NPS-HR-23-250

ACQUISITION RESEARCH PROGRAM SPONSORED REPORT SERIES

Clearing for Takeoff: A Comparative Analysis of the Latest Aviation Service Test Battery

March 2023

LCDR Eric M. Campbell, USN

Thesis Advisors: Dr. Maxim Massenkoff, Assistant Professor
Dr. Sae Young (Tom) Ahn, Assistant Professor

Department of Defense Management

Naval Postgraduate School

Approved for public release; distribution is unlimited.

Prepared for the Naval Postgraduate School, Monterey, CA 93943.

The research presented in this report was supported by the Acquisition Research Program of the Department of Defense Management at the Naval Postgraduate School.

To request defense acquisition research, to become a research sponsor, or to print additional copies of reports, please contact the Acquisition Research Program (ARP) via email, arp@nps.edu or at 831-656-3793.

ABSTRACT

Training Naval and Marine Corp Aviators is an expensive, time-consuming process, and both branches use the Aviation Service Test Battery (ASTB) to screen applicants before flight training. The ASTB helps ensure aviation applicants are prepared and capable of completing the rigorous aviation training pipeline to ensure resources are well spent. This test has undergone several revisions since its introduction in World War II and most recently in 2012 by substituting and revising portions of the test. Among the updates is a complex testing section using a joystick, throttle, and headphones meant to resemble actual flying in the Primary Flight Training closely. This study examines how effectively the ASTB-E reduces attrition within the aviation community and how each of the new sections contributes to assessing applicants.

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

My most sincere thank you goes out to so many. I was not alone in this.

First, to my classmates and, most especially, the HR officers who joined on this academic adventure: Thank you for your continuous support and encouragement.

Thank you to my advisors, Dr. Maxim Massenkoff and Dr. Tom Ahn. You have helped me solidify my understanding of the curriculum, apply concepts to this thesis and, more importantly, develop academically and professionally beyond my perceptions. I could ask for nothing more from a professor and leader.

To my beautiful daughter and adventure buddy, Anna: Your strength through our moves and unconditional kindness to others continue to inspire me. Thank you, Tiny.

For my wonderful wife, Jen, I offer my gratitude. Where would I be without you? Through highs and lows and any number of trials from Florida to California and from high school to now, you are my constant.

THIS PAGE INTENTIONALLY LEFT BLANK

NPS-HR-23-250

ACQUISITION RESEARCH PROGRAM SPONSORED REPORT SERIES

Clearing for Takeoff: A Comparative Analysis of the Latest Aviation Service Test Battery

March 2023

LCDR Eric M. Campbell, USN

Thesis Advisors: Dr. Maxim Massenkoff, Assistant Professor
Dr. Sae Young (Tom) Ahn, Assistant Professor

Department of Defense Management

Naval Postgraduate School

Prepared for the Naval Postgraduate School, Monterey, CA 93943.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

ANIT	Aviation/Nautical Information Test
API	Aviation Preflight Indoctrination
AQR	Academic Qualification Rating
ASTB	Aviation Selection Test Battery
ATT	Airplane Tracking Task
CUI	Controlled Unclassified Information
DLT	Dichotic Listening Task
DOR	Drop on Request
DOT	Direction Orientation Task
FOFAR	Flight Officer Aptitude Rating
MCT	Mechanical Comprehension Test
MST	Math Skills Test
NATFI	Naval Aviation Trait Facet Inventory
NROTC	Naval Reserve Officer Training Corps
OCS	Officer Candidate School
PBM	Performance-Based Measurements
PFAR	Pilot Flight Aptitude Rating
RCT	Reading Comprehension Test
SAT	Scholastic Aptitude Test
SNA	Student Naval Aviator
VTT	Vertical Tracking Task

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

Flight school is expensive and high-risk, and aspiring military aviators usually apply with little to no flight experience. These are some of the reasons the Navy and Marine Corps use a test, the Aviation Service Test Battery (ASTB), to assess a candidate's ability before flight training. Without such a test, attrition has been as high as 60% (North & Griffin, 1977). This test has been in use since World War II, helping the Department of Defense (DOD) select officers with the potential to succeed through rigorous flight training. The newest version of this test, implemented in December 2013, is called the ASTB-E. The previous version used five subtests to assess a candidate's cognitive ability and knowledge using math, reading, mechanical comprehension, aviation and nautical information test, and spatial awareness assessments. The ASTB-E includes four previously used subtests and adds five new ones, which involve using a joystick and throttle to mimic actual piloting challenges closely. These additions were meant to predict better the chance that a candidate will succeed in Primary Flight Training.

The ASTB-E is one of the first steps becoming a Navy or Marine Corps aviator and officer, in general. Most Navy and Marine Corps officers take this test as it also includes a section, the Officer Aptitude Rating (OAR), which includes reading, math and mechanical comprehension subtests, that assess their potential as an officer. The OAR is similar to how the Armed Service Vocational Aptitude test (ASVAB) assesses potential enlisted personnel. In addition to the OAR section, aviation candidates also complete three separate sections designed specifically for aviators and flight officers. These aviation sections include the Aviation and Nautical Information Test (ANIT), Performance-Based Measurements (PBM) and a personality assessment. This test takes 2–3 hours, with 1–2 hours devoted to the officer assessment.

Upon completion, applicants receive different scores based on the program they apply to. All earn an OAR, a score between 20 and 80 that grades their potential as an officer in Naval Officer Candidate School (OCS). This score is based on the first few subtests, including math, reading comprehension, and mechanical comprehension. Aviation applicants also receive an Academic Qualification Rating (AQR) score between

1 and 9 (Navy Medicine Operational Training Command, n.d.-b). The AQR grades their potential to succeed later in flight school's academic portions, including Aviation Preflight Indoctrination (API) and ground school events in Primary Flight Training. Last, aviation applicants receive a Flight Officer Aptitude Rating (FOFAR) and Pilot Flight Aptitude Rating (PFAR), which predicts Naval Flight Officers' and Naval Aviators' performance, respectively, in Primary Flight Training. (Navy Medicine Operational Training Command, n.d.-b) The scores required for different programs are shown in Figure 1.

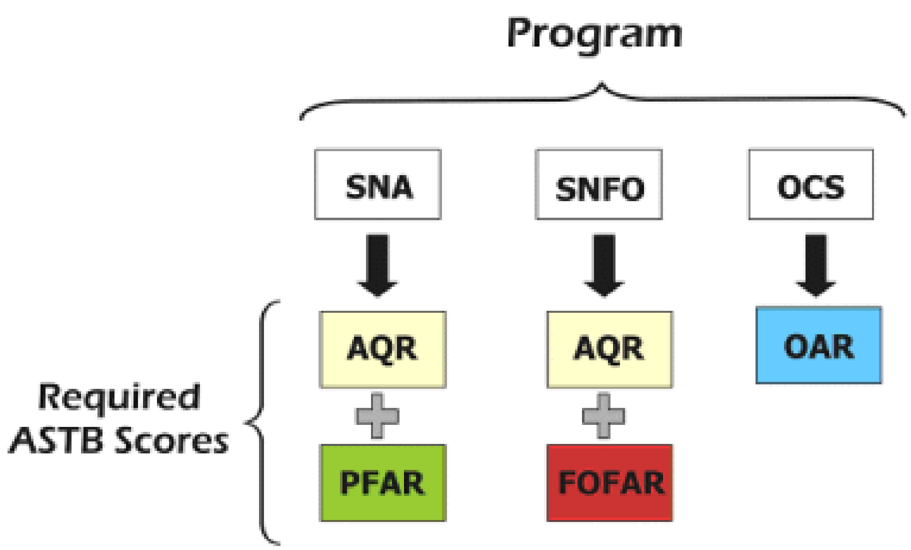


Figure 1. Score Components Used in the Selection Process. Source: Navy Medicine Operational Training Command (n.d.-b).

The price tag of $180,462 per student in Primary Flight Training (D. Reed, personal communication, February 17, 2023) necessitates an accurate and comprehensive assessment of as many capabilities and aptitudes related to aviation before investing heavily in the individual. This is where the ASTB-E comes in. In addition to revamping the spatial ability test, the test now includes several PBM subtests and the Naval Aviation Trait Facet Inventory (NATFI), a personality assessment. The PBM subtests include four tasks incorporating a joystick, throttle, and headphones. These tasks grow increasingly complex starting with individual tasks with the joystick and then the throttle, a dichotic listening test, and culminates with all previous task performed at once (Navy Medicine

Operational Training Command, n.d.-a). This allows an assessment beyond the previous cognitive and knowledge tests to incorporate psychomotor skills (Navy Medicine Operational Training Command, n.d.-a). Moreover, the NATFI assesses the applicant along nine personality traits that are shown pertinent to success in naval aviation (Navy Medicine Operation Training Command, n.d.-a). In total, the ASTB-E now “assesses the examinee’s ability to think in three dimensions, physical dexterity, eye-hand coordination, and ability to divide attention among different tasks” (Navy Medicine Operational Training Command, n.d.-a).

With these new advanced subtests come high expectations on predicting performance and reducing attrition. “The marked increase in incremental validity that results from the addition of PBM composites to the ASTB suggests that the addition of the PBM to Naval Aviation selection will significantly reduce attrition from the Naval Aviation training pipeline…” (Phillips et al., n.d., p. 141). Additionally, according to one estimate by Keiser et al. (2019), the ASTB-E will be 10% more accurate than its predecessor at forecasting the results of flight training. Last, an ASTB-E cost assessment projected the Navy and Marine Corps saving over $52 million a year in attrition costs (Navy Medicine Operational Training Command, n.d.-a).

The ASTB-E demonstrates the latest advancements in predicting flight performance and has been in service for nearly ten years. This research paper studies the impact of the ASTB-E.

A. PRIMARY RESEARCH QUESTIONS

(1) Was the ASTB-E successful at reducing attrition in Primary Flight Training?

By expanding the breadth of evaluation by incorporating PBM subtests and personality assessment, we expect a more thorough evaluation of a candidate’s ability to pilot aircraft. This means the ASTB-E Student Naval Aviators (SNAs)--i.e., flight students who took the new test-- should outperform those under the previous test, leading to lower attrition in Primary Flight Training.

(2) Was the ASTB-E better able to predict performance of future SNAs in Primary Flight Training?

Like research question one, the ASTB-E is expected to predict future SNAs' performance better. Therefore, ASTB-E SNAs should have, on average, a higher Navy Standard Score (NSS), a score used to grade flight performance, than the ASTB SNAs.

B. SECONDARY RESEARCH QUESTIONS

(1) To what extent do the new PBM subtests contribute to the predictiveness of the ASTB-E?

The previous subtests already provide a partial understanding of a candidate's ability but are limited to cognitive and knowledge testing. We expect a more complete picture of their capability with the new subtests.

(2) Is the Direction Orientation Taks (DOT) subtest continuing to contribute predictiveness to the ASTB-E between 2015 and 2018?

Coyne et al. (2022) researched the DOT subtest and concluded that "incremental validity is no longer present in applicants after 2015." The DOT in previous years was contributing to validity, but after 2015, it stopped. We use our data to recreate this test and answer the same question.

(3) Do the new PBM subtests create any adverse impacts among varying demographics?

One of the other goals of the ASTB-E was to reduce any adverse impact when testing creating a more diverse group of SNAs. The new PBM subtests are very different from the legacy subtest and may have varying impact on demographic subgroups. We will test for any impact present in the data.

