be required to undergo recertification of their training skills periodically.

TECHNICAL REPORT

INTRODUCTION

This report provides results of a barrier analysis of the air traffic control specialist (ATCS) centralized hiring process at the Federal Aviation Administration (FAA).

THE ORIGIN OF BARRIER ANALYSIS REQUIREMENTS

"MANAGEMENT DIRECTIVE 715 AND THE INSTRUCTIONS:

The Equal Employment Opportunity Commission (EEOC) issued MD 715 to provide guidance and establish standards for developing and maintaining effective programs of equal employment opportunity (EEO) under Section 717 of Title VII of the 1964 Civil Rights Act. This directive provides policy guidance and standards for establishing and maintaining effective affirmative programs of EEO under Section 717 of Title VII (PART A) and effective affirmative action programs under Section 501 of the Rehabilitation Act (PART B). The intent of this directive was to ensure that federal employees and applicants have equality of opportunity regardless of race, sex, national origin, color, religion, disability, or reprisal for engaging in a prior protected activity. Part A of MD-715 clarifies the concept of a barrier analysis and provides a set of instructions that guides and set standards to help maintain effective EEO affirmative programs.

SECTION II: BARRIER IDENTIFICATION AND ELIMINATION

Section II of these instructions contains operational guidance on how to identify barriers that tend to inhibit free and open workplace competition, and how to develop a meaningful plan to eliminate those barriers.

To develop a competitive and highly qualified workforce, federal agencies must fully use all workers' talents, without regard to race, color, religion, national origin, sex, disability, or reprisal for prior EEO activity. This goal cannot be accomplished when barriers to EEO persist in an agency's management/personnel policies, procedures, or practices.

As Section II of these Instructions explains, the barrier analysis process requires much more of agencies than has been asked in the past. The barrier analysis process cannot be guided solely by examining workforce statistics. While snapshot statistics can be useful as a starting point, statistics alone do not enable an agency to effectively identify workplace barriers.

Workforce statistics can serve to reveal symptoms of barriers to equal opportunity. It must be understood, however, that the statistics themselves are not the barriers. Therefore, when there is an indication, through statistical analyses or other means, that potential barriers exist in the workplace, an agency is responsible for undertaking a thorough examination of all

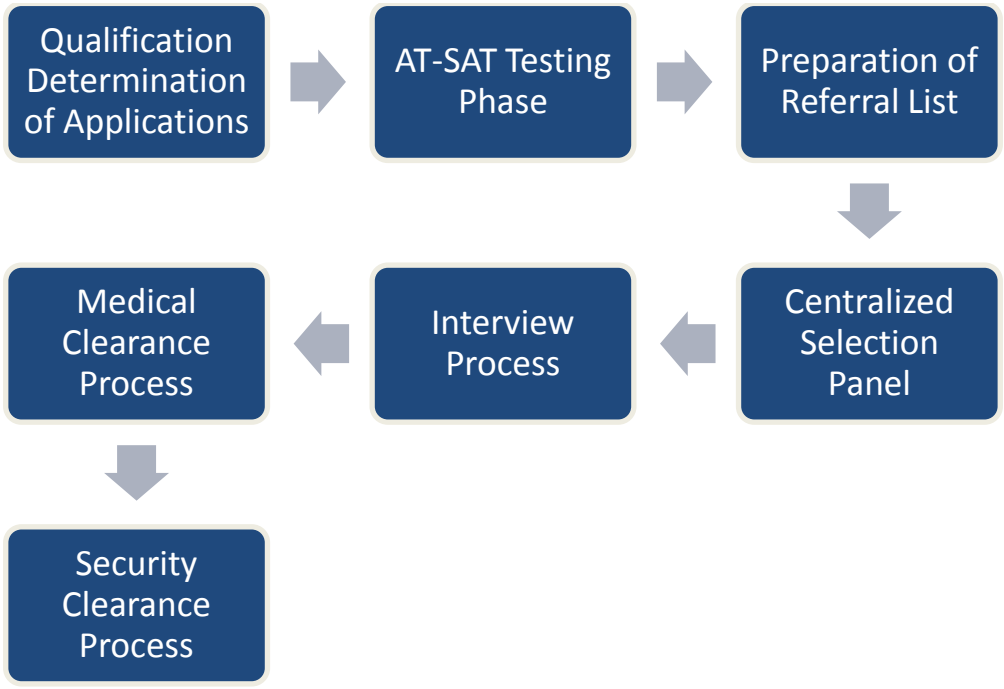
related policies, procedures, and practices to uncover whether a barrier to EEO exists. When an agency uncovers and understands the barrier then appropriate objectives can be implemented to eliminate it.

Lastly, in addition to analyzing workforce statistics, Section II requires agencies to explore a variety of sources to identify potential barriers to equal employment opportunity." [1]

[1] *The U.S. Equal Employment Opportunity Commission Instructions to Federal Agencies for EEO MD-715*

ATCS Centralized Hiring Process

Figure 1 shows the ATCS hiring process.



Based upon our review of the ATCS Centralized Hiring Process, we identified seven critical decision points at which barriers to racial and gender employment opportunities could occur. These seven decision points are:

1. Qualification determination of applications
2. AT-SAT testing phase
3. Preparation of the referral list of eligible and qualified applicant list
4. Centralized Selection Panel determination of applicants to proceed to interview
5. Interview
6. Medical clearance process
7. Security clearance process

We will consider each decision point separately and provide the root-cause analysis of the multivariance step analysis that led to our conclusions and recommendations for each of the identified barriers. To begin, applicants can come from several sources. The qualifications required for the various application sources are as follows:

- Application Source: Retired Military Controllers (RMC)
  - US citizen
  - Open to military air traffic controllers who are either on terminal leave pending retirement from active duty military service or who retired from active duty on or after September 17, 1999
  - Must be able to speak English clearly enough to be understood over radios, intercoms, and similar communications equipment
  - Appointments will not be made or extended past the last day of the month in which an applicant reaches the mandatory separation age of 56
  - Must possess experience in a military or civilian air traffic facility that requires the knowledge, skills, and abilities needed to perform the level of work of the position; this experience must have provided a comprehensive knowledge of appropriate air traffic control laws, rules, and regulations
  - Must have received either an Air Traffic Control Specialist Certification or a facility rating according to FAA standards
- Application Source: Veteran's Recruitment Appointment (VRA)
  - U.S. citizen
  - Under maximum entry age of 31 per Public Law 92-291
  - 52 Consecutive weeks of qualifying air traffic control experience
  - Must be able to speak English clearly enough to be understood over radios, intercoms, and similar communications equipment
  - VRA Eligibility
    - Veterans who have a compensable service connected disability rating of 10% or more; or
    - Veterans who served on active duty in the Armed Forces during a war or in a campaign or expedition for which a campaign badge or expeditionary medal was awarded; or
    - Veterans who served on active duty in the Armed Forces during a military operation for which an Armed Forces Service Medal (AFSM) was awarded on or after June 1, 1992; or
    - Veterans who have separated from active service (for other than training purposes) within the last three years.

    In addition to the above criteria, all veterans must have been discharged under the general or honorable conditions.

- Application Source: Control Tower Operator (CTO)
  - U.S. citizen

- Under maximum entry age of 31 per Public Law 92-291
- Must be able to speak English clearly enough to be understood over radios, intercoms, and similar communications equipment
- Must have a Control Tower Operators (CTO) License with a facility rating of Tower/Cab

- Application Source: Public

  - U.S. citizen
  - Under maximum entry age of 31 per Public Law 92-291
  - Must be able to speak English clearly enough to be understood over radios, intercoms, and similar communications equipment
  - May qualify based upon college education or work experience or a combination of work experience and college credits:
    - To qualify based on college education, applicants must have had a full four-year course of study leading to a Bachelor's degree.
    - To qualify based on work experience, applicants must have had three years of progressively responsible work experience that demonstrates the potential for learning and performing air traffic control work. Progressively responsible experience was described in examples as an increase in leadership duties, working with less supervision, or an increase in financial responsibility. Progression was required to be documented in detail on the application.
    - To qualify based on combination of work experience and college credits, applicants must have one year of undergraduate study (30 semester hours or 45 quarter hours is the equivalent of 9 months of general work experience).
  - After a qualifications review by HR, eligible applicants are required to take and pass the AT-SAT pre-employment test. AT-SAT scoring: 100% to 70% is passing, 100% to 85% is well qualified, 84.9% to 70% is qualified.

- Application Source: Reinstatement-Department of Defense-Certified Professional Controller (DOD CPC)

  - U.S. citizen
  - Reinstatement eligibility refers to applicants who held a career or career-conditional appointment; there is no time limit on reinstatement eligibility for those who either have veteran's preference or acquired career tenure by completing one year of substantially continuous service
  - Must possess 52 consecutive weeks of ATCS experience in a military or civilian air traffic control facility that requires the knowledge, skills, and abilities needed to perform the level of work; this experience must have provided a comprehensive knowledge of the laws, rules, regulations and procedures governing the movement of air traffic, knowledge of aircraft

separation standards and control techniques, and the ability to apply them properly, often under conditions of great stress

- If experience was gained as a civilian ATCS with the DOD, must have been appointed prior to the maximum entry age of 31
- Required to submit proof of reinstatement eligibility and age requirements by providing SF-50 Notification of Personnel Actions or its equivalent indicating appointment into the 2152 series as well as current or separation SF-50 actions.
- Veterans and preference eligibles who are honorably discharged from the Armed Forces after completing at least three years of continuous active service
  - Veteran and Preference eligible: A veteran or preference eligible, or a veteran spouse, widow, or mother who meets the requirements outlined in 5 USC 2108
  - Active Service for Enhanced Veterans Hiring Opportunities (EVHO) eligibility only: Active duty in the uniformed services and includes full-time training duty, annual training duty, full-time National Guard duty and attendance, while in the active service, at a school designated as a service school by law or by the Secretary of the military department concerned
  - Three years of continuous active duty service includes those individuals that were released from active duty a few days before completing a three-year tour if it was for the convenience of the government; the Human Resource Management Office (HRMO) makes this determination
  - Veterans' Preference in hiring does not apply on announcements with an area of consideration of "Current or Former Federal Employees and EVHO" or "Qualified Civil Service Employees and EVHO"; however, applicants who meet the above criteria will be considered in the hiring process; when the area of consideration includes EVHO, a preference eligible or eligible veteran may apply to a vacancy announcement even though he or she may be outside of the area of consideration, for example, an EVHO-eligible candidate is not limited by the commuting area for the purposes of being considered

- Application Source: College Training Initiative (CTI)

  - U.S. citizen
  - Under maximum entry age of 31 per Public Law 92-291
  - Must be able to speak English clearly enough to be understood over radios, intercoms, and similar communications equipment
  - Passed AT-SAT test
  - Received recommendation from the CTI school

- Two-year initial eligibility from graduation date
- Extensions could be requested/granted until maximum entry age is met

DIVERSITY AND APPLICANT SOURCE

Before examining the impact of the ATCS Centralized Hiring Process, we first examined the diversity within the different application sources used by the FAA as input into the hiring system. Prior to assessing the diversity of the applicant pools, the data had to be adjusted to the applicant level of analysis. That is, the data set originally represented applications, not applicants, in that a single applicant could apply multiple times throughout a fiscal year depending upon the number of posted announcements. The RNO and gender analysis that we are reporting required an assessment of individuals (i.e., applicants) and not applications. Thus, we aggregated the data by applicant unique identifier so that the applicants for each fiscal year could be identified. The data presented in Table 1 are a function of this applicant level of analysis. Specifically, Table 1 provides the ethnic and gender composition of the applicants by weighted average of FY 2007 through FY 2011 from the different application sources. The data in Table 1 is based on the weighted average of data from FY 2007 through FY 2011.

Table 1: Subgroup Race/Ethnicity and Gender as a Function of Weighted Average by Application Source

Ethnicity	Weighted Average FY2007–FY2011*											
	Veterans Recruitment Appointment (VRA)		Retired Military Controller (RMC)		Other (CTO)		Public		Reinstatement-DOD CPC		College Training Initiative (CTI)	
	Total	%	Total	%	Total	%	Total	%	Total	%	Total	%
Asian	414	2.1%	318	2.2%	239	2.2%	1136	2.8%	62	2.0%	154	4.6%
African-American	6,646	34.6%	6,100	42.5%	3291	30.8%	12,433	31.3%	844	27.9%	179	5.4%
Hawaiian	141	0.7%	67	0.5%	63	0.6%	274	0.7%	27	1.0%	16	0.9%
Hispanic	1,182	6.0%	849	5.9%	628	5.6%	2,390	5.8%	183	5.8%	219	6.5%
Multi	1,409	7.4%	920	6.4%	797	7.2%	2,920	7.0%	210	7.0%	231	6.9%
Native American	109	0.6%	72	0.5%	66	0.6%	225	0.5%	20	0.7%	9.0	0.3%
Unanswered	1,054	5.6%	707	5.1%	529	4.7%	1,890	4.5%	223	7.2%	191	5.7%
White	8,894	45.5%	5,546	38.8%	5,638	50.7%	20,400	49.8%	1,634	51.9%	2,373	70.4%
Gender												
Female	4,203	21.4%	3,522	24.3%	2,316	20.8%	9,846	23.9%	652	20.4%	577	17.2%
Male	13,321	67.3%	9,278	63.8%	7,595	67.7%	27,151	65.3%	2,114	66.1%	2,481	73.6%
Unanswered	2,325	11.9%	1,779	12.3%	1,340	12.1%	4,671	11.2%	437	13.8%	314	9.7%
Total	19,849		14,579		11,251		41,668		3,203		3,372	

*The weighted averages are based on available data from FY 2007 to FY 2011

Before drawing conclusions from the data in Table 1, three points should be noted. First, the percentage of applicants self-reporting RNO and gender status is high. This supports the premise that the samples reporting RNO and gender data are sufficiently high in relation to the total sample to indicate that there is little, if any, difference between the statistics from the self-report sample and what would be expected from the total sample. This is true even though more

applicants self-reported RNO status than reported gender status. Second, given that the centralized hiring process was not fully implemented until 2007, the data for 2006 include the pre-2007 hiring process together with the ATCS Centralized Hiring Process currently under examination. Thus, we present results from FY 2007 to FY 2011. Finally, we computed weighted averages shown in Table 1 to provide an aggregate picture of the RNO and gender diversity composition in the various applicant sources across the fiscal years. This weighted average is computed using the following formula:

$$\overline{p}_i = \frac{\sum_{t=1}^{k} (p_{ik} f_{ik})}{\sum_{i=1}^{k} f_{ik}}$$

where p_i is the weighted average percentage of applicants in RNO subgroup$_i$ across the fiscal time periods FY 2007 to FY 2011, p_{ik} represents the percentage of applicants in RNO subgroup$_i$ in fiscal year$_k$, and f_{ik} represents the frequency of applicants in the RNO subgroup(i) in fiscal year$_k$. We only included FY 2007 to FY 2011 data because, as indicated previously, the FY 2006 data was a mixture of the old and new hiring systems. The weighted average is a commonly used statistic found in other statistical analyses such as meta-analytic research (e.g., Rosenthal, 1991; Schmidt & Hunter, 1990) as well as in random coefficient modeling (Raudenbush and Bryk, 2002). It is often used to provide an accurate estimate of an effect when there are multiple samples estimating of the effect. The weighted average places more weight on those samples that have the most stability (i.e., the largest samples). Refer to Appendix A, Tables 1A through 1F for the data presented year by year.

APPLICANT SOURCE RESULTS

Examining the demographic information in Table 1, it is clear that the applicant sources vary in terms of the proportion of any one race, ethnic, or gender group. As an example, the proportion of African-Americans by source (based on the weighted average) ranges from a low of 5.4% for CTI to a high of 42.5% for Retired Military Controller. This means that the choice of applicant source can pose a significant barrier based on RNO. For example, Hispanic and female applicants are underrepresented relative to their numbers in the US population, regardless of the source used. Notably there is a higher proportion of Whites than any other applicant group. Moreover, the percentage of males in the AT-CTI applicant source is larger than the other applicant sources. In addition to the variability of RNO and gender representation across sources, it is evident that the FAA has not used all applicant sources throughout the years. Indeed, announcements generating applications from the public source were only posted during FY 2007, FY 2008, and FY 2009. This is the case despite the extensive representation of African-Americans in this application pool (average 31.3%).

This finding raises questions about the RNO and gender diversity of the CTI schools. A recent report titled “Air Traffic-Collegiate Training Initiative Diversity Data 2011-12” dated June 12,

2012, sought to determine the diversity of the potential applicant pool from the partner CTI colleges and universities. The authors of the report sought to establish a baseline that would be used to assess the effectiveness of future efforts to improve the diversity of the AT-CTI program. This report indicated that after combining the data across all colleges, universities, and years in the program: "The predominate race identification was 'White' 58.7% followed by 'Black/African-American' at 12%, 'Asian' 7.1%, and 'American Indian/Alaska Native' and 'Native Hawaiian/Other Pacific Islander' at less than 1%" (p. 4). The report indicated that Hispanic/Latino composed 17.8% of the AT-CTI student population. The conclusion of the report indicated that "There is a 3:1 ratio of male to female students pursuing degrees that qualify for application under the CTI announcement. The reporting of categories other than 'White' is 41% among the respondents" (p. 5).

Examining "Air Traffic-Collegiate Training Initiative Diversity Data 2011-12" report in greater detail reveals several issues that need highlighting. First, based on the information in the report on the four-year programs, there is a dropout of African-Americans from freshman year (16.9%) to the senior year (6.0%). This dropout rate of African-Americans from the program is troubling. There is not a similar dropout rate for other subgroups. We recommend that the FAA undertake efforts to understand the cause for this selective dropout rate and take reasonable countermeasures to reduce it in the future. The FAA should investigate this problem by contacting four-year CTI schools and work with them to identify the possible causes of the dropout. Second, it is important to note that while "Air Traffic-Collegiate Training Initiative Diversity Data 2011-12" estimated the percentage of African-Americans in the CTI schools to be 12%, that percentage was determined by collapsing all students, regardless of their matriculation credits.

For purposes of the FAA diversity assessment, the more important statistic is to examine the percentage of students from each ethnic subgroup who are in their senior year in their four-year programs. Specifically, 6% of the CTI seniors from universities are African-American, which more closely corresponds to the 2011 CTI applicant source results that we reported in Table 1F, Appendix A. Similar data for Whites show that they make up 58.7% of the CTI Seniors from four-year universities. Finally, it would be helpful in future reports for the RNO and gender results to be reported at the school level as well as the individual level of analysis. A school-level analysis will identify diversity problems at the college or university level. This issue is not addressed in the analysis reported in "Air Traffic-Collegiate Training Initiative Diversity Data 2011-12.

Qualities of Different Applicant Sources

Another potential barrier is the decision as to the pool from which applicants are drawn. The definition of the applicant pool affects the extent to which different RNO and gender groups are able to pass the minimum qualification criteria. Table 2 shows the proportion of applicants passing the qualifications stage as a function of application source specified by RNO and gender subgroup and fiscal year. Individual RNO and gender groups per year can be found in Tables 10A through 10F in the Auxiliary Tables Section (Appendix A). As can be seen from Table 2, there are differences in the proportion of applicants with minimum qualifications as a function of application source. The AT-CTI application source has the applicant pool with the highest proportion of applicants meeting the minimum qualification across ethnic subgroups. The AT-CTI applicants passed the minimum qualifications criteria from a low of 98.6% (Hispanics/Latinos) to a high of 99.6% (Whites). However, the next application source with the highest proportion of applicants passing the minimum qualifications was the General Public. The General Public applicants passed the minimum qualifications from a low of 74.7% (African-Americans) to a high of 86.5% (Whites). In contrast, the application source with the lowest proportion of minimum qualifications was the Retired Military Controller. This application source ranged from 4.2% (African-Americans) to 14.7% (Whites) minimum qualification pass rate.

Table 2: Proportion of Applicants Passing the Qualification Stage as a Function of Applicant Source

Race/Ethnicity	Weighted Average FY2007–FY2011*											
	Veterans Recruitment Appointment (VRA)		Retired Military Controller (RMC)		Other (CTO)		Public		Reinstatement-DOD CPC		College Training Initiative (CTI)	
	Total	%	Total	%	Total	%	Total	%	Total	%	Total	%
Asian	414	24.9%	318	7.2%	239	61.5%	1,136	84.2%	62	33.9%	154	99.3%
African-American	6,646	9.6%	6,100	4.2%	3,291	49.4%	12,433	74.7%	844	9.3%	179	99.5%
Hispanic	1,182	28.7%	849	9.8%	628	58.3%	2,390	78.5%	183	24.6%	219	98.6%
Multi	1,409	28.4%	920	7.4%	797	64.9%	2,920	82.8%	210	28.6%	231	99.1%
Race Not Identified	1,054	34.4%	707	11.3%	529	70.5%	1,890	84.9%	223	51.6%	191	99.5%
White	8,894	35.5%	5,546	14.7%	5,638	70.9%	20,400	86.5%	1,634	51.2%	2,373	99.6%
Gender												
Female	4203	19.0%	3,522	4.3%	2316	59.3%	9,846	82.0%	652	25.5%	577	99.3%
Male	13,321	27.5%	9,278	10.9%	7,595	64.6%	27,151	82.1%	2,114	38.5%	2,481	99.5%
Gender Not Identified	2,325	27.1%	1,779	10.0%	1,340	61.8%	4,671	81.6%	437	45.5%	314	99.7%

*The weighted averages are based on available data from FY 2007 to FY 2011

The critical point is that the application sources also differ substantially in terms of the percentage of applicants from the different racial/ethnic and gender groups within it. As an example, in the Public Source category the different subgroup percentages are Asian at 2.7%, African-American at 29.8%, Hispanic at 5.7%, Multiracial at 7.0%, White at 48.9%, and race unidentified at 4.5%. On the other hand, the percentage of each applicant group in the AT-CTI source category is Asian at 4.5%, African-American at 5.3%, Hispanic at 6.5%, Multiracial at 6.8%, White at 70.3%, and race unidentified at 5.7%. These figures show that if the public source is used, then 29.8% of the applicants entering this step are African-American. However, if the AT-CTI source is used, only 5.3% of the applicants entering step 1 are African-American. Given that the total numbers of applicants entering step 1 via the Public and AT-CTI are 41,688 and 3,372, respectively, using the AT-CTI pool severely reduces the diversity of the applicant pool at step 1 for African-Americans. Thus a policy that reduces the use of the General Public Source in favor of the AT-CTI source is acting as a barrier to African-Americans at the applicant source step in the hiring process. This is not the case for other subgroups.

Our finding that the General Public application source has a diverse RNO application pool as well as the second highest pass-rate of minimum qualifications is at odds with FAA policies such as not posting ATC jobs to the General Public since 2009 as well as comments from trainers such as "Please tell them not to send me any more public hires" (Barr et al., 2011, p. 16). The Barr et al. (2011) report interpreted this quote and similar data as indicating that the FAA "needs to review its hiring practices to take advantage of the AT-CTI system it has created" (p. 16). We question these policies and conclusions. Given the limited number of applicants and especially the diminished diversity in the AT-CTI application source, there are serious diversity consequences for not fully using the General Public application source.

When asked in the "Policy Questions and Answers" document (July 2012) how it was decided which application source would be used when an announcement is posted, the agency indicated that the determination as to which hiring sources will be used is made jointly by the Air Traffic Organization (ATO) and the Office of Human Resource Management (AHR).

All the sources except General Public require the applicants to have ATC training or experience. The premise for having the training and experience requirement is that the content is similar to that in the job. This is a content validation approach. This is despite the fact the FAA has a training school in place to train applicants and describes the position as requiring a high school diploma, three years of general work experience, or a combination of college and general work experience. The policy of requiring applicants to come into the job with skills for which the agency provides training violates the content validity provisions of the Uniform Guidelines on Employee Selection Procedures.

Lack of systematic consideration of RNO and gender diversity when making decisions regarding the applicant source for ATCS Centralized Hiring Process is a barrier. Why is the General Public announcement not issued unless there is likely to be a shortage of qualified applicants from other sources, given that this group has the second highest passing rate on minimum qualifications? It yields more diverse applicants than any other source. Thus, a root

cause of differences in qualification determinations by RNO is the lack of systematic consideration of RNO and gender diversity when determining the applicant sources. The application source is a barrier for women and minorities if the Public Source is not posted for applicants to apply and be considered for the ATCS position for each announcement.

ROOT-CAUSE ANALYSIS OF THE SEVEN DECISION POINTS IN THE ACTS CENTRALIZED HIRING PROCESS

DECISION POINT 1: QUALIFICATION DETERMINATION OF APPLICANTS

Our analysis indicates that the qualifications for the ATCS position vary greatly by the applicant source. The first step in the ATCS Centralized Hiring Process is the Qualification Determination. This step involves verifying whether the qualifications of the applicants match the minimum qualifications specified in the job announcements. The determination of whether someone matches the minimum qualifications for the position is either a system decision (i.e., the applicant is missing an objective qualification) or it is a decision made by HR personnel (i.e., a subjective decision regarding whether the person has three years of progressive work experience or the requisite combination of experience and education).

Race/Ethnicity: For each year from 2006 to 2011, we analyzed the qualification decision rates made as a function of the RNO and gender subgroups that composed more than two percent of the population of applicants. After reviewing Table 1, it was determined that Native American/American Indian/Alaska Native and Native Hawaiian/Other Pacific Islander subgroups would be excluded from all subsequent analyses in this report due to their less than two percent representation in the applicant population. The decision to exclude these two subgroups is consistent with the analysis recommendations of the Uniform Guidelines issued by the EEOC, Civil Service Commission, Department of Labor, and the Department of Justice in 1978.

Table 3 shows the proportion of applicants who were identified as qualified by RNO subgroup for each fiscal year. As can be seen in this table, the average qualified decision varies substantially as a function of self-reported RNO subgroup and year.

Table 3: Percentage of Applicants Passing the Qualification Stage of the ATCS as a Function of Race/Ethnicity

Race/Ethnicity		Qualified						
		2006	2007	2008	2009	2010	2011	Weighted Average 2007-2011+
Asian	Mean	0.45	0.61	0.68	0.73	0.56	0.78	0.68
	N	74	414	688	559	153	118	386.40
	Std. Deviation	0.50	0.49	0.47	0.44	0.50	0.42	0.47
African-American	Mean	0.50	0.48	0.46	0.52	0.29	0.59	0.47
	N	412	7470	9177	5956	1270	392	4853.00
	Std. Deviation	0.50	0.50	0.50	0.50	0.45	0.49	0.50
Hispanic	Mean	0.60	0.56	0.59	0.65	0.54	0.83	0.60
	N	227	1129	1571	1231	301	169	880.20
	Std. Deviation	0.49	0.50	0.49	0.48	0.50	0.37	0.49
Multi	Mean	*	0.58	0.63	0.71	0.57	0.81	0.65
	N	*	1147	1846	1502	374	295	1032.80
	Std. Deviation	*	0.49	0.48	0.46	0.50	0.39	0.47
White	Mean	0.55	0.66	0.73	0.77	0.68	0.88	0.73
	N	1707	7669	11776	10909	2819	2172	7069.00
	Std. Deviation	0.50	0.47	0.44	0.42	0.47	0.33	0.44

Note: * Multiracial individuals composed less than two percent of the applicant pool in 2006, and thus consistent with the 1978 Uniform Guidelines, were not analyzed for that year.

+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation ("N" stands for sample size).

Table 4 provides the results of statistical comparisons between the Table 3 qualification rates for the minority subgroups compared to the White majority subgroup. Excluding the 2006 results for reasons mentioned earlier in this report, as well as the single result for Hispanics/Latinos in 2011, there were significant differences in the qualification rates for all RNO subgroups shown in Table 2 when compared to the White majority group. In other words, the minority subgroups were classified as qualified at a significantly lower rate than the White majority group was.

