



AFRL-RH-WP-TR-2023-0009

PREDICTIVE VALIDITY OF THE OFFICER SUITABILITY MEASURE (OSM) AND FOUR ALTERNATIVE OFFICERSHIP COMPOSITES

Montana R. Woolley
Infoscitex, a DCS company

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

Amanda N. Mouton
Andrew R. Dereglia
Air Force Personnel Center
Strategic Research and Assessments Branch
JBSA-Randolph AFB, TH

March 2023
INTERIM REPORT

DISTRIBUTION A. Approved for public release.

**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

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the Air Force Research Laboratory Public Affairs Office and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (<http://www.dtic.mil>).

AFRL-RH-WP-TR-2023-0009 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION.

//signature//

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

//signature//

LOGAN A. WILLIAMS, PhD
Human Performance Product Area Lead
Product Development
Performance Optimization Branch
Airman Biosciences Division

This report is published in the interest of scientific and technical information. And its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE			Form OMB No. 0704-0188		Approved
The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.					
1. REPORT DATE (DD-MM-YY) 21-02-23		2. REPORT TYPE Interim		3. DATES COVERED (From - To) 10/2022 – 12/2022	
4. TITLE AND SUBTITLE Predictive Validity of the Officer Suitability Measure (OSM) and Four Alternative Officership Composites				5a. CONTRACT NUMBER FA8650-21-F-4104	
				5b. GRANT NUMBER	
				5c. PROGRAM ELEMENT NUMBER	
6. AUTHOR(S) Montana R. Woolley ^a , Thomas R. Carretta ^b , Amanda N. Mouton ^c , and Andrew R. Deregl ^a				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 711 th Human Performance Wing Airman Systems Directorate Airman Biosciences Division Performance Optimization Branch Wright-Patterson AFB, OH 45433				10. SPONSORING/MONITORING AGENCY ACRONYM(S) 711 HPW/RHBC	
				11.SPONSORING/MONITORING AGENCY REPORT NUMBER(S) AFRL-RH-WP-TR-2023-0009	
12. DISTRIBUTION/AVAILABILITY STATEMENT Distribution Statement A: Approved for public release.					
13. SUPPLEMENTARY NOTES Report contains color. AFRL-2023-1557, cleared 4 April 2023					
14. ABSTRACT The Officer Suitability Measure (OSM), which is a composite of facets from the Self-Description Inventory (SDI), was developed by Barron (2013, 2014) as an indicator of officership potential. This report details the revalidation of the OSM using three different Air Force samples (one from Officer Training School [OTS] and two Reserved Officer Training Corps [ROTC] samples). We also explored four alternative composites developed using the same methodology from Barron (2013) using the OTS data. These composites were developed to predict overall success in OTS, leadership performance, academic performance, and the combination of these three criteria. These SDI-based composites were cross-validated in the ROTC samples. Results indicated the OSM was a moderate predictor of leadership ability in OTS and provided incremental validity above and beyond the Air Force Officer Qualifying Test (AFOQT) Academic Aptitude cognitive composite. Although the alternative personality-based composites were stronger predictors of specific OTS training criteria, the OSM score remains the best predictor of success in ROTC training					
15. KEY WORDS Air Force, officer suitability measure, officership, validation					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT: SAR	18. NUMBER OF PAGES 22	19a. NAME OF RESPONSIBLE PERSON (Monitor) Thomas R. Carretta
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER (Include Area Code)

TABLE OF CONTENTS

EXECUTIVE SUMMARY	ii
1.0 INTRODUCTION.....	1
2.0 METHOD	4
2.1 Sample 1.....	4
2.2 Samples 2 and 3	7
2.3 Technical Approach.....	9
2.3.1. Step 1: Revalidating OSM in Sample 1	9
2.3.2. Step 2: Creating New SDI-O Composites in Sample 1	9
2.3.3. Step 3: Testing New Composites in Sample 1	9
2.3.4. Step 4: Cross-Validation in Sample 2.....	10
2.3.5. Step 5: Cross-Validation in Sample 3.....	10
3.0 RESULTS	10
3.1 Step 1: Results.....	10
3.2 Step 2: Results.....	10
3.3 Step 3: Results.....	11
3.4 Step 4: Results.....	12
3.5 Step 5: Results.....	14
4.0 DISCUSSION	15
5.0 REFERENCES.....	16
LIST OF SYMBOLS, ABBREVIATIONS, AND ACRONYMS	17

