Research Note 2021-03

Optimizing Officer Classification: Selection of Predictor Constructs and Measures

Emily S. Medvin, Editor
Human Resources Research Organization

Peter J. Legree, Editor
Mark C. Young, Editor
Robert N. Kilcullen, Editor
U.S. Army Research Institute

September 2021

United States Army Research Institute for the Behavioral and Social Sciences

U.S. Army Research Institute
for the Behavioral and Social Sciences

Department of the Army
Deputy Chief of Staff, G1

Authorized and approved:

MICHELLE L. ZBYLUT, PhD.
Director

Research accomplished under contract
for the Department of the Army by:

Human Resources Research Organization

Technical Review by:

Jessica R. Carre, U.S. Army Research Institute

DISPOSITION

This Research Note has been submitted to the
Defense Technical Information Center.

REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy) September 2021	2. REPORT TYPE Interim	3. DATES COVERED (from. . . to) September 30, 2019 - June 30, 2020
4. TITLE AND SUBTITLE Optimizing Officer Classification: Selection of Predictor Constructs and Measures		5a. CONTRACT OR GRANT NUMBER W911NF-19-C-0065 5b. PROGRAM ELEMENT NUMBER 622785
6. AUTHOR(S) Emily S. Medvin, Peter Legree, Mark C. Young, and Robert N. Kilcullen		5c. PROJECT NUMBER A790 5d. TASK NUMBER
		5e. WORK UNIT NUMBER 1011
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Human Resources Research Organization 66 Canal Center Plaza, Suite 700 Alexandria, Virginia 22314		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences 6000 6th Street (Bldg 1464/Mail Stop 5610) Ft. Belvoir, VA 22060-5586		10. MONITOR ACRONYM ARI 11. MONITOR REPORT NUMBER Research Note 2021-03

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES

ARI Research POC: Peter J. Legree, Selection and Assignment Research Unit

14. ABSTRACT *(Maximum 200 words)*:

To assign newly commissioned officers to branches, the U.S. Army uses a complex process that is primarily based on Cadet Order of Merit List rank, cadet branch preference, and cadet willingness to extend their active duty service obligation, as well as U.S. Army operational requirements and personnel policy. Previous analyses demonstrate that pre-commissioning and non-cognitive metrics could improve the in-unit performance of officers by improving the branch assignment process. However, those previous analyses used non-cognitive metrics that were designed to predict overall officer performance as opposed to officer branch performance. To better support the branch assignment process, the U.S. Army Research Institute for the Behavioral and Social Sciences is developing and validating psychological metrics that are likely to predict branch performance. Accordingly, this project identified attributes that are likely to predict officer branch performance based on job analysis and focus group data, as well as supplemental analyses of existing data. This information was used to propose measures to include in the Cadet Assessment Battery for ongoing research to improve officer classification and branch assignment procedures. The incremental validity of the Cadet Assessment Battery to improve the officer branch assignment process will be assessed in an upcoming project.

15. SUBJECT TERMS

Officer branching, Classification

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified	Unlimited Unclassified	63	Tonia S. Heffner 703-545-4408

Standard Form 298

Research Note 2021-03

Optimizing Officer Classification: Selection of Predictor Constructs and Measures

Emily S. Medvin, Editor
Human Resources Research Organization

Peter J. Legree, Editor
Mark C. Young, Editor
Robert N. Kilcullen, Editor
U.S. Army Research Institute

Selection and Assignment Research Unit
Tonia S. Heffner, Chief

September 2021

OPTIMIZING OFFICER CLASSIFICATION: SELECTION OF PREDICTOR CONSTRUCTS AND MEASURES

EXECUTIVE SUMMARY

Research Requirement:

The Optimizing Officer Classification project was designed to develop and evaluate a battery of cognitive and non-cognitive measures that could be used to assign newly commissioned officers to those branches in which they are likely to perform well. To support this objective, job analysis and focus group data were collected to identify promising predictors and create the Cadet Assessment Battery (CAB) for future applications supporting officer classification.

Procedure:

To create the CAB, we reviewed relevant research, including existing taxonomies of personal attributes relevant for U.S. Army officers, with focus on previous officer job analysis results and relevant research conducted with Soldier populations. A series of focus groups, which were comprised of Captain Career Course cadre from seven branches, reviewed these materials. This effort resulted in branch-specific lists of critical personality attributes selected to provide differential validity effects over branch.

In addition, we conducted supplemental analyses using data from the Army Officer Classification project (Ford et al., 2020). These analyses examined individual predictor contributions to the previous classification research and created branch-specific keys for the Leader Knowledge Test and Work Values Inventory. We assessed the utility of these keys to improve the prediction of performance within branch and examined the utility of using U.S. Army officer vocational interest profiles for predicting branch-specific performance outcomes.

Findings:

We synthesized conclusions from the literature review, focus groups, and analyses to identify and describe potential predictor constructs and measures to be included in the CAB. Recommended predictors included measures of cognitive abilities, personality traits and tendencies, as well as vocational interests, knowledge and skills.

Utilization and Dissemination of Findings:

Predictors included in the CAB were based on the strength of the theoretical and empirical rationale for each proposed predictor as well as considerations regarding their administration time, administration mode, and the adaptability of measures that had been originally designed for enlisted personal. The CAB will be administered to junior officers and its concurrent validity will be assessed in the next phase of the project.

OPTIMIZING OFFICER CLASSIFICATION: SELECTION OF PREDICTOR CONSTRUCTS AND MEASURES

CONTENTS

Page

APPENDICES

TABLES

CHAPTER 1: RESEARCH OVERVIEW

Laura A. Ford and Emily S. Medvin

The U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) has supported an ongoing research program to develop and validate personnel assessment measures that can be used to address U.S. Army officer selection and classification goals since 2007. This research program has primarily focused on validating non-cognitive predictors (e.g., the Cadet Background and Experience Form [CBEF]) against metrics of cadet performance and continuance in officer pre-commissioning programs (Baldwin & Young, 2020; Bynum & Young, 2020; Legree et al., 2014; Putka et al., 2009). Based on the validity findings for non-cognitive predictors, the U.S. Army Cadet Command (USACC) implemented the CBEF to help award scholarships to individuals who are likely to excel in pre-commissioning training programs and become commissioned officers.

More recent work has evaluated the utility of using pre-commissioning training performance and non-cognitive measures to guide the branch assignment process for newly commissioned officers (Ford et al., 2020; Legree et al., 2019). The U.S. Army Cadet Command (USACC) generally assigns newly commissioned officers to one of 17 branches. These branches include: Adjutant General, Air Defense Artillery, Armor, Aviation, Chemical, Engineers, Cyber, Field Artillery, Finance, Infantry, Medical Service, Military Intelligence, Military Police, Ordnance, Quartermaster, Signal, and Transportation. The current branch assignment process attempts to address multiple operational objectives that include: meeting Army end-strength manning requirements; distributing high quality leaders across branches; and balancing officer demographic characteristics over branches. In addition, the USACC attempts to assign cadets to those branches where they will likely perform well.

Although this process is designed to satisfy numerous operational criteria, there is little empirical evidence that it produces an officer corps that maximizes officer performance, continuance, and satisfaction. Therefore, research has evaluated the longitudinal validity of pre-commissioning training performance and non-cognitive measures against indices of officer performance within operational units (Ford et al., 2020). While correlational analyses established the long-term validity of pre-commissioning and non-cognitive measures against metrics of in-unit officer performance, hierarchical regression analyses also identified these measures as having differential levels of validity across officer branches. Specifically, subsequent analyses found that these pre-commissioning training and non-cognitive measures can be used to improve the match between the characteristics (e.g., talents) of newly commissioned officers and the occupational requirements of their initial branch assignments to enhance officer performance.

However, two important issues limit the utility of those analyses and results. First, the analyses projected performance gains through improved branch assignment for most, but not all, officer branches. Importantly, performance gains were lacking for key combat branches such as Infantry. Second, the available suite of non-cognitive measures had been developed to predict overall performance in pre-commissioning programs (i.e., these measures had not been developed to provide differential validity and support branch assignment; Ford et al., 2020).

Nonetheless, these results demonstrate the potential of using a variety of pre-commissioning and non-cognitive measures to improve the process by which USACC assigns newly commissioned officers to their initial branches, thereby improving the performance and career continuance of these officers.

In response to these results, ARI initiated the Optimizing Officer Classification (OOC) project to broaden the range and scope of cognitive and non-cognitive predictors that could be available to improve the officer branch assignment process. Our approach is consistent with tenets of personnel classification theory (Brogden, 1959) as well as Differential Assignment Theory (Zeidner, Johnson & Scholarios, 1997). These theories underscore that classification effects are highly dependent upon the use of measures that can be weighted to improve the match between the characteristics of individuals with the requirements of various occupations.

The initial phase of the OOC project focused on identifying a broad array of constructs that are likely to provide differential validity across officer branches and have the potential to improve the branch assignment process for newly commissioned officers. In addition, an important aspect of this effort was to identify metrics that could be used to assign officers to combat branches. We were less interested in developing predictors of overall officer performance because these measures cannot be used to identify officer branches that are most appropriate for specific individuals.

This report describes activities that were conducted to develop the Cadet Assessment Battery (CAB). As described above, the CAB was envisioned to include scales that are likely to provide differential validity across officer branches and to be used in later projects to enhance the USACC branch assignment process for newly commissioned officers.

Similar to conventional job analyses, we first reviewed relevant literature and analyzed available data to identify promising predictors of officer performance (Chapter 2). We then conducted focus group interviews to assess the potential of identified constructs to provide differential validity within and across officer branch clusters (Chapter 3). We then summarize the conceptual expectations regarding the identified predictors (Chapter 4). Finally, we considered practical limitations of these scales to develop the proposed CAB (Chapter 4).

In later phases of the OOC project, we expect to validate the CAB against branch-specific measures of officer performance and career continuance. If successful, this project will provide a battery of scales that can be used to improve the branch assignment process for newly commissioned officers.

CHAPTER 2: LITERATURE REVIEW AND BACKGROUND RESEARCH

Michael P. Wilmot, Teresa L. Russell, and Emily S. Medvin

To identify potential predictors of branch-specific performance, we reviewed existing literature and previous research projects examining the use of cognitive abilities, personality, and person-job fit metrics to predict the performance of military personnel.

Literature Review

The goal of the literature review was to develop a reasonably comprehensive taxonomy of personal attributes relevant for improving the Army officer classification process. Our primary sources of information included:

- Comprehensive and recent taxonomies of knowledge, skills, abilities and other characteristics (e.g., the U.S. Department of Labor's Occupational Informational Network [O*NET]; Stanek & Ones, 2017);
- Previous job analysis literature that describes the requisite skills, abilities and characteristics necessary to perform well in Army officer occupations (e.g., Paullin et al., 2014; Paullin et al., 2011); and
- Reviews focused on military uses for vocational interest assessment batteries (e.g., Farmer et al., 2006; Kirkendall et al., 2020).

Table 1 presents the Officer Classification Taxonomy, which is organized by four broad domains: (a) cognitive abilities, (b) personality traits and tendencies, (c) vocational interests and knowledge, and (d) skills. These four domains are then organized into relevant sub-domains and associated attributes. To balance our goal of a theoretically accurate and comprehensive taxonomy with practical assessment constraints, we selected personal attribute constructs at the "meso-level" of abstraction. That is, we included constructs that occupy the space between the "macro-level" and the myriad of "micro-level" attributes and measures.

Table 1. *Officer Classification Taxonomy of Personal Attributes*

Domains/Attributes	Definition
	Cognitive Abilities
Memory	Ability to retain and recall information
Perceptual Speed and Accuracy	Ability to perceive things quickly and accurately and to detect similarities or differences in objects, words, or numbers
Spatial Ability	Ability to manipulate objects in the mind's eye, imagine objects from different perspectives, and remain unconfused by different views
Verbal Reasoning	Ability to reason and draw conclusions based on verbal or written materials.
Quantitative Reasoning	Ability to use induction or deduction in reasoning with quantitative concepts (e.g., numbers, mathematical relations)
Psychomotor Ability	Ability to perform activities that require eye-hand coordination or coordinating the simultaneous movements of one's limbs
Attentiveness	Ability to focus on the problem or situation and shift attention between activities when appropriate

(continued)

Table 1. ***(Continued)***

Domains/Attributes	Definition
Situational Awareness	Ability to perceive events in the immediate environment and how information, events, and actions will impact goals and objectives
	Personality Traits and Tendencies
Emotional Stability	
Stress Tolerance	Tends to control emotions and keep cool in stressful situations
Composure	Tends to feel happy and self-confident
Agreeableness	
Compassion	Tends to show care, consideration, and altruism toward others
Politeness	Tends to consider and respect the needs and desires of others
Conscientiousness	
Industriousness	Tends to have high aspirations, initiative, work hard, and achieve goals
Orderliness	Tends to be neat, organized, planful, and detail oriented
Extraversion	
Assertiveness	Tends to be socially dominant, influential, energetic, and takes charge
Enthusiasm	Tends to experience positive emotions and enjoy the company of others
Openness/Intellect	
Openness	Tends to be open to art, culture, and imagination
Intellect	Tends to be quick, open to new ideas, and intellectually engaged
Compound Traits	
Health and Fitness Orientation	Tends to maintain good health and physical conditioning by prioritizing good nutrition, physical exercise, and adequate sleep
Initiative	Tends to rely on one's own abilities to overcome obstacles and be effective in situations that require a willingness to originate action or take independent action to achieve a goal
Learning Orientation	Tends to seek out learning opportunities and enjoy acquiring new knowledge and skills
Self-Efficacy	Tends to be confident in one's ability to succeed, effectively meet challenges, and overcome obstacles.
Team Orientation	Tends to enjoy being part of a team, have a strong identification with teammates, and feel a sense of commitment and obligation to the team
	Vocational Interests and Knowledge
Realistic Interests	Interests in practical, hands-on, concrete activities with physical objects
Investigative Interests	Interests in rational and systematic reasoning and working with facts, data, and abstract concepts
Artistic Interests	Interests in expressing oneself creatively
Social Interests	Interests in working with and helping others
Enterprising Interests	Interests in persuading people or exerting influence over others
Conventional Interests	Interests in organizing data, people, or physical environments
Occupation-Specific Knowledge	Depth and breadth of knowledge related to a specific occupation
	Skills
Social/Interpersonal	
Behavioral Flexibility	Skill in changing one's own behavior, approach, or interpersonal style as appropriate

(continued)

Table 1. ***(Continued)***

Domains/Attributes	Definition
Cultural Awareness	Skill in demonstrating acceptance and understanding of individuals from other cultural and social backgrounds
Perspective Taking	Skill in understanding how people interpret events and interpersonal interactions
Social Perceptiveness	Skill in accurately perceiving other people's motives, attitudes, and feelings based on what they do or say, and understanding one's own impact on the behavior of others
Social Sensitivity	Skill in displaying diplomacy and tact when interacting with others
Self-Management	Skill in managing the full range of work and non-work responsibilities
Leadership	
Directing and Supervising Others	Skill in assigning subordinates specific tasks and setting individual goals for work and assignments
Motivating Others	Skill in generating support, involvement, energy, and enthusiasm for the mission among subordinates
Delegating	Skill in appropriately delegating authority and responsibility for decision making and for planning and executing tasks
Shared Leadership	Skill in organizing and orienting team members to meet goals
Team Building	Skill in assembling a team of people who work together effectively, fostering group identity and cohesion
Training and Developing Others	Skill in determining the training needs of individual subordinates and providing the appropriate level of instruction, guidance, and developmental opportunities
Management	
Adaptability	Skill in modifying behaviors and/or plans as necessary to reach goals
Coordinating	Skill in coordinating the efforts of multiple, diverse groups to accomplish a mission
Innovation	Skill in developing and using new and creative methods or strategies to accomplish work or achieve goals when established methods and procedures are inapplicable or ineffective
Judgment and Decision Making	Skill in making decisions based on accurate and appropriate assessment of the costs/benefits and short- and long-term consequences of alternative actions and solutions
Planning and Organizing	Skill in defining the means for achieving the unit or organization goals, establishing priorities, anticipating important or critical events, identifying resource requirements, and assigning responsibility and performance expectations for work
Problem Solving	Skill in identifying complex problems, gathering related information, evaluating information relevance, and generating alternative solutions
Relationship Building	Skill in developing and maintaining effective working relationships

The first major attribute domain in the Officer Classification Taxonomy is *Cognitive Abilities*. Table 1 presents eight sub-domains of ability constructs, their definitions, and associated personal attributes. Most of the abilities are reported by Stanek and Ones (2017) and Paullin et al. (2011). Although, Paullin and colleagues also included military-specific abilities that were found to be important in job analyses.

The second major attribute domain in the Officer Classification Taxonomy corresponds to *Personality Traits and Tendencies*, which refers to characteristic patterns of thinking, feeling, and behaving across situations. The Big Five model (i.e., Emotional Stability, Agreeableness, Conscientiousness, Extraversion, and Openness/Intellect) is the most robustly supported descriptive taxonomy of traits (Stanek & Ones, 2017). Regarding the hierarchical organization outlined in Table 1, each Big Five dimension has two subordinate *aspect* traits that occupy the meso-level of abstraction. These 10 aspects appear useful for classification efforts because they show evidence of predictive utility over their parent Big Five trait and show differential external relations (Judge et al., 2013). In addition, there are also several personality constructs that cannot be reduced to a single Big Five dimension, but are useful for predicting key outcomes. These more complex traits are referred to as *compounds*. As Table 1 indicates, several compound traits identified by Paullin et al. (2011) are also included in the Officer Classification Taxonomy of personal attributes.