(4) Does the PFAR remain valid with increasing time between testing and commencing flight training?

PFAR predicts performance in Primary Flight Training, so a high PFAR score should correlate to a high score in Primary. Sometimes, however, years can pass between when a candidate tests and when they begin training. The current policy is that ASTB scores do not expire. We test if there is any significant decrease in PFAR validity over time.

THIS PAGE INTENTIONALLY LEFT BLANK

II. BACKGROUND

A. INTRODUCTION

This section will describe the process of taking the ASTB-E, continuing into Primary Flight Training, and possibly attriting from training. This includes a discussion of the nature of attrition in DOD, the process of taking the ASTB-E, what each ASTB-E subtest entails, and the initial phases of flight school. Continued training beyond Primary is not within the scope of this paper.

B. ATTRITION

Attrition is a challenge for all jobs in every branch of the military. Attrition is when personnel depart from a certain job or organization for any reason, whether voluntary or involuntary. In addition to other metrics, the DOD also has a test for enlisted personnel, the ASVAB, to help ensure enlistees have the right requisite knowledge and skills to perform their function. Without it, some personnel would not have the ability or motivation to stay in a role and would leave prematurely. Attrition is especially challenging for aviation designators because of the specific skills required and demand of the job.

One reason for attrition is called Drop on Request (DOR). The role of SNA is a strictly voluntary one. So, at any point in an aviator's career, he or she may drop from the flight program even as far out as after they have completed flight training and are at their operational squadron. A DOR can be for any reason, but given the nature of the profession, anxiousness, or anxiety about flying performance aircraft is assumed to be associated (Arnold & Phillips, 2008, p. 9).

Another reason is the medical requirements for aviation. The Federal Aviation Administration (FAA) and Naval Aerospace Medical Institution maintain high health standards for individuals to fly (North & Griffin, 1977). If at any point, an SNA is deemed not medically qualified, they could be dropped from the flight program. Although most cases are caught with an initial thorough medical screening, attrition can occur for this reason.

A similar medical reason for attrition is the body's ability to cope with the stress of flying a performance aircraft. Those selected for the tailhook or jet community are taught to cope with high g-forces in a centrifuge, but even before this, the stress can be too much for some individuals. The force of the aircraft can also cause personnel to quit because of motion sickness.

Last, a separate stress in flight training is to perform. Sometimes, combining the knowledge and skills to pilot performance aircraft in a high-stress environment can be too much, and grades drop leading to possible attrition. Several measures are in place to help avoid this as the investment has already been made in the individual. However, this is still a common reason for attrition within flight training.

C. ASTB HISTORY

Before World War II, the military and civilian industry recognized the need to evaluate pilots physically and mentally. This need was primarily driven by monetary costs and the high failure rate (North & Griffin, 1977, p. 1). One of the first aviation test validation studies took place in 1919, called the Kelley Field Study (North & Griffin, 1977, p. 5). Additionally, during this time, many psychologists felt that psychomotor testing, like the one currently used in the ASTB-E, would lead to more predictive results (North & Griffin, 1977, p. 4). This requirement became even more critical during World War II with the increased demand for aviators.

The earliest version of the ASTB was developed in 1941 by the Pensacola 1000 Aviator Study (North & Griffin, 1977, p. 5; Natali, 2018). The original test was called the Naval Aviation Questionnaire (NAQ) and consisted of three subtests: the Wonderlic Personnel Test, which is a measure of mental ability, the Bennett Mechanical Comprehension Test (BMCT), and a biographical inventory (BI) (Natali, 2018). The project tested successfully at reducing attrition and was recommended for operational use in November 1941. The Navy then added cutoff scores and combined two subtests, the BMCT and BI, to create the Flight Aptitude Rating (FAR).

The NAQ became the ASTB in the 1950s and continued to see several revisions and updates in the coming decades. In 1953, the first spatial apperception test was added including the Aviation Qualification Test (AQT). Research continued through the 1960s on the effects of stress on aviator performance and psychomotor tests to increase incremental validity with the ASTB. In the 1980s, psychomotor testing continued with measures to improve diversity (Natali, 2018). In 1992, scores were separated between SNAs and Student Naval Flight Officers (SNFOs) with the PFAR and FOFAR.

In the 2000s, two changes occurred to the ASTB before the introduction of the ASTB-E. For the first time, the ASTB was administered using the Automatic Pilot Examination (APEX), a web-based platform providing a central testing source. Also, the biographical inventory was removed from the ASTB. Then, in December 2013, the Navy and Marine Corp introduced the ASTB-E.

D. ASTB PROCESS AND SCORING

The ASTB-E can be taken nationwide at several locations and at select locations outside the United States, but the number of attempts limits applicants. The test can be administered at Navy and Marine Corp Officer recruiting stations and Naval Reserve Officer Training Corps (NROTC) units at many major universities such as Old Dominion University (ODU) or San Diego State University (SDSU), which have large NROTC commands (Navy Medicine Operational Training Command, n.d.-a). Additionally, the test can be taken at military institutes such as the Naval Academy. Applicants may only take this test three times and the results are good for life. However, only the most recent score counts, not the highest score. So, if an applicant retests and scores lower, the low score is used. Those requesting to test again must also wait 30 calendar days after completing the ASTB-E and complete the entire test again.

Four scores are earned after taking the whole ASTB-E: the OAR, AQR, PFAR, and AQR. As stated above, many non-aviation officers take this test only for the OAR score, which predicts how well an officer will perform in OCS and ranges from 20 to 80. The remaining scores are based on a stanine scale shown in Figure 2. On the stanine scale, examinees are scored against previous test takers and given a score of 1 through 9. For example, an examinee with a score of 7 is in the 77–88th percentile of previous testers.

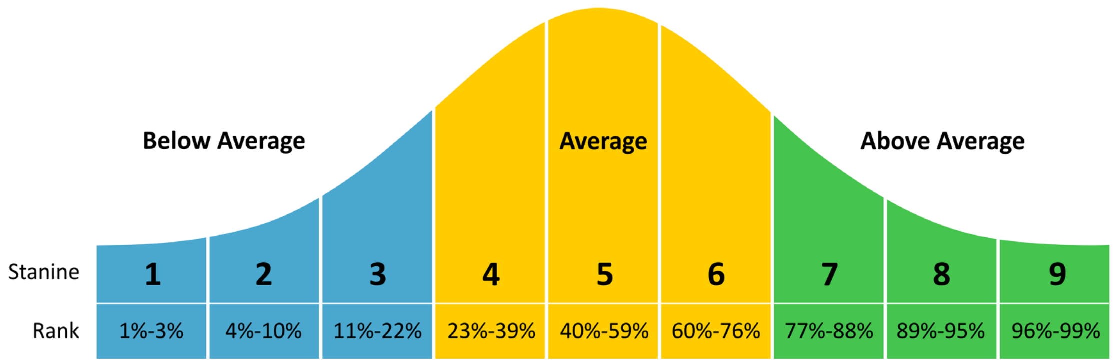


Figure 2. Stanine Scale. Adapted from Mometrix (2018).

Aviation candidates require two scores, the AQR and PFAR, and must make the minimum to be considered as a pilot. The AQR and PFAR are scores used to predict how well an applicant will perform in API and Primary ground school events, and Primary flight performance, respectively (Navy Medicine Operational Training Command, n.d.-b). So, a high AQR of 8 means the applicant will likely perform better in API than applicants with lower scores. There is the same correlation with a high PFAR score. The minimum score to apply as an aviator is a 4 in AQR and a 5 in PFAR. Performance in API and Primary Flight School is judged using the Navy Standard Score (NSS), which is discussed in detail later.

E. ASTB-E SUBTEST DESCRIPTIONS

The ASTB-E incorporates seven subtests, four from the original version and three new sections, as shown in Figure 3. The PBM section shown in Figure 3 is broken into five separate subtests. The new subtests aim to assess additional aptitudes and personal characteristics over the previous version to better predict training outcomes and NSS scores. Each of the subtests evaluates different abilities.

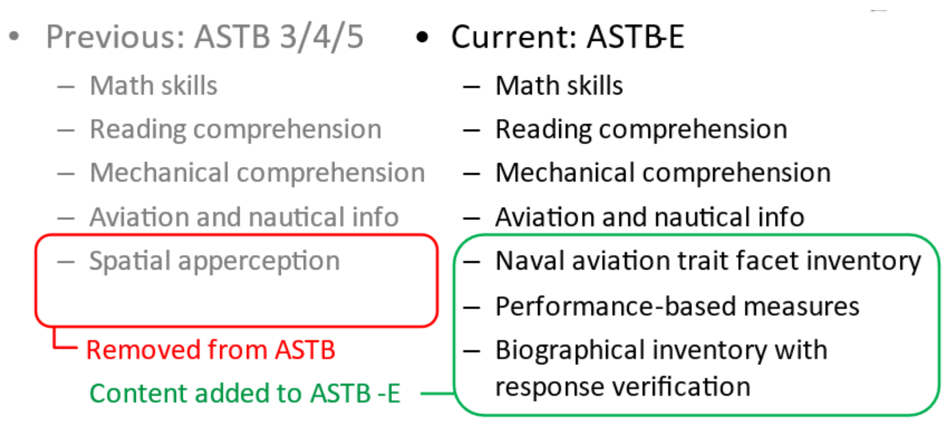


Figure 3. ASTB-E Subtests. Source: Natali (2018).

The first four subtests are multiple choice and test applicants' math, reading, and mechanical comprehension and basic aviation and nautical knowledge. The Math Skills Test (MST) is 30 questions with a time limit of 30 minutes. It consists of arithmetic, algebra, and certain geometry problems to include word problems. Altogether, basic mathematical operations, problem-solving with variables, fractions, roots, and exponents, and the computation of angles, areas, and geometric form perimeters are tested. Figure 4 contains a sample question.

3. *Easy*

If the hypotenuse of a right triangle is 5 inches long and one of the sides is 4 inches long, how long is the remaining side?

(A) 2 inches
(B) 3 inches
(C) 4 inches
(D) 9 inches

Figure 4. Math Skills Test (MST) Sample Question. Source: Navy Medicine Operational Training Command (2014).

In the Reading Comprehension Test (RCT), candidates are asked to interpret the meaning of several passages. This portion is 20 questions with a time limit of 30 minutes. Candidates are graded on their ability to develop sound conclusions with each written section. Although some answers appear factually true, candidates need to be aware that only a single response can be obtained based on the information given (Navy Medicine Operational Training Command, n.d.-b). Figure 5 contains a sample question.

4. *Difficult*

Military units shall assist civilian communities in providing medical emergency helicopter services beyond the capability of that community. Military units shall not compete for emergency medical evacuation missions in areas where support can be provided by civilian contractors.

(A) Medical emergency helicopter services are provided through Military units during an emergency in which civilian contractors can provide support.

(B) Civilian contractors can provide medical emergency helicopter services to aid military units that assist civilian communities during emergencies.

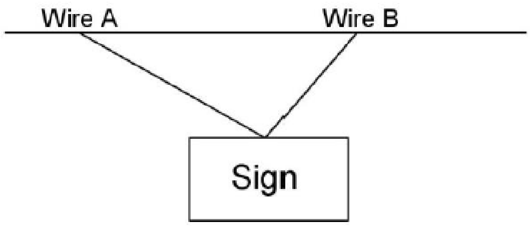
(C) Military units shall assist civilian contractors in all medical evacuation missions by providing medical emergency helicopter services.