LIST OF TABLES

Table 1. SDI+ Facet to Domain Linkages	2
Table 2. SDI-O Facet to Domain Linkages	3
Table 3. OTS Sample Demographics.....	4
Table 4. OTS Criteria Intercorrelations	6
Table 5. Sample 2 Criteria Intercorrelations.....	8
Table 6. Sample 3 Criteria Intercorrelations.....	8
Table 7. OSM, New Composites, and OTS Criteria Bivariate Correlations	11
Table 8. Variance Explained by OSM and Alternative Composites	12
Table 9. Sample 2 (ROTC) Bivariate Correlations.....	13
Table 10. Sample 2 (ROTC) Explained Variance in Field Training Rank.....	13
Table 11. Sample 3 (ROTC) Bivariate Correlations.....	14
Table 12: Explained Variance in Sample 3 (ROTC) Criteria.....	14

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 Officer Suitability Measure (OSM) is an indicator of adaptability and leadership potential. It was derived from personality facets measured by the Self-Description Inventory Plus (SDI+; Barron, 2013), a personality measure initially included in the Air Force Officer Qualifying Test (AFOQT) Form S in 2005. Twenty SDI+ facet measures were developed six years later as scores on the SDI+ (Manley, 2011). Barron used these facet scores to develop and validate the OSM against several Officer Training School (OTS) criteria including Basic Officer Training (BOT) completion and distinguished graduate status. Later, an updated personality inventory named the Self-Description Inventory for Officers (SDI-O) was developed and included with the rollout of the AFOQT Form T in 2014 (Manley & Weissmuller, 2017). Although there is overlap in content between the SDI+ and SDI-O, there were changes to items and facets with introduction of the new form. With the change from SDI+ to SDI-O, Barron (2014) developed an equitable OSM score which could be produced from items on the SDI-O.

The current study utilized three samples to revalidate the OSM score and examine the utility of four alternative personality composites in predicting several officer training criteria. The first sample consisted primarily of OTS students. It was used to evaluate the utility of the current OSM composite and evaluate the utility of four new composites designed to predict officer training criteria (i.e., Overall Success Predictor, Leadership Predictor, Academic Predictor, and Combined Predictor). The other two samples consisted of Reserve Officer Training Corps (ROTC) students and were used to cross-validate both the OSM, and the four new composites developed using the OTS sample.

The Overall Success Predictor was meant to predict overall success in OTS and was validated against the Distinguished Graduate Status criterion. The Leadership Predictor was used to predict leadership performance and was validated against the Leadership Criterion (i.e., a unit-weighted combination of Tactical Graded Leadership Position [GLP1], Leadership Reaction Course [LRC], Peer Ranking, and Instructor Ranking). The Academic Predictor was meant to predict academic performance, and therefore was validated against the Academic Criterion (i.e., a unit-weighted composite of grades on academic assessments, briefings, and papers). These three predictors are referred to as the *specific predictors*, because they are designed to predict specific outcomes. Finally, all of the facets that contributed to any of the three specific predictors were combined into a single unit-weighted composite (i.e., the Combined Predictor). It was hypothesized that the Combined Predictor should predict all three specific criteria, although likely would not explain as much variance as the specific composites with their linked criterion.

Correlations and linear regression were used to establish the criterion-related validity of the various composites. Analyses indicated that the OSM composite was a moderate predictor of leadership ability in OTS and provided incremental validity over the AFOQT Academic Aptitude cognitive composite. Although the alternative specific composites were stronger predictors of the specific officership criteria in OTS, the OSM composite was the best predictor of success in ROTC training.

1.0 INTRODUCTION

The Officer Suitability Measure (OSM) was developed by Barron (2013) for use as part of the United States Air Force (USAF) OTS officer commissioning process. It was originally based on 20 facets from the Self-Description Inventory + (SDI+; Christal, 1994; Manley 2011) which was included in Form S of the AFOQT beginning in 2005. The OSM composite was revised (Barron, 2014) with the implementation of AFOQT Form T in 2014, which included a revised version of the SDI (Self-Description Inventory for Officers [SDI-O], Manley & Weissmuller, 2017).

The original OSM composite was created by evaluating the relations of SDI+ facets to BOT completion/attrition (see Table 1 for all SDI+ facets). Next, facets that were significantly correlated with BOT completion were unit-weighted and utilized to predict Order of Merit criteria and Final Training Performance Report scores. Findings showed that the OSM provided significant incremental validity over the AFOQT cognitive ability measures (2.4-3.4 percent (%); Barron, 2013). Cross-validation of the OSM was completed using United States Air Force Academy (USAF) Military Performance Averages. Once again, the OSM showed significant incremental validity beyond current AFOQT cognitive selection measures (1.3-1.7%; Barron, 2013). The original OSM composite was comprised of five unit-weighted SDI+ facets. Two facets contributed positively toward the final OSM score (i.e., higher scores on these facets increase OSM score) and three facets contributed negatively toward the final OSM score (i.e., high scores on these facets decrease OSM score). This OSM composite was used in AFOQT Form S until the rollout of the new SDI-O in AFOQT Form T.