Table 4: Statistical Test for Differences in Qualified Decisions as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Proportions for Qualified Decisions						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Asian	1.74	1.88	2.81**	2.45**	3.19**	3.11**	2.69**
African-American	1.86	22.25**	40.22**	34.06**	23.47**	14.15**	26.83**
Hispanic	-1.31	6.50**	11.70**	9.98**	4.83**	1.66	6.93**
Multi	N/A	5.10**	8.87**	5.96**	4.24**	3.26**	5.49**

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of the qualified rate being greater for the White sample.

Finally, we provide two additional tables to help understand the magnitude of this effect. While a significant effect does not mean that the effect is necessarily meaningful in a practical sense, an estimate of the effect's magnitude must be taken into account to understand the importance of the finding (Hanges & Wang, 2012).

Table 5 shows the traditional adverse impact ratio used in fair employment practices cases. The adverse impact ratio (AI) is computed by dividing the proportion deemed qualified for a minority subgroup (e.g., one of the Table 1 qualification rates for a particular RNO minority subgroup) by the proportion for the majority subgroup. For this analysis, we choose the White subgroup as the majority group due to its percentage representation in the sample. Specifically, the adverse impact ratio is computed as follows:

$$AI = \frac{p_{\min}}{p_{majority}} \qquad (2)$$

Table 5: Adverse Impact Ratios of Minority Qualified to White Majority Qualified

Race/ Ethnicity	Adverse Impact Ratios (4/5ths Rule) for Difference in Proportions for Qualified Decisions						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Asian	0.81	0.93	0.93	0.94	0.82	0.89	0.90
African-American	0.91	0.73	0.63	0.67	0.42	0.67	0.62
Hispanic	1.08	0.85	0.81	0.83	0.80	0.95	0.85
Multi	N/A	0.88	0.86	0.91	0.84	0.92	0.88

Note: The White population was used as the majority comparison group.

In the formula above, p_{min} represents the percentage of a particular minority subgroup rated as qualified and $p_{majority}$ represents the percentage of the majority subgroup rated as qualified. AI ratios lower than 80% are considered indicative of underrepresentation or barriers in this instance (Hanges, Salmon, & Aiken, in press). In Table 5, only the AI ratio for the African-American subgroup is consistently below the 0.80 cutoff from FY 2007 to FY 2011.

Another statistic used to understand the magnitude of the difference between two groups is to compute what is known as the effect size or d-ratio. This is a common statistic used in the scientific literature. The d-ratio is computed by using the following formula:

$$d = \frac{\left(p_{\min} - p_{majority}\right)}{\sqrt{\left(\frac{p_{\min}(1-p_{\min})}{n_{\min}}\right)\left(\frac{p_{majority}(1-p_{majority})}{n_{majority}}\right)}} \quad (3)$$

Similar to Formula 2, p_{min} and $p_{majority}$ represent the percentage of minority subgroup and percentage of majority subgroup rated as qualified, respectively. Finally, n_{min} and $n_{majority}$ are the number rated as qualified. The d statistic is computed by dividing the difference in proportions by the pooled standard deviation for the two compared ethnicities. Table 6 shows these d-values as a function of RNO subgroups.

The benefit of using d-ratios is that there are standards that have been proposed to help researchers and practitioners interpret their magnitude. In particular, d-values are considered small if they are less than 0.30; they are considered moderate if they are close to 0.50; and they are generally considered large if they are close to 0.80 or above (Cohen, 1988). Smaller d-ratios values may be considered trivial whereas larger d-values are more problematic. This, however, is not always the case. Even small durations can indicate significant practical effects. As an example, the overall selection ratio (i.e., number of vacancies divided by the number of applicants) is very low. This means there are typically very few openings and too many applicants. In this situation even small differences in the selection rates for various applicant groups can have meaningful negative consequences and constitute a barrier. For this reason we used three criteria to determine what constitutes a barrier. Those criteria are based on the weighted average:

- A statistically significant difference between a minority or gender group and the majority group (in this instance Whites and males)
- An effect size of 0.20 or higher
- An adverse impact ratio below 0.80

If a selection step or component of the hiring process meets at least two of the three criteria above, we consider it a barrier. Examining the Tables 3 through 6, it can be seen that qualification decisions for African-Americans meet all three criteria therefore, constitute a barrier for that group.

Table 6: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for Qualification Decisions						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Asian/White	0.21	0.10	0.11	0.11	0.27	0.30	0.18
African-American/White	0.10	0.37	0.58	0.57	0.85	0.81	0.64
Hispanic/White	-0.09	0.21	0.32	0.30	0.30	0.13	0.25
Multi/White	N/A	0.16	0.22	0.16	0.24	0.20	0.20

Gender: A similar analysis was conducted for gender. Table 7 shows the proportion of qualified applicants as a function of gender per year. This table also provides a weighted average of qualified applicants as a function of gender across FY 2007 to FY 2011. As can be seen in this table, the average qualified decision varied over the years. Table 8 shows the results from our statistical comparison of the female-to-male qualification decisions. As shown in this table, there were differences in the qualification determination as a function of gender. Table 9 shows the AI ratios as a function of gender for each year for the qualification decisions. Table 10 presents the d-ratio for this comparison. As can be seen from Tables 7 through 10, the qualification determination is not a barrier based on gender.

Table 7: Proportions of Applicants Passing the Qualification Stage of the ATCS as a Function of Gender

Gender		Qualified						
		2006	2007	2008	2009	2010	2011	Weighted Average 2007–2011+
Female	Mean	0.49	0.56	0.59	0.66	0.50	0.82	0.60
	N	527	4,960	6,407	4,525	953	557	3,480.40
	Std. Deviation	0.50	0.50	0.49	0.47	0.50	0.39	0.49
Male	Mean	0.56	0.58	0.63	0.70	0.58	0.84	0.65
	N	1,940	1,1780	17,085	14,427	3,660	2,535	9,897.40
	Std. Deviation	0.50	0.49	0.48	0.46	0.49	0.37	0.47

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 8: Statistical Test for Differences in Proportions for Qualification Decisions as a Function of Gender

Gender	Statistical Test for Difference in Proportions for Qualified Decisions						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Female	2.90**	2.63**	5.75**	5.12**	4.10**	1.20	3.76**

Notes: * $p < 0.05$, ** $p < 0.01$
The male population was used as the majority comparison group. Positive Z-scores are indicative of the qualification rate being greater for the male sample.

Table 9: Adverse Impact Ratios Qualification Decisions as a Function of Gender

Gender	Adverse Impact Ratios (4/5ths Rule) for Qualified Decisions						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Female	0.87	0.96	0.93	0.94	0.87	0.97	0.93

Note: The male population was used as the majority comparison group.

Table 10: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for Qualification Decisions						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Female/Male	0.14	0.04	0.08	0.09	0.15	0.06	0.08

ROOT-CAUSE ANALYSIS OF BARRIERS AT DECISION POINT 1

Analyses revealed that there were systematic reductions in RNO diversity at the first stage of the ATCS Centralized Hiring Process for African-Americans, Hispanics, and Multiracial subgroups. The question to answer now is why this occurs. Our analysis indicates several possible root causes.

Nonstandardization of training for HR Specialists

Interviews of HR managers as well as a review of procedures, policies, and practices indicate that the HR specialists who review the applications do not receive sufficient training to perform their decisions. According to the "Policy Questions and Answers" document (July 2012):

> The Human Resource Specialists responsible for determining the qualifications determination for the Air Traffic Control Specialist (ATCS) position have received formal and informal training. The formal coursework that HR Specialists attends includes Basic Staffing from the United States Department of Agriculture (USDA) Graduate School and Delegated Examining Unit (DEU) training from the Office of Personnel Management. Informal training consists of information provided by Air Traffic Organization (ATO) subject matter experts and on-the-job training from HR Team Leads and experienced HR Specialists.

This indicates that the HR specialists may have a sufficient background in the academic side of staffing as well as specific FAA policies with regard to staffing; however, they do not receive sufficient training or practice in developing a common frame-of-reference to apply similar standards in a similar manner when making subjective assessments of applicant qualifications. As an example, one element of the qualifications determination is whether applicants have three years of work experience. Interviews with a number of people, including the current HR director responsible for managing the field HR managers, indicate that based on their experience at OPM, some HR specialists define the qualification standard as three years of general work experience while others apply a standard of three years of progressive work experience.

When asked to indicate how HR specialists evaluated whether applicants met the 3-year experience requirement, the director of the HR Management Office and the manager of the Aviation Career Division cited in the "Policy Questions and Answers" document (July 2012), making the following statement:

> The 3 years of experience is interpreted as a total of 36 months. The "progressively responsible" aspect was applied by assuming an individual who worked at the same job for 3 years received progressively more responsibility during their tenure with that employer. If the individual had a variety of different types of employment our HR

> Specialist would look for a pattern of increasing responsibility over the 3 year period among the different jobs. The practice described in this response is consistent with OPM policy.

Some of our interviews raised questions regarding the consistency with which the three years of progressively responsible experience standard is being applied. We see this inconsistency as a definite barrier. The variability in the standards used to assess applicants allows for errors and biases to enter the hiring process. Psychological research has shown that the more inconsistency in an evaluation process, the greater the likelihood that bias can occur (Roch et al., 2012).

DECISION POINT 2: AT-SAT TESTING PHASE

Following the Qualification Determination, applicants from the General Public and the CTI sources must complete the AT-SAT test. Applicants from other hiring sources do not take the AT-SAT test, but move to the next step of the hiring process, the preparation of the referral list of eligible and qualified applicants by state (See Figure 1). The AT-SAT consists of eight subtests. Some of these subtests are composed of multiple scales. Twenty-three AT-SAT scales are standardized, weighted, combined, and transformed into the final AT-SAT composite score.

First we examine the RNO and gender composition of the AT-SAT applicants in the highly qualified band followed by the qualified band. Table 11 shows the percentage of applicants in the highly qualified band by RNO. As can be seen from this table, the percentage of African-Americans passing at the highly qualified level is substantially lower than the percentage of Whites. The same is true for Hispanics/Latinos compared to Whites. There is evidence that this highly qualified band is resulting in substantial adverse impact against RNO minorities.

Table 11: Percentage of Ethnic Minorities and Gender Subgroups Passing the AT-SAT Composite Score When Cut Score Is 85 or Better (Highly Qualified)

	Demographic Characteristics of Applicants Scoring ≥ 85 on the AT-SAT Composite					
Race/Ethnicity	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**
Asian	0.45	0.67	0.55	0.55	0.74	0.78
African-American	0.22	0.34	0.29	0.37	0.36	0.38
Hispanic	0.17	0.41	0.41	0.48	0.58	0.41
Multi	0.33	0.53	0.56	0.63	0.55	0.77
White	0.53	0.67	0.67	0.68	0.74	0.78
Gender						
Female	0.32	0.38	0.38	0.42	0.60	0.59
Male	0.52	0.62	0.60	0.66	0.71	0.78
Percent Highly Qualified	0.55	0.57	0.55	0.61	0.64	0.69

Table 12 shows the percentage of applicants in the qualified band by RNO. As can be seen from this table, there are only minor differences in passing rates using the 70 pass score, the qualified band. These findings for the pass score of 70 are consistent with the findings of King, Manning, and Drechsler (2007) and Dattel and King (2006). The AT-SAT does not show adverse impact when the cut score is set at 70. However, that is probably due to the fact that, on average, approximately 95.3% of all applicants pass the exam. Results indicate that the AT-SAT is not useful for helping the FAA make the application process manageable.

Table 12: Percentage of Ethnic Minorities and Gender Subgroups Passing the AT-SAT Composite Score When Cut Score is 70 or Better (Qualified)

	Demographic Characteristics of Applicants Scoring ≥70 on the AT-SAT Composite					
Race/Ethnicity	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**
Asian	1.00	0.96	0.93	0.97	1.00	1.00
African-American	0.87	0.87	0.83	0.88	0.96	0.90
Hispanic	1.00	0.88	0.88	0.92	1.00	0.86
Multi	0.94	0.98	0.94	0.94	0.98	0.96
White	0.97	0.97	0.97	0.97	0.99	0.99
Gender						
Female	0.93	0.89	0.89	0.92	0.99	0.94
Male	0.96	0.96	0.94	0.96	0.99	0.99

When asked in the "Policy Questions and Answers" document submitted by the authors to the FAA (July 2012) for the rationale behind the highly qualified score on the AT-SAT, Civil Aerospace Medical Institute (CAMI) responded as follows:

> The 70 cutoff was used as the passing score. Given the large number of applicants, this results in thousands of applicants scoring 70 and above who must be selected through a merit-based process.
>
> The FAA had an alternative approach available which allowed selecting officials to consider the total individual, including such important factors as motivation, interest in the job, prior experience, etc. This approach was called category ranking. In category ranking, the candidates are sorted into two groups, qualified and well-qualified, in this case on the basis of their AT-SAT score. Selecting officials first select from among candidates in the well-qualified group.

To identify the dividing line between qualified and well-qualified, the AT-SAT contractor was asked to re-scale the test scores so that certain scores corresponded to cut points based on the applicants' predicted job performance. The measure of job performance was the composite peer

and supervisory rating used in the concurrent validation of AT-SAT. The composite peer and supervisory rating was the only validation criterion that included qualitative statements about whether the job performance was below average, average, or above average. The cutoff of 85 corresponds to those whose job performance was predicted to be above average. Those scoring below 85, but at least 70, have predicted job performance that is average.

Examining the information provided regarding the revision of the AT-SAT yielded the following results. First, thousands of people were passing the AT-SAT using the 70 passing score. This created an administrative burden on the centralized hiring process that the use of the highly qualified band was designed to eliminate. The rescoring of the AT-SAT was undertaken not to directly adjust the magnitude of the applicants passing the AT-SAT, but rather to address a problem with the original weighting scheme. Namely, only 62% of fully certified controllers would pass the test under the original weighting scheme, meaning 38% of good controllers would fail the test. The reweighting directly corrected this problem and as a result, the pass rate using a cut score of 70 for certified controllers was 98.8%. Second, in terms of future applicants, the report on the reweighting of the AT-SAT indicated that the expected pass rate would be 67.5%, whereas the report indicated that the expected rate of future applicants scoring 85 or better would be 19.2%. Comparing these expected rates to the rates shown in Tables 11 and 12 indicates that something is wrong with the AT-SAT. The current pass-rate of the AT-SAT averages around 95.3% and the percentage of people scoring 85 or higher on the test is averaging 60.2%. More troubling, there is evidence that White applicants scoring 85 or higher on the AT-SAT over the last three years have generally been increasing (see Table 11).

The departure from the predicted pass rates, in addition to the growth rate of applicants falling into the highly qualified band, is atypical and leads to a conclusion that the test probably has been compromised in some fashion. During our investigation, we found cases of items similar to actual test items being available as well as descriptions and suggestions for taking the AT-SAT on the StuckMic.com website. In addition, there is a published book that provides software to help readers gain experience with similar types of items on a computer (e.g., Mattson, 2006). The reviews by people who have taken the AT-SAT have indicated that they believe that the Mattson (2006) book's information about the content of the AT-SAT was helpful as was practice on some parts of the AT-SAT (e.g., AT Scenarios) but not others (e.g., Letter Factory). Overall, this raises issues of differential availability of this resource material as a function of previous connections with the FAA (e.g., StuckMic.com website) or resources. The possibility that the AT-SAT test has been compromised can explain the distortions in the distribution and the substantial growth in the distribution of the highly qualified band that occurred over the last three years in a test that supposedly measures general cognitive ability.

Finally, the FAA practice is that all applicants in the highly qualified band are reviewed and considered before any applicants in the qualified band are considered. Unfortunately, this practice is highly questionable given that it creates adverse impact and that there is no evidence that an applicant who scores in the highly qualified band on the AT-SAT is substantially superior to an applicant who scores in the qualified band.

Even more troubling, the criterion measure used in the corrected validation study seems to overlook the fact that the FAA trains applicants who are selected for the position, so training performance should also have been considered a criterion measure.

Our analysis indicates the policy of considering applicants in the highly qualified band before considering applicants in the qualified band is a root cause is the AT-SAT proves to be a barrier to racial and ethnic minority applicants, which is addressed later in this report.

When asked in the "Policy Questions and Answers" document (July 2012) why the AT-SAT was not updated, CAMI responded:

> The AT-SAT test was changed several times to update it for changes in the computer hardware and software environment and for computer programming issues. Although there is no evidence that the test is in any way out of date, there are research projects underway to replace the Experience Questionnaire and to add an Auditory Attention Task. We have been in competition with other research priorities for the resources needed for this work. It takes time, research participants, and access to ATCSs to complete these projects.
>
> The proposed replacement for the Experience Questionnaire is a Five Factor Model of Personality measure developed by the US Air Force. This measure, the Self-Description Inventory+ (SDI+), has been shown to have equivalence with the Goldberg Adjective Checklist (Goldberg, 1992). The SDI+ was included in the CoVATCH test battery; we will test whether it has greater predictive validity than the Experience Questionnaire when the contractor delivers the de-identified data set to the FAA.
>
> Additionally, we will conduct fairness analyses to determine what, if any, adverse impact the SDI+ has. The Auditory Attention Task will test the ability to track an auditory input while performing a visual tracking task. The Auditory Attention Task is still in development, but we hope to begin pilot testing this fall.
>
> An alternative version of the test was developed and has been held in reserve in case of a major test security breach. Experience with the old written test suggested that a breach was a significant possibility. The alternative version probably requires some software updating at this point. It was developed some time ago.

ROOT-CAUSE ANALYSIS OF OVERALL COMPOSITE AT-SAT

In addition to examining the AT-SAT bands, we also examined the transformed composite scores on the AT-SAT. While it is traditionally assumed that the Centralized Selection Panel participants do not have access to this information, our interviews with the Centralized Selection Panel participant revealed that applicants occasionally put their AT-SAT score on their application. In other words, the Centralized Selection Panel members do occasionally see an applicant's score. This practice can act as a barrier because Whites tend to have higher scores than African-Americans and all other racial groups. The applications should be screened so this information is not available for consideration. Indeed, more than one Centralized Selection Panel participant indicated that they encouraged applicants to provide their AT-SAT score on the application. The probability of this happening increases as the applicant's AT-SAT score gets closer to 100. Given that the Centralized Selection Panel members do see some overall composite scores of the AT-SAT, we examined the overall composite scores of the AT-SAT to understand RNO and gender differential of this test.

Race/Ethnicity: Table 13 shows the yearly AT-SAT composite score means and standard deviations for each racial/ethnic subgroup that composed more than two percent of the applicant population. Table 14 shows the results of the statistical significance test of the difference in means comparing each minority racial/ethnic subgroup to the majority White subgroup. As shown in Table 13, the AT-SAT Composite score means are significantly lower for African-Americans, Hispanics/Latinos, Asians, and Multiracial subgroups compared to Whites. The effect size of these subgroup differences are shown in Table 15. As shown in Table 15, the AT-SAT composite score has a large effect on African-Americans, a medium-to-large effect on Hispanics/Latinos, and a small effect on Asians and Multiracial subgroups. The AT-SAT meets all three criteria established to identify barriers. Therefore we conclude that it is a barrier to Asians, African-Americans, Hispanics/Latinos, and persons from Multiracial groups.

Table 13: Descriptive Statistics of the AT-SAT Composite Score as a Function of Race/Ethnicity

Race/Ethnicity		AT-SAT Composite Score						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Asian	Mean	*	87.55	84.91	86.13	90.21	*	86.13
	N	*	79	207	197	34	*	129.25
	Std. Deviation	*	8.99	8.87	8.28	7.26	*	8.59
African-American	Mean	79.14	80.43	79.37	81.25	83.30	*	80.15
	N	23	494	1,670	945	47	*	635.80
	Std. Deviation	8.74	9.01	9.46	9.54	8.99	*	9.41
Hispanic	Mean	*	82.46	81.76	83.80	85.65	*	82.79
	N	*	145	425	310	60	*	235.00
	Std. Deviation	*	9.88	9.56	9.20	7.50	*	9.39
Multi	Mean	82.13	86.01	85.28	86.37	86.05	*	85.76
	N	18	139	508	422	56	*	228.60
	Std. Deviation	8.97	9.07	9.03	9.15	7.95	*	9.03
White	Mean	85.48	87.92	87.87	88.39	89.52	*	88.14
	N	155	1,054	4,215	4,066	545	*	2,007.00
	Std. Deviation	8.07	8.45	8.41	8.38	7.39	*	8.35

Note: * Asian and Hispanic applicants composed less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year. In 2011, no racial or ethnic group comprised more than two percent of the applicant pool.

+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 14: Statistical Test for Differences in Scores for the AT-SAT Composite as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Differences in Mean Scores on the AT-SAT Composite						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian	N/A	0.35	4.70**	3.74**	-0.54	N/A	2.59**
African-American	3.28**	15.55**	32.08**	21.19**	4.61**	N/A	19.16**
Hispanic	N/A	6.34**	12.69**	8.52**	3.80**	N/A	8.36**
Multi	1.51	2.35**	6.15**	4.35**	3.13**	N/A	3.80**

Notes: * $p < 0.05$, ** $p < 0.01$

The White population was used as the majority comparison group. Positive Z-scores are indicative of higher AT-SAT composite scores in the White sample.

Table 15: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for AT-SAT Composite Score						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian/White	N/A	0.04	0.35	0.27	-0.09	N/A	0.24
African-American/White	0.78	0.87	0.97	0.83	0.83	N/A	0.93
Hispanic/White	N/A	0.63	0.72	0.54	0.52	N/A	0.63
Multi/White	0.41	0.22	0.31	0.24	0.47	N/A	0.28

Gender: Table 16 shows the yearly AT-SAT composite score means and standard deviations as a function of gender. Table 17 shows the statistical significance test of the difference in female/male AT-SAT composite means. There are significant female/male differences on the AT-SAT composite score over time. Finally, Table 18 shows the effect size of these gender differences over time. The AT-SAT Composite score has a medium effect on gender. Therefore, we conclude that the AT-SAT is a barrier based on gender.

Table 16: Descriptive Statistics of the AT-SAT Composite Score as a Function of Gender

Gender		AT-SAT Composite Score						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Female	Mean	82.27	81.82	81.83	82.77	86.75	85.84	82.42
	N	44	509	1,668	1,199	149	66	605.83
	Std. Deviation	7.73	9.04	9.27	8.94	7.77	8.31	9.04
Male	Mean	84.91	86.85	86.33	87.85	89.04	90.52	87.19
	N	160	1,318	4,981	4,445	568	341	1,968.83
	Std. Deviation	8.84	9.03	9.31	8.82	7.68	7.690	8.97

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 17: Statistical Test for Differences in Scores for the AT-SAT Composite as a Function of Gender

Gender	Statistical Test for Difference in Scores for the AT-SAT Composite						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female	1.94	10.67^{**}	17.14^{**}	17.51^{**}	3.21^{**}	4.24^{**}	11.39^{**}

Notes: * $p < 0.05$, ** $p < 0.01$
The male population was used as the majority comparison group. Positive Z-scores are indicative of higher AT-SAT composite scores in the male sample.

Table 18: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for AT-SAT Composite						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female/Male	0.31	0.56	0.48	0.58	0.30	0.60	0.53

To understand the source of the adverse impact produced by the test and its potential contribution as a barrier to RNO and gender diversity was examined.

AT-SAT SCALES: DIALS

The Dials scale measures how quickly and accurately applicants can read dials on an instrument panel. The applicant must quickly identify the correct reading for a specific dial from a list of five possible answers.

Race/Ethnicity: Table 19 shows the scale means and standard deviations for each racial/ethnic subgroup that composed more than two percent of the applicant population. Table 20 shows the results of the statistical significance test of the difference in means comparing each minority racial/ethnic subgroup to the majority White subgroup. As shown in Table 20, the scale's means are significantly lower for African-Americans, Hispanics/Latinos, and Multiracial subgroups compared to Whites. The effect sizes for these subgroup differences are shown in Table 21. As shown in this table, this scale has a large effect on African-Americans, a medium-to-large effect on Hispanics/Latinos, and a small effect on Asians and Multiracial subgroups. Tables 20 and 21 show that the Dials subtest is a barrier to African-Americans, Hispanics/Latinos, and persons from Multiracial groups.

Table 19: Descriptive Statistics of the Dials Subtest as a Function of Race/Ethnicity

Race/Ethnicity		Dials Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Asian	Mean	*	19.00	18.83	18.81	19.24	*	18.88
	N	*	79	207	197	34	*	129.25
	Std. Deviation	*	1.38	1.49	1.33	0.96	*	1.39
African-American	Mean	16.13	17.61	17.37	17.53	17.72	*	17.45
	N	23	494	1,670	945	47	*	635.80
	Std. Deviation	2.93	2.39	2.60	2.56	2.29	*	2.56
Hispanic	Mean	*	17.78	17.83	18.11	18.82	*	17.98
	N	*	145	425	310	60	*	235.00
	Std. Deviation	*	2.68	2.31	2.00	1.71	*	2.24
Multi	Mean	18.06	18.36	18.32	18.50	18.50	*	18.39
	N	18	139	508	422	56	*	228.60
	Std. Deviation	2.01	2.18	1.78	1.84	1.57	*	1.85
White	Mean	18.66	18.90	18.80	18.89	19.12	*	18.86
	N	155	1,054	4,215	4,066	545	*	2,007.00
	Std. Deviation	1.33	1.27	1.40	1.33	1.13	*	1.34

Note: * Asian and Hispanic applicants composed less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year. In 2011, no racial or ethnic group comprised more than two percent of the applicant pool.

+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 20: Statistical Test for Differences in Scores on the Dials Subtest as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Mean Scores on the Dials Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian	N/A	-0.27	0.82	-0.69	-0.63	N/A	-0.10
African-American	4.08**	11.26**	21.36**	15.82**	4.13**	N/A	13.37**
Hispanic	N/A	4.95**	8.55**	6.74**	1.33	N/A	5.93**
Multi	1.24	2.85**	5.92**	4.32**	2.86**	N/A	3.74**

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of higher Dials subtest scores in the White sample.

Table 21: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for the Dials Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian/White	N/A	-0.079	-0.02	0.06	-0.105	N/A	-0.02
African-American/White	1.564	0.753	0.788	0.831	1.106	N/A	0.82
Hispanic/White	N/A	0.74	0.647	0.561	0.25	N/A	0.60
Multi/White	0.427	0.383	0.336	0.286	0.524	N/A	0.34

Gender: Table 22 shows the yearly Dial scale score means and standard deviations as a function of gender. Table 23 shows that there are significant female/male differences on the AT-SAT composite score across years. Table 24 shows the effect size of these gender differences across years. The applied Dials subtest has a small-to-medium effect as a function of gender. Tables 23 and 24 show that the Dials subtest is a barrier based on gender.

Table 22: Descriptive Statistics of the Dials Subtest as a Function of Gender

Gender		Dials Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Female	Mean	17.68	17.98	17.84	18.07	18.77	18.80	17.99
	N	44	509	1,668	1,199	149	66	605.83
	Std. Deviation	1.93	2.30	2.36	2.16	1.45	1.57	2.24
Male	Mean	18.34	18.64	18.55	18.76	19.02	19.13	18.67
	N	160	1,318	4,981	4,445	568	341	1,968.83
	Std. Deviation	2.16	1.67	1.76	1.56	1.35	1.26	1.65

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 23: Statistical Test for Differences in Score for the Dials Subtest as a Function of Gender

Gender	Statistical Test for Difference in Mean Scores on the Dials Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female	1.95	5.94**	11.28**	10.3**	1.93	1.61	6.98**

Notes: * $p < 0.05$, ** $p < 0.01$
The male population was used as the majority comparison group. Positive Z-scores are indicative of higher Dials subtest scores in the male sample.

Table 24: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for Dials Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female/Male	0.31	0.36	0.37	0.40	0.19	0.25	0.38

AT-SAT SCALES: APPLIED MATH

The Applied Math scale measures the applicant's ability to mentally apply mathematics to solve problems involving the traveling speed, time, and distance of aircraft. For the test, applicants must select the correct response to math word problems from lists of four possible answers.

