

Technical Report 1359

Tier One Performance Screen Initial Operational Test and Evaluation: 2013 Annual Report

Bethany H. Bynum, Editor

Human Resources Research Organization

Heather M. Mullins, Editor

U.S. Army Research Institute



July 2017

**United States Army Research Institute
for the Behavioral and Social Sciences**

Approved for public release; distribution is unlimited

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

**Department of the Army
Deputy Chief of Staff, G1**

Authorized and approved:

**MICHELLE SAMS, Ph.D.
Director**

Research accomplished under contract W911NF13C0088
for the Department of the Army by

Human Resources Research Organization

Technical Review by

Jonathan F. Kochert, U.S. Army Research Institute
Cristina D. Kirkendall, U.S. Army Research Institute

NOTICES

DISTRIBUTION: This Technical Report has been submitted to the Defense Information Technical Center (DTIC). Address correspondence concerning ARI reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, Attn: DAPE-ARI-ZXM, 6000 6th Street Building 1464 / Mail Stop: 5610, Fort Belvoir, VA 22060-5610.

FINAL DISPOSITION: Destroy this Technical Report when it is no longer needed. Do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this Technical Report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

REPORT DOCUMENTATION PAGE					
1. REPORT DATE (dd-mm-yy) July 2017		2. REPORT TYPE Interim		3. DATES COVERED (from. . . to) August 2009 to December 2013	
4. TITLE AND SUBTITLE Tier One Performance Screen Initial Operational Test and Evaluation: 2013 Annual Report				5a. CONTRACT OR GRANT NUMBER W911NF-13-C-0088	
				5b. PROGRAM ELEMENT NUMBER 633007	
6. AUTHOR(S) Bethany H Bynum, Editor Heather M. Mullins, Editor				5c. PROJECT NUMBER A790	
				5d. TASK NUMBER 329a	
				5e. WORK UNIT NUMBER	
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 6 th Street (Bldg 1464/Mail Stop 5610) Ft. Belvoir, VA 22060-5610				10. MONITOR ACRONYM ARI	
				11. MONITOR REPORT NUMBER Technical Report 1359	
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES ARI Research POC: Dr. Heather Wolters, Personnel Assessment Research Unit					
14. ABSTRACT (<i>Maximum 200 words</i>): In addition to educational, physical, and moral screens, the U.S. Army relies on the Armed Forces Qualification Test (AFQT), a composite score from the Armed Services Vocational Aptitude Battery (ASVAB), to select new Soldiers into the Army. Although the AFQT is useful for selecting new Soldiers, other personal attributes are important to Soldier performance and retention. Based on previous U.S. Army Research Institute (ARI) investigations, the Army selected one promising measure, the Tailored Adaptive Personality Assessment System (TAPAS), for an initial operational test and evaluation (IOT&E), beginning administration to applicants in 2009. Criterion data are being compiled at 6-month intervals from administrative records, schools for selected military occupational specialties (MOS), and Soldiers in units. This is the eighth in a series of planned evaluations of the TAPAS. Similar to prior research, the cumulative results thus far suggest that several TAPAS scales significantly predict a number of criteria of interest, indicating that the measure holds promise for both selection and classification purposes.					
15. SUBJECT TERMS Personality Assessment Personnel Manpower Selection and classification					
SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT Unlimited	20. NUMBER OF PAGES 119	21. RESPONSIBLE PERSON Tonia Heffner Chief 703-545-4408
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified			

Technical Report 1359

**Tier One Performance Screen Initial Operational Test
and Evaluation: 2013 Annual Report**

Bethany H. Bynum, Editor
Human Resources Research Organization

Heather M. Mullins, Editor
U.S. Army Research Institute

Personnel Assessment Research Unit
Tonia S. Heffner, Chief

July 2017

Approved for public release; distribution is unlimited

ACKNOWLEDGEMENTS

There are individuals not listed as authors who made significant contributions to the research described in this report. First and foremost are the Army cadre who support criterion data collection efforts at the schoolhouses. These noncommissioned officers (NCOs) ensure that trainees are scheduled to take the research measures and provide ratings of their Soldiers' performance in training. Those Army personnel who support the in-unit data collections are also instrumental to this research program. Dr. Deirdre Knapp (HumRRO) provides project oversight and assisted in editing this report. Thanks also go to Dr. Tonia Heffner, Dr. Len White, and Ms. Sharon Meyers (ARI) and Mr. Doug Brown, Mr. Chris Graves, and Ms. Charlotte Campbell (HumRRO) for their important contributions to this research effort.

We also extend our appreciation to the Army Test Program Advisory Team (ATPAT), a group of senior NCOs who periodically meet with ARI researchers to help guide this work in a manner that ensures its relevance to the Army and who help enable the Army support required to implement the research. Members of the ATPAT are:

CW4 Mary Negron
CSM Lamont Christian
CSM Thomas Clingel
CSM Michael Evans
CSM John Pack Jr.
SGM Donnie Isaacs
SGM Robert Trawick Jr.
MSG Patricio Cardonavega
MSG Fredrick Clayton
MSG Daniel Gutierrez
MSG Thomas Morgan
SFC Eli Aguilar
SFC Timothy Castelow
SFC Thomas Hill
SFC Vanda Johnson
SFC James Swenson
SSG William Cheek
SSG Teddy Hayes
SSG Suhun Sung
Mr. Tom Bowles
Mr. Clifford McMillan
Mr. William Palya
Mr. Gerald Purcell

TIER ONE PERFORMANCE SCREEN INITIAL OPERATIONAL TET AND EVALUATION: 2013 ANNUAL REPORT

EXECUTIVE SUMMARY

Research Requirement:

In addition to educational, physical, and moral screens, the U.S. Army relies on the Armed Forces Qualification Test (AFQT), a composite score from the Armed Services Vocational Aptitude Battery (ASVAB), to select new Soldiers into the Army. Although the AFQT has proven to be and will continue to serve as a useful metric for selecting new Soldiers, there is a growing recognition of the need to consider whole person assessment that takes other personal attributes, in particular non-cognitive attributes (e.g., temperament, interests, and values) into consideration. Non-cognitive attributes are important to entry-level Soldier performance and retention (e.g., Campbell & Knapp, 2001; Ingerick, Diaz, & Putka, 2009; Knapp & Heffner, 2009, 2010; Knapp & Tremble, 2007). Based on previous research (Knapp & Heffner, 2010), the Army selected one particularly promising measure, the Tailored Adaptive Personality Assessment System (TAPAS), as the basis for an initial operational test and evaluation (IOT&E) of the *Tier One Performance Screen (TOPS)*. The TAPAS capitalizes on the latest advances in testing technology to assess motivation through the measurement of personality characteristics.

Procedure:

In May 2009, the Military Entrance Processing Command (MEPCOM) began administering the TAPAS on the computer adaptive platform for the ASVAB (CAT-ASVAB) at Military Entrance Processing Stations (MEPS). To evaluate the TAPAS, outcome (criterion) data are being collected at multiple points in time from Soldiers who took the TAPAS at entry. Specifically, initial military training (IMT) criterion data are being collected at schools for Soldiers in eight military occupational specialties (MOS). Project teams also are collecting criterion data from Soldiers (regardless of MOS) in their units in multiple waves of site visits during the course of the IOT&E.

The criterion measures include job knowledge tests, an attitudinal assessment (the Army Life Questionnaire), and performance rating scales completed by the Soldiers' cadre members (in IMT) or supervisors (in units). Course grades, completion rates, and attrition status are obtained from administrative records for all Soldiers.

The data presented in this report come from TAPAS data collected through September 2013 and criterion data collected through December 2013. It consists of a total of 486,310 applicants who took the TAPAS; 443,229 of these individuals were in the TOPS Applicant Sample. The Applicant Sample (used for analysis purposes) excluded those ineligible for service based on education requirements or extremely low AFQT scores and prior service applicants. The validation sample sizes were considerably smaller, with the IMT Validation Sample comprising 23,495 Soldiers, the In-Unit Validation Sample comprising 1,965 Soldiers, and the

Administrative Validation Sample (which includes Soldiers with criterion data [e.g., attrition] from at least one administrative source) comprising 226,055 Soldiers.

Data from the job knowledge tests, performance-rating scales provided by cadre or supervisor, attitudinal assessment, and administrative sources were combined to yield an array of scores representing important Soldier outcomes. In general, the criterion scores exhibited acceptable and theoretically consistent psychometric properties. The exception to this was the performance-rating scales, which exhibit low inter-rater reliability. Results involving the rating scales may underestimate relationships with other variables.

Our approach to analyzing the TAPAS' incremental predictive validity was consistent with previous evaluations of this measure and similar experimental non-cognitive predictors (e.g., Ingerick et al., 2009; Knapp & Heffner, 2009, 2010, 2011). In brief, this approach involved testing a series of hierarchical regression models, regressing scores for each criterion measure onto Soldiers' AFQT scores, followed by their TAPAS composite or TAPAS scale scores in the second step. The resulting increment in the multiple correlation value (ΔR) when the TAPAS composite or TAPAS scale scores were added to the baseline regression models served as our index of incremental validity. Correlations between TAPAS scale scores and selected criteria were also examined. Analyses used the TAPAS Will-Do, Can-Do, and Adaptation composite scores.

To further examine the relationships between the TAPAS and the various IMT, in-unit, and attrition criteria, we conducted analyses that assessed implementation of the TAPAS with respect to AFQT categories and TAPAS score percentiles. For some of these analyses, AFQT Category IIIB/IV Soldiers were classified as either IIIB/IV Pass or IIIB/IV Fail based on their TAPAS Will-Do composite scores, with Soldiers scoring among the bottom 10% classified as IIIB/IV Fail.

Findings:

Consistent with previous TOPS IOT&E reports, the results of this evaluation suggest TAPAS holds promise for new Soldier selection. Results of the incremental validity analyses indicate that the TAPAS predicts important first-term criteria over and above the AFQT, especially measures tapping motivational aspects of Soldier performance, such as physical fitness, adjustment to Army life, commitment and fit, and discipline. Further, examination of AFQT categories and quintile splits of predictor composites showed a clear linear improvement in favor of higher scoring individuals. Individuals in the lowest AFQT categories performed the worst.

The Will-Do composite, a combination of TAPAS scales that predict motivation-based outcomes, was associated with the greatest incremental validity gains compared to other TAPAS composites. This was especially true for the prediction of physical fitness, Will Do Performance, and Army Life Adjustment. When examining outcomes by AFQT category, a clear distinction was seen when comparing the IIIB/IV Pass group and the IIIB/IV Fail group. The largest difference was for disciplinary incidents and training restarts where the IIIB/IV Fail group had approximately 21% and 25% more disciplinary incidents and restarts, respectively, compared to the IIIB/IV Pass group. Results showed consistently higher attrition among the IIIB/IV Fail AFQT category. The

IIIB/IV Pass group tended to have 40% lower attrition than the IIIB/IV Fail group. The Adaptation composite generally provided small incremental validity gains for predicting attrition, showing relatively larger gains for predicting attrition later in the enlistment term. Even these small gains in validity are important, particularly given the modest relationship with the AFQT.

Utilization and Dissemination of Findings:

The research findings will be used by the Army Deputy Chief of Staff, G-1; U.S. Army Recruiting Command; Assistant Secretary of the Army (Manpower and Reserve Affairs); and Training and Doctrine Command to evaluate the effectiveness of tools used for Army applicant selection and assignment. With each successive set of findings, the TAPAS can be revised and refined to meet Army needs and requirements.

TIER ONE PERFORMANCE SCREEN INITIAL OPERATIONAL TEST AND EVALUATION: 2013 ANNUAL REPORT

CONTENTS

	Page
CHAPTER 1: INTRODUCTION	1
Deirdre J. Knapp (HumRRO), Tonia S. Heffner, and Leonard A. White (ARI)	
Background	1
The Tier One Performance Screen (TOPS).....	2
Evaluating TOPS	3
Overview of Report	3
CHAPTER 2: SAMPLE CHARACTERISTICS	4
D. Matthew Trippe, Joseph P. Caramagno, and Chris R. Graves (HumRRO)	
Data Sources.....	4
Sample Filters.....	5
Description of Analysis Samples	6
Summary	9
CHAPTER 3: THE TAILORED ADAPTIVE PERSONALITY ASSESSMENT SYSTEM (TAPAS)	11
Stephen Stark, O. Sasha Chernyshenko, Christopher Nye, and Fritz Drasgow (Drasgow Consulting Group)	
Description	11
Psychometric Properties of TAPAS Test Versions.....	12
TAPAS Composites	15
Summary	15
CHAPTER 4: DESCRIPTION AND PSYCHOMETRIC PROPERTIES OF CRITERION MEASURES	16
Thomas B. Kiger, Joseph P. Caramagno, and Matthew C. Reeder (HumRRO)	
Job Knowledge Tests (JKTs)	17
Performance Rating Scales (PRS).....	19
Army Life Questionnaire (ALQ).....	23
Administrative Criteria.....	26
Criterion Composites.....	29
Summary	32

CONTENTS (CONTINUED)

	Page
CHAPTER 5: EVIDENCE FOR THE PREDICTIVE VALIDITY OF THE TAPAS	33
Michael G. Hughes, Bethany H. Bynum, (HumRRO), and Heather M. Mullins (ARI)	
Analytic Approach	33
Findings	35
Summary	56
CHAPTER 6: SUMMARY AND A LOOK AHEAD	57
Deirdre J. Knapp, Bethany H. Bynum (HumRRO), Heather M.K. Wolters, and Tonia S. Heffner (ARI)	
Summary of the TOPS IOT&E Method.....	57
Summary of Evaluation Results to Date	57
Looking Ahead	58
Related Research	59
Conclusion.....	59
REFERENCES	60
APPENDICES	
APPENDIX A: PREDICTOR MEASURE PSYCHOMETRIC PROPERTIES IN THE APPLICANT SAMPLE	A-1
APPENDIX B: CORRELATIONS AMONG CRITERION MEASURES IN THE IMT AND IN-UNIT VALIDATION SAMPLES	B-1
APPENDIX C: CRITERION PSYCHOMETRIC PROPERTIES IN THE FULL IMT AND IN-UNIT SAMPLES	C-1
APPENDIX D: SUMMARY OF BIVARIATE CORRELATIONS BETWEEN TAPAS SCALES AND SELECTED CRITERIA	D-1
CONTENTS (CONTINUED)	

Page

List of Tables

Table 2.1. Full TAPAS Sample Characteristics.....	7
Table 2.2. Background and Demographic Characteristics of the TOPS Samples	8

Table 3.1. TAPAS Facets Names and Definitions.....	13
Table 3.2. TAPAS Versions by Administration Date.....	14
Table 4.1. Summary of IMT and In-Unit Criterion Measures	16
Table 4.2. Reliability Estimates of the Job Knowledge Tests (JKTs) in the IMT and In- Unit Validation Samples	17
Table 4.3. Descriptive Statistics for the Job Knowledge Tests (JKTs) in the IMT Validation Sample by Education Tier	18
Table 4.4. Descriptive Statistics for Job Knowledge Tests (JKTs) in the In-Unit Validation Sample by Education Tier	19
Table 4.5. Interrater Reliability Estimates for the IMT Performance Rating Scales (PRS)	20
Table 4.6. Descriptive Statistics for the Performance Rating Scales (PRS) in the IMT Validation Sample by Education Tier	21
Table 4.7. In-Unit Army-Wide Performance Rating Scale Dimensions and Composite Scores.....	22
Table 4.8. Descriptive Statistics for Army-Wide Performance Rating Scales (PRS) in the In-Unit Validation Sample	23
Table 4.9. Army Life Questionnaire (ALQ) Likert-Type Scales.....	24
Table 4.10. Descriptive Statistics and Reliability Estimates for the Army Life Questionnaire (ALQ) in the IMT Validation Sample by Education Tier	25
Table 4.11. Descriptive Statistics and Reliability Estimates for the Army Life Questionnaire (ALQ) in the In-Unit Validation Sample by Education Tier	26
Table 4.12. Base Rates for Attrition Criteria for Regular Army Soldiers in the Validation Sample by Education Tier	28
Table 4.13. Base Rates or Descriptive Statistics for Administrative IMT Criteria in the Validation Sample.....	28
Table 4.14. IMT and In-Unit Criterion Scores	29
Table 4.15. Descriptive Statistics for Criterion Composites in the IMT and In-Unit Validation Samples by Education Tier	31
Table 5.1. Summary of Regression Models Evaluated for each Criterion	34
Table 5.2. Incremental Validity Estimates for the TAPAS over AFQT for Predicting IMT Technical by Education Tier	37
Table 5.3. Incremental Validity Estimates for the TAPAS over AFQT for Predicting IMT Adjustment, Commitment & Fit, Fitness, and Retention Criteria by Education Tier	38
Table 5.4. Incremental Validity Estimates for the TAPAS over AFQT for Predicting IMT Performance Rating Criteria by Education Tier	39
Table 5.5. Incremental Validity Estimates for the TAPAS over AFQT for Predicting In- Unit Technical Criteria by Education Tier.....	42
Table 5.6. Incremental Validity Estimates for the TAPAS over AFQT for Predicting In- Unit Adjustment, Commitment & Fit, Fitness, and Retention Criteria by Education Tier.....	43

CONTENTS (CONTINUED)

	Page
Table 5.7. Incremental Validity Estimates for the TAPAS over AFQT for Predicting In-Unit Performance Rating Criteria by Educational Tier	44
Table 5.8. Incremental Validity Estimates for the TAPAS Composites over AFQT for Predicting Dichotomous Criteria by Education Tier	47
Table 5.9. Incremental Validity Estimates for the TAPAS Composite Scores over AFQT for Predicting Cumulative Attrition through 24 Months of Service by Education Tier (Regular Army Only)	49

List of Figures

Figure 2.1. Overview of TOPS data file merging and nested sample generation process.....	4
Figure 5.1. Increase in prediction of IMT criteria using Will-Do TAPAS scales for the combined Tier 1 and 2 sample.	40
Figure 5.2. Increase in prediction of IMT criteria using Can-Do TAPAS scales for the combined Tier 1 and 2 sample.	40
Figure 5.3. Increase in prediction of in-unit criteria using Will-Do TAPAS scales for the combined Tier 1 and 2 sample.	45
Figure 5.4. Increase in prediction of in-unit criteria using Can-Do TAPAS scales for the combined Tier 1 and 2 sample.	45
Figure 5.5. Predictive accuracy of the AFQT and Will-Do TAPAS composite in the discrimination of both IMT and In-Unit Disciplinary Incidents and IMT Restarts for the combined Tier 1 and 2 sample.....	51
Figure 5.6. Predictive accuracy of the AFQT and Adaptation TAPAS composite in the discrimination of attrition outcomes for Regular Army Soldiers in the combined Tier 1 and 2 sample.	52
Figure 5.7. Tier 1 Soldier outcomes for selected IMT criteria by AFQT category and TAPAS pass/fail status.	54
Figure 5.8. Tier 1 Soldier attrition by AFQT category and TAPAS pass/fail status.	55
Figure 5.9. Tier 1 Soldier outcomes for selected in-unit criteria by TAPAS percentile score categories.	56

TIER ONE PERFORMANCE SCREEN INITIAL OPERATIONAL TEST AND EVALUATION: 2013 ANNUAL REPORT

CHAPTER 1: INTRODUCTION

Deirdre J. Knapp (HumRRO), Tonia S. Heffner, and Leonard A. White (ARI)

Background

The Personnel Assessment Research Unit (PARU) of the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) is responsible for conducting personnel research for the Army. The focus of PARU's research is maximizing the potential of the individual Soldier through effective selection, classification, and retention strategies.