Not all facets and items used in the original OSM composite were retained in AFOQT Form T. (see Table 2 for all SDI-O facets). Therefore, a comparable OSM score was created using items from the SDI-O, and then the SDI+ based OSM composite and new SDI-O based composite were equated (Barron, 2014). These new facets were found to be equitable, and therefore make up the AFOQT Form T unit-weighted OSM composite.

Table 1. SDI+ Facet to Domain Linkages

Domain	Facet #	Facet Name
Agreeableness	A1	Team Player
Agreeableness	A2	Pleasant
Agreeableness	A3	Considerate
Agreeableness	A4	Helpful-Altruistic
Agreeableness	A5	Hyper-Competitive
Conscientiousness	C1	Achievement Striving
Conscientiousness	C2	Order
Extraversion	E1	Unassertive
Extraversion	E2	Sociable
Extraversion	E3	Dominance
Machiavellianism	M1	Envious
Machiavellianism	M2	Individualistic
Machiavellianism	M3	Self-serving
Neuroticism	N1	Stress Under Pressure
Neuroticism	N2	Temperamental
Neuroticism	N3	Worry
Openness	O1	Creative
Openness	O2	Reflective
Openness	O3	Scientific Interest
Openness	O4	Cultured

Table 2. SDI-O Facet to Domain Linkages

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

The purpose of the current research was to revalidate the OSM composite used for the AFOQT Form T SDI-O, as well as to explore alternative composites that could be utilized in the future. Based on prior research (McHenry et al., 1990) it is expected that personality-based composites like the OSM will demonstrate the strongest relations to training performance criteria reflecting level of effort, leadership, and discipline and relatively low validity for criteria assessing technical proficiency (e.g., academic grades).

2.0 METHOD

The current research utilizes data from three separate sources. A sample of OTS applicants was the main source for OSM revalidation and for exploring potential alternative composites. Two separate samples of ROTC graduates were utilized to cross-validate the OSM and new composites developed in the OTS sample. For all three sources, officership criteria and AFOQT data were matched using each Airman's SPARK ID (i.e., a unique 9-character alphanumeric code). For all three samples, SDI-O data, as well as AFOQT verbal and quantitative cognitive scores were included. The verbal and quantitative scores used were the test-takers score of record. The SDI-O scores were first-time administration scores. The SDI-O data were cleaned for missingness and carelessness consistent with the methods used by Woolley et al., (2022). In short, all missingness was removed listwise, all participants flagged as careless responders were removed, and all SDI-O data collected prior to August 17, 2016 were removed due to an administrative error. Criteria data were retained, and pairwise deletion was used on a case-by-case basis for any missingness present within criteria scores.

2.1 Sample 1

Sample 1 primarily consisted of OTS trainees. The sample came from four main accession sources: OTS Civilians, OTS Active Duty, Air National Guard (ANG), and AF Reserves. There were also 12 participants from the ROTC. See Table 3 for a breakdown of this sample's demographics. Data were collected from August 2016 to February 2020 (sample size, $N = 1,416$). All demographic data were

Table 3. OTS Sample Demographics

Variable	N	%
Sex		
Male	1139	80
Female	277	20
Unknown	0	0
TOTAL	1416	100%
Race		
Non-Hispanic White	950	67
Non-Hispanic Black	64	5
Non-Hispanic Asian	63	4
Non-Hispanic AIAN	2	0
Non-Hispanic NHPI	4	0
Non-Hispanic Multiracial	133	9
Hispanic, Any Race (besides White Only, Black Only)	55	4
Hispanic Only (No race selected)	29	2
Hispanic White	102	7
Hispanic Black	5	0
Any ethnic / racial minority	463	33
Unknown	3	0
TOTAL	1416	132%¹

Race and Gender Intersectionality		
Non-Hispanic White Men	785	88
Any ethnic / racial / gender minority	111	12
NA	0	0
TOTAL	896	100%
Socio-Economic Status		
Much higher than average	53	4
Somewhat higher than average	318	22
Average	606	43
Somewhat lower than average	286	20
Much lower than average	108	8
Declined to respond	44	3
Unknown	1	0%
TOTAL	1416	100%
Education		
Completed 12	2	0
Completed 13	6	0
Completed 14	54	4
Completed 15	289	20
Completed 16	736	52
Completed 17	152	11
Completed 18	117	8
Completed 19	27	2
Completed 20	13	1
Completed 21+	18	1
Unknown	2	0
TOTAL	1416	100%
Academic Degree		
High School Diploma	140	10
Associates Degree	293	21
Bachelor's Degree	840	59
Master's Degree	128	9
Doctoral Degree	13	1
Unknown	2	0%
TOTAL	1416	100%
Accession Source		
OTS-CIV	405	29
OTS-AD	434	31
AACP	0	0
USAFA	0	0
ROTC	12	1
ANG	426	30
AFRES	139	10
TOTAL	1416	100%
Age		
	Mean	SD
	27.32	4.46
<i>Note.</i> ¹ The total percent for race/ethnicity is greater than 100% because respondents could indicate more than one race. SD = Standard Deviation		

collected during AFOQT administration. This sample was primarily male (80%), and White (67%), with an average age of 27.32 years ($SD = 4.46$). Most of the participants in this sample reported an average or higher socioeconomic status (60%). Seventy-five percent of the sample completed at least 16 years of education.