Race/Ethnicity: Table 25 shows the means and standard deviations for each racial/ethnic subgroup that compose more than two percent of the applicant population. Table 26 shows the results of the statistical significance test of the difference in means comparing each minority racial/ethnic subgroup to the majority White subgroup. As shown in Table 26, the means are significantly lower for African-Americans, Hispanics/Latinos, and Multiracial subgroups compared to Whites. The effect size of these subgroup differences are shown in Table 27. As shown in this table, this scale has a large effect for African-Americans and Hispanics/Latinos, and a medium-to-large effect for Multiracial subgroups. Tables 26 and 27 show that the Applied Math subtest is a barrier to African-Americans, Hispanics/Latinos, and persons from Multiracial groups. It is not a barrier to Asians.

Table 25: Descriptive Statistics of the Applied Math Subtest as a Function of Race/Ethnicity

Race/Ethnicity		Applied Math Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Asian	Mean	*	19.00	17.66	18.75	20.27	*	18.45
	N	*	79	207	197	34	*	129.25
	Std. Deviation	*	4.73	5.17	4.34	4.32	*	4.75
African-American	Mean	12.78	13.80	13.07	13.81	15.34	*	13.44
	N	23	494	1,670	945	47	*	635.80
	Std. Deviation	4.54	5.36	5.40	5.43	5.56	*	5.40
Hispanic	Mean	*	14.97	14.81	15.86	17.00	*	15.32
	N	*	145	425	310	60	*	235.00
	Std. Deviation	*	5.34	5.46	5.57	4.38	*	5.42
Multi	Mean	14.72	17.63	17.36	17.54	17.16	*	17.41
	N	18	139	508	422	56	*	228.60
	Std. Deviation	5.31	5.72	5.38	5.51	4.70	*	5.44
White	Mean	17.77	18.82	18.95	19.20	19.76	*	19.06
	N	155	1,054	4,215	4,066	545	*	2,007.00
	Std. Deviation	5.03	5.01	4.84	4.85	4.53	*	4.85

Note: * Asian and Hispanic applicants composed less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year. In 2011, no racial or ethnic group comprised more than two percent of the applicant pool.

+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 26: Statistical Test for Differences in Scores on the Applied Math Subtest as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Mean Scores on the Applied Math Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian	N/A	-0.33	3.53**	1.40	-0.67	N/A	1.42
African-American	4.85**	17.50**	38.73**	28.01**	5.30**	N/A	23.44**
Hispanic	N/A	8.18**	15.06**	10.26**	4.61**	N/A	10.12**
Multi	2.32*	2.33*	6.37**	5.94**	3.95**	N/A	4.40**

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of higher Applied Math subtest scores in the White sample.

Table 27: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for Applied Math Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian/White	N/A	-0.04	0.27	0.09	-0.11	N/A	0.13
African-American/White	1.00	0.98	1.17	1.09	0.96	N/A	1.13
Hispanic/White	N/A	0.76	0.85	0.68	0.61	N/A	0.76
Multi/White	0.60	0.23	0.33	0.34	0.57	N/A	0.34

Gender: Table 28 shows the yearly Applied Math scale score means and standard deviations as a function of gender. Table 29 shows that there are significant female/male differences on the AT-SAT composite score Applied Math score across years. Table 30 shows the effect size of these gender differences across years. Tables 29 and 30 show that the Applied Math scale is a barrier as a function of gender.

Table 28: Descriptive Statistics of the Applied Math Subtest as a Function of Gender

Gender		Applied Math Subtest						
		2006	2007	2008	2009	2010	2011	Average Across Years+
Female	Mean	13.75	14.43	14.50	14.84	16.77	16.47	14.73
	N	44	509	1,668	1,199	149	66	605.83
	Std. Deviation	4.70	5.36	5.43	5.42	4.81	5.06	5.38
Male	Mean	17.45	18.17	18.02	18.86	19.68	20.60	18.50
	N	160	1,318	4,981	4,445	568	341	1,968.83
	Std. Deviation	5.30	5.39	5.44	5.07	4.63	4.35	5.23

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 29: Statistical Test for Differences in Score for the Applied Math Subtest as a Function of Gender

Gender	Statistical Test for Difference in Mean Scores on the Applied Math Subtest						
	2006	2007	2008	2009	2010	2011	Weighted Average Across Years
Female	4.50	13.33**	22.88**	23.04**	6.61**	6.21**	15.19**

Notes: * $p < 0.05$, ** $p < 0.01$
The male population was used as the majority comparison group. Positive Z-scores are indicative of higher Applied Math subtest scores in the male sample.

Table 30: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for Applied Math Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female/Male	0.71	0.69	0.65	0.78	0.62	0.92	0.72

AT-SAT SCALES: SCAN

In the Scan scale, multiple moving blocks of data appear at random on a computer screen. The test measures the applicant's ability to scan the screen and identify targets with numbers outside a specified valid range.

Race/Ethnicity: Table 31 shows the scale means and standard deviations for each racial/ethnic subgroup that composed more than two percent of the applicant population. Table 32 shows the results of the statistical significance test of the difference in means comparing each minority racial/ethnic subgroup to the majority White subgroup. As shown in Table 32, the scale means are significantly lower for African-Americans and Hispanics/Latinos subgroups compared to Whites. The effect size of these subgroup differences are shown in Table 33. As shown in this table, this subtest has a medium effect for African-Americans and a small effect for Hispanics/Latinos.

Table 31: Descriptive Statistics of the Scan Subtest as a Function of Race/Ethnicity

Race/Ethnicity		Scan Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Asian	Mean	*	180.04	180.03	180.40	188.15	*	180.71
	N	*	79	207	197	34	*	129.25
	Std. Deviation	*	20.04	21.48	26.27	9.13	*	22.75
African-American	Mean	162.83	166.99	167.57	171.54	179.04	*	168.80
	N	23	494	1670	945	47	*	635.80
	Std. Deviation	25.06	31.71	30.73	26.45	11.76	*	29.45
Hispanic	Mean	*	173.80	171.18	175.55	181.95	*	173.71
	N	*	145	425	310	60	*	235.00
	Std. Deviation	*	24.64	29.95	25.08	11.14	*	26.77
Multi	Mean	173.94	181.09	176.94	178.97	182.50	*	178.42
	N	18	139	508	422	56	*	228.60
	Std. Deviation	11.97	14.60	20.75	22.77	9.63	*	20.41
White	Mean	174.95	179.18	179.39	180.44	183.70	*	179.96
	N	155	1,054	4,215	4,066	545	*	2,007.00
	Std. Deviation	14.42	22.14	23.04	21.58	11.58	*	21.76

Note: * Asian and Hispanic applicants composed less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year. In 2011, no racial or ethnic group composed more than two percent of the applicant pool.

+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 32: Statistical Test for Differences in Scores on the Scan Subtest as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Mean Scores on the Scan Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian	N/A	-0.37	-0.42	0.02	-2.71**	N/A	-0.36
African-American	2.26**	7.70**	14.21**	9.63**	2.61**	N/A	8.82**
Hispanic	N/A	2.49**	5.49**	3.34**	1.15	N/A	3.45**
Multi	0.33	-1.35	2.48**	1.27	0.87	N/A	1.07

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of higher Scan subtest scores in the White sample.

Table 33: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for Scan Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian/White	N/A	-0.04	-0.03	0.00	-0.39	N/A	-0.03
African-American/White	0.75	0.48	0.46	0.39	0.40	N/A	0.47
Hispanic/White	N/A	0.24	0.35	0.22	0.15	N/A	0.28
Multi/White	0.07	-0.09	0.11	0.07	0.11	N/A	0.07

Tables 32 and 33 show that the Scan subtest is a barrier to African-American and Hispanic/Latino applicants.

Gender: Table 34 shows the yearly Scan scale means and standard deviations as a function of gender. Table 35 shows that there is only one significant female/male difference (FY 2009) on this scale. Table 36 shows the effect size of these gender differences over time. This scale has a trivial difference as a function of gender. The scan subtest is not a barrier based on gender.

Table 34: Descriptive Statistics of the Scan Subtest as a Function of Gender

Gender		Scan Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Female	Mean	172.82	175.66	175.17	177.29	183.70	185.45	176.45
	N	44	509	1,668	1,199	149	66	605.83
	Std. Deviation	18.92	25.09	26.88	24.05	9.19	9.11	24.91
Male	Mean	173.18	175.98	176.28	179.06	182.95	184.85	177.82
	N	160	1318	4981	4445	568	341	1,968.83
	Std. Deviation	16.21	24.86	25.39	22.80	13.96	15.39	23.58

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 35: Statistical Test for Differences in Score for the Scan Subtest as a Function of Gender

Gender	Statistical Test for Difference in Mean Scores on the Scan Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female	0.11	0.24	1.48	2.28*	-0.78	-0.43	1.20

Notes: $^{*}p < 0.05$, $^{**}p < 0.01$
The male population was used as the majority comparison group. Positive Z-scores are indicative of higher Scan subtest scores in the male sample.

Table 36: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for Scan Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female/Male	0.02	0.01	0.04	0.08	-0.06	-0.04	0.06

AT-SAT SCALE: ANGLES

The Angles scale measures the applicant's ability to recognize and identify angles. The scale format is multiple choice.

Race/Ethnicity: Table 37 shows this scale means and standard deviations for each racial/ethnic subgroup that composed more than two percent of the applicant population. Table 38 shows the results of the statistical significance test of the difference in means comparing each minority racial/ethnic subgroup to the majority White subgroup. As shown in Table 38, the subtest means are significantly lower for African-Americans, Hispanics/Latinos, and Multiracial subgroups compared to Whites. The effect size of these subgroup differences are shown in Table 39. As shown in this table, this subtest has a large effect for African-Americans and a medium-to-large effect for Hispanics/Latinos and Multiracial subgroups. The results shown in Tables 38 and 39 show that the Angles subtest is a barrier to African-Americans, Hispanics/Latinos, and applicants from Multiracial groups.

Table 37: Descriptive Statistics of the Angles Subtest as a Function of Race/Ethnicity

Race/Ethnicity		Angles Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Asian	Mean	*	27.27	26.23	26.76	27.77	*	26.69
	N	*	79	207	197	34	*	129.25
	Std. Deviation	*	2.46	3.69	3.11	2.05	*	3.22
African-American	Mean	22.83	22.53	22.25	23.20	25.60	*	22.63
	N	23	494	1,670	945	47	*	635.80
	Std. Deviation	3.38	5.06	5.27	4.80	3.53	*	5.07
Hispanic	Mean	*	24.70	23.80	25.42	26.82	*	24.67
	N	*	145	425	310	60	*	235.00
	Std. Deviation	*	4.72	4.82	3.86	2.25	*	4.38
Multi	Mean	24.28	25.60	25.37	25.97	26.71	*	25.67
	N	18	139	508	422	56	*	228.60
	Std. Deviation	3.98	4.15	4.06	4.08	2.71	*	4.03
White	Mean	26.88	26.57	26.60	26.90	27.68	*	26.78
	N	155	1,054	4,215	4,066	545	*	2,007.00
	Std. Deviation	2.58	3.17	3.21	3.03	2.09	*	3.07

Note: * Asian and Hispanic applicants composed less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year. In 2011, no racial or ethnic group comprised more than two percent of the applicant pool.

+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 38: Statistical Test for Differences in Scores on the Angles Subtest as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Mean Scores on the Angles Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian	N/A	-2.37**	1.44	0.64	-0.24	N/A	0.33
African-American	5.52**	16.32**	31.50**	22.72**	3.98**	N/A	19.56**
Hispanic	N/A	4.63**	11.73**	6.61**	2.83**	N/A	7.20**
Multi	2.71**	2.66**	6.62**	4.55**	2.58**	N/A	4.06**

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of higher Angles subtest scores in the White sample.

Table 39: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for Angles Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian/White	N/A	-0.22	0.12	0.05	-0.04	N/A	0.03
African-American/White	1.51	1.04	1.11	1.08	0.93	N/A	1.14
Hispanic/White	N/A	0.55	0.83	0.48	0.41	N/A	0.65
Multi/White	0.95	0.30	0.37	0.30	0.45	N/A	0.35

Gender: Table 40 shows this scale means and standard deviations on the Angles subtest as a function of gender. Table 41 shows that there are significant female/male differences for this scale across years. Table 42 shows the effect size of these gender differences over time. This scale has a medium effect as a function of gender. We conclude from Tables 41 and 42 that the Angles Subtest is a barrier based on gender.

Table 40: Descriptive Statistics of the Angles Subtest as a Function of Gender

Gender		Angles Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Female	Mean	24.66	23.75	23.62	24.51	26.82	26.64	24.13
	N	44	509	1,668	1,199	149	66	605.85
	Std. Deviation	3.44	4.80	4.91	4.47	2.59	3.00	4.64
Male	Mean	26.43	25.98	25.84	26.63	27.60	27.74	26.30
	N	160	1,318	4,981	4,445	568	341	19,68.83
	Std. Deviation	3.29	3.85	4.06	3.39	2.25	2.14	3.67

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 41: Statistical Test for Differences in Score for the Angles Subtest as a Function of Gender

Gender	Statistical Test for Difference in Mean Scores on the Angles Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female	3.05^{**}	9.37^{**}	16.66^{**}	15.29^{**}	3.34^{**}	2.84^{**}	10.55^{**}

Notes: * $p < 0.05$, ** $p < 0.01$

The male population was used as the majority comparison group. Positive Z-scores are indicative of higher Angles Subtest scores in the male sample.

Table 42: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for Angles Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female/Male	0.53	0.54	0.52	0.58	0.33	0.48	0.55

AT-SAT SCALE: THE LETTER FACTORY

The Letter Factory test measures three abilities required to perform the air traffic control job:

1. Planning and deciding what action to take in a given situation
2. Thinking ahead to avoid problems before they occur
3. Maintaining awareness of the work setting

In the Letter Factory scenarios, letters move down four conveyor belts shown on a computer screen. Applicants must place the colored letters into colored boxes, load empty boxes into the loading area, order new boxes as needed, and answer multiple choice situational awareness questions. There are 18 scenarios in the test, and two scores are calculated for each applicant.

The Letter Factory Situational Awareness scale is calculated based on the number of correct answers provided for the multiple choice situational awareness items.

Race/Ethnicity: Table 43 shows the scale means and standard deviations on the Letter Factory Situational Awareness scale for each racial/ethnic subgroup that composed more than two percent of the applicant population. Table 44 shows the results of the statistical significance test of the difference in means, comparing each minority racial/ethnic subgroup to the majority White subgroup. As shown in Table 44, the subtest means are significantly lower for African-Americans, Asian, Hispanics/Latinos, and Multiracial subgroups compared to Whites. The effect size of these subgroup differences are shown in Table 45. As shown in this table, this subtest has a large effect for African-Americans and Hispanics/Latinos, a medium effect for Multiracial subgroups, and a small effect for the Asian subgroup. This subtest meets our definition of a barrier for all racial and ethnic groups studied.

Table 43: Descriptive Statistics of the Letter Factory Situational Awareness Subtest as a Function of Race/Ethnicity

Race/Ethnicity		Letter Factory Situational Awareness Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Asian	Mean	*	41.75	40.89	42	45.68	*	41.76
	N	*	79	207	197	34	*	129.25
	Std. Deviation	*	8.82	8.19	8.12	7.45	*	8.23
African-American	Mean	36.7	37.32	36.79	37.15	39.68	*	37.02
	N	23	494	1,670	945	47	*	635.80
	Std. Deviation	9.29	7.53	8.01	8.09	7.59	*	7.97
Hispanic	Mean	*	38.75	38.51	39.81	39.77	*	39.05
	N	*	145	425	310	60	*	235.00
	Std. Deviation	*	8.38	8.28	7.39	7.54	*	7.97
Multi	Mean	39.44	42.63	41.35	42.04	39.89	*	41.66
	N	18	139	508	422	56	*	228.60
	Std. Deviation	8.99	8.71	8.62	8.28	7.16	*	8.45
White	Mean	42.83	43.29	43.37	43.52	44.53	*	43.48
	N	155	1,054	4,215	4,066	545	*	2,007.00
	Std. Deviation	7.87	7.77	7.87	7.98	7.58	*	7.89

Note: * Asian and Hispanic applicants composed less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year. In 2011, no racial or ethnic group comprised more than two percent of the applicant pool.

+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 44: Statistical Test for Differences in Scores on the Letter Factory Situational Awareness Subtest as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Mean Scores on the Letter Factory Situational Awareness Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian	N/A	1.51	4.26**	2.57**	-0.87	N/A	2.31*
African-American	3.01**	14.40**	28.55**	21.87**	4.20**	N/A	17.85**
Hispanic	N/A	6.18**	11.60**	8.48**	4.64**	N/A	8.05**
Multi	1.53	0.85	5.03**	3.50**	4.59**	N/A	3.10**

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of higher Letter Factory Situational Awareness subtest scores in the White sample.

Table 45: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for Letter Factory Situational Awareness Subtest Score						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian/White	N/A	0.20	0.31	0.19	-0.15	N/A	0.22
African-American/White	0.76	0.78	0.83	0.80	0.64	N/A	0.82
Hispanic/White	N/A	0.58	0.62	0.47	0.63	N/A	0.56
Multi/White	0.42	0.08	0.25	0.18	0.62	N/A	0.23

Gender: Table 46 shows the scale means and standard deviations as a function of gender. Table 47 shows that there are statistically significant female/male differences on this scale. Table 48 shows the effect size of these gender differences over time. This scale has a small effect as a function of gender. Based on the data in Tables 47 and 48, we conclude that the Letter Factory Situational Awareness subtest is not a barrier based on gender.

Table 46: Descriptive Statistics of the Letter Factory Situational Awareness Subtest Score as a Function of Gender

Gender		Letter Factory Situational Awareness Subtest Score						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Female	Mean	42.02	40.04	40.19	40.90	43.45	42.17	40.59
	N	44	509	1,668	1,199	149	66	605.83
	Std. Deviation	7.76	7.83	8.20	8.20	7.74	8.24	8.13
Male	Mean	41.76	41.83	41.68	42.51	43.70	44.72	42.20
	N	160	1,318	4,981	4,445	568	341	1,968.83
	Std. Deviation	8.59	8.43	8.52	8.29	7.75	7.74	8.37

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 47: Statistical Test for Differences in Score for the Letter Factory Situational Awareness Subtest as a Function of Gender

Gender	Statistical Test for Difference in Mean Scores for the Letter Factory Situational Awareness Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female	-0.19	4.28**	6.37**	6.03**	0.35	2.32*	4.21**

Notes: * $p < 0.05$, ** $p < 0.01$

The male population was used as the majority comparison group. Positive Z-scores are indicative of higher Letter Factory Situational Awareness subtest scores in the male sample.

Table 48: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for Letter Factory Situational Awareness Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female/Male	-0.03	0.22	0.18	0.20	0.03	0.33	0.19

AT-SAT SCALE: THE LETTER FACTORY PERFORMANCE SCALE

The Letter Factory Performance Scale is calculated using the total number of boxes correctly placed by the applicant, the total number of failures to order boxes when needed, the total number of attempts to place unneeded boxes, the correct box placement latency, and the box placement failure latency.

Race/Ethnicity: Table 49 shows the means and standard deviations on this subtest for each racial/ethnic subgroup that composed more than two percent of the applicant population. Table 50 shows the results of the statistical significance test of the difference in means comparing each minority racial/ethnic subgroup to the majority White subgroup. As shown in this table, subtest means are significantly lower for African-Americans, Hispanics/Latinos, and Multiracial subgroups compared to Whites. The effect size of these subgroup differences are shown in Table 51. As shown in this table, this subtest has a large effect for African-Americans, a medium effect for Hispanics/Latinos, and a small effect for Multiracial subgroups. Based on the data in Tables 50 and 51, we conclude that the Letter Factory Performance Subtest is a barrier to African-Americans, Hispanics/Latinos, and persons from Multiracial groups.

Table 49: Descriptive Statistics of the Letter Factory Performance Subtest as a Function of Race/Ethnicity

Race/Ethnicity		Letter Factory Performance Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Asian	Mean	*	0.25	0.25	0.27	0.3	*	0.26
	N	*	79	207	197	34	*	129.25
	Std. Deviation	*	0.09	0.09	0.08	0.07	*	0.08
African-American	Mean	0.21	0.19	0.2	0.2	0.22	*	0.20
	N	23	494	1,670	945	47	*	635.80
	Std. Deviation	0.09	0.06	0.07	0.06	0.06	*	0.06
Hispanic	Mean	*	0.22	0.23	0.24	0.25	*	0.23
	N	*	145	425	310	60	*	235.00
	Std. Deviation	*	0.07	0.07	0.06	0.06	*	0.07
Multi	Mean	0.27	0.26	0.24	0.26	0.27	*	0.25
	N	18	139	508	422	56	*	228.60
	Std. Deviation	0.1	0.08	0.07	0.08	0.08	*	0.07
White	Mean	0.3	0.26	0.27	0.27	0.28	*	0.27
	N	155	1,054	4,215	4,066	545	*	2,007.00
	Std. Deviation	0.07	0.07	0.07	0.07	0.07	*	0.07

Note: * Asian and Hispanic applicants composed less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year. In 2011, no racial or ethnic group composed more than two percent of the applicant pool.

+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 50: Statistical Test for Differences in Scores on the Letter Factory Performance Subtest as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Mean Scores on the Letter Factory Performance Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian	N/A	1.00	3.00**	0.69	-1.64	N/A	1.27
African-American	4.82**	20.34**	34.86**	29.39**	6.09**	N/A	23.27**
Hispanic	N/A	7.11**	11.62**	9.32**	3.19**	N/A	8.47**
Multi	0.97	0.71	7.70**	3.40**	1.28	N/A	3.52**

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of higher Letter Factory Performance subtest scores in the White sample.

Table 51: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for Letter Factory Performance Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian/White	N/A	0.14	0.25	0.06	-0.27	N/A	0.14
African-American/White	1.22	1.05	0.97	0.98	0.85	N/A	1.03
Hispanic/White	N/A	0.61	0.58	0.49	0.40	N/A	0.57
Multi/White	0.32	0.06	0.36	0.18	0.21	N/A	0.29

Gender: Table 52 shows the scale means and standard deviations as a function of gender. Table 53 shows three statistically significant female/male differences from FY 2008 to FY 2010 but not on the other years and not overall. Table 54 shows that these gender differences are trivial and in favor of women. We conclude therefore that the Letter Factory Performance Subtest is not a barrier based on gender.

Table 52: Descriptive Statistics of the Letter Factory Performance Subtest as a Function of Gender

Gender		Letter Factory Performance Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Female	Mean	0.29	0.25	0.25	0.26	0.29	0.29	0.26
	N	44	509	1,668	1,199	149	66	605.83
	Std. Deviation	0.07	0.08	0.08	0.08	0.07	0.08	0.08
Male	Mean	0.28	0.24	0.25	0.26	0.27	0.29	0.25
	N	160	1,318	4,981	4,445	568	341	1,968.83
	Std. Deviation	0.08	0.07	0.08	0.08	0.07	0.08	0.08

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 53: Statistical Test for Differences in Score for the Letter Factory Performance Subtest as a Function of Gender

Gender	Statistical Test for Difference in Mean Scores on the Letter Factory Performance Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female	-0.78	-1.68	-2.12*	-2.21*	-3.62**	0.32	-1.26

Notes: * $p < 0.05$, ** $p < 0.01$

The male population was used as the majority comparison group. Positive Z-scores are indicative of higher Letter Factory Performance Subtest scores in the male sample.

Table 54: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for Letter Factory Performance Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female/Male	-0.12	-0.09	-0.06	-0.07	-0.32	0.05	-0.05

AT-SAT SCALES: AIR TRAFFIC SCENARIOS TEST (ATST)

The Air Traffic Scenarios test simulates air traffic situations and is designated to measure the ability to safely and efficiently guide airplanes. This low-fidelity simulation of a controller's job tests the applicant's ability to learn and follow simple procedures, visualize and project paths in three dimensions, monitor several objects at once, plan ahead, and remember to execute elements of the plan at the appropriate time. The test consists of four operational scenarios, and three scores are calculated for each applicant.

The ATST Efficiency score is based on the speed with which aircraft are accepted into the airspace (handoff latency) and reach their destinations.

Race/Ethnicity: Table 55 shows subtest means and standard deviations on the ATST Subtest for each racial/ethnic subgroup that composed more than two percent of the applicant population. Table 56 shows the results of the statistical significance test of the difference in means comparing each minority racial/ethnic subgroup to the majority White subgroup. As shown in this table, the subtest means are significantly lower for African-Americans, Asians, Hispanics/Latinos, and Multiracial subgroups compared to Whites. The effect sizes of these subgroups differences are shown in Table 57. As shown in this table, this subtest has a large effect for African-Americans, and a medium to large effect for Hispanics/Latinos, and a medium effect for Asian and Multiracial subgroups. We conclude that the ATST subtest is a barrier to African-Americans, Hispanics/Latinos, Asians, and Multiracial applicants.

Table 55: Descriptive Statistics of the ATST Efficiency Subtest as a Function of Race/Ethnicity

Race/Ethnicity		ATST Efficiency Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Asian	Mean	*	58.70	59.95	64.28	71.78	*	62.19
	N	*	79	207	197	34	*	129.25
	Std. Deviation	*	15.74	15.98	14.85	10.66	*	15.25
African-American	Mean	49.92	45.93	46.98	49.5	62.14	*	47.81
	N	23	494	1,670	945	47	*	635.80
	Std. Deviation	13.25	13.21	13.9	14.55	15.64	*	14.02
Hispanic	Mean	*	54.78	55.5	60.98	67.51	*	57.96
	N	*	145	425	310	60	*	235.00
	Std. Deviation	*	14.78	15.49	14.62	13.78	*	15.01
Multi	Mean	61.52	61.3	58.74	63.73	68.32	*	61.41
	N	18	139	508	422	56	*	228.60
	Std. Deviation	16.77	15.32	14.86	16.48	13.2	*	15.50
White	Mean	66.64	64.56	65.04	67.8	72.90	*	66.56
	N	155	1,054	4,215	4,066	545	*	2,007.00
	Std. Deviation	13.04	14.35	14.38	14.36	12.46	*	14.25

Note: * Asian and Hispanic applicants composed less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year. In 2011, no racial or ethnic group composed more than two percent of the applicant pool.

+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 56: Statistical Test for Differences in Scores on the ATST Efficiency Subtest as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Mean Scores on the ATST Efficiency Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian	N/A	3.21**	4.50**	3.25**	0.59	N/A	3.17**
African-American	5.66**	25.15**	44.50**	34.92**	4.59**	N/A	29.27**
Hispanic	N/A	7.50**	12.19**	7.92**	2.90**	N/A	8.35**
Multi	1.25	2.38**	9.06**	4.88**	2.48**	N/A	4.80**

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of higher ATST Efficiency Subtest scores in the White sample.

Table 57: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for ATST Efficiency Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian/White	N/A	0.41	0.35	0.24	0.09	N/A	0.30
African-American/White	1.28	1.33	1.27	1.27	0.84	N/A	1.32
Hispanic/White	N/A	0.68	0.66	0.47	0.43	N/A	0.60
Multi/White	0.38	0.23	0.44	0.28	0.37	N/A	0.36

Gender: Table 58 shows the scale means and standard deviations as a function of gender. Table 59 shows that there are statistically significant female/male differences on this scale across years. Table 60 shows that these gender differences are of a medium effect. We conclude that this subtest is a barrier based on gender.

Table 58: Descriptive Statistics of the ATST Efficiency Subtest as a Function of Gender

Gender		ATST Efficiency Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Female	Mean	56.54	51.21	52.05	56.63	66.69	69.46	54.41
	N	44	509	1,668	1,199	149	66	605.83
	Std. Deviation	12.22	15.21	15.21	15.92	12.85	12.91	15.29
Male	Mean	65.82	61.42	62.16	66.25	72.83	74.99	64.55
	N	160	1,318	4,981	4,445	568	341	1,968.83
	Std. Deviation	14.37	15.75	15.84	15.35	12.77	12.25	15.40

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 59: Statistical Test for Differences in Score for the ATST Efficiency Subtest as a Function of Gender

Gender	Statistical Test for Difference in Mean Scores on the ATST Efficiency Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female	4.29**	12.74**	23.25**	18.72**	5.20**	3.21**	14.25**

Notes: * $p < 0.05$, ** $p < 0.01$
The male population was used as the majority comparison group. Positive Z-scores are indicative of higher ATST Efficiency subtest scores in the male sample.