In addition to educational, physical, and moral screens, the U.S. Army relies on the Armed Forces Qualification Test (AFQT), a composite score from the Armed Services Vocational Aptitude Battery (ASVAB), to select new Soldiers into the Army. Although the AFQT has proven to be, and will continue to serve as, a useful metric for selecting new Soldiers, other personal attributes, in particular non-cognitive attributes (e.g., temperament, interests, and values), are important to entry-level Soldier performance and retention (e.g., Knapp & Tremble, 2007).

In December 2006, the Department of Defense (DoD) ASVAB review panel—a panel of experts in the measurement of human characteristics and performance—released their recommendations (Drasgow, Embretson, Kyllonen, & Schmitt, 2006), several of which focused on supplementing the ASVAB with additional measures for use in selection and classification decisions. The ASVAB review panel further recommended that the use of these measures be validated against performance criteria.

Just prior to the release of the ASVAB review panel's findings, ARI had initiated a longitudinal research effort, *Validating Future Force Performance Measures (Army Class)*, to examine the prediction potential of several non-cognitive measures (e.g., temperament and person-environment fit) for Army outcomes (e.g., performance, attitudes, attrition). The Army Class research project was a six-year effort conducted with contract support from the Human Resources Research Organization ([HumRRO]; Allen, Knapp, & Owens, 2013; Ingerick, Diaz, & Putka, 2009; Knapp & Heffner, 2009). Experimental predictors were administered to new Soldiers in 2007 and early 2008. Army Class collected school-based criterion data on a subset of the Soldier sample as they completed job training. Job performance criterion data were collected from Soldiers in the Army Class longitudinal validation sample in 2009 with a second round of data collections in Soldiers' units completed in April 2011 (Knapp, Owens, & Allen, 2012). Final analysis and reporting of this program of research is complete (Allen et al., 2013).

After the Army Class research began, ARI initiated the *Expanded Enlistment Eligibility Metrics (EEEM)* project (Knapp & Heffner, 2010). The EEEM goals were similar to Army Class, but the focus was specifically on Soldier selection and the time horizon was much shorter. Specifically, EEEM required identification of one or more promising new predictor measures for immediate

implementation. The EEEM project capitalized on the existing Army Class data collection procedure and, thus, the EEEM sample was a subset of the Army Class sample.

As a result of the EEEM findings, Army policy-makers approved an initial operational test and evaluation (IOT&E) of the *Tier One Performance Screen (TOPS)*. This report is the latest in a series presenting continuing analyses from the IOT&E of TOPS.

The Tier One Performance Screen (TOPS)

Six experimental pre-enlistment measures were included in the EEEM research (Allen, Cheng, Putka, Hunter, & White, 2010). These included several temperament measures, a situational judgment test, and two person-environment fit measures based on values and interests. The most promising measures recommended to the Army for implementation were identified based on the following considerations:

- Incremental validity over AFQT for predicting important performance and retention-related outcomes,
- Minimal subgroup differences,
- Low susceptibility to response distortion (e.g., faking optimal responses), and
- Minimal administration time requirements.

The Tailored Adaptive Personality Assessment System ([TAPAS]; Stark, Chernyshenko, & Drasgow, 2010) surfaced as the top choice¹. The TAPAS is a measure of personality characteristics (e.g., achievement, sociability) that capitalizes on the latest advances in psychometric theory and provides a good indicator of personal motivation.

In May 2009, the Military Entrance Processing Command (MEPCOM) began administering the TAPAS on the computer adaptive platform for the ASVAB (CAT-ASVAB). Initially, the TAPAS was to be administered only to Education Tier 1, non-prior service applicants.² This limitation to Education Tier 1 was removed early in CY2011 so the Army could evaluate the TAPAS across all types of applicants.

TOPS uses the TAPAS to identify applicants who would likely perform differently (higher or lower) than would be predicted by their ASVAB scores. As part of the TOPS IOT&E, TAPAS scores are being used to screen out a small number of AFQT Category IIIB/IV applicants.³ Recently, the IOT&E was broadened to include all Tier 2 applicants, regardless of AFQT category.

The initial conceptualization for the IOT&E was to use the TAPAS as a tool for “screening in” Education Tier 1 applicants with lower AFQT scores and TAPAS scores that suggest they would perform well in the Army. However, changing economic conditions spurred a reconceptualization that led to using the TAPAS as a tool that screens out low motivated

¹ Other promising assessments include the Work Preferences Assessment ([WPA]; Putka & Van Iddekinge, 2007) and the Information/Communications Technology Literacy test ([ICTL]; Russell & Sellman, 2009).

² Applicant educational credentials are classified as Tier 1 (primarily high school diploma), Tier 2 (primarily non-diploma graduate), and Tier 3 (no educational credentials).

³ Examinees are classified into categories based on their AFQT percentile scores (Category I = 93-99, Category II = 65-92, Category IIIA = 50-64, Category IIIB = 31-49, Category IV = 10-30, Category V = 1-9).

applicants, thus making the selection criteria to enter the Army more stringent. As recruiting conditions continue to shift, the IOT&E, as well as operational implementations, can adjust to fit the applicant market. That is, TAPAS composite scores and cut points can be modified to fit recruiting market conditions.

Evaluating TOPS

To evaluate the TAPAS, the Army is collecting training criterion data on Soldiers in multiple target military occupational specialties (MOS) as they complete initial military training (IMT).⁴ The criterion measures include job knowledge tests (JKTs); an attitudinal assessment, the Army Life Questionnaire (ALQ); and performance rating scales (PRS) completed by the Soldiers' cadre. Course grades and completion rates are obtained from administrative records for all Soldiers who take the TAPAS, regardless of MOS.

Criterion data are also being collected from Soldiers and their supervisors during data collection trips to major Army installations. These proctored "in-unit" data collections began in January 2011 and target all Soldiers who took the TAPAS prior to enlistment. The in-unit criterion measures include JKTs, the ALQ attitudinal assessment, and PRS completed by the Soldiers' supervisors. The data collection model closely mirrors that which was used in the Army Class research program (Knapp et al., 2012). Separation status of all Soldiers who took the TAPAS prior to enlistment is tracked throughout the course of the research.

This report describes the eighth iteration of the criterion-related validation through the TOPS IOT&E initiative. Prior evaluations are described in a series of technical reports (Knapp & Heffner, 2011, 2012; Knapp, Heffner, & White, 2011; Knapp & LaPort, 2013a, 2013b, 2014) and internal memoranda. Additional validation analyses will be prepared and conducted at six-month intervals throughout the multi-year IOT&E period.

Overview of Report

Chapter 2 explains how the evaluation analysis data files are constructed and then describes characteristics of the current sample. Chapter 3 describes the TAPAS, including content, scoring, and psychometric characteristics. Chapter 4 describes the IMT and in-unit criterion scores used in this evaluation, including their psychometric characteristics. Criterion-related validation analyses for the TAPAS are presented in Chapter 5. The report concludes with Chapter 6, which summarizes our continuing efforts to evaluate TOPS and looks toward plans for future iterations of these evaluations.

⁴ The target MOS are Infantryman (11B), Armor Crewman (19K), Military Police (31B), Human Resources Specialist (42A), Health Care Specialist (68W), Motor Transport Operator (88M), and Light Wheel Vehicle Mechanic (91B). These MOS were selected to include large, highly critical MOS as well as to represent the diversity of work requirements across MOS.

CHAPTER 2: SAMPLE CHARACTERISTICS

D. Matthew Trippe, Joseph P. Caramagno, and Chris R. Graves (HumRRO)

This chapter describes characteristics of the samples used in the TOPS IOT&E evaluation analyses. We begin with a brief summary of data sources, describe how Soldier data were filtered for analysis, and then describe multiple subsamples that were created to support various types of analyses.

Data Sources

An illustrative view of the TOPS sources of predictor and criteria data is provided in Figure 2.1. The lighter boxes within the figure represent sources of data, and the darker boxes represent samples on which descriptive or inferential analyses are conducted. The leftmost column in the figure summarizes the predictor data sources used to derive the TOPS Applicant Sample. The other columns summarize the research-only (i.e., non-administrative) and administrative criterion data. Predictor and criterion data are merged to form the IMT, In-Unit and Administrative Validation Samples.

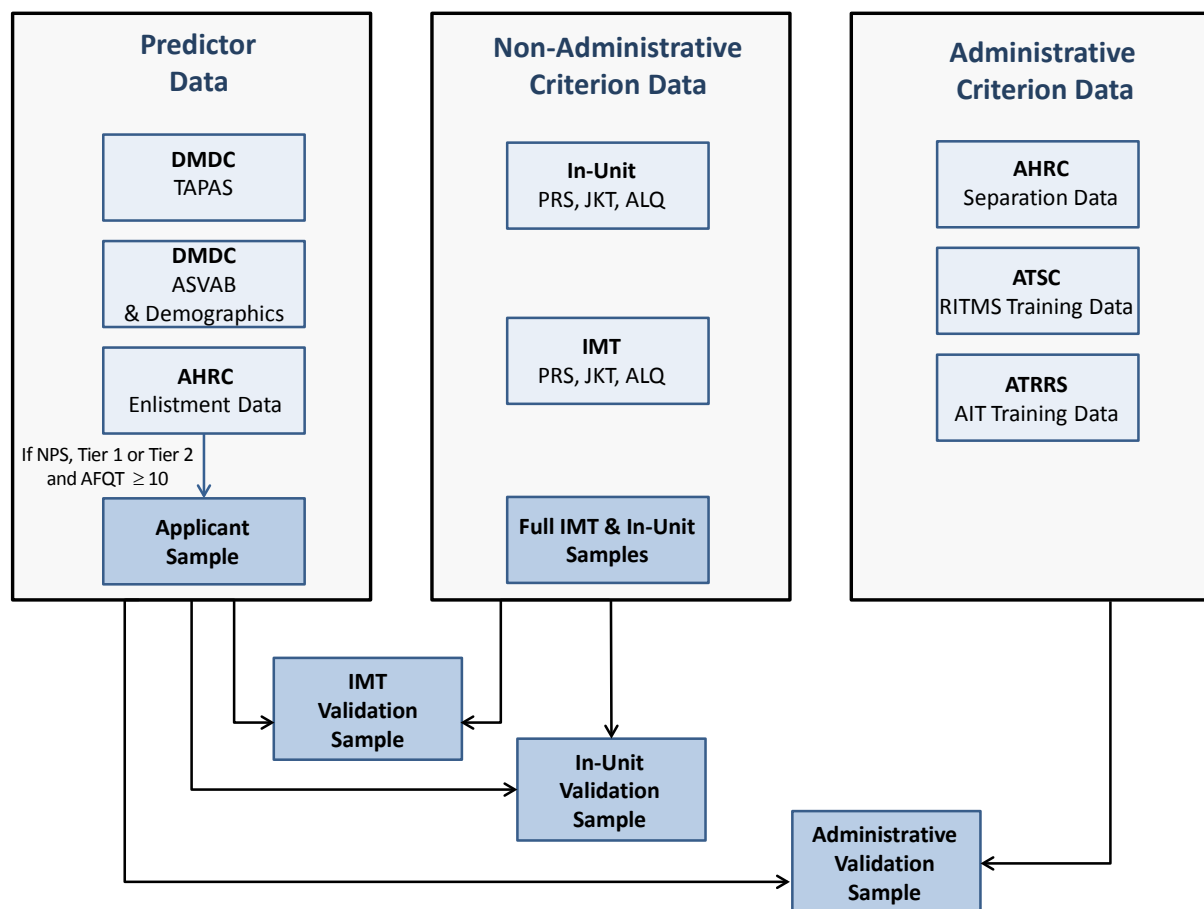


Figure 2.1. Overview of TOPS data file merging and nested sample generation process.

Sample Filters

For the purpose of evaluating the effectiveness of TAPAS, exclusions to the analysis samples were imposed based on AFQT score, education level, service history, and component.

AFQT Category

The ASVAB is a multiple aptitude battery of tests administered by the MEPCOM. Most military applicants take the computer adaptive version of ASVAB (i.e., the CAT-ASVAB). Scores on the ASVAB tests are combined to create composite scores for use in selecting applicants into the Army and classifying them into an MOS. The AFQT, the composite used for selecting applicants into the Army, comprises the Verbal Expression⁵ (VE), Arithmetic Reasoning (AR), and Math Knowledge (MK) tests ($AFQT = 2*VE + AR + MK$). Applicants must meet a minimum AFQT score to be eligible to serve in the military, and the Services favor high-scoring applicants for enlistment. AFQT percentile scores are divided into the following categories:⁶

- Category I (93-99)
- Category II (65-92)
- Category IIIA (50-64)
- Category IIIB (31-49)
- Category IV (10-30)
- Category V (1-9)

AFQT Category V Soldiers are not eligible for enlistment. Category IV accessions are greatly restricted, some restriction is placed on accessing Category IIIB accessions, and priority is given to Category I-IIIA accessions. The Applicant Sample excludes Soldiers with an AFQT score of less than 10 (i.e., Category V; $n = 8,564$).

For classification, scores on the ASVAB tests are combined to form 10 Aptitude Area (AA) composites. An applicant must receive a minimum score on the MOS-relevant AA composite(s) to qualify for classification to that MOS. For example, applicants must score a 95 in both the Electronics (EL) and Signal Communications (SC) AA composites to qualify as a Signal Support Specialist (25U). Descriptive statistics for the AFQT, ASVAB tests, and AA composites are reported in Appendix A. AFQT category frequencies are reported in Tables 2.1 and 2.2.

Education Tier

In the early 1980s, the Department of Defense initiated a detailed study of the relationship between educational credentials, other background characteristics, and adaptability for military service. The results supported a three-tier classification of educational credentials including:

⁵ Verbal Expression is a scaled combination of the Word Knowledge (WK) and Paragraph Comprehension (PC) tests.

⁶ For more information on ASVAB scoring, see the official website of the ASVAB, www.officialasvab.com.

- Tier 1 – Primarily high school diploma and higher (e.g., individuals currently in high school or college, college graduates, adult/alternative diplomas, home school diplomas)⁷
- Tier 2 – Primarily non-diploma graduate (e.g., GED certificants, vocational-technical certificants, non-traditional high school credential holders)
- Tier 3 – No educational credential (i.e., individuals not currently attending high school and do not possess a high school diploma or alternate credential)

Consistent with Army policy, which specifies that Soldiers classified as Tier 3 are ineligible for accession, the Applicant Sample excludes Tier 3 Soldiers and those with unknown values ($n = 9,462$).⁸ Subsequent analyses report results separately for Tier 1 and Tier 2 Soldiers.

Service History

Because the TOPS program is designed to predict first term Soldier performance, individuals with prior service history are excluded from the analysis samples ($n = 9,520$).

Service Component

The Applicant Sample includes Soldiers from all Army components – Regular Army (RA), U.S. Army Reserve (USAR), and U.S. Army National Guard (ARNG). For most analyses, Soldiers from all components are included. However, for analyses involving separation data, results are only presented for the Regular Army Soldiers.

Description of Analysis Samples

Table 2.1 summarizes the full TAPAS sample by the key variables that were used to create the analysis samples. Among the 486,310 applicants in the total unfiltered sample, 443,229 (91.1%) met the criteria for the Applicant Sample (i.e., non-prior service, Education Tier 1 or 2, and minimum AFQT score of 10). A detailed breakout of background and demographic characteristics observed in the analysis samples appears in Table 2.2.

⁷ In 2012, the Department of Defense announced that applicants who score 50 or higher on the AFQT and possess diplomas from home schools, virtual/distance learning, and adult/alternative schools will receive Tier 1 enlistment priority.

⁸ Starting with the December 2013 TOPS data file, we reconfigured the data sources used to best capture a Soldier's education tier status at the time of his or her accession. DMDC records now serve as the primary source of this information. As a result, figures for education tier reported in the current report differ from corresponding figures in previous reports. The differences are generally minor and do not impact the overall results.

Table 2.1. Full TAPAS Sample Characteristics

Variables	<i>n</i>	% of Total Sample (<i>N</i> = 486,310)
<i>Education Tier</i>		
Tier 1	451,371	92.8
Tier 2	25,477	5.2
Tier 3	6,633	1.4
Unknown	2,829	0.6
<i>Prior Service</i>		
Yes	9,520	2.0
No	476,790	98.0
<i>Military Occupational Specialty</i>		
11B/11C/11X/18X	42,074	8.7
19K	2,938	0.6
25B	2,869	0.6
25N	714	0.1
25U	4,768	1.0
31B	11,569	2.4
42A	6,868	1.4
68W	13,439	2.8
88M	12,429	2.6
91B	12,937	2.7
Other	155,198	31.9
Unknown ^a	220,507	45.3
<i>AFQT Category</i> ^b		
I	28,825	5.9
II	132,828	27.3
IIIA	91,967	18.9
IIIB	151,947	31.2
IV	72,138	14.8
V	8,564	1.8
Unknown ^a	41	0.0
<i>Contract Status</i>		
Signed	286,368	58.9
Not signed	199,942	41.1
Applicant Sample ^c	443,229	91.1

^a Generally, when the MOS or AFQT Category is unknown, it is either because the information was not yet available in the data sources on which the December 2013 data file was based or because the respondent did not access into the Army.

^b AFQT Categories IIIB and IV are oversampled. Values presented are not representative of Army accessions.

^c The Applicant Sample size is smaller than the total TAPAS sample because it is limited to non-prior service, Education Tier 1 and 2, and AFQT ≥ 10 applicants.