The third major attribute domain is labelled *Vocational Interests and Knowledge*, which encompasses patterns of preferences for the type of work that individuals prefer, as well as the knowledge needed for specific types of work. The most frequently used model for vocational interests is the Holland's (1997) RIASEC model.[1] The history of vocational interest research indicates that these measures predict performance and retention (Van Iddekinge, Putka, & Campbell, 2011), and show promise for improving the job assignment process (Kirkendall et al., 2020). In addition, the O*NET used the RIASEC model to provide interest ratings for all its occupations (Rounds et al., 2010), and RIASEC based-ratings are available for each of the U.S. Army officer branches (Ford et al., 2020). Accordingly, we included the six RIASEC interests in the Officer Classification Taxonomy (see Table 1). We also included occupation-specific knowledge as an attribute to target any specific knowledge needed to successfully perform in a particular branch.

The fourth domain corresponds to *Skills*, which involves attributes in the sub-domains of social/interpersonal, self-management, leadership, and management. We included attributes for each of these Skills sub-domains based on conclusions provided by Paullin and her colleagues (2011).

Enlisted Selection and Classification Research

While the literature review primarily focused on results collected for officer populations, we supplemented this information using in-depth analyses conducted for enlisted populations. In the 1980s and 1990s, each of the Armed Services conducted Job Performance Measurement (JPM) projects to evaluate the validity of the Armed Services Vocational Aptitude Battery (ASVAB) for predicting in-unit job performance. Before that time, the ASVAB had been primarily validated against training performance metrics. Each Service gathered its own predictor and criterion data and conducted its own validation analyses. In addition, the Army and Navy used the opportunity to evaluate new predictors. Because of the similarity of military officer and enlisted populations, these projects represent an important potential source of measures that could be used for officer branch assignment purposes.

[1] Realistic (R), Investigative (I), Artistic (A), Social (S), Enterprising (E), and Conventional (C)

The Army's JPM project was called Project A. It included new spatial tests, psychomotor/perceptual tests, a biodata/personality measure, and a measure of occupational interests. Classification analyses were done comparing expected improvements in mean predicted performance with predictor composites, not individual test scores (Rosse, Campbell, & Peterson, 2001).

The Army compared the classification potential of the ASVAB to that of the ASVAB and spatial tests, the ASVAB and psychomotor/perceptual tests, the ASVAB and the personality measure, and the ASVAB and the interest measure. While all of the new test composites increased mean predicted performance, occupational interests and psychomotor/perceptual tests provide the greatest increases in predictive validity.

As a follow-on to the JPM projects, the U.S. Department of Defense sponsored the Enhanced Computer Administered Test (ECAT) project, as a joint-Service project in order to evaluate the most promising scales from the Army and Navy JPM projects using joint-service data. The resultant test battery consisted of 10 ASVAB subtests, three graphical stimuli tests from Project A, one perception test from Project A, two psychomotor tests from Project A, a test of sequential memory from the Navy, and two tests of working memory from the Navy.

The battery of tests was administered to over 9,000 trainees in 17 military occupations. End-of-training criterion data included Final School Grades (FSG) and scores on hands-on performance tests. Three types of indices were used in evaluating test batteries that could be formed from the full set of 19 tests: absolute validity across jobs, classification potential (mean predicted job performance – Brogden Index), and subgroup differences (Sager, Peterson, Oppler, Rosse & Walker, 1997). These indices were calculated for every possible combination of tests.

The general ability tests were found to be best for maximizing absolute validity, while the best tests for maximizing mean predicted performance (Brogden, 1959) tended to be tests of specific skills (e.g., target identification, auto and shop information, spatial orientation). Of these tests, some appear more relevant to officer occupations than others. For example, target identification and auto and shop information are likely less relevant, whereas spatial orientation and figural reasoning are likely more relevant. We included in our consideration of potential predictors for the current project those specific abilities that maximized mean predicted performance and were likely relevant to officer occupations.

Talent Management Supplement

To broaden our review of existing relevant research, we obtained analyses and materials from the ARI Talent Management Job Analysis Project. The goal of this project was to identify the talents, skills, and abilities required for officers at different stages in their career for different branches and functional areas. The Talent Management Job Analysis data were collected as part of an online survey administered to officers in the Active Army component in 2017. The survey was designed to quantify the importance of various talents and the frequency and importance of various skills and abilities required for officers to effectively perform their jobs.

While analyses conducted using these data identified some talents, skills, and abilities (e.g., Communicator, Interpersonal, Problem Solver, Cooperation/Teamwork, and Active Listening) as important across nearly all branches, the analyses also identified talent, skill, and ability requirements that were unique to particular branches. For example, the talent of Innovating Technology was rated "critical" for the Cyber branch, but it was rated as "least important" for many other branches. Appendix A provides the Talent Management Importance ratings by branch, and Appendix B provides the Talent Management Skill and Ability criticality ratings by branch.

Crosswalk: Officer Classification Taxonomy by Talent Management Project

To combine attributes from the Officer Classification Taxonomy with the talents, skills, and abilities from the Talent Management Job Analysis Project, we developed a crosswalk to link the two sources of information. We compared the definitions of attributes in the Officer Classification Taxonomy to definitions of the constructs in the Talent Management dataset to identify conceptually similar constructs and attributes. Of the 49 attributes in the Officer Classification Taxonomy, 28 were conceptually linked to talents, skills, and abilities in the Talent Management work. The remaining 21 attributes included vocational interests, and abilities such as memory, as well as personality traits such as assertiveness and politeness

Appendix C presents the crosswalk of all attributes and corresponding definitions that were included in our focus group materials, including attributes that were found critical in the Talent Management work that were not originally included in the Officer Classification Taxonomy. These materials were used to conduct the focus groups that are described in Chapter 3.

CHAPTER 3: FOCUS GROUP RESEARCH

Emily S. Medvin, Christopher R. Graves, and Brenda Ellis

We conducted focus groups with representatives from different branches to gather their perspectives on personal attributes most critical to success in each branch, as well as attributes that were likely most differentiating across branches (i.e., those attributes that would predict success in some branches but not others). Focus groups were conducted with Captains Career Course training cadre from the following seven branches: Armor, Infantry, Field Artillery, Air Defense Artillery, Quartermaster, Transportation Corps, and Ordnance.[2] These branches were selected because the branch assignment analyses indicated that additional metrics are required to ensure that officer performance gains would be obtained for officers assigned to combat branches through improved branch assignment algorithms (Ford et al., 2020). The number of cadre per branch ranged from 2 to 6, with an average of 5 cadre per branch. Participating cadre included 33 senior captains and 1 lieutenant.

Focus Group Planning and Execution

To guide the focus groups, we created branch-specific attribute lists based on the results from the literature review and the analysis of the Talent Management job analysis results. The starting point for each branch's attribute list was the Talent Management Attributes data described previously. These findings identified those talents that were considered "very important" for each branch, as well as those skills and abilities that were considered "highly critical" for each branch.

A talent was considered very important if its mean rating of importance was between 3.00 and 4.00, on a scale of 0–4. A skill or ability was considered highly critical if its criticality score was between 12.00 and 20.00, on a scale of 0–20. The criticality score was calculated by multiplying respondents' frequency and importance ratings from the skill or ability. The very important talents and highly critical skills and abilities, displayed with their definitions, were consolidated into one *critical attribute* list for each branch. We also included interests from the RIASEC model in a branch's list if the mean importance for that interest was greater than or equal to 4.00, on a scale of 1–7 (Bayer, McLean, & Salyer, 2017). Finally, we developed a list of *potential new* attributes and definitions for each branch. These lists included attributes that were included in the Officer Classification Taxonomy, but did not have a counterpart in the Talent Management job analysis survey.

Each focus group session began with a brief discussion on junior officer performance across unit types in the relevant branch. This discussion helped ensure that focus group participants were considering the full scope of conditions and performance requirements for junior officers in their branch when looking at the attribute lists. After this discussion, we provided each participant with the critical attribute list for their branch. Participants reviewed the list and discussed their beliefs regarding the most and least important attributes for junior officer performance in their branch. We then provided participants with a list of potential new attributes and asked them to identify attributes that should also be included in the critical list for their branch.

[2] We were unable to conduct focus groups with additional branches due to COVID-19-related travel restrictions.

As a final step in the focus groups, participants documented critical incidents for specific attributes to provide examples of the presence or absence of specific attributes affecting officer job performance. Participants were assigned attributes for this task based on those endorsed during the first part of the focus groups. Facilitators prioritized attribute assignment in accordance with previous data identifying attributes with differential importance across branches (i.e., attributes with a standard deviation greater than 0.90 across the branches from the Talent Management data). Facilitators also prioritized the assignment of attributes from the potential new attribute list that were nominated as critical by the participants.

Focus Group Analyses

We consolidated the data from each focus group into one spreadsheet, noting the number of participants by branch, who endorsed each attribute as critical, as well as the number of participants who described each attribute as least important. These data were used to identify consistencies and inconsistencies (i.e., agreement vs. disagreement) with the existing Talent Management job analysis data. For each attribute on a branch's list of critical attributes, we noted agreement if at least one focus group participant in a particular branch endorsed the attribute as among the most critical for successful branch performance. We also documented disagreement if none of the participants endorsed an attribute, or if a participant identified that attribute as among the least important. Across the seven branches, there was 61.3% agreement in terms of the criticality of attributes.

To identify agreement with job analysis data regarding the difference in importance of attributes across branches, we calculated the range of endorsement for each attribute across branches. For example, four participants in one branch (Field Artillery) endorsed the situational awareness attribute as highly critical, but participants from other branches did not endorse this attribute as critical, thereby resulting in a range of "4." Similarly, the attribute of delegating was endorsed as critical by 1 participant in two branches (Quartermaster and Transportation) but was identified as among the least important by 3 participants in another branch (Armor), also resulting in a range of "4".

For attributes with job analysis data, we identified those with the most differentiation potential by determining which had high standard deviations in the data as well as a relatively large range from the focus group findings. For the attributes without data (i.e., those on the potential new attributes list), we identified those attributes with high ranges of endorsement from the focus group members for potential inclusion in the predictor battery.

Chapter 4: Army Officer Classification Analyses

Oren R. Shewach, Christopher R. Huber, and Brenda Ellis

While the predictor measures used in the Army Officer Classification project were primarily developed to augment the selection process of four-year ROTC scholarship recipients, the data analyses indicated that those constructs also had potential for differentiating performance across branch cluster (Ford et al., 2020). For clarification purposes, we refer to branch "clusters" rather than branches in most of the analyses described below. The branch clusters had been created to compensate for low sample sizes obtained for specific branches (Ford et al., 2020). For example, data for Infantry, Armor, and Aviation officers were merged into the Maneuver cluster for analytic purposes. See Figure 1.

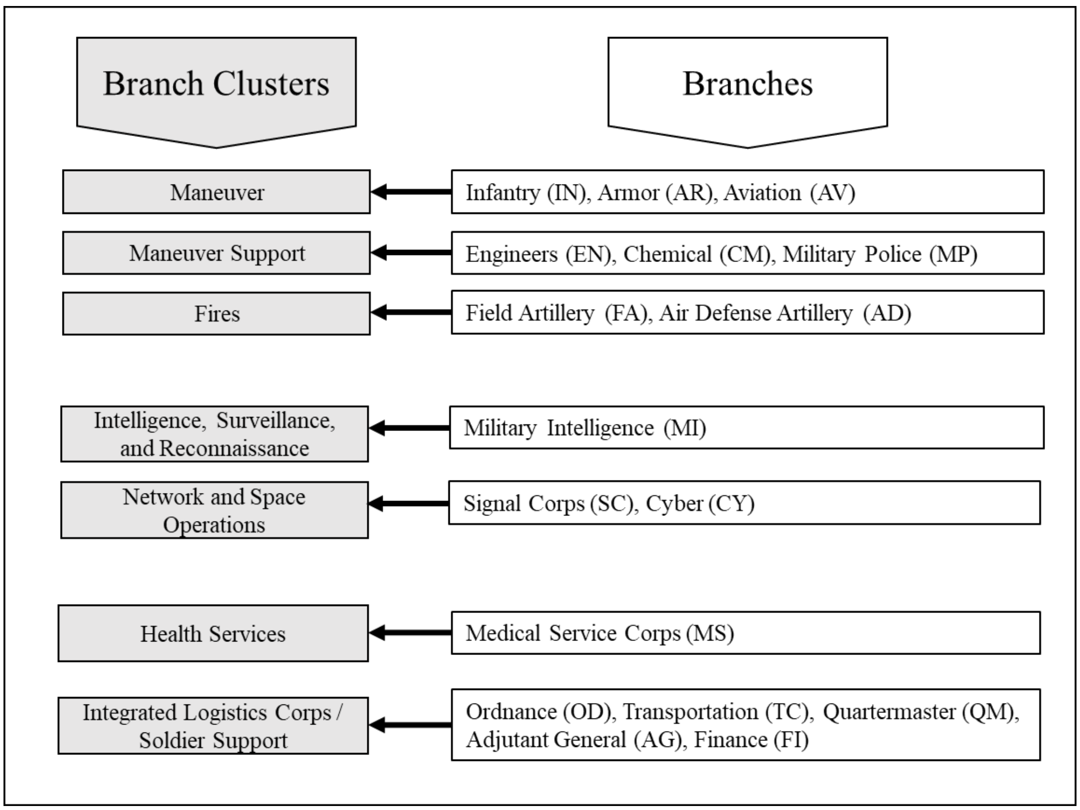

Figure 1. ***Organization of officer branches for the Army Officer Classification project.***

We conducted several additional analyses using the Army Officer Classification project data to further explore these results, and these analyses are described below. The first set of these analyses explored the utility of each of the 14 predictors used for the Army Officer Classification project. The second set of analyses explored the potential use of within-branch scoring keys to improve the utility of the Leader Knowledge Test (LKT) and Work Values Inventory (WVI) predictor scales. The final set of analyses examined possible refinements to the procedure used to compute vocational interest scores based on the computation of RIASEC scores using branch cluster standards.

Classification Algorithm: Predictor Evaluation

As described above, ARI developed a classification algorithm for the Army Officer Classification project (Ford et al., 2020). This algorithm modelled the cadet branching process by considering cadet branch preferences, and Army requirements, as well as incorporating optimization options across multiple criteria. Although the classification algorithm's overall effectiveness was evaluated, underlying validity analyses did not examine the contribution of individual predictors to the classification algorithm. To expand those results and guide this project, we sought information about individual predictor effectiveness to determine which predictors were most promising for further development. Therefore, the 14 predictors that were used for the Army Officer Classification project were more fully examined for this chapter. (Due to the design of the Army Officer Classification project, some of these predictors can also be described as predictor composites because they were based or more than one predictor scale.) These 14 predictors included:

- Three (3) broad predictor measures underlying the Cadet Outcome Metric Score (OMS);
- Six (6) predictors derived from the CBEF biodata scales;
- Two (2) predictors derived from the LKT;
- The WVI Profile Similarity Index;
- The Cadet college quality index; and
- The officer branch fit measure, which was based on the college major.

Ford et al. (2020) provides details regarding these predictors/predictor composites. For this project, we conducted additional analyses to evaluate the relative effectiveness of the predictors in contributing to the Army Officer Classification results.

Methodological Overview

To understand the rationale for our approach, it is important that gains in classification efficiency are approximated by the Brogden Index of Classification Efficiency:

$$R\sqrt{1-r},$$

where *R* is the average multiple correlation between the tests in a battery and performance for each branch cluster, and *r* is the average intercorrelation between predicted scores across branch clusters (Brogden, 1959; Sager et al., 1997).

Based on Brogden's Index, we can infer that the most useful predictors are those that (a) have the largest impact on validity, and (b) help to reduce the correlation between predicted values for different branches or branch clusters. We refer to these two factors as *predictor validity* and *predictor differentiation*. We then used the existing Army Officer Classification validation data to evaluate the contribution of each predictor to classification. Ultimately, a two-pronged approach was used to quantify predictor validity and predictor differentiation, as discussed below.

Predictor Validity Dominance Analyses. Conceptually, predictors that have the largest impact on validity will tend to have the greatest impact on classification efficiency. However, each predictor does not exist in a vacuum, and it is important to consider the effect of an individual predictor, holding all other predictors constant. Therefore, we followed a regression-based approach to evaluating predictor validity.

For the Army Officer Classification validation analyses, a Bayesian Model Averaging (BMA) regression approach had been used to estimate the full battery prediction models (Ford et al., 2020). BMA estimates a regression model for every possible combination of predictors (Russell et al., 2017). Ford et al. (2020) generated BMA regression weights and predictor criticality values for each branch cluster and used this information to model potential classification gains that could be obtained through improvements to the branch assignment process.

While this approach was logical for the Army Officer Classification validation analyses, predictor criticality values do not quantify relative strength of prediction for a given predictor, but rather the likelihood that predictor is included in the "best" sub-models. In addition, the use of regression weights as relative strength indices is inappropriate when predictors are correlated (Budescu, 1993; Johnson & LeBreton, 2004).