Table 60: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for ATST Efficiency Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female/Male	0.67	0.65	0.64	0.62	0.48	0.45	0.66

The ATST Safety scale is based on avoidance of separation errors and crashes.

Race/Ethnicity: Table 61 shows the scale means and standard deviations on the ATST Safety Subtest for each racial/ethnic subgroup that composed more than two percent of the applicant population. Table 62 shows the results of the statistical significance test of the

difference in means comparing each minority racial/ethnic subgroup to the majority White subgroup. As shown in Table 62, the scale means are significantly lower for African-Americans, Hispanics/Latinos, and Multiracial subgroups compared to Whites.

The effect size of these subgroup differences are shown in Table 63. As shown in this table, this subtest has a large effect for African-Americans and small but nontrivial effect for Hispanics/Latinos and Multiracial subgroups. We conclude that the ATST Safety Subtest is a barrier for African-Americans as well as Hispanic/Latinos and persons from Multiracial groups.

Table 61: Descriptive Statistics of the ATST Safety Subtest as a Function of Race/Ethnicity

Race/Ethnicity		ATST Safety Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Asian	Mean	*	59.28	60.94	63.77	77.09	*	62.83
	N	*	79	207	197	34	*	129.25
	Std. Deviation	*	17.85	16.63	17.15	10.04	*	16.72
African-American	Mean	51.48	48.59	49.79	51.23	59.7	*	50.19
	N	23	494	1,670	945	47	*	635.80
	Std. Deviation	17.2	18.37	18.02	18.42	17.08	*	18.18
Hispanic	Mean	*	56.22	57.09	60.69	66.08	*	58.72
	N	*	145	425	310	60	*	235.00
	Std. Deviation	*	18.54	17.45	17.43	15.67	*	17.53
Multi	Mean	64.31	60.46	58.19	62.99	67.43	*	60.78
	N	18	139	508	422	56	*	228.60
	Std. Deviation	14.87	18.54	17.32	18.36	16.17	*	17.78
White	Mean	66.93	64.01	63.19	65.37	72.55	*	64.73
	N	155	1,054	4,215	4,066	545	*	2,007.00
	Std. Deviation	14.48	16.54	16.77	16.75	14.14	*	16.58

Note: * Asian and Hispanic applicants composed less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year. In 2011, no racial or ethnic group composed more than two percent of the applicant pool.

+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 62: Statistical Test for Differences in Scores on the ATST Safety Subtest as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Mean Scores on the ATST Safety Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian	N/A	2.28*	1.91	1.28	-2.49**	N/A	1.25
African-American	4.10**	15.89**	26.22**	21.62**	5.01**	N/A	17.94**
Hispanic	N/A	4.80**	6.90**	4.57**	3.06**	N/A	5.00**
Multi	0.71	2.15*	6.18**	2.56**	2.28*	N/A	3.20**

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of higher ATST Safety subtest scores in the White sample.

Table 63: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for ATST Safety Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian/White	N/A	0.29	0.14	0.10	-0.33	N/A	0.11
African-American/White	1.04	0.90	0.78	0.83	0.89	N/A	0.86
Hispanic/White	N/A	0.46	0.36	0.28	0.45	N/A	0.36
Multi/White	0.18	0.21	0.30	0.14	0.36	N/A	0.24

Gender: Table 64 shows the scale means and standard deviations as a function of gender. Table 65 shows that there are statistically significant female/male differences on this scale across years. Table 66 shows that these gender differences are of a medium effect. This subtest is a barrier based on gender.

Table 64: Descriptive Statistics of the ATST Safety Subtest as a Function of Gender

Gender		ATST Safety Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Female	Mean	58.89	51.97	52.29	55.05	65.97	69.36	54.11
	N	44	509	1,668	1,199	149	66	605.83
	Std. Deviation	13.38	19.11	18.20	18.26	15.92	17.55	18.21
Male	Mean	66.16	61.88	61.56	64.83	72.58	75.17	63.81
	N	160	1318	4,981	4,445	568	341	1,968.83
	Std. Deviation	15.97	17.26	17.47	17.19	14.30	14.51	17.10

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 65: Statistical Test for Differences in Score for the ATST Safety Subtest as a Function of Gender

Gender	Statistical Test for Difference in Mean Scores on the ATST Safety Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female	3.05**	10.20**	18.17**	16.66**	4.60**	2.53*	11.63**

Notes: * $p < 0.05$, ** $p < 0.01$

The male population was used as the majority comparison group. Positive Z-scores are indicative of higher ATST Safety subtest scores in the male sample.

Table 66: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for ATST Safety Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female/Male	0.47	0.56	0.52	0.56	0.45	0.39	0.56

The ATST Procedures scale is based on following the rules for correct destination and landing/exit conditions.

Race/Ethnicity: Table 67 shows the scale means and standard deviations for each racial/ethnic subgroup that composed more than two percent of the applicant population. Table 68 shows the results of the statistical significance test of the difference in means comparing each minority racial/ethnic subgroup to the majority White subgroup. As shown in Table 68, the scale means are significantly lower for African-Americans, Hispanics/Latinos, and Multiracial subgroups compared to Whites. The effect size of these subgroup differences are shown in Table 69. As shown in this table, this subtest has a medium effect for African-Americans and a small but nontrivial effect for Hispanics/Latinos and Multiracial subgroups. We conclude, therefore, that the ATST Procedures Subtest is a barrier for African-Americans, Hispanics/Latinos, and persons from Multiracial groups.

Table 67: Descriptive Statistics of the ATST Procedures Subtest as a Function of Race/Ethnicity

Race/Ethnicity		ATST Procedures Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Asian	Mean	*	50.38	51.68	56.15	61.53	*	53.83
	N	*	79	207	197	34	*	129.25
	Std. Deviation	*	14.74	16.41	15.80	12.32	*	15.72
African-American	Mean	50.91	41.98	42.77	45.74	51.79	*	43.72
	N	23	494	1,670	945	47	*	635.80
	Std. Deviation	21.04	18.42	18.51	18.27	14.63	*	18.39
Hispanic	Mean	*	48.57	48.03	51.00	58.40	*	49.75
	N	*	145	425	310	60	*	235.00
	Std. Deviation	*	17.65	17.49	16.72	15.77	*	17.17
Multi	Mean	59.78	48.95	47.96	51.24	55.45	*	49.85
	N	18	139	508	422	56	*	228.60
	Std. Deviation	18.12	14.94	16.05	15.4	14.66	*	15.66
White	Mean	61.24	50.85	51.38	54.01	58.08	*	52.90
	N	155	1,054	4,215	4,066	545	*	2,007.00
	Std. Deviation	13.04	14.64	14.8	14.93	14.27	*	14.78

Note: * Asian and Hispanic applicants composed less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year. In 2011, no racial or ethnic group composed more than two percent of the applicant pool.

+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 68: Statistical Test for Differences in Scores on the ATST Procedures Subtest as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Mean Scores on the ATST Procedures Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian	N/A	0.27	-0.26	-1.86	-1.57	N/A	-0.65
African-American	2.29*	9.40**	16.98**	12.95**	2.83**	N/A	11.47**
Hispanic	N/A	1.49	3.82**	3.08**	-0.15	N/A	2.70**
Multi	0.33	1.41	4.57**	3.52**	1.28	N/A	2.81**

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of higher ATST Procedures subtest scores in the White sample.

Table 69: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for ATST Procedures Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian/White	N/A	0.03	-0.02	-0.14	-0.24	N/A	-0.06
African-American/White	0.72	0.56	0.54	0.53	0.44	N/A	0.58
Hispanic/White	N/A	0.15	0.22	0.20	-0.02	N/A	0.21
Multi/White	0.11	0.13	0.23	0.19	0.18	N/A	0.20

Gender: Table 70 shows the scale means and standard deviations as a function of gender. Table 71 shows that there are statistically significant female/male differences on this scale across years. Table 72 shows that these gender differences are of a small effect. We concluded that the ATST Procedures Subtest is not a barrier based on gender.

Table 70: Descriptive Statistics of the ATST Procedures Subtest as a Function of Gender

Gender		ATST Procedures Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Female	Mean	60.22	45.94	47.58	50.49	58.03	56.62	49.06
	N	44	509	1,668	1,199	149	66	605.83
	Std. Deviation	15.37	17.17	18.28	17.83	15.33	13.54	17.76
Male	Mean	60.17	49.13	49.43	53.05	57.60	58.84	51.57
	N	160	1,318	4,981	4,445	568	341	1,968.83
	Std. Deviation	14.10	15.84	15.79	15.38	14.22	14.24	15.51

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 71: Statistical Test for Differences in Score for the ATST Procedures Subtest as a Function of Gender

Gender	Statistical Test for Difference in Mean Scores on the ATST Procedures Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female	-0.02	3.63**	3.69**	4.53**	-0.31	1.21	3.13**

Notes: * $p < 0.05$, ** $p < 0.01$
The male population was used as the majority comparison group. Positive Z-scores are indicative of higher ATST Procedures subtest scores in the male sample.

Table 72: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for ATST Procedures Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female/Male	0.00	0.20	0.11	0.16	-0.03	0.16	0.16

AT-SAT SCALES: ANALOGIES TEST

The Analogies test assesses inductive reasoning and information-processing skills using three item types: nonsemantic words, semantic words, and nonsemantic visuals. For each test item, the Analogy program separates the computer screen into three sections. The applicant then uses the mouse to view each section of the screen and answer the item. Two scores are calculated for each applicant. The Analogies Correct Subtest is scored on the number of items the applicant answers correctly.

Race/Ethnicity: Table 73 shows the means and standard deviations on the Analogies Correct subtest for each racial/ethnic subgroup that composed more than two percent of the applicant population. Table 74 shows the results of the statistical significance test of the difference in means comparing each minority racial/ethnic subgroup to the majority White subgroup. As shown in Table 74, means are significantly lower for African-Americans, Asians, Hispanics/Latinos, and Multiracial subgroups compared to Whites. The effect size of these subgroup differences are shown in Table 75. As shown in this table, this subtest has a large effect for African-Americans and Hispanics/Latinos, a small nontrivial effect for Asian and Multiracial subgroups. Therefore, the Analogies Correct subtest is a barrier to African-Americans, Hispanics/Latinos, and persons from Multiracial groups and Asians.

Table 73: Descriptive Statistics of the Analogies Correct Subtest as a Function of Race/Ethnicity

Race/Ethnicity		Analogies Correct Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Asian	Mean	*	26.6	25.77	27.31	28.94	*	26.69
	N	*	79	207	197	34	*	129.25
	Std. Deviation	*	7.87	7.09	7.31	5.23	*	7.21
African-American	Mean	19.22	23.24	22.23	22.97	24.49	*	22.62
	N	23	494	1,670	945	47	*	635.80
	Std. Deviation	5.74	6.65	7.22	7.06	7.51	*	7.09
Hispanic	Mean	*	23.01	22.72	23.65	25.12	*	23.22
	N	*	145	425	310	60	*	235.00
	Std. Deviation	*	7.68	7.46	6.6	5.91	*	7.14
Multi	Mean	22.94	27.36	26.13	27.33	26.00	*	26.66
	N	18	139	508	422	56	*	228.60
	Std. Deviation	6.74	7.35	7.05	6.87	7.3	*	7.03
White	Mean	25.94	28.04	28.01	28.27	28.80	*	28.13
	N	155	1,054	4,215	4,066	545	*	2,007.00
	Std. Deviation	5.36	6.40	6.62	6.62	5.87	*	6.54

Note: * Asian and Hispanic applicants composed less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year. In 2011, no racial or ethnic group composed more than two percent of the applicant pool.

+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 74: Statistical Test for Differences in Scores on the Analogies Correct Subtest as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Mean Scores on the Analogies Correct Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian	N/A	1.59	4.45**	1.81	-0.16	N/A	2.21*
African-American	5.28**	13.39**	28.34**	21.01**	3.83**	N/A	17.40**
Hispanic	N/A	7.53**	14.07**	11.88**	4.58**	N/A	10.05**
Multi	1.82	1.04	5.73**	2.69**	2.78**	N/A	3.01**

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of higher Analogies Correct subtest scores in the White sample.

Table 75: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for Analogies Correct Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian/White	N/A	0.22	0.34	0.15	-0.03	N/A	0.22
African-American/White	1.24	0.74	0.85	0.79	0.72	N/A	0.83
Hispanic/White	N/A	0.77	0.79	0.70	0.63	N/A	0.74
Multi/White	0.54	0.10	0.28	0.14	0.47	N/A	0.22

Gender: Table 76 shows the scale means and standard deviations as a function of gender. Table 77 shows that there are statistically significant female/male differences on this scale across years. Table 78 shows that these gender differences are small but nontrivial. We conclude that the Analogies Correct Subtest is a barrier based on gender.

Table 76: Descriptive Statistics of the Analogies Correct Subtest as a Function of Gender

Gender		Analogies Correct Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Female	Mean	23.93	25.45	24.91	26.06	27.83	27.17	25.51
	N	44	509	1,668	1,199	149	66	605.83
	Std. Deviation	5.16	7.30	7.43	7.19	6.19	5.85	7.23
Male	Mean	24.68	26.69	26.45	27.36	28.10	29.21	26.96
	N	160	1,318	4,981	4,445	568	341	1,968.83
	Std. Deviation	6.39	6.89	7.25	6.96	6.29	6.20	7.02

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 77: Statistical Test for Differences in Score for the Analogies Correct Subtest as a Function of Gender

Gender	Statistical Test for Difference in Mean Scores on the Analogies Correct Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female	0.80	3.30**	7.37**	5.62**	0.47	2.57*	4.32**

Notes: * $p < 0.05$, ** $p < 0.01$
The male population was used as the majority comparison group. Positive Z-scores are indicative of higher Analogies Correct Subtest scores in the male sample.

Table 78: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for Analogies Correct Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female/Male	0.12	0.18	0.21	0.19	0.04	0.33	0.20

The Analogies Window Return score is based on the total number of times the applicant selects each screen.

Race/Ethnicity: Table 79 shows the means and standard deviations on the Analogies Window Return test for each racial/ethnic subgroup that composed more than two percent of the applicant population. Table 80 shows the results of the statistical significance test of the difference in means comparing each minority racial/ethnic subgroup to the majority White subgroup. As shown in Table 80, the scale means are significantly lower for African-Americans, Asians, Hispanics/Latinos, and Multiracial subgroups compared to Whites. The effect size of these subgroup differences are shown in Table 81. As shown in this table, this subtest has small effects for all subgroups except Asians; therefore it is a barrier to Asians.

Table 79: Descriptive Statistics of the Analogies Window Return Subtest as a Function of Race/Ethnicity

Race/Ethnicity		Analogies Window Return Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Asian	Mean	*	5.87	5.38	5.15	6.03	*	5.41
	N	*	79	207	197	34	*	129.25
	Std. Deviation	*	4.30	4.52	4.11	3.48	*	4.28
African-American	Mean	6.44	6.66	6.67	6.54	7.36	*	6.64
	N	23	494	1670	945	47	*	635.80
	Std. Deviation	2.73	4.63	4.81	4.53	5.47	*	4.70
Hispanic	Mean	*	6.12	6.74	6.36	5.67	*	6.45
	N	*	145	425	310	60	*	235.00
	Std. Deviation	*	5.26	4.79	4.41	5.29	*	4.78
Multi	Mean	7.61	6.76	6.62	6.33	6.32	*	6.53
	N	18	139	508	422	56	*	228.60
	Std. Deviation	2.64	4.56	4.57	4.30	4.17	*	4.43
White	Mean	7.51	7.16	7.34	7.18	6.60	*	7.22
	N	155	1,054	4,215	4,066	545	*	2,007.00
	Std. Deviation	2.40	4.41	4.64	4.54	3.97	*	4.51

Note: * Asian and Hispanic applicants composed less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year. In 2011, no racial or ethnic group composed more than two percent of the applicant pool.

+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 80: Statistical Test for Differences in Scores on the Analogies Window Return Subtest as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Mean Scores on the Analogies Window Return Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian	N/A	2.57**	6.09**	6.75**	0.92	N/A	4.65**
African-American	1.79	2.04	4.86**	3.92**	-0.94	N/A	2.74**
Hispanic	N/A	2.27*	2.50**	3.15**	1.32	N/A	2.35**
Multi	-0.16	1.00	3.34**	3.87**	0.48	N/A	2.23*

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of higher Analogies Window Return Subtest scores in the White sample.

Table 81: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for Analogies Window Return Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian/White	N/A	0.29	0.42	0.45	0.14	N/A	0.40
African-American/White	0.44	0.11	0.14	0.14	-0.19	N/A	0.13
Hispanic/White	N/A	0.23	0.13	0.18	0.23	N/A	0.17
Multi/White	-0.04	0.09	0.16	0.19	0.07	N/A	0.15

Gender: Table 82 shows this scale means and standard deviations as a function of gender. Table 83 shows that there are statistically significant female/male differences on this scale across years. Table 84 shows that these gender differences are small and in favor of women. The Analogies Window Return Subtest is not a barrier based on gender.

Table 82: Descriptive Statistics of the Analogies Window Return Subtest as a Function of Gender

Gender		Analogies Window Return Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Female	Mean	8.05	7.75	7.81	8.06	7.93	6.80	7.87
	N	44	509	1,668	1,199	149	66	605.83
	Std. Deviation	2.88	4.82	4.90	4.80	4.61	3.97	4.81
Male	Mean	7.06	6.45	6.78	6.62	6.20	6.18	6.64
	N	160	1,318	4,981	4,445	568	341	1,968.83
	Std. Deviation	2.42	4.40	4.62	4.40	4.04	3.84	4.44

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 83: Statistical Test for Differences in Score for the Analogies Window Return Subtest as a Function of Gender

Gender	Statistical Test for Difference in Mean Scores on the Analogies Window Return Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female	-2.07*	-5.32**	-7.51**	-9.40**	-4.17**	-1.17	-5.61**

Notes: * $p < 0.05$, ** $p < 0.01$
The male population was used as the majority comparison group. Positive Z-scores are indicative of higher Analogies Window Return Subtest scores in the male sample.

Table 84: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for Analogies Window Return Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female/Male	-0.39	-0.29	-0.22	-0.32	-0.41	-0.16	-0.27

AT-SAT SCALES: EXPERIENCE QUESTIONNAIRE (EQ)

The Experience Questionnaire (EQ) assesses whether the applicants possess certain work-related attributes by asking questions about their past experiences. Each question is written as a statement about the applicants' past experience. This questionnaire requires applicants to indicate their level of agreement with each statement on a five-point scale. The questionnaire produces 11 scores for each applicant.

The EQ Composure score indicates the applicant's ability to think clearly in stressful situations.

Race/Ethnicity: Table 85 shows the scale means and standard deviations for each racial/ethnic subgroup that composed more than two percent of the applicant population. Table 86 shows the results of the statistical significance test of the difference in means, comparing each minority racial/ethnic subgroup to the majority White subgroup. As shown in Table 86, the scale's means only showed significant differences only for Asians. Therefore, the EQ Composure Subtest is a barrier to Asians. The effect sizes of these subgroups differences are shown in Table 87.

Table 85: Descriptive Statistics of the EQ Composure Subtest as a Function of Race/Ethnicity

Race/Ethnicity		EQ Composure Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Asian	Mean	*	81.00	77.46	76.98	78.09	*	77.86
	N	*	79	207	197	34	*	129.25
	Std. Deviation	*	12.69	12.63	13.04	13.57	*	12.88
African-American	Mean	79.26	80.51	80.23	81.95	79.29	*	80.76
	N	23	494	1,670	945	47	*	635.80
	Std. Deviation	12.38	12.47	12.8	12.43	13.03	*	12.64
Hispanic	Mean	*	80.98	80.32	80.46	79.42	*	80.41
	N	*	145	425	310	60	*	235.00
	Std. Deviation	*	12.86	12.04	11.31	12.66	*	11.99
Multi	Mean	78.5	79.33	80.00	80.90	80.15	*	80.24
	N	18	139	508	422	56	*	228.60
	Std. Deviation	10.65	12.81	12.03	12.12	12.66	*	12.18
White	Mean	75.38	80.61	80.33	80.35	79.46	*	80.25
	N	155	1,054	4,215	4,066	545	*	2,007.00
	Std. Deviation	12.20	11.74	11.65	11.71	12.84	*	11.76

Note: * Asian and Hispanic applicants composed less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year. In 2011, no racial or ethnic group composed more than two percent of the applicant pool.

+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 86: Statistical Test for Differences in Scores on the EQ Composure Subtest as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Mean Scores on the EQ Composure Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian	N/A	-0.26	3.21**	3.56**	0.58	N/A	2.05*
African-American	-1.40	0.16	0.30	-3.61**	0.09	N/A	-0.91
Hispanic	N/A	-0.32	0.02	-0.16	0.03	N/A	-0.20
Multi	-1.16	1.12	0.58	-0.89	-0.39	N/A	0.01

Notes: $^{*}p < 0.05$, $^{**}p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of higher EQ Composure Subtest scores in the White sample.

Table 87: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for EQ Composure Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian/White	N/A	-0.03	0.25	0.29	0.11	N/A	0.20
African-American/White	-0.32	0.01	0.01	-0.14	0.01	N/A	-0.04
Hispanic/White	N/A	-0.03	0.00	-0.01	0.00	N/A	-0.01
Multi/White	-0.26	0.11	0.03	-0.05	-0.05	N/A	0.00

Gender: Table 88 shows the means and standard deviations as a function of gender. Table 89 shows that there are statistically significant female/male differences on this scale across years. Table 90 shows that these gender differences are small. The EQ Composure Subtest is not a barrier based on gender.

Table 88: Descriptive Statistics of the EQ Composure Subtest as a Function of Gender

Gender		EQ Composure Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Female	Mean	79.31	79.51	78.88	79.50	80.32	76.69	79.20
	N	44	509	1,668	1,199	149	66	605.83
	Std. Deviation	11.08	12.38	12.17	11.85	12.16	12.90	12.10
Male	Mean	76.41	81.14	80.56	80.78	79.07	79.25	80.54
	N	160	1,318	4,981	4,445	568	341	1,968.83
	Std. Deviation	12.29	11.85	11.92	11.95	12.84	12.99	12.01

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 89: Statistical Test for Differences in Score for the EQ Composure Subtest as a Function of Gender

Gender	Statistical Test for Difference in Mean Scores on the EQ Composure Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female	-1.50	2.54*	4.91**	3.32**	-1.11	1.47	2.40*

Notes: * $p < 0.05$, ** $p < 0.01$

The male population was used as the majority comparison group. Positive Z-scores are indicative of higher EQ Composure Subtest scores in the male sample.

Table 90: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for EQ Composure Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female/Male	-0.24	0.14	0.14	0.11	-0.10	0.20	0.11

The EQ Consistency of Work Behavior score indicates the applicant's ability to behave consistently at work (e.g., dealing with coworkers in a consistent manner; consistently using the correct phraseology).

Race/Ethnicity: Table 91 shows the scale means and standard deviations on the EQ Consistency of Work Behavior Subtest for each racial/ethnic subgroup that composed more than two percent of the applicant population. Table 92 shows the results of the statistical significance test of the difference in means comparing each minority racial/ethnic subgroup to the majority White subgroup. As shown in Table 92, the subtest means showed occasional significant differences but only one significant difference to the disadvantage of a minority group, Asians. The effect sizes of these subgroup differences are shown in Table 93, and they are trivial in size except in the case of Asians. This subtest is a barrier to Asians.

Table 91: Descriptive Statistics of the EQ Consistency of Work Behavior Subtest as a Function of Race/Ethnicity

Race/Ethnicity		EQ Consistency of Work Behavior Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Asian	Mean	*	84.67	82.76	81.84	85.23	*	82.86
	N	*	79	207	197	34	*	129.25
	Std. Deviation	*	10.83	10.54	11.17	9.36	*	10.78
African-American	Mean	86.92	85.92	85.55	86.97	84.82	*	86.03
	N	23	494	1,670	945	47	*	635.80
	Std. Deviation	10.69	10.46	10.48	10.22	9.87	*	10.40
Hispanic	Mean	*	85.24	84.47	85.03	85.59	*	84.84
	N	*	145	425	310	60	*	235.00
	Std. Deviation	*	10.77	10.77	9.95	10.43	*	10.50
Multi	Mean	83.89	84.97	84.48	85.27	85.04	*	84.85
	N	18	139	508	422	56	*	228.60
	Std. Deviation	12.32	10.22	10.23	10.13	10.00	*	10.22
White	Mean	83.31	85.48	84.95	85.1	85.77	*	85.08
	N	155	1,054	4,215	4,066	545	*	2,007.00
	Std. Deviation	9.50	9.63	9.97	10.00	9.90	*	9.94

Note: * Asian and Hispanic applicants composed less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year. In 2011, no racial or ethnic group composed more than two percent of the applicant pool.

+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 92: Statistical Test for Differences in Scores on the EQ Consistency of Work Behavior Subtest as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Mean Scores on the EQ Consistency of Work Behavior Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian	N/A	0.65	2.92**	4.03**	0.32	N/A	2.28*
African-American	-1.53	-0.80	-2.02*	-5.07**	0.63	N/A	-2.01*
Hispanic	N/A	0.25	0.87	0.13	0.12	N/A	0.33
Multi	-0.19	0.55	0.98	-0.32	0.52	N/A	0.34

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of higher EQ Consistency of Work Behavior Subtest scores in the White sample.

Table 93: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for EQ Consistency of Work Behavior Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian/White	N/A	0.08	0.22	0.33	0.05	N/A	0.22
African-American/White	-0.37	-0.05	-0.06	-0.19	0.10	N/A	-0.10
Hispanic/White	N/A	0.02	0.05	0.01	0.02	N/A	0.02
Multi/White	-0.06	0.05	0.05	-0.02	0.07	N/A	0.02

Gender: Table 94 shows the scale means and standard deviations as a function of gender. Table 95 shows that there are two statistically significant female/male differences (FY 2006 and FY 2010) on this scale. Table 96 shows that the gender difference is trivial and in favor of women. Therefore, the EQ Consistency of Work Behavior Subtest is not a barrier based on gender.

Table 94: Descriptive Statistics of the EQ Consistency of Work Behavior Subtest as a Function of Gender

Gender		EQ Consistency of Work Behavior Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years[+]
Female	Mean	86.80	85.07	85.34	85.36	87.16	86.60	85.43
	N	44	509	1,668	1,199	149	66	606.83
	Std. Deviation	9.34	9.87	9.53	9.63	8.79	8.66	9.57
Male	Mean	83.52	85.78	84.84	85.29	85.23	85.58	85.14
	N	160	1,318	4,981	4,445	568	341	1,968.83
	Std. Deviation	10.09	10.03	10.35	10.22	9.89	9.83	10.23

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 95: Statistical Test for Differences in Score for the EQ Consistency of Work Behavior Subtest as a Function of Gender

Gender	Statistical Test for Difference in Mean Scores on the EQ Consistency of Work Behavior Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female	-2.02*	1.37	-1.82	-0.23	-2.32*	-0.86	-0.64

Notes: * $p < 0.05$, ** $p < 0.01$
The male population was used as the majority comparison group. Positive Z-scores are indicative of higher EQ Consistency of Work Behavior Subtest scores in the male sample.

Table 96: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for EQ Consistency of Work Behavior Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female/Male	-0.33	0.07	-0.05	-0.01	-0.20	-0.11	-0.03

The EQ Concentration score indicates the applicant's ability to focus on job activities amid distractions for short periods of time.