Table 2.2. Background and Demographic Characteristics of the TOPS Samples

Characteristics	Applicant ^a		Administrative Validation ^b		IMT Validation ^c		In-Unit Validation ^d	
	<i>n</i> = 443,229		<i>n</i> = 226,055		<i>n</i> = 23,495		<i>n</i> = 1,965	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
<i>Component</i>								
Regular	252,781	57.0	131,820	58.3	14,506	61.7	1,960	99.7
ARNG	135,335	30.5	66,040	29.2	6,772	28.8	5 ^e	0.3
USAR	55,113	12.4	28,188	12.5	2,217	9.4	--	--
<i>Education Tier</i>								
Tier 1	419,426	94.6	217,610	96.3	22,608	96.2	1,902	96.6
Tier 2	23,803	5.4	8,445	3.7	887	3.8	63	3.2
<i>Military Occupational Specialty</i>								
11B/11C/11X/18X	39,087	8.8	35,899	15.9	9,104	38.7	468	23.8
19K	2,804	0.6	2,594	1.1	846	3.6	41	2.1
25B	2,680	0.6	2,377	1.1	912	3.9	16	0.8
25N	673	0.2	631	0.3	143	0.6	11	0.6
25U	4,541	1.0	4,002	1.8	17	0.1	33	1.7
31B	10,824	2.4	9,545	4.2	3,364	14.3	55	2.8
42A	6,453	1.5	5,629	2.5	1,304	5.6	53	2.7
68W	12,691	2.9	11,572	5.1	3,877	16.5	82	4.2
88M	11,649	2.6	10,136	4.5	2,836	12.1	91	4.6
91B	12,134	2.7	10,500	4.6	551	2.3	103	5.2
Other	145,477	32.8	132,875	58.8	541	2.3	1,012	51.5
Unknown	194,216	43.8	295	0.1	--	--	--	--
<i>AFQT Category^f</i>								
I	25,976	5.9	15,013	6.6	1,660	7.1	109	5.5
II	122,456	27.6	71,466	31.6	8,725	37.1	562	28.6
IIIA	85,239	19.2	47,733	21.1	4,921	20.9	435	22.1
IIIB	142,334	32.1	80,188	35.5	7,177	30.5	765	38.9
IV	67,224	15.2	11,655	5.2	1,012	4.3	94	4.8
<i>Gender</i>								
Female	90,896	20.5	38,943	17.2	3,093	13.2	250	12.7
Male	341,243	77.0	181,525	80.3	19,997	85.1	1,689	86.0
Missing	11,090	2.5	5,587	2.5	405	1.7	26	1.3
<i>Race</i>								
African American	95,539	21.6	42,951	19.0	3,452	14.7	402	20.5
American Indian	3,560	0.8	1,706	0.8	187	0.8	13	0.7
Asian	16,161	3.6	8,202	3.6	766	3.3	79	4.0
Hawaiian/Pacific	1,847	0.4	1,016	0.4	118	0.5	10	0.5
Caucasian	310,954	70.2	166,368	73.6	18,153	77.3	1,385	70.5
Multiple	1,804	0.4	1,002	0.4	102	0.4	7	0.4
Declined to Answer	13,364	3.0	4,810	2.1	717	3.0	69	3.5
<i>Ethnicity</i>								
Hispanic/Latino	67,600	15.3	33,219	14.7	2,960	12.6	243	12.4
Not Hispanic	362,801	81.9	188,783	83.5	19,932	84.8	1,667	84.8
Declined to Answer	12,828	2.9	4,053	1.8	603	2.6	55	2.8

^a Limited to applicants who had no prior service, Education Tier 1 or 2, and AFQT ≥ 10 ; served as the core analysis sample.

^b Soldiers in Applicant Sample with at least one criterion record (i.e., schoolhouse, in-unit, ATRRS, RITMS, or attrition).

^c Soldiers in Applicant Sample with criterion data collected at schoolhouses.

^d Soldiers in Applicant Sample with criterion data collected in units.

^e We believe these Soldiers were on active duty when the in-unit data collections were taking place.

^f AFQT Categories IIIB and IV are oversampled. Values presented are not representative of Army accessions.

Across all analysis samples, a majority of the Soldiers are Regular Army, Education Tier 1. In addition, the samples are predominantly male, White, and non-Hispanic; however, a large number of Soldiers declined to provide information on race or ethnicity. Of the targeted MOS, 11B predominates, with nearly three times as many Soldiers as the next largest group (i.e., 68W). The least represented MOS include 19K, 25B, and 25N Soldiers.

The Administrative Validation Sample includes 226,055 Soldiers who meet all of the inclusion criteria for the TOPS Applicant Sample and also have at least one record in an administrative criterion data source (e.g., Army Training Requirements and Resources System [ATRRS], Resident Individual Training Management System [RITMS]). There are 62,176 Soldiers with IMT criteria data; however, only 23,495 were linked to an administrative TAPAS record and included in the IMT Validation Sample. Similarly, there are 5,493 Soldiers with in-unit data but only 1,965 of these Soldiers have matching TAPAS data and were included in the In-Unit Validation Sample. There are 339 Soldiers with a TAPAS record and both IMT and in-unit criterion data.

There are two primary reasons for the diminution of sample sizes between the Applicant Sample and the Administrative Validation samples. First, is the fact that many of the applicants did not access into the Army. Second, we rely on self-reported name and date of birth to match TAPAS records to the criterion data, which often results in unsuccessful matches. Further, fewer than half of the total number of Soldiers for whom we have IMT and in-unit criterion data are in the IMT and In-Unit Validation samples. In addition to cases lost due to unreliable reporting of the matching variables (name and date of birth), criterion testing started early in 2009 before TAPAS was being widely administered to applicants.

Sample sizes reported in all subsequent chapters and appendices are generally smaller than the figures reported here because of further data filtering or disaggregation that occurs for each particular analysis. For example, predictor and criterion scores were determined to be valid if they passed multiple data quality screens intended to identify unmotivated responding. Additional screens are analysis specific and have not yet been applied to the descriptive analysis of the samples described in this chapter. Further, a relatively small number of Soldiers in the Applicant Sample ($n = 1,646$) were administered an early version of the TAPAS and were excluded from analyses because of conceptual dissimilarities with subsequent TAPAS forms.

Summary

The TOPS analysis samples represent a combination of administrative, IMT, and in-unit data obtained from Soldiers, their supervisors and cadre, and archival sources at multiple points in time using a variety of data collection methods. The December 2013 full sample includes 486,310 applicants who took the TAPAS; however, some of them did not access into the Army or were ineligible for inclusion in the analyses based on their education status, AFQT score, component, or service history.

After excluding Education Tier 3, AFQT Category V, and prior service applicants from the full sample, the remaining 443,229 Soldiers were included in the TOPS Applicant Sample. This sample represents Soldiers who possess qualities that are most representative of applicants to the

Army. A majority of the Soldiers included in the sample are listed as Regular Army, Education Tier 1; and are predominantly male, White and non-Hispanic.

Additional analysis samples were created based on this initial sample; however, they include fewer Soldiers. Of the full Applicant Sample, 226,055 (51.0%) had a record in at least one of the administrative criterion data sources; 23,495 (5.3%) had IMT data collected from the schoolhouse and 1,965 (.04%) had in-unit criterion data. The applicant sample and validation samples were used in subsequent analyses presented in Chapter 4, 5, and 6 and associated appendixes.

CHAPTER 3: THE TAILORED ADAPTIVE PERSONALITY ASSESSMENT SYSTEM (TAPAS)

Stephen Stark, O. Sasha Chernyshenko, Christopher Nye, and Fritz Drasgow
(Drasgow Consulting Group)

Description

The purpose of this chapter is to describe the primary predictor measure being investigated in the TOPS IOT&E, the TAPAS (Drasgow, et al., 2012; Stark et al., 2010). The TAPAS is a personality measurement tool originally developed by Drasgow Consulting Group (DCG) under the Army's Small Business Innovation Research (SBIR) program. The system builds on the foundational work of the Assessment of Individual Motivation ([AIM]; White & Young, 1998) by incorporating features designed to promote resistance to faking and by measuring narrow personality constructs (i.e., facets) that are known to predict outcomes in work settings. The TAPAS uses methods from item response theory (IRT) to construct and score items. It can be administered in multiple formats: (a) as a fixed length, *non-adaptive test* where examinees respond to the same sequence of items or (b) as an *adaptive test* where each examinee responds to a unique sequence of items selected to maximize measurement accuracy for that specific examinee.

The TAPAS uses an IRT model for multidimensional pairwise preference (MDPP) items (Stark, Chernyshenko, & Drasgow, 2005) as the basis for constructing, administering, and scoring personality tests that are designed to reduce response distortion (i.e., faking) and yield normative scores even with tests of high dimensionality (Stark, Chernyshenko, & Drasgow 2012). TAPAS items consist of pairs of personality statements for which a respondent's task is to choose the one that is "more like me." The two statements constituting each item are matched in terms of social desirability and statement location (extremity), and often represent different personality facets. This approach makes it more difficult for examinees to determine which answers are better from the Army's perspective, and thus it is harder to "fake good" on all facets throughout the course of a test than it is with a single-statement Likert-type personality test. Stark et al. (2014) reported small mean differences in scores of individuals who might be motivated to increase their scores (i.e., Army applicants who were told that their score might affect their enlistment eligibility) compared to individuals not so motivated (Air Force applicants who were asked to complete the TAPAS for research purposes only). In short, the TAPAS' features make it more difficult for respondents to distort their responses to obtain more desirable scores.

The use of an IRT model also greatly increases the flexibility of the assessment process. A variety of test versions can be constructed to measure personality facets that are relevant to specific work contexts, and the measures can be administered via paper-and-pencil or computerized formats. If test content specifications (i.e., test blueprints) are comparable across versions, the respective scores can be readily compared because the metric of the statement parameters has already been established by calibrating response data obtained from a base or reference group (e.g., Army recruits). The same principle applies to adaptive testing, wherein each examinee receives a different set of items chosen specifically to reduce the error in his or

her facet scores at points throughout the exam. Adaptive item selection enhances test security because there is less overlap across examinees in terms of the items presented.

Another important feature of the TAPAS is that pools of statements representing over two dozen narrow personality facets are available. The initial TAPAS trait taxonomy was developed using the results of several large scale factor-analytic studies with the goal of identifying a comprehensive set of non-redundant narrow traits. Since then, additional facets have been added and these narrow traits, if necessary or desired, can be combined to form either the Big Five (the most common organization scheme for narrow personality traits) or any other number of broader traits (e.g., Integrity or Positive Core Self-Evaluations). This is advantageous for applied purposes because TAPAS versions can be created to fit a wide range of applications (both pre- and post-enlistment) and are not limited to a particular service branch or criterion. Selection of specific TAPAS facets can be guided by consulting the results of a meta-analytic study performed by DCG that mapped TAPAS facets to several important organizational criteria for military and civilian jobs (e.g., task proficiency, training performance, attrition) (Chernyshenko & Stark, 2007), as well as subsequent validation research. Table 3.1 presents the names of the TAPAS facets together with a description of a typical high scoring individual.

Scoring details and the criterion-related validation work that led to the inclusion of TAPAS in the TOPS IOT&E can be found in the *Expanded Enlistment Eligibility Metrics* report (Knapp & Heffner, 2010) and in earlier evaluation reports in this series (e.g., Knapp & Heffner, 2011).

Psychometric Properties of TAPAS Test Versions

As part of the TOPS IOT&E, nine versions of the TAPAS have been administered (see Table 3.2). The different versions have allowed ARI to explore the value of alternative facets and to retire the statement pools that were exposed in research settings. Currently, MEPS testing uses a statement pool developed solely for use by ARI and those agencies it authorizes (e.g., the other US military services). All versions created in August 2011 or later use ARI-owned statement pools. In the present report, the analyses reported in Chapters 5 and 6 are based on the five 15D versions, each administering 120 items (i.e., pairs of statements).⁹

As a test security measure, form equivalence information is provided in a limited distribution addendum. Scores have been standardized within TAPAS versions to enable cross-version analyses.

⁹ The three latest 13D versions (v9, v10, v11) were not filtered from the data file, but there were only six cases for which we have criterion data.

Table 3.1. TAPAS Facets Names and Definitions

Facet Name	Brief Description
Achievement	High scoring individuals are seen as hard working, ambitious, confident, and resourceful.
Adjustment	High scoring individuals are well adjusted, worry free, and handle stress well.
Adventure Seeking	High scoring individuals enjoy participating in extreme sports and outdoor activities.
Aesthetics	High scoring individuals appreciate various forms of art and music and participate in art-related activities more than most people.
Attention Seeking	High scoring individuals tend to engage in behaviors that attract social attention. They are loud, loquacious, entertaining, and even boastful.
Commitment to Serve	High scoring individuals identify with the military and have a strong desire to serve their country.
Consideration	High scoring individuals are affectionate, compassionate, sensitive, and caring.
Cooperation	High scoring individuals are pleasant, trusting, cordial, non-critical, and easy to get along with.
Courage	High scoring individuals stand up to challenges and are not afraid to face dangerous situations.
Curiosity	High scoring individuals are inquisitive and perceptive; they are interested in learning new information and attend courses and workshops whenever they can.
Dominance	High scoring individuals are domineering, “take charge” and are often referred to by their peers as “natural leaders.”
Even Tempered	High scoring individuals tend to be calm and stable. They don’t often exhibit anger, hostility, or aggression.
Ingenuity	High scoring individuals are inventive and can think “outside of the box.”
Intellectual Efficiency	High scoring individuals believe they process information and make decisions quickly; they see themselves (and they may be perceived by others) as knowledgeable, astute, or intellectual.
Non-Delinquency	High scoring individuals tend to comply with rules, customs, norms, and expectations, and they tend not to challenge authority.
Optimism	High scoring individuals have a positive outlook on life and tend to experience joy and a sense of well-being.
Order	High scoring individuals tend to organize tasks and activities and desire to maintain neat and clean surroundings.
Physical Conditioning	High scoring individuals tend to engage in activities to maintain their physical fitness and are more likely participate in vigorous sports or exercise.
Responsibility	High scoring individuals are dependable, reliable, and make every effort to keep their promises.

Table 3.1. (Continued)

Facet Name	Brief Description
Self Control	High scoring individuals tend to be cautious, levelheaded, able to delay gratification, and patient.
Selflessness	High scoring individuals are generous with their time and resources.
Situational Awareness	High scoring individuals pay attention to their surroundings and rarely get lost or surprised.
Sociability	High scoring individuals tend to seek out and initiate social interactions.
Team Orientation	High scoring individuals prefer working in teams and make people work together better.
Tolerance	High scoring individuals are interested in other cultures and opinions that may differ from their own. They are willing to adapt to novel environments and situations.
Virtue	High scoring individuals strive to adhere to standards of honesty, morality, and “good Samaritan” behavior.

Descriptive statistics and intercorrelations of individual TAPAS scale scores and composite scores are provided in Appendix A (Tables A.1-A.5). Also reported there are correlations of TAPAS scales with AFQT, ASVAB subtests, and Aptitude Area composites (Table A.13). Because most of the observed correlations between TAPAS scales and ASVAB subtests were in the -.20 to +.20 range, the two measures are judged to provide non-redundant information about applicants’ dispositions, which is advantageous in selection and classification contexts.

Table 3.2. TAPAS Versions by Administration Date

TAPAS Version	Dates Administered	# of Facets	Adaptive	# of Items
13D-CAT	May 4, 2009 to July 10, 2009	13	Yes	104
15D-Static	July 2010 to August 2011	15	No	120
15D-CAT v4	July 2010 to August 2011	15	Yes	120
15D-CAT v5	August 2011 to September 2013	15	Yes	120
15D-CAT v7	August 2011 to September 2013	15	Yes	120
15D-CAT v8	August 2011 to September 2013	15	Yes	120
13D-CAT v9	September 2013 – present	13	Yes	120
13D-CAT v10	September 2013 – present	13	Yes	120
13D-CAT v11	September 2013 – present	13	Yes	120

TAPAS Composites

An initial Education Tier 1 performance screen was developed from the TAPAS-95s scales for the purpose of testing in an applicant setting (Allen et al., 2010).¹⁰ This was accomplished by (a) identifying key criteria of most interest to the Army, (b) sorting these criteria into “can do” and “will do” categories (see below), and (c) selecting composite scales corresponding to the can do and will do criteria, taking into account both theoretical rationale and empirical results. Two unit-weighted TAPAS composites were initially developed: (a) Can-Do (for predicting technical training performance and completion) and (b) Will-Do (for predicting attrition and motivation-based performance). These composites were used operationally from January 2010 – September 2013.

A subsequent set of composites was developed by DCG and includes three regression-weighted scores: (a) Can-Do, (b) Will-Do, and (c) Adaptation (for predicting attrition). These scores became available for Army decision-making in September 2013. More information about how the new composites were developed is provided in a limited distribution addendum. Those interested in obtaining a copy of this addendum should contact the editors for further information. The specific facet scales comprising each TAPAS composite are close-hold information given the operational nature of this measure.

The criterion-related validation analyses in Chapter 5 use the new composite scores. Not all versions of the TAPAS include the scales comprising the Can-Do and Adaptation composites, so the sample sizes for analyses involving those scores are substantially lower than sample sizes for the Will-Do composite-related validation analyses.

Summary

The purpose of this chapter was to describe the primary predictor measure being evaluated in the TOPS IOT&E. The TAPAS is unique among personality measures because it uses forced-choice pairwise items and IRT to promote resistance to faking. Promising initial validation research conducted as part of EEEM has been followed by additional research showing the validity of TAPAS in operational settings (Nye et al., 2012).

¹⁰ TAPAS-95s was a paper-and-pencil, static version of the TAPAS used in the Army Class research.

CHAPTER 4: DESCRIPTION AND PSYCHOMETRIC PROPERTIES OF CRITERION MEASURES

Thomas B. Kiger, Joseph P. Caramagno, and Matthew C. Reeder (HumRRO)

Criterion scores to validate the TAPAS were derived from measures administered for purposes of this research and from administrative records. The research measures included data provided by the Soldiers in the form of job knowledge tests (JKTs) and a questionnaire measuring self-reported attitudes and performance (Army Life Questionnaire [ALQ]). Additionally, research measures include data provided by the Soldiers' cadre and supervisors through performance rating scales (PRS) created for research purposes. Criterion scores drawn from Soldiers' administrative records included separation status (i.e., attrition), Initial Military Training (IMT) completion, and IMT grades. Table 4.1 provides a description of each of these measures.