The OTS sample included a variety of criteria. For an overview of all criteria, see Attachment 1, the OTS Course Syllabus. The best criteria to describe overall success within OTS were Final Course Score, Distinguished Graduate Status, and Top Graduate Status. Additionally, two criterion composites were created.

The first criterion composite was meant to indicate leadership success (referred to as the Leadership Criterion). This criterion is a unit-weighted combination of Tactical GLP1, LRC, Peer Ranking, and Instructor Ranking. All four criteria are meant to assess leadership abilities. Please note that the Operational GLP2 and the Warrior Expeditionary Leadership Problem Solving (WELPS) scores are also meant to assess leadership ability but had too much missingness in this data set to be utilized. Furthermore, GLP2 and WELPS had lower correlations with the other leadership criteria (see Table 4).

Table 4. OTS Criteria Intercorrelations

	AA1	AA2	B1	B2	GLP1	GLP2	WEL PS	P1	P2	LRC	Peer Rank	Inst. Rank	Final Score	Dist. Grad
AA2	<u>.39</u>													
B1	<u>.22</u>	<u>.19</u>												
B2	<u>.18</u>	<u>.14</u>	.38											
GLP1	.10	.10	.17	.18										
GLP2	.15	.14	.18	.17	.17									
WELPS	.12	.14	.14	.09	.15	.16								
P1	<u>.17</u>	<u>.17</u>	<u>.22</u>	<u>.11</u>	.07	.18	.09							
P2	<u>.23</u>	<u>.15</u>	<u>.21</u>	<u>.25</u>	.13	.14	.12	.20						
LRC	.06	.05	.10	.09	.18	.09	.11	.00	.05					
Peer Rank.	.20	.20	.26	.27	.26	.25	.14	.19	.18	.17				
Inst. Rank.	<u>.20</u>	<u>.21</u>	.26	.27	.29	.27	.15	.21	.18	.17	.82			
Final Score	.03	.11	.13	.04	.07	.16	.12	.35	.03	-.03	.11	.12		
Dist. Grad.	<u>.28</u>	<u>.25</u>	<u>.23</u>	<u>.21</u>	<u>.25</u>	<u>.22</u>	.18	<u>.21</u>	<u>.20</u>	.17	.32	.32	.14	
Top Grad.	.06	.05	.04	.03	.07	.08	.06	.05	.06	.05	.08	.08	.04	.16

Note. $N = 1,416$. Yellow color $p < .05$; Light green color $p < .01$; Dark green color $p < .001$.

Bold values indicate intercorrelations among criteria used to create the Leadership Criterion composite.

Underlined values indicate intercorrelations among the criteria used to create the Academic Criterion composite.

The second criterion composite was meant to assess academic success (referred to as Academic Criterion). This criterion is a unit-weighted combination of Academic Assessment 1 (AA1),

Academic Assessment 2 (AA2), Informative Briefing (B1), Advocacy Briefing (B2), Informative Paper (P1) and Advocacy Paper (P2).

For additional descriptive statistics reviewing all predictors and potential criteria (from all samples), see Attachment 2.

2.2 Samples 2 and 3

The second sample was utilized for cross-validation. This sample came from ROTC graduates between the years 2015 and 2017 ($N = 1,675$). The sample was greatly diminished due to only using AFOQT scores after Aug 17, 2016. Criteria from this sample include subjective ratings on Leadership (Lead), Professional Qualities (ProQ), Communication (Comm), Judgement and Decision Making (JD), Warrior Ethos (WE), and Followership (Foll). Each category was scored on an anchored rating form, with rating options including Unsatisfactory, Marginal, Satisfactory, Excellent, and Outstanding. There are corresponding Lead, ProQ, Comm, JD, WE, and Foll scores for Air Expeditionary Force (AEF) training. Additionally, there is a final rank variable which indicates overall success in ROTC. Table 5 presents the Sample 2 criteria intercorrelations.

The third sample was also utilized for cross-validation. It consisted of ROTC graduates between the years 2018 and 2020 ($N = 5,403$). Some of the criteria and rating systems had changed between 2017 and 2018, therefore the two ROTC samples were analyzed separately. The main criterion for this sample is Field Training Rank (FT Rank), which is the overall order of merit ranking. This is a composite combining Preparation (20%), Fitness (10%), Active Learning (10%), Leadership (30%), and Field Training Officer (FTO) rank (30%). Please note that this variable is normally reversed, so that intercorrelations between it and other criteria are negative. In this report, the negative correlations were reversed for ease of understanding. Each of these five categories' scores are made up of several components (see Attachment 3). There were three smaller components which were of particular interest. These include FTO Flight Ranking (part of FTO Rank score), the LRC scores (part of Leadership score) and the Mission Scenario scores (part of Leadership score). Therefore, Mission Scenario and LRC scores both indicate leadership success. See Attachment 3 for an example of the Field Training Performance Report and Attachment 4 for the ROTC Field Training Leadership Evaluation Form. Table 6 provides the Sample 3 criteria intercorrelations.