Race/Ethnicity: Table 97 shows the scale means and standard deviations for each racial/ethnic subgroup that composed more than two percent of the applicant population. Table 98 shows the results of the statistical significance test of the difference in means, comparing each minority racial/ethnic subgroup to the majority White subgroup. As shown in Table 98, the scale means only showed occasional significant differences, which for the most part are in favor of minorities. The effect sizes of these subgroup differences are shown in Table 99, and they are small in size. The EQ Concentration Subtest is not a barrier for any racial or ethnic group.

Table 97: Descriptive Statistics of the EQ Concentration Subtest as a Function of Race/Ethnicity

Race/Ethnicity		EQ Concentration Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Asian	Mean	*	85.17	83.38	82.02	85.22	*	83.25
	N	*	79	207	197	34	*	129.25
	Std. Deviation	*	12.78	12.99	12.67	10.91	*	12.73
African-American	Mean	84.78	86.39	85.75	87.05	86.58	*	86.24
	N	23	494	1,670	945	47	*	635.80
	Std. Deviation	13.16	11.57	11.54	11.71	9.49	*	11.58
Hispanic	Mean	*	86.08	85.68	85.95	83.49	*	85.69
	N	*	145	425	310	60	*	235.00
	Std. Deviation	*	13.29	11.70	11.14	13.35	*	11.90
Multi	Mean	84.46	84.61	84.97	85.72	84.88	*	85.19
	N	18	139	508	422	56	*	228.60
	Std. Deviation	11.45	10.92	11.89	11.33	11.45	*	11.55
White	Mean	82.02	85.28	84.7	84.84	83.94	*	84.74
	N	155	1,054	4,215	4,066	545	*	2,007.00
	Std. Deviation	11.98	11.17	11.26	11.48	12.62	*	11.43

Note: * Asian and Hispanic applicants composed less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year. In 2011, no racial or ethnic group composed more than two percent of the applicant pool.
+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 98: Statistical Test for Differences in Scores on the EQ Concentration Subtest as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Mean Scores on the EQ Concentration Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian	N/A	0.08	1.44	3.07**	-0.65	N/A	1.29
African-American	-0.95	-1.77	-3.17**	-5.24**	-1.78	N/A	-2.86**
Hispanic	N/A	-0.69	-1.65	-1.68	0.25	N/A	-1.17
Multi	-0.85	0.68	-0.49	-1.50	-0.58	N/A	-0.56

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of higher EQ Concentration Subtest scores in the White sample.

Table 99: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for EQ Concentration Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian/White	N/A	0.01	0.12	0.25	-0.10	N/A	0.13
African-American/White	-0.23	-0.10	-0.09	-0.19	-0.21	N/A	-0.13
Hispanic/White	N/A	-0.07	-0.09	-0.10	0.04	N/A	-0.08
Multi/White	-0.21	0.06	-0.02	-0.08	-0.08	N/A	-0.04

Gender: Table 100 shows the scale means and standard deviations as a function of gender. Table 101 shows that there are only two statistically significant female/male differences (FY 2006 and FY 2010) on this scale, and they favor women. Table 102 shows that these gender differences are trivial. The EQ Concentration Subtest is not a barrier based on gender.

Table 100: Descriptive Statistics of the EQ Concentration Subtest as a Function of Gender

Gender		EQ Concentration Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Female	Mean	86.80	85.76	85.25	85.47	85.94	84.63	85.43
	N	44	509	1,668	1,199	149	66	605.83
	Std. Deviation	10.32	11.27	11.25	11.15	11.51	12.47	11.24
Male	Mean	82.03	85.63	84.88	85.10	83.64	84.03	84.93
	N	160	1,318	4,981	4,445	568	341	1,968.83
	Std. Deviation	12.82	11.56	11.48	11.72	12.32	11.99	11.66

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 101: Statistical Test for Differences in Score for the EQ Concentration Subtest as a Function of Gender

Gender	Statistical Test for Difference in Mean Scores on the EQ Concentration Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female	-2.57*	-0.22	-1.15	-1.01	-2.14*	-0.36	-0.96

Notes: * $p < 0.05$, ** $p < 0.01$
The male population was used as the majority comparison group. Positive Z-scores are indicative of higher EQ Concentration Subtest scores in the male sample.

Table 102: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for EQ Concentration Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female/Male	-0.39	-0.01	-0.03	-0.03	-0.19	-0.05	-0.04

The EQ Decisiveness Subtest indicates the applicant's ability to make effective decisions in a timely manner.

Race/Ethnicity: Table 103 shows the scale means and standard deviations for each racial/ethnic subgroup that composed more than two percent of the applicant population. Table 104 shows the results of the statistical significance test of the difference in means, comparing each minority racial/ethnic subgroup to the majority White subgroup. As shown in Table 104, the scale means only showed occasional significant differences, most of which favor minorities when considering the average across years, with the exception being for Asians. The effect sizes of these subgroup differences are shown in Table 105. This scale is a barrier to Asians.

Table 103: Descriptive Statistics of the EQ Decisiveness Subtest as a Function of Race/Ethnicity

Race/Ethnicity		EQ Decisiveness Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Asian	Mean	*	81.31	80.30	79.59	79.73	*	80.14
	N	*	79	207	197	34	*	129.25
	Std. Deviation	*	13.52	12.52	12.83	11.66	*	12.76
African-American	Mean	84.11	83.57	82.72	84.19	81.77	*	83.28
	N	23	494	1,670	945	47	*	635.80
	Std. Deviation	12.33	12.11	11.96	12.06	10.76	*	12.00
Hispanic	Mean	*	83.09	82.49	82.53	82.77	*	82.61
	N	*	145	425	310	60	*	235.00
	Std. Deviation	*	12.76	11.63	11.24	12.20	*	11.73
Multi	Mean	81.04	82.13	82.53	83.56	82.23	*	82.82
	N	18	139	508	422	56	*	228.60
	Std. Deviation	10.28	11.11	11.77	11.14	13.35	*	11.52
White	Mean	79.57	83.31	82.57	82.48	82.42	*	82.55
	N	155	1,054	4,215	4,066	545	*	2,007.00
	Std. Deviation	11.09	10.87	10.93	11.20	11.8	*	11.08

Note: * Asian and Hispanic applicants composed less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year. In 2011, no racial or ethnic group composed more than two percent of the applicant pool.

+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 104: Statistical Test for Differences in Scores on the EQ Decisiveness Subtest as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Mean Scores on the EQ Decisiveness Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian	N/A	1.29	2.56**	3.10**	1.31	N/A	2.10*
African-American	-1.67	-0.40	-0.44	-3.98**	0.40	N/A	-1.36
Hispanic	N/A	0.20	0.13	-0.08	-0.21	N/A	-0.07
Multi	-0.57	1.19	0.07	-1.90	0.11	N/A	-0.33

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of higher EQ Decisiveness Subtest scores in the White sample.

Table 105: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for EQ Decisiveness Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian/White	N/A	0.18	0.21	0.26	0.23	N/A	0.22
African-American/White	-0.40	-0.02	-0.01	-0.15	0.06	N/A	-0.07
Hispanic/White	N/A	0.02	0.01	-0.01	-0.03	N/A	-0.01
Multi/White	-0.13	0.11	0.00	-0.10	0.02	N/A	-0.02

Gender: Table 106 shows this scale means and standard deviations as a function of gender. Table 107 shows that there are three statistically significant female/male differences (FY 2006, FY 2009, FY 2010) on this scale across years. Table 108 shows that these gender differences are of a trivial-to-small effect and favor women, indicating there is no barrier based on gender.

Table 106: Descriptive Statistics of the EQ Decisiveness Subtest as a Function of Gender

Gender		EQ Decisiveness Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years[+]
Female	Mean	83.85	83.97	82.94	83.43	83.89	81.56	83.27
	N	44	509	1,668	1,199	149	66	605.83
	Std. Deviation	10.76	11.24	11.11	10.83	11.46	13.10	11.08
Male	Mean	79.90	83.02	82.37	82.44	81.79	81.67	82.39
	N	160	1,318	4,981	4,445	568	341	1,968.83
	Std. Deviation	11.54	11.43	11.41	11.64	11.80	11.88	11.53

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 107: Statistical Test for Differences in Score for the EQ Decisiveness Subtest as a Function of Gender

Gender	Statistical Test for Difference in Mean Scores on the EQ Decisiveness Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female	-2.12*	-1.62	-1.80	-2.74**	-1.99*	0.06	-1.70

Notes: * $p < 0.05$, ** $p < 0.01$
The male population was used as the majority comparison group. Positive Z-scores are indicative of higher EQ Decisiveness Subtest scores in the male sample.

Table 108: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for EQ Decisiveness Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female/Male	-0.35	-0.08	-0.05	-0.09	-0.18	0.01	-0.08

The EQ Self-Confidence Subtest assesses the applicant's belief that he/she is the right person for the job.

Race/Ethnicity: Table 109 shows the scale means and standard deviations for each racial/ethnic subgroup that composed more than two percent of the applicant population. Table 110 shows the results of the statistical significance test of the difference in means, comparing each minority racial/ethnic subgroup to the majority White subgroup. As shown in Table 110, with the exception of Asians, the scale means only showed occasional significant differences. The effect sizes of these subgroup differences are shown in Table 111, and they are below 0.20, with the exception of Asians. Therefore, the EQ Confidence Subtest is a barrier only to Asians.

Table 109: Descriptive Statistics of the EQ Self-Confidence Subtest as a Function of Race/Ethnicity

Race/Ethnicity		EQ Self-Confidence Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Asian	Mean	*	84.87	83.67	82.40	83.86	*	83.38
	N	*	79	207	197	34	*	129.25
	Std. Deviation	*	12.93	12.55	11.89	11.73	*	12.33
African-American	Mean	86.16	87.17	85.98	87.26	86.76	*	86.56
	N	23	494	1,670	945	47	*	635.80
	Std. Deviation	9.81	10.11	10.62	10.33	10.11	*	10.44
Hispanic	Mean	*	87.10	86.44	86.98	86.33	*	86.71
	N	*	145	425	310	60	*	235.00
	Std. Deviation	*	11.01	10.08	9.43	9.11	*	9.97
Multi	Mean	88.56	85.50	85.86	86.00	84.64	*	85.85
	N	18	139	508	422	56	*	228.60
	Std. Deviation	8.95	10.62	11.38	10.52	11.62	*	10.96
White	Mean	82.84	86.61	85.96	85.94	85.57	*	85.95
	N	155	1,054	4,215	4,066	545	*	2,007.00
	Std. Deviation	11.11	10.11	10.05	10.30	10.82	*	10.22

Note: * Asian and Hispanic applicants composed less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year. In 2011, no racial or ethnic group composed more than two percent of the applicant pool.

+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 110: Statistical Test for Differences in Scores on the EQ Self-Confidence Subtest as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Mean Scores on the EQ Self-Confidence Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian	N/A	1.18	2.58**	4.11**	0.83	N/A	2.32*
African-American	-1.49	-1.01	-0.08	-3.52**	-0.77	N/A	-1.28
Hispanic	N/A	-0.50	-0.94	-1.86	-0.60	N/A	-1.10
Multi	-2.50**	1.17	0.18	-0.10	0.57	N/A	0.13

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of higher EQ Self-Confidence Subtest scores in the White sample.

Table 111: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for EQ Self-Confidence Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian/White	N/A	0.17	0.22	0.34	0.16	N/A	0.25
African-American/White	-0.30	-0.06	0.00	-0.13	-0.11	N/A	-0.06
Hispanic/White	N/A	-0.05	-0.05	-0.10	-0.07	N/A	-0.08
Multi/White	-0.52	0.11	0.01	-0.01	0.09	N/A	0.01

Gender: Table 112 shows this scale means and standard deviations as a function of gender. Table 113 shows that there are no statistically significant female/male differences on this scale across years. Table 114 shows that the effect is below 0.20. The EQ Confidence Subtest is not a barrier based on gender.

Table 112: Descriptive Statistics of the EQ Self-Confidence Subtest as a Function of Gender

Gender		EQ Self-Confidence Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Female	Mean	85.46	86.33	85.62	85.80	86.56	83.06	85.77
	N	44	509	1,668	1,199	149	66	605.83
	Std. Deviation	9.99	10.49	10.42	10.33	10.42	12.83	10.45
Male	Mean	83.65	86.85	86.00	86.14	85.19	85.76	86.07
	N	160	1,318	4,981	4,445	568	341	1,968.83
	Std. Deviation	11.27	10.28	10.34	10.48	10.68	10.31	10.41

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 113: Statistical Test for Differences in Score for the EQ Self-Confidence Subtest as a Function of Gender

Gender	Statistical Test for Difference in Mean Scores on the EQ Self-Confidence Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female	-1.03	0.96	1.29	0.99	-1.42	1.61	0.62

Notes: $^{*}p < 0.05$, $^{**}p < 0.01$
The male population was used as the majority comparison group. Positive Z-scores are indicative of higher EQ Self-Confidence Subtest scores in the male sample.

Table 114: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for EQ Self-Confidence Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female/Male	-0.16	0.05	0.04	0.03	-0.13	0.25	0.03

The EQ Interpersonal Tolerance Subtest score indicates the applicant's ability to accommodate or deal with criticisms, interpersonal conflicts, and differences in personalities in the work environment.

Race/Ethnicity: Table 115 shows the scale means and standard deviations for each

racial/ethnic subgroup that composed more than two percent of the applicant population. Table 116 shows the results of the statistical significance test of the difference in means comparing each minority racial/ethnic subgroup to the majority White subgroup. As shown in Table 116, the scale means only showed a significant advantage for the African-Americans and the Multiracial subgroup over the White subgroup. The effect sizes of these subgroup differences are shown in Table 117. They are small and favor minority groups. This subtest is not a barrier to any racial or ethnic minority group.

Table 115: Descriptive Statistics of the EQ Interpersonal Tolerance Subtest as a Function of Race/Ethnicity

Race/Ethnicity		EQ Interpersonal Tolerance Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Asian	Mean	*	88.76	87.16	87.18	90.29	*	87.62
	N	*	79	207	197	34	*	129.25
	Std. Deviation	*	10.12	9.75	9.50	7.14	*	9.58
African-American	Mean	90.52	89.98	89.56	90.75	88.68	*	89.97
	N	23	494	1,670	945	47	*	635.80
	Std. Deviation	7.91	8.69	8.52	8.19	9.36	*	8.46
Hispanic	Mean	*	89.64	88.56	89.54	90.33	*	89.16
	N	*	145	425	310	60	*	235.00
	Std. Deviation	*	9.03	8.85	8.32	8.30	*	8.68
Multi	Mean	88.89	89.44	89.32	89.94	89.75	*	89.58
	N	18	139	508	422	56	*	228.60
	Std. Deviation	11.57	8.35	8.85	8.59	8.88	*	8.74
White	Mean	86.34	88.17	88.15	88.33	88.92	*	88.24
	N	155	1,054	4,215	4,066	545	*	2,007.00
	Std. Deviation	9.43	8.89	8.93	9.09	9.02	*	9.01

Note: * Asian and Hispanic applicants composed less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year. In 2011, no racial or ethnic group composed more than two percent of the applicant pool.

+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 116: Statistical Test for Differences in Scores on the EQ Interpersonal Tolerance Subtest as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Mean Scores on the EQ Interpersonal Tolerance Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian	N/A	-0.51	1.44	1.66	-1.07	N/A	0.72
African-American	-2.31*	-3.81**	-5.64**	-8.03**	0.17	N/A	-4.44**
Hispanic	N/A	-1.85	-0.90	-2.47**	-1.24	N/A	-1.54
Multi	-0.90	-1.68	-2.81**	-3.65**	-0.66	N/A	-2.19*

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of higher EQ Interpersonal Tolerance Subtest scores in the White sample.

Table 117: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for EQ Interpersonal Tolerance Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian/White	N/A	-0.07	0.11	0.13	-0.15	N/A	0.07
African-American/White	-0.45	-0.21	-0.16	-0.27	0.03	N/A	-0.20
Hispanic/White	N/A	-0.17	-0.05	-0.14	-0.16	N/A	-0.10
Multi/White	-0.26	-0.14	-0.13	-0.18	-0.09	N/A	-0.15

Gender: Table 118 shows the scale means and standard deviations as a function of gender. Table 119 shows that there are two statistically significant female/male differences on this scale, both in favor of women. However, these were years with a substantial number of applicants, and the average difference is statistically significant. Table 120 shows that these gender differences are trivial, indicating that this subtest is not a barrier based on gender.

Table 118: Descriptive Statistics of the EQ Interpersonal Tolerance Subtest as a Function of Gender

Gender		EQ Interpersonal Tolerance Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Female	Mean	89.23	89.54	89.08	89.53	90.19	88.82	89.34
	N	44	509	1,668	1,199	149	66	605.83
	Std. Deviation	8.61	8.27	8.40	8.52	8.49	8.10	8.43
Male	Mean	86.48	88.75	88.34	88.67	88.80	88.23	88.50
	N	160	1,318	4,981	4,445	568	341	1,968.83
	Std. Deviation	9.64	8.93	9.02	9.02	8.98	8.88	9.02

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 119: Statistical Test for Differences in Score for the EQ Interpersonal Tolerance Subtest as a Function of Gender

Gender	Statistical Test for Difference in Mean Scores on the EQ Interpersonal Tolerance Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female	-1.83	-1.79	-3.08**	-3.06**	-1.76	-0.53	-2.09*

Notes: * $p < 0.05$, ** $p < 0.01$
The male population was used as the majority comparison group. Positive Z-scores are indicative of higher EQ Interpersonal Tolerance Subtest scores in the male sample.

Table 120: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for EQ Interpersonal Tolerance Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female/Male	-0.29	-0.09	-0.08	-0.10	-0.16	-0.07	-0.09

The EQ Execution Subtest score indicates the applicant's ability to take timely action to avoid problems and to solve existing problems.

Race/Ethnicity: Table 121 shows the scale means and standard deviations for each racial/ethnic subgroup that composed more than two percent of the applicant population. Table

122 shows the results of the statistical significance test of the difference in means, comparing each minority racial/ethnic subgroup to the majority White subgroup. As shown in Table 122, the scale means only showed one statistically significant difference and that favored African-Americans. The effect size for this subgroup difference is shown in Table 123. The EQ Execution Subtest is not a barrier based on race or ethnicity.

Table 121: Descriptive Statistics of the EQ Execution Subtest as a Function of Race/Ethnicity

Race/Ethnicity		EQ Execution Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Asian	Mean	*	84.51	84.61	83.31	84.82	*	84.11
	N	*	79	207	197	34	*	129.25
	Std. Deviation	*	12.05	10.46	10.86	9.71	*	10.84
African-American	Mean	85.13	84.78	83.65	84.80	84.26	*	84.19
	N	23	494	1,670	945	47	*	635.80
	Std. Deviation	11.00	10.65	11.22	10.89	10.25	*	11.03
Hispanic	Mean	*	86.54	84.60	84.81	84.60	*	84.97
	N	*	145	425	310	60	*	235.00
	Std. Deviation	*	10.80	10.83	10.71	10.78	*	10.79
Multi	Mean	84.56	83.93	83.89	84.99	86.14	*	84.42
	N	18	139	508	422	56	*	228.60
	Std. Deviation	9.55	10.01	10.91	10.56	11.39	*	10.68
White	Mean	81.61	84.58	84.10	84.04	84.14	*	84.09
	N	155	1,054	4,215	4,066	545	*	2,007.00
	Std. Deviation	11.23	10.44	10.38	10.51	11.33	*	10.50

Note: * Asian and Hispanic applicants composed less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year. In 2011, no racial or ethnic group composed more than two percent of the applicant pool.

+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 122: Statistical Test for Differences in Scores on the EQ Execution Subtest as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Mean Scores on the EQ Execution Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian	N/A	0.05	-0.68	0.93	-0.39	N/A	-0.02
African-American	-1.43	-0.35	1.41	-1.95*	-0.07	N/A	-0.20
Hispanic	N/A	-2.06	-0.90	-1.22	-0.31	N/A	-1.18
Multi	-1.21	0.72	0.42	-1.76	-1.25	N/A	-0.44

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of higher EQ Execution Subtest scores in the White sample.

Table 123: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for EQ Execution Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian/White	N/A	0.01	-0.05	0.07	-0.06	N/A	0.00
African-American/White	-0.31	-0.02	0.04	-0.07	-0.01	N/A	-0.01
Hispanic/White	N/A	-0.19	-0.05	-0.07	-0.04	N/A	-0.08
Multi/White	-0.27	0.06	0.02	-0.09	-0.18	N/A	-0.03

Gender: Table 124 shows this scale means and standard deviations as a function of gender. Table 125 shows that there are only two years for which statistically significant female/male differences were found, and they favor women. Table 126 shows that these gender differences are trivial based on the average across years. The EQ Execution Subtest is not a barrier based on gender.

Table 124: Descriptive Statistics of the EQ Execution Subtest as a Function of Gender

Gender		EQ Execution Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Female	Mean	85.14	85.41	84.08	85.21	86.56	85.48	84.78
	N	44	509	1,668	1,199	149	66	605.83
	Std. Deviation	10.20	10.52	10.70	10.34	10.34	9.47	10.51
Male	Mean	81.76	84.62	84.01	83.92	83.71	83.81	83.99
	N	160	1,318	4,981	4,445	568	341	1,968.83
	Std. Deviation	11.50	10.52	10.62	10.68	11.16	10.98	10.68

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 125: Statistical Test for Differences in Score for the EQ Execution Subtest as a Function of Gender

Gender	Statistical Test for Difference in Mean Scores on the EQ Execution Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female	-1.89	-1.44	-0.25	-3.81**	-2.94**	-1.28	-1.61

Notes: * $p < 0.05$, ** $p < 0.01$

The male population was used as the majority comparison group. Positive Z-scores are indicative of higher EQ Execution Subtest scores in the male sample.

Table 126: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for EQ Execution Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female/Male	-0.30	-0.08	-0.01	-0.12	-0.26	-0.16	-0.07

The EQ Task Closure/Thoroughness Subtest score indicates the applicant's ability to continue an activity to completion through the coordination and inspection of work.

Race/Ethnicity: Table 127 shows the scale means and standard deviations for each racial/ethnic subgroup that composed more than two percent of the applicant population. Table 128 shows the results of the statistical significance test of the difference in means comparing each minority racial/ethnic subgroup to the majority White subgroup. As shown in Table 128, the scale means showed an advantage for African-Americans over the White subgroup based on the average across the years. The effect sizes of these subgroup differences are shown in Table 129, and they are small in size. This subtest is not a barrier for racial/ethnic minorities.

Table 127: Descriptive Statistics of the EQ Task Closure/Thoroughness Subtest as a Function of Race/Ethnicity

Race/Ethnicity		EQ Task Closure/Thoroughness Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Asian	Mean	*	83.90	81.72	81.00	80.93	*	81.73
	N	*	79	207	197	34	*	129.25
	Std. Deviation	*	11.50	11.76	11.68	11.16	*	11.67
African-American	Mean	86.39	85.45	84.59	85.64	84.65	*	85.05
	N	23	494	1,670	945	47	*	635.80
	Std. Deviation	11.11	10.93	10.99	11.00	8.38	*	10.95
Hispanic	Mean	*	84.98	83.68	84.44	82.03	*	84.02
	N	*	145	425	310	60	*	235.00
	Std. Deviation	*	11.42	10.93	10.36	10.81	*	10.83
Multi	Mean	85.09	83.78	83.81	83.71	84.32	*	83.81
	N	18	139	508	422	56	*	228.60
	Std. Deviation	8.56	10.44	10.98	10.65	11.11	*	10.77
White	Mean	79.95	84.36	83.33	83.28	83.27	*	83.36
	N	155	1,054	4,215	4,066	545	*	2,007.00
	Std. Deviation	10.81	10.39	10.63	10.73	11.52	*	10.70

Note: * Asian and Hispanic applicants composed less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year. In 2011, no racial or ethnic group composed more than two percent of the applicant pool.

+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 128: Statistical Test for Differences in Scores on the EQ Task Closure/Thoroughness Subtest as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Mean Scores on the EQ Task Closure/Thoroughness Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian	N/A	0.35	1.93	2.69**	1.18	N/A	1.55
African-American	-2.60**	-1.86	-4.00**	-5.97**	-1.05	N/A	-3.41**
Hispanic	N/A	-0.62	-0.63	-1.90	0.84	N/A	-0.89
Multi	-2.34**	0.62	-0.93	-0.79	-0.67	N/A	-0.60

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of higher EQ Task Closure/Thoroughness Subtest scores in the White sample.

Table 129: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for EQ Task Closure/Thoroughness Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian/White	N/A	0.04	0.15	0.21	0.20	N/A	0.15
African-American/White	-0.59	-0.10	-0.12	-0.22	-0.12	N/A	-0.16
Hispanic/White	N/A	-0.06	-0.03	-0.11	0.11	N/A	-0.06
Multi/White	-0.48	0.06	-0.05	-0.04	-0.09	N/A	-0.04

Gender: Table 130 shows the scale means and standard deviations as a function of gender. Table 131 shows that there are statistically significant female/male differences on this scale across years in favor of women. Table 132 shows that these gender differences are small. The EQ Task Closure/Thoroughness Subtest is not a barrier based on gender.

Table 130: Descriptive Statistics of the EQ Task Closure/Thoroughness Subtest as a Function of Gender

Gender		EQ Task Closure/Thoroughness Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Female	Mean	85.97	86.45	85.57	85.53	86.52	85.38	85.72
	N	44	509	1,668	1,199	149	66	605.83
	Std. Deviation	8.96	9.65	9.98	9.87	9.65	10.50	9.88
Male	Mean	80.38	84.20	82.92	83.13	82.37	82.10	83.06
	N	160	1,318	4,981	4,445	568	341	1,968.83
	Std. Deviation	10.99	10.78	10.98	11.04	11.34	11.86	11.03

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 131: Statistical Test for Differences in Score for the EQ Task Closure/Thoroughness Subtest as a Function of Gender

Gender	Statistical Test for Difference in Mean Scores on the EQ Task Closure/Thoroughness Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female	-3.48**	-4.32**	-9.14**	-7.30**	-4.50**	-2.27*	-5.64**

Notes: * $p < 0.05$, ** $p < 0.01$
The male population was used as the majority comparison group. Positive Z-scores are indicative of higher EQ Task Closure/Thoroughness Subtest scores in the male sample.

Table 132: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for EQ Task Closure/Thoroughness Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female/Male	-0.53	-0.21	-0.25	-0.22	-0.38	-0.28	-0.25

The EQ Flexibility Subtest score indicates the applicant's ability to adapt to changing situations or conditions.

Race/Ethnicity: Table 133 shows the scale means and standard deviations for each racial/ethnic subgroup that composed more than two percent of the applicant population. Table 134 shows the results of the statistical significance test of the difference in means, comparing each minority racial/ethnic subgroup to the majority White subgroup. As shown in Table 134, the scale means show that the Asian group was significantly lower than the White subgroup on this scale, but African-Americans and Hispanics/Latinos were significantly higher on this scale in most years than the White subgroup. The effect sizes of these subgroup differences are shown in Table 135, and they are small in size. This subtest is a barrier only for Asians.

Table 133: Descriptive Statistics of the EQ Flexibility Subtest as a Function of Race/Ethnicity

Race/Ethnicity		EQ Flexibility Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Asian	Mean	*	83.81	82.98	82.69	82.82	*	82.73
	N	*	79	207	197	34	*	129.25
	Std. Deviation	*	11.39	10.59	11.17	10.38	*	11.09
African-American	Mean	85.33	85.98	85.75	86.77	85.02	*	85.91
	N	23	494	1,670	945	47	*	635.80
	Std. Deviation	13.08	10.28	10.05	9.90	9.25	*	10.21
Hispanic	Mean	*	86.01	85.14	85.61	85.98	*	85.19
	N	*	145	425	310	60	*	235.00
	Std. Deviation	*	10.75	9.83	9.37	9.26	*	10.01
Multi	Mean	84.98	85.10	85.45	86.38	86.00	*	85.44
	N	18	139	508	422	56	*	228.60
	Std. Deviation	6.71	9.55	9.80	9.12	9.45	*	9.70
White	Mean	82.61	85.72	85.16	85.34	85.23	*	84.94
	N	155	1,054	4,215	4,066	545	*	2,007.00
	Std. Deviation	9.53	9.95	9.84	9.89	10.56	*	10.04

Note: * Asian and Hispanic applicants composed less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year. In 2011, no racial or ethnic group composed more than two percent of the applicant pool.