Table 4.1. Summary of IMT and In-Unit Criterion Measures

Criterion Measure	Description
<i>Soldier/ Cadre/ Supervisor Reported</i>	
Job Knowledge Tests (JKT)	The Warrior Tasks and Battle Drills (WTBD) JKT measures knowledge that is general to all enlisted Soldiers. MOS-specific JKTs measure Soldiers' knowledge of basic facts, principles, and procedures required of Soldiers in training for a particular MOS. Each JKT includes a mix of item formats (e.g., multiple-choice, multiple-response, and rank order).
Army Life Questionnaire (ALQ)	The ALQ measures Soldiers' self-reported attitudes and experiences in the Army. The IMT and in-unit versions are very similar.
Performance Rating Scales (PRS)	The IMT PRS measure Soldiers' performance in two domains: (a) MOS-specific (e.g., learns preventive maintenance checks and services, learns to troubleshoot vehicle and equipment problems) and (b) Army-wide (e.g., exhibits effort, supports peers, demonstrates physical fitness). The IMT PRS are completed by training cadre. In-unit PRS cover Army-wide dimensions only and are completed by supervisors.
<i>Administrative</i>	
Attrition	Separation data are obtained on participating Soldiers beginning at 3-months and at regular 3-month intervals thereafter. Attrition data through 36 months were available for the current sample.
Initial Military Training (IMT) Criteria	These data provide information about whether Soldiers restarted IMT and for what reasons, the number of times Soldiers restarted training, graduation status, and final school grades for Soldiers in Advanced Individual Training AIT.

In this chapter, we describe the criterion measures and composites, along with their distributional and psychometric properties. The descriptive statistics and correlations among the criteria (shown in Appendix B) are based on the Validation Sample (i.e., Education Tier 1 and 2, non-prior service, AFQT Category IV or above Soldiers with matching criterion data). Descriptive statistics and psychometric properties of the criterion measures for the full IMT and in-unit samples are reported in Appendix C.

Job Knowledge Tests (JKTs)

There are multiple JKTs used in the current research. All participating Soldiers are given a generic JKT called the Warrior Tasks and Battle Drills (WTBD). Additionally, Soldiers in some MOS are given MOS-specific JKTs. Two such JKTs were developed for this research project and all others (including the WTBD JKT) were developed through previous ARI research projects: Select21 (Collins, Le, & Schantz, 2005) or Army Class (Moriarty, Campbell, Heffner, & Knapp, 2009). Currently, there are MOS-specific JKTs for the following jobs: Infantry (11B/C/X + 18X), Armor (19K), Military Police (31B), Health Care Specialist (68W), Light Wheel Vehicle Mechanic (91B), Motor Transport Operator Soldiers (88M) Signal Support Systems Specialist (25U), and Human Resources Specialist (42A). Most of the JKT items are in a multiple-choice format with two to four response options. However, other formats, such as multiple-response (i.e., check all that apply), rank ordering, and matching are also used. The items use visual images to make them more realistic and reduce reading requirements for the test.

A single, overall raw score was computed for each JKT by summing the total number of points Soldiers earned across the JKT items and computing a percent correct score based on the maximum number of points that could be obtained on each test. For the criterion-related validity analyses, we converted the total raw score to a standardized score (or z-score) by standardizing the scores *within* each MOS. A JKT score was flagged not included in analysis if the Soldier (a) omitted more than 10% of the assessment items, (b) took fewer than 5 minutes to complete the entire assessment, or (c) selected an implausible response to one of the careless responding items (Knapp et al., 2012). Table 4.2 lists the reliability estimates for the MOS-Specific JKTs and the WTBD JKT for the IMT and in-unit samples.

Table 4.2. Reliability Estimates of the Job Knowledge Tests (JKTs) in the IMT and In-Unit Validation Samples

Domain/JKT	<i>n</i>	<i>α</i>
IMT		
<i>MOS-Specific</i>		
11B/C/X + 18X	6,821	.78
19K	731	.78
31B	2,973	.76
42A	1,037	.75
68W	3,251	.88
88M	2,165	.76
91B	369	.89
<i>WTBD (Army-Wide)^a</i>	21,198	.67
In-Unit		
<i>MOS-Specific</i>		
11B/C/X + 18X	403	.67
<i>WTBD (Army-Wide)^a</i>	1,893	.55

Note. WTBD = Warrior Tasks and Battle Drills. Statistics based on fewer than 100 cases are not reported.

^a The WTBD JKTs are more heterogeneous in content than the MOS-specific JKTs, so would be expected to have lower alphas.

Table 4.3 lists the descriptive statistics for the IMT WTBD and MOS-specific JKTs by education tier and Table 4.4 lists the descriptive statistics for the in-unit WTBD and MOS-specific JKTs by education tier.

Table 4.3. Descriptive Statistics for the Job Knowledge Tests (JKTs) in the IMT Validation Sample by Education Tier

Domain/JKT	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>r_{WTBD}</i>	<i>r_{AFQT}</i>
Tier 1 + 2 (Combined)							
<i>MOS-Specific</i>							
11B/C/X + 18X	6,821	60.28	10.56	25.58	86.96	.59	.44
19K	731	61.65	11.77	20.29	86.15	.49	.30
31B	2,973	67.14	8.50	33.33	91.26	.50	.47
42A	1,037	54.58	12.43	16.67	85.19	.54	.42
68W	3,251	72.39	10.42	26.00	92.39	.51	.27
88M	2,165	63.67	9.96	31.94	88.89	.56	.40
91B	369	58.36	13.08	27.37	90.72	.54	.26
All MOS Combined ^a	17,347	63.82	11.58	16.67	92.39	.57	.47
<i>WTBD (Army-Wide)</i>	21,198	64.01	12.44	6.45	97.30		.43
Tier 1							
<i>MOS-Specific</i>							
11B/C/X + 18X	6,517	60.31	10.55	25.58	86.96	.59	.44
19K	709	61.81	11.73	20.29	86.15	.49	.30
31B	2,890	67.08	8.51	33.33	91.26	.49	.46
42A	1,012	54.59	12.47	16.67	85.19	.54	.42
68W	3,138	72.42	10.44	26.00	92.39	.51	.27
88M	2,068	63.67	9.94	31.94	88.89	.55	.40
91B	356	58.19	13.02	27.37	90.72	.53	.27
All MOS Combined ^a	16,690	63.85	11.58	16.67	92.39	.57	.47
<i>WTBD (Army-Wide)</i>	20,390	63.99	12.46	6.45	97.30		.43
Tier 2							
<i>MOS-Specific</i>							
11B/C/X + 18X	304	59.73	10.81	26.09	86.96	.55	.30
68W	113	71.61	9.87	33.70	84.78	.45	.22
All MOS Combined ^a	657	63.26	11.61	26.09	87.38	.58	.40
<i>WTBD (Army-Wide)</i>	808	64.64	12.05	16.13	93.55		.34

Note. *Ms*, *SDs*, *Min*, and *Max* reflect percent correct. WTBD = Warrior Tasks and Battle Drills. *r_{WTBD}* = correlation with WTBD JKT scores. *r_{AFQT}* = correlation with AFQT scores. Statistics based on fewer than 100 cases are not reported. All correlations are statistically significant ($p < .05$, one-tailed).

^a Includes 11B/C/X + 18X, 19K, 31B, 42A, 68W, 88M, and 91B.

Table 4.4. Descriptive Statistics for Job Knowledge Tests (JKTs) in the In-Unit Validation Sample by Education Tier

Domain/JKT	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>r</i> _{WTBD}	<i>r</i> _{AFQT}
Tier 1 + 2 (Combined)							
<i>MOS-Specific</i>							
11B/C/X + 18X	403	61.94	8.47	26.76	83.10	.56	.45
All MOS Combined ^a	763	64.02	10.06	26.76	90.74	.49	.44
<i>WTBD Army Wide</i>	1,893	64.60	11.13	22.22	96.15		.45
Tier 1							
<i>MOS-Specific</i>							
11B/C/X + 18X	389	62.02	8.34	35.05	83.10	.55	.44
All MOS Combined ^a	737	64.04	10.00	35.05	90.74	.49	.44
<i>WTBD Army Wide</i>	1,835	64.63	11.13	22.22	96.15		.45

Note. *Ms*, *SDs*, *Min*, and *Max* reflect percent correct. WTBD = Warrior Tasks and Battle Drills. *r*_{WTBD} = correlation with WTBD JKT scores. *r*_{AFQT} = correlation with AFQT scores. Results based on fewer than 100 cases are not reported. Tier 2 results are not reported because sample sizes are less than 100 for all variables. All correlations are statistically significant ($p < .05$, one-tailed).

^a Includes 11B/C/X + 18X, 19K, 31B, 42A, 68W, 88M, and 91B.

Performance Rating Scales (PRS)

The PRS, like the JKTs, also were adapted from or based on previous research (see Moriarty et al., 2009 for details). The IMT and in-unit PRS are fairly different, so they will be described separately.

IMT PRS

The IMT PRS target two domains of Soldier performance requirements: (a) Army-wide and (b) MOS-specific. The IMT PRS were completed by cadre members (drill sergeants, trainers) of participating Soldiers.

Over the course of the TOPS IOT&E, two versions of the IMT PRS were administered. Early IOT&E evaluations noted low interrater reliability estimates for the PRS (Moriarty & Bynum, 2011). Accordingly, several changes were made to the IMT instruments in an attempt to improve their psychometric characteristics. First, the number of scales for the Army-wide PRS was reduced from eight to five, paralleling the five scores generated from the original scales (Sparks & Peddie, 2013). No changes were made to the MOS-specific PRS scales; the number of scales ranged from five to nine, and a composite score was computed by averaging ratings across the individual scales for each MOS. Second, the rating scales for both the Army-wide and MOS-specific PRS was changed from a 7-point Behaviorally Anchored Rating Scale (BARS) to a 5-point relative scale format with scales ranging from 1 (Among the Weakest) to 5 (Among the Best). Ratings on the initial PRS rating scale were re-scaled to reflect the new 5-point scale. After doing so, we combined ratings data across the two versions of PRS to create a single PRS score. All IMT PRS results reported are based on data from both the initial and revised PRS, and

are expressed on a 5-point scale¹¹. Finally, cadre members also indicated their opportunity to observe each Soldier being rated using a 4-point “familiarity” scale. The initial PRS used a 3-point familiarity scale. This was changed to a 4-point scale to enable raters to more clearly indicate their ability to judge each Soldier’s performance.

Table 4.5 compares the estimated interrater reliability for the initial and revised versions of the IMT PRS. The interrater reliability estimates were generally low (.30 or less) for the initial version of the PRS. As hoped, the revised PRS showed increased interrater reliability with many of the reliability estimates more than doubling. However, the estimates were still low. We attribute these low coefficients to several interrelated issues. First, the number of ratees per rater is high, averaging about 22 with 40% of the raters providing the maximum allowable number of ratings (30 ratees). As a result, raters likely became fatigued during the rating task. Second, within-rater variance was generally limited, perhaps reflecting raters’ inability to differentiate among individual Soldiers. Third, these data collections were not proctored, unlike prior research (e.g., Knapp & Heffner, 2009, 2010). While generally the estimates were still low, in a review across 10 military studies, Van Iddekinge, Roth, Putka, and Lanivich (2011) found that the average reliability among raters was .45 compared to .66 in civilian samples. Given the complication associated with collecting these ratings, the interrater reliability estimates for several of the scales are not drastically lower than what has been reported in other studies. The low IRRs in this, and other samples, may reflect raters viewing different samples of Soldiers’ performance, and as such some of the low consistency between-raters may reflect true differences in performance (e.g., Putka, Hoffman, & Carter, 2014).

Table 4.5. Interrater Reliability Estimates for the IMT Performance Rating Scales (PRS)

<i>Army-Wide</i>	Initial PRS		Revised PRS	
	<i>n_t</i>	<i>IRR</i>	<i>n_t</i>	<i>IRR</i>
Adjustment to the Army	12,551	.14	4,053	.28
Effort & Personal Discipline	12,278	.17	4,053	.37
MOS Qualification Knowledge	11,096	.10	3,521	.25
Physical Fitness & Bearing	12,179	.19	4,043	.18
Working with Others	12,211	.15	4,049	.29
Overall Performance	12,082	.32	3,965	.31
<i>MOS-Specific</i>				
11B/C/X + 18X	3,986	.18	710	.25
19K	158	.41	115	.04
31B	1,799	.11	619	.49
68W	3,095	.01	328	.05
88M	671	.00	0	--
91B	255	.11	0	--

Note. *n_t* = number of Soldiers who have been rated. IRR = Interrater reliability. Interrater reliability was assessed using *G(q,k)*, a reliability metric designed specifically for studies like TOPS where the measurement design is ill-structured (Putka, Le, McCloy, & Diaz, 2008). IRR estimates were not estimated if 30 or fewer Soldiers were rated by more than one cadre member.

Table 4.6 summarizes the descriptive statistics for IMT PRS by education tier. A Soldier’s PRS

¹¹ The initial rating scale was converted from a 7-point scale to a 5-point scale by identifying meaningful cuts along the 7-point scale and comparing percentiles of the initial PRS to the new PRS to ensure the cuts points produced consistent percentiles in each group. The following conversions were used: 1.00-2.99 = 1; 3.00-4.99 = 2; 5.00-5.99 = 3; 6.00-6.99 = 4; 7.00 = 5.

ratings were not included in the analyses if the rater (a) indicated he or she had little opportunity to observe this Soldier, (b) omitted more than 10% of the assessment items, (c) indicated that he or she had not observed the Soldier on more than 50% of the dimensions, or (d) engaged in “flat responding”—that is, if the rater rated 10 or more Soldiers on a particular scale and 90% or more of those rating profiles were exactly the same. Mean ratings were above the mid-point, a consistent finding in prior Army research involving performance ratings (e.g., Campbell & Knapp, 2001; Knapp & Tremble, 2007; Moriarty & Bynum, 2011). The IMT PRS were also highly intercorrelated (see Appendix B).

Table 4.6. Descriptive Statistics for the Performance Rating Scales (PRS) in the IMT Validation Sample by Education Tier

Domain/PRS	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
Tier 1 + Tier 2 (Combined)					
<i>Army-Wide</i>					
Adjustment to the Army	6,739	3.46	0.98	1.00	5.00
Effort & Personal Discipline	6,742	3.29	0.97	1.00	5.00
MOS Qualification					
Knowledge	6,253	3.40	0.93	1.00	5.00
Physical Fitness & Bearing	6,726	3.34	0.98	1.00	5.00
Working with Others	6,731	3.30	0.97	1.00	5.00
Overall Performance	6,670	3.57	0.84	1.00	5.00
<i>MOS-Specific</i>					
11B/C/X + 18X	1,719	3.21	0.83	1.00	5.00
19K	214	3.44	0.68	1.71	5.00
31B	1,130	3.32	0.74	1.00	5.00
42A	344	3.71	0.68	2.00	5.00
68W	881	3.10	0.90	1.00	5.00
88M	119	2.88	0.79	1.20	5.00
All MOS Combined ^a	4,460	3.26	0.83	1.00	5.00
Tier 1					
<i>Army-Wide</i>					
Adjustment to the Army	6,489	3.46	0.98	1.00	5.00
Effort & Personal Discipline	6,492	3.29	0.97	1.00	5.00
MOS Qualification					
Knowledge	6,014	3.40	0.93	1.00	5.00
Physical Fitness & Bearing	6,476	3.34	0.98	1.00	5.00
Working with Others	6,481	3.30	0.97	1.00	5.00
Overall Performance	6,423	3.58	0.84	1.00	5.00
<i>MOS-Specific</i>					
11B/C/X + 18X	1,643	3.22	0.84	1.00	5.00
19K	205	3.45	0.69	1.71	5.00
31B	1,102	3.33	0.75	1.00	5.00
42A	337	3.72	0.68	2.00	5.00
68W	848	3.10	0.89	1.00	5.00
88M	113	2.88	0.78	1.20	5.00
All MOS Combined ^a	4,300	3.26	0.83	1.00	5.00

Table 4.6. (Continued)

Domain/PRS	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
Tier 2					
<i>Army-Wide</i>					
Adjustment to the Army	250	3.36	0.98	1.00	5.00
Effort & Personal Discipline	250	3.19	0.96	1.00	5.00
MOS Qualification					
Knowledge	239	3.35	0.87	1.00	5.00
Physical Fitness & Bearing	250	3.14	0.95	1.00	5.00
Working with Others	250	3.27	1.00	1.00	5.00
Overall Performance	247	3.43	0.85	1.00	5.00
<i>MOS-Specific</i>					
All MOS Combined ^a	160	3.12	0.77	1.00	5.00

Note. Ratings on PRS range from 1 and 5. PRS ratings from supervisors with a familiarity rating of 1 (“I have had little opportunity to observe this Soldier”) were excluded from analyses. Results based on fewer than 100 cases are not reported.

^a Includes 11B/C/X + 18X, 11B, 19K, 31B, 42A, 68W, 88M, and 91B.

In-Unit PRS

The in-unit PRS only target Army-wide dimensions of performance (i.e., there are no MOS-specific in-unit PRS) and include 13 performance dimensions, plus a Leadership Potential scale (see Table 4.7). One scale with poor psychometric properties was replaced in 2011 with the Adjustment to Army Life scale, comparable to the corresponding IMT scale. Ratings on several of the individual scales were combined to form four PRS composites and three scales were left as single-item dimensions. Cronbach’s alpha coefficients for the in-unit PRS composite scales are reported in Table 4.7. The in-unit PRS have consistently employed the 7-point BARS format used for the initial IMT scales. The revised 4-point “familiarity” scale used in the new IMT PRS also is used with the in-unit PRS. The majority of Soldiers in units were rated by only one supervisor, so interrater reliability estimates were not calculated. Table 4.8 reports the basic descriptive statistics for the in-unit Army-wide PRS by performance domain and education tier.

Table 4.7. In-Unit Army-Wide Performance Rating Scale Dimensions and Composite Scores

In-Unit Rating Composites	<i>α</i>
<i>Can Do</i>	.90
Performing Core Warrior Tasks	
Performing MOS-Specific Tasks	
Processing Information	
Solving Problems	
<i>Effort & Personal Discipline</i>	.80
Exhibiting Effort	
Exhibiting Personal Discipline	
<i>Working with Others</i>	.66
Communicating with Others	
Contributing to the Team	
<i>Self-Management</i>	.77
Following Safety Procedures	
Developing Own Skills	
Managing Personal Matters	

Table 4.7. (Continued)

In-Unit Rating Single-Item Dimensions
Adjustment to Army Life
Physical Fitness and Bearing
Overall Leadership Potential Rating

Note. Of the seven performance ratings used in analyses, four are composites of multiple dimensions and three are single dimension ratings.

