AFRL-RH-WP-TR-2023-0007

DEVELOPMENT AND PSYCHOMETRIC EVALUATION OF PREDICTIVE SUCCESS MODELS FOR US AIR FORCE RATED CAREER FIELDS

Montana R. Woolley
Julia L. Walsh
Infoscitex, a DCS company

Thomas R. Carretta
Air Force Research Laboratory
Wright-Patterson AFB, OH

Amanda N. Mouton
Andrew R. Deregla
Air Force Personnel Center
Strategic Research & Assessments Branch
JBSA-Randolph AFB, TX

MARCH 2023
Interim Report

DISTRIBUTION A. Approved for public release: Distribution unlimited.

AIR FORCE RESEARCH LABORATORY
711TH HUMAN PERFORMANCE WING,
AIRMAN SYSTEMS DIRECTORATE,
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

NOTICE AND SIGNATURE PAGE

AFRL-RH-WP-TR-2023-0007 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION.

CARRETTA.THOMAS.R.1228984929 Digitally signed by CARRETTA.THOMAS.R.1228984929 Date: 2023.05.15 14:09:08 -04'00'

THOMAS R. CARRETTA, PhD
Work Unit Manager
Performance Optimization Branch
Airman Biosciences Division

WILLIAMS.LOGAN.ANDREW.1273597634 Digitally signed by WILLIAMS.LOGAN.ANDREW.1273597634 Date: 2023.05.15 15:14:52 -04'00'

LOGAN R. WILLIAMS, DR-III, PhD
Human Performance Area Lead
Operational Product Section
Product Development Branch
Airman Biosciences Division

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE (*DD-MM-YY*)	2. REPORT TYPE	3. DATES COVERED (*From - To*)
21-02-23	Interim	April 2022 – September 2022

4. TITLE AND SUBTITLE
Development and Psychometric Evaluation of Predictive Success Models for US Air Force Rated Career Fields

5a. CONTRACT NUMBER
FA8650-21-F-4104

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
00000F

6. AUTHOR(S)
Montana R. Woolley[a], Julia L. Walsh[a], Thomas R. Carretta[b], Amanda N. Mouton[c], and Andrew R. Deregla[c]

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER
H12Q

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Infoscitex, a DCS Company[a]
4027 Colonel Glenn Highway, Suite 210
Dayton, OH 45431-1672

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Materiel Command[b]
Air Force Research Laboratory
711th Human Performance Wing
Airman Systems Directorate
Airman Biosciences Division
Performance Optimization Branch
Wright-Patterson AFB, OH 45433

Air Force Personnel Center[c]
Strategic Research and Assessments Branch
550 C Street West, Ste. 10
JBSA – Randolph AFB, TX 78150

10. SPONSORING/MONITORING AGENCY ACRONYM(S)
711 HPW/RHBC

11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)
AFRL-RH-WP-TR-2023-0007

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution Statement A. Approved for public release.

13. SUPPLEMENTARY NOTES
Report contains color. AFRL-2023-1433, cleared 27 March 2023

14. ABSTRACT
The Air Force Officer Qualifying Test (AFOQT) is used for officer commissioning and for qualification of already commissioned officers into aircrew training programs, including manned aircraft pilot, Remotely Piloted Aircraft (RPA) pilot, Combat Systems Officer (CSO), and Air Battle Manager (ABM). While the current AFOQT composites have demonstrated strong predictive validities against aircrew training outcomes, they also demonstrate moderate to large mean score subgroup differences for females and racial/ethnic minorities. Using regression and Pareto Optimization (PO) methods, the current study investigated the utility of new Predictive Success Models (PSMs) that combined the AFOQT cognitive subtests and the Self-Description Inventory – Officers (SDI-O) facets, with the goal of maintaining (or improving) predictive validities while reducing mean score subgroup differences, resulting in higher qualification of females and racial/ethnic minorities in aircrew training programs. The results showed that the new PSMs hold promise to improve (or maintain) predictive validities while decreasing (or maintaining) mean score subgroup differences.

15. SUBJECT TERMS
Air Force Officer Qualifying Test, AFOQT, Self-Description Inventory – Officers, SDI-O, predictive success models

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT:	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON (Monitor) / 19b. TELEPHONE NUMBER (*Include Area Code*)
Unclassified	Unclassified	Unclassified	SAR	26	Thomas Carretta / N/A

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1.0 EXECUTIVE SUMMARY

**Please note that the public release (Distribution A) version of this document has limited information. Please contact AFPC/DSYX (POC) for the complete version of this technical report.*

The Air Force Officer Qualifying Test (AFOQT) is used to qualify applicants for officer commissioning through the Reserve Officer Training Corps (ROTC) and Officer Training School (OTS) and for aircrew training qualification (manned aircraft pilot, Remotely Piloted Aircraft [RPA] pilot, Combat Systems Officer [CSO], and Air Battle Manager [ABM]). Although the rated career field aptitude composites have shown predictive validity against initial aircrew skills training for many years, they also have demonstrated moderate to large subgroup differences (SGDs) for females and racial and ethnic minorities. Historically, the AFOQT aptitude composites have been computed from a combination of the cognitive subtests. The purpose of this study was to examine the utility of new Predictive Success Models (PSMs) for aircrew training qualification based on the AFOQT cognitive subtests and personality facets from the Self-Description Inventory – Officers (SDI-O) to be used for pilots, CSOs, and ABMs. The new PSMs utilized the current cognitive subtest composites as a baseline. Analyses were performed to determine the utility of adding SDI-O personality facets to the models to improve predictive validity and diversity and inclusion of females and racial/ethnic minorities.

Three statistical methods were utilized to create potential new PSMs: non-linear multiple regression, corrected linear multiple regression (NLMR), and corrected Pareto optimization (PO). The best performing models created from each statistical method were then tested against each other and against the original cognitive composites. These models were subject to criterion-related validation and SGD examinations, with the goal of maintaining or increasing predictive validity, while decreasing SGDs for women and underrepresented racial/ethnic groups.

The new models were found to be successful in increasing criterion-related validity and decreasing (or maintaining) SGDs. The new PSMs will be used as alternate aptitude composites alongside the current cognitive-based aptitude composites for AFOQT Form T Version 2 when it is implemented in 2023.