(D) A military unit may assist civilian communities in providing medical emergency helicopter services if civilian contractors cannot provide support.

Figure 5. Reading Comprehension Test (RCT) Sample Question. Source: Navy Medicine Operational Training Command (2014).

The Mechanical Comprehension Test (MCT) is 30 questions with a 15-minute time limit. This subtest assesses the candidate's ability to see and comprehend the nature of physical relationships to solve real-world challenges with mechanical concepts (Navy Medicine Operational Training Command, n.d.-b). Topics within this section most closely relate to those found in early high school physics courses and their function in multiple scenarios. Some examples include the properties of pressure, volume and velocity, engine performance, principles of gears and weight distribution, and simple devices like pulleys and fulcrums (Navy Medicine Operational Training Command, n.d.-b). Figure 6 contains a sample question.

12. *Intermediate*



A sign hangs from a ceiling by two pieces of wire. Which of the two wires bears the greater tension?

(A) Wire A bears the greater tension
(B) Wire B bears the greater tension
(C) Both wires experience equal tension

Figure 6. Mechanical Comprehension Test (MCT) Sample Question. Source: Navy Medicine Operational Training Command (2014).

The Aviation and Nautical Information Test (ANIT) is one of the more straightforward subtests as it is primarily knowledge-based as shown in Figure 7. This section tests the candidate’s “familiarity with aviation history, nautical terminology and procedures, and aviation-related concepts such as aircraft components, aerodynamic principles, and flight rules and regulations” (Navy Medicine Operational Training Command, n.d.-b). Since one’s score is based on prior studying, this section’s score is easily improved. Study material can include several items including material related to flight, navigation, and FAA documents (Navy Medicine Operational Training Command, n.d.-b). There are also numerous study guides available.

4. *Intermediate*

Pressure altitude is an altitude measured from the standard sea level pressure of ________.

(A) 29.89
(B) 29.92
(C) 29.97
(D) 29.99

Figure 7. Aviation and Nautical Information Test (ANIT) Sample Question. Source: Navy Medicine Operational Training Command (2014).

(1) PBM Subtests

The other five subtests are the new PBM assessments. Aside from the DOT subtest, they all get progressively difficult as an applicant works through several tasks with the joystick and throttle. This begins with testing the applicant on the joystick and throttle separately, continues with dichotic listening, and ends with combining all the previous tasks.

The first PBM subtest is the Directional Orientation Task (DOT), designed to assess the applicant's ability to quickly judge an object's physical orientation. In this scenario, applicants are shown two pictures: a tracker map displaying the heading of an unmanned aerial vehicle (UAV), and the camera view from the UAV shown in Figure 8. Then, they are asked to identify one of four parking lots. While some individuals may look at Figure 8 and choose the parking lot on the right in line with the cardinal direction, that is not the view from the UAV. Examinees must quickly orient themselves with the camera view and then pick the correct parking lot. The ability of pilots to orient themselves within a given space is a critical skill, and this tests that ability (Coyne et al., 2022).

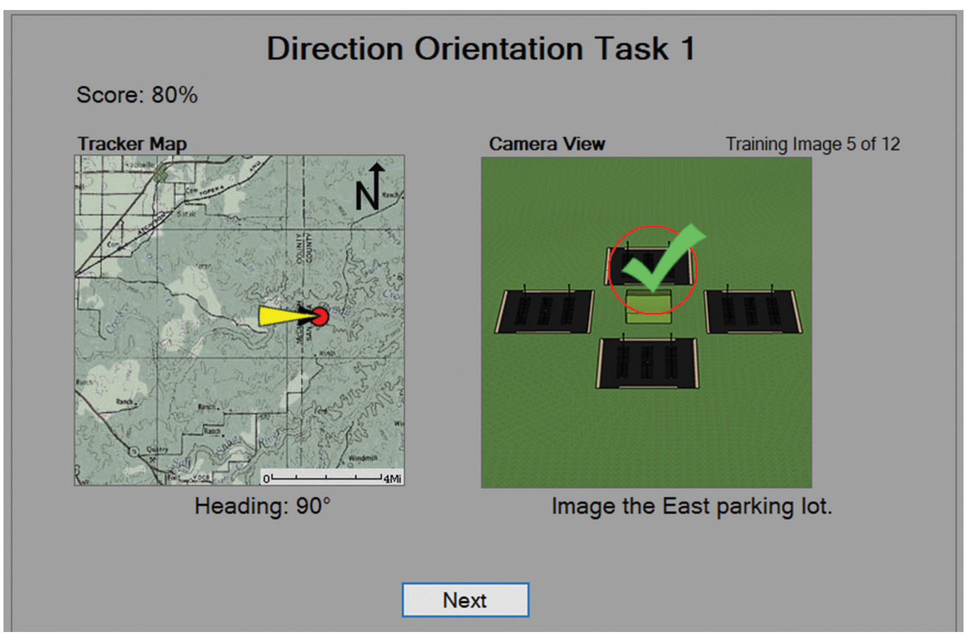


Figure 8. DOT Subtest Sample. Source: Natali (2018).

Examinees then use the joystick in the Airplane Tracking Task (ATT). In this portion, they are graded on their ability to track a moving object both vertically and horizontally. Specifically, a small yellow airplane image appears on the screen, which the examinee must follow closely with the red crosshairs using the joystick, shown in Figure 9. The airplane will continuously move randomly across the screen. The score is based on the average distance the crosshairs are from the airplane and how many "hits" occur (Natali, 2018).

Figure 9. Airplane Tracking Task (ATT) (Airplane and Crosshairs on the Right). Source: Natali (2018).

Following the joystick use, examinees perform a similar challenge but with the throttle in the Vertical Tracking Task (VTT). This portion tests the examinee's ability to track an object in one dimension vertically, shown in Figure 10. Again, the score is based on the average distance from the airplane and the number of "hits" (Natali, 2018).

Figure 10. Vertical Tracking Task (VTT) (Airplane and Crosshairs on the Left). Source: Natali (2018).

At this point in the exam, examinees put on headphones and are asked to respond to specific audio cues. A sequence of numbers and letters is played in both ears, as shown in Figure 11. The examinee must identify particular characters that are heard in just one ear and then respond with the appropriate control input. This task is done independently and while also performing the ATT and VTT simultaneously.

Figure 11. Dichotic Listening Task (DLT) Sample. Source: Natali (2018).

The last PBM subtest is the Emergency Scenario, which combines all three previous elements. Examinees must simultaneously perform the ATT and VTT while responding to specific simulated aviation emergencies, as shown in Figure 12. Upon the onset of the emergency, the examinee must make the correct control inputs similar to what they accomplished in the DLT. Examiners record the score as the "Skill Factor."

Figure 12. Emergency Scenarios Task Sample. Source: Natali (2018).

The final subtest after the PBM portion is the Naval Aviation Trait Facet Inventory (NATFI). During this subtest, candidates respond to several paired statements that examiners use to assess nine personality traits related to the successful completion of flight training, shown in Figure 13 (Natali, 2018). The specific traits considered are proprietary information maintained by Naval Medical Operational Training Command (NMOTC) and are not released. Each set of statements is designed to be difficult to choose between (Navy Medicine Operational Training Command, n.d.-b). Selecting one does not necessarily mean the candidate is willing to engage in the stated behavior, only that they prefer that option over the other (Navy Medicine Operational Training Command, n.d.-b).

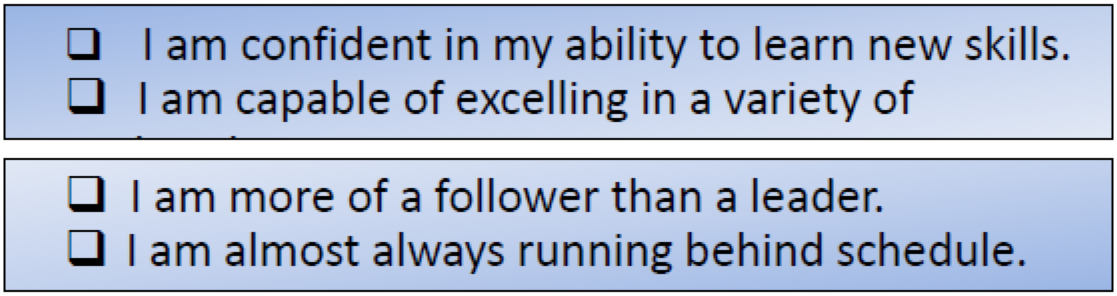


Figure 13. Naval Personality Trait Facet Inventory (NATFI) Sample Question. Source: Navy Medicine Operational Training Command (2014).

F. AVIATION PREFLIGHT INDOCTRINATION

Before Primary Flight Training, SNAs must complete an intense academic regiment called Aviation Preflight Indoctrination. This is essentially the beginning of flight school, with only a brief flying experience before this called Introductory Flight Screening (IFS). This school challenges students mentally and physically. The curriculum includes aerodynamics, navigation, aircraft engines, meteorology, and basic flight regulations. Each subject lasts about one week with a comprehensive test weekly. While completing the academic portion, SNAs are trained in water survival techniques, culminating in a 1-mile swim test while wearing a flight suit. It is during this portion that the first significant portion of attrition occurs.

There are two primary ways an SNA fails out of this training: failing academically or the water survival training. The option to DOR is also available. A test score under 80% triggers academic counseling, and the third failure results in dismissal from the flight program. Additionally, water survival is challenging; some may fail during training or the final swim test.

G. PRIMARY FLIGHT TRAINING

After completing API, SNAs transfer to either Whiting Field, FL, or Corpus Christi TX for Primary Flight Training. This portion is specifically called Joint Primary Pilot Training (JPPT) and is a phase of training that equips the SNA with the basic airmanship skills associated with military aircraft and establishes a firm foundation. This is also the first opportunity for SNAs to fly a high-performance aircraft, the T-6B Texan II, shown in Figure 14. JPPT is divided into several training blocks and has standardized grading.

Figure 14. T-6B Texan II. Adapted from CNATRA (2021).

SNAs complete the JPPT phase through different blocks of training designed with varying learning objectives. There are four main blocks: Contact or Visual Flight Rules (VFR) training, Instrument or Instrument Flight Rules (IFR) training, Navigation, and Formation (Chief of Naval Aviation Training [CNATRA], 2021). There is also a portion in the Contact block for aerobatic flight. For most blocks, the training is a sequence of ground school classes, simulator events, and flight events. Each event has a minimum passing score called a Maneuver Item File (MIF) on a scale from 1 to 5. The minimum begins lower and gradually builds as training progresses. For example, the MIF for a G-maneuver, where the SNAs turn the aircraft aggressively in a circle to generate G-force, is 3 in the early phases of Contact, but will increase to 4 later.

After completing the JPPT phase, the SNAs score is "the sum of the [SNA's] grades for maneuvers performed (item grades) divided by the sum of the MIF for those maneuvers" (CNATRA, 2022, p. G-1). This score is then standardized to correct any possible non-normality in the distribution of phase scores (CNATRA, 2022, p. G-4). The group size used for the standardization is 200. SNAs that attrite from the program are excluded from this calculation. After this, the SNA has a final NSS for continued training.

H. CONCLUSION

The path from testing to Primary is challenging. The ASTB-E is a critical portion of that path to help assess candidates before investing. Flying military aircraft is an extraordinarily complex and demanding undertaking that requires a high level of cognitive and physical abilities. Military aviation is also inherently dangerous, so it's essential to ensure that only those with the right skills and abilities are selected for flight training. The ASTB-E contains several new testing methods to help better evaluate a candidate's potential, which should yield lower attrition over the legacy test.