Therefore, we used relative importance analyses to quantify the strength of individual predictors in the regression analyses developed to predict performance. Dominance analysis is a prominent method of conducting relative importance analysis (Budescu, 1993). Dominance analysis consists of estimating the incremental R^2 contribution of each predictor in the model across all possible sub-models involving that predictor. Dominance weights are ultimately the aggregated incremental R^2 contribution across all possible sub-models. We conducted relative importance analysis on data from each of the seven branch clusters using the 14 predictors. The criterion was an overall job performance score that combined a number of specific criterion measures (e.g., supervisor ratings, merit awards). Overall performance was standardized within branch cluster, and the analyses were conducted in the "Dominance Analysis R" package (Navarrette & Soares, 2020).

Predictor Differentiation. The second factor affecting classification gains is predictor differentiation. Predictors that reduce the correlation between predicted values across branch clusters are more "differentiating," and are therefore more useful for the classification process. Predictors that can distinguish predicted performance within different branches or branch clusters are inherently valuable to classification efficiency.

We assessed predictor differentiation via a leave-one-out approach (Hastie, Tibshirani, & Friedman, 2008). Each of the 14 predictors was iteratively left out of the regression model. Excluding a predictor allows for examining that predictor's effect on differentiation. For each of the seven branch clusters, 15 regression models were estimated: one overall model that included all 14 predictors, and 14 regression models with each model based on 13-predictor models so that one predictor was left out of each model. The regression models were estimated using the actual samples for each of the seven branch clusters, consisting of sample sizes from n = 127 (Health Services) to n = 473 (Integrated Logistics / Soldier Support).

We then generated seven sets of predicted performance scores for each officer in the validation sample (i.e., one set per branch cluster; $N = 1{,}940$). We repeated this process for each of the 15 regression models. We then computed intercorrelations between the predicted scores for the branch clusters for each leave-one-out regression model. Following this approach, a higher average intercorrelation represents a predictor composite that is more differentiating. For example, if the average intercorrelation of predicted scores across branch clusters increased by $r = +.05$ from the baseline model when leaving out the LKT Skills, then the LKT Skills predictor would have a positive differentiating effect. This result would show that there is less differentiation between branch clusters when LKT Skills predictor is excluded and would support retaining the LKT Skills predictor in the classification battery.

Results

Predictor Validity Dominance. Table 2 presents results from the dominance analyses for the 14 predictors. The OMS Leadership Component had the highest predictive validity by a substantial margin. The OMS Leadership measure had an average incremental R^2 contribution of .04, accounting for 23.78% of the explainable criterion variance across branch clusters.

Figure 2 presents dominance weight percentage values for the branch clusters: Fires; Health Services; Integrated Logistics Corps/Soldier Support (ILC/SS); Intelligence, Surveillance, and Reconnaissance (ISR); Maneuver; Maneuver Support; and Network & Space Operations. Figure 2 shows OMS Leadership was the top predictor for each branch cluster with the exception of the Fires branch cluster, in which it had a trivial weight percentage. This result underscores the importance of maintaining the OMS Leadership predictor in the classification battery.

Table 2. ***Prediction Validity: Dominance Analysis Results***

Predictor Composite	Mean Dominance Weight % (*SD*)	Dominance Weight % Range	Mean Raw Weight	Mean Weight Rank
OMS Academic Component Score	7.77 (7.25)	.48, 18.05	.014	6.57
OMS Leadership Component Score	23.78 (13.22)	1.10, 39.20	.040	2.43
OMS Physical Component Score	6.25 (5.78)	1.66, 17.27	.008	6.86
College Quality Index	8.54 (7.39)	.30, 18.92	.017	6.86
CBEF Achievement	6.16 (3.99)	1.06, 13.57	.010	6.71
CBEF Fit/Commitment	7.36 (8.40)	.96, 24.68	.010	7.71
CBEF Response Distortion	4.64 (5.36)	1.31, 16.50	.005	9.00
CBEF Peer Leadership	6.06 (3.63)	.52, 12.62	.011	6.86
CBEF Fitness Motivation	4.92 (2.50)	.76, 7.34	.007	7.43
CBEF Stress Tolerance	7.06 (6.29)	.97, 20.09	.010	6.71
WVI Profile Similarity Index	1.47 (1.25)	.45, 4.10	.002	12.14
LKT Characteristics	4.43 (3.83)	.86, 12.16	.008	8.14
LKT Skills	8.37 (6.15)	2.05, 18.05	.014	6.29
Predicted Interest-Branch Fit	3.17 (5.38)	.06, 14.08	.007	11.29

Note. All results are averaged across branch clusters. Mean raw weight reflects the average incremental R^2 contribution for a given predictor across all sub-models. Mean weight rank reflects each predictor composite's rank-order (by weight percentage) within a branch cluster, averaged across all branch clusters.
Predictor Composite: OMS = Outcome Metric Score; CBEF = Cadet Background and Experience Form; WVI = Work Values Inventory; LKT = Leader Knowledge Test

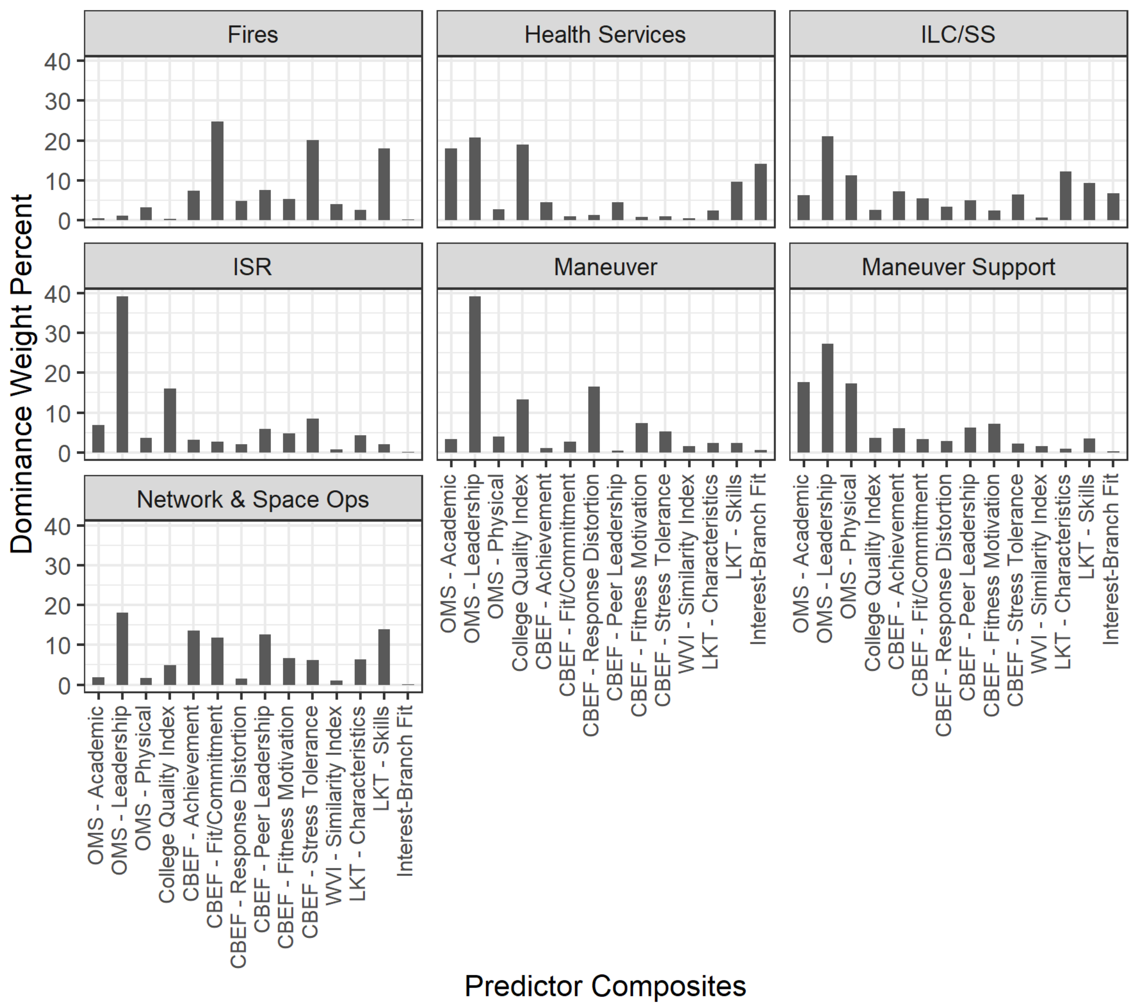

Figure 2. ***Dominance weight profiles by branch cluster.***

Behind OMS Leadership predictor, the strongest performance predictors include the College Quality Index ($M_{\text{dominance-weight}}$ = 8.54%, SD = 7.39%), the LKT Skills Score ($M_{\text{dominance-weight}}$ = 8.37%, SD = 6.15%), and the OMS Academic predictor ($M_{\text{dominance-weight}}$ = 7.77%, SD = 7.25%). The CBEF predictors had average dominance weight percentages ranging from 4.64% (Response Distortion) to 7.06% (Stress Tolerance). The WVI Profile Similarity Index ($M_{\text{dominance weight}}$ = 1.47%, SD = 1.25%) and Predicted Interest-Branch Fit ($M_{\text{dominance weight}}$ = 3.17%, SD = 5.38%) displayed the lowest relative strength among the predictor battery. Figure 2 graphically illustrates that the relative strength of individual predictors varied substantially across branch clusters.

Predictor Differentiation. Table 3 presents results from the predictor composite differentiation analyses. Overall, the average predicted score ($\hat{y}$) intercorrelation was r = .44, which represents the baseline differentiation level. For five predictors, the $\hat{y}$ intercorrelation increased (i.e., differentiation decreased) when removing the predictor composite of interest, suggesting that these predictors are valuable for differentiation among branch clusters. The predictors that displayed the most value in differentiation were LKT Characteristics

($M_{\hat{y}\text{ Intercorrelation } \textit{Excluding Predictor}}$ = .49, *SD* = .20), LKT Skills ($M_{\hat{y}\text{ Intercorrelation } \textit{Excluding Predictor}}$ = .49, *SD* = .17),[3] Predicted Interest-Branch Fit ($M_{\hat{y}\text{ Intercorrelation } \textit{Excluding Predictor}}$ = .46, *SD* = .20), and CBEF Response Distortion ($M_{\hat{y}\text{ Intercorrelation } \textit{Excluding Predictor}}$ = .46, *SD* = .20). Several excluded-predictor models displayed intercorrelation values clustering around the baseline value of .44.

Table 3. ***Predictor Differentiation: Predicted Score Intercorrelations across Branch Clusters***

Predictor	Average ŷ Intercorrelation Excluding Predictor (SD)	Average Intercorrelation Difference from Full Model	Rank
Overall Model (All Predictors)	.44 (.20)		
OMS Academic Component Score	.43 (.20)	-.01	11
OMS Leadership Component Score	.38 (.24)	-.06	14
OMS Physical Component Score	.44 (.21)	.00	10
College Quality Index	.45 (.21)	.01	5
CBEF Achievement	.44 (.20)	.00	8
CBEF Fit/Commitment	.43 (.21)	-.01	12
CBEF Response Distortion	.46 (.20)	.02	4
CBEF Peer Leadership	.44 (.20)	.00	9
CBEF Fitness Motivation	.44 (.20)	.00	6
CBEF Stress Tolerance	.42 (.23)	-.02	13
WVI Profile Similarity Index	.44 (.20)	.00	7
LKT Characteristics	.49 (.20)	.05	2
LKT Skills	.49 (.17)	.05	1
Predicted Interest-Branch Fit	.46 (.20)	.02	3

Note. Average ŷ intercorrelation excluding predictor reflects the average predicted score intercorrelation across all branch clusters, excluding the predictor of interest. Average intercorrelation difference simply subtracts this value from the overall model intercorrelation (r = .44).
Predictor Composite: OMS = Outcome Metric Score; CBEF = Cadet Background and Experience Form; WVI = Work Values Inventory; LKT = Leader Knowledge Test.

Overall Effectiveness Results

Table 4 presents overall classification effectiveness results. The predictor validity and predictor differentiation analyses were converted to *Z*-score metrics and averaged to form an overall classification effectiveness index. LKT Skills (Z = 1.94) and Characteristics (Z = 1.04) displayed the highest value for the classification algorithm, according to Brogden's Index. The OMS Leadership Component (Z = .81) was next, largely because of its high predictive validity on job performance. College Quality (Z = .46) also displayed an above average contribution to classification efficiency. We note that negative *Z*-scores do not mean negative contribution to the algorithm; rather, that a predictor with a negative value scored below average on classification efficiency, relative to the other predictors. Overall, evaluating the predictors in terms of their contribution to classification efficiency supports our effort to focus on and expand

[3] Due to concerns that the LKT Skills and LKT Characteristics scales' high correlation might have confounded results, we re-ran analyses excluding both LKT predictors. The average ŷ intercorrelation increased to r = .55. This result suggests that the intercorrelation did not confound results, but rather that both LKT predictors have a valuable differentiation effect on the classification results.

the role of the most promising predictors. This result led us to consider alternate scoring approaches for predictors ranked high on this metric (See Table 4).

Table 4. ***Overall Classification Effectiveness Results***

Predictor	Dominance Weight *Z*-Score	ŷ Intercorrelation *Z*-Score	Average *Z*-Score
LKT - Skills	.24	1.70	1.94
LKT - Characteristics	-.52	1.56	1.04
OMS - Leadership Component Score	3.21	-2.40	.81
College Quality Index	.27	.19	.46
CBEF Response Distortion	-.48	.53	.04
Predicted Interest-Branch Fit	-.77	.56	-.20
CBEF Achievement	-.19	-.02	-.21
OMS Academic Component Score	.12	-.39	-.27
CBEF Peer Leadership	-.21	-.15	-.36
OMS Physical Component Score	-.17	-.20	-.38
CBEF Fitness Motivation	-.43	.04	-.39
CBEF Fit/Commitment	.04	-.52	-.48
CBEF Stress Tolerance	-.02	-.89	-.91
WVI Profile Similarity Index	-1.09	.00	-1.10

Note. Dominance weight *Z*-score standardizes the mean dominance weight % column from Table 6.1. Similarly, ŷ intercorrelation *Z*-score standardizes the average ŷ intercorrelation from Table 6.2. Average *Z*-score is the average across columns one and two.
Predictor Composite: LKT = Leader Knowledge Test; OMS = Outcome Metric Score; CBEF = Cadet Background and Experience Form; WVI = Work Values Inventory

Limitations

The major limitation of our use of Brogden's Index to assess classification efficiency is that this method is indirect. We are not directly modeling the classification algorithm. Specifically, this approach does not account for branch preferences and rule-based placements according to cadet preference and OMS. Nonetheless, this approach addresses the effects of each predictor without imposing the constraints of the actual classification process.

Leader Knowledge Test and Work Values Inventory: Branch Specific Keying

Based on the above results, we conducted analyses to assess the potential of using branch-specific keys to enhance the predictive validity of the LKT and the WVI beyond the use of non-branch specific keys. We created branch-specific keys using officer data collected for a concurrent validity project (Russell et al., 2017). After creating the keys, we computed branch-specific shape scores for the officers in the Army Officer Classification project database (Ford et al., 2020). (Shape scores are described below.) Although sample size constraints did not permit evaluation of the predictive validity of scores based on these keys, we evaluated the differences in keys, key reliabilities, and shape scores by keying method to shed light on the potential usefulness of within-branch expert keying.

Leader Knowledge Test

The LKT was developed to assess understandings of implicit leadership theory (Legree et al., 2010). This test contains two subscales: LKT Characteristics and LKT Skills. The LKT uses a 10-point Likert-type scale and requires examinees to rate the importance of 30 characteristics and 30 skills for effective military leadership.

The LKT yields separate shape scores for the LKT Characteristics and Skills subscales with each shape score computed as the product moment correlation between an examinee's ratings and the values in the scoring key. The LKT scoring keys had been consensually developed based on the opinions of junior officers who were assigned to operational units (Legree et al., 2010).

Consensus keying procedures were originally developed to create scoring keys for emerging constructs that lack a well-developed knowledge corpus or readily available experts (Legree, 1996). These domains include social intelligence (Legree, 1996), emotional intelligence (Mayer, Caruso & Salovey, 1999), and tacit knowledge for military leaders (Hedlund et al., 2003). Consensus keying reflects expectations that initiate, journeyman, and expert opinion reflects both true and error variance and the assumption that expert opinion reflects greater levels of true variance and lower levels of error variance. Based on these assumptions, scoring keys based on the analysis of opinion data collected from a large number of initiates and journeymen should generally be more accurate than scoring keys based on a small number of experts. Analyses generally confirm these theoretical expectations (Legree et al., 2005).

However, analyses have not explored the possibility that implicit leadership theories may vary in the military context across officer branches. For example, this might occur if branches fundamentally differ in their requirements for demonstrating either direct leadership skills or technical expertise. Therefore, we explored the utility of using branch incumbents to refine the scoring keys for the LKT scales.

Work Values Inventory

The WVI uses a two-step process in which examinees are asked to: (1) rank 11 job characteristics in order of desirability, and (2) select all characteristics from the same list that would need to be present in their ideal job. The WVI is scored as the rank order correlation between an examinee's responses and an expert response profile. We also explored the use of branch specific keying to enhance the utility of the WVI.

Methods and Results

To conduct these analyses, we identified a concurrent validation (CV) dataset that could be used to create alternate keying samples for the LKT and WVI (Russell et al., 2017). This CV sample include 875 active duty Army officers from U.S. Army Forces Command (FORSCOM), and it contained sufficient quantities of useable data for: the LKT, $n = 620$; and the WVI, $n = 640$.