+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 134: Statistical Test for Differences in Scores on the EQ Flexibility Subtest as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Mean Scores on the EQ Flexibility Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian	N/A	1.45	2.91**	3.26**	1.31	N/A	2.20*
African-American	-0.96	-0.47	-2.05*	-4.01**	0.15	N/A	-2.09*
Hispanic	N/A	-0.30	0.05	-0.49	-0.59	N/A	-0.36
Multi	-1.35	0.71	-0.62	-2.21*	-0.58	N/A	-0.74

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of higher EQ Flexibility Subtest scores in the White sample.

Table 135: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for EQ Flexibility Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian/White	N/A	0.19	0.22	0.27	0.23	N/A	0.22
African-American/White	-0.27	-0.03	-0.06	-0.15	0.02	N/A	-0.10
Hispanic/White	N/A	-0.03	0.00	-0.03	-0.07	N/A	-0.03
Multi/White	-0.26	0.06	-0.03	-0.11	-0.07	N/A	-0.05

Gender: Table 136 shows the scale means and standard deviations as a function of gender. Table 137 shows that there are only two statistically significant female/male differences on this scale. Table 138 shows that these gender differences are trivial. This subtest is not a barrier based on gender.

Table 136: Descriptive Statistics of the EQ Flexibility Subtest as a Function of Gender

Gender		EQ Flexibility Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Female	Mean	86.92	86.11	85.57	85.68	86.56	85.76	85.74
	N	44	509	1,668	1,199	149	66	605.83
	Std. Deviation	8.47	10.14	9.66	9.75	10.09	9.93	9.77
Male	Mean	82.63	85.71	85.17	85.50	84.72	84.79	85.29
	N	160	1,318	4,981	4,445	568	341	1,968.83
	Std. Deviation	10.13	10.00	9.99	9.94	10.22	9.99	9.98

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 137: Statistical Test for Differences in Score for the EQ Flexibility Subtest as a Function of Gender

Gender	Statistical Test for Difference in Mean Scores on the EQ Flexibility Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female	-2.85**	-0.77	-1.45	-0.56	-1.97*	-0.72	-1.00

Notes: * $p < 0.05$, ** $p < 0.01$
The male population was used as the majority comparison group. Positive Z-scores are indicative of higher EQ Flexibility Subtest scores in the male sample.

Table 138: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for EQ Flexibility Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female/Male	-0.44	-0.04	-0.04	-0.02	-0.18	-0.10	-0.05

The ***EQ*** Self-Awareness Subtest score indicates the applicant's ability to maintain an internal awareness of his/her actions and attitudes, including knowing his/her limitations.

Race/Ethnicity: Table 139 shows the scale means and standard deviations for each racial/ethnic subgroup that composed more than two percent of the applicant population. Table 140 shows the results of the statistical significance test of the difference in means, comparing each minority racial/ethnic subgroup to the majority White subgroup. As shown in Table 140, the scale means show significant differences between African-Americans and Asians when compared to the White subgroup. The effect sizes of these subgroup differences are shown in Table 141, and they are small in size. This subtest is not a barrier based on race or ethnicity.

Table 139: Descriptive Statistics of the EQ Self-Awareness Subtest as a Function of Race/Ethnicity

Race/Ethnicity		EQ Self-Awareness Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Asian	Mean	*	80.73	80.84	82.18	83.59	*	81.56
	N	*	79	207	197	34	*	129.25
	Std. Deviation	*	10.08	9.60	9.55	9.77	*	10.11
African-American	Mean	81.65	80.53	80.51	81.11	82.21	*	81.44
	N	23	494	1,670	945	47	*	635.80
	Std. Deviation	14.94	12.08	11.41	11.05	10.42	*	11.35
Hispanic	Mean	*	81.72	80.18	81.56	82.37	*	81.61
	N	*	145	425	310	60	*	235.00
	Std. Deviation	*	10.86	10.23	9.33	10.75	*	10.24
Multi	Mean	84.89	81.83	81.30	82.71	81.75	*	82.27
	N	18	139	508	422	56	*	228.60
	Std. Deviation	7.24	10.11	10.72	9.42	9.81	*	10.21
White	Mean	82.13	82.14	81.97	81.93	83.35	*	82.27
	N	155	1,054	4,215	4,066	545	*	2,007.00
	Std. Deviation	8.46	9.85	9.85	9.91	10.32	*	10.02

Note: * Asian and Hispanic applicants composed less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year. In 2011, no racial or ethnic group composed more than two percent of the applicant pool.
+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 140: Statistical Test for Differences in Scores on the EQ Self-Awareness Subtest as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Mean Scores on the EQ Self-Awareness Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian	N/A	1.45	2.91**	3.26**	1.31	N/A	2.20*
African-American	-0.96	-0.47	-2.05*	-4.01**	0.15	N/A	-2.09*
Hispanic	N/A	-0.30	0.05	-0.49	-0.59	N/A	-0.36
Multi	-1.35	0.71	-0.62	-2.21*	-0.58	N/A	-0.74

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of higher EQ Self-Awareness Subtest scores in the White sample.

Table 141: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for EQ Self-Awareness Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian/White	N/A	0.19	0.22	0.27	0.23	N/A	0.07
African-American/White	-0.27	-0.03	-0.06	-0.15	0.02	N/A	0.08
Hispanic/White	N/A	-0.03	0.00	-0.03	-0.07	N/A	0.07
Multi/White	-0.26	0.06	-0.03	-0.11	-0.07	N/A	0.00

Gender: Table 142 shows the scale means and standard deviations as a function of gender. Table 143 shows that there are three statistically significant female/male differences on this scale and that the overall average is statistically significant in favor of women. Table 144 shows that these gender differences are small. Therefore, we conclude that the EQ Self-Awareness Subtest is not a barrier based on gender.

Table 142: Descriptive Statistics of the EQ Self-Awareness Subtest as a Function of Gender

Gender		EQ Self-Awareness Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years[+]
Female	Mean	84.05	82.56	82.39	82.38	84.11	83.82	82.53
	N	44	509	1,668	1,199	149	66	605.83
	Std. Deviation	8.72	10.03	10.26	9.65	9.43	9.88	9.97
Male	Mean	81.39	81.27	81.16	81.72	82.87	83.24	81.53
	N	160	1,318	4,981	4,445	568	341	1,968.83
	Std. Deviation	9.67	10.73	10.34	10.11	10.48	8.93	10.26

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 143: Statistical Test for Differences in Score for the EQ Self-Awareness Subtest as a Function of Gender

Gender	Statistical Test for Difference in Mean Scores on the EQ Self-Awareness Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female	-1.75	-2.42*	-4.24**	-2.08*	-1.39	-0.44	-2.14*

Notes: * $p < 0.05$, ** $p < 0.01$

The male population was used as the majority comparison group. Positive Z-scores are indicative of higher EQ Self-Awareness Subtest scores in the male sample.

Table 144: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for EQ Self-Awareness Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female/Male	-0.28	-0.12	-0.12	-0.07	-0.12	-0.06	-0.10

The EQ Sustained Attention Subtest score indicates the applicant's ability to stay focused on tasks for periods of time over 60 minutes.

Race/Ethnicity: Table 145 shows the scale means and standard deviations for each racial/ethnic subgroup that composed more than two percent of the applicant population. Table 146 shows the results of the statistical significance test of the difference in means comparing each minority racial/ethnic subgroup to the majority White subgroup. As shown in Table 146, the scale means showed no statistically significant differences based upon the average across years. The effect sizes of these subgroup differences are shown in Table 147, and they are trivial. The EQ Sustained Attention Subtest is not a barrier based on race/ethnicity.

Table 145: Descriptive Statistics of the EQ Sustained Attention Subtest as a Function of Race/Ethnicity

Race/Ethnicity		EQ Sustained Attention Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Asian	Mean	*	87.06	85.28	83.96	85.88	*	85.09
	N	*	79	207	197	34	*	129.25
	Std. Deviation	*	12.22	11.70	11.47	12.10	*	11.74
African-American	Mean	85.30	86.12	85.60	87.16	84.30	*	86.12
	N	23	494	1,670	945	47	*	635.80
	Std. Deviation	16.18	12.50	12.36	11.72	11.6	*	12.21
Hispanic	Mean	*	86.57	85.28	85.96	83.90	*	85.61
	N	*	145	425	310	60	*	235.00
	Std. Deviation	*	11.89	12.10	11.22	12.84	*	11.84
Multi	Mean	84.44	85.96	86.04	86.98	85.86	*	86.34
	N	18	139	508	422	56	*	228.60
	Std. Deviation	11.20	10.57	11.48	10.49	11.25	*	11.00
White	Mean	81.97	86.14	85.47	85.43	84.91	*	85.44
	N	155	1,054	4,215	4,066	545	*	2,007.00
	Std. Deviation	11.55	10.56	10.78	10.98	11.80	*	10.91

Note: * Asian and Hispanic applicants composed less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year. In 2011, no racial or ethnic group composed more than two percent of the applicant pool.
+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 146: Statistical Test for Differences in Scores on the EQ Sustained Attention Subtest as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Mean Scores on the EQ Sustained Attention Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian	N/A	-0.65	0.24	1.76	-0.46	N/A	0.33
African-American	-0.95	0.04	-0.38	-4.13**	0.34	N/A	-1.26
Hispanic	N/A	-0.41	0.31	-0.80	0.58	N/A	-0.22
Multi	-0.88	0.19	-1.05	-2.87**	-0.60	N/A	-1.17

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of higher EQ Sustained Attention Subtest scores in the White sample.

Table 147: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for EQ Sustained Attention Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Asian/White	N/A	-0.09	0.02	0.13	-0.08	N/A	0.03
African-American/White	-0.27	0.00	-0.01	-0.16	0.05	N/A	-0.06
Hispanic/White	N/A	-0.04	0.02	-0.05	0.09	N/A	-0.02
Multi/White	-0.22	0.02	-0.05	-0.14	-0.08	N/A	-0.08

Gender: Table 148 shows the scale means and standard deviations as a function of gender. Table 149 shows that there is no statistically significant female/male difference on this scale based upon the average across years. Table 150 shows that gender differences are trivial. This subtest is not a barrier based on gender.

Table 148: Descriptive Statistics of the EQ Sustained Attention Subtest as a Function of Gender

Gender		EQ Sustained Attention Subtest						
		2006	2007	2008	2009	2010	2011	Weighted Average Across Years+
Female	Mean	86.73	86.19	85.65	85.74	86.26	86.79	85.81
	N	44	509	1,668	1,199	149	66	605.83
	Std. Deviation	10.45	11.07	11.26	10.91	10.79	10.88	11.08
Male	Mean	82.56	86.34	85.47	85.77	84.43	84.70	85.57
	N	160	1,318	4,981	4,445	568	341	1,968.83
	Std. Deviation	12.41	11.23	11.31	11.19	11.96	11.83	11.32

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 149: Statistical Test for Differences in Score for the EQ Sustained Attention Subtest as a Function of Gender

Gender	Statistical Test for Difference in Mean Scores on the EQ Sustained Attention Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female	-2.24*	0.26	-0.55	0.10	-1.79	-1.41	-0.47

Notes: * $p < 0.05$, ** $p < 0.01$
The male population was used as the majority comparison group. Positive Z-scores are indicative of higher EQ Sustained Attention Subtest scores in the male sample.

Table 150: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for EQ Sustained Attention Subtest						
	2006	2007	2008	2009	2010	2011	Average Across Years
Female/Male	-0.35	0.01	-0.02	0.00	-0.16	-0.18	-0.02

BARRIER ANALYSIS OF AT-SAT:

On the basis of our review, we conclude that current use of the AT-SAT is a barrier to RNO and gender diversity for the FAA. The following components of the AT-SAT are barriers based on either race/ethnicity or gender:

- Dials Subtest
- Applied Math Subtest
- Angles Test Subtest
- The Letter Factory Subtest
  - The Letter Performance Subtest
- Air Traffic Scenarios Test (ATST)
  - ATST Efficiency Subtest
  - ATST Safety Subtest
  - ATST Procedures Subtest
- Analogies Correct Subtest

In addition, the decision to develop and use the highly qualified band has resulted in a substantial reduction in minority and gender representations in the FAA from the application sources that are required to take this test. The organization appeared not to question the consequences of this change in its score cut decision. Given that (a) the validity report indicated that the cut score of 70 was justified, (b) it warned about the adverse consequences of using the AT-SAT if alternative diversity recruiting methods were not adopted, and (c) several reports indicated that the AT-SAT was performing well in terms of adverse impact if a cut score of 70 was maintained, the FAA was not justified from a diversity perspective to use the highly qualified band to resolve a problem of too many applicants passing the test. The high passing rate was probably indicative of more underlying problems with the exam than inappropriate pass scores.

Further, as revealed in our results, several of the scales of the AT-SAT show substantial problems with regard to RNO and gender diversity. Specifically, the Dials Subtest, the Applied Math Subtest, the Angles Subtest, and the AT-Scenarios Efficiency and Safety scales caused problems with regard to both RNO and gender diversity. The Scan Subtest, the Letter Factory Situation Awareness and Performance Scales, the AT-SAT Scenarios Procedure scale, and the Analogies Correct Scale caused problems for RNO diversity. Overall, this is not surprising given that this portion of the AT-SAT is measuring cognitive ability. Steps should be explored to reduce the severity of the adverse impact of this instrument. For example, the 2001 validity report assesses the incremental validity for all of the AT-SAT scales. It should be noted that the validity report shows that not all the AT-SAT scales exhibit incremental validity (Waugh, 2001a). Specifically, as noted in the footnote for Table 5.5.2: “For $p < 0.05$ level of

significance, the incremental validity for a single scale must be greater than about 0.06 (the critical value varies from 0.055 to 0.062, depending upon the column)."

Examining Table 5.5.2 in Volume 2 of the original validity report reveals that the following AT-SAT Subtests do not show incremental validity for either the computer-based performance measure (CBPM) or the ratings criteria:

- ATST Safety Subtest
- Analogies Latency Subtest
- Dials Subtest
- Letter Factory Situational Awareness Subtest
- Letter Factory Performance Subtest

Focusing on just the rating criterion,[2] the number of cognitive Subtests that show incremental validity decreases. Thus, at the very least, the ATST: Safety scale, the Analogies: Latency, Dials Subtest, Letter Factory: Situational Awareness, and Letter Factory: Performance should be removed from the computation of the AT-SAT composite score. The 2001 validity document indicates that these scales do not provide incremental validity and these scales contribute to reduced RNO and gender diversity. The adverse impact of the final AT-SAT test could be reduced further if only the scales with incremental validity on the rating criteria were examined. The FAA should take corrective action by exploring the elimination of subtests that do not show incremental validity, and assessing the validity of various combinations of cognitive and noncognitive components in terms of subgroup differences.

Additionally, in our response to comments from the Agency regarding our initial draft report, we indicate that we are worried about the AT-SAT's validity due to lack of use of alternative forms, updating items, and potential breaches of test security. We have noted a problematic distribution shift in the AT-SAT not seen in other measures of cognitive ability (i.e., in contrast to predictions after reweighting of AT-SAT, over 90% of the applicants pass the test and the percentage of applicants passing the highly qualified cut-off has substantially increased especially in the last three years). While the underlying cause of this distribution shift is uncertain at this time, we strongly urge that all explanations for this shift warrant serious and immediate attention. It was because of the administrative burden caused by the magnitude of applicants passing the AT-SAT at this stage of the ATCS Centralized Hiring Process that the FAA adopted the highly qualified band and minimized use of the General Public application

[2] According to Borman, Hedge, Hanson, Bruskiewicz, Mogilka, & Manning (2001), the CBPM is "patterned after the situational judgment test method. The basic idea was to have an air traffic scenario appear on the computer screen, allow a little time for the problem to evolve, and then freeze the screen and ask the examinee a multiple choice question about how to respond to the problem" (p. 2). Thus, the CBPM can be conceived as a cognitive multiple choice test of air traffic job knowledge. We contend that the CBPM criterion will be more cognitively loaded and reflect maximal job performance when the ATCSs know that they are being assessed. In contrast, while they have problems, job performance ratings are reflective of typical job performance.

source, and thereby inadvertently increased the adverse impact of the ATCS Centralized Hiring Process.

DECISION POINT 3: PREPARATION OF THE REFERRAL LIST OF ELIGIBLE AND QUALIFIED APPLICANTS

The next step of the ATCS Centralized Hiring Process is the preparation of the referral list of eligible and qualified applicants.

Race/Ethnicity: Table 151 shows the descriptive statistics for the referrals for each racial/ethnic subgroup that composed more than two percent of the population over the period from 2006 to 2011. Table 152 shows the results of the statistical significance tests comparing the referral rate of each minority racial/ethnic subgroup to the majority White subgroup. As shown in this table, the White subgroup is referred significantly more often than African-Americans, Hispanics/Latinos, and the Multiracial subgroups. This is not the case for Asians.

Table 153 shows the adverse impact ratios of the referral rates. The effect sizes of these subgroup differences are shown in Table 154. They are medium for African-Americans and Multiracial subgroups but small for Asians and Hispanics/Latinos. For African Americans only, there is a statically significant difference in referral rates compared to Whites (based on the average across years). The adverse impact ratio associated with this difference is 0.61 (Tables 153) and the effect size is 0.48, we conclude that the process for preparation of the referral list of eligible and qualified applicants is a barrier only to African-Americans.

Table 151: Proportion of Applicants Passing the Referral Stage of the ATCS as a Function of Race/Ethnicity

Race/ Ethnicity		Referral						
		2006	2007	2008	2009	2010	2011	Weighted Average 2007–2011+
Asian	Mean	0.97	0.50	0.43	0.44	0.71	0.91	0.50
	N	33	254	469	408	85	92	261.60
	Std. Deviation	0.174	0.50	0.50	0.50	0.46	0.28	0.49
African-American	Mean	0.937	0.26	0.33	0.24	0.39	0.69	0.29
	N	205	3,586	4,202	3,094	363	231	2,295.20
	Std. Deviation	0.244	0.44	0.47	0.43	0.49	0.46	0.45
Hispanic	Mean	0.941	0.47	0.45	0.35	0.71	0.82	0.46
	N	135	632	926	796	163	141	531.60
	Std. Deviation	0.237	0.50	0.50	0.48	0.46	0.39	0.49
Multi	Mean	*	0.39	0.47	0.39	0.63	0.83	0.46
	N	*	668	1,165	1059	213	239	668.80
	Std. Deviation	*	0.49	0.50	0.49	0.48	0.38	0.49
White	Mean	0.94	0.51	0.54	0.46	0.68	0.83	0.54
	N	937	5,050	8,611	8,449	1,917	1,907	5,186.80
	Std. Deviation	0.237	0.50	0.50	0.50	0.47	0.38	0.49

Notes: * Multiracial applicants were less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year.
+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 152: Statistical Test for Differences in Proportions for Referral Decisions as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Proportions for Referral Decisions						
	2006	2007	2008	2009	2010	2011	Weighted Average 2007–2011
Asian	-0.72	0.31	4.91**	0.88	-0.50	-2.17**	0.69
African-American	0.16**	23.35**	22.14**	21.46**	10.48**	5.06**	16.50**
Hispanic	-0.05	1.80	5.56**	6.19**	-0.68	0.30	2.63**
Multi	N/A	5.44**	4.69**	4.08**	1.51	-0.12	3.12**

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of the referral rate being greater for the White sample.

Table 153: Adverse Impact Ratios for Referral Decisions as a Function of Race/Ethnicity

Race/ Ethnicity	Adverse Impact Ratios (4/5ths Rule) for Difference in Proportions for Referral Decisions						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Asian	1.03	0.98	0.79	0.95	1.04	1.11	0.97
African-American	1.00	0.51	0.62	0.52	0.58	0.83	0.61
Hispanic	1.00	0.92	0.82	0.75	1.04	0.99	0.90
Multi	N/A	0.78	0.87	0.86	0.93	1.00	0.89

Note: The White population was used as the majority comparison group.

Table 154: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for Referral Decisions						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Asian/White	-0.13	0.02	0.23	0.04	-0.06	-0.23	0.00
African-American/White	0.02	0.53	0.43	0.46	0.62	0.35	0.48
Hispanic/White	0.00	0.07	0.19	0.23	-0.05	0.03	0.09
Multi/White	N/A	0.22	0.15	0.13	0.11	-0.01	0.12

Gender: Table 155 shows the descriptive statistics for the referrals as a function of gender from 2006 to 2011. Table 156 shows the results of the statistical significance tests comparing the referral rate for males and females. As shown in this table, males and females significantly differ in their referral rates; however, the adverse impact ratio is above 0.80 and the effect size is below 0.20. Therefore, we conclude that the process for preparation of the referral list of eligible and qualified applicants is not a barrier based on gender.

Table 157 shows the adverse impact ratios for these referral rates. These ratios are all within the 4/5ths rule. The effect sizes of the gender differences are shown in Table 158, and they are small effects. We conclude, therefore, that the referral process for establishing eligible and qualified applicants is not a barrier based on gender.

Table 155: Proportion of Applicants Passing the Referral Stage of the ATCS as a Function of Gender

Gender		Referral						
		2006	2007	2008	2009	2010	2011	Weighted Average 2007–2011+
Female	Mean	0.49	0.35	0.41	0.345	0.59	0.82	0.40
	N	527	2,774	3,750	2,970	477	454	2,085.00
	Std. Deviation	0.50	0.48	0.49	0.475	0.49	0.39	0.48
Male	Mean	0.56	0.44	0.49	0.414	0.66	0.82	0.49
	N	1,940	6,842	10,701	10,053	2,106	2,118	6,364.00
	Std. Deviation	0.50	0.50	0.50	0.493	0.47	0.39	0.49

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 156: Statistical Test for Differences in Proportions for Referral Decisions as a Function of Gender

Gender	Statistical Test for Difference in Proportions for Referral Decisions						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Female	2.90**	8.57**	8.55**	6.72**	2.93**	-0.05	5.34**

Notes: * $p < 0.05$, ** $p < 0.01$
The male population was used as the majority comparison group. Positive Z-scores are indicative of the referral rate being greater for the male sample.

Table 157: Adverse Impact Ratios for Referral Decisions as a Function of Gender

Gender	Adverse Impact Ratios (4/5ths Rule) for Difference in Proportions for Referral Decisions						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Female	0.87	0.78	0.84	0.83	0.89	1.00	0.87

Note: The male population was used as the majority comparison group.

Table 158: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for Referral Decisions						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Female/Male	0.08	0.19	0.16	0.14	0.15	0.00	0.13

DECISION POINT 4: CENTRALIZED SELECTION PANEL

The next step of the ATCS Centralized Hiring Process is the evaluation of applicants by the Centralized Selection Panel, which involves selecting the candidates who can proceed to the interview stage.

Race/Ethnicity: Table 159 shows the descriptive statistics for the referrals for each racial/ethnic subgroup that composed more than two percent of the population over the period from 2006 to 2011. Table 160 shows the results of the statistical significance tests comparing the referral rate of each minority racial/ethnic subgroup to the majority White subgroup. As shown in this table, the White subgroup is referred significantly more often than African-Americans based on the average across years. In addition, Table 161 shows that the adverse impact ratio for this difference is below 0.80. Table 162 shows that effect size is above 0.20 based on the average across years. Therefore, we conclude the Centralized Selection Panel process is a barrier for African-Americans.

Table 159: Proportion of Applicants Passing the Selection Stage as a Function of Race/Ethnicity

Race/ Ethnicity		Selected						
		2006	2007	2008	2009	2010	2011	Weighted Average 2007-2011+
Asian	Mean	0.31	0.48	0.40	0.20	0.20	0.48	0.35
	N	32	126	200	179	60	84	129.8
	Std. Deviation	0.47	0.50	0.49	0.40	0.40	0.50	0.46
African-American	Mean	0.17	0.31	0.22	0.13	0.18	0.36	0.23
	N	192	918	1404	742	142	159	673.00
	Std. Deviation	0.38	0.46	0.41	0.34	0.39	0.48	0.41
Hispanic	Mean	0.21	0.42	0.29	0.22	0.22	0.41	0.31
	N	127	296	413	276	115	115	243.00
	Std. Deviation	0.41	0.50	0.46	0.42	0.41	0.49	0.46
Multi	Mean	*	0.42	0.40	0.23	0.22	0.42	0.34
	N	*	263	546	418	134	198	311.80
	Std. Deviation	*	0.50	0.49	0.42	0.41	0.50	0.47
White	Mean	0.213	0.44	0.41	0.20	0.19	0.49	0.35
	N	881	2553	4669	3894	1304	1575	2799.00
	Std. Deviation	0.41	0.50	0.49	0.40	0.39	0.50	0.46

Notes: * Multiracial applicants were less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year.

+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 160: Statistical Test for Differences in Selection Decisions as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Proportions for Selection Decisions						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Asian	-1.35	-0.71	0.37	0.22	-0.25	0.16	-0.04
African-American	1.27	6.87**	13.22**	4.64**	0.12	3.03	5.58**
Hispanic	0.00	0.72	4.84**	-0.89	-0.79	1.57	1.09
Multi	N/A	0.68	0.63	-1.54	-0.82	1.75	0.14

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of the selected rate being greater for the White sample.

Table 161: Adverse Impact Ratios for Selection Decisions as a Function of Race/Ethnicity

Race/ Ethnicity	Adverse Impact Ratios (4/5ths Rule) for Difference in Proportions for Selection Decisions						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Asian	1.47	1.07	0.97	0.97	1.07	0.98	1.01
African-American	0.81	0.71	0.53	0.64	0.98	0.74	0.72
Hispanic	1.00	0.95	0.70	1.11	1.16	0.84	0.95
Multi	N/A	0.95	0.97	1.16	1.16	0.86	1.02

Note: The White population was used as the majority comparison group.

Table 162: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for Selection Decisions						
	2006	2007	2008	2009	2010	2011	Average 2007-2011
Asian/White	-0.24	-0.07	0.03	0.02	-0.03	0.02	-0.01
African-American/White	0.10	0.27	0.41	0.19	0.01	0.25	0.23
Hispanic/White	0.00	0.04	0.25	-0.06	-0.08	0.15	0.06
Multi/White	N/A	0.50	0.49	-0.08	0.39	0.50	0.36

Gender: Table 163 shows the descriptive statistics for the referrals as a function of gender from 2006 to 2011. Table 164 shows the results of the statistical significance tests comparing the

referral rate for males and females. As shown in this table, there is a statistically significant difference based on the average across years. There are no real patterns of statistical significance across years.

Table 163: Proportion of Applicants Passing the Referral Stage of the ATCS as a Function of Gender

Gender		Selected						
		2006	2007	2008	2009	2010	2011	Weighted Average 2007-2011+
Female	Mean	0.26	0.38	0.30	0.17	0.23	0.45	0.30
	N	238	957	1,545	1,025	282	372	836.20
	Std. Deviation	0.44	0.49	0.46	0.38	0.42	0.50	0.45
Male	Mean	0.20	0.43	0.39	0.20	0.19	0.47	0.34
	N	1,025	3,007	5,280	4,160	1,394	1,732	3,114.60
	Std. Deviation	0.40	0.50	0.49	0.40	0.39	0.50	0.46

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 164: Statistical Test for Differences in Selection as a Function of Gender

Gender	Statistical Test for Difference in Proportions for Selection Decisions						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Female	-1.91	2.79**	6.81**	2.24**	-1.43	0.46	2.17**

Notes: * $p < 0.05$, ** $p < 0.01$
The male population was used as the majority comparison group. Positive Z-scores are indicative of the selected rate being greater for the male sample.