III. LITERATURE REVIEW

A. INTRODUCTION

There is limited research to address both primary research questions. However, the available data is still very useful in understanding the effect the ASTB-E may have on attrition and its impact on racial subgroups. Specifically, there is light research on the introduction of computer testing on the previous version of the ASTB, which carries over to the ASTB-E, and on the differences in flight and ASTB performance between races and gender. Additionally, there is also work on the Navy's implementation of PBM, which make up the bulk of novel ASTB subtests, and a problem with one subtest, the Directional Orientation Task (DOT), which has decreasing validity in predicting flight training outcomes and may be negatively affecting attrition.

B. TREND FROM DIGITAL TRANSITION

The Aviation Service Test Battery has been regularly updated since its first introduction in World War II. Before the most recent update in 2013, researchers first explored and updated the ASTB to include computer testing. The advent of low-cost microchips allowed this method to become the new affordable testing method over paper and pencil (Kennedy et al., 1998). While this study focused on the introduction of computer use, it did not draw any conclusions on the effect of predicting flight training performance. However, it did show some variation when subjects conducted the test on different computers. This variation is later shown with the ASTB-E in that screen size and joystick model affected scores. However, neither our data provider, the source of this information, nor current research specifies how scores are affected.

C. CURRENT RESEARCH

More recent research by the Naval Medical Research Unit has been centered around testing and implementing PBM to the ASTB mimicking the addition of the Test of Basic Aviation Skills (TBAS) to the Air Force Officer Qualifying Test (AFOQT). The ASTB-E seeks to broaden the Navy's aviation selection capabilities beyond the knowledge, skills, and abilities drawn from the current selection batteries (Phillips et al., 2011). The TBAS includes several new assessment measures, several of which are now part of the ASTB-E, including: Airplane Tracking Test (ATT), Vertical Tracking Test (VTT), a version of the Unmanned Aerial Vehicle (UAV) test called Directional Orientation Test (DOT), a version of the Three Digit Listening Test called the Dichotic Listening Test (DLT), and the Emergency Scenario Test (EST) which combines several of the tests above (Phillips et al., 2011). Across the tests, the tasks are arranged in order of increasing cognitive load stress leading up to the final subtest, the EST. In addition to base test scores, they are also graded on the speed of their response in many subtests.

Some studies have analyzed differences in performance by gender and race using older versions of the ASTB and in flight school. There are no significant differences in attrition between men and women, but there can be with racial groups (Gibson & Gibson, 2005). One study examined the impact of raising the minimum acceptable score for minorities on the ASTB. Currently, the Navy only accepts candidates who score a minimum of 5 out of a possible nine on the Pilot Flight Aptitude Rating (PFAR), but the Marine Corps raises their minimum to 6. While raising the minimum score did increase performance results for a randomly selected group, it did so at the cost of decreasing minority representation (Dean, 1996). Furthermore, the randomly non-selected group performed at a lower level while also experiencing higher attrition (Dean, 1996). Although this study is older, it provides insight into the historical research on previous ASTB versions concerning race.

Another study looked at the role of gender and race in completing Naval Flight Training. While there was no significant difference in attrition between men and women, X groups had higher attrition rates than Y (Gibson & Gibson, 2005). Consistent with previous research, the authors also found that white aviators had higher success rates and African American aviators completed flight school at lower rates compared to their white aviators (Gibson & Gibson, 2005). One significant limitation in this study is the lack of control variables that may be correlated with race and subsequent performance. For example, suppose white aviators have more prior flight training, simulator practice, and education than African American aviators. In that case, those differences may contribute to the difference in outcomes between the two groups. My research will address this problem by including a rich set of controls.

The most current research on the ASTB-E has centered on the Directional Orientation Task (DOT). This portion of the test involves 48 spatial orientation questions, each with four possible responses, and is the only part of the exam that tests spatial orientation. Test takers must quickly respond to two images shown in Figure 15: "a map depicting an aircraft on a specific heading and a forward-facing view from that aircraft showing a building surrounded by four parking lots situated at right angles to each other" (Phillips et al., 2011). However, the nature of this test portion lends itself to problems vice the psychometric portions involving the use of a joystick and throttle.

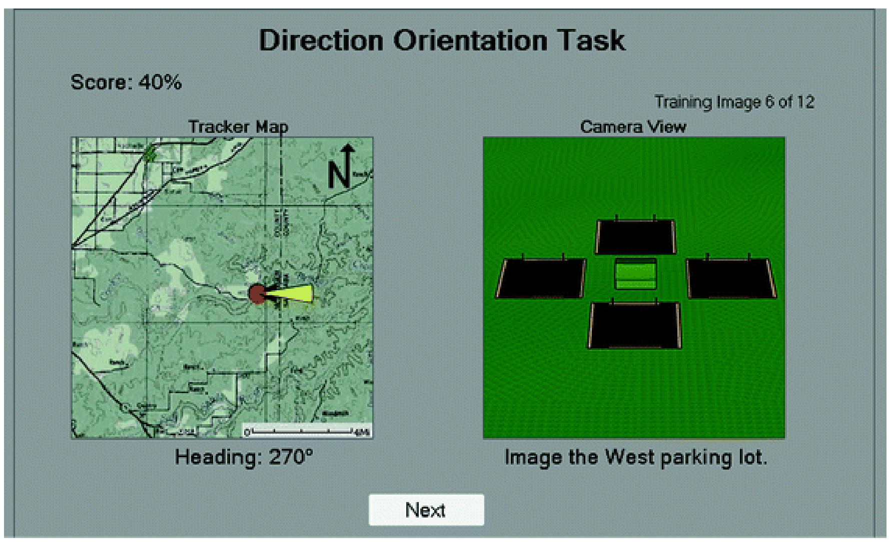


Figure 15. DOT Sample Question 2. Source: Coyne et al. (2022).

A Naval Aerospace Medical Institute (NAMI) group first identified issues with the DOT in 2019 and again in 2022 by another separate NAMI group (Coyne et al., 2022; Keiser et al., 2019). The first group was concerned about a steady rise in DOT scores, the multiple-choice nature of the test, and testers' ability to memorize the directional combinations before the test. There are 12 possible directions, 30-degree increments between 0 and 360, with four possible multiple-choice answers. To address these problems, they tested a newer version of the test they called DOT2. This new test is very similar to the first but increases the number of possible responses from four to twelve and the possible combinations from 48 to 144. These changes resulted in significantly lower scores on DOT2 versus DOT. Test takers took longer to complete the test, and there was a decrease in the probability of guessing the correct answer (Keiser et al., 2019). Overall, the DOT2 is more difficult and has more variance in the scores creating a more valid prediction in flight performance (Keiser et al., 2019).

Unlike the 2019 NAMI group, the second group at NAMI focused on the root of the problem. The test's increasing scores were becoming less valuable in separating applicants based on spatial aptitude (Coyne et al., 2022). Specifically, their concern was that test takers use non-spatial strategies to solve a spatial problem, gain an edge through practice, and use visual aids in the test. Using DOT scores from December 2013 until September 2020, they found that correct scores averaged 67.8% across test-takers in the first full year of DOT in 2014 compared to 75.6% across applicants in the first nine months of 2020 (Coyne et al., 2022). Furthermore, they conducted a hierarchical regression for each year from 2013 to 2020 with two models. One model contained all subtests except the DOT, and the other had all. After 2015, the DOT became significantly less relevant to predicting Naval Standard Scores of aviators, while the other subtests remained fairly constant yearly. This regression model demonstrates the effect of learned practices over natural applicant ability. Keiser et al. (2019) found a similar increase in test scores.

This increase in DOT scoring does not reflect a better spatial capability of the candidate as they are using math, compass rose drawings and practice instead of nature ability. In addition to specific mathematical testing strategies, the authors note that a YouTube video demonstrating how to rotate a drawing of a compass rose on scrap paper to solve DOT problems has more views than ASTB-E test takers (Coyne et al., 2022). Last, Coyne et al. (2022) state that although Keiser et al. (2019) made two attempts to revise the DOT and make it more difficult, both revised versions have the same potential strategy problems as the version of the DOT that is currently being used for selection. Further testing should determine whether the DOT is measuring spatial ability sufficiently or whether new substitute tests need to be developed (Coyne et al., 2022).

D. SUMMARY

Overall, there is possibly still disparity in racial groups regarding flight performance and ASTB scores, but, more importantly, there are effects of the PBM portion of the ASTB we can analyze with a large sample group. Racial differences are noted in Naval Aviation's past, and our data may contain information that shows a similar trend. Unlike some previous research, our data has numerous flight candidates of varying demographics. This will help us accurately assess the effect of the ASTB-E on racial groups if one exists. This is also the first time the ASTB-E's effectiveness against attrition will be assessed. Previous research has focused on one subtest, the DOT, but our research includes all subtests and the ASTB-E when examining historical attrition.

THIS PAGE INTENTIONALLY LEFT BLANK

IV. METHODOLOGY

A. INTRODUCTION

The following section discusses the source and content of our data, some research limitations, details on required data cleaning, and a brief overview of our methods. This provides a foundational understanding of our work and foreshadows future areas of continued research.

B. DATA

NMOTC provided an initial dataset with 11,278 rows of data with 52 variables. Each row contains either a USN or USMC student aviator who attended flight training between 2008 and 2018, demographic information, test identification showing whether the student aviator took either the ASTB or ASTB-E, ASTB scores and sub-scores, and flight performance data.

The same source provided a second data set that contained two excel sheets. One sheet had 2,893 rows and 43 variables, was specifically ASTB-E test takers, and contained similar variables as the previous one with the exemption of demographic information. Also, the attrition and NSS values are more complete than the previous data set. The second has 8,643 rows, 67 variables, and mainly ASTB test takers. This set also contained similar variables to the last one.

All three excel sheets were converted to a Stata Dataset and merged. The ASTB-E-specific rows were isolated from the original dataset, appended to the ASTB-E dataset, and then merged back into the original set. This created more complete ASTB-E rows with the demographic information from the first set and the NSS and attrition values from the second. The second, with mostly ASTB rows, was merged into the original dataset, filling in several missing values.

C. LIMITATIONS OF RESEARCH

While the data obtained is comprehensive, our research had a few limitations. First, much of our demographic information is self-reporting, leading to possible reduced accuracy in the answers. For example, there were several variables for prior flight experience, including the number of flight hours. Some may provide an estimated value, and others appeared to have used crew hours instead of flight hours. Other examples include flight simulator experience, HOTAS experience, and flight simulator hours in the past month.

Second, no dataset contained age information, so we could not account for flight performance differences between younger and older candidates. However, two factors help alleviate this. One, applicants have a maximum age. Pilot candidates must have been commissioned before their 27th birthday with a two-year waiver of up to 29 with prior enlisted experience (Office of Chief of Naval Operations [CNO], 2009). Second, our variable, "Prior Enlisted," provides insight into the candidate's age as prior-enlisted applicants are usually older because of their previous military experience. Altogether, there is a limited window in age to enter flight training, and we have some idea of the candidate's age with a separate variable.

Third, no variable defines the location where the candidate attends Primary Flight Training. SNAs attend this training in either Corpus Christi, TX, or Milton, FL. Factors within that location may affect performance that are not accounted for in any other variable. Without that variable, we cannot include location-based fixed effects as a control.