2.0 INTRODUCTION

The Air Force Personnel Center Strategic Research and Assessments Branch (AFPC/DSYX) initiated a contract to document the psychometric properties of the SDI-O and the SDI+. During the execution of the project, AFPC/DSYX expanded the scope to develop and test new PSMs that would include cognitive subtests from the AFOQT Form T and personality facets from the SDI-O. This expansion of the scope of work was sparked by ongoing United States Air Force (USAF) efforts to predict important organizational outcomes and widen the selection aperture for historically underrepresented subgroups. This technical report describes the results of the mathematical combination of the AFOQT cognitive subtests and the SDI-O facets to form new predictive composites. A comprehensive psychometric evaluation of the SDI-O and SDI+ are documented in several separate technical reports (e.g., Woolley et al., 2022, Mann et al., in press). The introductory section briefly overviews the history of the AFOQT and its composites.

2.1 AFOQT Brief History

The AFOQT has been an important component of the Air Force Personnel Testing Program (AFPTP) since 1953. It is a critical tool for officer selection and aircrew training classification and is widely accepted among military personnel selection communities as a useful and cost-effective instrument. Historically, the AFOQT has been the primary selection test for the Air Force ROTC, OTS, and the Airman Education and Commissioning Program (AECP). It is also used in the selection process for Undergraduate Pilot Training (UPT), Undergraduate RPA Training (URT), CSO training, and ABM training. Since its inception, the AFOQT has undergone several revisions to improve both its performance prediction and officer classification (see Drasgow et al., 2010 for a complete history of the AFOQT).

2.2 AFOQT Composites

Originally, the AFOQT contained only cognitive subtests that were designed to measure several different aspects of cognitive ability (i.e., verbal, math, spatial, perceptual speed, and aircrew knowledge). The cognitive subtests were combined mathematically to form composites (see Table 1 for the AFOQT Form T composites). Through the decades, the cognitive composites have demonstrated a strong track record of criterion-related validity (Aguilar, 2017; Carretta et al., 2016).

Table 1: AFOQT Form T Composites

Subtest	Abbreviation	Composites					
		Verbal	Quantitative	Academic Aptitude	Pilot	CSO	ABM
Verbal Analogies	VA	X		X			X
Arithmetic Reasoning	AR		X	X			
Word Knowledge	WK	X		X		X	
Math Knowledge	MK		X	X	X	X	X
Reading Comprehension	RC	X		X			
Situational Judgment Test	SJT						
Self-Description Inventory-Officers	SDI-O						
Physical Science	PS						
Table Reading	TR				X	X	X
Instrument Comprehension	IC				X		X
Block Counting	BC					X	X
Aviation Information	AI				X		X

Note. SJT, SDI-O, and PS were included as experimental subtests in AFOQT Form T.

In 2005, with the implementation of the AFOQT Form S, an experimental non-cognitive measure of personality, the Self-Description Inventory Plus (SDI+), was added. In the subsequent psychometric evaluations, it became clear that revisions to the SDI+ were necessary (Manley & Weissmuller, 2017). In 2015, the SDI+ was replaced by the SDI-O with the implementation of the AFOQT Form T. The SDI-O improved upon the SDI+ by including additional personality facets and by demonstrating higher reliability (see Woolley et al., 2022). An additional non-cognitive measure, a situational judgment test (SJT) was also included in the AFOQT Form T in 2015. The description of the SJT is beyond the scope of the current technical report and those with interest are encouraged to read Barron (2013), Walsh et al. (2022), and Sizemore et al. (2022). Generally, the goals of including non-cognitive measures in the AFOQT were to (1) broaden the assessment of critical officer and aircrew attributes and (2) combine the non-cognitive scores with cognitive scores such that predictive validity is maintained or improved, and SGDs are minimized, resulting in improved diversity within the USAF.

The most recent psychometric evaluations of the AFOQT Form T revealed that further revisions to the cognitive subtests were in order (see Kantrowitz et al., 2022; Walsh, Brady, et al., 2022; Walsh, Woolley, et al., 2022). These revisions included replacing problematic items in the cognitive subtests with new items designed to optimize the diversity-validity relation. These changes will be implemented in the AFOQT Form T Version 2 (see Kantrowitz et al., in press). Additionally, AFPC/DSYX decided to explore the potential of combining existing cognitive composites with SDI-O facets. This decision was influenced by recent personnel selection literature suggestions that compensatory models (e.g., higher personality scores being used to offset lower cognitive scores) can help minimize the diversity-validity dilemma (Rupp, Song, & Strah, 2020). Thus, the next section describes the methods for creating the alternative composites.

3.0 METHOD

The present research leveraged NLMR, corrected linear multiple regression (MR), and corrected PO methods to create and test alternative AFOQT composites for pilots, CSOs, and ABMs. Exploration of alternative composites used to predict success for specific types of pilots (i.e., fighter, mobility, RPA pilots) was also of interest. Using the same methods, we explored creating an experimental tailored composite that predicts RPA pilot success and could potentially be incorporated in the AFOQT in the future. Participant demographics and the statistical methods are discussed below.

3.1 Participants

We leveraged archival AFOQT Form T data collected between 2016 and 2020. Of that data, most test-takers wanted to become manned aircraft pilots (*N* = 1,187 [N - Sample Size]), followed by those who wanted to become RPA pilots (*N* = 719), and then those who wanted to become CSOs (*N* = 658). Examinees who wanted to become ABMs were the smallest portion of the sample (*N* = 267). Most of examinees across all samples were male (73 percent (%)-92%) and White (69%-78%; see Table 2).