Table 165 shows the adverse impact ratios with regard to these referral rates. As shown, these ratios are all within the 4/5ths rule. The effect sizes of these gender differences are shown in Table 166, and they are trivial based on the average across years.

Table 165: Adverse Impact Ratios for Differences in Proportions for Selection Decisions as a Function of Gender

Gender	Adverse Impact Ratios (4/5ths Rule) for Difference in Proportions for Selection Decisions						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Female	1.28	0.88	0.76	0.85	1.19	0.97	0.93

Note: The male population was used as the majority comparison group.

Table 166: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for Selection Decisions						
	2006	2007	2008	2009	2010	2011	Average 2007-2011
Female/Male	-0.14	0.10	0.20	0.08	-0.09	0.03	0.06

Based on the results shown in table 164 through 166, we conclude that the Centralized Selection Panel process is not a barrier based on gender.

BARRIER ANALYSIS OF THE CENTRALIZED SELECTION PANELS (CSP)

We interviewed 17 participants who were involved with the CSPs. From these interviews, we learned that new CSP members do not receive formal training. Rather they receive on-the-job training from the lead CSP participant. Not surprisingly, the nature of this informal training varied. Some of the interviewees described themselves as full participants the first time, others indicated that they "reviewed applications and the lead CSP person checked in occasionally to see how it was going" whereas others indicated that they were only observers the first time. It should be noted that all participants in our discussions remarked that the HR specialists do a good job with their briefing at the beginning of the CSP sessions and that the HR specialists are available to answer any questions during the CSP sessions. One CSP interview participant mentioned that the HR specialists watched over the participant when as a new member to ensure appropriate adherence to HR policies and practices.

There were other problems identified in this decision point besides unstandardized training. First, some CSP participants indicated that they called applicant references, whereas others said that they never called references. Second, the CSP participants indicated that some applicants provided their actual GPA and AT-SAT test score on their applications. Indeed, some of the CSP participants indicated that they encourage applicants to do so when the applicant has a high AT-SAT or GPA so that they would "stick out" from the pool of applicants. Third, some CSP participants indicated that the applicant location preference was

controversial. In particular, some participants indicated the FAA might lose talented applicants if the location preference requirement was strictly adhered to, meaning they sometimes would not apply the rule. Other CSP participants whom we interviewed indicated that they understood the reasons why the location preferences requirement was instituted and that they stick to the applicant's stated preferences. At the very least, the location preference and its limitation could potentially create RNO diversity issues. Further, there are issues with regard to which CSP participants sometimes do not pay attention to the applicants stated location preferences. Finally, the CSP participants that we interviewed indicated that there are lists of names that are brought to the CSP meeting of high potential applicants. One of the CSP participants indicated during our interview that he/she has reviewed an applicant that he/she had known previously. A more formal policy regarding this potential conflict of interest needs to be established, or if already established, more formally reinforced.

DECISION POINT 5: INTERVIEW

The 17 CSP participants were also asked about their experience with the interview. Overall, CSP participants indicated that it is a situational interview. They used to receive a manual to ensure they could be certified as an interviewer. Training is now online. Several CSP participants indicated that all the interview questions are "out" and applicants come to the interview completely prepared. Indeed, one CSP participant indicated that the length of the interviews has shortened over the years because applicants are so well prepared and the interview questions have not been updated over the years. Another problem with the interview is the quality of note taking by the interviewers and the extent to which they can document their ratings. A few of the participants we interviewed indicated that when interviewers' ratings are challenged, the notes often do not justify the actual ratings that were given to the applicant. Thus, the interview is not an effective tool for the FAA in its current form. However, research shows that it can be used effectively. This requires more consistent updating of questions as well as a rigorous certification process that is maintained by a body within the FAA.

Our review of the interview data indicated that there was no adverse impact. This was due to the fact that the pass rates on this decision point are over 85% for all RNO and gender subgroups.

DECISION POINT 6: MEDICAL CLEARANCE

Table 167 shows the analyses for the medical clearance. No adverse impact was found for either RNO or gender diversity.

Table 167: Pass Rate for Medical Clearance

Race/Ethnicity	Number of People who Reached Stage	Pass Rate for General Medical	Pass Rate for MMPI
African-American	431	0.97	0.97
Hispanic	174	0.98	0.97
Multi	300	0.98	0.97
White	2,728	0.87	0.8
Gender			
Female	694	0.99	0.99
Male	2,907	0.99	0.98
Total	4,015	0.99	0.98

DECISION POINT 7: SECURITY CLEARANCE

The final decision point for applicants is the security clearance process. Applicants who successfully pass the security clearance are offered a job. This section has two components: the conditional suitability approval and the final suitability approval. For the conditional suitability, applicants pass an initial screening such as a credit check. The final suitability approval is the complete clearance including a background check.

SECURITY CLEARANCE: CONDITIONAL SUITABILITY

Race/Ethnicity: Table 168 shows the descriptive statistics for the conditional security clearance for each racial/ethnic subgroup that composed more than two percent of the population over the period from 2006 to 2011. Table 169 shows the results of the statistical significance tests comparing the conditional security clearance rate comparing each minority racial/ethnic subgroup to the majority White subgroup. As shown in this table, there were some significant differences between the White subgroup and African-Americans. The difference for African-Americans is statistically significant based on the average across years.

Table 168: Proportion of Applicants Passing the Security Clearance: Conditional Suitability Stage as a Function of Race/Ethnicity

Race/ Ethnicity		Security Clearance: Conditional Suitability						
		2006	2007	2008	2009	2010	2011	Weighted Average 2007–2011+
Asian	Mean	0.00	0.10	0.51	0.75	0.86	0.64	0.49
	N	3	62	91	40	14	42	49.80
	Std. Deviation	0.00	0.30	0.50	0.44	0.36	0.49	0.44
African-American	Mean	0.05	0.12	0.41	0.59	0.65	0.58	0.34
	N	21	390	419	132	37	65	208.60
	Std. Deviation	0.22	0.32	0.49	0.49	0.48	0.50	0.43
Hispanic	Mean	0.11	0.10	0.46	0.68	0.70	0.69	0.44
	N	9	125	168	87	30	49	91.80
	Std. Deviation	0.33	0.30	0.50	0.47	0.47	0.47	0.44
Multi	Mean	*	0.08	0.51	0.73	0.79	0.71	0.49
	N	*	157	291	122	38	94	140.40
	Std. Deviation	*	0.27	0.50	0.45	0.41	0.46	0.44
White	Mean	0.022	0.11	0.51	0.74	0.80	0.63	0.48
	N	92	1,418	2,543	985	289	819	1,210.80
	Std. Deviation	0.147	0.32	0.50	0.44	0.40	0.48	0.45

Notes: * Multi-ethnic applicants were less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year.

+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 169: Statistical Test for Differences in Security Clearance: Conditional Suitability Decisions as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Proportions for Security Clearance: Conditional Suitability						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Asian	0.26	0.25	0.00	-0.14	-0.55	-0.13	-0.11
African-American	-0.67	-0.55	3.57**	3.61**	2.08*	0.80	1.90**
Hispanic	-1.50	0.34	1.18	1.22	1.28	-0.85	0.63
Multi	N/A	1.15	-0.13	0.24	0.14	-1.53	-0.03

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of the conditional security clearance rate being greater for the White sample.

Table 170 shows the adverse impact ratios of the conditional clearance rates. None of the adverse impact ratios are below 0.80 and all are either 1 or close to 1. The effect sizes of these subgroup differences are shown in Table 171, and they are all small-to-medium effects. Given that the adverse impact ratios are 1 or close to 1 and the effect sizes are generally small, we conclude that the Security Clearance: Conditional Suitability is not a barrier based on race/ethnicity.

Table 170: Adverse Impact Ratios for in Security Clearance: Conditional Suitability Decisions as a Function of Race/Ethnicity

Race/ Ethnicity	Adverse Impact Ratios (4/5ths Rule) for Difference in Proportions for Security Clearance: Conditional Suitability						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Asian	0.00	0.91	1.00	1.01	1.08	1.02	1.00
African-American	2.18	1.09	0.81	0.80	0.81	0.92	0.89
Hispanic	5.05	0.91	0.91	0.92	0.88	1.10	0.94
Multi	N/A	0.73	1.01	0.99	0.99	1.13	0.97

Note: The White population was used as the majority comparison group.

Table 171: Effect Size Estimates of the Adverse Impact Effects

Race/Ethnicity	D-ratios for Security Clearance: Conditional Suitability						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Asian/White	0.15	-0.07	0.00	-0.02	-0.15	-0.15	-0.08
African-American/White	-0.16	0.27	0.19	0.34	0.36	0.36	0.30
Hispanic/White	-0.53	0.04	0.09	0.14	0.23	0.23	0.15
Multi/White	N/A	0.04	-0.01	0.03	0.02	0.02	0.02

Gender: Table 172 shows the descriptive statistics for the conditional clearance as a function of gender from 2006 to 2011. Table 173 shows the results of the statistical significance tests comparing the conditional clearance rate for males and females. As shown in this table, no pattern of statistical significant difference emerges over time. There is no statistically significant difference based on the average across years.

Table 172: Proportion of Applicants Passing the Security Clearance: Conditional Suitability Stage as a Function of Gender

Gender		Security Clearance: Conditional Suitability						
		2006	2007	2008	2009	2010	2011	Weighted Average 2007–2011+
Female	Mean	0.04	0.12	0.48	0.71	0.69	0.64	0.45
	N	27	422	600	234	84	185	305
	Std. Deviation	0.19	0.33	0.50	0.46	0.47	0.48	0.45
Male	Mean	0.03	0.11	0.50	0.72	0.81	0.64	0.47
	N	100	1,641	2,712	1,071	314	872	1,322
	Std. Deviation	0.17	0.31	0.50	0.45	0.39	0.48	0.44

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 173: Statistical Test for Differences in Security Clearance: Conditional Suitability Decisions as a Function of Gender

Gender	Statistical Test for Difference in Proportions for Security Clearance: Conditional Suitability						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Female	-0.26	-0.58	0.89	0.31	2.38**	0.00	0.60

Notes: * $p < 0.05$, ** $p < 0.01$
The male population was used as the majority comparison group. Positive Z-scores are indicative of the conditional security clearance rate being greater for the male sample.

Table 174 shows the adverse impact ratios for conditional clearance rates. As shown, these ratios are all within the 4/5ths rule. The effect sizes of these gender differences are shown in Table 175, and they are 0.10 based upon the average across years. Therefore, we conclude that the Conditional Security Clearance is not a barrier based on gender.

Table 174: Adverse Impact Ratios for Security Clearance: Conditional Suitability Decisions as a Function of Gender

Gender	Adverse Impact Ratios (4/5ths Rule) for Difference in Proportions for Security Clearance: Conditional Suitability						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Female	1.33	1.09	0.96	0.99	0.85	1.00	0.98

Note: The male population was used as the majority comparison group.

Table 175: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for Security Clearance: Conditional Suitability						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Female/Male	-0.04	-0.04	0.04	0.04	0.30	0.30	0.13

SECURITY CLEARANCE: FINAL SUITABILITY

Race/Ethnicity: Table 176 shows the descriptive statistics for the Security Clearance: Final Suitability for each racial/ethnic subgroup that composed more than two percent of the population over the period from 2006 to 2011. Table 177 shows the results of the statistical significance tests comparing the final suitability security clearance rate, comparing each minority racial/ethnic subgroup to the majority White subgroup. As shown in this table, there were no consistent patterns of significance. None of the differences are statistically significant based on the average across years.

Table 176: Proportion of Applicants Passing the Security Clearance: Final Suitability Stage as a Function of Race/Ethnicity

Race/ Ethnicity		Security Clearance: Final Suitability						
		2006	2007	2008	2009	2010	2011	Weighted Average 2007–2011+
Asian	Mean	0.00	0.05	0.29	0.53	0.64	0.29	0.29
	N	3	62	91	40	14	42	49.80
	Std. Deviation	0.00	0.22	0.45	0.51	0.50	0.46	0.42
African-American	Mean	0.05	0.11	0.42	0.67	0.62	0.37	0.34
	N	21	380	419	132	37	65	206.60
	Std. Deviation	0.22	0.31	0.50	0.47	0.49	0.49	0.43
Hispanic	Mean	0.11	0.09	0.40	0.59	0.40	0.37	0.35
	N	9	125	168	87	30	49	91.80
	Std. Deviation	0.33	0.28	0.49	0.50	0.50	0.49	0.45
Multi	Mean	*	0.07	0.44	0.50	0.39	0.33	0.35
	N	*	157	291	122	38	94	140.40
	Std. Deviation	*	0.26	0.50	0.50	0.50	0.47	0.45
White	Mean	0.022	0.08	0.45	0.66	0.66	0.39	0.40
	N	92	1,418	2,543	985	289	819	1,210.80
	Std. Deviation	0.147	0.28	0.50	0.48	0.48	0.49	0.45

Notes: * Multiethnic applicants were less than two percent of the applicant pool in 2006 and thus, consistent with the 1978 Uniform Guidelines, were not analyzed for that year.

+ The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 177: Statistical Test for Differences in Security Clearance: Final Suitability as a Function of Race/Ethnicity

Race/ Ethnicity	Statistical Test for Difference in Proportions for Security Clearance: Final Suitability						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Asian	0.26	0.86	3.00**	1.70	0.15	1.30	1.40
African-American	-0.67	-1.85	0.88	-0.23	0.48	0.32	-0.08
Hispanic	-1.50	-0.39	1.16	1.32	2.82**	0.28	1.04
Multi	N/A	0.44	0.16	3.48	3.24**	1.13	1.69

Notes: * $p < 0.05$, ** $p < 0.01$
The White population was used as the majority comparison group. Positive Z-scores are indicative of the final security clearance rate being greater for the White sample.

Table 178 shows the adverse impact ratios of the final suitability clearance rates. The adverse impact ratios are near or above 0.80. The effect sizes of these subgroup differences are shown in Table 179, and they are all small effects. We conclude that the Security Clearance: Final Suitability is not a barrier based on race/ethnicity.

Table 178: Adverse Impact Ratios for Security Clearance: Final Suitability Decisions as a Function of Race/Ethnicity

Race/ Ethnicity	Adverse Impact Ratios (4/5ths Rule) for Difference in Proportions for Security Clearance: Final Suitability						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Asian	0.00	0.63	0.64	0.80	0.97	0.74	0.76
African-American	2.18	1.38	0.95	1.02	0.94	0.95	1.05
Hispanic	5.05	1.13	0.90	0.89	0.61	0.95	0.90
Multi	N/A	0.88	0.99	0.76	0.59	0.85	0.81

Note: The White population was used as the majority comparison group.

Table 179: Effect Size Estimates of the Adverse Impact Effects

Race/ Ethnicity	D-ratios for Security Clearance: Final Suitability						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Asian/White	0.15	0.11	0.32	0.28	0.03	0.03	0.15
African-American/White	-0.16	-0.11	0.05	-0.02	0.08	0.08	-0.02
Hispanic/White	-0.53	-0.04	0.09	0.15	0.54	0.54	0.26
Multi/White	N/A	0.04	0.01	0.33	0.55	0.55	0.30

Gender: Table 180 shows the descriptive statistics for the conditional clearance as a function of gender from 2006 to 2011. Table 181 shows the results of the statistical significance tests comparing the conditional clearance rate for males and females. As shown in this table, no statistically significant differences were found. The adverse impact ratio is above 0.80 and the effect size is near zero. We conclude that the Security Clearance: Final Suitability is not a barrier based on gender.

Table 180: Proportion of Applicants Passing the Security Clearance: Final Suitability Stage as a Function of Gender

Gender		Security Clearance: Final Suitability						
		2006	2007	2008	2009	2010	2011	Weighted Average 2007–2011+
Female	Mean	0.04	0.08	0.41	0.61	0.60	0.42	0.36
	N	27	422	600	234	84	185	305
	Std. Deviation	0.19	0.28	0.49	0.49	0.49	0.49	0.44
Male	Mean	0.03	0.08	0.44	0.64	0.61	0.36	0.38
	N	100	1,641	2,712	1,071	314	872	1,322
	Std. Deviation	0.17	0.27	0.50	0.48	0.49	0.48	0.45

Note: + The average per year is a weighted average computed by summing the yearly proportion multiplied by the sample size for that year. This summation is divided by the total sample size for an ethnicity across the six years of data. The average standard deviation was computed by first converting the yearly entries into variances and multiplying them by their degrees of freedom. This summed product was then divided by the total sample size minus 6 and the square root of the result was obtained to yield the average standard deviation.

Table 181: Statistical Test for Differences in Security Clearance: Final Suitability Decisions as a Function of Gender

Gender	Statistical Test for Difference in Proportions for Security Clearance: Final Suitability						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Female	-0.26	0.00	1.34	0.86	0.17	-1.53	0.17

Notes: $^{*}p < 0.05$, $^{**}p < 0.01$
The male population was used as the majority comparison group. Positive Z-scores are indicative of the final security clearance rate being greater for the male sample.

Table 182 shows the classic adverse impact ratios with regard to these conditional clearance rates. As shown, these ratios are all within the 4/5ths rule. The effect sizes of these gender differences are shown in Table 183, and they are small effects.

Table 182: Adverse Impact Ratios for Security Clearance: Final Suitability Decisions as a Function of Gender

Gender	Adverse Impact Ratios (4/5ths Rule) for Difference in Proportions For Security Clearance: Final Suitability						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Female	1.33	1.00	0.93	0.95	0.98	1.17	1.01

Note: The male population was used as the majority comparison group.

Table 183: Effect Size Estimates of the Adverse Impact Effects

Gender	D-ratios for Security Clearance: Final Suitability						
	2006	2007	2008	2009	2010	2011	Average 2007–2011
Female/Male	-0.04	-0.01	0.06	0.09	0.03	0.03	0.04

RECOMMENDED ROOT CAUSE CORRECTIVE ACTIONS TO ELIMINATE BARRIERS

Evidence of barriers was found for racial/ethnic minorities at the first four decision points.

1. Qualification determination of applications
2. AT-SAT testing phase
3. Preparation of the referral list of eligible and qualified applicants
4. Centralized Selection Panel determination of applicants to proceed to interview

No evidence of barriers was found for the other decision points.

5. Interview
6. Medical clearance process

7. Security clearance process

Based upon our review, we recommend that a positive climate for RNO and gender diversity be established. RNO and gender diversity needs to be considered a high priority and any changes to the ATCS selection system should be documented prior to adoption. We further recommend the following corrective actions:

- The technical quality of the entire ATCS Centralized Hiring Process should be improved.
  - The process needs standardization, monitoring, and overall improvement on an on-going basis.
  - A committee should be established that is responsible for monitoring and improving the ATCS Centralized Hiring Process.
    - The committee could be an existing FAA unit/division or a committee of persons drawn from various units throughout the FAA who have some responsibility for input with regard to ATCS hiring.

- There should be continued efforts to reach out to the national community to educate populations about the job.

- Qualification Decision Point:
  - There needs to be a more standardized process for HR specialists to evaluate the qualifications of applicants.
  - CTI School placement needs to be reconsidered with regard to diversity
  - CTI School Differential Effectiveness
    - Barr et al. (2011) report that not all AT-CTI programs are equal. As indicated in that report, the FAA needs to track the success of the graduates from each school. AT-CTI schools differ in equipment (e.g., simulations time and equipment) and the length of training (e.g., two year versus four year) given to the students. The Barr et al. (2011) report recommends categorizing the AT-CTI schools by levels based upon the strength of the program's curriculum. Analyzing the quality of these schools as well as their diversity is highly recommended.

- RNO and gender diversity should be explicitly considered when determining the sources for the applicants in each upcoming recruitment year.

- Efforts need to continue to reach out to diverse RNO groups and to females/males to provide information about the nature of ATCS position.

- AT-SAT Decision Point:

  - The agency should not use a primarily cognitively loaded test to select applicants from the General Public for a job in which required competencies are trained.

- There is little if any evidence that the AT-SAT predicts subsequent on-the-job performance of actual applicants, therefore, evidence of its validity for applicants should be generated.
- Use of a highly qualified band creates RNO and gender diversity problems and is not justified. It is recommended that a cut score of 70 be used.
- There are alternative ways of combining multiple predictors to handle applicant volume.
- Our overall recommendation is to revise the way AT-SAT is used for the short-term, and long-term to replace AT-SAT.
  - Update items for those components of the test that can be shown to be valid.
  - Revise the test so that it assesses cognitive abilities in the context of how those abilities are used on the job while keeping RNO and gender diversity in mind.
  - Conduct a search for additional predictors to add to the selection protocol.
  - Apply a multihurdle approach to reduce adverse impact.
  - Complete any validation projects associated with AT-SAT immediately and use that information to revise the hiring process.

- CSP Decision Point:

  - Remove inconsistencies in the CSP process
    - Either prohibit calls to applicant references or mandate that all references be called.
  - Ensure that everyone understands the goals of the CSP
    - Some CSP participants assume that applicants have already been screened (are already qualified for job), and think their job is to identify facility/geographical placement.
  - Reconsider the use of location preferences of applicants
  - Develop a standardized training process for new CSP members
  - Control the information that the CSP receives
    - Exploratory items on the application should be eliminated so that CSP participants do not see them.
    - Applicants' self-entry of AT-SAT and other scores on their application should be prevented.
  - A policy needs to be established and/or enforced regarding conflicts of interest in CSP recruitment (e.g., forwarding an applicant to next stage of process because a CSP member knows the applicant personally).

- Interview Decision Point:

  - The interview needs to be standardized and brought up to professional standards.
  - New questions need to be written and updated frequently.

- Rigorous frame of reference training needs to be implemented.
- Standards need to be established for who is qualified and motivated to be an interviewer.
- Periodic recertification of interviewer skills is needed.

Study Limitations/Challenges

- The barrier analysis was underway for several years and, due to a variety of circumstances, the final product was rendered unacceptable. It did not meet the needs of Federal Aviation Administration and/or the other stakeholders.

  Outtz and Associates were brought in on a very short turnaround and faced with a great deal of challenges at the onset of this undertaking. The ATCS Centralized Hiring Process roles and responsibilities are channeled throughout several organizations that are subsets to ATO. Many are housed within ATO but several are divisions within the human resources community and have a unique relationship to the hiring process. All these partner and stakeholder roles were not always apparent or well defined. The data that were necessary to complete the barrier analysis were drawn from all these sources and identifying which source retained the data was time consuming and challenging. As such, there will be lessons learned from this process that will serve the FAA as it proceeds with the barrier analysis process.

- Finding the expert in various stages of the process and being able to convey the stage of the process with clarity and ownership was not available to the contractors early on in the collection of data.

- The data was not available at the time the contract was awarded for a great number of reasons. We learned that the data and the data collection process changed as decisions involving the hiring sources changed. Identifying all of the changes became problematic.

- We had to identify the principals, roles, and responsibilities early on and acquiring participation from all parties (information offices) was necessary to alleviate contract work stoppage. The FAA intervened several times to ensure that it was working on locating the data and getting it to us within the time constraints.

- The Secretary of Transportation and the Acting Administrator of the Federal Aviation Administration were contacted by leading advocates of civil rights organizations on the lack of progress in producing any tangible reports of a barrier analysis on a major occupation that has such vital national interest of the national airspace. This exacerbated the time limitations under which the study had to be conducted.

- There were a number of external stakeholders who expressed their desire to have the barrier analyses completed through Congress, national forums, and organizations that advocate actively against employment barriers and discrimination. Several meetings were held with these groups on an abbreviated timeline to complete the barrier analysis due to the concerns that each of these groups brought to the FAA. We heard from other stakeholders, congressional offices representing their constituents, and two congressmen. These stakeholders were dissatisfied with the lack of timeliness of producing results and placed stringent time lines to complete the product in fiscal year 2012.

REFERENCES

Barr, M., Brady, T, Koleszar, G., New, M., & Pounds, J. (2011). *FAA Independent Review Panel on the Selection, Assignment, and Training of Air Traffic Control Specialists (Final Report).* September 22, 2011.

Borman, W.C., Hedge, J.W., Hanson, M.A., Bruskiewicz, K.T., Mogilka, H., & Manning, C. (2001). Development of Criterion Measures of Air traffic Controller Performance. 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). Washington, DC: Federal Aviation Administration Office of Aviation Medicine. http://www.faa.gov/data_research/research/med_humanfacs/oamtechreports/2000s/2001/.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences (Second Ed).* Hillsdale, NJ: Lawrence Erlbaum.

Dattel, A.R. & King, R.E. (2006). Reweighting AT-SAT to Mitigate Group Score Differences (DOT/FAA/AM-06/16). Washington, DC: FAA Office of Aerospace Medicine.http://www.faa.gov/data_research/research/med_humanfacs/oamtechreports/2000s/media/200616.pdf.

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.

Hanges, P. J., Salmon, E.D., & Aiken, J.R. (in press). Legal Issues in Industrial Testing and Assessment. In K.F. Geisinger (Ed.-in-Chief), B.A. Bracken, J.F. Carlson, J.C. Hansen, N.R. Kuncel, S.P. Reise, & M.C. Rodriguez (Assoc. Eds.), *APA Handbooks in Psychology: APA Handbook of Testing and Assessment in Psychology*: *Vol. 1. Test Theory and Testing and Assessment in Industrial and Organizational Psychology*. Washington, DC: American Psychological Association.

Hanges, P.J. & Wang, M. (2012). Seeking the Holy Grail in Organizational Science: Uncovering Causality through Research Design. (pp. 79-116). In S.W.J. Kozlowski (Ed.), *The Oxford Handbook of Organizational Psychology.* New York: Oxford University Press.

King, R.E., Manning, C.A. & Dreschler, G.K. (2007). Operational Use of the Air Traffic Selection and Training Battery, (DOT/FAA/AM-07/14).Washington, DC: FAA Office of Aerospace Medicine. http://www.faa.gov/data_research/research/med_humanfacs/oamtechreports/2000s/media/200714.pdf

Mattson, P. R. (2006). *Air Traffic Control Career Prep: A Comprehensive Guide to One of the Best-Paying Federal Government Careers, Including Test Preparation for the Initial ATC Exams*. Newcastle, WA: Aviation Supplies and Academics.

Raudenbush, S. W. & Bryk, A.S. (2002). Hierarchical Linear Models: Applications and Data Analysis Methods (Advanced Quantitative Techniques in the Social Sciences). Thousand Oaks, CA: SAGE Publications.

Roch, Woehr, Mishra & Kieszczynska. (2012).

Rosenthal, R. (1991). *Meta-Analytic Procedures for Social Research*. Newbury Park, CA: SAGE Publications.

Schmidt, F.L. & Hunter, J.E. (1990). *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*. Newbury Park, CA: SAGE Publications.

Waugh, G. (2001a). 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). Washington, DC: Federal Aviation Administration Office of Aviation Medicine. http://www.faa.gov/data_research/research/med_humanfacs/oamtechreports/2000s/2001/.

Waugh, G. (2001b). Analyses of Group Differences and Fairness. 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). Washington, DC: Federal Aviation Administration Office of Aviation Medicine. Available from http://www.faa.gov/data_research/research/med_humanfacs/oamtechreports/2000s/2001/.

GLOSSARY OF TERMS

Adverse Impact – A substantially different rate of selection in hiring, promotion, or other employment decision that works to the disadvantage of members of a race, sex, or ethnic group.

Air Traffic Collegiate Training Initiative (AT-CTI) – The FAA has partnerships with many colleges and universities. These schools offer two- and four-year nonengineering aviation degrees that teach basic courses in air traffic control. The program is designed to provide qualified applicants to fill developmental air traffic control specialist positions.

Air Traffic Control Specialist (ATCS) – A civilian employee who is actively engaged in the separation and control of air traffic; or providing preflight, in-flight, or airport advisory service to aircraft operators. Also included in this definition is an employee who is the immediate (first-level) supervisor of an air traffic controller as defined above.

Analysis – The process of identifying a question or issue to be addressed and examining the issue, investigating the results, interpreting the results, and possibly making a recommendation. Analysis typically involves using scientific or mathematical methods for evaluation.