Table 5. Sample 2 Criteria Intercorrelations

	Lead	ProfQ	Comm	JD	WE	Foll	AEFLead	AEFProfQ	AEFComm	AEFJD	AEFWE	AEFFoll
Professional Qualities	.61											
Communication	.65	.54										
Judgement.Decis	.64	.62	.62									
Warrior.Ethos	.63	.57	.60	.61								
Followership	.51	.52	.52	.55	.56							
AEF Leadership	.48	.41	.40	.41	.42	.33						
AEF Prof. Qualities	.41	.37	.36	.35	.36	.30	.70					
AEF Communication	.44	.37	.39	.38	.40	.32	.76	.69				
AEF Judgement.Decis	.43	.37	.37	.38	.37	.31	.77	.71	.74			
AEF Warrior.Ethos	.43	.35	.38	.36	.39	.29	.71	.64	.68	.68		
AEF Followership	.39	.31	.34	.34	.33	.28	.64	.63	.62	.63	.64	
FT_RANK	.57	.47	.49	.49	.50*	.39 ^{ns}	.63	.57	.60	.58	.57	.51

Note. $N = 1,675$. AEF = Air Expeditionary Force. All correlations were significant at $p < .001$; unless indicated with * $p < .05$; or ns, $p > .05$.

Table 6. Sample 3 Criteria Intercorrelations

Variable	FTO Rank	LRC 1	Scenario 1
LRC1	.31		
Scenario1	.45	.22	
FT Rank	.94	.42	.54

Note. $N = 5,403$. $p < .001$; FTO Rank = Field Training Officer Rank

2.3 Technical Approach

This study goes through multiple steps. We began by revalidating the original OSM composite (Barron, 2013, 2014) with the OTS sample. We also created alternative officership composites using the SDI-O facets and criteria from the OTS sample. The performance of the new composites in the OTS sample were examined and compared to that of the original OSM composite. Finally, we conducted separate cross-validations of the original OSM and new composites in the second (1st ROTC) and third (2nd ROTC) samples.

2.3.1. Step 1: Revalidating OSM in Sample 1

To begin revalidating the OSM composite, correlations between the OSM score and the OTS sample criteria were calculated. Next, we tested if the OSM score provided incremental explained variance above that of the AFOQT Academic Aptitude cognitive composite (composite of the three verbal and two math subtests). This was done using multiple regression and testing for significant additional variance.

2.3.2. Step 2: Creating New SDI-O Composites in Sample 1

After the revalidation steps, we explored the possibility of other SDI-O composites that could be used in place of the OSM composite. This was done following the methodology used by Barron (2013). First, we examined statistically significant correlations between the 30 SDI-O facets and the five OTS criteria (i.e., Final Course Score, Distinguished Graduate Status, Top Graduate Status, Leadership Criterion, and Academic Criterion). Significantly correlated facets were then combined into three different unit-weighted composites. The first new composite, referred to as the Overall Success Predictor, was meant to predict overall success in OTS, and utilized the Distinguished Graduate Status criterion. The second composite, referred to as the Leadership Predictor, was meant to predict leadership performance, and therefore used the Leadership Criterion previously described. The third, referred to as the Academic Predictor, was meant to predict academic success, and therefore was validated against the Academic Criterion previously described. These three composites together will be referred to as the *specific predictors*, meaning that the Overall Success Predictor should predict overall success best (i.e., Distinguished Graduate Status), the Leadership Predictor should predict the Leadership Criteria best, and the Academic Predictor should predict the Academic Criteria best.

Finally, facets from all three of these new SDI-O composites were combined into a single unit-weighted composite, utilizing all facets included in any of the three specific predictors. This composite will be referred to as the Combined Predictor. This should predict all three criteria, although likely will not explain as much variance as the specific predictors with their linked criterion.

2.3.3. Step 3: Testing New Composites in Sample 1

Next, the newly created officership composites were tested. First using simple regression, we examined how much variance was explained by each composite. For the specific composites (Overall Success Predictor, Leadership Predictor, and Academic Predictor), we examined how

much variance they could explain in their linked criteria. For the Combined Predictor, we examined explained variance in all three criteria, then compared this to the OSM composite performance. Additionally, we examined incremental variance explained above and beyond that explained by the AFOQT Academic Aptitude cognitive composite and compared these results to the OSM composite.