Last, we could not use certain variables or test the combination of variables used to compute PFAR and AQR due to classification level. The ASTB-E includes variables from the NATFI, which has nine personality traits significant to completing flight training. However, the identity of these traits is Controlled Unclassified Information (CUI), so we could not analyze them explicitly outside of the ASTB-E performance as a whole. Additionally, the specific combination of PBM subtests used to calculate PFAR and AQR is also CUI, so our analysis will focus on each PBM subtest individually or as a group.

D. DATA CLEANING

Many issues with missing values were solved when merging the datasets but, the data still required a thorough examination and cleaning. We resolved most problems by dropping data but adjusted a few following logical conclusions. One initial set of rows dropped contained AQR Raw values very different from the mean of .66, in which we dropped 272 rows. We also dropped 81 rows as the year the individual attended API was less than their test year, which is impossible. Concerning flight hours, we adjusted the number on 94 rows to zero as these rows had aircrew hours listed for aviation experience but no mention of flight lessons or a license. Therefore, we concluded the values listed were aircrew hours and not applicable to our research. Last, we made logical conclusions when adjusting the attrition status in Primary. 78 rows showed SNAs complete in a follow-on training phase, but their Primary status was blank. So, it is logical to conclude that the individual also completed Primary, and we adjusted the empty values accordingly. We used a similar strategy to annotate 34 rows complete in API.

E. METHODS

A regression-based analysis was the most suited method for our research for several reasons. Previous methods discussed in our literature review all used linear regressions but occasionally required more observations. For example, some authors could not make conclusions as the subgroups were too small (Dean, 1996; Gibson & Gibson, 2005). However, we can complete regression-based research in these fields as even our smallest subgroup, Pacific Islanders, has 55 observations, enough to detect any differences. Moreover, our data contains numerous control variables relating specifically to our research field, reducing the possibility of omitted variable bias. For example, flight experience, degree majors, and branch of service correlate with flight training performance. We discuss further details of our methods in the results section.

THIS PAGE INTENTIONALLY LEFT BLANK

V. RESULTS AND RECOMMENDATIONS

A. INTRODUCTION

Below we present our findings, beginning with an overview of our data, some initial findings based on raw data, then a regression analysis of our primary and secondary research questions.

B. DESCRIPTIVE TABLES

The two groups may differ in ways other than just the test they took, which could pose a challenge. For example, recruiting impediments in later years may mean that education is lower for the ASTB-E test takers, which could bias the comparison. Table 1 explores this potential. To illustrate with an example, the first row shows that 91% of ASTB test takers were male compared to 88% of ASTB-E test takers. Column (3) shows that this is a 3-percentage point difference and that it is highly significant.

There are many significant differences between the groups. The ASTB-E group is slightly more diverse. They are 3pp more likely to be female, 4pp less likely to be white, and 3pp less likely to be Hispanic SNAs. The racial subgroup declines are accounted for in the other five groups, notably with an increase of nearly 2pp and 3pp in African and Asian SNAs, respectively.

There were also marked changes in the prior experience and education of the two testing groups. The ASTB-E consists of many more Navy student aviators with an increase of 10pp, but 4pp fewer prior-enlisted SNAs. Prior average flight hours between the two groups remain relatively unchanged. The ASTB group had greater averages of those with higher education, specifically in postgraduate education. There is also a considerable increase of 11pp of those with engineering degrees in ASTB-E test takers, the largest educational major among both groups. However, it should be noted that nearly 40% of the majors in the legacy group are unknown, creating a difference of 31pp.

Last, there are some variances in flight training outcomes between test takers, including attrition, completion, and Naval Standard Score in Primary Flight Training.

While there is no significant change in attrition rates, there is a 2pp decrease in completion among modern test takers. The average NSS of modern test takers also markedly increased by 0.48 points. Overall, it appears that, although NSS is higher for modern test takers, there is early evidence that the ASTB-E was not more effective at predicting flight performance.

The differences between ASTB and ASTB-E candidates noted above provide us with reasons to use controls and carefully interpret the results. For example, there may be flight performance differences between varying racial subgroups, and the ASTB-E population is more diverse than the legacy group. Additionally, Coyne et al. (2022) notes that science, technology, engineering, and mathematics (STEM) majors have high spatial abilities, which could lead to higher scores on one of the subtests, the DOT, and there are marked differences in the number of ASTB-E candidates with certain degree majors in our data.

Table 1. Balance Table

	(1) ASTB	(2) ASTB-E	(3) Difference (1)-(2)
Demographics			
Male	0.91	0.88	0.03***
White	0.90	0.86	0.04***
Hispanic	0.04	0.01	0.03***
African	0.02	0.04	-0.01***
Asian	0.02	0.05	-0.02***
Native American	0.00	0.01	-0.01***
Pacific Islander	0.00	0.01	-0.01***
Other	0.01	0.02	-0.01***
USN	0.64	0.74	-0.10***
Prior Enlisted	0.09	0.06	0.04***
Licensed Pilot	0.09	0.09	0.00
Lessons	0.13	0.14	-0.01
Formal Flt Hours	24.72	24.38	0.34
Education			
Postgraduate Educ	0.01	0.02	-0.01***
Master's Degree	0.01	0.00	0.01***
PhD	0.00	0.00	0.00
Arts Major	0.01	0.02	-0.00*
Business Major	0.08	0.10	-0.02***
Computer Science Major	0.01	0.03	-0.01***
Engineering Major	0.18	0.28	-0.11***
Humanities Major	0.04	0.06	-0.02***
Life Science Major	0.03	0.06	-0.03***
Math Major	0.02	0.04	-0.02***
Natural Science Major	0.05	0.09	-0.05***
Social Science Major	0.08	0.12	-0.04***
Other Major	0.10	0.12	-0.01**
Unknown Major	0.39	0.08	0.31***
Flight Training Outcomes and Scores			
Primary Attrition	0.08	0.08	-0.00
Primary Complete	0.92	0.90	0.02***
Primary NSS	49.40	49.88	-0.48**
Observations	8394	2893	11287

C. INITIAL RESULTS

In December 2013, the Navy began solely testing with the ASTB-E, as shown in Figure 16. The series shows the fraction of test-takers in that year who took the ASTB-E. The line is steady from 2004 to 2012, with only legacy candidates. Then, there is an initial increase in 2013, with just 11 ASTB-E candidates testing in December 2013, before sharply increasing to 100% in 2014. This figure illustrates what we expected with implementation, but the year ASTB-E SNAs began flight training demonstrates the gap between when a candidate tests and when they started flight training.

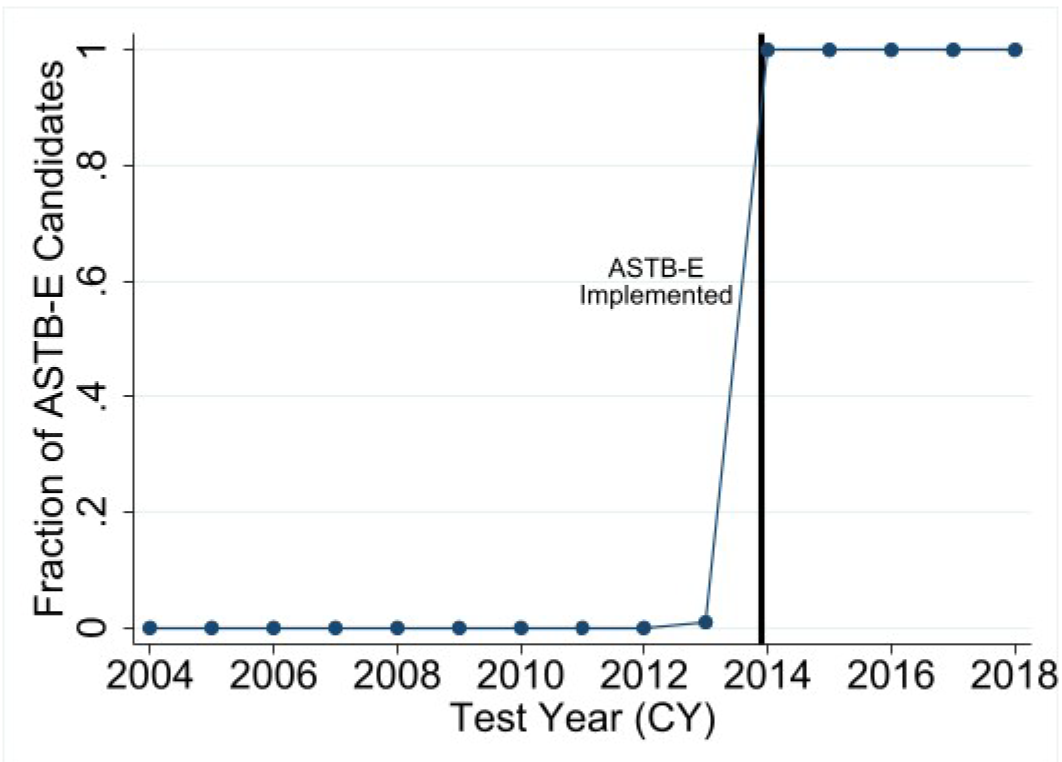


Figure 16. ASTB-E Implementation from 2004 to 2018

Figure 17 shows ASTB-E SNAs beginning API from FY 2008 to 2018. Specifically, the ASTB-E candidates in Figure 16 have now started the initial phase of flight school. For example, the dot in 2015 shows that 21% of SNAs in fiscal year 2015 had taken the ASTB-E, while the remainder had taken the ASTB. The fraction rises gradually rather than spikes because ASTB results of any version remain valid indefinitely. Such as, someone who took the ASTB in 2011 could start flight school in 2016. In the starting year, less than 1% of student aviators had taken the ASTB-E. However, in 2016, the average ASTB-E taker rise to nearly 70%, topping out at 100% in 2018. Subsequent regressions include API year fixed effects but cannot apply test year fixed effects as that would leave no variation in legacy and ASTB-E candidates.