Table 2: Demographic Breakdowns

Variable	Pilot		CSO		ABM		RPA Pilot	
Sex	***N***	**%**	***N***	**%**	***N***	**%**	***N***	**%**
Male	1,055	89	534	81	196	73	665	92
Female	132	11	124	19	71	27	54	8
Unknown	0	0	0	0	0	0	0	0
TOTAL	**1,187**	**100%**	**658**	**100%**	**267**	**100%**	**719**	**100%**
Race	***N***	**%**	***N***	**%**	***N***	**%**	***N***	**%**
Non-Hispanic White	921	78	474	72	184	69	508	71
Non-Hispanic Black	30	3	15	2	16	6	22	3
Non-Hispanic Asian	31	3	18	3	12	4	20	3
Non-Hispanic AIAN	1	0	0	0	0	0	1	0
Non-Hispanic NHPI	1	0	1	0	0	0	3	0
Non-Hispanic Multiracial	91	8	57	9	21	8	79	11
Hispanic, Any Race (besides White Only, Black Only)	13	1	17	3	8	3	15	2
Hispanic Only (No race selected)	17	1	12	2	1	0	14	2
Hispanic White	64	5	53	8	22	8	48	7
Hispanic Black	5	0	2	0	1	0	3	0
Any Ethnic / Racial Minority	259	22	181	28	83	31	210	29
Unknown	7	1%	3	0	0	0	1	0
TOTAL	**1,180**	**122%***	**655**	**127%***	**267**	**130%***	**719**	**129%***
Race and Sex	***N***	**%**	***N***	**%**	***N***	**%**	***N***	**%**
Non-Hispanic White Men	823	70	387	59	140	52	508	71
Any Ethnic / Racial / Gender Minority	357	30	268	41	127	48	210	29
Unknown	0	0	0	0	0	0	1	0
TOTAL	**1,180**	**100%**	**655**	**100%**	**267**	**100%**	**718**	**100%**
Socio-Economic Status	***N***	**%**	***N***	**%**	***N***	**%**	***N***	**%**
Much higher than average	94	8	20	3	6	2	32	4
Somewhat higher than average	488	41	196	30	86	32	194	27
Average	386	33	280	43	123	46	286	40
Somewhat lower than average	117	10	106	16	34	13	128	18
Much lower than average	26	2	28	4	10	4	40	6
Declined to respond	74	6	26	4	8	3	38	5
Unknown	2	0	2	0	0	0	1	0
TOTAL	**1,187**	**100%**	**658**	**100%**	**267**	**100%**	**719**	**100%**
Education	***N***	**%**	***N***	**%**	***N***	**%**	***N***	**%**
Completed 12	2	0	6	1	1	0	5	1
Completed 13	190	16	151	23	52	19	96	13
Completed 14	466	39	98	15	36	13	179	25
Completed 15	238	20	92	14	33	12	162	23
Completed 16	239	20	241	37	113	42	215	30
Completed 17	27	2	47	7	25	9	29	4
Completed 18	16	1	17	3	6	2	26	4
Completed 19	3	0	3	0	0	0	3	0
Completed 20	3	0	1	0	0	0	1	0
Completed 21+	2	0	0	0	1	0	1	0
Unknown	1	0	2	0	0	0	2	0
TOTAL	**1,187**	**100%**	**658**	**100%**	**267**	**100%**	**719**	**100%**
Academic Degree	***N***	**%**	***N***	**%**	***N***	**%**	***N***	**%**
High School Diploma	855	72	291	44	109	41	360	50
Associates Degree	88	7	75	11	21	8	122	17
Bachelor's Degree	224	19	275	42	129	48	217	30
Master's Degree	17	1	15	2	8	3	19	3
Doctoral Degree	0	0	0	0	0	0	0	0
Unknown	3	0	2	0	0	0	1	0
TOTAL	**1,187**	**100%**	**658**	**100%**	**267**	**100%**	**719**	**100%**
Accession Source	***N***	**%**	***N***	**%**	***N***	**%**	***N***	**%**
OTS – Civilian	148	12	222	34	91	35	87	12
OTS – Active Duty	50	4	95	15	37	14	222	31
AECP	0	0	0	0	1	0	1	0
USAFA	517	44	32	5	18	7	189	26
ROTC	304	26	259	40	84	32	167	23
ANG	110	9	22	3	26	10	41	6
AFRES	58	5	23	4	3	1	10	1
TOTAL	**1,187**	**100%**	**653**	**100%**	**260**	**100%**	**717**	**100%**
Age	***M***	***SD***	***M***	***SD***	***M***	***SD***	***M***	***SD***
	21.63	2.85	22.7	3.32	22.89	3.32	24.4	5.01

*The total percent for race/ethnicity is greater than 100% because test takers could indicate more than one race. AFRES = Air Force Reserves.

3.2 Measures

3.2.1 Criteria Measures

Criteria data came from various aircrew training programs. The main criteria utilized in the present research were the Merit Assignment Selection System (MASS) scores. MASS scores are composites that indicate the overall assessment of a trainee's airmanship based on indicators of performance, such as academic grades, check flight scores, daily flight scores, and flight commander ratings. MASS scores range from 0 to 100.

Overall, the MASS scores were chosen as the primary criteria based on many discussions with AFPC/DSYX personnel. For manned aircraft and RPA pilots, the main criteria were the MASS scores obtained at the end of Specialized Undergraduate Pilot Training (SUPT) Primary training. The MASS scores from Initial Flight Training (IFT) and from SUPT Advanced training were used to cross-validate the new pilot composites. For CSOs, the main criteria were the MASS score obtained at the end of the CSO Primary training. The MASS scores from IFT and from CSO Advanced training were used to cross-validate the new CSO composites. For ABMs, we only had access to the MASS scores that came from Tyndall Air Force Base (AFB). No cross-validation was possible for ABM training. Finally, for RPA pilots, we only had access to the IFT MASS scores, and cross-validation was not possible.

All criteria data contained a SPARK ID (Strategic Policy Analysis Resource and Knowledgebase Identification), a unique AFPC-generated alphanumeric indicator used to identify each trainee. The SPARK ID was used to match demographic data and predictor measures to the criteria data.

3.2.2 Predictor Measures

The predictor measures consisted of cognitive and personality scores obtained from the AFOQT Form T test takers between 2016 and 2020. The cognitive scores included the AFOQT composite scores of record for pilots, CSOs, and ABMs. The personality scores included the SDI-O facet-level scores from the test-takers' first attempt on the AFOQT. The SDI-O consists of 30 facets (see Table 3), but only 26 facets were used. This was due to an administrative error which affected the scores on these four facets. The data (N = 60,066) were cleaned for missingness and carelessness (see Arias et al., 2020; Bowling et al., 2021; DeSimone et al., 2018; Huang et al., 2012). Carelessness was identified using several post-hoc statistical procedures: (1) longstring analysis (identified 187 careless responders); (2) intra-individual response variability (identified 61 careless responders); and (3) odd even consistency (identified 381 careless responders). See Woolley et al., 2022 for a detailed review of data cleaning procedures.