The CV dataset contained substantial variance in officer rank and an indicator of officer performance that might affect response patterns on the LKT and WVI. Therefore, we also extracted a subsample of experienced officers that contained all officers ranked captain or higher (ranks O3 – O6) for the LKT (*n* = 210) and WVI (*n* = 212). We also identified a high performance officer sub-sample that included all officers performing at the 50th percentile or higher in general performance for the LKT (*n* = 297) and WVI (*n* = 305). We explored other expert samples based on specific facets of performance (leadership performance, technical performance), but found that scoring based on these samples produced very similar LKT and WVI scores as the overall CV sample. Therefore, we report results for three keying samples: the overall sample, the experienced officer sample, and the high performance sample.

We then divided the keying samples into branches and calculated consensus keys based on the mean response for each item within each sample. We combined data for any branch sample with fewer than 10 officers into its corresponding branch cluster to ensure a sufficient number of raters (see Figure 1). Upon calculation of key vectors, we examined the intercorrelation between the 16 branches' key vectors for each measure. High intercorrelations across key vectors suggested that within-branch consensus keying is not likely to produce differentiated LKT or WVI scores. We also examined the reliability of the scorings keys by using the intraclass correlation coefficient (ICC). Specifically, we used ICC2*k* because there was a fixed sample of raters and we were interested in consistency across raters (Landers, 2015).

For the LKT Characteristics and Skills subscales, the derived across-branch keys were highly correlated, and the reliabilities were similarly high for the LKT scales. This pattern of results suggests that officers were responding to the LKT similarly across branches and provides initial evidence that within-branch keying of the LKT is unlikely to prove beneficial. Table 5 reports mean key intercorrelations and ICC results by measure and keying sample.

In contrast, the key intercorrelations and rater reliability estimates were much lower for the WVI. This pattern of results suggests that there may not be sufficient rater agreement for the WVI to support branch-specific keying.

Table 5. *LKT and WVI Expert Key Results by Keying Sample*

Sample	LKT-Characteristics	LKT-Skills	WVI
	Mean (SD)	Mean (SD)	Mean (SD)
	Key Intercorrelation		
Overall CV Sample	.96 (.03)	.91 (.10)	.39 (.29)
Experienced Officer Sample	.93 (.07)	.90 (.08)	.18 (.31)
Performance 50%+ Sample	.91 (.08)	.88 (.13)	.28 (.26)
	Intraclass Coefficient (ICC2k)		
Overall Expert Sample	.98 (.02)	.96 (.03)	.69 (.13)
Experienced Officer Sample	.94 (.06)	.91 (.06)	.49 (.24)
Performance 50%+ Sample	.96 (.02)	.93 (.04)	.55 (.27)

Note. Across 16 branches, branch key sample size ranged from 4 to 157 (*M* = 39, *SD* = 37). When branch sample size was less than 10, the sample was aggregated to the branch cluster for larger *n*. One branch cluster *n* was still less than 10.

We then applied the branch keys to cases in the Army Officer Classification sample according to officer branch; these data were collected when the officers had been cadets (Ford et al., 2020). We calculated shape scores for the LKT and WVI by correlating each branch key with the cadet's response for each measure. Following this approach, a cadet has four alternate scores for each measure. These four scores were based on:

- Non-branch specific keying;
- Branch specific keying based on the overall sample;
- Branch specific keying based on the experienced officer sample; and
- Branch specific keying based on the high performance sample.

Table 6 presents intercorrelations among scores using these alternate keying procedures. These results indicate that the alternate LKT scores were highly correlated across the four keying methods. This result indicates that branch-specific keying is unlikely to add to predictor validity or differentiation for the LKT. However, we do know from the results of the classification algorithm examination that the LKT was valuable for officer classification (See Table 4). Therefore, we tentatively conclude that knowledge of leadership attributes is useful *across* rather than *within* branches when classifying officers. We outline another method for scoring the LKT in the next section.

Table 6. *LKT and WVI Intercorrelations by Keying Methods*

	Non-specific Key	Overall Expert Key	Experienced Officer Key
		LKT-Characteristics	
Overall Expert Key	.99		
Experienced Officer Key	.98	.99	
Performance 50%+ Key	.98	.99	.99
		LKT-Skills	
Overall Expert Key	.97		
Experienced Officer Key	.95	.97	
Performance 50%+ Key	.96	.99	.96
		Work Values Inventory	
Overall Expert Key	.33		
Experienced Officer Key	.40	.66	
Performance 50%+ Key	.32	.69	.56

Note. All correlations are uncorrected. For the LKT and WVI, pairwise *n*'s ranged from 4,291 – 4,512 and 4,322 – 4,512, respectively.

In contrast to the results for the LKT scales, the alternate WVI score correlations were low to moderate by keying method. This result is more supportive of attempts to differentiate scores by keying method to provide beneficial results. In addition, correcting for key unreliability using the mean WVI reliability values in Table 5, correlation estimates of $\rho = .40$, .57, and .43 were found between the non-specific key and the overall expert, experienced officer, and performance key samples. Further research examining criterion-related validity is necessary to address whether branch-specific keying is beneficial for the WVI.

Profile Similarity Metric Scoring

For the LKT, we also examined profile similarity metric (PSM) scoring to explore whether additional scoring methods improve the prediction of job performance across branches. Based on this initial analysis, one might not expect large improvements in validity from PSM scoring when predicting officer performance. However, Legree et al. (2020) found PSM scoring to produce substantial validity improvements from traditional distance-based metrics. Overall, results suggest that PSM scoring is worthy of further exploration after administration of the CAB. See Appendix D for an in-depth discussion of this analysis and results.

Vocational Interests

As described above, vocational interests are a useful predictor of performance and continuance (Van Iddekinge et al., 2011) and hold promise for branch classification (Kirkendall et al., 2020). Therefore, we reanalyzed data from the Army Officer Classification/officer branch assignment project to examine the utility of Army officers' vocational interest profiles for predicting outcomes relevant to performance and continuance, and explore their utility for branch classification.

First, we examined the validity of RIASEC scale scores inferred from the officer's college major. Second, we created a person-cluster fit index by computing the correlation between each officer's RIASEC scores and the RIASEC scores of each branch cluster. Finally, we calculated another series of fit indices using a regression-based approach that focused on predicting specific criteria.

The results indicated modest prediction; however, there are a few caveats to this conclusion. First, the analyses were conducted at the branch cluster level; therefore, lack of homogeneity of interest profiles within clusters contributed error to the analyses. Second, the officer RIASEC scores were inferred from college majors rather than self-reported. Given the coarseness and unknown construct validity of the predictor scores, achieving any prediction at all is encouraging. Therefore, it is reasonable to assume that interest scores collected from a validated RIASEC inventory might perform substantially better, and it is possible that the two sources of information combined will predict better than either one could alone. In addition, accumulating larger sample sizes should improve our ability to construct regression-based congruence models without overfitting the data. Therefore, follow-up research is planned to expand on these initial results. Appendix D provides an in-depth discussion of these analyses and results.

Chapter 5: Initial Recommendations for Predictor Constructs

Emily S. Medvin and Laura A. Ford

We synthesized findings from the research streams described previously (e.g., literature review, focus groups, additional analyses) to identify the most promising attributes for differentiation in performance. A brief summary of the rationale for each of these initial attribute recommendations is provided.

Cognitive Abilities

General mental ability has consistently been shown to be a strong predictor of performance in a variety of contexts (Schmidt & Hunter, 1998). In analysis of predictors in the officer branch assignment project, mental ability, as represented by the Academic Component Score was the fourth most important predictor (Ford et al., 2020). While this scale did not show promising differentiation results in this analysis, previous research has shown that general mental ability is a stronger predictor of performance for jobs higher in complexity (Hunter & Hunter, 1984). Because some branches require more cognitively taxing, technical training and work, general mental ability could show differentiation in the prediction of branch-specific performance.

The Memory attribute was reviewed by focus group participants. While Memory was endorsed by participants from certain branches (e.g., Ordnance and Field Artillery), it was not endorsed by participants from other branches. Furthermore, Memory was identified by focus group participants as among the least important contributors to successful performance for the Transportation branch. Therefore, Memory may hold potential to differentially predict branch performance.

Situational awareness was identified as critical for many branches by the Talent Management job analyses. However, this attribute was rated as less important for several branches (e.g., Cyber, Signal). This range in criticality assessment was also supported by our focus group data (i.e., this attribute was endorsed by nearly all of participants for the Field Artillery and Infantry branches, but was not identified as critical by participants for the Quartermaster and Transportation branches).

Spatial Orientation ability was one of the specific abilities found to maximize mean predicted performance of enlisted populations (Sager et al., 1997). While this ability was not included in our focus group protocols, because it did not meet the critical threshold for any branch, the standard deviation in importance ratings for Spatial Orientation from the Talent Management job analysis survey was large (1.60; see Appendix B), thus showing promise as a differentiating predictor.

Personality Traits and Tendencies

Composure, Industriousness, and Politeness are aspects of the Big Five personality traits of Emotional Stability, Conscientiousness, and Agreeableness. Each of these attributes was reviewed by focus group participants and was endorsed as critical in some branches, but not others. Politeness was heavily endorsed as critical in the Field Artillery branch, but was identified as among the least important attributes for the Ordnance and Transportation branches. Industriousness was endorsed in several branches (e.g., Artillery, Infantry, Ordnance), but not others (e.g., Field Artillery, Quartermaster). Similarly, Composure was endorsed in some branches (e.g., Air Defense, Artillery), but not others (e.g., Ordnance, Transportation).

Initiative was included in the list of "potential other attributes" and was endorsed in all branches, but to varying degrees. Four participants in Air Defense and three participants in Field Artillery endorsed Initiative as critical, but only one participant in each Logistics branch (i.e., Ordnance, Quartermaster, and Transportation) endorsed this attribute.

Learning Orientation was included in the list of "potential other attributes" in the focus groups and was heavily endorsed by some branches (e.g., Armor, Field Artillery) but not others (e.g., Ordnance, Quartermaster).

Tolerance for Injury was also included in the list of "potential other attributes" in the focus groups. It was endorsed as critical by participants in some branches (e.g., Field Artillery) but was identified as among the least important in other branches (e.g., Ordnance).

Vocational Interests and Knowledge

Leadership orientation, in terms of accuracy of individuals' beliefs about important characteristics and skills for an Army officer to have, showed the most differentiation value in the Army Officer Classification algorithm analysis (Ford et al., 2020). This value should increase as we create branch-specific keys to more accurately define what leadership entails in each branch (see Chapter 4).

Vocational interests, or patterns of preferences for the type of work that individuals want to do, were one of the top differentiating predictors in the Army Officer Classification algorithm analysis (Ford et al., 2020). Additionally, three interests in the RIASEC model (Conventional, Enterprising, and Realistic) showed differentiation promise from the focus group data, being endorsed as critical in some branches (e.g., Conventional and Enterprising in Transportation) and being identified as among the least important for success in others (e.g., Enterprising and Realistic in Infantry).

Work Values, in terms of the match between an individual's values and those provided in Army officer positions (e.g., selfless service, recognition, pay), showed moderate differentiation value in the Army Officer Classification algorithm analysis (Ford et al., 2020). However, individual scores in this analysis were based on similarity with the values the Army provides overall, not by branch. Because branches do differ on some of the values assessed (e.g., leadership opportunities, variety, comfortable work environment), our creation of branch-specific keys for work values may result in added differentiation value (see Chapter 4).

Skills

Social Sensitivity was found in the Talent Management job analysis to be critical in some branches (e.g., Adjutant General, Transportation), but not other branches. The standard deviation in the importance ratings for this attribute across branches was 1.30. This differentiation value was supported by our focus group findings. Social sensitivity was endorsed as critical by participants in some, but not all, branches, and was also identified as among the least important attributes in the Quartermaster branch.

Written communication, while identified as critical for all branches in the Talent Management job analysis, still had a relatively large standard deviation of its importance ratings across branches (SD = 0.92). Additionally, participants in some branches (e.g., Air Defense, Armor, Ordnance) agreed that this attribute is critical, while participants in other branches (e.g., Infantry, Quartermaster) identified it as among the least important.

Preliminary Cadet Assessment Battery Map

In Table 7, we present a test map of these initial recommendations, as well as potential measures for each and practical considerations of these measures (e.g., administration mode, administration time).

Table 7. ***Cadet Assessment Battery Test Map***

	Potential Predictor Measures															
Attribute	GPA	SAT/ACT Scores	Working Memory Measure (e.g., Mental Counters)	SAGAT	Map Test	Big Five Aspects Scale	Proactive Personality Scale (short version)	General Learning and Performance Goal Orientation Scale	CBEF	LKT	AVID	College Major (coded to RIASEC)	WPS	O*NET Interest Profiler	WVI	Background Data Measure of Social Intelligence
Cognitive Abilities																
General Mental Ability	X	X														
Memory			X													
Situational Awareness				X												
Spatial Orientation					X											
Personality Traits and Tendencies																
Composure						X										
Industriousness						X										
Politeness						X										
Initiative							X									
Learning Orientation								X								
Tolerance for Injury									X							
Vocational Interests and Knowledge																
Leadership Orientation										X						
Vocational Interests											X	X	X	X		
Work Values															X	
Skills																
Social Sensitivity																X
Written Communication									X							
Administration Properties																
Estimated Time (Min)	NA	NA	30	TBD	12	10	3	3	5	15	10	NA	15	15	15	15
Mode	Arch.	Arch.	Cmp.	Sim.	P/P	Any	Any	Any	Any	Any	Any	Arch.	Any	Any	Any	Any

Note. Potential Measures: GPA = Grade Point Average, SAGAT = Situation Awareness Global Assessment Technique (Endsley, 1998), CBEF = Cadet Background and Experience Form, LKT =Leader Knowledge Test, AVID = Army Vocational Interest Diagnostic (Nye et al., 2018), WPS = Work Preferences Survey (Van Iddekinge et al., 2011), WVI = Work Values Inventory.
Administration Mode: Arch. = Archival, Cmp. = Computer-Based, Sim. = Simulation, P/P = Paper/Pencil.

CHAPTER 6: FINAL CADET ASSESSMENT BATTERY

Emily S. Medvin, Laura A. Ford, Brenda Ellis, Teresa L. Russell, Peter J. Legree, Robert N. Kilcullen, and Mark C. Young

To finalize the predictor constructs and corresponding measures to include in the CAB, we reviewed conceptual and empirical support for the prediction and differentiation value of the attributes, as well as the feasibility of administering measures given our data collection requirements and constraints. Based on this review, some attributes were removed from consideration. For example, we confirmed that situational awareness is best measured via simulations that are frozen at randomly selected times while respondents answer questions regarding their perception of the situation at that point in time (e.g., Situation Awareness Global Assessment Technique - SAGAT). Unfortunately, such simulations are not possible given our data collection constraints. Therefore, situational awareness was excluded from the CAB.

We also considered any conceptual overlap between the constructs. As Brogden indicates, classification efficiency increases as the predictors of performance are less correlated (Brogden, 1959). Regarding our list of potential constructs, working memory has been shown to be highly correlated with high-level cognitive abilities such as reasoning and problem-solving (Kyllonen & Christal, 1990). Due to this overlap, we compared the feasibility of measures of working memory to those of general mental ability, and elected to exclude working memory from the CAB, in favor of including the more-easily collected measures of general mental ability (e.g., self-reported GPA and test scores).

The final attributes included in the CAB are listed below, along with a description of the associated measure(s) and any development efforts that were taken to finalize the measure.

Cognitive Abilities

General Mental Ability

General mental ability will be measured using a set of metrics from both administrative records and self-reported information. The first measure will be grade point average (GPA) from the undergraduate institution (provided by USACC). However, because cadets attend hundreds of different colleges with different academic standards, college quality will be used to correct GPA. Previous research has found GPA corrected for college quality is a better predictor of performance than GPA alone (e.g., Koch et al., 2013). We will follow the approach used by Koch and colleagues (2013) to create the College Quality Index.

Other metrics of general mental ability will be self-reported SAT and ACT scores. Analyses have demonstrated that SAT and ACT scores can be converted to a measure of *Psychometric g* (Frey & Detterman, 2004). We also note that the CAB is being specifically developed for research purposes. If measures of general mental ability prove to be useful for classification purposes, future iterations of the CAB will likely include additional maximum performance measures.

Spatial Orientation

To measure spatial orientation, we will use the Map Test that was originally administered in Project A (Russell et al., 2001). In Project A, the map test showed acceptable levels of reliability (Split half = .90; Test-retest = .78). Additionally, this measure represents a replacement for land navigation assessments that were removed from the U.S. Army Cadet Command Advanced Camp field training exercise (FTX). To facilitate potential future computer administration, we converted the paper-based graphics in the Project A Map Test to an electronic format for electronic administration.

Personality Traits and Tendencies

Composure, Industriousness, and Politeness

To measure these three aspects of Big Five personality traits, we opted to use the three corresponding scales (i.e., Volatility, Industriousness, and Politeness) included in the public domain Big Five Aspects Scale (DeYoung, Quilty, & Peterson, 2007). Each scale contains 10 items and have been shown to have acceptable reliabilities (Volatility Alpha = .85; Industriousness Alpha = .81; Politeness Alpha = .75).

Initiative

Existing measures of initiative or overlapping constructs (e.g., proactive personality; Bateman & Crant, 1993) are either not in the public domain or do not sufficiently differentiate between initiative and its strong correlates (e.g., achievement orientation, industriousness, self-efficacy; Bateman & Crant, 1993; Borman et al., 2001; Frese et al., 1997; Roberts et al., 2005).