Applicant – Any individual who is a candidate for initial employment into an ATCS position.

Applicant Flow Data – Information reflecting characteristics of the pool of individuals applying for an employment opportunity.

Assessment – Process of measuring or judging the value or level of something.

Attrition – Number of new hires who are no longer in the ATC training program.

AVIATOR – Automated Vacancy Information Access Tool for Online Referral – Generates vacancy announcements and automatically posts them to the FAA and USAJOBS websites.

Barrier – A policy, practice, or procedure that limits or tends to limit employment opportunities for members of a particular race, gender, ethnic background, or because of a disability.

Barrier Analysis – A process that examines relevant data, trends and benchmarks to identify a policy, practice or procedure that limits or tends to limit employment opportunities.

Centralized Selection Panel (CSP) – A panel consisting of managers from selected air traffic facilities that assigns candidates using a referral list.

Certified Professional Controller (CPC) – An applicant becomes a CPC once he/she has completed all the required training and has been tested and certified on the necessary air traffic instruments.

Civilian Labor Force (CLF) – Data collected and compiled by the U.S. Census Bureau for persons 16 years of age and over, except those in the armed forces, who are employed or are unemployed and seeking work. This information is to be used as the benchmark to compare and analyze the command/activity workforce as part of the barrier analysis process.

Control Tower Operator (CTO) – A person who performs air traffic control duties at an air traffic control tower.

Facility – Generally, any installation of equipment designated to aid in the navigation, communication, or control of air traffic. Specifically, the term denotes the total electronic equipment, power generation, or distribution systems, and any structure used to house, support, and/or protect this equipment and systems. A facility may include a number of systems, subsystems, and equipment.

Hire Source – Announcement hiring source.

Hires – Number of hires from source.

Human Factors – A multidisciplinary effort to generate and compile information about human capabilities and limitations and apply that information to equipment, systems, facilities, procedures, jobs, operations, environments, training, staffing, and personnel management for safe, comfortable, efficient, and effective human performance.

Major Occupations – Agency occupations that are mission related and heavily populated relative to other populations within the agency.

Medical Examination – Any and all examinations performed under the ATCS Health Program.

National Airspace System (NAS) – Included are system components shared jointly with the military. The system's present configuration is a reflection of the technological advances concerning the speed and altitude capability of jet aircraft, as well as the complexity of microchip, and satellite-based navigation equipment.

Professional Air Traffic Controllers Organization (PATCO) – An independent labor union in the United States. It is certified by the NLRB and currently represents air traffic controllers who work in private sector control towers.

Qualitative Data – Subjective data that is expressed as a measure of quality; nominal data.

Quantitative Data – Objective data expressed as a quantity, number, or amount; allows for rational analysis and substantiation of findings.

Reinstatement Candidate (REIN) – Any former ATCS seeking reemployment into an operational ATCS position.

Relevant Civilian Labor Force (RCLF) – The source from which an agency draws or recruits applicants for employment or an internal selection, such as a promotion, will determine a more precise benchmark to use to compare the command/activity workforce.

Retired Military Controller (RMC) – Also known as PC-20, PHX-20, or Phoenix 20.

Stakeholder – A group or individual who is affected by or is in some way accountable for the outcome of an undertaking; an interested party having a right, share, or claim in a product or service, or in its success in possessing qualities that meet that party's needs and/or expectations.

Validation – The process of proving that the right system is being built, i.e., that the system requirements are unambiguous, correct, complete, and verifiable.

Verification – The process that ensures that the system requirements have been met by the design solution and the system is ready to be used in the operational environment for which it is intended.

Veterans Recruitment Appointment (VRA) – A special authority that permits agencies to appoint certain veterans noncompetitively up to FG-11 or equivalent.

APPENDIX A

AUXILIARY TABLES

TABLES 1A THROUGH 1F

Table 1A: Subgroup Ethnicity and Gender as a Function of Application Source by Year

Ethnicity	Source: Veterans Recruitment Appointment (VRA)													
	2006		2007		2008		2009		2010		2011		Weighted Average[++]	
	F_i[+]	%	F_i	%	F_i	%	F_i	%	F_i	%	F_i	%	Total	%
Asian	58	2.4%	123	2.1%	136	2.1%	86	1.9%	46	2.5%	23	2.1%	414	2.1%
African-American	308	12.8%	2,039	35.2%	2,522	38.0%	1,506	33.3%	435	24.0%	144	13.1%	6,646	34.6%
Hawaiian	23	1.0%	41	0.7%	39	0.6%	33	0.7%	15	0.8%	13	1.2%	141	0.7%
Hispanic	186	7.7%	370	6.4%	405	6.1%	263	5.8%	93	5.1%	51	4.7%	1,182	6.0%
Multi	0	0.0%	326	5.6%	461	7.0%	335	7.4%	165	9.1%	122	11.1%	1409	7.4%
Native American	20	0.8%	33	0.6%	41	0.6%	18	0.4%	11	0.6%	6	0.5%	109	0.6%
Unanswered	380	15.7%	411	7.1%	309	4.7%	197	4.4%	90	5.0%	47	4.3%	1054	5.6%
White	1,438	59.6%	2,444	42.2%	2717	41.0%	2,086	46.1%	957	52.8%	690	63.0%	8,894	45.5%

Table 1A Continued: Subgroup Ethnicity and Gender as a Function of Application Source by Year

Gender	Source: Veterans Recruitment Appointment (VRA)													
	2006		2007		2008		2009		2010		2011		Weighted Average[++]	
	F_i[+]	%	F_i	%	F_i	%	F_i	%	F_i	%	F_i	%	Total	%
Female	438	18.2%	1,304	22.5%	1,524	23.0%	577	20.0%	297	16.4%	173	15.8%	4,203	21.4%
Male	1,595	66.1%	3,714	64.2%	4,345	65.5%	2,481	68.7%	1,305	72.0%	847	77.3%	13,321	67.3%
Unanswered	380	15.7%	769	13.3%	761	11.5%	314	11.3%	210	11.6%	76	6.9%	2,325	11.9%
Total	2,413		5,787		6,630		4,524		1,812		1,096		19,849	

Note: [+]F_i. = Frequency of applicants self-identified in each category.
[++]The Weighted Averages are based on available data from FY 2007 to FY 2011.
[a] Applications from CTI applicants were received in FY 2006 and FY 2007; however, they were not entered into the Aviator system. Thus, no data were available for our analyses for the CTI application source for FY 2006 and FY 2007.

Table 1B: Subgroup Ethnicity and Gender as a Function of Application Source by Year

Ethnicity	Source: Retired Military Controller (RMC)													
	2006		2007		2008		2009		2010		2011		Weighted Average[++]	
	F_i[+]	%	F_i	%	F_i	%	F_i	%	F_i	%	F_i	%	Total	%
Asian	42	2.9%	78	1.9%	109	2.1%	93	2.3%	28	3.2%	10	2.7%	318	2.2%
African-American	228	15.8%	1,901	45.7%	2,262	43.6%	1,596	40.1%	265	30.1%	76	20.4%	6,100	42.5%
Hawaiian	8	0.6%	20	0.5%	24	0.5%	19	0.5%	3	0.3%	1	0.3%	67	0.5%
Hispanic	106	7.3%	233	5.6%	331	6.4%	215	5.4%	56	6.4%	14	3.8%	849	5.9%
Multi	0	0.0%	213	5.1%	346	6.7%	268	6.7%	67	7.6%	26	7.0%	920	6.4%
Native American	8	0.6%	23	0.6%	21	0.4%	21	0.5%	6	0.7%	1	0.3%	72	0.5%
Unanswered	264	18.3%	279	6.7%	211	4.1%	171	4.3%	32	3.6%	14	3.8%	707	5.1%
White	788	54.6%	1,414	34.0%	1,881	36.3%	1,597	40.1%	423	48.1%	231	61.9%	5,546	38.8%

Table 1B Continued: Subgroup Ethnicity and Gender as a Function of Application Source by Year

Gender	Source: Retired Military Controller (RMC)													
	2006		2007		2008		2009		2010		2011		Weighted Average[++]	
	F_i[+]	%	F_i	%	F_i	%	F_i	%	F_i	%	F_i	%	Total	%
Female	239	16.6%	1,078	25.9%	1,305	25.2%	915	23.0%	160	18.2%	64	17.2%	3522	24.4%
Male	941	65.2%	2,527	60.7%	3,267	63.0%	2,580	64.8%	621	70.6%	283	75.9%	9,278	63.8%
Gender Not Identified	264	18.3%	556	13.4%	613	11.8%	485	12.2%	99	11.3%	26	7.0%	1,779	12.3%
Total	1,444		4,161		5,185		3,980		880		373		14,579	

Note: [+]F_i. = Frequency of applicants self-identified in each category.
[++]The Weighted Averages are based on available data from FY 2007 to FY 2011.
[a] Applications from CTI applicants were received in FY 2006 and FY 2007; however, they were not entered into the Aviator system. Thus, no data were available for our analyses for the CTI application source for FY 2006 and FY 2007.

Table 1C: Ethnicity and Gender as a Function of Application Source by Year

Ethnicity	Source: Other (CTO)													
	2006		2007		2008		2009		2010		2011		Weighted Average[++]	
	F_i[+]	%	F_i	%	F_i	%	F_i	%	F_i	%	F_i	%	Total	%
Asian			4	3.4%	49	1.6%	103	2.2%	55	2.6%	28	2.4%	239	2.2%
African-American			23	19.5%	1,196	37.9%	1,376	29.3%	508	24.3%	188	15.9%	3,291	30.8%
Hawaiian			0	0.0%	14	0.4%	26	0.6%	13	0.6%	10	0.8%	63	0.6%
Hispanic			6	5.1%	181	5.7%	271	5.8%	122	5.8%	48	4.1%	628	5.6%
Multi			7	5.9%	193	6.1%	337	7.2%	144	6.9%	116	9.8%	797	7.2%
Native American			1	0.8%	19	0.6%	24	0.5%	11	0.5%	11	0.9%	66	0.6%
Unanswered			7	5.9%	139	4.4%	221	4.7%	93	4.4%	69	5.8%	529	4.7%
White			70	59.3%	1,364	43.2%	2,346	49.9%	1,147	54.8%	711	60.2%	5,638	50.7%

Table 1C Continued: Ethnicity and Gender as a Function of Application Source by Year

Gender														
Female			24	20.3%	732	23.2%	969	20.6%	401	19.2%	190	16.1%	2316	20.8%
Male			80	67.8%	2,014	63.8%	3,174	67.5%	1,434	68.5%	893	75.6%	7,595	67.7%
Unanswered			14	11.9%	409	13.0%	561	11.9%	258	12.3%	98	8.3%	1,340	12.1%
Total			118		3,155		4,704		2,093		1,181		11,251	

Note: $^{+}F_{i.}$ = Frequency of applicants self-identified in each category.

$^{++}$The Weighted Averages are based on available data from FY 2007 to FY 2011.

[a] Applications from CTI applicants were received in FY 2006 and FY 2007; however, they were not entered into the Aviator system. Thus, no data were available for our analyses for the CTI application source for FY 2006 and FY 2007.

Table 1D: Ethnicity and Gender as a Function of Application Source by Year

Ethnicity	Source: Public													
	2006		2007		2008		2009		2010		2011		Weighted Average++	
	F_i+	%	F_i	%	F_i	%	F_i	%	F_i	%	F_i	%	Total	%
Asian			297	2.3%	481	2.8%	358	3.1%					1,136	2.8%
African-American			4,993	38%	5,055	29.8%	2,385	20.6%					12,433	31.3%
Hawaiian			89	0.7	112	0.7%	73	0.6%					274	0.7%
Hispanic			754	5.7%	992	5.9%	644	5.6%					2,390	5.8%
Multi			856	6.5%	1238	7.3%	826	7.1%					2,920	7.0%
Native American			69	0.5	86	0.5%	70	0.6%					225	0.5%
Unanswered			623	4.7%	733	4.3%	534	4.6%					1,890	4.5%
White			5,469	41.6%	8242	48.7%	6,689	57.8%					20,400	49.8%

Table 1D Continued: Ethnicity and Gender as a Function of Application Source by Year

Gender														
Female			3,454	26.3%	4,043	23.9%	2,349	20.3%					9,846	23.9%
Male			8,157	62.0%	11,037	65.2%	7,957	68.7%					27,151	65.3%
Unanswered			1,539	11.7%	1,859	11.0%	1,273	11.0%					4,671	11.2%
Total			13,150		16,939		11,579						41,668	

Note: $^{+}F_{i.}$ = Frequency of applicants self-identified in each category.

$^{++}$The Weighted Averages are based on available data from FY 2007 to FY 2011.

[a] Applications from CTI applicants were received in FY 2006 and FY 2007; however, they were not entered into the Aviator system. Thus, no data were available for our analyses for the CTI application source for FY 2006 and FY 2007.

Table 1E: Ethnicity and Gender as a Function of Application Source by Year

Ethnicity	Source: Reinstatement-DOD CPC													
	2006		2007		2008		2009		2010		2011		Weighted Average[++]	
	F_i[+]	%	F_i	%	F_i	%	F_i	%	F_i	%	F_i	%	Total	%
Asian			0	0.0%	18	2.3%	13	1.4%	22	2.1%	9	2.3%	62	2.0%
African-American			5	13.2%	211	27.3%	277	29.9%	311	29.0%	40	10.2%	844	27.9%
Hawaiian			0	0.0%	3	0.4%	8	0.9%	11	1.0%	5	1.3%	27	1.0%
Hispanic			2	5.3%	48	6.2%	57	6.2%	60	5.6%	16	4.1%	183	5.8%
Multi			1	2.6%	30	3.9%	77	8.3%	74	6.9%	28	7.1%	210	7.0%
Native American			0	0.0%	7	0.9%	4	0.4%	6	0.6%	3	0.8%	20	0.7%
Unanswered			5	13.2%	61	7.9%	52	5.6%	78	7.3%	27	6.9%	223	7.2%
White			25	65.8%	394	51.0%	438	47.3%	511	47.6%	266	67.5%	1634	51.9%
Gender														
Female			6	15.8%	167	21.6%	191	20.6%	221	20.6%	67	17.0%	652	20.4%
Male			24	63.2%	490	63.5%	618	66.7%	700	65.2%	282	71.6%	2,114	66.1%
Unanswered			8	21.1%	115	14.9%	117	12.6%	152	14.2%	45	11.4%	437	13.8%
Total			38		772		926		1,073		394		3,203	

Note: [+]F_i. = Frequency of applicants self-identified in each category.

[++]The Weighted Averages are based on available data from FY 2007 to FY 2011.

[a] Applications from CTI applicants were received in FY 2006 and FY 2007; however, they were not entered into the Aviator system. Thus, no data were available for our analyses for the CTI application source for FY 2006 and FY 2007.

Table 1F: Ethnicity and Gender as a Function of Application Source by Year

Ethnicity	Source: College Training Initiative (CTI)													
	2006		2007		2008		2009		2010		2011		Weighted Average[++]	
	F_i[+]	%	F_i	%	F_i	%	F_i	%	F_i	%	F_i	%	Total	%
Asian	[a]	[a]	[a]	[a]	25	5.0%	20	3.6%	42	4.2%	67	5.1%	154	4.6%
African-American	[a]	[a]	[a]	[a]	34	6.8%	33	5.9%	51	5.1%	61	4.6%	179	5.4%
Hawaiian	[a]	[a]	[a]	[a]	0	0.0%	8	1.4%	2	0.2%	6	0.5%	16	0.9%
Hispanic	[a]	[a]	[a]	[a]	28	5.6%	41	7.3%	62	6.2%	88	6.7%	219	6.5%
Multi	[a]	[a]	[a]	[a]	28	5.6%	33	5.9%	65	6.5%	105	7.9%	231	6.9%
Native American	[a]	[a]	[a]	[a]	2	0.4%	0	0.0%	4	0.4%	3	0.2%	9	0.3%
Unanswered	[a]	[a]	[a]	[a]	31	6.2%	26	4.7%	66	6.6%	68	5.1%	191	5.7%
White	[a]	[a]	[a]	[a]	349	70.2%	398	71.2%	702	70.6%	924	69.9%	2373	70.4%
Gender														
Female	[a]	[a]	[a]	[a]	98	19.7%	93	16.6%	160	16.1%	226	17.1%	577	17.2%
Male	[a]	[a]	[a]	[a]	335	67.4%	412	73.7%	734	73.8%	1,000	75.6%	2481	73.6%
Unanswered	[a]	[a]	[a]	[a]	64	12.9%	54	9.7%	100	10.1%	96	7.3%	314	9.7%
Total					497		559		994		1,322			

Note: [+]F_i. = Frequency of applicants self-identified in each category.

[++]The Weighted Averages are based on available data from FY 2007 to FY 2011.

[a] Applications from CTI applicants were received in FY 2006 and FY 2007; however, they were not entered into the Aviator system. Thus, no data were available for our analyses for the CTI application source for FY 2006 and FY 2007.

TABLES 10A THROUGH 10F:

Table 10A: Proportion of Applicants Passing the Qualification Stage as a Function of Application Source by Year

Ethnicity	Source: Veterans Recruitment Appointment (VRA)													
	2006		2007		2008		2009		2010		2011		Weighted Average FY2007–2011++	
	F_i+	%	F_i	%	F_i	%	F_i	%	F_i	%	F_i	%	Total	%
Asian	58	53.4%	123	28.5%	136	16.2%	86	22.1%	46	30.4%	23	56.5%	414	24.9%
African-American	308	54.2%	2,039	10.0%	2,522	5.6%	1,506	6.3%	435	23.7%	144	63.9%	6,646	9.6%
Hispanic	186	66.1%	370	33.5%	405	21.0%	263	17.9%	93	50.5%	51	70.6%	1,182	28.7%
Multi	0	0.0%	326	23.3%	461	18.4%	335	20.6%	165	47.9%	122	75.4%	1,409	28.4%
Race Not Identified	380	46.8%	411	38.9%	309	27.2%	197	19.8%	90	50.0%	47	74.5%	1,054	34.4%
White	1438	56.7%	2,444	38.1%	2,717	25.6%	2,086	23.9%	957	53.5%	690	75.5%	8,894	35.5%
Gender														
Female	438	53.9%	1,304	19.9%	1524	11.7%	905	13.9%	297	37.0%	173	71.1%	4203	19.0%
Male	1595	58.3%	3,714	28.8%	4345	18.0%	3,110	18.3%	1,305	46.8%	847	74.5%	13,321	27.5%
Gender Not Identified	380	46.8%	769	29.6%	761	21.7%	509	17.7%	210	44.8%	76	68.4%	2,325	27.1%
Total	2413		5787		6630		4525		1812		1096			

Table 10B: Proportion of Applicants Passing the Qualification Stage as a Function of Application Source by Year

Ethnicity	Source: Retired Military Controller (RMC)													
	2006		2007		2008		2009		2010		2011		Weighted Average FY2007–2011++	
	F_i+	%	F_i	%	F_i	%	F_i	%	F_i	%	F_i	%	Total	%
Asian	42	16.7%	78	3.8%	109	3.7%	93	6.5%	28	17.9%	10	50.0%	318	7.2%
African-American	228	21.5%	1,901	3.1%	2,262	2.8%	1,596	3.1%	265	16.6%	76	53.9%	6,100	4.2%
Hispanic	106	23.6%	233	14.6%	331	6.0%	215	4.2%	56	23.2%	14	50.0%	849	9.8%
Multi	0	0.0%	213	3.8%	346	5.5%	268	6.3%	67	14.9%	26	53.8%	920	7.4%
Race Not Identified	264	25.4%	279	11.1%	211	10.0%	171	8.8%	32	21.9%	14	42.9%	707	11.3%
White	788	25.5%	1,414	13.1%	1,881	11.7%	1,597	9.6%	423	30.5%	231	56.3%	5,546	14.7%
Gender														
Female	239	18.4%	1,078	3.4%	1,305	3.1%	915	4.0%	160	10.0%	64	34.4%	3,522	4.3%
Male	941	26.0%	2,527	9.7%	3,267	7.9%	2,580	6.8%	621	27.2%	283	59.0%	9,278	10.9%
Gender Not Identified	264	25.4%	556	7.6%	613	8.6%	485	9.1%	99	24.2%	26	57.7%	1,779	10.0%
Total	1,444		4,161		5,185		3,980		880					

Table 10C: Proportion of Applicants Passing the Qualification Stage as a Function of Application Source by Year

Ethnicity	Source: Other (CTO)													
	2006		2007		2008		2009		2010		2011		Weighted Average FY2007–2011[++]	
	F_i[+]	%	F_i	%	F_i	%	F_i	%	F_i	%	F_i	%	Total	%
Asian			4	75.0%	49	10.2%	103	85.4%	55	60.0%	28	64.3%	239	61.5%
African-American			23	78.3%	1,196	6.4%	1,376	90.3%	508	40.0%	188	45.2%	3,291	49.4%
Hispanic			6	83.3%	181	16.0%	271	87.8%	122	48.4%	48	72.9%	628	58.3%
Multi			7	100.0%	193	13.0%	337	95.0%	144	59.7%	116	68.1%	797	64.9%
Race Not Identified			7	100.0%	139	36.7%	221	91.4%	93	62.4%	69	79.7%	529	70.5%
White			70	97.1%	1364	24.8%	2,346	94.5%	1,147	67.1%	711	84.7%	5,638	70.9%
Gender														
Female			24	91.7%	732	14.2%	969	92.4%	401	54.4%	190	70.0%	2,316	59.3%
Male			80	92.5%	2,014	17.3%	3,174	93.1%	1,434	58.9%	893	76.6%	7,595	64.6%
Gender Not Identified			14	92.9%	409	19.8%	561	88.6%	258	63.2%	98	75.5%	1,340	61.8%
Total			1,18		3,155		4,704		2,093		1,181			

Table 10D: Proportion of Applicants Passing the Qualification Stage as a Function of Application Source by Year

Ethnicity	Source: Public													
	2006		2007		2008		2009		2010		2011		Weighted Average FY2007–2011++	
	F_i+	%	F_i	%	F_i	%	F_i	%	F_i	%	F_i	%	Total	%
Asian			297	75.8%	481	87.9%	358	86.3%					1136	84.2%
African-American			4,993	68.3%	5,055	78.7%	2,385	79.6%					12,433	74.7%
Hispanic			754	69.5%	992	82.3%	644	83.2%					2,390	78.5%
Multi			856	73.5%	1,238	85.7%	826	88.1%					2,920	82.8%
Race Not Identified			623	75.1%	733	88.9%	534	90.8%					1,890	84.9%
White			5,469	77.5%	8,242	89.9%	6,689	89.8%					20,400	86.5%
Gender														
Female			3,454	74.2%	4,043	85.8%	2,349	86.9%					9,846	82.0%
Male			8,157	72.6%	11,037	85.4%	7,957	87.1%					27,151	82.1%
Gender Not Identified			1,539	72.4%	1,859	85.5%	1,273	86.9%					4,671	81.6%
Total			13,150		16,939		11,579							

Table 10E: Proportion of Applicants Passing the Qualification Stage as a Function of Application Source by Year

Ethnicity	Source: Reinstatement-DOD CPC													
	2006		2007		2008		2009		2010		2011		Weighted Average FY2007–2011++	
	F_i+	%	F_i	%	F_i	%	F_i	%	F_i	%	F_i	%	Total	%
Asian			0	0.0%	18	55.6%	13	23.1%	22	27.3%	9	22.2%	62	33.9%
African-American			5	40.0%	211	9.0%	277	4.0%	311	11.6%	40	25.0%	844	9.3%
Hispanic			2	100.0%	48	16.7%	57	24.6%	60	30.0%	16	18.8%	183	24.6%
Multi			1	0.0%	30	16.7%	77	20.8%	74	32.4%	28	53.6%	210	28.6%
Race Not Identified			5	100.0%	61	59.0%	52	34.6%	78	47.4%	27	70.4%	223	51.6%
White			25	76.0%	394	57.9%	438	41.1%	511	53.8%	266	50.4%	1,634	51.2%
Gender														
Female			6	50.0%	167	18.6%	191	18.3%	221	28.5%	67	50.7%	652	25.5%
Male			24	70.8%	490	44.9%	618	28.0%	700	40.3%	282	43.3%	2,114	38.5%
Gender Not Identified			8	100.0%	115	53.0%	117	32.5%	152	40.1%	45	68.9%	437	45.5%
Total			38		772		926		1,073		394			

Table 10F: Proportion of Applicants Passing the Qualification Stage as a Function of Application Source by Year

Ethnicity	Source: College Training Initiative (CTI)													
	2006		2007		2008		2009		2010		2011		Weighted Average FY2007–2011++	
	F_i+	%	F_i	%	F_i	%	F_i	%	F_i	%	F_i	%	Total	%
Asian	a	a	a	a	25	100.0%	20	100.0%	42	100.0%	67	98.5%	154	99.3%
African-American	a	a	a	a	34	100.0%	33	100.0%	51	100.0%	61	98.4%	179	99.5%
Hispanic	a	a	a	a	28	100.0%	41	97.6%	62	100.0%	88	97.7%	219	98.6%
Multi	a	a	a	a	28	92.9%	33	100.0%	65	100.0%	105	100.0%	231	99.1%
Race Not Identified	a	a	a	a	31	100.0%	26	100.0%	66	98.5%	68	100.0%	191	99.5%
White	a	a	a	a	349	99.7%	398	99.0%	702	99.7%	924	99.8%	2,373	99.6%
Gender														
Female	a	a	a	a	98	100.0%	93	97.8%	160	100.0%	226	99.1%	577	99.3%
Male	a	a	a	a	335	99.1%	412	99.3%	734	99.7%	1,000	99.6%	2,481	99.5%
Gender Not Identified	a	a	a	a	64	100.0%	54	100.0%	100	99.0%	96	100.0%	314	99.7%
Total					497		559		994		1,322			

APPENDIX B

FAA JOB ANNOUNCEMENT SUMMARY 2006 – 2011

WHO MAY BE CONSIDERED:	COUNT
Individuals with air traffic control experience who are eligible for appointment under the Veteran's Readjustment Appointment (VRA) authority and who have been discharged from active duty or who are on terminal leave.	**231 (57.6%)**
Military air traffic controllers who are either on terminal leave pending retirement from active duty military service or who retired from active duty on or after September 17, 1999, may apply.	**40 (10%)**
U. S. Citizens	**88 (21.9%)**
This announcement is open to reinstatement eligible and employees of other federal agencies that are eligible for transfer. This vacancy announcement is also open to individuals who are eligible for Selection Priority Program (SPP) consideration. All SPP eligible must submit an application and a copy of their RIF Notice with this job announcement referenced to receive consideration under SPP. We are not accepting applications from noncitizens.	**2 (0.4%)**
All sources may apply under this announcement. Former employees eligible for SPP consideration must submit an application, a copy of their RIF Notice, and any other information specified in this announcement to receive consideration as an SPP candidate.	**1 (0.02%)**
Current or former employees – qualified civil service employees – This announcement is open to non-FAA employees who are eligible for transfer or reinstatement and also open to former FAA employees eligible for SPP consideration. All SPP eligible must submit a copy of their RIF Notice and any other information specified in this announcement to receive consideration under SPP.	**13 (0.2%)**
Students of Silver State CTO programs only	**5 (0.1)%**
AT-CTI graduates	**9 (0.2%)**

All sources – applicants must possess the following selective placement factor to qualify: A control tower operators (CTO) license with a facility rating of Tower Cab.	**4 (0.1%)**
Current or Former Federal Employees & EVHO – this announcement is open to former FAA CPC-level air traffic controllers or those who qualify under the Expanded Veterans Hiring Opportunity (EVHO).	**2 (0.01%)**
Current or former federal employees & EVHO – This announcement is open to civilian Department of Defense (non-FAA employees) who are eligible for transfer or reinstatement or those who qualify under the Expanded Veterans Hiring Opportunity (EVHO)	**5 (0.1%)**
TOTAL	**401**