Table 3. SDI-O Facet-to-Domain Linkages

Domain	Facet #	Facet Name
Agreeableness	F01	Team Player
	F07	Pleasant
	F13	Helpful Altruistic
	F25	Optimist
	F29	Well-Adjusted
Conscientiousness	F04	Achievement Striving
	F10	Order
	F16	Self-Discipline
	F22*	Deliberation*
	F26	Unconventional
Extraversion	F03	Reserved
	F09	Dominance-Leader
	F15	Excitement Seeking
	F21*	High-Intensity Pleasure*
	F27*	Spontaneous-Variety*
	F28	Activity
Machiavellianism	F06	Interpersonal Tactics
	F12	Cynical View
	F18	Envious
	F19	Independent
	F24	Influence Tactics
Neuroticism	F02	Stress Under Pressure
	F08	Temperamental
	F14	Worry
	F20	Angry Hostility
Openness	F05	Creative
	F11	Reflective
	F17	Scientific Interest
	F23	Cultured
	F30*	Imagination*

*Indicates the SDI-O facets that were excluded from the analyses

3.3 Technical Approach

As previously described, the goal of this research was to create and test new composites for manned aircraft pilots, CSOs, ABMs, and RPA pilots using the original AFOQT cognitive composites and any of the 26 SDI-O personality facets. To accomplish this goal, we first narrowed down the number of facets to be considered for each career field by examining their inter-correlations, theoretical linkages, and inclusion in stepwise regression models. Next, using the short list of the facets, we applied three different approaches to creating the new composites: (1) NLMR, (2) linear MR with range restriction corrections, and (3) PO with range restriction corrections. The paragraphs below describe each approach in greater detail and discuss its advantages and disadvantages.

3.3.1 Non-linear Multiple Regression

NLMR was utilized because of interest in the potential non-linear relationship between personality and performance (e.g., Benson & Campbell, 2007). These analyses were limited to only those facets that displayed significant quadratic relationships to the training criteria. First, the facet scores were transformed into *z*-scores (for mean centering). Next, the linear terms were entered into the regression Model 1. Then, the quadratic terms were entered into regression Model 2. If Model 2 outperformed Model 1 ($p < .10$), then the quadratic term was considered for inclusion in the new composite scores. It should be noted that if the quadratic term was used in a composite score, it was always used in conjunction with its linear counterpart.

One methodological limitation of this method is that the predictors were not corrected for range restriction, due to violations of the linearity assumption underlying range restriction corrections (Lawley, 1943). Although failing to account for range restriction can result in biased validity coefficients and underestimated explained variance, we believe that the range restriction corrections will have only small effects on the personality facets. This is because the only predictors affected by direct range restriction were the cognitive composite scores used for commissioning and aircrew training classification. The personality facet scores were affected only by indirect range restriction. Additionally, the correlations between the cognitive composites and the personality facets were weak, meaning that the indirect range restriction should have had minimal impact. This speculation is further supported by examining the changes in correlations after correcting for range restriction between predictors and criteria (see Table 4).

Table 4: Uncorrected and Corrected Correlations

Pilots				ABMs			
Variable	Uncorrected Correlation	Corrected Correlation	Difference	Variable	Uncorrected Correlation	Corrected Correlation	Difference
Pilot Comp	.15	.23	.08	ABM Comp	.29	.36	.07
F01	.04	.02	-.02	F01	.05	.04	-.01
F02	-.05	-.04	.01	F02	-.01	-.04	-.03
F03	.00	.01	.01	F03	-.02	-.05	-.04
F04	.05	.02	-.03	F04	.05	.04	-.01
F05	.06	.03	-.02	F05	.03	.03	.00
F06	.02	.05	.03	F06	-.05	-.05	.00
F07	.00	-.02	-.02	F07	-.04	-.01	.03
F08	-.05	-.03	.02	F08	-.01	-.03	-.02
F09	.03	.00	-.02	F09	.02	.06	.03
F10	.04	.01	-.03	F10	.03	.02	-.02
F11	-.03	-.03	.00	F11	.01	-.04	-.04
F12	.06	.09	.02	F12	-.05	-.04	.01
F13	-.01	-.04	-.03	F13	.00	.05	.05
F14	.05	.03	-.02	F14	-.06	-.02	.04
F15	.02	.08	.06	F15	.02	.04	.03
F16	.05	.01	-.05	F16	.08	.06	-.02
F17	.05	.07	.02	F17	-.09	.03	.13
F18	-.01	.03	.04	F18	-.01	-.03	-.02
F19	-.07	-.04	.02	F19	.02	-.01	-.03
F20	-.04	-.02	.02	F20	-.04	-.09	-.05
F23	-.12	-.14	-.02	F23	.01	-.04	-.05
F24	-.05	-.03	.02	F24	.03	.00	-.03
F25	.00	-.02	-.02	F25	.02	.02	.00
F26	-.02	.02	.05	F26	.01	-.02	-.02
F28	.04	.02	-.02	F28	.08	.07	-.01
F29	.00	.00	.00	F29	.03	.03	-.01

CSOs				RPA Pilots			
Variable	Uncorrected Correlation	Corrected Correlation	Difference	Variable	Uncorrected Correlation	Corrected Correlation	Difference
CSO Comp	.21	.27	.06	RPA Comp	.31	.44	.13
F01	-.03	-.04	.00	F01	-.07	-.08	-.01
F02	-.01	-.03	-.02	F02	.00	-.04	-.04
F03	.07	.09	.02	F03	.02	.01	-.01
F04	.00	.01	.01	F04	.00	-.03	-.03
F05	-.08	-.08	.00	F05	-.06	-.07	-.01
F06	.04	.03	-.01	F06	-.04	.02	.06
F07	-.07	-.10	-.02	F07	-.06	-.06	.00
F08	.05	.05	.00	F08	.02	.03	.01
F09	-.05	-.05	-.01	F09	-.09	-.05	.04
F10	-.07	-.07	.00	F10	-.06	-.08	-.02
F11	-.07	-.07	.00	F11	-.20	-.19	.01
F12	.07	.06	-.01	F12	-.08	-.03	.04
F13	-.03	-.05	-.01	F13	-.07	-.10	-.03
F14	-.04	-.07	-.03	F14	-.08	-.06	.01
F15	-.05	.00	.04	F15	.01	.09	.08
F16	-.01	-.01	.01	F16	-.01	-.04	-.04
F17	.03	.08	.05	F17	-.04	.01	.05
F18	.09	.10	.01	F18	.06	.14	.08
F19	.04	.04	.01	F19	.03	.03	.00
F20	.01	.01	.00	F20	.05	.05	.00
F23	-.11	-.11	.00	F23	-.18	-.23	-.05
F24	.01	-.02	-.03	F24	-.10	-.08	.03
F25	-.04	-.05	-.01	F25	-.06	-.06	.00
F26	.07	.07	.00	F26	.02	.07	.05
F28	-.04	-.04	.01	F28	-.05	-.02	.04
F29	-.03	-.03	.00	F29	-.09	-.08	.01

Note. Displayed correlations are with each sample's main MASS criterion.
Pilot N = 1,187; CSO N = 658; ABM N = 267; RPA Pilot N = 719.