2.3.4. Step 4: Cross-Validation in Sample 2

In step 4, we examined the correlations between the five predictors (i.e., the OSM, the three specific predictors, and the Combined Predictor) and the criteria from Sample 2 (1st ROTC sample). We also calculated the variance explained in the FT Rank variable by each composite, as well as the incremental variance explained above and beyond the AFOQT Academic Aptitude composite.

2.3.5. Step 5: Cross-Validation in Sample 3

Finally, we mimicked the procedure in Step 4 using the third sample (2nd ROTC sample). We examined the correlations between the five predictors (i.e., the OSM, the three specific predictors, and the Combined Predictor) and the criteria from Sample 3. We also calculated the variance explained in all criteria variables by each composite, as well as the incremental predictive validity beyond the AFOQT Academic Aptitude composite.

3.0 RESULTS

3.1 Step 1: Results

The OSM composite did not predict overall success in OTS or academic performance. However, it did significantly correlate with the Leadership Criterion ($r = .12, p < .001$). Unsurprisingly, the OSM composite was significantly correlated with all four components of the Leadership Criterion as well: GLP1 ($r = .08, p < .01$), LRC ($r = .08, p < .01$), Peer Ranking ($r = .11, p < .001$) and Instructor Ranking ($r = .11, p < .001$).

The OSM composite explained 1.38% of the variance in the Leadership Criterion ($b = .41, p < .001$), and provided 1.53% incremental variance above and beyond the AFOQT Academic Aptitude cognitive composite ($b_{\text{academic}} = .22, p < .05$; $b_{\text{OSM}} = .43, p < .001$). The OSM composite alone did not explain variance in the other criteria. However, when controlling for the AFOQT Academic Aptitude cognitive composite, the OSM score explained a statistically significant, although trivial, amount of incremental variance in Distinguished Graduate Status ($\Delta R^2 = .003, p < .001$).

3.2 Step 2: Results

To create alternative composites, first we examined correlations between the SDI-O facets and the OTS criteria. There was only one statistically significant correlation for Final Score, and no

significant correlations for Top Graduate Status. There were five significant correlations between the SDI-O facets and Distinguished Graduate Status, therefore we created a unit-weighted composite which included all five facets. This is referred to as the Overall Success Predictor. For the alternative composite meant to predict academic success (i.e., Academic Predictor), we used all five facets which were significantly correlated with the Academic Criterion. There were 11 facets that significantly correlated with the Leadership Criterion. Therefore, we only used the eight facets that were correlated with at least $p < .01$ to develop the alternative composite meant to predict leadership success (i.e., Leadership Predictor). These three composites are the specific predictors.

As previously described, we also combined these three specific predictors into one Combined Predictor. There were no directionality disagreements, therefore we were able to include all 13 facets used in the specific predictors to create the unit-weighted composite.

3.3 Step 3: Results

Next, we tested the performance of the new composites, and compared them to the OSM composite. Final Score in OTS had the strongest correlation with the Combined Predictor ($r = .07$, $p < .05$). Top Graduate Status was not significantly correlated with any of the composites. As expected, Distinguished Graduate Status had the strongest correlation with the Overall Success Predictor and the Academic Criterion had the strongest correlation with the Academic Predictor. However, the Leadership Criterion had equally strong correlations with both the Leadership Predictor and the OSM ($r = .12$, $p < .001$). See Table 7 for all correlations.

Table 7. OSM, New Composites, and OTS Criteria Bivariate Correlations

Criteria	OSM	Overall Success Predictor	Leadership Predictor	Academic Predictor	Combined Predictor
Final Score	.04	.05	.06	.06	.07
Dist. Grad	.04	.14	.07	.08	.10
Top Grad	.03	.02	.03	.01	.03
Leadership Criterion	.12	.08	.12	.08	.11
Academic Criterion	.04	.10	.05	.15	.11

Note. $N = 1,416$.

Dist. Grad = Distinguished Graduate Status, Top Grad = Top Graduate Status.

Yellow color $p < .05$; Light green color $p < .01$; Dark green color $p < .001$.

The same patterns emerged when examining incremental validity over the AFOQT Academic Aptitude cognitive composite. The most incremental variance for the Distinguished Graduate Status criterion and the Academic Predictor came from their respective specific predictors, while both OSM and the Leadership Predictor explained the same amount of variance in the Leadership Criterion. Similar patterns emerged when we examined results stratified by accession source (see Table 8).