Using the Bateman and Crant (1993) measure of proactive personality as a marker measure during initial data collection, we began to develop an initiative scale for Army use. To this end, project staff generated and reviewed a pool of items and prepared them for pilot testing in the initial data collection. The new items use the same Likert-type response agreement scale as the marker measure. The data from these final items will be used to develop the final initiative scale for Army use.

A team of internal item writers used the marker measure items, the definition of initiative used in our prior focus group data collection, and the description of the marker measure construct (Bateman & Crant, 1993), as guidance in developing a pool of 45 items. These items were initially reviewed for grammar, bias, and face validity. A second team of project SMEs then reviewed the items for construct mapping. We presented the SMEs with the drafted items and a set of seven constructs: the target construct and six related constructs. SMEs rated each item on how strongly it appeared to tap each construct on a scale of 0 (not at all) to 4 (very much).

We then refined the item pool via a series of steps. First, we used item ratings to remove items that did not sufficiently tap the target construct. Items were excluded if they (a) received an average rating of 2 or less on mapping to the target construct or (b) had an average rating on the mapping to the target construct that was less than the average of the ratings for the mapping to all other constructs, excluding the constructs rated as 0. For example, if an item tapped only three of the six related constructs, the average rating for that item was the sum of those three ratings divided by three. This method ensured that retained items were rated as (a) tapping the target construct more than "somewhat" and (b) tapping the target construct at least as heavily as the average of the other constructs. This process resulted in retaining 30 items.

After completing the steps described above, some items were still retained that tapped other constructs more heavily than the target construct. For example, for an item with a target rating of 3.25 and a set of related constructs ratings of 1, 4, and 4, the average rating on the related constructs would be 3.0. Thus, the item would be retained because its target rating (3.25) exceeds the average rating to related constructs. However, this item taps two related constructs more heavily than it is tapping the target construct.

Thus, as a next step, we removed items if they had an average rating on a related construct that was more than 0.5 greater than the average rating on the target construct. Given the example above, an item with a target rating of 3.25 and even one related construct rating of 4.0 would be excluded. This left 26 viable items. Incidentally, all remaining items had average ratings of 3 or greater on the target construct.

Finally, we assessed items for social desirability. Socially desirable responding is a response tendency in which a person endorses items to present a positive image (Paulhus, 2002). Items that are deemed highly socially desirable may entice respondents to strongly endorse them, regardless of how well the respondent believes the item reflects their own personality. Because initiative is generally seen as a desirable trait, it was important to assess the degree to which each item would be subject to socially desirable response bias. Items with an average rating greater than 3 on social desirability were removed in an attempt to mitigate the effects from this form of response bias. One item was removed at this step, leaving 25 final items.

Given that pilot response data analyses will yield information with which to further refine the scale, we erred on the side of leaving more items in than the number anticipated in the final operational scale, while keeping in mind survey administration time constraints. The 25 developed items will be piloted in the initial CAB survey data collection, along with the marker measure. Once sufficient data have been collected for an initial data extract, we will conduct item and scale analyses and further refine the item set for the final Initiative scale.

Learning Orientation

As with Initiative scales, existing learning orientation scales are either not in the public domain (e.g., Vandewalle, 1997), have been developed for student populations rather than for working adults (e.g., mastery orientation, Elliot & McGregor, 2001), or tap intellectual pursuit rather than broader learning orientation (e.g., love of learning, http://ipip.ori.org). Thus, we followed the same procedure used for initiative to develop items for learning orientation, using Vandewalle's (1997) learning goal orientation scale as a marker measure.

Internal staff generated and reviewed a pool of items and prepared them for pilot testing in the initial data collection. The new items use the same Likert-type response agreement scale as the marker measure. Using items from Vandewalle's (1997) scale as a reference, as well as the definition of learning orientation used in our prior focus group data collection and Vandewalle's (1997) description of the marker measure construct, item writers generated 35 items for construct mapping.

A team of staff researchers then rated the items on how well they tapped the target construct as well as six related constructs. We used the same scoring method for these items as for the Initiative items. After removing items which failed the first two scoring rules, 24 items remained. After removing any remaining items for which the average rating on any related construct was more than 0.5 greater than the average rating on the target construct, 22 items remained. After removing items with average social desirability ratings greater than 3, 21 items were retained. Given that the marker measure consists of six items, we refined our items further by removing any items that had an average target construct mapping rating less than 3. This removed an additional 2 items, resulting in a final item set of 19 items.

The set of items will be included in the CAB along with the marker measure. After item analyses from pilot data and convergent validity analyses with the marker measure, the items will be further refined into a learning orientation scale for the Army.

Tolerance for Injury

Tolerance for Injury is a CBEF attribute that was identified as particularly promising for differentiation. Because this measure has been consistently administered in Advanced Camp FTX, we will use the previously collected data for this project, merging existing data for this project's sample into the current database.

Vocational Interests and Knowledge

Leadership Orientation

Leadership Orientation will be assessed using the LKT. As mentioned previously, the LKT is 60-item test which assesses an officer's implicit leadership theories, which are beliefs and assumptions about the attributes necessary for effective leadership (Legree et al., 2010). Respondents rate on a 10-point Likert-type scale the relative importance of characteristics and skills for military leadership (e.g., Legree et al., 2010). Because the LKT has not been consistently administered in Advanced Camp FTX since 2013, we will administer it as part of our data collection effort for this project.

Vocational Interests

We identified six candidate interest measures that the Army could easily access: O*NET Interest Profiler (IP), Interest Item Pool (IIP) Basic Interest Markers (BIM), IIP RIASEC Markers, AVID, Army Vocational Interest Career Examination (AVOICE), and WPS. We evaluated each based on psychometric and qualitative criteria. While all measures have reasonably good psychometric properties, they differ in the level of specificity of the scales. The O*NET IP, IIP RIASEC Markers, and WPS all yield scores for each of the six RIASEC factors. However, the RIASEC level is likely not specific enough to provide sufficient discrimination across branches. Therefore, we focused on the instruments that provided more specific basic interest scales: IIP BIM, AVID, and AVOICE. Because the AVID is an updated version of the AVOICE, we eliminated the AVOICE from consideration.

We then compared the IIP BIM and AVID scales. While there was much overlap between the scales, some AVID scales appeared at a lower level than would be relevant for an Officer population (e.g., Food Service). Additionally, the AVID scales were more heavily weighted to the Realistic facet in the RIASEC model. We also noted that the IIP BIM included more scales that could potentially differentiate between branches. For example, while the AVID has one overall "science" scale, the IIP BIM makes distinctions between physical, life, and social sciences, which are likely differentially related to occupations.

Therefore, we decided to move forward using the IIP BIM as the starting point for our measure of vocational interests in the CAB (i.e., the Preferred Activity Questionnaire [PAQ]). However, because no Combat scale is included in the IIP BIM, we culled combat-related items from the AVOICE and AVID Combat scales as a starting point for an additional scale for the CAB. Because the AVOICE/AVID Combat items were targeted to enlisted soldiers, we created a Military Operations scale with activities more aligned with Officer tasks.

Our final step was to prune and finalize our items. To this end, we first selected the items within each scale that were most relevant to the military context. We also compared similar scales on the AVOICE and IIP BIM and made changes for consistency and relevance. For example, we replaced the IIP BIM Outdoors-Agriculture scale with items from the AVOICE Outdoors scale, because the IIP BIM was too agriculture-focused. We also merged some scales into more overarching categories. For example, we merged Business, Office Work, and Management into a Business Administration category.

To confirm that the number of items included in the PAQ could be completed in a reasonable amount of time relative to our overall data collection administration time, we conducted an internal timing trial in which we asked three individuals to answer each item, arranged in a random order, and report the time required to do so. Individuals reported an average of 11.67 minutes to complete the PAQ, therefore all items were retained.

Additionally, we plan to create the predicted interest-branch fit variable described earlier and used for the Army Officer Classification project analyses (Ford et al., 2020). That variable infers interests based on the officer's college major. We plan to investigate the relationship between interests as measured by a 15-minute survey (the PAQ) and interests inferred by college major.

Work Values

Work Values will be assessed using the WVI. As mentioned previously, the WVI is a two-item survey which asks respondents to rank a list of job characteristics in order of desirability, and then asks respondents to select all characteristics from the same list that would need to be present in their ideal job. Because the WVI has not been consistently administered in Advanced Camp FTX since 2013, we will administer it as part of our data collection effort for this project.

Skills

Written Communication

Written Communication is a CBEF attribute that was identified as particularly promising for differentiation. Because this measure has been consistently administered in Advanced Camp FTX, we will use the previously collected data for this project, merging existing data for this project's sample into the current database.

Summary

The attributes and measures described above represent the CAB, which we developed to add classification differentiation beyond what is currently available (i.e., not contained in readily available predictor batteries). However, when evaluating classification, we will also analyze available data that we know is useful for classification (e.g., OMS, CBEF measures). In the next phase of the project, we will administer the CAB to junior officers and assess measure performance and concurrent validity.

REFERENCES

Baldwin, S., & Young, M.C. (Eds.). (2020). *Research on the Cadet Background and Experience Form (CBEF) to Support Army ROTC Personnel Assessment (2018-2019)* (Technical Report 1358). U.S. Army Research Institute for the Behavioral and Social Sciences. (DTIC No. AD1109745).

Bateman, T. S., & Crant, J. M. (1993). The proactive component of organizational behavior: A measure and correlates. *Journal of Organizational Behavior, 14*(2), 103–118. https://doi.org/10.1002/job.4030140202

Bayer, J., McLean, K., & Salyer, S. (2017). *Careers in the military taxonomy and website update report* (Unpublished manuscript). ASVAB Career Exploration Program, Defense Personnel Assessment Center, Office of People Analytics, Department of Defense.

Borman, W. C., Penner, L. A., Allen, T. D., & Motowidlo, S. J. (2001). Personality predictors of citizenship performance. *International Journal of Selection and Assessment, 9*(1/2), 52–69. https://doi.org/10.1111/1468-2389.00163

Brogden, H. E. (1959). Efficiency of classification as a function of number of jobs, percent rejected, and the validity and intercorrelation of job performance estimates. *Educational and Psychological Measurement, 19*(2), 181–190. https://doi.org/10.1177/001316445901900204

Budescu, D. V. (1993). Dominance analysis: A new approach to the problem of relative importance of predictors in multiple regression. *Psychological Bulletin*, *114*(3), 542–551. 10.1037/0033-2909.114.3.542

Bynum, B. H., & Young, M. C . (Eds.). (2020). *Validation of the Cadet Background and Experience Form (CBEF) to Support Army ROTC Personnel Assessment (2015-2018)* (Technical Report 1383). U.S. Army Research Institute for the Behavioral and Social Sciences. (DTIC No. AD1107158).

DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology, 93*(5)*,* 880-896. https://doi.org/10.1037/0022-3514.93.5.880

Edwards, J. R. (1993). Problems with the use of profile similarity indices in the study of congruence in organizational research. *Personnel Psychology, 46*(3), 641–665. http://dx.doi.org/10.1111/j.1744-6570.1993.tb00889.x

Elliot, A. J., & McGregor, H. A. (2001). A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology, 80*(3), 501–519. 10.1037/0022-3514.80.3.501

Endsley, M. R., Selcon, S. J., Hardiman, T. D., & Croft, D. G. (1998). A comparative analysis of SAGAT and SART for evaluations of situation awareness. *Proceedings of the Human*

Factors and Ergonomics Society Annual Meeting, 42(1), 82–86. https://doi.org/10.1177/154193129804200119

Farmer, W. L., Fedak, G. E., Bearden, R. M., Lightfoot, M. A., McBride, J. R., Heggestad, E. D., Alley, W. E., Harmon, L. W., & Rounds, J. (2006). *Navy interest inventory: Approach development* (No. NSPRST-TN-07-04). Navy Personnel Research, Studies, and Technology Division (NPRST/PERS-1) Bureau of Naval Personnel.

Ford, L. A., Russell, T. L., Legree, P. J., Kilcullen, R. N., & Young, M. C. (In preparation). *Investigating Army officer classification: Validation and utility of pre-commissioning metrics to enhance junior officer performance* (Technical Report). U.S. Army Research Institute for the Behavioral and Social Sciences.

Frese, M., Fay, D., Hilburger, T., Leng, K., & Tag, A. (1997). The concept of personal initiative: Operationalization, reliability and validity in two German samples. *Journal of Occupational and Organizational Psychology, 70*(2), 139–161. https://doi.org/10.1111/j.2044-8325.1997.tb00639.x

Frey, M. C., & Detterman, D. K. (2004). Scholastic assessment or g? The relationship between the scholastic assessment test and general cognitive ability. *Psychological Science, 15*(6), 373-378. http://dx.doi.org/10.1111/j.0956-7976.2004.00687.x

Hastie, T., Tibshirani, R., & Friedman, J. (2008). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer Science & Business Media.

Hedlund, J., Forsythe, G. B., Horvath, J. A., Williams, W. M, Snook, S., & Sternberg, R. J. (2003). Identifying and assessing tacit knowledge: Understanding the practical intelligence of military leaders. *The Leadership Quarterly*, *14*(2), 117–140. http://dx.doi.org/10.1016/S1048-9843(03)00006-7

Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments* (3rd ed.). Psychological Assessment Resources.

Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin, 96*(1), 72–98. http://dx.doi.org/10.1037/0033-2909.96.1.72

Johnson, J. W., & LeBreton, J. M. (2004). History and use of relative importance indices in organizational research. *Organizational research methods*, *7*(3), 238–257. http://dx.doi.org/10.1177/1094428104266510

Judge, T. A., Rodell, J. B., Klinger, R. L., Simon, L. S., & Crawford, E. R. (2013). Hierarchical representations of the five-factor model of personality in predicting job performance: Integrating three organizing frameworks with two theoretical perspectives. *Journal of Applied Psychology, 98*(6), 875–925. http://dx.doi.org/10.1037/a0033901

Kirkendall, C. D., Christopher, C. D., Rounds, J., Drasgow, F., Chernyshenko, O. S., & Stark, S. (2020). Adaptive vocational interest diagnostic: Informing and improving the job assignment process. *Military Psychology*, *32*(1), 91–100. http://dx.doi.org/10.1080/08995605.2019.1652480

Koch, A. J., Bynum, B. H., Young, M. C., Babin, N. E., & Russell, T. L. (2013, April 10-13). *Predictive validity of undergraduate GPA adjusted for college quality* [Poster presentation]. Society for Industrial & Organizational Psychology 28th Annual Conference, Houston, TX, United States.

Kyllonen, P.C., & Christal, R.E. (1990). Reasoning ability is (little more than) working-memory capacity?!. *Intelligence, 14*(4), 389-433. https://doi.org/10.1016/S0160-2896(05)80012-1

Landers, R. N. (2015). Computing intraclass correlations (ICC) as estimates of interrater reliability in SPSS (*The Winnower 2*:e143518.81744) [Computer software]. DOI: 10.15200/winn.143518.81744

Legree, P. J. (1995). Evidence for an oblique social intelligence factor established with a Likert-based testing procedure. *Intelligence, 21*(3), 247–266. http://dx.doi.org/10.1016/0160-2896(95)90016-0

Legree, P. J., Kilcullen, R. N, Psotka, J., Putka, D., & Ginter, R. N. (2010). *Scoring situational judgment tests using profile similarity metrics* (Technical Report 1272). U.S. Army Research Institute for the Behavioral and Social Sciences. (DTIC No. ADA530091).

Legree, P. J., Kilcullen, R. N., Putka, D. J., & Wasko, L. E. (2014). Identifying the leaders of tomorrow: Validating predictors of leader performance. *Military Psychology, 25*(4), 292-309. https://doi.org/10.1037/mil0000054

Legree, P. J., Ness, A. M., Kilcullen, R. N., & Koch, A. J. (2020). *Enhancing the Validity of Rating-Based Tests* (Technical Report 1371). U.S. Army Research Institute for the Behavioral and Social Sciences. (DTIC No. AD1067951).

Legree, P. J., Psotka, J., Tremble, T., & Bourne, D. R. (2005). Using consensus based measurement to assess emotional intelligence. In R. Schulze & R. D. Roberts (Eds.), *Emotional Intelligence: An International Handbook* (pp. 155–180). Hogrefe & Huber.

Legree, P. J., Purl, J., Kilcullen, R. N., & Young, M. C. (2019). *Cadet training and personality metrics predict officer in-unit performance: R= .37* (Technical Report 1377). U.S. Army Research Institute for the Behavioral and Social Sciences. (DTIC No. AD1081235).

Mayer, J. D., Caruso, D. R., & Salovey, P. (1999). Emotional intelligence meets traditional standards for an intelligence. *Intelligence, 27*(4), 267-298. https://doi.org/10.1016/S0160-2896(99)00016-1

Navarrette, C. B., & Soares, F. P. (2020). *Dominanceanalysis R package.* https://cran.r-project.org/web/packages/dominanceanalysis/dominanceanalysis.pdf

Nye, C. D., Butt, S. M., Bradburn, J., & Prasad, J. (2018). Interests as predictors of performance: An omitted and underappreciated variable. *Journal of Vocational Behavior, 108,* 178–189. http://dx.doi.org/10.1016/j.jvb.2018.08.003

Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 49–69). Lawrence Erlbaum Associates.