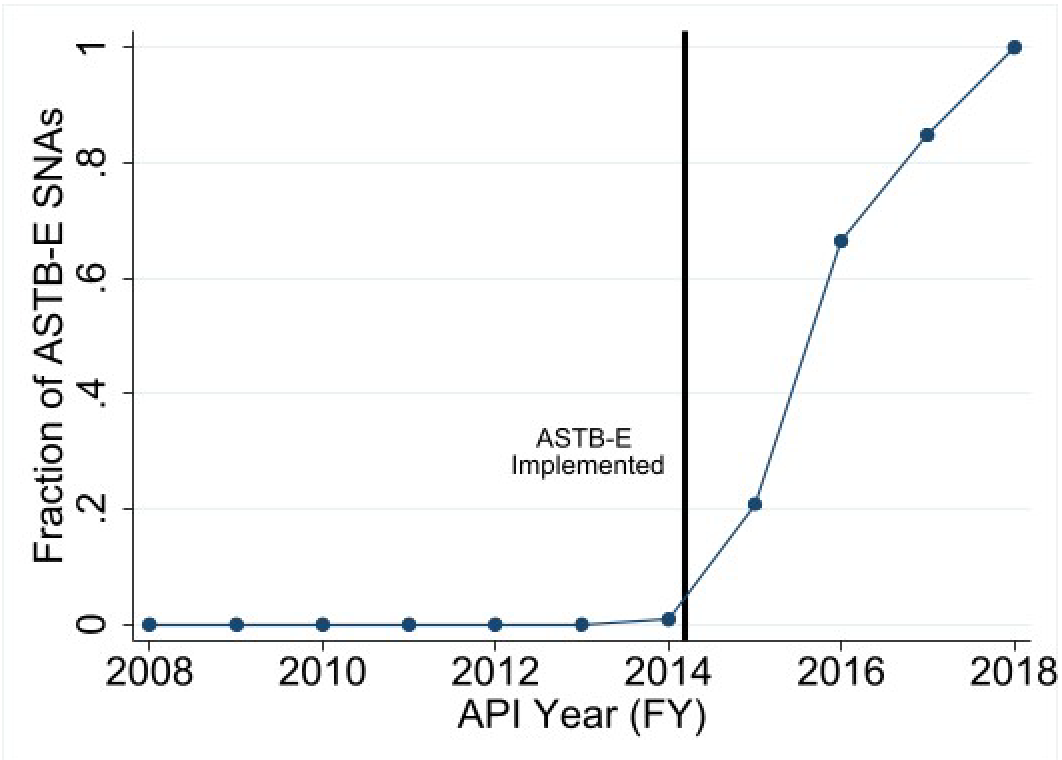


Figure 17. ASTB-E Implementation from FY 2008 to 2018

If the ASTB-E significantly impacted attrition, we should see attrition decrease as students transition to the ASTB-E. Figure 18 tests this, showing the average yearly attrition rate of USN and USMC student aviators from CY (Calendar Year) 2004 to 2018. Attrition, in this case, is defined as those involuntarily leaving the aviation pipeline for reasons other than medical disqualification. The figure shows attrition by test year, but the actual attrition occurs much later in Primary Flight Training. For example, the average attrition in CY 2006 is about .07 or 7%. Attrition rates slowly decrease from CY 2012 to 2015, which coincide with the ASTB-E start, but then climb again. Two years after the ASTB-E began, attrition rates are highest for three consecutive years, peaking at about .11 or 11%. Overall, however, attrition is generally unaffected. The average attrition rate in Primary Flight Training is 8.1% for ASTB candidates and 8.2% for ASTB-E candidates.

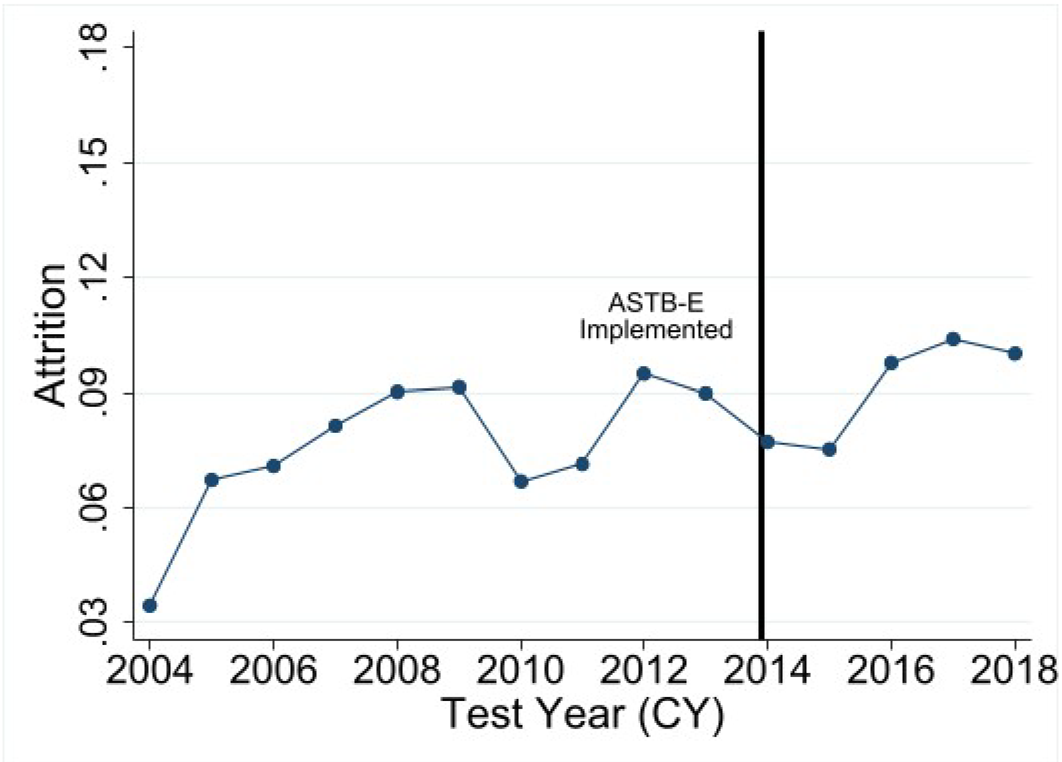


Figure 18. Average Attrition from FY 2008 to 2018

D. PRIMARY RESEARCH QUESTIONS

The averages from Figure 18 could be misleading if there were substantial changes to the population groups. In the following regression, we compensate for possible differences using multivariable regressions.

(1) ASTB-E Effects on Attrition

Table 2 measures the effect of the ASTB-E on attrition using linear regressions. We added several controls to account for demographics, prior experience, and education that may affect attrition. For example, someone with previous flight experience will likely perform better than the average candidate. We added controls in the following sequence:

Column 1: No controls

Column 2: Demographics, prior experience, and education

Column 3: Year-fixed effects

Column 4: Combination of Columns 2 and 3

Column 5: Combination of Columns 2 and 3 and educational majors

The below linear probability model was used to assess the effects of the ASTB-E on attrition.

$$\text{Attrition} = \beta_0 + \beta_1(\text{ASTBE}) + \text{Controls} + \varepsilon$$

The outcome is an indicator of attrition. A "1" represents attrition, and a "0" represents completion. Therefore, positive numbers in Table 2 identify controls contributing to attrition, while negative numbers show controls that reduce attrition. Next, ASTBE is an indicator for which version of the ASTB the candidate took. A "1" represents an ASTB-E candidate, and a "0" represents a legacy candidate. The controls are varied across columns. If the ASTBE reduces attrition, we should find that B1 is negative.

We find no significant effect of taking the ASTB-E on attrition probability. However, based on the 95 percent confidence intervals, we cannot reject that the new test is associated with up to a 3.48pp decline in attrition using the point estimate from column (5), the most controlled regression. The initial regression without control variables shown in column 1 shows a slight positive correlation with attrition at just 0.1pp. When controlling for demographics, the effect is the opposite, showing a decrease in attrition of 0.2pp. However, when solely controlling for API year effects, the effect reverses again with an increase in attrition of 0.4pp. This effect increases to -0.5pp and -1.2pp when controlling for demographics and year effects in column 4 and all controls in column 5, respectively. The ASTB-E is never statistically significant with any of the controls, but several control variables were significant to predicting attrition.

There are some interesting results in the control variables. Being a licensed pilot, prior-enlisted, or a USN Aviator are all statistically significant in each regression, some highly significant with $p<0.01$ when predicting attrition. Being a licensed pilot is associated with a 4.9pp smaller chance of attrition with all controls (column 5) and still a -4.3pp chance with limited controls against those who are not licensed (column 5), all with a p-value less than 0.01. USN student aviators are also 2.0pp more likely to attrite than Marines. Last, those who are prior enlisted are 4.9pp more likely to attrite than those without prior-enlisted time.

Table 2. Effects of ASTB-E on Attrition in Primary Training

	(1)	(2)	(3)	(4)	(5)
ASTB-E	0.001	-0.002	0.004	-0.005	-0.012
	(0.006)	(0.008)	(0.010)	(0.012)	(0.012)
Male		-0.021*		-0.020*	-0.018
		(0.010)		(0.010)	(0.010)
Hispanic		0.036*		0.035*	0.034*
		(0.017)		(0.017)	(0.017)
African		0.039*		0.037	0.035
		(0.019)		(0.019)	(0.019)
Asian		0.068***		0.067***	0.072***
		(0.019)		(0.019)	(0.019)
USN		0.017**		0.018**	0.020***
		(0.005)		(0.006)	(0.006)
Prior Enlisted		0.054***		0.055***	0.049***
		(0.011)		(0.011)	(0.011)
Licensed Pilot		-0.043***		-0.043***	-0.049***
		(0.009)		(0.009)	(0.009)
Sim Expert		-0.018		-0.021	-0.028*
		(0.012)		(0.012)	(0.013)
HOTAS Expert		-0.014		-0.011	-0.001
		(0.032)		(0.032)	(0.032)
Business Major					0.049***
					(0.011)
Humanities Major					0.060***
					(0.015)
Social Science Major					0.048***
					(0.011)
Outcome mean	0.08	0.08	0.08	0.08	0.08
R-squared	0.000	0.011	0.003	0.014	0.018
Observations	11,024	11,024	11,024	11,024	11,024

Standard errors in parentheses
For outcome in all specifications, 1=Attrite and 0=Completed
Prior Flight Hours is calculated per 100 flight hours
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Additional controls as follows: Col 2: Native American, Pacific Islander, Some Masters, master's degree, PhD, Flight Hours, Flight Lessons, Sim Novice, Sim Intermediate, HOTAS Novice, HOTAS Intermediate. Col 3: Year Fixed Effects from 2008–2018. Col 4: All Col 2 & 3. Col 5: All Col 2 & 3, Arts Major, Computer Science Major, Life Science Major, Math Major, Natural Science Major, Other Major, Unknown Major

(2) ASTB-E Effect on NSS

The ASTB-E also affected NSS in API and Primary, as shown in Table 3. Contrary to before, values shown here represent whole point differences on a scale from 20 to 80. Although we applied the same controls used in Table 2, here, they have different effects based on the nature of the training. For example, API is mainly an academic phase, so controls such as being a licensed pilot have less impact on the outcome than on Primary. Additionally, this table highlights variables' effects on flight performance over the attrition table, as NSS is a continuous variable while attrition is binary.

Here, ASTB-E SNAs are associated with having a lower API and Primary NSS when compared to ASTB SNA when applying all controls. Columns 2 and 4 show statistically significant results below 0.1% of a 1.013- and 1.276-point lower NSS for ASTB-E candidates for API and Primary, respectively. Without controls, as shown in columns 1 and 3, ASTB-E SNAs have, on average, a lower API score while having a slightly higher Primary score of about 0.6 points.

There are also several notable results with the control variables. Males, on average, score higher than females in API by 2.04 points. Every racial subgroup except American Indians (not shown) also scores lower than Caucasians in both phases, the results of which are highly statistically significant. Prior enlisted SNAs score lower than non-prior enlisted in both phases as well. Several degree majors make a difference in scores in API and Primary. SNAs with business, humanities, or social science degrees, on average, score lower than those with engineering degrees.