3.3.2 Linear Multiple Regression with Range Restriction Corrections

Correcting for range restriction is an important step designed to ensure the observation of accurate validity coefficients. Therefore, we performed Lawley's (1943) multivariate correction for range restriction (Carretta & Ree, 2022) and ran linear MR models to find the optimal alternative composites. When creating regression equations, we entered all the predictors in the model at the same time and looked for the statistical significance of each predictor to determine which ones would matter the most. Limitations of this method include the inability to assess non-linear relationships.

3.3.3 Corrected Pareto Optimization with Range Restriction Corrections

PO is a newer statistical technique which can help mitigate the diversity-validity dilemma (De Corte et al., 2011; 2022). The goals for the new composites were to maintain their validities while also reducing mean score SGDs. Using PO would allow us to generate regression weights that would optimize achievement of both these objectives. Multivariate correction for range restriction was performed prior to running the PO analyses. Limitations of PO include the inability to assess non-linear relationships and all predictors need to positively predict the outcome. To circumvent the latter limitation, we reversed the signs of any negative correlations so that they might be utilized.

4.0 RESULTS

The following sections detail the results of the analyses. Sections 3.1, 3.2, and 3.3 detail the results of each of the three methods utilized. In section 3.4, the highest-performing models were compared and contrasted with one another, and the best-performing models were identified for each sample. In section 3.5, we discuss the results of a subject matter expert (SME) review by manned aircraft pilots which led to a change in the final manned aircraft pilot model. The change is implemented, and the new final pilot model, as well as all other final models, are discussed in Section 3.6.

4.1 Non-Linear Multiple Regression Results

As previously described, if quadratic components were found to be statistically significant, both the quadratic and linear components were kept in the model. If the quadratic component was not statistically significant, it was not included in any analyses, but the linear component was kept in the model. Initial models included all possible SDI-O facet scores (26), as well as the current aircrew cognitive composite for each career field (pilot, RPA, CSO, or ABM). Each successive model dropped any predictors that were not statistically significant (with the exception of linear components when the quadratic term was significant). This process continued until only statistically significant predictors were retained in the model. No range restriction corrections were performed on these models, therefore the amount of variance explained by the cognitive composite score is likely underestimated, and the increase in variance associated with the addition of personality facets is likely overestimated.

For manned aircraft pilots, the cognitive composite alone explained 2.2% ($p < .001$) of the variance in the SUPT Primary MASS score. The final model that included the pilot cognitive composite *and* the personality facets explained 5.8% of the variance in the criterion ($p < .001$). No quadratic components were included in the final model.

For CSOs, the cognitive composite alone explained 4.3% ($p < .001$) of the variance in the CSO Primary MASS score. The final model that included the CSO cognitive composite *and* the SDI-O personality facets explained 7.4% of the variance in this criterion ($p < .001$). In the final model, two quadratic components were included.

For ABMs, the cognitive composite alone explained 8.3% ($p < .001$) of the variance in the ABM MASS score. The final model that included the ABM cognitive composite *and* the SDI-O personality facets explained 13.3% of the variance in this criterion ($p < .001$). In the final ABM model, two quadratic components were included.

For RPA pilots, the Pilot cognitive composite alone explained 9.67% ($p < .001$) of the variance in the RPA IFT MASS score. However, when testing for non-linear relationships, we found that the cognitive composite has a positive quadratic relationship with the RPA IFT MASS score, such that those with an above average cognitive score perform better, but for those scoring below average, lower cognitive scores do not influence performance (see Figure1). Therefore, we included the quadratic components of the AFOQT Pilot composite in these models. Inclusion of the quadratic component increased explained variance to 10.9% ($p < .001$). The final RPA model which included the pilot cognitive composite *and* the personality facets explained 16.4% of the variance ($p < .001$). The SDI-O personality facets had no significant non-linear components.

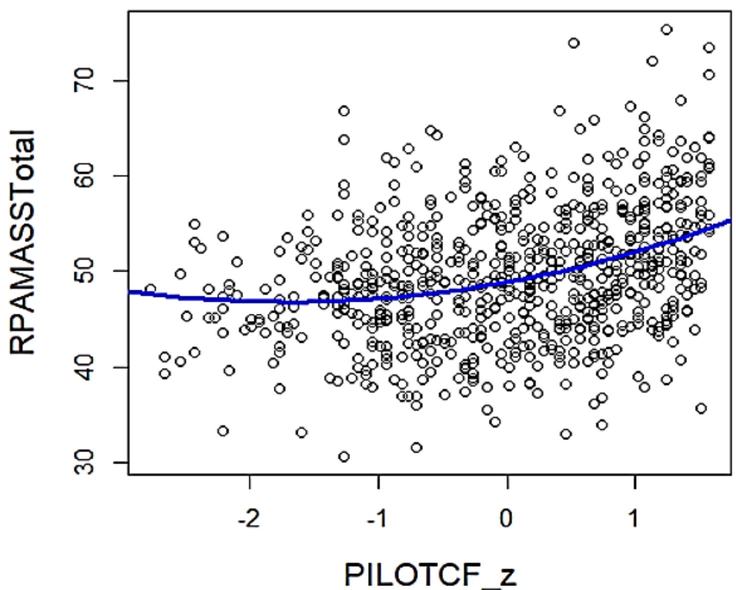

Figure 1. Non-Linear Relationship between Pilot Cognitive Composite and RPA MASS.
Note. PILOTCF_z = the cognitive pilot composite score, transformed into a z-score.
RPAMASSTotal = RPA Pilot Main Criterion.

4.2 Linear Regression Results

Initial models included all possible SDI-O facet scores and the AFOQT cognitive composites. Each successive model dropped any predictors that were not statistically significant. This process was iterative and continued until only statistically significant predictors were included in the model. In this report, only the best performing models are included. Therefore, sometimes models will retain nonsignificant predictors. For linear MR, we used data which had been corrected for range restriction using Lawley's (1943) multivariate formula. Therefore, variance estimates will be more accurate than in the non-linear MR analyses.

For pilots, the cognitive Pilot composite alone explained 5.2% ($p < .001$) of the variance in the SUPT Primary MASS score. The final model that included the pilot cognitive composite *and* the personality facets explained 8.2% of the variance in this criterion ($p < .001$).

For CSOs, the cognitive CSO composite alone explained 7% ($p < .001$) of the variance in the CSO Primary MASS score. The final model that included CSO cognitive composite *and* the SDI-O personality facets explained 8.2% of the variance in this criterion ($p < .001$).

For ABMs, the cognitive ABM composite alone explained 12.6% ($p < .001$) of the variance in the ABM MASS score. The final model explained 12.8% of the variance in the criterion ($p < .001$).