Paullin, C., Legree, P. J., Sinclair, A. L., Moriarty, K. O., Campbell, R. C., & Kilcullen, R. N. Russell, T. L. (2014). Delineating officer performance and its determinants. *Military Psychology, 26*(4), 259-277. https://doi.org/10.1037/mil0000051

Paullin, C., Sinclair, A. L., Moriarty, K. O., Vasilopoulos, N. L., Campbell, R. C., & Russell, T. L. (2011). *Army officer job analysis: Identifying performance requirements to inform officer selection and assignment* (Technical Report 1295). U.S. Army Research Institute for the Behavioral and Social Sciences. (DTIC No. ADA548872).

Roberts, B. W., Chernyshenko, O. S., Stark, S., & Goldberg, L. R. (2005). The structure of conscientiousness: An empirical investigation based on seven major personality questionnaires. *Personnel Psychology, 58*(1), 103–139. https://doi.org/10.1111/j.1744-6570.2005.00301.x

Rosse, R., Campbell, J., & Peterson, N. (2001). Personnel classification and differential job assignments: Estimating classification gains. In J. P. Campbell & D. J. Knapp (Eds.), *Exploring the limits in personnel selection and classification* (pp. 453-506). Lawrence Erlbaum Associates.

Russell, T. L., Paullin, C. J., Legree, P. J., Kilcullen, R. N., & Young, M. C. (2017). *Identifying and validating selection tools for predicting officer performance and retention* (Research Note 2017-01). U.S. Army Research Institute for the Behavioral and Social Sciences. (DTIC No. AD1038674).

Russell, T. L., Peterson, N. G., Rosse, R. L., Hatten, J. L. T., McHenry, J. J., & Houston, J. S. (2001). The measurement of cognitive, perceptual, and psychomotor abilities. In J. P. Campbell & D. J. Knapp (Eds.), *Exploring the limits in personnel selection and classification* (pp. 71–109). Lawrence Erlbaum Associates.

Sager, C., Peterson, N., Oppler, S., Rosse, R., & Walker, C. (1997). An Examination of five indexes of test battery performance: Analysis of the ECAT battery. *Military Psychology*, *9*(1), 97–120. https://doi.org/10.1207/s15327876mp0901_6

Stanek, K., & Ones, D. S. (2017). Taxonomies and compendia of cognitive ability and personality constructs and measures relevant to Industrial, Work and Organizational Psychology. In D. S. Ones, N. Anderson, C. Viswesvaran, & H. K. Sinangil (Eds.), *Personnel Psychology and Employee Performance: Vol. 1. The SAGE handbook of* *industrial, work and organizational psychology* (2nd ed., pp. 366–407). Sage Publications.

Van Iddekinge, C. H., Roth, P. L., Putka, D. J., & Lanivich, S. E. (2011). Are you interested? A meta-analysis of relations between vocational interests and employee performance and turnover. *Journal of Applied Psychology, 96*(6), 1167–1194. http://dx.doi.org/10.1037/a0024343

Vandewalle, D. (1997). Development and validation of a work domain goal orientation instrument. *Educational and Psychological Measurement, 57*(6), 995–1015. https://doi.org/10.1177/0013164497057006009

Zeidner, J., Johnson, C. D., & Scholarios, D. M. (1997). Evaluating military selection and classification systems in the multiple job context. *Military Psychology, 9(2)*, 169-186. https://doi.org/10.1207/s15327876mp0902_4

APPENDIX A

TALENT MANAGEMENT IMPORTANCE RATINGS ACROSS BRANCHES

Table A.1. ***Talent Ratings by Branch***

Talent	Mean Branch Rating																		*SD*
	AD	AG	AR	AV	CM	CY	EN	FA	FM	IN	LG	MI	MP	MS	OD	QM	SC	TC	
Bodily-Kinesthetic	1.30	1.39	1.87	1.90	1.42	0.54	1.50	1.71	1.04	2.31	1.46	1.06	1.91	1.37	1.68	1.58	1.40	1.35	0.39
Communicator	3.52	3.68	3.64	3.59	3.52	3.58	3.64	3.58	3.58	3.66	3.72	3.71	3.66	3.63	3.56	3.62	3.62	3.54	0.06
Cross-Culturally Fluent	2.54	2.81	2.48	2.13	2.28	2.05	2.46	2.41	2.45	2.54	2.77	2.95	2.8	2.67	2.49	2.69	2.58	2.70	0.24
Detail-Focused	3.39	3.58	3.22	3.30	3.30	3.49	3.28	3.42	3.63	3.28	3.46	3.36	3.33	3.38	3.38	3.36	3.43	3.28	0.11
Domain-Specific Education	1.75	1.90	1.53	1.95	1.74	3.64	2.04	1.92	2.46	1.58	2.22	1.80	1.91	2.33	1.73	1.93	2.43	1.69	0.48
Innovative	2.69	2.88	2.88	2.66	2.73	3.31	2.83	2.78	2.76	2.83	3.14	2.74	2.84	2.91	2.98	2.91	3.04	2.99	0.16
Inspirational Leader	2.70	3.05	3.01	2.81	2.72	2.56	2.75	2.75	2.73	3.04	3.10	2.67	2.93	2.88	2.90	3.01	2.99	2.95	0.16
Interdisciplinary	2.73	2.75	2.73	2.79	2.88	2.85	2.86	3.91	2.81	2.89	3.20	3.03	2.77	2.96	2.84	3.01	2.86	2.94	0.27
Interpersonal	3.21	3.57	3.43	3.21	3.19	2.9	3.39	3.3	3.34	3.43	3.46	3.31	3.42	3.40	3.35	3.53	3.29	3.34	0.15
Introspective	2.62	2.72	2.66	2.51	2.45	2.41	2.44	2.41	2.49	2.70	2.74	2.55	2.58	2.56	2.61	2.64	2.49	2.58	0.10
Logical/Analytical	3.22	3.27	3.23	3.21	3.02	3.51	3.14	3.17	3.34	3.20	3.51	3.34	3.31	3.24	3.32	3.33	3.32	3.27	0.12
Mentally Tough	3.12	3.29	3.43	3.2	3.14	2.72	2.99	3.21	3.04	3.38	3.38	3.05	3.45	3.12	3.16	3.17	3.25	3.07	0.18
Multi-Tasker	3.32	3.52	3.32	3.36	3.42	2.97	3.25	3.33	3.38	3.35	3.45	3.18	3.47	3.25	3.25	3.38	3.31	3.27	0.12
Perceptive	2.77	3.09	2.93	2.92	2.88	2.77	2.90	2.90	2.80	2.95	3.17	3.12	3.14	2.88	2.88	3.03	2.94	2.84	0.12
Physically Fit	1.67	1.75	2.38	1.83	1.96	0.64	1.94	2.26	1.53	2.88	2.04	1.59	2.35	1.82	2.10	2.06	2.00	1.88	0.45
Problem Solver	3.12	3.27	3.44	3.25	3.19	3.64	3.38	3.32	3.27	3.47	3.45	3.17	3.29	3.28	3.39	3.41	3.46	3.33	0.13
Process Disciplined	3.04	3.27	2.74	2.96	2.90	2.90	2.86	3.16	3.18	2.83	3.28	2.79	3.05	3.00	2.89	3.07	3.02	2.9	0.16
Project Manager	3.00	3.33	3.13	3.07	3.02	3.26	3.21	3.1	3.03	3.15	3.32	3.10	3.06	3.15	3.13	3.36	3.31	3.16	0.11
Prudent Risk Taker	2.17	2.21	2.66	2.66	2.20	2.44	2.33	2.46	1.97	2.81	2.72	2.18	2.53	2.32	2.59	2.5	2.42	2.35	0.22
Spatially Intelligent	2.65	2.55	2.66	2.90	2.52	2.59	2.47	2.77	2.39	2.81	2.99	2.67	2.58	2.65	2.72	2.77	2.65	2.58	0.15
Technologically Adept	2.84	2.89	2.39	2.75	2.60	3.79	2.53	2.60	2.88	2.37	2.81	2.70	2.42	2.59	2.58	2.65	3.12	2.48	0.33

Note. Branches: AD = Air Defense, AG = Adjutant General, AR = Armor, AV = Aviation, CM = Chemical, CY = Cyber, EN = Engineer, FA = Field Artillery, FM = Financial Management, IN = Infantry, LG = Logistics, MI = Military Intelligence, MP = Military Police, MS = Medical Services, OD = Ordnance, QM = Quartermaster, SC = Signal, TC = Transportation.
St. Dev = Standard Deviation.

APPENDIX B

TALENT MANAGEMENT SKILL AND ABILITY CRITICALITY RATINGS BY BRANCH

Table B.1. ***Skills and Abilities Criticality Ratings by Branch***

Talent	Mean Branch Rating																		*SD*
	AD	AG	AR	AV	CM	CY	EN	FA	FM	IN	LG	MI	MP	MS	OD	QM	SC	TC	
Psychomotor Ability	3.68	4.23	4.91	7.73	3.45	1.47	3.32	4.06	3.07	5.96	4.19	4.19	4.19	4.19	4.19	4.19	4.19	4.19	1.25
Control Precision	2.99	1.73	4.14	7.58	2.36	1.28	2.14	3.03	1.38	3.46	2.71	2.71	2.71	2.71	2.71	2.71	2.71	2.71	1.36
Reaction Time	4.10	2.26	4.61	7.48	2.66	1.28	2.54	3.52	1.89	5.05	3.01	3.01	3.01	3.01	3.01	3.01	3.01	3.01	1.37
Physical Strength	2.75	3.46	6.58	3.50	4.10	0.74	4.38	5.81	2.50	8.65	5.07	5.07	5.07	5.07	5.07	5.07	5.07	5.07	1.70
Physical Endurance	3.36	3.46	7.28	4.22	4.88	1.23	5.24	6.31	3.07	9.65	5.23	5.23	5.23	5.23	5.23	5.23	5.23	5.23	1.76
Focus	10.13	8.91	12.03	11.02	11.26	11.32	10.22	11.37	7.38	12.65	10.46	9.61	10.17	9.26	9.86	10.37	9.61	9.20	1.24
Attentiveness	13.65	14.58	14.81	13.72	13.86	13.78	13.60	14.20	12.00	14.98	13.71	13.90	14.37	13.62	14.42	14.22	13.57	14.07	0.65
Precision	13.56	15.36	13.11	13.11	13.62	14.82	12.75	13.37	14.12	12.87	13.34	13.14	13.42	13.59	13.37	13.36	12.70	13.53	0.67
Pattern Recognition	8.58	9.37	9.03	8.99	9.88	12.39	8.07	8.15	9.18	9.36	8.34	10.92	8.90	8.68	9.67	9.47	8.55	8.93	1.04
Processes Information and Data	11.58	13.66	10.75	10.79	11.85	14.66	10.90	11.20	14.59	9.96	12.76	11.95	10.02	12.01	12.01	11.82	11.66	11.63	1.33
Analyze Data or Information	12.31	13.74	12.02	11.25	12.36	15.92	11.85	12.12	14.94	11.33	14.08	13.92	10.99	12.89	12.98	12.88	12.39	13.09	1.30
Spatial Visualization	6.79	4.08	6.76	8.75	6.54	4.72	5.87	6.96	3.61	7.22	6.91	6.40	5.35	5.47	6.32	6.90	5.17	5.95	1.23
Spatial Orientation	6.73	4.42	8.25	9.42	6.38	3.49	5.82	7.58	3.34	8.75	6.78	6.11	6.32	5.75	6.41	6.68	5.21	6.35	1.60
Situational Awareness	12.77	12.04	14.11	13.25	12.03	10.16	12.10	13.19	10.81	13.86	13.47	13.11	13.48	12.59	12.55	13.44	11.82	12.25	1.03
Verbal Reasoning	13.47	14.93	14.48	13.13	14.47	12.24	14.50	13.43	13.82	14.50	14.94	15.43	14.98	14.86	14.59	14.61	14.16	14.78	0.79
Quantitative Reasoning	7.65	9.75	7.86	7.98	9.17	12.68	9.15	9.15	14.33	8.29	10.96	7.38	6.59	9.68	10.03	9.95	8.52	9.71	1.88
Interdisciplinary Reasoning	10.28	10.51	9.94	10.40	11.48	12.16	11.24	11.10	11.12	10.88	12.72	12.65	9.80	11.65	11.76	11.82	11.34	11.90	0.85
Analytical Thinking	13.20	14.13	13.83	13.42	14.05	16.22	14.05	13.55	14.78	13.81	15.19	15.60	13.28	14.40	14.32	14.54	14.33	14.28	0.79
Systems Thinking	11.41	11.55	11.76	11.27	11.71	16.14	11.84	11.97	12.90	11.72	13.40	11.59	9.77	12.27	11.99	12.47	13.32	12.03	1.28
Strategic Thinking	11.69	11.04	10.92	10.55	11.41	12.32	10.19	10.80	11.37	10.82	13.21	12.22	10.69	11.72	11.43	12.04	11.93	12.21	0.77
Structured Problem Solving	11.84	12.50	12.25	11.38	12.40	13.39	11.69	11.68	12.66	11.95	13.73	12.22	11.55	12.70	12.48	13.17	12.68	12.81	0.65
Unstructured Problem Solving	10.48	11.21	11.99	11.41	11.46	14.16	11.38	10.73	10.21	11.55	11.68	11.80	10.78	11.72	11.93	11.55	11.39	11.44	0.83
Creative Problem Solving	10.25	11.00	11.42	10.52	11.64	13.89	10.80	10.72	10.35	11.43	11.20	11.07	10.58	11.39	11.80	11.60	11.24	11.38	0.80
Assessing and Mitigating Harm	8.25	6.34	9.67	10.65	9.35	6.92	8.62	9.60	6.76	10.23	10.15	7.28	10.06	8.87	9.37	9.66	8.40	8.47	1.29

(continued)

Table B.1. ***Skills and Abilities Criticality Ratings by Branch (Continued)***

Talent	Mean Branch Rating																		*SD*
	AD	AG	AR	AV	CM	CY	EN	FA	FM	IN	LG	MI	MP	MS	OD	QM	SC	TC	
Judgement and Decision Making	11.79	11.66	13.09	13.64	11.25	11.34	12.44	12.74	10.58	13.69	13.03	11.91	12.70	12.76	13.04	12.74	12.45	12.08	0.84
Awareness of Cognitive Biases	9.02	10.42	10.16	9.03	9.53	7.34	9.32	8.84	7.86	10.04	9.45	11.39	10.50	9.42	9.40	10.70	9.27	9.71	0.97
Reflective Thinking	9.84	11.36	10.04	9.24	10.61	8.89	9.16	9.66	9.11	10.05	10.75	11.53	10.74	10.33	10.11	11.13	10.07	10.51	0.78
Written Communication	13.53	16.08	13.27	13.55	14.77	13.76	14.78	13.79	14.55	13.81	15.78	15.99	15.81	15.71	14.65	14.77	14.55	15.19	0.92
Oral Communication	15.47	16.48	16.70	15.84	15.90	13.32	15.97	16.05	14.85	16.69	16.13	16.28	16.36	16.21	15.73	16.21	16.10	16.40	0.79
Intercultural Communication	8.29	10.35	8.85	6.45	8.88	5.13	8.03	8.37	8.88	8.42	9.69	8.91	8.89	9.51	9.15	10.53	9.02	10.19	1.31
Active Listening	13.28	15.40	14.27	13.64	14.07	11.14	14.00	13.79	12.94	14.46	14.62	14.66	14.78	14.61	14.02	15.09	14.59	14.54	0.96
Encourages Discourse	10.68	11.79	11.52	10.54	11.01	8.84	10.68	10.74	10.42	12.05	11.94	12.11	10.96	11.24	11.42	12.49	11.23	11.95	0.85
Social Sensitivity	10.45	12.85	10.97	9.52	10.59	7.39	10.13	10.26	10.25	10.84	12.35	11.39	11.96	11.57	11.18	12.65	10.93	12.26	1.30
Relationship Building	13.08	15.18	14.75	13.18	13.44	11.87	13.57	13.80	13.03	14.36	15.00	14.02	13.75	14.30	13.46	15.01	13.76	14.60	0.85
Cooperation/Teamwork	14.04	15.49	15.39	14.71	14.09	12.84	13.98	14.45	14.17	15.14	15.51	14.70	14.93	14.76	14.41	15.16	14.57	15.07	0.66
Conflict Management	9.55	10.93	10.74	10.19	10.60	8.37	9.80	9.55	8.56	11.16	11.38	9.93	11.39	10.88	10.19	10.80	10.25	10.49	0.86
Social Perceptiveness	9.79	11.77	10.84	9.76	10.75	7.95	9.80	9.57	9.11	11.00	11.46	11.16	11.06	10.51	9.87	11.49	10.07	10.89	0.97
Cultural Awareness	8.31	10.59	8.81	7.07	8.36	4.89	7.42	8.05	8.23	8.62	9.98	9.49	9.50	8.78	8.69	10.62	8.37	9.79	1.35
Joint, Interagency, Inter-governmental, and Multinational (JIIM) Perspective	8.86	7.79	6.33	7.31	8.21	10.84	6.11	7.40	7.84	6.75	9.26	9.48	7.19	7.96	7.82	8.35	8.15	8.28	1.15
Working in Multidisciplinary Contexts	8.90	8.10	6.67	7.51	9.11	11.05	8.08	7.17	8.91	7.40	10.48	10.04	8.14	10.08	8.37	8.72	9.53	9.44	1.20
Working with the Public	5.37	7.15	4.54	4.36	4.87	3.47	5.50	4.56	6.67	5.03	6.86	4.32	7.44	6.90	5.87	6.17	5.33	5.55	1.12
Motivating Others	11.69	13.73	13.32	12.10	11.13	11.13	11.29	11.85	11.56	13.60	12.80	12.21	12.28	12.20	12.61	12.74	12.68	11.98	0.79
Team Building	11.09	12.58	12.67	11.80	10.55	9.89	10.76	11.59	10.77	12.91	12.57	11.27	11.86	11.43	12.31	12.42	12.04	11.80	0.84
Planning and Organizing	13.05	13.98	14.20	13.85	14.39	12.92	13.27	13.61	11.42	14.67	14.49	13.15	13.84	13.65	13.81	14.14	14.00	13.96	0.75
Directing and Supervising Others	12.68	14.14	14.24	13.12	10.98	11.70	12.42	12.76	10.98	14.07	13.71	13.36	13.03	13.62	13.26	13.43	13.74	13.29	0.98