Table 4.8. Descriptive Statistics for Army-Wide Performance Rating Scales (PRS) in the In-Unit Validation Sample

PRS Dimensions/Composites	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
Tier 1 + Tier 2 (Combined)					
Can Do ^a	1,367	4.88	1.30	1.00	7.00
Effort & Personal Discipline ^a	1,367	5.25	1.38	1.00	7.00
Physical Fitness & Bearing	1,361	5.30	1.57	1.00	7.00
Self-Management ^a	1,364	5.33	1.16	1.00	7.00
Working with Others ^a	1,367	5.30	1.24	1.00	7.00
Adjustment to Army Life	1,145	5.32	1.53	1.00	7.00
Overall Leadership Potential	1,327	4.78	1.68	1.00	7.00
Tier 1					
Can Do ^a	1,323	4.89	1.29	1.00	7.00
Effort & Personal Discipline ^a	1,323	5.26	1.36	1.00	7.00
Physical Fitness & Bearing	1,317	5.32	1.56	1.00	7.00
Self-Management ^a	1,321	5.34	1.15	1.00	7.00
Working with Others ^a	1,323	5.31	1.24	1.00	7.00
Adjustment to Army Life	1,107	5.33	1.52	1.00	7.00
Overall Leadership Potential	1,285	4.78	1.67	1.00	7.00

Note. Ratings on PRS range from 1 and 7. PRS ratings from supervisors with a familiarity rating of 1 (“I have had little opportunity to observe this Soldier”) were excluded from analyses. Tier 2 results are not reported because sample sizes are less than 100 for all PRS.

^a Ratings composites comprise two or more Army-wide PRS.

Army Life Questionnaire (ALQ)

The ALQ was designed to measure Soldiers’ self-reported attitudes and experiences in the Army. Earlier forms of the training and in-unit versions of the ALQ (Van Iddekinge, Putka, & Sager, 2005) were modified slightly for use in the TOPS IOT&E. The ALQ includes scales that cover (a) Soldiers’ commitment and retention-related attitudes and (b) Soldiers’ performance and adjustment. Each ALQ scale is measured in a variety of different ways depending on the scale. The Army Physical Fitness Test (APFT) score is a write-in single item. Training Achievements, Training Failure, (both of which appear only on the IMT version of the ALQ), and Disciplinary Incidents are simply a count of the “yes” responses to associated items. The remaining scales (see Table 4.9) are composed of Likert-type response scales and are scored by computing the mean of the constituent item scores. Most scales appear on both the IMT and in-unit versions of the scales, though the IMT version has two unique Likert-based scales (i.e., Normative Commitment and Army Life Adjustment) and the in-unit version has one unique Likert-based scale (i.e., MOS satisfaction).

ALQ data were flagged as unusable if the Soldier (a) omitted more than 10% of the assessment items, (b) took fewer than 5 minutes to complete the entire assessment, or (c) chose an implausible response to the careless responding item. Tables 4.10 and 4.11 summarize the descriptive statistics for the ALQ scales by education tier for the IMT and in-unit samples, respectively.

Table 4.9. Army Life Questionnaire (ALQ) Likert-Type Scales

Scale Name	Description	Number of Items	Example Item	Likert Scale Anchors	IMT α	In-Unit α
Affective Commitment	Measures Soldiers' emotional attachment to the Army.	7	I feel like I am part of the Army 'family.'	1 (strongly disagree) to 5 (strongly agree)	.86	.88
Normative Commitment ^a	Measures Soldiers' feelings of obligation toward staying in the Army until the end of their current term of service.	5	I would feel guilty if I left the Army before the end of my current term of service.	1 (strongly disagree) to 5 (strongly agree)	.78	
Career Intentions	Measures Soldiers' intentions to reenlist and to make the Army a career.	3	How likely is it that you will make the Army a career?	Varies by item: 1 (strongly disagree) to 5 (strongly agree); 1 (not at all confident) to 5 (extremely confident); 1 (extremely unlikely) to 5 (extremely likely)	.91	.93
Reenlistment Intentions	Measures Soldiers' intention to reenlist in the Army.	4	How likely is it that you will leave the Army after completing your current term of service?	1 (strongly disagree) to 5 (strongly agree)	.83	.81
Attrition Cognitions	Measures the degree to which Soldiers think about attriting before the end of their first term.	4	How likely is it that you will complete your current term of service?	Varies by item: 1 (strongly disagree) to 5 (strongly agree); 1 (never) to 5 (very often)	.75	.80
Army Life Adjustment ^a	Measures Soldiers' transition from civilian to Army life.	9	Looking back, I was not prepared for the challenges of training in the Army.	1 (strongly disagree) to 5 (strongly agree)	.87	
MOS Fit	Measures Soldiers' perceived fit with their MOS.	9	My MOS provides the right amount of challenge for me.	1 (strongly disagree) to 5 (strongly agree)	.92	.93
Army Fit ^b	Measures Soldiers' perceived fit with the Army.	8	The Army is a good match for me.	1 (strongly disagree) to 5 (strongly agree)	.86	.77
MOS Satisfaction ^c	Measures Soldiers' satisfaction with their MOS	9	My MOS allows me to perform the kind of work I want to do.	1 (strongly disagree) to 5 (strongly agree)		.93

Note. α = coefficient alpha.

^a Appears only on the IMT ALQ.

^b Scale has 6 items on the in-unit ALQ.

^c Appears only on the in-unit ALQ.

Table 4.10. Descriptive Statistics and Reliability Estimates for the Army Life Questionnaire (ALQ) in the IMT Validation Sample by Education Tier

Domain/Scale	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
Tier 1 + Tier 2 (Combined)					
<i>Retention</i>					
Army Career Intentions	22,171	3.19	1.09	1.00	5.00
Affective Commitment	22,171	3.89	0.67	1.00	5.00
Army Fit	22,171	4.09	0.60	1.00	5.00
Army Reenlistment Intentions	22,171	3.48	0.95	1.00	5.00
Attrition Cognition	22,171	1.51	0.59	1.00	5.00
MOS Fit	22,171	3.78	0.83	1.00	5.00
Army Life Adjustment	22,171	4.09	0.66	1.00	5.00
Normative Commitment	22,171	4.18	0.69	1.00	5.00
<i>Achievement/Performance</i>					
Disciplinary Incidents (#)	20,488	0.28	0.62	0.00	7.00
Disciplinary Incidents (Y/N)	20,488	0.21	0.41	0.00	1.00
APFT Score	21,958	252.02	30.57	104.00	300.00
Training Achievement (#)	22,161	0.40	0.61	0.00	2.00
Training Failure (#) ^a	21,486	0.08	0.28	0.00	2.00
Tier 1					
<i>Retention</i>					
Army Career Intentions	21,335	3.18	1.09	1.00	5.00
Affective Commitment	21,335	3.89	0.67	1.00	5.00
Army Fit	21,335	4.08	0.60	1.00	5.00
Army Reenlistment Intentions	21,335	3.48	0.95	1.00	5.00
Attrition Cognition	21,335	1.51	0.59	1.00	5.00
MOS Fit	21,335	3.78	0.83	1.00	5.00
Army Life Adjustment	21,335	4.09	0.66	1.00	5.00
Normative Commitment	21,335	4.17	0.69	1.00	5.00
<i>Achievement/Performance</i>					
Disciplinary Incidents (#)	19,699	0.27	0.61	0.00	7.00
Disciplinary Incidents (Y/N)	19,699	0.21	0.41	0.00	1.00
APFT Score	21,134	252.29	30.53	104.00	300.00
Training Achievement (#)	21,325	0.41	0.61	0.00	2.00
Training Failure (#) ^a	20,681	0.07	0.28	0.00	2.00
Tier 2					
<i>Retention</i>					
Army Career Intentions	836	3.37	1.09	1.00	5.00
Affective Commitment	836	4.00	0.66	1.00	5.00
Army Fit	836	4.14	0.61	1.25	5.00
Army Reenlistment Intentions	836	3.59	0.95	1.00	5.00
Attrition Cognition	836	1.47	0.58	1.00	5.00
MOS Fit	836	3.84	0.82	1.00	5.00
Army Life Adjustment	836	4.12	0.66	1.22	5.00
Normative Commitment	836	4.25	0.65	1.80	5.00
<i>Achievement/Performance</i>					
Disciplinary Incidents (#)	789	0.33	0.72	0.00	6.00
Disciplinary Incidents (Y/N)	789	0.23	0.42	0.00	1.00
APFT Score	824	245.17	30.69	136.00	300.00
Training Achievement (#)	836	0.35	0.56	0.00	2.00
Training Failure (#) ^a	805	0.08	0.31	0.00	2.00

^a Training Failure is based on the total number of affirmative responses to whether a Soldier restarted from Basic Combat Training (BCT) or One Station Unit Training (OSUT) or whether a Soldier repeated a block or module at AIT or OSUT.

Table 4.11. Descriptive Statistics and Reliability Estimates for the Army Life Questionnaire (ALQ) in the In-Unit Validation Sample by Education Tier

Domain/Setting/Scale	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
Tier 1 + Tier 2 (Combined)					
<i>Retention</i>					
Army Career Intentions	1,926	2.60	1.21	1.00	5.00
Affective Commitment	1,926	3.57	0.80	1.00	5.00
Army Fit	1,926	3.90	0.70	1.50	5.00
Army Reenlistment Intentions	1,926	2.99	1.18	1.00	5.00
Attrition Cognition	1,926	1.69	0.75	1.00	5.00
MOS Fit	1,926	3.30	0.93	1.00	5.00
MOS Satisfaction	1,926	3.55	0.90	1.00	5.00
<i>Achievement/Performance</i>					
Disciplinary Incidents (#)	1,926	0.40	0.86	0.00	7.00
Disciplinary Incidents (Y/N)	1,926	0.25	0.43	0.00	1.00
APFT Score	1,888	248.73	32.10	118.00	300.00
Tier 1					
<i>Commitment & Fit</i>					
Army Career Intentions	1,864	2.60	1.20	1.00	5.00
Affective Commitment	1,864	3.56	0.80	1.00	5.00
Army Fit	1,864	3.90	0.70	1.50	5.00
Army Reenlistment Intentions	1,864	2.97	1.17	1.00	5.00
Attrition Cognition	1,864	1.69	0.74	1.00	5.00
MOS Fit	1,864	3.30	0.93	1.00	5.00
MOS Satisfaction	1,864	3.54	0.90	1.00	5.00
<i>Achievement/Performance</i>					
Disciplinary Incidents (#)	1,864	0.38	0.80	0.00	7.00
Disciplinary Incidents (Y/N)	1,864	0.24	0.43	0.00	1.00
APFT Score	1,827	248.73	32.09	118.00	300.00

Note. Tier 2 results are not reported because sample sizes are less than 100 for all ALQ scales.

Administrative Criteria

Attrition

Attrition is a broad category that encompasses involuntary and voluntary separations for a variety of reasons (e.g., underage enlistment, conduct, family concerns, drugs or alcohol, performance, physical standards or weight, mental disorder, or violations of the Uniform Code of Military Justice [UCMJ]). The reason for separation was determined by the Soldiers' Separation Program Designator (SPD) code. Soldiers who left the Army for reasons outside of their or the Army's control (e.g., death or serious injury incurred while performing one's duties) were excluded from our analyses. Separation data are reported for Regular Army Soldiers only. The current analyses cover attrition through 36 months of service. Table 4.12 summarizes the basic descriptive statistics for attrition by education tier.

AIT Grade

Soldiers' final AIT course grades were extracted from RITMS (Resident Individual Training Management System). Final grades from One Station Unit Training (OSUT) courses were excluded from the data file because the variance in the grades was highly restricted or based on a pass-fail metric that was redundant with the data from ATRRS (Army Training Requirements and Resources System). Final AIT Grade represents the cumulative grade across all courses administratively recorded for the Soldier.¹² A standardized version of Final AIT Grade was computed by standardizing each course grade for courses with 15 or more Soldiers. Table 4.13 summarizes the distribution of Final AIT Grade by education tier.

Training Restarts

Soldiers' IMT completion status and whether they graduated from IMT with training restarts or training failures were extracted from ATRRS. Training restarts identify Soldiers with at least one restart (i.e. must begin training again) during IMT. Training failures identify Soldiers that graduated IMT with at least one failure (i.e. failed a component of training). Failures are further divided into failures that were due to academic reasons versus those that were due to pejorative reasons. Soldiers who had not had an opportunity to complete their IMT at the time data were extracted were excluded from analyses. Table 4.13 presents the base rates of Soldiers with at least one training restart or training failure during IMT.

¹² In the rare instance where a Soldier has more than one administratively recorded AIT course grade, scores are averaged across courses for that Soldier.

Table 4.12. Base Rates for Attrition Criteria for Regular Army Soldiers in the Validation Sample by Education Tier

<i>Cumulative Attrition</i>	Tier 1 + Tier 2 (Combined)			Tier 1			Tier 2		
	<i>n</i>	<i>n</i> _{Attrit}	% _{Attrit}	<i>n</i>	<i>n</i> _{Attrit}	% _{Attrit}	<i>n</i>	<i>n</i> _{Attrit}	% _{Attrit}
6-Month	114,458	12,191	10.7	110,559	11,646	10.5	3,899	545	14.0
12-Month	89,315	11,678	13.1	85,908	11,068	12.9	3,407	610	17.9
24-Month	47,157	8,394	17.8	45,915	8,078	17.6	1,242	316	25.4
36-Month	17,452	4,084	23.4	17,236	4,011	23.3	216	73	33.8

Note. *n* = number of Soldiers with attrition data at the time data were extracted. *n*_{Attrit} = number of Soldiers who attrited at the specified months of service. %_{Attrit} = percentage of Soldiers who attrited through the specified months of service [$(n_{\text{Attrit}} / n) \times 100$].

Table 4.13. Base Rates or Descriptive Statistics for Administrative IMT Criteria in the Validation Sample

<i>Restarted Initial Military Training (IMT)</i>	Tier 1 + Tier 2 (Combined)			Tier 1			Tier 2		
	<i>n</i> ^a	<i>n</i> _{Restarted}	% _{Restarted}	<i>n</i> ^a	<i>n</i> _{Restarted}	% _{Restarted}	<i>n</i> ^a	<i>n</i> _{Restarted}	% _{Restarted}
IMT Restart	224,654	10,429	4.6	216,266	9,960	4.6	8,388	469	5.6
IMT Failure	129,517	17,126	13.2	124,109	16,205	13.1	5,408	921	17.0
For Pejorative Reasons	127,924	15,480	12.1	122,596	14,640	11.9	5,328	840	15.8
For Academic Reasons	124,839	12,436	10.0	119,755	11,840	9.9	5,084	596	11.7
<i>Final AIT School Grades</i>	<i>n</i> ^b	<i>M</i>	<i>SD</i>	<i>n</i> ^b	<i>M</i>	<i>SD</i>	<i>n</i> ^b	<i>M</i>	<i>SD</i>
Overall Average (Unstandardized)	42,373	91.67	8.18	40,509	91.67	8.20	1,864	91.62	7.72
Overall Average (Standardized within Course)	42,070	0.05	0.80	40,217	0.05	0.80	1,853	-0.03	0.81

Note. *n*^a = number of Soldiers with IMT data at the time data were extracted. *n*_{Restarted} = number of Soldiers who restarted at least once during IMT. %_{Restarted} = percentage of Soldiers who restarted at least once during IMT [$(n_{\text{Restarted}} / n) \times 100$]. Standardization excludes MOS courses with insufficient sample size ($n < 15$). *n*^b = number of Soldiers with AIT school grade data at time data were extracted.

Criterion Composites

A number of the criterion scales measure similar underlying constructs. Composites of these criterion scales were developed to reduce the number of criteria used to validate the TAPAS and simplify the interpretation of results, without sacrificing information. The four composites of Overall Performance, Commitment & Fit, Retention Cognitions, and Knowledge & Skill were constructed using theoretical rationale and examined using confirmatory factor analysis (CFA) (Bynum & Beatty, 2014). Can Do Performance and Will Do Performance composites were constructed using rational judgment, exploratory factor analyses (EFA), and confirmatory factor analysis (CFA), and were intended to encapsulate performance criteria associated with the ability to do the job and the motivation to consistently perform well, respectively (as referenced in Chapter 3, page 15). Table 4.14 lists the IMT and in-unit criterion composites, the scales included in each composite, and a brief description of how the composite was constructed. Descriptive statistics for the IMT and in-unit criterion composites are shown in Table 4.15. All of the composites were included in the validation analyses.

Table 4.14. IMT and In-Unit Criterion Scores

<i>IMT</i>		
Criterion Score	Scales	Description
Overall Performance	PRS: Army Adjustment PRS: Effort and Discipline PRS: MOS Qualification PRS: Physical Fitness PRS: Working with Others PRS: MOS Specific	General effort/ motivation criterion. Scales are averaged to form the composite.
Commitment & Fit	ALQ: Affective Commitment ALQ: Normative Commitment ALQ: General MOS Fit ALQ: Needs Supplies Army Fit	General commitment to and fit with the Army. Scales are averaged to form the composite.
Retention Cognitions	ALQ: Army Career Intentions ALQ: Army Re-enlistment ALQ: Attrition Cognition	General intentions of continuance in the Army. Scales are averaged to form the composite.
Knowledge & Skill	WTBD JKT MOS JKT AIT Grade	WTBD JKT and MOS JKT are averaged to form an overall knowledge/skill composite. For those that do not have an MOS JKT score, AIT grade is substituted.
Can Do Performance	WTBD JKT MOS JKT	Total score of WTBD JKT and MOS JKT
Will Do Performance	Average Army-Wide PRS ^a MOS-Specific PRS APFT Score Average ALQ ^b Training Achievement Training Failure (#) Disciplinary Incidents (#)	The seven scales are standardized and then summed to produce an overall will do performance score.

Table 4.14. (Continued)

<i>In-Unit</i>		
Criterion Score	Scales	Description
Overall Performance	PRS: Can Do PRS: Effort and Personal Discipline PRS: Working with Others PRS: Self-Management PRS: Physical Fitness and Bearing	General effort/ motivation criterion. Scales are averaged to form the composite.
Commitment & Fit	ALQ: Affective Commitment ALQ: General MOS Fit ALQ: Needs Supplies Army Fit	General commitment to and fit with the Army. Scales are averaged to form the composite.
Retention Cognitions	ALQ: Army Career Intentions ALQ: Army Re-enlistment ALQ: Attrition Cognition	General intentions of continuance in the Army. Scales are averaged to form the composite.
Knowledge & Skill	WTBD JKT MOS JKT	WTBD JKT and MOS JKT are averaged to form an overall knowledge/skill composite. For those that do not have MOS JKT, only WTBD JKT scores are used.

^a Army-Wide PRS scales included in the average: Army Adjustment, Effort and Discipline, MOS Qualification, Physical Fitness, Working with Others, and Overall Performance.

^b ALQ scales included in the average: Affective Commitment, Normative Commitment, Career Intentions, Re-enlistment Intentions, Attrition Cognition, Army Life Adjustment, and MOS Fit.