For RPA pilots, the AFOQT cognitive Pilot composite alone explained 19.7% of the variance in the main criterion (RPA IFT MASS; $p < .001$). The final model explained 24.1% of the variance in this criterion ($p < .001$).

4.3 Pareto Optimization Results

Using the PO method, we attempted to maximize validity while minimizing mean score SGDs. We calculated two PO models: the first looking at male versus female mean score SGDs, and the second looking at racial majority (i.e., White and Non-Hispanic) versus racial minority (i.e., underrepresented races; those who are not in the racial majority) mean score SGDs. There was one exception: the applicant gender ratio was lower than the ABM gender ratio, therefore no gender model was run for ABMs. The first set of these models included all potential SDI-O facets and then analyses were re-run using only the facets included in the first set of models.

We examined all potential models and selected the ones that (1) did not reduce validity provided by the cognitive composites alone and (2) provided the highest possible adverse impact (AI) ratio (i.e., the lowest possible mean score SGDs). Therefore, all PO models do not change validity in comparison to the cognitive composites alone.

4.4 Model Testing

Next, we compared the highest-performing models identified from each of the three methods against one another. Models were tested several ways. First, validity coefficients were produced. These coefficients were produced for the main criterion for each sample, as well as for any alternative criteria previously described in Section 2.2.1. Next, the effect sizes for the mean score SGDs were computed. *These e*ffect sizes, expressed as Cohe*n*'s *d*, were produced for both gender (male vs female) and race (majority vs minority). Further, these SGDs were calculated in both the incumbent samples and in the applicant sample across all models. See Table 13 through Table 16 for all model comparisons for manned aircraft pilots, CSOs, ABMs, and RPA pilots.

In some cases, the best models were clear. In other cases, they were not. For manned aircraft pilots (Table 13), the linear MR model emerged as the best performer. This is because of the interest in predicting the SUPT Primary MASS criterion (as opposed to the alternative criteria of Pilot IFT MASS and SUPT Advanced MASS), as well as mostly lower SGDs. For CSOs and ABMs, the best performing models were clearly the non-linear regression models. These had the strongest validity coefficients and lower or lowest SGDs in comparison to all other models. For RPA pilots, the non-linear MR model showed the highest validity coefficient and lowest SGDs for gender. However, the non-linear MR model had slightly higher racial mean score SGDs ($d = .56$) compared to the cognitive composite alone ($d = .53$) and PO ($d = .53$) models. Notwithstanding, we observed improved SGDs across the new models and higher validity coefficients compared to the existing cognitive composites.

Table 5: Model Comparisons for Manned Aircraft Pilot

Model	**IFT MASS**	**Main Criterion SUPT Primary**	**SUPT Advanced**	**Pilot Sample**		**Applicant Sample**	
	R^2	R^2	R^2	MinMaj *d*	Gender *d*	MinMaj *d*	Gender *d*
AFOQT Pilot Composite	13.7	2.2	4.45	.33	.80	.53	.87
Nonlinear MR	4.95	6.31	2.29	.27	.91	.45	.92
Linear MR	5.89	5.95	2.74	.25	.82	.44	.87
PO	13.4	2.3	4.44	.32	.80	.53	.87

Note. Pilot sample $N = 1,187$; Applicant sample $N = 46,440$.
Nonlinear MR = the top performing NLMR model; Linear MR = the top performing corrected linear MR model; PO = the top performing PO model; MinMaj = the majority vs minority Racial subgroups; *d* = Cohen's *d*; IFT = Initial Flight Training; SUPT = Supervised Undergraduate Pilot Training

Table 6: Model Comparisons for CSO

Model	**IFT MASS**	**Main Criterion CSO Primary**	**CSO Advanced**	**CSO Sample**		**Applicant Sample**	
	R^2	R^2	R^2	MinMaj *d*	Gender *d*	MinMaj *d*	Gender *d*
AFOQT CSO Composite	2.79	4.33	5.51	.30	.16	.45	.33
Nonlinear MR	6.41	8.09	6.82	.37	.07	.47	.17
Linear MR	4.62	7.31	4.36	.49	.45	.54	.57
PO	2.85	4.41	5.47	.30	.16	.44	.33

Note. CSO Sample $N = 658$; Applicant Sample $N = 46,440$.
Nonlinear MR = the top performing nonlinear MR model; Linear MR = the top performing corrected linear MR model; PO = the top performing PO model; MinMaj = the majority vs minority Racial subgroups; *d* = Cohen's *d*; IFT = Initial Flight Training; CSO = Combat Systems Officer.

Table 7: Model Comparisons for ABM

Model	Main Criterion	ABM Sample		Applicant Sample	
	R^2	MinMaj *d*	Gender *d*	MinMaj *d*	Gender *d*
AFOQT ABM Composite	8.31	.31	.37	.49	.59
Nonlinear MR	15.3	.27	.09	.27	.26
Linear MR	8.74	.30	.51	.44	.52
PO	8.38	.31	.37	.49	.58

Note. ABM Sample $N = 267$; Applicant Sample $N = 46{,}440$.
Nonlinear MR = the top performing NLMR model; Linear MR = the top performing corrected linear MR model; PO = the top performing PO model; MinMaj = the majority vs minority Racial subgroups; *d* = Cohen's *d*; ABM = Air Battle Manager.

Table 8: Model Comparisons for RPA

Model	Main Criterion	RPA Pilot Sample		Applicant Sample	
	R^2	MinMaj *d*	Gender *d*	MinMaj *d*	Gender *d*
AFOQT Pilot Composite	9.67	.29	.58	.53	.87
Nonlinear MR	17.2	.40	.51	.56	.67
Linear MR	15.7	.43	.55	.61	.79
PO	9.66	.29	.58	.53	.87

Note. RPA Pilot Sample $N = 267$; Applicant Sample $N = 46{,}440$.
Nonlinear MR = the top performing MLMR model; Linear MR = the top performing corrected linear MR model; PO = the top performing PO model; MinMaj = the majority vs minority Racial subgroups; *d* = Cohen's *d*; RPA = Remotely Piloted Aircraft.

4.5 Subject Matter Expert (SME) Review

Upon reviewing the highest-performing models for manned aircraft pilots, CSOs, ABMs, and RPA pilots, there was some concern for the relationships found within the Pilot model. Therefore, these relationships were subjected to a short pilot SME review (N = 3) to establish potential reasoning for the relationships between the various personality facets and training outcomes for pilots.

The pilot SMEs provided adequate reasoning for every relationship found, except for one facet. Because of these findings, this facet was removed from the pilot model (see Section 3.6).