(continued)

Table B.1. ***Skills and Abilities Criticality Ratings by Branch (Continued)***

Talent	Mean Branch Rating																		*SD*
	AD	AG	AR	AV	CM	CY	EN	FA	FM	IN	LG	MI	MP	MS	OD	QM	SC	TC	
Delegating	12.40	13.34	13.95	12.66	10.88	10.76	12.13	12.08	10.64	13.80	13.18	12.42	12.54	13.43	12.94	13.02	13.84	12.80	1.01
Training and Developing Others	11.59	11.44	12.02	11.31	10.31	10.97	10.26	11.17	9.97	12.65	11.69	11.08	10.72	11.30	11.71	11.79	11.72	10.95	0.67
Coordinating Multiple Groups	10.80	9.43	10.86	11.19	10.15	10.11	10.21	10.89	8.34	12.01	11.66	10.35	10.39	10.56	11.26	11.18	11.03	11.24	0.85
Adaptability	12.97	13.47	14.04	14.12	13.88	13.59	13.26	13.54	12.03	14.02	13.95	13.64	13.41	13.80	13.81	14.04	13.46	13.89	0.50
Cognitive Flexibility	12.14	12.14	12.89	12.26	12.73	13.38	11.75	11.89	11.36	12.62	12.84	12.70	11.83	12.88	12.46	12.69	12.07	12.37	0.50
Tolerating Pressure	12.47	13.11	13.81	13.52	13.77	11.24	12.31	13.34	11.56	13.79	13.42	12.89	13.12	13.30	13.18	12.79	12.29	13.15	0.73
Tolerating Uncertainty	11.85	12.29	13.56	13.18	13.93	13.73	12.29	13.10	11.47	13.40	12.95	12.92	12.35	12.48	13.53	13.05	12.31	13.07	0.67
Juggling Competing Demands	13.03	13.92	14.13	14.72	14.45	13.32	13.74	13.76	13.29	14.16	13.81	13.80	13.07	14.01	14.37	13.98	13.64	13.72	0.46
Knowledge of Procedures	12.29	14.11	11.90	13.49	12.49	10.70	11.63	12.48	13.67	11.66	13.35	11.89	11.84	12.49	12.70	12.65	12.60	12.51	0.83
Evaluating Compliance	9.76	12.45	9.49	10.38	10.54	9.08	9.45	9.65	12.86	10.13	11.41	10.23	10.05	10.33	10.69	10.96	10.85	10.68	0.98
Specialized Expertise	8.12	7.89	5.71	9.03	9.02	14.94	6.97	7.49	10.19	6.27	9.68	6.60	5.68	8.69	8.17	7.29	8.71	7.25	2.12
Financial Management	3.22	4.94	4.07	4.37	4.48	2.00	4.40	4.03	16.39	4.49	6.94	3.33	4.84	6.57	5.40	5.60	4.83	5.43	3.00
Proficiency with Weapons Systems	7.19	2.96	5.79	4.11	3.13	4.64	2.67	5.72	2.57	6.57	5.03	2.51	4.68	1.99	3.53	3.77	3.28	2.93	1.51
Inspecting Equipment, Objects, Structures, or Materials	5.24	3.36	5.94	5.36	5.15	3.27	4.37	5.61	3.26	5.92	5.86	3.04	4.79	3.60	4.71	4.78	5.62	4.12	1.01
Mechanically and Technologically Savvy	5.83	3.95	5.16	6.76	5.24	10.95	4.67	5.05	4.13	4.74	5.88	3.37	3.42	3.88	4.95	4.72	7.74	3.60	1.83
Expertise with Information Technology	6.45	8.24	6.82	6.77	7.30	15.89	6.76	6.26	8.70	6.19	8.37	6.72	5.39	6.76	6.90	6.66	11.46	6.47	2.45
Learning New Technology	6.21	6.83	5.87	6.24	5.64	15.94	5.38	6.16	7.51	5.66	7.42	5.69	4.77	5.72	6.61	5.75	9.42	5.80	2.50
Innovating Technology	4.16	4.93	3.59	3.23	4.18	13.81	3.34	3.91	5.51	3.75	5.58	3.41	3.18	4.02	4.12	4.01	6.94	3.92	2.46

Note. Branches: AD = Air Defense, AG = Adjutant General, AR = Armor, AV = Aviation, CM = Chemical, CY = Cyber, EN = Engineer, FA = Field Artillery, FM = Financial Management, IN = Infantry, LG = Logistics, MI = Military Intelligence, MP = Military Police, MS = Medical Services, OD = Ordnance, QM = Quartermaster, SC = Signal, TC = Transportation.
St. Dev = Standard Deviation.

APPENDIX C

OFFICER CLASSIFICATION TAXONOMY BY TALENT MANAGEMENT ATTRIBUTES CROSSWALK

Table C.1. ***Crosswalk of the Officer Classification Taxonomy and Talent Management Attributes and Corresponding Definitions***

Officer Classification Taxonomy	Talent Management Attributes	Definition
		Cognitive Abilities
Memory	NA	Ability to retain and recall information.
Perceptual Speed and Accuracy	Pattern Recognition	Perceives things quickly and accurately, and detects similarities or differences in objects, words or numbers.
Verbal Reasoning	Verbal Reasoning	Reasons and draws conclusions based on verbal or written materials.
Quantitative Reasoning	Quantitative Reasoning	Solves problems that involve mathematical concepts or numbers.
Attentiveness	Attentiveness	Focuses on the problem or situation and shifts attention between activities when appropriate.
Situational Awareness	Situational Awareness	Perceives what is happening in the immediate environment, and how information, events, and actions will impact current and near-term goals and objectives.
NA	Oral Communication	Speaks clearly and effectively in one-on-one and group settings, appropriately using gestures and other forms of nonverbal communication.
NA	Written Communication	Communicates written information and ideas to others in a clear, accurate, concise, grammatically correct, and well-organized manner.
NA	Active Listening	Carefully attends to and understands both the overt and implied meaning of oral communications from others by accurately perceiving the content, context, and tone of the speaker.
		Personality Traits and Tendencies
Stress Tolerance	Tolerating Pressure	Deals calmly and effectively with high-stress or volatile situations.
	Mentally Tough Talent	Stress tolerant and emotionally mature. Performs well even under extreme psychological duress.
Composure	NA	Tends to feel happy and self-confident.
Compassion	Interpersonal Talent	Skilled in developing appropriate relationships. Able to connect with others to affect positive results.
Politeness	NA	Tends to consider and respect the ne7u7u and desires of others.
Industriousness	NA	Tends to have high aspirations, initiative, work hard, and achieve goals.
Orderliness	Precision	Attentive to detail and thorough, accurate, and precise in completing a task.
	Detail-Focused Talent	Thorough, perceptive, and precise in all matters. Possesses a keen eye and notices everything.
Assertiveness	NA	Tends to be social dominant, influential, energetic, and take charge.
Enthusiasm	NA	Tends to experience positive emotions and enjoy the company of others.
Openness	NA	Tends to be open to art, culture, and imagination.
Intellect	Cognitive Flexibility	Willing to entertain new approaches to solving problems. Creates new plans and ideas. Initiates and accepts change and innovation.
Fitness Motivation	NA	Tends to be motivated for and engage in behaviors that promote physical fitness.

(continued)

Table C.1. ***Crosswalk of the Officer Classification Taxonomy and Talent Management Attributes and Corresponding Definitions (Continued)***

Officer Classification Taxonomy	Talent Management Attributes	**Definition**
		Personality Traits and Tendencies (continued)
Initiative	NA	Tends to rely on one's own abilities to overcome obstacles and be effective in situations that require a willingness to originate action or take independent action to achieve a goal.
Learning Orientation	NA	Tends to seek out learning opportunities and enjoy acquiring new knowledge and skills.
Self-Efficacy	NA	Tends to be confident in one's ability to succeed, effectively meet challenges, and overcome obstacles.
Team Orientation	NA	Tends to enjoy being part of a team, have a strong identification with teammates, and feel a sense of commitment and obligation to the team.
NA	Tolerating Uncertainty	Feels comfortable and excels in unstructured, complex, or rapidly changing work environments.
		Vocational Interests and Knowledge
Realistic Interests	NA	Interests in practical, hands-on, concrete activities with physical objects.
Investigative Interests	NA	Interests in rational and systematic reasoning and working with facts, data, and abstract concepts.
Artistic Interests	NA	Interests in expressing oneself creatively.
Social Interests	NA	Interests in working with and helping others.
Enterprising Interests	NA	Interests in persuading people or exerting influence over others.
Conventional Interests	NA	Interests in organizing data, people, or physical environments.
Occupation-Specific Knowledge	Specialized Expertise	Depth and breadth of knowledge related to a specific occupation.
NA	Knowledge of Procedures	Employs appropriate technical procedures or bureaucratic processes to accomplish tasks.
		Skills
Behavioral Flexibility	NA	Skill in changing one's own behavior, approach, or interpersonal style as appropriate.
Cultural Awareness	Cultural Awareness	Skill in demonstrating acceptance and understanding of individuals from other cultural and social backgrounds.
Perspective Taking	NA	Skill in understanding how people interpret events and interpersonal interactions.
Social Sensitivity	Social Sensitivity	Displays diplomacy and tact when interacting with others.
NA	Cooperation/Teamwork	Works collaboratively with others to solve problems and achieve group goals and objectives.
Self-Management	Juggling Competing Demands	Effectively manages the full range of one's work and nonwork responsibilities (e.g., setting and prioritizing goals, allocating effort and personal resources, and assessing own performance).
Directing and Supervising Others	Directing and Supervising Others	Assigns subordinates specific tasks, and sets individual work/assignment goals. Ensures assignments are clearly understood. Monitors subordinate performance and gives appropriate feedback.

(continued)

Table C.1. ***Crosswalk of the Officer Classification Taxonomy and Talent Management Attributes and Corresponding Definitions (Continued)***

Officer Classification Taxonomy	Talent Management Attributes	Definition
		Skills (continued)
Motivating Others	Motivating Others	Skill in generating support, involvement, energy, and enthusiasm for the mission among subordinates.
Delegating	Delegating	Skill in appropriately delegating authority and responsibility for decision making and for planning and executing tasks.
Shared Leadership	NA	Skill in organizing and orienting team members to meet goals.
Team Building	Team Building	Assembles a team of people that work together effectively. Identifies and effectively utilizes the appropriate mix of mission-relevant skills. Fosters group identity and cohesion by clearly communicating team goals, and encouraging and rewarding cooperation among team members.
Training and Developing Others	Training and Developing Others	Skill in determining the training needs of individual subordinates and providing the appropriate level of instruction, guidance, and developmental opportunities.
Adaptability	Adaptability	Rapidly adapts to new information, changing conditions and strategy, or unexpected obstacles, processes, and requirements.
Coordinating	Coordinating Multiple Groups	Skill in coordinating the efforts of multiple, diverse groups to accomplish a mission.
Innovation	Innovative Talent	Skill in the developing and using new and creative methods or strategies to accomplish work or achieve goals when established methods and procedures are inapplicable or ineffective.
Judgment and Decision Making	Judgment and Decision Making	Skill in making decisions based on accurate and appropriate assessment of the costs/benefits and short- and long-term consequences of alternative actions and solutions.
Planning and Organizing	Planning and Organizing	Defines the means for achieving the unit or organization goals, establishes priorities, anticipates important or critical events, identifies resource requirements, and assigns responsibility and performance expectations for specific work.
Problem Solving	Problem Solver Talent	Able to choose between best practices and unorthodox approaches to reach a solution. Accomplishes the task.
	Structured Problem Solving	Analyzes readily obtained information and evaluates results to select the best solution from a set of existing approaches to solve a problem.
	Unstructured Problem Solving	Identifies complex problems, gathers related information, evaluates information relevance, evaluates the credibility of alternative information sources, and generates alternative solutions.
	Creative Problem Solving	Develops and utilizes new and creative methods or strategies to accomplish work or achieve goals when established methods and procedures are inapplicable or ineffective.
Relationship Building	Relationship Building	Skill in developing and maintaining effective working relationships.

(continued)

Table C.1. ***Crosswalk of the Officer Classification Taxonomy and Talent Management Attributes and Corresponding Definitions (Continued)***

Officer Classification Taxonomy	Talent Management Attributes	Definition
		Skills (continued)
NA	Analytical Thinking	Analyzes information and uses logic to address work-related issues and problems.
NA	Focus	Mentally processes multiple sources of sensory information/data at the same time while avoiding distractions (e.g., flying a helicopter or commanding an armored vehicle).
NA	Analyzing Data or Information	Identifies underlying principles, reasons, or facts by breaking down information or data into separate parts.
NA	Multi-Tasker Talent	Rapidly processes and prioritizes multiple demands simultaneously. Takes appropriate action when multiple things compete for his or her attention.
NA	Logical/Analytical Talent	Uses reason and thinks in terms of cause and effect. Able to deconstruct and solve complex problems.
NA	Process Disciplined Talent	Diligently abides by procedures designed to ensure accuracy, effectiveness, and safety.
NA	Project Manager Talent	Able to determine requirements, develop work processes, delegate responsibilities, and lead teams to desired outcomes.
NA	Innovating Technology	Creates new technologies or adapts existing technologies to perform new functions.
NA	Processes Information and Data	Compiles, codes, categorizes, calculates, tabulates, audits, or verifies information or data.
NA	Strategic Thinking	Develops a complex, systems-level understanding of the relationship between his/her Army unit or organization and the broader environment and uses that understanding to envision a desirable future state for the unit/organization.
NA	Interdisciplinary Reasoning	Understands and integrates information from multiple professional disciplines to complete tasks and projects.
NA	Systems Thinking	Conceptualizes and understands relationships and arrangements within and between relevant components and structures.
NA	Learning New Technology Systems	Learns how to use and apply advances in technologies or technological systems.
NA	Expertise with Information Technology	Uses computers and computer systems (including hardware and software) to program, write software, set up functions, create new databases, or develop knowledge management systems.
NA	Evaluating Compliance	Uses relevant information, knowledge, and individual judgement to determine whether events or processes comply with laws, regulations, or standards.
NA	Financial Management	Uses financial resources effectively to set priorities and accomplish goals.
NA	Encourages Discourse	Promotes discussion and recognizes the importance of considering input from diverse perspectives.

Note. NA = Not Applicable

APPENDIX D

IN-DEPTH ANALYSES AND RESULTS

LKT Profile Similarity Metrics

Within the Army Officer Classification project, LKT shape scores were used in officer classification (Ford et al., 2020). However, Legree and colleagues (2010; 2014) outline an approach to scoring rating-based tests called profile similarity metrics (PSM), for which shape scores are only one metric. PSM calculates four metrics from a respondent's vector of LKT responses in relation to the scoring key: shape (the correlation between the respondent's vector of responses and the key vector, which can be thought of as the correspondence of relative ratings between the respondent and the key), scatter (a respondent's tendency to give—or not give—extreme responses), elevation (the tendency of a respondent to use one end of the scale), and delta (the squared difference between the respondent's mean rating and the mean keyed value). We provide below the formulas for each of these metrics, along with the more traditional distance-based score, where *x* represents a respondent, *k* represents the key, and *i* represents an individual item:

Shape: $r_{x,k}$
Scatter: sd_x^2
Elevation: X_{mean}
Delta: $(X_{mean} - K_{mean})^2$
Distance: $\sum_{i=1}^{n} |X_i - K_i| / n$

These respondent-level metrics are then used in regression modeling to find the optimal weight to assign each response profile metric to predict the criterion of interest. Legree et al. (2020) found prediction of a performance rating outcome of $R = .33$ for PSM scoring, as compared with $R = .13$ for distance-based scoring.

We used the Army Officer Classification data to evaluate PSM scoring, with usable LKT data from the 2010 – 2012 Leader Development Assessment Course (LDAC) cohorts. There were $N = 814$ cases with usable LKT data and overall performance scores. We ran analyses collapsing across branches/branch clusters to preserve sample size. The 2010 cohort ($n = 176$) responded to the LKT on a 10-point response scale, whereas the 2011 and 2012 cohorts ($n = 638$) responded to the LKT on a 25-point scale. We tested multiple methods of collapsing the two different response scales' LKT data. Ultimately, we found that standardizing item responses within response scale type (i.e., 2010 and 2011/2012) along with standardizing the key vector produced the most similar PSM metric distributions across cohorts. Item responses displayed Cronbach's α's of .88 and .93 for the LKT-Characteristics and Skills, respectively. Upon standardizing, we calculated each of the five metrics (shape, scatter, elevation, delta, distance) for the LKT-Skills and Characteristics subscales, producing 10 predictors total.