Table 4.15. Descriptive Statistics for Criterion Composites in the IMT and In-Unit Validation Samples by Education Tier

Domain/Measure	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
Tier 1 + Tier 2 (Combined)					
<i>IMT</i>					
Overall Performance	5,744	3.38	0.78	1.00	5.00
Commitment & Fit	20,439	3.99	0.58	1.13	5.00
Retention Cognitions	20,439	2.73	0.58	1.00	4.33
Knowledge & Skill ^a	16,040	0.04	0.85	-3.73	2.27
Can Do Performance ^a	21,378	-0.04	1.63	-7.46	4.53
Will Do Performance ^a	3,679	0.15	3.11	-14.91	8.46
<i>In-Unit</i>					
Overall PRS	1,250	5.22	1.18	1.00	7.00
Commitment & Fit	1,769	3.59	0.67	1.51	5.00
Retention Cognitions	1,769	2.42	0.67	1.00	4.14
Knowledge & Skill ^a	1,745	0.01	0.94	-4.00	2.25
Tier 1					
<i>IMT</i>					
Overall Performance	5,529	3.38	0.78	1.00	5.00
Commitment & Fit	19,682	3.98	0.58	1.13	5.00
Retention Cognitions	19,682	2.72	0.58	1.00	4.33
Knowledge & Skill ^a	15,442	0.04	0.85	-3.73	2.27
Can Do Performance ^a	20,564	-0.04	1.63	-7.46	4.53
Will Do Performance ^a	3,545	0.17	3.11	-14.91	8.46
<i>In-Unit</i>					
Overall Performance	1,214	5.23	1.16	1.00	7.00
Commitment & Fit	1,717	3.59	0.67	1.51	5.00
Retention Cognitions	1,717	2.42	0.67	1.00	4.14
Knowledge & Skill ^a	1,697	0.01	0.93	-4.00	2.25
Tier 2					
<i>IMT</i>					
Overall Performance	215	3.31	0.76	1.52	4.92
Commitment & Fit	757	4.06	0.58	1.69	5.00
Retention Cognitions	757	2.82	0.58	1.00	4.00
Knowledge & Skill ^a	598	0.06	0.86	-2.98	2.15
Can Do Performance ^a	814	0.03	1.64	-5.96	4.31
Will Do Performance ^a	134	-0.44	3.01	-7.66	7.32

Note. Tier 2 in-unit results are not reported because sample sizes are less than 100. Overall Performance scores for IMT Soldiers are on a 5-point scale. Overall Performance scores for in-unit Soldiers are on a 7-point scale.

^a The variables that are included in the criterion composites are reported on a standardized z-score scale (mean = 0).

Summary

Criterion data, such as attrition, training restarts, and AIT course grades, were gathered from administrative records. In addition, three types of criterion measures were adapted from previous Army research to validate the TAPAS: (a) the JKTs, (b) the PRS, and (c) the ALQ. The JKTs measure WTBD (Army-wide) and (for multiple target MOS) MOS-specific knowledge. These were combined with administrative records of AIT grades to form a Knowledge/Skill composite, intended to measure a Soldier's task-specific knowledge. The PRS are completed by training cadre (IMT) or supervisors (in-unit) and measure Army-wide constructs such as effort and leadership and (for selected IMT MOS) MOS-specific competence. The PRS were combined to form an Overall Performance composite intended to measure cadre and/or supervisor ratings of a Soldier's general performance level. The ALQ asks Soldiers to complete verifiable self-report performance items (e.g., their APFT scores) and self-report attitudinal items (e.g., adjustment to Army life). For the validation analyses, the ALQ scales were combined to form a Commitment & Fit composite and a Retention Cognitions composite. Finally, two criterion composites were created to measure general "can do" and "will do" performance.

In general, the criterion measures described in this chapter exhibited acceptable and theoretically consistent psychometric properties. Nearly all reliability estimates for the JKTs and ALQ scales are acceptable and correlations among the scales are all in theoretically consistent direction (see Appendixes B and C). The correlations between MOS-specific JKTs and AFQT scores ranged from moderate ($r = .26$) to strong ($r = .47$). In addition, MOS-specific JKTs correlated strongly with WTBD JKT scores. The exception to this was the Army-wide and MOS-specific PRS, which continued to exhibit low interrater reliability coefficients. Lower interrater reliability is not uncommon in military samples (Van Iddekinge et al., 2011) and despite the low reliability, both the current and past research has shown meaningful relationships between non-cognitive predictors and performance ratings (e.g., Knapp & Heffner, 2009; McHenry, Hough, Toquam, Hanson, & Ashworth, 1990). Because unreliability can attenuate correlations, this should be considered when interpreting results involving the PRS.

Regarding the criterion scores used to validate the TAPAS predictor composites, the intent was to use criterion scores that would best measure performance aspects predicted by each of the three predictor composite scores. As further discussed in Chapter 3, the Can-Do predictor composite was constructed to predict technical training performance, which is captured by JKT scores and AIT grades. The Will-Do predictor composite was constructed to predict motivation-based performance, which is captured by AFQT, the PRS, and the ALQ criteria. The Adaptation predictor composite was constructed to predict attrition. Attrition is captured through administrative records of attrition and self-report measures of retention cognition. Chapter 5 summarizes the validation results examining the relationships between the predictor composites and these criteria.

CHAPTER 5: EVIDENCE FOR THE PREDICTIVE VALIDITY OF THE TAPAS

Michael G. Hughes, Bethany H. Bynum, (HumRRO), and Heather M. Mullins (ARI)

This chapter evaluates the potential of the TAPAS to predict Soldiers' performance and retention through their first unit of assignment. We begin with a brief description of our analytic approach. Next, we summarize the main findings from incremental validity analyses of the (a) current TAPAS composites (Can-Do, Will-Do, Adaptation) and (b) criterion-specific TAPAS scales. Lastly, we discuss analyses that examined the implementation of TAPAS screens that were based on AFQT and TAPAS scores. Specifically, these analyses focused on (a) IMT and attrition outcomes based on Soldiers' AFQT categories and (b) in-unit outcomes based on Soldiers' percentile scores on the TAPAS.

Analytic Approach

To evaluate the TAPAS' potential to enhance new Soldier selection, we examined the incremental validity of the TAPAS over the AFQT in predicting first-term outcomes important to the Army. Consistent with the Army's personnel goals, we examined performance and retention-related outcomes that, as a group, provide representative coverage of the criterion space (Campbell, Hanson, & Oppler, 2001; Campbell, McHenry, & Wise, 1990; Knapp & Tremble, 2007; Strickland, 2005).

Our analysis approach was generally consistent with previous evaluations of the TAPAS and similar experimental non-cognitive predictors (e.g., Ingerick et al., 2009; Knapp & Heffner, 2009; 2010; Trippe, Caramagno, Allen, & Ingerick, 2011). In brief, this approach involved testing a series of two-step hierarchical regression models, where scores on each criterion measure or composite were regressed onto Soldiers' AFQT scores in the first step, followed by scores on either the (a) TAPAS composites or (b) TAPAS scales in the second step. In each case, we evaluated the degree to which adding the predictor(s) in the second step provided incremental validity beyond the AFQT with respect to the criterion of interest.

A series of four regression models were estimated for each criterion measure. Specifically, three of the four models each consisted of one TAPAS composite added in the second step of the model. The fourth model, however, varied depending on the specific criterion being examined. For Knowledge & Skill, Warrior Tasks and Battle Drills Job Knowledge Test (WTBD JKT), and Can Do Performance (i.e., Can Do criteria), the Can-Do TAPAS scales were included in the second step. For Army Fit, Army Life Adjustment, Commitment & Fit, Retention Cognitions, APFT Score, Will Do Performance, Disciplinary Incidents, IMT Restarts, and the Performance Rating Scales (PRS) of Effort and Discipline, Adjustment to the Army, Physical Fitness and Bearing, Working with Others, and Overall Performance (i.e., Will Do criteria), the Will-Do TAPAS scales were included in the second step. For the attrition criteria examined at 6, 12, and 24 months; the Adaptation TAPAS scales were added in the second step.

Estimates for the fourth model consisting of criterion-specific TAPAS scales represent the best-case scenario of the TAPAS predictive potential. Estimates for the TAPAS composite models capture the predictive potential of the composites as configured for operational usage. All models were estimated using a combined Tier 1 and 2 sample as well as separately by education tier where sample sizes

were greater than 100. In addition, attrition criteria were examined separately for Regular Army and Guard/Reserve Soldiers. Table 5.1 provides a summary of each of the regression models.

Table 5.1. Summary of Regression Models Evaluated for each Criterion

Model	Step 2 Predictors	# of Predictors added in Step 2	Description
1	TAPAS Can-Do Composite	1	The TAPAS Can-Do composite is a single variable based on scores from multiple TAPAS scales, and it was added to the model in Step 2.
2	TAPAS Will-Do Composite	1	The TAPAS Will-Do composite is a single variable based on scores from multiple TAPAS scales, and it was added to the model in Step 2.
3	TAPAS Adaptation Composite	1	The TAPAS Adaptation composite is a single variable based on scores from multiple TAPAS scales, and it was added to the model in Step 2.
4	TAPAS Facet Scales		
	Can Do TAPAS Facets	6	For models predicting Can Do criteria, the TAPAS scales that comprise the TAPAS Can-Do composite were added in Step 2.
	Will Do TAPAS Facets	4	For models predicting Will Do criteria, the TAPAS scales that comprise the TAPAS Will-Do composite were added in Step 2.
	Adaptation TAPAS Facets	4	For models predicting Attrition, the TAPAS scales that comprise the TAPAS Adaptation composite were added in Step 2.

Note. DV: Dependent Variable. All regression models included the AFQT as the only predictor in Step 1. The TAPAS Can-Do, Will-Do, and Adaptation composites each represent single variables comprised of multiple TAPAS scales. Models 1 through 3 were conducted for every criterion variable. For Model 4 (TAPAS Scales), the predictors added in Step 2 varied depending on the type of criteria. For security reasons, the specific TAPAS scales that form each composite are not provided here.

In the present report, models predicting continuously scaled criteria were estimated using Ordinary Least Squares (OLS) regression. Logistic regression was used for the dichotomous criteria (i.e., attrition, Disciplinary Incidents, IMT Restarts).¹³ Note that because different regression methods are required for different types of criteria (i.e., continuous vs. dichotomous), the statistical indices used to evaluate the OLS and logistic models also are different. Additional details concerning the specific indices presented for the logistic regression analyses are provided in the section on dichotomous outcomes later in the chapter.

For continuously scaled criteria examined using OLS regression, we computed cross-validity estimates to adjust the observed R and ΔR for shrinkage. These estimates enable comparisons of results across models from different samples of education tier levels. Specifically, we adjusted the observed R estimates associated with each step in the models using Burket's (1964) formula for population cross-validity (cf. Schmitt & Ployhart, 1999):

$$\rho_c = (NR^2 - k)/[R(N - k)] \quad (1)$$

¹³ The dichotomous version of Disciplinary Incidents (0 = no disciplinary incidents; 1 = one or more disciplinary incidents) was used for all analyses due to a low base rate beyond one incident.

where ρ_c equals the estimated population cross-validity (i.e., shrinkage-adjusted R), R equals the observed multiple correlation, k equals the number of predictors in the model, and N equals the sample size.

Next, we computed the difference between the adjusted R estimates by subtracting the adjusted R associated with the AFQT-only model from the adjusted R obtained from the full model (e.g., the AFQT + TAPAS composite model). Note that these adjustments were made only for continuously scaled criteria because true R values are not available from logistic regression analyses of categorical outcomes.

In addition to the incremental validity analyses, we examined the predictive validity of the individual TAPAS scales based on the bivariate correlations between scores on the TAPAS scales and the selected criterion measures. The results of these analyses are presented in Appendix D.

Findings

Results of these analyses are organized by criterion domain: (a) IMT performance, (b) in-unit performance, and (c) dichotomous outcomes. A few notes related to interpretation of the findings are in order:

- The results for Tier 2 Soldiers should be interpreted with caution at this stage of the TOPS evaluation because of limited criterion data for those Soldiers. Accordingly, our discussion primarily focuses on the analyses of the combined Tier 1 and 2 sample of Soldiers.¹⁴
- Results of OLS regression analyses are discussed with respect to the shrinkage-adjusted R values for all continuous criteria.
- Results of logistic regression analyses are discussed with respect to odds ratios (ORs) in combination with likelihood ratio χ^2 tests of change in model fit (i.e., deviance).
- Much of our discussion focuses on the TAPAS composite models because these models best evaluate the TAPAS' current operational format as well as its potential future format. Similarly, tables of results included in this chapter include models of the TAPAS composites only. Results of the criteria-specific TAPAS scales models are graphically displayed and briefly discussed.

¹⁴ Due to the large proportion of Tier 1 Soldiers relative to Tier 2 Soldiers, results for the combined Tier 1 and Tier 2 sample were generally comparable to Tier 1 only results.

Predicting IMT Performance

Tables 5.2 to 5.4 summarize the incremental validity results of the TAPAS composites for predicting IMT performance criteria over and above the AFQT. Overall, the results suggest that both the Will-Do and Adaptation TAPAS composites can enhance the Army's ability to predict a number of important outcomes. Below, we describe the specific outcomes for which the TAPAS composites demonstrated notable predictive gains beyond the AFQT alone.

With respect to the motivation-based performance criteria, the TAPAS composites exhibited gains in predictive validity over the AFQT in predicting several outcomes. In the combined Tier 1 and 2 sample, the TAPAS Will-Do composite enhanced the prediction of the Will Do Performance composite (Adj. $\Delta R = .25$), APFT Score (Adj. $\Delta R = .17$), Commitment & Fit (Adj. $\Delta R = .12$), and Army Life Adjustment (Adj. $\Delta R = .14$); and to a lesser extent, the prediction of Army Fit (Adj. $\Delta R = .09$) and the PRS criteria of Physical Fitness and Bearing (Adj. $\Delta R = .09$), Adjustment to the Army (Adj. $\Delta R = .06$), and Overall Performance (Adj. $\Delta R = .05$).

The TAPAS Adaptation composite also demonstrated incremental validity in the prediction of the Will Do Performance composite (Adj. $\Delta R = .15$), APFT Score (Adj. $\Delta R = .09$), Army Life Adjustment (Adj. $\Delta R = .05$), and PRS: Physical Fitness and Bearing (Adj. $\Delta R = .05$) in the combined sample. The Can-Do composite did not demonstrate any notable incremental validity in either the combined or Tier 1 samples for any of the motivation-based outcomes. However, it is not designed to predict Will Do criteria.

Consistent with expectations and previous analyses, the TAPAS composites (Can-Do, Will-Do, and Adaptation) evidenced no notable increments over the AFQT in predicting scores on the composite measure of Knowledge & Skill (Adj. $\Delta R < .01$). Similarly, the TAPAS composites did not show incremental validity in the prediction of the other Can Do criteria, namely WTBD JKT and the Can Do Performance composite (Adj. ΔR 's $< .01$). These results are not surprising given that AFQT is an established strong predictor of Can Do outcomes.

Table 5.2. Incremental Validity Estimates for the TAPAS over AFQT for Predicting IMT Technical by Education Tier

IMT Criterion Measure / Model	Tier 1 + 2 (Combined)				Tier 1				Tier 2			
	AFQT R	AFQT + TAPAS R	ΔR	Adj ΔR	AFQT R	AFQT + TAPAS R	ΔR	Adj ΔR	AFQT R	AFQT + TAPAS R	ΔR	Adj ΔR
<i>Knowledge & Skill</i>	<i>n = 11,524 - 16,040</i>				<i>n = 11,106 - 15,442</i>				<i>n = 418 - 598</i>			
Can-Do ^a	.45	.46	.00	.00	.46	.46	.00	.00	.35	.35	.00	.00
Will-Do	.45	.45	.00	.00	.45	.45	.00	.00	.38	.38	.00	.00
Adaptation ^a	.45	.45	.00	.00	.46	.46	.00	.00	.35	.35	.00	.00
<i>WTBD JKT</i>	<i>n = 13,821 - 19,556</i>				<i>n = 13,313 - 18,824</i>				<i>n = 508 - 732</i>			
Can-Do ^b	.43	.43	.00	.00	.43	.43	.00	.00	.31	.31	.00	.00
Will-Do	.43	.43	.00	.00	.43	.43	.00	.00	.35	.35	.00	.00
Adaptation ^a	.43	.43	.00	.00	.43	.43	.00	.00	.31	.31	.00	.00
<i>Can Do Performance</i>	<i>n = 13,923 - 19,720</i>				<i>n = 13,410 - 18,982</i>				<i>n = 513 - 738</i>			
Can-Do ^a	.45	.45	.00	.00	.45	.45	.00	.00	.33	.33	.00	.00
Will-Do	.44	.44	.00	.00	.44	.45	.00	.00	.37	.37	.00	.00
Adaptation ^a	.45	.45	.00	.00	.45	.45	.00	.00	.33	.33	.00	.00

Note. WTBD JKT = Warrior Tasks and Battle Drills Job Knowledge Test. *R* = multiple correlations between the AFQT and selected TAPAS composite scales with the targeted criterion measure. ΔR = Increment in *R* from adding the selected TAPAS composite scale to the regression model [(AFQT + TAPAS) – AFQT Only]. Adj ΔR = Increment in estimated population cross-validity from adding the selected TAPAS composite scales over AFQT to the regression model. Bolded values indicate $p < .05$.

^a Because the Adaptation and Can-Do composites are based on a subset of the data, these results may underestimate their incremental validity.