Table 3. ASTB-E Effects on NSS in API and Primary

	API NSS	API NSS	Pri NSS	Pri NSS
ASTB-E	-0.090	-1.013**	0.559*	-1.276**
	(0.153)	(0.309)	(0.241)	(0.449)
Male		2.040***		4.510***
		(0.237)		(0.343)
Hispanic		-2.967***		-4.357***
		(0.397)		(0.607)
African		-4.473***		-5.042***
		(0.459)		(0.630)
Asian		-1.025**		-2.409***
		(0.379)		(0.599)
Pacific Islander		-3.325***		-4.183***
		(0.969)		(1.249)
USN		-0.984***		-1.082***
		(0.144)		(0.225)
Prior Enlisted		-1.518***		-1.766***
		(0.240)		(0.398)
Licensed Pilot		0.619*		5.662***
		(0.313)		(0.473)
Lessons		0.111		0.969**
		(0.212)		(0.312)
Business Major		-3.592***		-1.823***
		(0.267)		(0.400)
Humanities Major		-3.772***		-3.667***
		(0.335)		(0.498)
Social Science Major		-3.685***		-2.422***
		(0.254)		(0.386)
Outcome mean	50.03	50.03	49.56	49.56
R-squared	0.000	0.075	0.001	0.158
Observations	11,103	11,103	8,799	8,799

Standard errors in parentheses
For outcome in all specifications, NSS is scored from 20-80
Prior Flight Hours is calculated per 100 flight hours
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Additional controls as follows: Col 2 & 4: Native American, Some Masters, master's degree, PhD, Flight Hours, Flight Lessons, Sim Novice, Sim Intermediate, Sim Expert, HOTAS Novice, HOTAS Intermediate, HOTAS Expert, Year Fixed Effects from 2008–2018, Arts Major, Computer Science Major, Life Science Major, Math Major, Natural Science Major, Other Major, Unknown Major

E. SECONDARY RESEARCH QUESTIONS

(1) DOT Subtest Validity

Table 4 tests if the DOT subtest is continuing to add validity in addition to the other four PBM subtests yearly from 2015 to 2018. This is a recreation of the previous DOT validity analysis (Coyne et al., 2022, p. 390). In the previous research, Coyne et al. (2022), demonstrated that the DOT added incremental validity when predicting NSS in 2013, 2014, and 2015 but failed to in 2016 and 2017. It is possible, Coyne et al. (2022) concluded, that applicants are, among other issues, using non-spatial strategies to solve a spatial problem, which leads to a less accurate assessment. Validity refers to how accurately the research method measures what it intended to and is gauged by R-squared, as in Coyne et al. (2022). In this case, the ASTB-E subtests are designed to measure a candidate's ability to score well in Primary Flight Training, shown by a high NSS, and our method tests if the DOT subtest is increasing that accuracy.

Our method uses a nested regression which compares the R-squared of two groups of variables. In our first group, we have the subtests that use HOTAS. Our second group adds in the DOT, which measures spatial ability. Then, we compared the R-squared between those two groups of variables.

The R-squared change from adding the DOT varies yearly with no noticeable trend. There is a slight R-squared increase in 2015 of 0.004, which rises to a 0.013 increase in 2016. Our most significant increase is 0.019 in 2017. For example, in 2017, the legacy test components explain 5.7% of the variation in NSS. Adding the new subtests increases the R-squared by 2 pp, to 7.6%. This change in R-squared is highly significant as shown by the p-value. The R-squared difference drops to 0.001 in 2018.

DOT validity rises between 2015 to 2017, demonstrated by an F-stat increase from 1.22 in 2015 to 14.13 in 2017 before sharply dropping off in 2018 to an F-stat of 0.25, the lowest value of any year. The DOT subtest is also statistically significant in 2016 and 2017 with a p-value < 0.01 and < 0.001, respectively. The ATT subtest is also highly statistically significant through 2015 to 2017 with a p-value less than 0.001, which drops to <0.05 in

2018. This regression fails to recreate the previous finding of decreasing validity of the DOT subtest.

Table 4. Hierarchical Regression of DOT Validity

	2015	2016	2017	2018
HOTAS Ability				
ATT	0.177***	0.089***	0.107***	0.047*
	(0.039)	(0.021)	(0.020)	(0.020)
VTT	0.005	0.052	-0.003	0.052
	(0.052)	(0.030)	(0.027)	(0.028)
DLT	0.002	0.108	0.135*	-0.034
	(0.099)	(0.060)	(0.059)	(0.056)
Skill Factor	0.061	0.011	-0.043	0.046
	(0.062)	(0.046)	(0.044)	(0.040)
Spatial Ability				
DOT	0.025	0.042**	0.053***	0.007
	(0.022)	(0.014)	(0.014)	(0.014)
R-squared 1	0.165	0.060	0.057	0.030
R-squared 2	0.169	0.073	0.076	0.031
F-stat (Spatial Ability)	1.22	8.77	14.13	0.25
P-value	0.270	0.003	0.000	0.618
N	208	640	813	837

Standard errors in parentheses
Note 1: Outcome is NSS
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(2) PBM Subtest Validity

The results in Table 4 suggest that there is no distinct decrease in the validity of the DOT subtest over time, in contrast to the findings in Coyne et al. (2022). We can employ the same regression model to measure the validity added by the entire group of PBM subtests. Table 5 tests whether the PBM subtests contribute to validity when predicting NSS. The R-squared values from two regressions are compared using a nested regression. The first regression includes just the legacy subtests shared between the two recent versions of the ASTB. The second then adds the PBM subtests. The outcome is NSS. The new subtests demonstrate an increase in R-squared from 0.124 to 0.152, suggesting an increase in validity. The result is also highly statistically significant based on the F-test of the R-squared difference.

Table 5. PBM Contribution to ASTB-E Predictiveness

	(1) NSS
Legacy Subtests	
MST	0.569
	(0.346)
RCT	0.458
	(0.299)
MCT	1.698***
	(0.312)
ANIT	3.426***
	(0.271)
Novel Subtests	
DOT	0.020**
	(0.007)
ATT	0.055***
	(0.011)
VTT	0.041**
	(0.015)
DLT	0.067*
	(0.029)
Skill Factor	0.002
	(0.022)
R-squared 1	0.124
R-squared 2	0.152
F-stat (Group 2)	17.00
P-value	0.000
N	2552

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(3) Adverse Impact of ASTB-E

Another purpose of the ASTB-E is to reduce adverse impacts as the DOD diversifies its force. Table 6 tests whether the PBM subtests adversely impact demographic subgroups. In columns 1–5, the outcome is a standardized version of the associated subtest score and the independent variables are gender and racial demographics. Column 6 has the same independent variables, but the outcome is the PFAR score on a scale from 1–9. Our test of adverse impacts focuses on whether the coefficients significantly differ between demographics.

As the subtests in Table 6 and PFAR predict performance in Primary, our table also includes Primary outcomes in the form of attrition and NSS from Primary. For example, a subgroup may perform well on certain subtests compared to Caucasians, creating the expectation that they would perform well in Primary. Then this expectation is proven with high scores in Primary. This example can also demonstrate that those specific subtests could have an adverse impact on the other subgroups.

This analysis presents several notable results. Subtest scores, columns 1–5, have been standardized, so each value represents the number of standard deviations from the mean. Columns 7–8, in contrast, show the raw score. First, males, on average, perform better on the ATT, 0.95 SDs, and VTT, 0.82 SDs, subtests, and PFAR than females and, correspondingly, score higher in Primary by about 2 points. Concerning racial subgroups, African Americans score lower on average than Caucasians on the ATT, -0.24 SDs, and have a lower PFAR, -0.47. This trend continues in Primary, where African Americans score 5.85 points lower on average than Caucasians (column 8). Asians have a striking contrast in scores. While Asians, on average, score higher than Caucasians on the ATT, 0.91 SDs, a subtest shown in Table 5 to be highly predictive of NSS, they score 3.3 points lower in Primary than Caucasians. Last, no other significant results exist for different races' performance in the subtests; however, nearly every subgroup except Native Americans performs lower than Caucasians in Primary (column 8).

Table 6. Adverse Impact of PBM Subtests with PFAR, Attrition, and NSS

	(1) DOT	(2) ATT	(3) VTT	(4) DLT	(5) Skill Factor	(6) PFAR	(7) Primary Attrition	(8) Primary NSS
Male	0.161*	0.953***	0.822***	-0.018	0.071	0.603***	0.004	2.022***
	(0.063)	(0.042)	(0.054)	(0.057)	(0.059)	(0.055)	(0.018)	(0.591)
Hispanic	0.143	-0.050	0.010	-0.067	-0.080	0.059	0.020	-4.370**
	(0.146)	(0.139)	(0.143)	(0.168)	(0.152)	(0.143)	(0.049)	(1.492)
African	-0.117	-0.238**	-0.073	-0.116	-0.163	-0.436***	0.026	-5.852***
	(0.117)	(0.088)	(0.105)	(0.113)	(0.109)	(0.104)	(0.032)	(0.973)
Asian	0.081	0.191*	0.166	0.059	0.118	0.077	0.045	-3.265***
	(0.088)	(0.097)	(0.114)	(0.086)	(0.091)	(0.095)	(0.028)	(0.872)
Native American	0.116	0.225	-0.020	-0.118	-0.329	-0.048	0.010	1.590
	(0.163)	(0.156)	(0.145)	(0.198)	(0.201)	(0.185)	(0.052)	(1.595)
Pacific Islander	-0.097	0.317	0.247	-0.130	-0.043	-0.054	-0.041	-3.029
	(0.221)	(0.193)	(0.195)	(0.194)	(0.226)	(0.208)	(0.036)	(1.837)
USN	-0.115**	0.135***	0.261***	-0.081	0.048	-0.002	0.036**	-0.702
	(0.042)	(0.037)	(0.043)	(0.043)	(0.043)	(0.039)	(0.011)	(0.413)
Outcome mean	0.00	0.00	0.00	0.00	-0.00	6.87	0.08	49.96
R-squared	0.028	0.233	0.109	0.012	0.015	0.123	0.031	0.210
Observations	2,804	2,804	2,804	2,804	2,804	2,804	2,796	2,504

Standard errors in parentheses
Note 1: For all specifications in columns 1-5, results have been standardized and represent a number of standard deviations rather than normal test score
Note 2: For all specifications in column 7, 1=Attrite and 0=Completed
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(4) Long-Term PFAR Validity

Candidates sometimes start flight school years after testing. Consequently, we question if the test score from the ASTB remains valid long term. Table 7 examines whether the PFAR score remains valid in predicting NSS with extended time between when an applicant tests and when they start API. Specifically, this assesses the applicability of the policy that ASTB scores never expire. Each column represents the number of years between when an applicant tests and begins API. Column one includes those SNAs with a one-year gap and those that began the same year, as this group was too small to examine alone. For this test, we place close attention to the R-squared value. A practical example of a significant reduction would be half the year one R-squared in year three, a value of 0.060. In year one, it begins at 0.120 and only slightly decreases until year four at 0.089. This is not a significant reduction in R-squared. Furthermore, year five shows the highest value in R-squared of 0.133. Overall, there appears to be no substantial reduction over time of PFAR validity.

Table 7. PFAR Long-Term Validity

	1 Year	2 Years	3 Years	4 Years	5 Years	6 Years
PFAR	7.832***	7.331***	7.042***	7.441***	9.536***	7.005**
	(0.510)	(0.347)	(0.516)	(0.839)	(1.325)	(2.645)
Outcome mean	49.23	48.64	50.47	51.66	51.25	51.70
R-squared	0.120	0.103	0.090	0.089	0.133	0.067
Observations	1,734	3,899	1,886	810	341	100

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

VI. CONCLUSION AND AREAS OF FUTURE RESEARCH

Our research compared the old ASTB to the newest version, the ASTB-E, in several ways. The new version contained novel subtests that were expected to significantly reduce attrition and better predict performance in Primary Flight Training. Our primary research questions test whether the ASTB-E performed as expected regarding attrition and NSS. Our secondary research questions explored further aspects of the ASTB-E including recreating previous subtest research, testing what validity was added by the PBM subtests, checking the new subtests for any adverse impacts, and verifying ASTB-E test score expiration policy. The following section summarizes our results and offers recommendations on policy and continued research.