Table 8. Variance Explained by OSM and Alternative Composites

Sample	Criteria	OSM		Specific Predictors		Combined Predictor	
		Variance Explained (%)	Incremental Variance (%)	Variance Explained (%)	Incremental Variance (%)	Variance Explained (%)	Incremental Variance (%)
Full Sample (n = 1414)	Dist. Grad	ns	ns	1.79	Δ1.59	0.96	Δ1.03
	Leadership	1.43	Δ1.59	1.43	Δ1.59	1.13	Δ1.16
	Academic	ns	ns	2.11	Δ1.7	1.04	Δ1.2
OTS CIV (n = 404)	Dist. Grad	ns	ns	1.63	Δ1.37	0.91	Δ0.99
	Leadership	ns	ns	ns	ns	ns	ns
	Academic	ns	ns	3.39	Δ2.4	0.92	Δ1.2
OTS AD (n = 433)	Dist. Grad	ns	ns	2.12	Δ1.92	1.23	Δ1.35
	Leadership	1.38	Δ1.48	1.49	Δ1.62	1.08	Δ1.12
	Academic	ns	ns	2.14	Δ1.95	1.25	Δ1.46
ANG (n = 426)	Dist. Grad	ns	ns	ns	ns	ns	ns
	Leadership	0.68	Δ0.91	1.02	Δ1.34	ns	ns
	Academic	ns	ns	ns	ns	ns	ns
RESERVES (n = 129)	Dist. Grad	ns	ns	ns	Δ2.0	ns	ns
	Leadership	ns	ns	ns	ns	ns	ns
	Academic	ns	Δ2.9	ns	ns	2.57	Δ2.0

Note. Table shows variance explained in the full OTS sample, as well as the sample stratified by accession sources. ns = non-significant; Dist. Grad = Distinguished Graduate Status; Comp = Composite.

3.4 Step 4: Results

We cross-validated the OSM and new composites using two different samples. First, we used Sample 2, the ROTC data collected from 2015 to 2017. As shown in Table 9, the OSM composite had stronger correlations with all the criteria in comparison to the alternative composites. Four individual facets emerged as the best predictors of ROTC success. The OSM composite also explained more incremental variance when used with the AFOQT Academic Aptitude cognitive composite for predicting FT Rank than the other composites (see Table10).

Table 9. Sample 2 (ROTC) Bivariate Correlations

Predictors	Lead	Prof Q	Comm	JD	WE	Foll	AEF Lead	AEF Prof Q	AEF Comm	AE FJ D	AEF WE	AEF Foll	FT_R ANK
OSM	.20	.12	.17	.14	.11	.1	.17	.14	.17	.14	.18	.11	-.18
Overall Predictor	.02	.05	.05	.05	.05	.0	.01	.02	.00	.00	-.03	.03	-.03
Leadership Predictor	.13	.09	.13	.11	.11	.1	.10	.09	.09	.08	.09	.07	-.11
Academic Predictor	-.02	.04	-.01	.03	.02	.0	-.04	-.04	-.05	-.04	-.09	-.03	.02
Combined Predictor	.06	.07	.07	.07	.07	.0	.03	.04	.02	.02	.01	.03	-.05

Note. $N = 1,675$. Lead = Leadership, ProfQ = Professional Qualities, Comm = Communication, JD = Judgement and Decision Making, WE = Warrior Ethos, Foll = Followership, AEF = Air Expeditionary Force, FT Rank = Field Training Rank, AEF = Air Expeditionary Force.

Yellow color $p < .05$; Light green color $p < .01$

Table 10. Sample 2 (ROTC) Explained Variance in Field Training Rank

Predictor	Variance Explained (%)	Incremental Variance (%)
OSM	3.32	3.69
Overall Success Predictor	ns	ns
Leadership Predictor	1.17	1.50
Academic Predictor	ns	ns
Combined Predictor	0.19	0.23

Note. $N = 1,675$

3.5 Step 5: Results

Finally, we cross-validated the composites in Sample 3, the ROTC sample collected from 2018-2020. Once again, we observed that the OSM composite had higher correlations with all four criteria compared to the other alternative predictors and contributed more incremental variance beyond the AFOQT Academic Aptitude cognitive composite. Similar to Sample 2, two individual facets emerged as the strongest facets predicting ROTC success. Table 11 summarizes the Sample 3 bivariate correlations. Table 12 shows the percent of variance explained by the predictors in the Sample 3 criteria and their incremental variance when used with the Academic Aptitude cognitive composite.

Table 11. Sample 3 (ROTC) Bivariate Correlations

Predictors	FTO Rank	LRC 1	Scenario 1	FT Rank
OSM	.22	.10	.13	-.21
Overall Success Predictor	.01	.00	.01	-.01
Leadership Predictor	.15	.05	.08	-.15
Academic Predictor	-.04	-.02	.00	.02
Combined Predictor	.06	.01	.03	-.06

Note. $N = 5,403$. Yellow color $p < .05$; Light green color $p < .01$.