Table 5.3. Incremental Validity Estimates for the TAPAS over AFQT for Predicting IMT Adjustment, Commitment & Fit, Fitness, and Retention Criteria by Education Tier

IMT Criterion Measure / Model	Tier 1 + 2 (Combined)				Tier 1				Tier 2			
	AFQT R	AFQT + TAPAS R	ΔR	Adj ΔR	AFQT R	AFQT + TAPAS R	ΔR	Adj ΔR	AFQT R	AFQT + TAPAS R	ΔR	Adj ΔR
<i>Army Fit</i>	<i>n = 14,421 - 20,439</i>				<i>n = 13,894 - 19,682</i>				<i>n = 527 - 757</i>			
Can-Do ^a	.04	.05	.01	.01	.04	.05	.01	.01	.04	.06	.02	.00
Will-Do	.04	.13	.09	.09	.04	.13	.09	.09	.02	.14	.13	.13
Adaptation ^a	.04	.05	.01	.01	.04	.05	.01	.01	.04	.04	.01	.00
<i>Army Life Adjustment</i>	<i>n = 14,421 - 20,439</i>				<i>n = 13,894 - 19,682</i>				<i>n = 527 - 757</i>			
Can-Do ^a	.07	.09	.03	.03	.07	.09	.03	.03	.00	.04	.04	.00
Will-Do	.06	.21	.14	.14	.07	.21	.14	.14	.01	.20	.19	.19
Adaptation ^a	.07	.11	.05	.05	.07	.11	.04	.04	.00	.09	.09	.05
<i>Commitment & Fit</i>	<i>n = 14,421 - 20,439</i>				<i>n = 13,894 - 19,682</i>				<i>n = 527 - 757</i>			
Can-Do ^a	.01	.03	.02	.02	.01	.03	.02	.02	.01	.03	.01	.00
Will-Do	.00	.12	.11	.12	.00	.12	.11	.12	.00	.14	.14	.12
Adaptation ^a	.01	.04	.03	.03	.01	.04	.03	.04	.01	.03	.02	.00
<i>Retention Cognitions</i>	<i>n = 14,421 - 20,439</i>				<i>n = 13,894 - 19,682</i>				<i>n = 527 - 757</i>			
Can-Do ^a	.12	.13	.01	.01	.12	.13	.01	.01	.13	.13	.00	-.01
Will-Do	.12	.13	.00	.00	.12	.13	.00	.00	.12	.13	.00	.00
Adaptation ^a	.12	.12	.00	.00	.12	.12	.00	.00	.13	.13	.00	-.01
<i>APFT Score</i>	<i>n = 14,285 - 20,245</i>				<i>n = 13,765 - 19,498</i>				<i>n = 520 - 747</i>			
Can-Do ^a	.09	.09	.01	.01	.09	.09	.01	.01	.06	.06	.00	.00
Will-Do	.09	.26	.17	.17	.09	.26	.16	.16	.08	.29	.21	.22
Adaptation ^a	.09	.18	.09	.09	.09	.18	.09	.09	.06	.14	.08	.08
<i>Will Do Performance</i>	<i>n = 2,584 - 3,337</i>				<i>n = 2,501 - 3,218</i>				<i>n = 83 - 119</i>			
Can-Do ^a	.03	.03	.01	.00	.03	.04	.01	.00	.06	.08	.02	.00
Will-Do	.05	.29	.25	.25	.05	.30	.24	.25	.05	.22	.17	.14
Adaptation ^a	.03	.16	.14	.15	.03	.16	.13	.14	.06	.11	.05	.00

Note. R = multiple correlations between the AFQT and selected TAPAS composite scales with the targeted criterion measure. ΔR = Increment in R from adding the selected TAPAS composite scale to the regression model [(AFQT + TAPAS) – AFQT Only]. Adj ΔR = Increment in estimated population cross-validity from adding the selected TAPAS composite scales over AFQT to the regression model. Bolded values indicate $p < .05$.

^a Because the Adaptation and Can-Do composites are based on a subset of the data, these results may underestimate their incremental validity.

Table 5.4. Incremental Validity Estimates for the TAPAS over AFQT for Predicting IMT Performance Rating Criteria by Education Tier

IMT Criterion Measure / Model	Tier 1 + 2 (Combined)				Tier 1				Tier 2			
	AFQT R	AFQT + TAPAS R	ΔR	Adj ΔR	AFQT R	AFQT + TAPAS R	ΔR	Adj ΔR	AFQT R	AFQT + TAPAS R	ΔR	Adj ΔR
<i>PRS: Effort and Discipline</i>	<i>n = 4,599 - 6,193</i>				<i>n = 4,447 - 5,967</i>				<i>n = 152 - 226</i>			
Can-Do ^a	.08	.08	.00	.00	.08	.08	.00	.00	.02	.12	.10	.01
Will-Do	.08	.12	.03	.03	.09	.12	.03	.03	.00	.12	.11	.04
Adaptation ^a	.08	.09	.01	.01	.08	.09	.01	.01	.02	.11	.09	.00
<i>PRS: Adjustment to the Army</i>	<i>n = 4,594 - 6,190</i>				<i>n = 4,442 - 5,964</i>				<i>n = 152 - 226</i>			
Can-Do ^a	.03	.03	.00	.00	.03	.03	.00	.00	.05	.06	.01	.00
Will-Do	.04	.10	.06	.06	.04	.10	.06	.06	.05	.09	.04	.00
Adaptation ^a	.03	.06	.03	.03	.03	.06	.03	.03	.05	.09	.04	.00
<i>PRS: Physical Fitness and Bearing</i>	<i>n = 4,584 - 6,178</i>				<i>n = 4,432 - 5,952</i>				<i>n = 152 - 226</i>			
Can-Do ^a	.06	.06	.01	.00	.06	.06	.01	.00	.00	.09	.09	.00
Will-Do	.06	.15	.09	.09	.06	.15	.09	.09	.03	.16	.14	.11
Adaptation ^a	.06	.11	.05	.05	.06	.11	.05	.05	.00	.15	.15	.07
<i>PRS: Working with Others</i>	<i>n = 4,586 - 6,182</i>				<i>n = 4,434 - 5,956</i>				<i>n = 152 - 226</i>			
Can-Do ^a	.06	.06	.00	.00	.06	.06	.00	.00	.01	.10	.09	.00
Will-Do	.07	.10	.03	.03	.07	.10	.03	.03	.01	.10	.09	.01
Adaptation ^a	.06	.08	.02	.01	.06	.08	.02	.01	.01	.07	.06	.00
<i>PRS: Overall Performance</i>	<i>n = 4,325 - 5,744</i>				<i>n = 4,179 - 5,529</i>				<i>n = 146 - 215</i>			
Can-Do ^a	.08	.08	.00	.00	.08	.08	.00	.00	.03	.08	.06	.00
Will-Do	.08	.13	.05	.05	.08	.13	.05	.05	.05	.12	.08	.05
Adaptation ^a	.08	.10	.02	.02	.08	.10	.02	.02	.03	.06	.03	.00

Note. PRS = Performance Rating Scales. R = multiple correlations between the AFQT and selected TAPAS composite scales with the targeted criterion measure. ΔR = Increment in R from adding the selected TAPAS composite scale to the regression model [(AFQT + TAPAS) – AFQT Only]. Adj ΔR = Increment in estimated population cross-validity from adding the selected TAPAS composite scales over AFQT to the regression model. Bolded values indicate $p < .05$.

^a Because the Adaptation and Can-Do composites are based on a subset of the data, these results may underestimate their incremental validity.

Figures 5.1 and 5.2 provide results of the combined Tier 1 and 2 analyses examining the incremental validity of criteria-specific TAPAS scales over and above the AFQT for Will Do and Can Do criteria, respectively. Similar to the results of the TAPAS composite models, the Will-Do scales provided the largest gains over the AFQT. In particular, increases in R were largest for the Will Do Performance composite, APFT Score, Army Life Adjustment, Commitment & Fit, Disciplinary Incidents, Army Fit, and PRS: Working with Others criteria ($\Delta R \geq .10$). Results of the Can-Do TAPAS scales showed no notable increments over the AFQT in predicting the Can Do criteria.

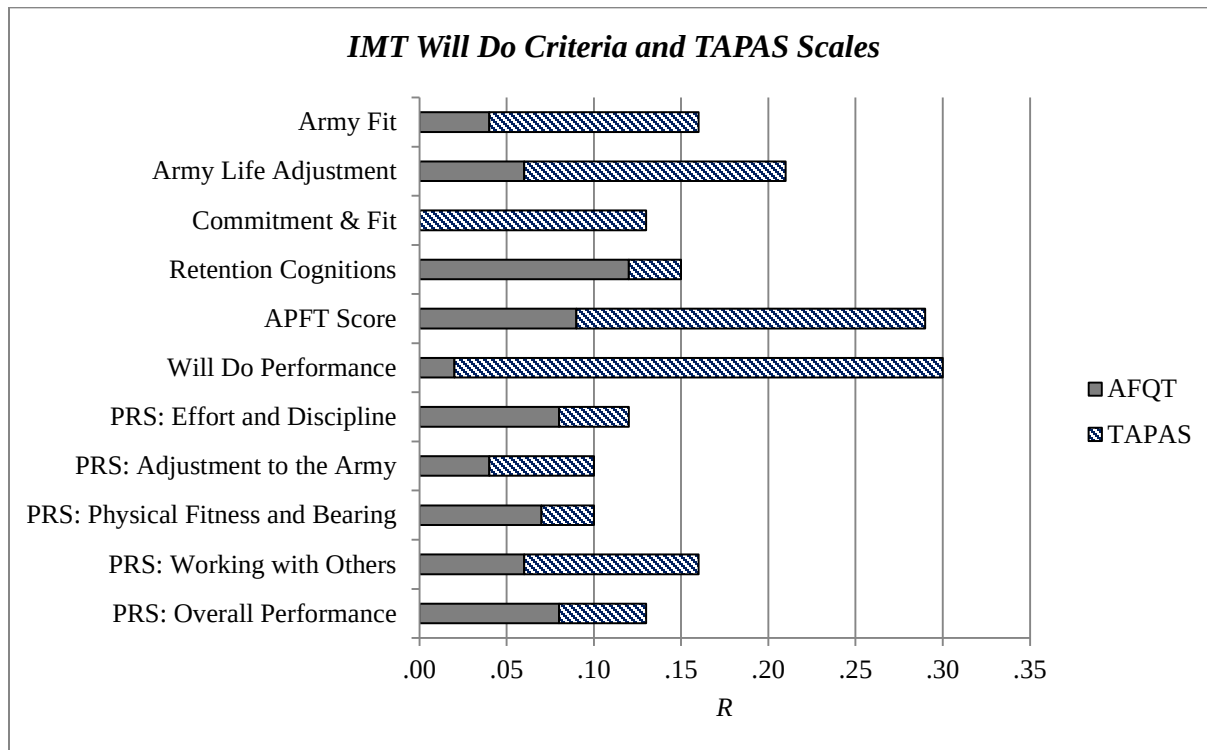


Figure 5.1. Increase in prediction of IMT criteria using Will-Do TAPAS scales for the combined Tier 1 and 2 sample.

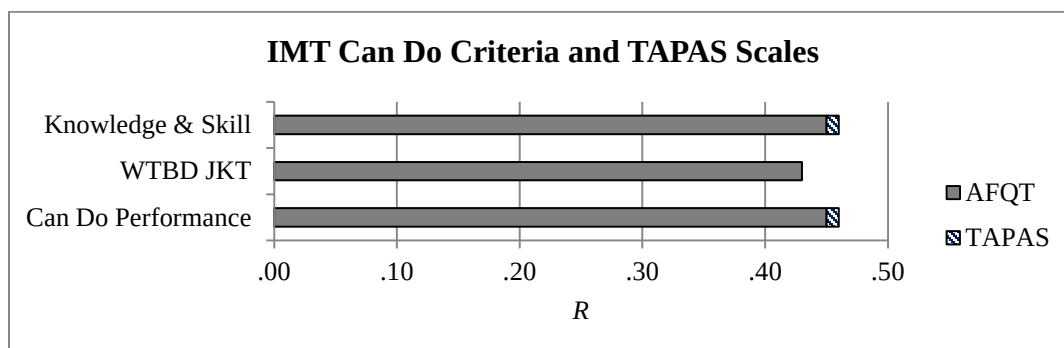


Figure 5.2. Increase in prediction of IMT criteria using Can-Do TAPAS scales for the combined Tier 1 and 2 sample.

Predicting In-Unit Performance

The incremental validity results for predicting in-unit performance are presented in Tables 5.5 to 5.7. Similar to the results for the IMT performance criteria, these results also suggest that the TAPAS composites are useful predictors of in-unit performance outcomes. For multiple outcomes, the Will-Do and Adaptation TAPAS composites both exhibited enhanced prediction beyond the AFQT alone. Results for which notable predictive gains were observed are detailed below. Note that separate analyses were not conducted for Tier 2 Soldiers due to limited in-unit criterion data available for those Soldiers (n 's < 100).

With respect to the combined Tier 1 and 2 sample results, none of the TAPAS composite predictors demonstrated incremental validity beyond the AFQT in the prediction of Knowledge & Skill or WTBD JKT scores ($\text{Adj. } \Delta R^2 \leq .01$). However, for motivation-based criteria, the Will-Do composite showed increases beyond the AFQT in predicting the APFT Score ($\text{Adj. } \Delta R = .23$), PRS: Physical Fitness and Bearing ($\text{Adj. } \Delta R = .13$), PRS: Overall Performance ($\text{Adj. } \Delta R = .07$), and PRS: Leadership Potential ($\text{Adj. } \Delta R = .06$).

The TAPAS Adaptation composite also demonstrated incremental validity for predicting the APFT Score ($\text{Adj. } \Delta R = .14$) and PRS: Physical Fitness and Bearing ($\text{Adj. } \Delta R = .07$) in the combined sample. Similar to the analyses of the IMT criteria, the Can-Do composite did not provide incremental validity in the prediction of any in-unit criteria. However, this result is expected given that the Can-Do composite is not intended to be related to Will Do outcomes, and the AFQT is an established strong predictor of Can Do outcomes.

Table 5.5. Incremental Validity Estimates for the TAPAS over AFQT for Predicting In-Unit Technical Criteria by Education Tier

In-Unit Criterion Measure / Model	Tier 1 + 2 (Combined)				Tier 1			
	AFQT <i>R</i>	AFQT + TAPAS <i>R</i>	ΔR	Adj ΔR	AFQT <i>R</i>	AFQT + TAPAS <i>R</i>	ΔR	Adj ΔR
<i>Knowledge & Skill</i>	<i>n</i> = 1,497 - 1,745				<i>n</i> = 1,461 - 1,697			
Can-Do ^a	.45	.46	.01	.01	.45	.46	.01	.01
Will-Do	.46	.46	.00	.00	.46	.46	.00	.00
Adaptation ^a	.45	.45	.00	.00	.45	.45	.00	.00
<i>WTBD JKT</i>	<i>n</i> = 1,494 - 1,741				<i>n</i> = 1,458 - 1,693			
Can-Do ^a	.44	.45	.01	.01	.44	.45	.01	.01
Will-Do	.45	.45	.00	.00	.45	.45	.00	.00
Adaptation ^a	.44	.45	.00	.00	.44	.45	.00	.00

Note. WTBD JKT = Warrior Tasks and Battle Drills Job Knowledge Test. *R* = multiple correlations between the AFQT and selected TAPAS composite scales with the targeted criterion measure. ΔR = Increment in *R* from adding the selected TAPAS composite scale to the regression model [(AFQT + TAPAS) – AFQT Only]. Adj ΔR = Increment in estimated population cross-validity from adding the selected TAPAS composite scales over AFQT to the regression model. Tier 2 results are not reported because sample sizes are less than 100. Bolded values indicate $p < .05$.

^a Because the Adaptation and Can-Do composites are based on a subset of the data, these results may underestimate their incremental validity.

Table 5.6. Incremental Validity Estimates for the TAPAS over AFQT for Predicting In-Unit Adjustment, Commitment & Fit, Fitness, and Retention Criteria by Education Tier

In-Unit Criterion Measure / Model	Tier 1 + 2 (Combined)				Tier 1			
	AFQT R	AFQT + TAPAS R	ΔR	Adj ΔR	AFQT R	AFQT + TAPAS R	ΔR	Adj ΔR
<i>Army Fit</i>	<i>n = 1,515 - 1,769</i>				<i>n = 1,476 - 1,717</i>			
Can-Do ^a	.05	.07	.02	.01	.05	.07	.02	.01
Will-Do	.04	.09	.05	.05	.04	.09	.05	.05
Adaptation ^a	.05	.06	.00	.00	.05	.06	.00	.00
<i>MOS Fit</i>	<i>n = 1,515 - 1,769</i>				<i>n = 1,476 - 1,717</i>			
Can-Do ^a	.00	.01	.01	.00	.01	.01	.00	.00
Will-Do	.00	.05	.04	.02	.01	.04	.04	.02
Adaptation ^a	.00	.04	.04	.00	.01	.03	.02	.00
<i>Commitment & Fit</i>	<i>n = 1,515 - 1,769</i>				<i>n = 1,476 - 1,717</i>			
Can-Do ^a	.06	.06	.01	.00	.05	.06	.01	.00
Will-Do	.04	.08	.04	.04	.04	.08	.04	.04
Adaptation ^a	.06	.06	.00	.00	.05	.06	.00	.00
<i>Retention Cognitions</i>	<i>n = 1,515 - 1,769</i>				<i>n = 1,476 - 1,717</i>			
Can-Do ^a	.12	.12	.00	.00	.12	.12	.00	.00
Will-Do	.11	.12	.00	.00	.11	.12	.00	.00
Adaptation ^a	.12	.12	.00	.00	.12	.12	.00	.00
<i>APFT Score</i>	<i>n = 1,486 - 1,737</i>				<i>n = 1,447 - 1,685</i>			
Can-Do ^a	.04	.05	.00	.00	.04	.05	.00	.00
Will-Do	.03	.24	.21	.23	.04	.24	.20	.21
Adaptation ^a	.04	.17	.13	.14	.04	.16	.12	.13

Note. R = multiple correlations between the AFQT and selected TAPAS composite scales with the targeted criterion measure. ΔR = Increment in R from adding the selected TAPAS composite scale to the regression model [(AFQT + TAPAS) – AFQT Only]. Adj ΔR = Increment in estimated population cross-validity from adding the selected TAPAS composite scales over AFQT to the regression model. Tier 2 results are not reported because sample sizes are less than 100. Bolded values indicate $p < .05$.

^a Because the Adaptation and Can-Do composites are based on a subset of the data, these results may underestimate their incremental validity.