A. PRIMARY RESEARCH QUESTIONS

(1) Was the ASTB-E successful at reducing attrition in Primary Flight Training?

The ASTB-E did not dramatically reduce attrition in Primary Flight Training. If there were a significant reduction in attrition related to the novel test, our ASTB-E variable would have been negative, indicating a reduction in attrition, and be significant below the 5% level, indicating the result is not by random chance. Our results do suggest a reduction in attrition with a negative coefficient of 0.012 in Table 2, column (5); however, the value is not statistically significant. At most under the 95% confidence interval, we can attribute the ASTB-E to a 3.48pp decline in attrition with all control variables considered.

(2) Was the ASTB-E better able to predict performance of future SNAs in Primary Flight Training?

Similar to our attrition question, the ASTB-E did not better predict performance in Primary Flight Training shown in Table 3. In fact, ASTB-E SNAs were associated with lower performance in API and Primary Flight Training when compared to legacy SNAs as indicated by their NSS. Contrary to our attrition testing, our ASTB-E coefficient was statistically significant in both the API and Primary regressions with controls.

B. SECONDARY RESEARCH QUESTIONS

(1) Is the DOT subtest continuing to contribute predictiveness to the ASTB-E between 2015 and 2018?

Our DOT model does not show the decreasing validity trend of the previous research. In the previous model by Coyne et al. 2022, fig. 3, the gain in R-squared steadily decreased between 2013 and 2017. Our model in Table 4 showed an increase in validity from 2015 to 2017 before it became negligible in 2018. Further, matching years show different R-squared values from their model to ours, but we also have different numbers of observations per year.

(2) To what extent are the new PBM subtests contributing to the predictiveness of the ASTB-E?

Several groups, including NMOTC, expect that the new PBM subtests will help assess more variance in predicting flight training outcomes leading to the selection of better flight candidates (Navy Medicine Operation Training Command, n.d.-a; Keiser et al., 2019; Phillips et al., n.d.). The PBM tests are, in fact, significantly adding incremental validity when predicting NSS with an R-squared increase of 18.4%. This is demonstrated by a 0.028 increase in R-squared, which is highly statistically significant with a p-value less than 0.000. Additionally, the ATT, DOT, and VTT are all statistically significant when predicting NSS with p-values less than 0.001, 0.01, and 0.01, respectively. We also note that prior research has shown significant overlap between the PBM subtests, and spatial and mechanical testing present in the legacy subtests, which may help explain why there wasn't a larger increase in R-squared (North & Griffin, 1977).

(3) Do the new PBM subtests create any adverse impacts among varying demographics?

We notice several adverse impacts of differing degrees among demographics. In particular, women and African Americans scored lower, on average, than males and other racial subgroups, respectively. As shown in Table 6, females score .953 and 0.822 SDs lower than males on the ATT and VTT, respectively, and both results are highly statistically significant. Also, African Americans score 0.238 SDs lower on the ATT and

.436 points lower on the PFAR than Caucasians. Although, it should again be noted that the formula for calculating the final PFAR is controlled information and lower performance in one subtest can be offset by high performance in another (Navy Medicine Operational Training Command, n.d.-a). For instance, the ATT and VTT may adversely impact females but may not be included when determining PFAR. Additionally, while females may, on average, perform lower on the ATT, that score could be made up by a high score on the DLT, a score not demonstrating adverse impact.

Of particular importance is the opposing scores for Asians. Asians, on average, actually perform better on the ATT, but then perform lower in Primary. In other words, the subtest used to predict performance in Primary predicts higher scores for one subgroup, but then that same subgroup's actual performance in Primary is higher. Our research, however, focuses on the adverse impact of the subtests and not Primary itself.

(4) Does the PFAR remain valid with increasing time between testing and commencing flight training?

Yes, the PFAR remains valid through prolonged periods between testing and starting flight training. As the PFAR is a single score used to predict performance in Primary and PFAR scores, as per the current policy, do not expire, this score must remain valid through the sometimes years that pass between predicted and actual performance. If the score's validity drops over time, we would expect the R-squared value to drop significantly 2–3 years after testing. Although we note some R-squared drop in Table 7, it does not suggest a flaw in the current policy.

C. CONTINUED RESEARCH AND RECOMMENDATIONS

Upon review of our results and conclusions, we also foresee several research areas to explore further and have two recommendations. This includes another similar analysis completed with a more recent data pool, further research into adverse impacts in Primary, and testing what effect malfunctioning hardware and software may have on the ASTB-E test outcomes.

While we had a large data pool, it would benefit decision-makers to have similar research recreated but with more current observations, in particular as we get further from the end of the covid pandemic. Our data pool stopped in 2018, just one year before the Covid-19 outbreak. The pandemic may have had untold effects on test scores and performance in Primary. In the near future, another research could pull more current data that is likely clear of any effects from Covid-19 and produce more up-to-date results and conclusions.

Where our research scope stopped at testing, another could continue work with adverse impacts in Primary. We noted certain minor adverse impacts in the subtests, but there were significant impacts in Primary Flight Training based on NSS. More research is needed into what other adverse impacts may exist in Primary, such as follow-on pipeline selection and, more importantly, why they are occurring and what solutions are there to correct the score imbalance.

Last, MEPS Operation Officers and individuals on social media have mentioned malfunctions with ASTB-E test equipment and the link to APEX (Airwarriors, 2021; M. Ashley, personal communication, September 28, 2022). If there are such problems, this creates a research opportunity into the effect and possible solutions. As stated above, the PBM subtests use a joystick and throttle to complete most sections, and test completion depends upon a constant internet connection to APEX, the server used to grade subtests. A malfunction in the connection or equipment could affect the examinee's score. This effect could be positive when an examinee gets more time to test as the connection is reestablished or negative if the joystick input does not match the aircraft motion on the screen.

We also have two recommendations: reexamine how the PFAR score is calculated and closely research what effect adversely impacted NSS scores have on attrition.

The PFAR score is calculated using specific subtest scores, the combination of which may not be as effective as it could be when predicting performance. As stated before, the formula for PFAR is controlled information and was not released; however, as the ASTB-E was ineffective at significantly reducing attrition, this presents an opportunity to reevaluate the PFAR score. Specific subtests are highly predictive of Primary outcomes and may be underutilized when calculating PFAR, while others overutilized. This becomes especially challenging, though, when balancing predictiveness with low overall adverse impacts. For example, the PFAR could more heavily weigh the ATT, which is highly predictive of NSS but is also shown to have adverse impacts on African Americans and females.

The adverse impacts on racial subgroups shown in Table 6 may also affect attrition not shown in our study, prompting a recommendation to examine those effects closely. While this may be considered another area of further research, we recommend this because the ASTB-E was shown not to affect attrition, and, at the same time, we proved the PBM subtests increase the validity of the ASTB-E. The dual goals of decreased attrition and adverse impact with the ASTB-E may have conflicted. Reducing the test’s adverse impact created a more diverse population of SNAs, shown in Table 1; however, we demonstrate that most racial subgroups also experience lower NSS in Primary, shown in Table 6. Therefore, with an increased population of diverse groups experiencing lower NSS, on average, this may have affected attrition, but was not apparent in our study. Further analysis could help prove that the ASTB-E was reducing attrition, but that adverse impacts negatively affected those results.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Airwarriors. (2021, July 19). *ASTB changes?*

Arnold, R., & Phillips, J. (2008). *Causes of student attrition in U.S. naval aviation training: A five year review from FY 2003 to FY 2007*. Naval Aerospace Medical Research Laboratory.

Chief of Naval Air Training [CNATRA]. (2021*). T-6B joint primary pilot training* (CNATRAINST 1542.166C). Department of the Navy.

Chief of Naval Air Training [CNATRA]. (2022*). Naval flight student training administration manual* (CNATRAINST 1500.4K). Department of the Navy.

Coyne, J. T., Dollinger, S., Brown, N., Foroughi, C., Sibley, C., & Phillips, H. (2022). Limitations of current spatial ability testing for military aviators. *Military Psychology*, *34*(1), 33–46. https://doi.org/10.1080/08995605.2021.1965786

Dean, B. (1996). *Aviation selection testing: The effect of minimum scores on minorities* [Master's thesis, Naval Postgraduate School]. Defense Technical Information Center. https://apps.dtic.mil/sti/pdfs/ADA307314.pdf

Gibson, S. G., & Gibson, J. R. (2005). Completion of naval flight training: Do gender or race matter*? Equal Opportunities International*, *24*(1), 35–43. https://doi.org/10.1108/02610150510787944

Keiser, H. N., Moclaire, C. M., King, K. M., Brown, N. L., Foroughi, C. K., Sibley, C., & Coyne, J. T. (2019). Updating the Direction Orientation Task: An aviation selection tool. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *63*(1), 1414–1418. https://doi.org/10.1177/1071181319631451

Kennedy, R. S., Lane, N. E., Drexler, J. M., & Bramble, W. J. (1998). An automated test battery for advanced aviator assessment. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 42*(11), 792–796. https://doi.org/10.1177/154193129804201105

Mometrix. (2018, December 18). *ASTB-E study guide questions*. https://www.mometrix.com/academy/astb-e/

Natali, M. W. (2018). *History, evolution, and validity of the ASTB.* United States Naval Aerospace Community, Pensacola, Fl.

Navy Medicine Operational Training Command. (n.d.-a). *ASTB-FAQ*. Retrieved February 9, 2023, from https://www.med.navy.mil/Navy-Medicine-Operational-Training-Command/Naval-Aerospace-Medical-Institute/ASTB-FAQ/#:~:text=Time%20to%20complete%20the%20ASTB,complete%20the%20OAR%20in%20APEX.

Navy Medicine Operational Training Command. (n.d.-b). *Department 53OP – Operational psychology.* Retrieved February 7, 2023, from https://www.med.navy.mil/Navy-Medicine-Operational-Training-Command/Naval-Aerospace-Medical-Institute/Department-53-Operational-Psychology/

Navy Medicine Operational Training Command. (2014). *Aviation Selection Test Battery sample questions*. https://www.med.navy.mil/Portals/62/Documents/NMFSC/NMOTC/NAMI/Clinical%20Medicine/Operational%20Psychology/ASTB_SampleQuestions_13May14.pdf?ver=8TT5lJfvAFsbSc3vezTe8w%3d%3d&timestamp=1629210016194

North, R., & Griffin, G. (1977). *Aviator selection 1919–1977* (Special Report No. 77-2). Naval Medical Research and Development Command.

Office of Chief of Naval Operations [CNO]. (2009). *Enlisted to officer commissioning program application administrative manual*.

Ostein, S. (2007). *An assessment of the Performance-Based Measurement Battery (PBMB), the Navy's psychomotor supplement to the Aviation Selection Test Battery (ASTB)* [Master's thesis, Naval Postgraduate School]. NPS Archive: Calhoun. https://calhoun.nps.edu/handle/10945/3124

Phillips, J. B., Chernyshenko, S., Drasgow, F., & Phillips, H. L., IV. (n.d.). *Development of scoring procedures for the Performance Based Measurement (PBM) Test: Psychometric and criterion validity investigation* (No. 12–10).

Acquisition Research Program
Naval Postgraduate School
555 Dyer Road, Ingersoll Hall
Monterey, CA 93943

www.acquisitionresearch.net