4.6 Final Models

As discussed in Section 5, pilot SMEs were concerned about a particular facet included in the Pilot model. Therefore, we went back to the highest-performing pilot training model (i.e., the corrected linear MR model) and dropped that facet as a predictor. The resulting model explained 4.92% of the variance in IFT MASS, 8.69% of the variance in SUPT Primary MASS, and 3.69% of the variance in SUPT Advanced MASS. Subgroup differences in the pilot sample remained the same for racial subgroups (d = .33) when compared to the AFOQT cognitive Pilot composite alone and decreased for the gender subgroup (d =.75). Finally, SGDs decreased as compared to the AFOQT Pilot composite for both racial (d =.52) and gender (d =.86) subgroups in the applicant sample.

The final model for the other three groups were all the non-linear regression models. As previously mentioned, the RPA pilot model will not be used in the upcoming AFOQT Form T Version 2 implementation. Table 19 provides a summary of comparisons between the best performing models and the AFOQT cognitive composites.

Table 9: Summary of Model Comparisons

Sample	Differences in Validity for Main Criterion	Differences in SGD Effects Size (*d*)	
	ΔR^2	Racial/Ethnic Minorities	Gender Minority
Pilots	+6.49%	+.01	+.01
CSOs	+3.76%	-.02	+.16
ABMs	+6.99%	+.22	+.33
RPA Pilots	+7.53%	-.03	+.20

Note. Pilot N = 1,187, CSO N = 658, ABM N = 267, RPA Pilot N = 719, Applicant N = 46,440. CSO's and RPA Pilots had trivial increases in SGDs for racial/ethnic minorities.

When comparing the new Pilot PSM to the current AFOQT cognitive-only composite, we found a 6.49% *increase* in criterion-related validity for the SUPT Primary MASS score, a .76% *decrease* in criterion-related validity for the SUPT Advanced MASS score, and an 8.78% *decrease* in criterion-related validity for IFT MASS score. When examining the qualification rates in the applicant sample (N = 46,440; using the current minimum qualifying score of the 25th percentile), there were no changes in qualification rates for any subgroups. Ultimately the new Pilot PSM provided improvements only for predicting the SUPT Primary MASS score.

When comparing the new CSO PSM to the current AFOQT cognitive-only composite, we found a 3.76% *increase* in criterion-related validity for the CSO Primary MASS score, a 1.31% *increase* in criterion-related validity for the CSO Advanced MASS score, and a 3.62% *increase* in criterion-related validity for the IFT MASS score. When examining the qualification rates in the applicant sample (N = 46,440; using the current minimum qualifying score of the 25th percentile), 4% more women qualified for CSO training, and 5% more women of a racial/ethnic minority obtained passing scores.

When comparing the new ABM PSM to the current AFOQT cognitive-only composite, we found a 7.0% *increase* in criterion-related validity for the ABM MASS score. When examining the training qualification rates in the applicant sample (N = 46,440; using the current minimum qualifying score of the 25th percentile), 5% more racial/ethnic minorities obtained a passing score, 8% more women obtained passing scores, and 14% more women of a racial/ethnic minority obtained passing scores.

When comparing the new RPA pilot PSM to the current AFOQT pilot cognitive-only composite, we found a 7.53% *increase* in criterion-related validity for the RPA MASS score. When examining the training qualification rates in the applicant sample (N = 46,440; using the current minimum qualifying score of the 25th percentile), 6.7% more women obtained passing scores, and 5.7% more women of a racial/ethnic minority obtained passing scores.

5.0 DISCUSSION

The purpose of the current research was to create new AFOQT PSMs for rated career fields (i.e., Pilots, CSOs, and ABMs) to be used on the new AFOQT Form T Version 2 and to examine PSMs for future use for RPA training. The new PSMs contained the same existing composites but also added relevant SDI-O personality facets. The new PSMs reflect a compensatory approach to selection and classification. Some individuals may be a good personality fit for a specific career field, but their cognitive composite scores are not high enough. Using a compensatory method, these individuals could still qualify for a particular career by using their higher scores on relevant personality facets to offset their low cognitive composite scores.

New PSMs were created for manned aircraft pilots, CSOs, ABMs and RPA pilots (for potential future use). The first three models are recommended for the operational use in the AFOQT Form T Version 2. To clarify, the intent of the new PSMs is not to replace the existing cognitive-only composites, but to compute additional scores to examine if more applicants can classify for the rated career fields. In other words, test takers may obtain a passing score on either the current cognitive-only composites or on the new combination cognitive and personality composites.

The new PSMs provide an increase in criterion-related validity when compared to the cognitive-only counterparts, and either maintain or decrease the SGDs for women and racial/ethnic minorities. These new PSMs should be revalidated again in the future as additional criterion data become available. Additionally, they can provide as foundational guidance for any future changes to the PSMs used on the AFOQT.

6.0 REFERENCES

Aguilar, I. D. (2017). *Comparison of the Air Force Officer Qualifying Test Form T and Form S: Initial item-and subtest-level analyses*. AFCAPS-TR-2017-0002. Randolph AFB, TX: Air Force Personnel Center, Strategic Research and Assessment Branch.

Arias, V. B., Garrido, L. E., Jenaro, C., Martínez-Molina, A., & Arias, B. (2020). A little garbage in, lots of garbage out: Assessing the impact of careless responding in personality survey data. *Behavior Research Methods, 52*(6), 2489-2505. https://doi.org/10.3758/s13428-020-01401-8

Barron, L. G. (2013). *Air Force Officer Qualifying Test (AFOQT) Situational Judgment Test (SJT) development*, AFCAPS-TR-2013-0005. Randolph AFB, TX: Air Force Personnel Center, Strategic Research and Assessment Branch.

Bowling, N. A., Gibson, A. M., Houpt, J. W., & Brower, C. K. (2021). Will the questions ever end? Person-level increases in careless responding during questionnaire completion. *Organizational Research Methods, 24*(4), 718-738. https://doi.org/10.1177/10944281209477

Benson, M. J., & Campbell, J. P. (2007). To be, or not to be, linear: An expanded representation of personality and its relationship to leadership performance. *International Journal of Selection and Assessment, 15*(2), 232-249.

Carretta, T. R., & Ree, M. J. (2022). Correction for range restriction: Lessons from 20 research scenarios. *Military Psychology*, 1-19.

Carretta, T. R., Rose, M. R., & Trent, J. D. (2016). *Air Force Officer Qualifying Test Form T: Initial item-, test-, factor-, and composite-level analyses*, AFRH-RH-WP-TR-2016-0093. Wright-Patterson AFB, OH: Air Force Research Laboratory, 711 Human Performance Wing, Warfighter Interface Division.