Following Legree and colleagues' (2020) methodology, we regressed the criterion of interest, overall performance, onto the ten predictors using a linear model. This produced an optimally weighted LKT score. To explore if this scoring method added predictive validity beyond the LKT scoring method used in Army Officer Classification project, we tested this against a model using shape score only. We conducted cross-validation analyses by splitting the data into two-thirds training and one-third testing samples. In the training sample, we fit the regression models to the data. In the testing sample, we applied the fitted regression model to this new data to evaluate how well the model developed in the training sample cross-validated. To address sampling error inherent in this process we conducted 10,000 bootstrapped iterations, resampling (without replacement) the training and testing samples for each iteration. Table D.1 provides the results of this analysis.

Table D.1. ***Cross-validation Results of PSM Scoring***

	Shape Scores Only	All PSM Metrics
	Training Sample	
Mean *R*	.111	.196
SD *R*	.026	.025
Min/Max *R*	.015/.214	.091/.293
Mean Δ*R*	--	.085
Mean Adj. *R*	.094	.141
Mean *N*	543	543
	Testing Sample	
Mean *R*	.094	.108
SD *R*	.052	.050
Min/Max *R*	-.093/.279	-.104/.292
Mean Δ*R*	--	.015
Mean *N*	271	271
Bootstrap Resamples	10,000	10,000
Predictors	2	10

Note. Multiple *R*'s were calculated by correlating the predicted values with the criterion. Adj. *R* was extracted from the training sample regression models.

The validity improvement in PSM scoring was more evident in the training samples than in the testing samples, although there were still slight predictive gains from PSM scoring in the testing samples. Based on this initial analysis, one might not expect large improvements in validity from PSM scoring when predicting officer performance. However, Legree et al. (2020) found PSM scoring to produce substantial validity improvements from traditional distance-based metrics. Overall, results suggest that PSM scoring is worthy of further exploration after administration of the CAB.

Vocational Interests

We examined the validity of RIASEC scale scores inferred from the officer's college major. Then, we created a person-cluster fit index by computing the correlation between each officer's RIASEC scores and the RIASEC scores of each branch cluster. Finally, we calculated another series of fit indices using a regression-based approach that focused on predicting specific criteria.

Validity of RIASEC Scale Scores

Officers' self-reported vocational interests were not included as part of the Army Officer Classification project; however, it was possible to infer these interests using their reported college majors. Each college major has an associated code based on the Classification of

Instructional Programs (CIP) system, a taxonomy of fields of study developed by the U.S. Department of Education's National Center for Education Statistics. The O*NET Resource Center provides a crosswalk of these CIP codes to the Standard Occupational Classification (SOC) codes for its reported occupations (https://www.onetonline.org/crosswalk/CIP/). Each SOC code has an associated set of vocational interest ratings based on the Holland RIASEC model (1985). Accordingly, we used RIASEC ratings for each SOC code to determine the interest profiles for each major.[4] For officers who reported more than one major, we took the average RIASEC profile across majors.

Each Army branch also has associated RIASEC ratings. We used these branch-specific ratings to create RIASEC interest score profiles for each of the branch clusters. Because certain military-specific branches (e.g., Infantry) do not map well onto O*NET's SOC codes, we obtained permission to access RIASEC ratings gathered for military occupations for the Careers in The Military (CITM) project (Bayer et al., 2017). To facilitate our reanalysis of the Army Officer Classification data, we created RIASEC ratings for each branch cluster based on scores of its constituent branches (see Table D.2). See Table D.3 for intercorrelations among the officer (i.e., person) and cluster RIASEC scores in the Army Officer Classification data.

Table D.2. ***Mean RIASEC Ratings by Branch Cluster***

Scale	Fires	Health	ILCSS	ISR	Maneuver	Maneuver Support	Netspace
Realistic	4.00	2.69	3.30	1.33	4.11	3.34	4.70
Investigative	3.67	5.98	3.07	6.67	3.95	4.37	4.39
Artistic	1.00	1.24	1.03	1.00	1.09	1.19	1.07
Social	3.33	5.07	3.34	3.67	2.94	3.48	2.24
Enterprising	6.00	3.97	4.54	4.67	5.10	4.04	2.96
Conventional	3.00	3.25	4.52	3.00	2.54	3.34	3.98

Note. ILCSS = Integrated Logistics Corps/Soldier Support, ISR = Intelligence, Surveillance, and Reconnaissance.

Table D.3. ***Intercorrelations of Person-Level and Cluster-Level RIASEC Scores***

		Person						Cluster					
		R	I	A	S	E	C	R	I	A	S	E	C
Person	R												
	I	**.38**											
	A	**-.35**	**.24**										
	S	**-.34**	.02	**.52**									
	E	**-.22**	**-.75**	**-.57**	**-.43**								
	C	**-.03**	**-.51**	**-.74**	**-.60**	**.67**							
Cluster	R	**.11**	**.04**	**-.03**	**-.08**	**-.03**	**.05**						
	I	**-.03**	**.05**	**.04**	**.03**	**-.05**	**-.05**	**-.70**					
	A	**.19**	**.15**	**-.03**	**-.03**	**-.08**	**-.08**	**.16**	**.12**				
	S	**-.05**	.00	.01	**.13**	.01	**-.11**	**-.65**	**.42**	**.33**			
	E	**-.15**	**-.13**	-.01	**.09**	**.09**	-.01	**-.04**	**-.19**	**-.43**	**.14**		
	C	.01	**-.02**	.00	-.01	.02	**.04**	.01	**-.45**	**-.20**	**-.05**	**-.45**	

Note. Bold = $p < .05$. R = Realistic; I = Investigative; A = Artistic; S = Social; E = Enterprising; C = Conventional

[4] https://www.onetcenter.org/dl_files/database/db_23_1_excel/Interests.xlsx

Table D.4 presents the validation sample sizes for each branch cluster. Sample sizes for the Separation from Service criterion are presented separately because attrition data were available for a much larger subset of the Army Officer Classification sample.

Table D.4. ***Sample Sizes for RIASEC Validation Analyses***

Branch Cluster	Main Sample	Separation Sample
Fires	260	1009
Health Services	161	487
Integrated Logistics Corps / Soldier Support	545	2282
ISR	287	1229
Maneuver	482	1918
Maneuver Support	366	1378
Network and Space Operations	182	877
Total	2283	9180

Note. Sample sizes vary slightly for individual analyses due to partial missing data.

We began by calculating validity coefficients for each RIASEC scale to establish the baseline predictive utility of individual interests. Correlations between officers' interest scores and various criterion measures are presented in Table D.5. These criterion measures come from the previous Army Officer Classification project and include performance composites (e.g., Overall Performance, Can-Do Performance, and Will-do Performance), attitudinal composites (e.g., Branch Satisfaction, Career Ambition, and Career Intentions), and separation from the military.

These validity coefficients were generally small, ranging from .00 to .06 in absolute magnitude. This finding was fairly unsurprising because the analysis considered each interest score in isolation rather than an overall profile of RIASEC scores and did not account for differences in interest profiles across branch clusters. Even so, the RIASEC scales did yield small but significant validities for predicting performance constructs. RIASEC scale scores did not show promise for predicting attitudinal criteria or separation.

Table D.5. ***Validity Coefficients for Individual RIASEC Scales***

Interest	Overall Performance	Can-Do	Will-Do	Branch Satisfaction	Career Ambition	Career Intentions	Separation
Realistic	-0.04	-0.04	**-0.04**	0.00	0.01	0.00	0.00
Investigative	**-0.06**	**-0.04**	**-0.05**	0.00	-0.03	-0.03	-0.01
Artistic	-0.01	-0.03	0.01	0.01	-0.02	0.00	0.00
Social	0.00	0.02	**0.05**	-0.01	0.03	0.03	0.00
Enterprising	**0.05**	0.03	0.02	0.02	0.02	0.02	0.01
Conventional	**0.05**	0.04	0.01	0.01	0.01	0.00	0.00

Note. Bold = $p < .05$.

Person-Cluster Fit Scores (Correlation-Based)

Our next analysis assessed the fit between individual and cluster-level RIASEC profiles. Specifically, we calculated the correlation between each officer's RIASEC scores and the RIASEC scores for their current branch cluster; this correlation coefficient served as the index of person-cluster fit. In addition to an officer's current cluster, we calculated the same fit index for all other branch clusters to determine whether each officer was optimally assigned based on their interest profile. We then created a dichotomous variable indicating whether each officer was currently assigned to their best-fit cluster (0 = no; 1 = yes).[5]

Person-cluster fit correlations ranged from -.95 to .99 overall, which is nearly the full possible range of the Pearson correlation coefficient, and the mean difference between a given officer's highest and lowest fit indices (i.e., their best-fit and worst-fit clusters) was .69. These findings suggest the correlational index produced substantial gradations in fit. Only 21.28% of officers were assigned to their best-fit cluster. As such, there may be significant room for improvement in terms of matching officer vocational interests with their career paths. On the other hand, the validation results suggest that improving this correlational fit index would not be a productive endeavor.

We computed validity coefficients for the current-cluster fit index, as well as the dichotomous best-fit cluster variable, in the overall validation and attrition samples. As shown in Table D.6, neither indicator was an effective predictor of the outcome variables. One potential explanation for this finding is that the correlational fit index accounts for an individual's (and cluster's) *relative* standing on each RIASEC trait but not their *absolute* levels of the traits. In other words, the index would not effectively differentiate between an officer with a variety of strong interests and an officer with none.

Table D.6. ***Validity Coefficients for the Correlational RIASEC Fit Index***

Criterion	Person-Current Cluster Fit	Fit Maximized (Dichotomous)
Overall Performance	-.02	.01
Can-Do	.00	-.01
Will-Do	-.03	-.01
Branch Satisfaction	.01	.02
Career Ambition	.03	.04
Career Intentions	.01	.02
Separation	-.01	.01
Average	.00	.01

Note. No correlations reached statistical significance at the .05 level. Person-Current Cluster Fit = correlation between an officer's RIASEC profile and the RIASEC profile of their current branch cluster. Fit Maximized = is the officer currently in their best-fit branch cluster? (0 = no; 1 = yes).

Person-Cluster Fit Scores (Regression-Based)

To explore RIASEC fit using a regression-based framework, we fit a series of ordinary least squares regression models to predict each criterion using RIASEC scores. We then saved the predicted criterion scores from each model, which served as the fit index (Nye et al., 2018). The models varied in terms of their complexity (i.e., whether they included cluster-level and/or polynomial terms), and we ran each model in two ways: (a) within the overall sample and (b)

[5] If two or more clusters tied for the highest fit index, we counted all the tied options as best-fit clusters.

within each cluster. Table D.7 summarizes the predictor terms included in the various models, each of which is described in greater detail below. We produced seven sets of predicted criterion scores, using one per criterion, for every model in Table D.7. This resulted in a total of 35 fit estimates.

Table D.7. *Summary of the RIASEC Regression Models*

Term	Level of Analysis	Overall Sample Models			Within-Cluster Models	
		Person-Only	Main Effects	Polynomial	Main Effects[a]	Polynomial
Intercept	–	x	x	x	x	x
Realistic	Person	x	x	x	x	x
Investigative	Person	x	x	x	x	x
Artistic	Person	x	x	x	x	x
Social	Person	x	x	x	x	x
Enterprising	Person	x	x	x	x	x
Conventional	Person	x	x	x	x	x
Realistic	Cluster		x	x		
Investigative	Cluster		x	x		
Artistic	Cluster		x	x		
Social	Cluster		x	x		
Enterprising	Cluster		x	x		
Conventional	Cluster		x	x		
Realistic^2	Person			x		x
Investigative^2	Person			x		x
Artistic^2	Person			x		x
Social^2	Person			x		x
Enterprising^2	Person			x		x
Conventional^2	Person			x		
Realistic	Person x Cluster			x		
Investigative	Person x Cluster			x		
Artistic	Person x Cluster			x		
Social	Person x Cluster			x		
Enterprising	Person x Cluster			x		
Conventional	Person x Cluster			x		

Note. Terms with an “x” are included in the specified model.
a. The main effects and person-only models were identical when run within clusters since the main effects of cluster-level RIASEC scores could not be modeled within a single cluster. We used the main effects label in this case since the key difference between the two within-cluster models was the inclusion or exclusion of polynomial terms.

The simplest model we fit was a basic person-only model, which included the six person-level RIASEC scores, in the overall sample. Although this model ignored any potential differences between branch clusters, it was helpful to establish a baseline level of prediction without accounting for environmental effects. Next, we tested a main effects model, which included both person- and cluster-level RIASEC scores. This model served as an intermediate step to account for differences between branch cluster interest profiles without adding the complexity of polynomial terms.

In addition to the overall sample, we also ran a main effects model within each branch cluster. Since cluster-level variables have no variance within a single cluster, it was not possible to include the cluster RIASEC scores in this analysis. As such, the within-cluster main effects model included the same terms as the overall sample person-only model. However, it still effectively accounted for cluster effects by fitting the data for each cluster individually.

In addition to modeling the main effects of officer and cluster RIASEC scores, we ran polynomial regression analyses to obtain a more complete model of person-environment congruence (Edwards, 1993; Nye et al., 2018). In doing so, we followed Nye et al.'s (2018) procedure for computing a polynomial fit index. First, we mean-centered the 12 RIASEC variables within the overall sample[6] to reduce multicollinearity between the main effect and polynomial terms. Next, we squared each RIASEC score to account for nonlinear effects and computed a person-cluster interaction term for each of the six interests. Finally, we regressed the main effect, squared, and interaction variables onto each criterion scale.

The within-cluster version of the polynomial models excluded main effects and squared terms for cluster-level RIASEC scores, once again due to their lack of variance within a single cluster. It was also necessary to exclude the interaction terms from the within-cluster models because of their redundancy with the person-level main effects. Similarly, we were ultimately unable to include the squared cluster-level terms in the overall sample models due to perfect multicollinearity with the other predictors.

Validity coefficients for the regression-based fit indices are presented in Table D.8. For the overall sample indices, these coefficients simply represent the multiple correlation for predicting each outcome. Their interpretation is conceptually similar for the within-cluster indices, except that these indices were constructed piecemeal from seven regression models run within individual clusters, whereas a multiple R is typically based on predicted values from a single model. A key limitation in both cases is that there is no penalty for capitalizing on chance. As a result, the validity coefficients in Table D.8 increase monotonically as more terms are added and when the analyses are run on individual clusters - both of which increase the potential for overfitting.

Table D.8. ***Validity Coefficients for the Regression-Based RIASEC Fit Indices***

	Overall Sample			Within-Cluster	
Criterion	Person-Only	Main Effects	Polynomial	Main Effects	Polynomial
Overall Performance	.09	.09	.11	.14	.18
Can-Do	.07	.08	.14	.16	.23
Will-Do	.08	.08	.13	.15	.22
Branch Satisfaction	.04	.04	.08	.15	.22
Career Ambition	.08	.08	.12	.16	.21
Career Intentions	.08	.08	.13	.17	.22
Separation	.01	.05	.06	.08	.11
Average	.06	.07	.12	.15	.21

[6] We chose not to center the variables within clusters because an officer's future branch cluster would not be known when using RIASEC scores to guide branch assignments.

To address this issue, we calculated adjusted multiple R's for each fit index. First, we squared each correlation in order to apply the formula for adjusted R^2, which penalizes a squared multiple correlation proportional to the number of predictors and inversely proportional to the sample size. After calculating adjusted R^2 for each fit index, we took the square root of that value to transform it back into a correlation. If the adjusted R^2 was negative, we set the correlation to zero since (a) the square root of a negative value is an irrational number and (b) a negative multiple correlation does not lend itself to meaningful interpretation.

Unfortunately, it was not straightforward to apply this correction to the within-cluster fit indices since they represented seven sets of predicted values from subsamples of varying sizes. Simply using the overall sample size would produce a severe undercorrection by understating the degree of sampling error in the individual clusters. As a result, we explored two alternate methods to apply the adjustment. First, we divided the overall sample size for each correlation by the number of clusters to produce an average cluster N, which we then input into the adjusted R^2 formula (Method 1). Second, we calculated an adjusted R^2 within each cluster and computed the mean of those values (Method 2).

Adjusted validity coefficients are presented in Table D.9. Among the overall sample analyses, the main effects models performed worse on average than the person-only and polynomial models. Since the main effects model added substantial complexity without truly accounting for person-cluster congruence, it is perhaps unsurprising that it performed poorly. The within-cluster fit indices were generally more predictive than the overall sample versions, and the polynomial model fared especially well. Notably, this trend did not hold for the Overall Performance criterion. The simple person-only model was the best predictor of this outcome, suggesting the more complex models did not add enough prediction to outweigh the potential drawbacks of overfitting.

Table D.9. ***Adjusted Validity Coefficients for the Regression-Based RIASEC Fit Indices***

	Overall Sample			Within-Cluster			
Criterion	Person-Only	Main Effects	Polynomial	Main Effects (Method 1)	Main Effects (Method 2)	Polynomial (Method 1)	Polynomial (Method 2)
Overall Performance	.07	.05	.04	.01	.04	.00	.00
Can-Do	.05	.02	.10	.08	.08	.14	.13
Will-Do	.06	.03	.08	.07	.05	.10	.08
Branch Satisfaction	.00	.00	.00	.05	.00	.09	.05
Career Ambition	.05	.00	.05	.08	.09	.06	.08
Career Intentions	.06	.03	.07	.09	.08	.08	.08
Separation	.00	.03	.03	.07	.08	.10	.11
Average	.04	.02	.05	.06	.06	.08	.08

Note. Methods 1 and 2 refer to methods for performing the adjustment. See text for details.