+

Table 12. Explained Variance in Sample 3 (ROTC) Criteria

Predictor	FTO Rank		LRC 1		Scenario 1		FT Rank	
	Variance Explained (%)	Incremental Variance (%)	Variance Explained (%)	Incremental Variance (%)	Variance Explained (%)	Incremental Variance (%)	Variance Explained (%)	Incremental Variance (%)
OSM	4.65	4.83	0.99	1.06	1.64	1.72	4.53	4.76
Overall Success Predictor	ns	ns	ns	ns	ns	ns	ns	ns
Leadership Predictor	2.17	2.14	0.25	0.31	0.65	0.75	2.14	2.45
Academic Predictor	0.12	0.23	ns	0.08	ns	ns	ns	0.16
Combination Predictor	0.32	0.32	ns	ns	0.10	0.10	0.36	0.35

Note. $N = 5,403$

4.0 DISCUSSION

The purpose of this study was to revalidate the OSM composite and explore alternative composite scores for use in predicting officer performance. The OSM score was a moderate predictor of OTS leadership performance but did not predict OTS academic performance or final training scores. This is consistent with prior studies (McHenry et al., 1990) showing personality to have relatively little predictive validity for criteria related to technical proficiency, but strong relation to criteria reflecting leadership, effort, and discipline.

We explored four alternative SDI-O composites: an overall success composite (i.e., Overall Success Predictor), a leadership composite (i.e., Leadership Predictor), an academic performance composite (i.e., Academic Predictor), and a composite that combines the facets from these three together (i.e., Combined Predictor). These composites were effective at predicting success in OTS, including academic abilities and leadership abilities. These findings alone could support the development of a new composite score for use in officer selection.

We cross-validated the OTS results in two different ROTC samples. In both samples, OSM was the strongest predictor of success. We found one stand-alone SDI-O facet that sometimes outperformed the OSM composite. However, these correlations were only marginally stronger than the OSM correlations. These results taken together suggest that the OSM composite remains a useful tool in officer selection.

Although these results are encouraging, additional analyses are needed to determine how to use these personality-based officership indicators along with current indicators of US Air Force aptitude. For example, the OSM and other personality-based officership composites could be used as an additional requirement to the current AFOQT cognitive composites used for officer commissioning (Verbal ≥ 15 and Quantitative ≥ 10). Alternatively, a personality-based officership composite could be used to qualify applicants for officer commissioning programs who did not meet the AFOQT Verbal and/or Quantitative minimums. Yet another alternative would be to replace the current AFOQT Verbal and Quantitative requirements with an officership composite that combines AFOQT cognitive subtests and SDI-O facets.

5.0 REFERENCES

- Barron, L. G (2013). *Development and Validation of the Officer Suitability Measure (OSM Form S)*, AFCAPS-TR-2013-0008. Randolph AFB, TX: Air Force Personnel Center, Strategic Research and Assessment Branch.
- Barron, L. G. (2014). *Development and equating of Officer Suitability Measure (OSM) based on revised AFOQT Form T Self-Description Inventory (SDI+)*, AFCAPS-TR-2014-0002. Randolph AFB, TX: Air Force Personnel Center, Strategic Research and Assessment Branch.
- Christal, R. E. (1994). *R&D Summary report F33615-91-D-OO10*. Brooks AFB, TX: Armstrong Laboratories, Manpower and Personnel Research Division.
- Manley, G. G. (2011). Development of domain and facet-level scales for the Self-Description Inventory (AFCAPS-TR-2011-0007). Randolph AFB, TX: Air Force Personnel Center, Strategic Research and Assessment Branch.
- 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.
- McHenry, J. J., Hough, L. M., Toquam, J. L., Hanson, M. A., & Ashworth, S. (1990). Project A validity results: The relationship between predictor and criterion domains. *Personnel Psychology*, 43, 335-354.
- Woolly, M. R., Mann, K. J., Wilson, R. T., & Carretta, T. R. (2022). *Self-Description Inventory – Officer (SDI-O) 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.

LIST OF SYMBOLS, ABBREVIATIONS, AND ACRONYMS

<i>b</i>	Beta (Regression Weight)
<	Less Than
%	Percent
AA1	Academic Assessment 1
AA2	Academic Assessment 2
AEF	Air Expeditionary Force
AFOQT	Air Force Officer Qualifying Test
ANG	Air National Guard
B1	Informative Briefing
B2	Advocacy Briefing
BOT	Basic Officer Training
Comp	Composite
Comm	Communication
Foll	Followership
FT Rank	Field Training Rank
FTO Rank	Field Training Officer Rank
GLP1	(Tactical) Graded Leadership Position
GLP2	(Operational) Graded Leadership Position
JD	Judgement and Decision Making
Lead	Leadership
LRC	Leadership Reaction Course
<i>N</i>	Sample Size
OSM	Officer Suitability Measure
OTS	Officer Training School
<i>p</i>	Probability
P1	Informative Paper
P2	Advocacy Paper
ProfQ	Professional Qualities
<i>r</i>	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)
USAF	United States Air Force
USAF A	United States Air Force Academy
WE	Warrior Ethos
WELPS	Warrior Expeditionary Leadership Problem Solving