De Corte, W., Lievens, F., & Sackett, P. R. (2007). Combining predictors to achieve optimal trade-offs between selection quality and adverse impact. *Journal of Applied Psychology, 92*(5), 1380.

De Corte, W., Lievens, F., & Sackett, P. R. (2022). A comprehensive examination of the cross-validity of pareto-optimal versus fixed-weight selection systems in the bi-objective selection context. *Journal of Applied Psychology*, *107*(8), 1243.

DeSimone, J. A., DeSimone, A. J., Harms, P. D., & Wood, D. (2018). The differential impacts of two forms of insufficient effort responding. *Applied Psychology*, *67*(2), 309-338. https://doi.org/10.1111/apps.12117

Drasgow, F., Nye, C. D., Carretta, T. R., & Ree, M. J. (2010). Factor structure of the Air Force Officer Qualifying Test Form S: Analysis and comparison with previous forms. *Military Psychology, 22*(1), 68-85.

Huang, J. L., Curran, P. G., Keeney, J., Poposki, E. M., & DeShon, R. P. (2012). Detecting and deterring insufficient effort responding to surveys. *Journal of Business and Psychology*, 27, 99–114. http://dx.doi.org/10.1007/s10869-011-9231-8.

Kantrowitz, T., Kingry, D., Engelsted, J., Travinin, G., Lovering, E., & Gould, M. (2022). *Air Force Officer Qualifying Test (AFOQT) Form T evaluation: Job analysis linkages, validity, and subgroup differences for current and alternative composites,* AFRL-RH-WP-TR-2022-0038. Wright-Patterson AFB, OH: 711 Human Performance Wing, Airman Biosciences Division, Performance Optimization Branch.

Kantrowitz, T., Kingry, D., Segall, D., & Valone, A. (in press). *Development of Air Force Officer Qualifying Test (AFOQT) Form T, version 2,* AFRL-RH-WP-TR-2023-xxxx. Wright-Patterson AFB, OH: 711 Human Performance Wing, Airman Biosciences Division, Performance Optimization Branch.

Lawley, D. N. (1943). A note on Karl Pearson's selection formulae. *Proceedings of the Royal Society of Edinburgh, Section A, 62(*1), 28–30. https://doi.org/10.1017/S0080454100006385

Manley, G. G., & Weissmuller, J. (2017, April). *Development of the United States Air Force's Self-Description Inventory* [Poster]. Society for Industrial Organizational Psychology Annual Conference, Orlando, FL.

Mann, K. J., Woolley, M. R., LeRoy, V., Carretta, T. R., Mouton, A. N., & Deregla, A. R. (in press). *Self-Description Inventory Plus (SDI+): Item-, facet-, and domain-level analyses,* AFRL-RH-WP-TR-2022-xxxx. Wright-Patterson AFB, OH: 711 Human Performance Wing, Airman Biosciences Division, Performance Optimization Branch.

Rupp, D. E., Song, C., & Strah, N. (2020). Addressing the so-called validity-diversity trade-off: Exploring the practicalities and legal defensibility of Pareto-optimization for reducing adverse impact within personnel selection. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 13,* 246-271.

Sizemore, S.J., Wilson, R. Mann, K. J., & Carretta, T. R. (2022). *Air Force Officer Qualifying Test (AFOQT) Form T evaluation: Situational Judgment Test (SJT) revision*, AFRL-RH-WP-TR-2022-0076. Wright-Patterson AFB, OH: 711 Human Performance Wing, Airman Biosciences Division, Performance Optimization Branch.

Walsh, J. L., Brady, M. F., Woolley, M. R., & Carretta, T. R. (2022). *Air Force Officer Qualifying Test (AFOQT) Form T evaluation: Item-level analyses*, AFRL-RH-WP-TR-2022-0037. Wright-Patterson AFB, OH: 711 Human Performance Wing, Airman Biosciences Division, Performance Optimization Branch.

Walsh, J. L., Woolley, M. R., Brady, M. F., & Carretta, T. R. (2022). *Air Force Officer Qualifying Test (AFOQT) Form T evaluation: Subtest-level analyses*, AFRL-RH-WP-TR-2022-0022. Wright-Patterson AFB, OH: 711 Human Performance Wing, Airman Biosciences Division, Performance Optimization Branch.

Walsh, J. L., Woolley, M. R., Brady, M. F., Melick, S. R., & Carretta, T. R. (2022). *Air Force Officer Qualifying Test (AFOQT) Form T evaluation: The Situational Judgment Test,* AFRL-RH-WP-TR-2021-0080. Wright-Patterson AFB, OH: 711 Human Performance Wing, Airman Biosciences Division, Performance Optimization Branch.

Woolley, M. R., Walsh, J. L., Mann, K. J., Wilson, R. T., Carretta, T. R., Mouton, A. N., & Deregla, A. R. (2022). *Self-Description Inventory – Officer (SDI-O): Item-, facet-, and domain-level analyses,* AFRL-RH-WP-TR-2022-0091. Wright-Patterson AFB, OH: 711 Human Performance Wing, Airman Biosciences Division, Performance Optimization Branch.

LIST OF SYMBOLS, ABBREVIATIONS, AND ACRONYMS

<	Less Than
Δ	Change
%	Percent
ABM	Air Battle Manager
AECP	Airman Education and Commissioning Program
AFB	Air Force Base
AFPC	Air Force Personnel Center
AFOQT	Air Force Officer Qualifying Test
AI	Adverse Impact
ANG	Air National Guard
b	Uncorrected Regression (Beta) Weight
Comp	Composite
CSO	Combat Systems Operator
d	Cohen's *d* (effect size)
ID	Identification
IFT	Initial Flight Training
IRV	Intra-Individual Response Variability
MASS	Merit Assignment Selection System
MinMaj	Minority Majority
MR	Multiple Regression
N	Sample Size
NLMR	Non-Linear Multiple Regression
OTS	Officer Training School
p	Probability
PO	Pareto Optimization
PSM	Predictive Success Model
R	Multiple Correlation
ROTC	Reserve Officer Training Corps
SD	Standard Deviation
SDI+	Self-Description Inventory+ (AFOQT Form S)
SDI-O	Self-Description Inventory – Officer (AFOQT Form T)
SGD	Subgroup Differences
SJT	Situational Judgment Test
SME	Subject Matter Expert
SPARK	Strategic Policy Analysis Resource and Knowledgebase
SUPT	Specialized Undergraduate Pilot Training
RPA	Remotely Piloted Aircraft
USAF	United States Air Force
USAFA	United States Air Force Academy