Table 5.7. Incremental Validity Estimates for the TAPAS over AFQT for Predicting In-Unit Performance Rating Criteria by Education Tier

In-Unit Criterion Measure / Model	Tier 1 + 2 (Combined)				Tier 1			
	AFQT <i>R</i>	AFQT + TAPAS <i>R</i>	ΔR	Adj ΔR	AFQT <i>R</i>	AFQT + TAPAS <i>R</i>	ΔR	Adj ΔR
<i>PRS: Effort and Discipline</i>	<i>n</i> = 1,086 - 1,252				<i>n</i> = 1,059 - 1,216			
Can-Do ^a	.10	.11	.01	.01	.09	.10	.01	.00
Will-Do	.10	.14	.04	.04	.09	.13	.04	.03
Adaptation ^a	.10	.10	.00	.00	.09	.10	.00	.00
<i>PRS: Working with Others</i>	<i>n</i> = 1,086 - 1,252				<i>n</i> = 1,059 - 1,216			
Can-Do ^a	.13	.14	.01	.00	.13	.14	.01	.00
Will-Do	.14	.18	.04	.04	.14	.17	.03	.03
Adaptation ^a	.13	.15	.02	.01	.13	.14	.01	.00
<i>PRS: Physical Fitness and Bearing</i>	<i>n</i> = 1,081 - 1,246				<i>n</i> = 1,054 - 1,210			
Can-Do ^a	.01	.03	.01	.00	.01	.02	.01	.00
Will-Do	.02	.14	.12	.13	.02	.13	.12	.12
Adaptation ^a	.01	.09	.08	.07	.01	.07	.06	.05
<i>PRS: Leadership Potential</i>	<i>n</i> = 1,059 - 1,217				<i>n</i> = 1,032 - 1,181			
Can-Do ^a	.08	.09	.00	.00	.08	.08	.00	-.01
Will-Do	.09	.15	.06	.06	.08	.14	.06	.06
Adaptation ^a	.08	.11	.03	.02	.08	.10	.02	.02
<i>PRS: Overall Performance</i>	<i>n</i> = 1,085 - 1,250				<i>n</i> = 1,058 - 1,214			
Can-Do ^a	.09	.10	.01	.01	.09	.10	.01	.00
Will-Do	.09	.16	.07	.07	.09	.15	.06	.06
Adaptation ^a	.09	.11	.02	.01	.09	.10	.01	.00

Note. PRS = Performance Rating Scales. *R* = multiple correlations between the AFQT and selected TAPAS composite scales with the targeted criterion measure. ΔR = Increment in *R* from adding the selected TAPAS composite scale to the regression model [(AFQT + TAPAS) – AFQT Only]. Adj ΔR = Increment in estimated population cross-validity from adding the selected TAPAS composite scales over AFQT to the regression model. Tier 2 results are not reported because sample sizes are less than 100. Bolded values indicate $p < .05$.

^a Because the Adaptation and Can-Do composites are based on a subset of the data, these results may underestimate their incremental validity.

The results of the combined Tier 1 and 2 incremental validity analyses for criteria-specific TAPAS scales are shown in Figures 5.3 and 5.4 for Will Do and Can Do criteria, respectively. Similar to the results of the TAPAS composite models predicting both IMT and in-unit criteria, the Will-Do scales provided the largest gains over the AFQT. Increases in R were largest for the APFT Score, PRS: Physical Fitness and Bearing, and Disciplinary Incidents criteria ($\Delta R_s \geq .10$). Results of the model including the Can-Do TAPAS scales exhibited negligible increments beyond AFQT in predicting either Knowledge & Skill or the WTBD JKT.

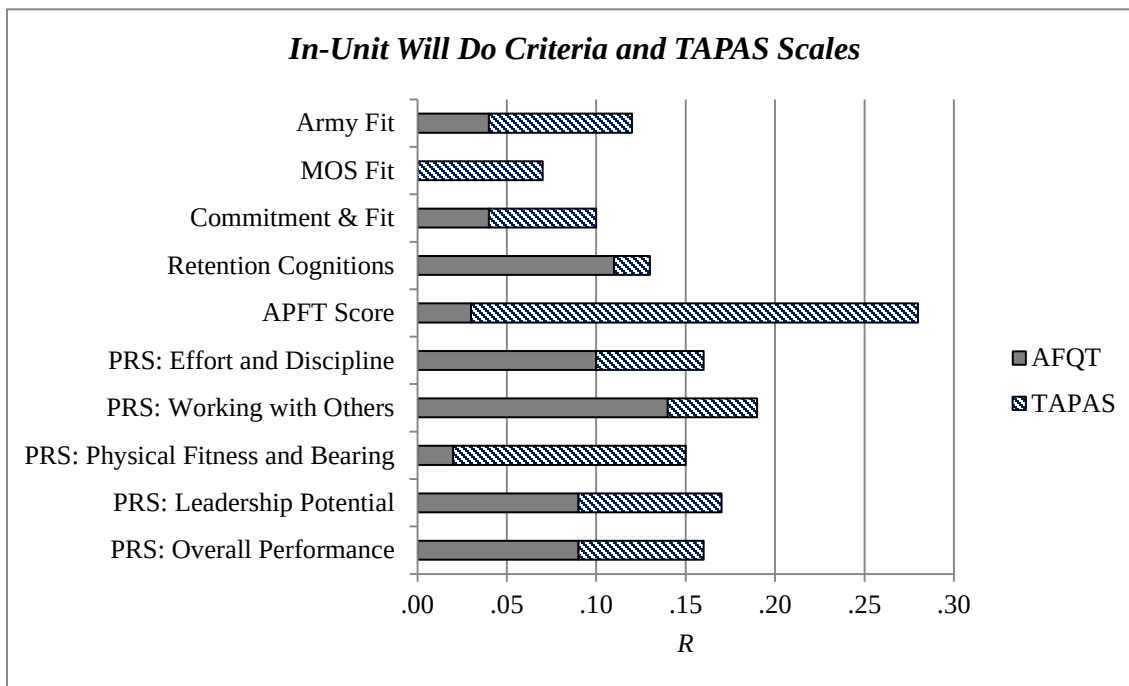


Figure 5.3. Increase in prediction of in-unit criteria using Will-Do TAPAS scales for the combined Tier 1 and 2 sample.

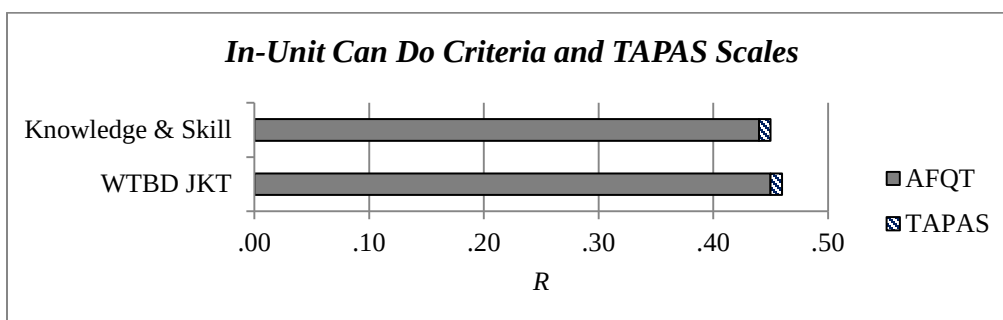


Figure 5.4. Increase in prediction of in-unit criteria using Can-Do TAPAS scales for the combined Tier 1 and 2 sample.

Predicting Dichotomous Outcomes

In addition to the OLS regression analyses of IMT and in-unit criteria, we conducted logistic regression analyses of the dichotomous outcomes, including Disciplinary Incidents (measured at both IMT and In-Unit), IMT Restarts, and attrition at 6, 12, and 24 months. For these models, we estimated odds ratios (ORs) for the predictors as well as the corresponding confidence intervals

(CIs). Additionally, we computed point biserial correlations (r_{pb}) and conducted χ^2 tests of the change in model deviance (i.e., negative two log likelihood; -2LL) from the AFQT-only to the AFQT + TAPAS composite models.

Odds ratios can be used to assess the likelihood (or odds) of a given outcome depending on change in a predictor. Specifically, for a given logistic regression model, a unique OR is estimated for each predictor, and represents the amount of change in the odds of the outcome that is associated with change in the given predictor. For the present analyses, the ORs represent the amount of change in the likelihood of each outcome that can be attributed to every 1.0 change in the predictor score. Note that ORs equal to 1.0 reflect no relationship between a given predictor and outcome, ORs greater than 1.0 reflect positive relationships, and ORs between 0.0 and less than 1.0 reflect negative relationships (i.e., decreasing odds of the outcome with increasing values of the predictor). For ORs below 1.0, values closer to 0.0 indicate stronger negative relationships. Although values of ORs cannot fall below 0.0, there is no upper limit for ORs (Cohen, Cohen, West, & Aiken, 2003). In addition, we computed 95% CIs for the ORs, which can be interpreted as an index of statistical significance for each. That is, a CI that contains 1.0 suggests that the relationship between the associated predictor and outcome is not significant.

Point biserial correlations represent the correlation between a Soldier's predicted probability of exhibiting a selected behavior and his or her actual behavior (e.g., being involved in a disciplinary incident; Tabachnick & Fidell, 2013). As such, stronger point biserial correlations reflect stronger relationships between predicted and observed outcomes, and thus are indicative of better-fitting models. Model deviance (i.e., -2LL) also provides an index of model fit. Moreover, the difference in deviances obtained from nested logistic regression models can be tested using likelihood ratio χ^2 tests to determine the statistical significance of change in model fit between models. In the present application, statistically significant likelihood ratio χ^2 tests of the change in deviances suggest that the inclusion of a given TAPAS composite to a regression model provides significantly better prediction of the outcome than the AFQT alone.

Results of the analyses examining Disciplinary Incidents (IMT and in-unit) and IMT Restarts are provided in Table 5.8. For the combined Tier 1 and 2 sample, both the Will-Do and Adaptation TAPAS composites enhanced the prediction of IMT Disciplinary Incidents and IMT Restarts beyond the AFQT-only models. Specifically, for both Disciplinary Incidents ($OR_{Will-Do} = .986$; $OR_{Adaptation} = .990$) and IMT Restarts ($OR_{Will-Do} = .993$; $OR_{Adaptation} = .994$), ORs associated with both composites were below 1.0, indicating that as scores on these composites went up, the likelihood of the outcome went down. For in-unit Disciplinary Incidents, only the Will-Do composite had a significant relationship ($OR_{Will-Do} = .988$) and resulted in better model fit over the AFQT. The Can-Do composite did not predict either IMT or in-unit Disciplinary Incidents (as evidence by CIs that included 1.000 for the associated ORs). Furthermore, the addition of the Can-Do composite to the AFQT-only model did not lead to an improvement in fit for models predicting Disciplinary Incidents (IMT and in-unit) or IMT Restarts.

Tables 5.9 presents the results of the logistic regression analyses examining attrition through 6, 12, and 24 months of service for Regular Army. For Regular Army, the Will-Do ($.990 \leq OR_{Will-Do} \leq .991$) and Adaptation ($.990 \leq OR_{Adaptation} \leq .991$) composites were negatively related to attrition at all three time points, and their respective inclusion in the models resulted in significantly better fit over the AFQT alone. Additionally, the TAPAS Can-Do composite was not related to attrition at any of the time points for Regular Army Soldiers.

Table 5.8. Incremental Validity Estimates for the TAPAS Composites over AFQT for Predicting Dichotomous Criteria by Education Tier

Criterion Measure / Model	Tier 1 + 2 (Combined)				Tier 1				Tier 2			
	OR _{AFQT} (CI)	OR _{TAPAS} (CI)	r_{pb}	Δ -2LL	OR _{AFQT} (CI)	OR _{TAPAS} (CI)	r_{pb}	Δ -2LL	OR _{AFQT} (CI)	OR _{TAPAS} (CI)	r_{pb}	Δ -2LL
<i>IMT Disciplinary Incidents</i>	<i>n = 12,812 - 18,822</i>				<i>n = 12,330 - 18,111</i>				<i>n = 482 - 711</i>			
Can-Do ^a												
AFQT	.997 (.995-.999)		.03		.997 (.994-.999)		.03		.999 (.986-1.011)		.01	
AFQT+TAPAS	.997 (.995-1.000)	.998 (.996-1.000)	.03	2.74	.997 (.995-1.000)	.998 (.996-1.000)	.03	2.49	1.001 (.987-1.015)	.996 (.985-1.008)	.03	0.38
Will-Do												
AFQT	.997 (.996-.999)		.02		.997 (.996-.999)		.02		1.003 (.992-1.013)		.02	
AFQT + TAPAS	.999 (.997-1.000)	.986 (.984-.988)	.11	237.72	.998 (.997-1.000)	.986 (.984-.987)	.11	231.50	1.003 (.993-1.013)	.989 (.980-.999)	.09	5.17
Adaptation ^a												
AFQT	.997 (.995-.999)		.03		.997 (.994-.999)		.03		.999 (.986-1.011)		.01	
AFQT + TAPAS	.998 (.996-1.000)	.990 (.988-.992)	.08	78.09	.998 (.996-1.000)	.990 (.987-.992)	.08	76.75	1.000 (.987-1.013)	.993 (.981-1.004)	.06	1.52
<i>IMT Restarts</i>	<i>n = 121,060 - 208,822</i>				<i>n = 116,604 - 201,132</i>				<i>n = 4,456 - 7,690</i>			
Can-Do ^a												
AFQT	1.002 (1.001-1.003)		.01		1.002 (1.001-1.003)		.01		1.007 (1.000-1.015)		.03	
AFQT+TAPAS	1.002 (1.000-1.003)	1.001 (1.000-1.003)	.01	3.41	1.002 (1.000-1.003)	1.001 (1.000-1.003)	.01	2.86	1.007 (.999-1.014)	1.001 (.995-1.008)	.03	0.18
Will-Do												
AFQT	1.001 (1.000-1.002)		.00		1.001 (1.000-1.002)		.00		1.004 (.999-1.010)		.02	
AFQT + TAPAS	1.001 (1.000-1.002)	.993 (.992-.995)	.03	145.05	1.001 (1.000-1.002)	.993 (.992-.994)	.03	146.65	1.004 (.999-1.010)	.998 (.993-1.004)	.02	0.36
Adaptation ^a												
AFQT	1.002 (1.001-1.003)		.01		1.002 (1.001-1.003)		.01		1.007 (1.000-1.015)		.03	
AFQT + TAPAS	1.003 (1.002-1.004)	.994 (.992-.995)	.03	87.52	1.003 (1.002-1.004)	.994 (.992-.995)	.03	85.58	1.008 (1.000-1.015)	.996 (.989-1.002)	.03	1.67

Table 5.8. (Continued)

Criterion Measure / Model	Tier 1 + 2 (Combined)				Tier 1				Tier 2			
	OR _{AFQT} (CI)	OR _{TAPAS} (CI)	r_{pb}	Δ -2LL	OR _{AFQT} (CI)	OR _{TAPAS} (CI)	r_{pb}	Δ -2LL	OR _{AFQT} (CI)	OR _{TAPAS} (CI)	r_{pb}	Δ -2LL
<i>In-Unit Disciplinary Incidents</i>	<i>n = 1,515 - 1,769</i>				<i>n = 1,476 - 1,717</i>				<i>n = 39-52</i>			
Can-Do ^a												
AFQT	.997 (.992-1.003)		.02		.997 (.991-1.003)		.03		--		--	
AFQT+TAPAS	.999 (.993-1.006)	.996 (.990-1.002)	.04	1.96	.999 (.993-1.006)	.996 (.989-1.002)	.04	1.99	--	--	--	--
Will-Do												
AFQT	.996 (.991-1.002)		.03		.996 (.990-1.001)		.04		--		--	
AFQT + TAPAS	.996 (.991-1.002)	.988 (.982-.994)	.10	16.84	.996 (.990-1.001)	.989 (.984-.995)	.09	12.68	--	--	--	--
Adaptation ^a												
AFQT	.997 (.992-1.003)		.02		.997 (.991-1.003)		.03		--		--	
AFQT + TAPAS	.998 (.992-1.004)	.996 (.990-1.002)	.04	1.52	.998 (.992-1.003)	.997 (.991-1.003)	.03	0.76	--	--	--	--

Note. OR = odds ratio for each predictor. CI = 95% confidence interval of the odds ratio. r_{pb} = point biserial correlation between the observed outcome and predicted probability. Δ -2LL = change in negative two log likelihood (deviance) from adding the selected TAPAS composite score to the AFQT-only logistic regression model. Odds ratios equal to 1.0 (or confidence intervals of the odds ratio that include 1.0) indicate no relationship between the predictor and criterion. Odds ratios less than 1.0 indicate a negative relationship between the predictor and criterion. Odds ratios greater than 1.0 indicate a positive relationship between the predictor and criterion. For Δ -2LL, bolded values indicate significant change in model fit based on a Likelihood Ratio χ^2 test, $p < .05$. Tier 2 results are not reported for in-unit Disciplinary Incidents because sample sizes are less than 100.

^a Because the Adaptation and Can-Do composites are based on a subset of the data, these results may underestimate their incremental validity.

Table 5.9. Incremental Validity Estimates for the TAPAS Composite Scores over AFQT for Predicting Cumulative Attrition through 24 Months of Service by Education Tier (Regular Army Only)

Criterion Measure / Model	Tier 1 + 2 (Combined)				Tier 1				Tier 2			
	OR _{AFQT} (CI)	OR _{TAPAS} (CI)	r_{pb}	Δ -2LL	OR _{AFQT} (CI)	OR _{TAPAS} (CI)	r_{pb}	Δ -2LL	OR _{AFQT} (CI)	OR _{TAPAS} (CI)	r_{pb}	Δ -2LL
6 Month	<i>n</i> = 69,410 - 105,419				<i>n</i> = 67,147 - 101,921				<i>n</i> = 2,263 - 3,498			
Can-Do ^a												
AFQT	.990 (.989-.991)		.06		.989 (.988-.991)		.06		1.000 (.992-1.008)		.00	
AFQT+TAPAS	.989 (.988-.991)	1.002 (1.000-1.003)	.06	5.70	.989 (.988-.990)	1.001 (1.000-1.003)	.06	2.94	.996 (.988-1.004)	1.008 (1.001-1.014)	.05	5.53
Will-Do												
AFQT	.990 (.989-.991)		.06		.990 (.989-.991)		.06		.997 (.991-1.004)		.01	
AFQT + TAPAS	.990 (.989-.991)	.991 (.990-.992)	.08	315.46	.990 (.989-.991)	.991 (.990-.992)	.08	310.71	.997 (.991-1.003)	.995 (.990-1.000)	.04	4.04
Adaptation ^a												
AFQT	.990 (.989-.991)		.06		.989 (.988-.991)		.06		1.000 (.992-1.008)		.00	
AFQT + TAPAS	.991 (.990-.992)	.991 (.990-.993)	.08	174.00	.991 (.990-.992)	.991 (.989-.992)	.08	181.81	1.000 (.992-1.007)	1.001 (.995-1.007)	.01	0.09
12 Month	<i>n</i> = 63,559 - 81,305				<i>n</i> = 61,420 - 78,278				<i>n</i> = 2,139 - 3,027			
Can-Do ^a												
AFQT	.992 (.991-.993)		.06		.991 (.990-.992)		.06		1.001 (.994-1.008)		.01	
AFQT+TAPAS	.991 (.990-.992)	1.001 (1.000-1.003)	.06	4.70	.991 (.989-.992)	1.001 (1.000-1.002)	.06	3.04	1.000 (.992-1.007)	1.003 (.997-1.008)	.02	0.77
Will-Do												
AFQT	.992 (.991-.993)		.05		.992 (.990-.993)		.06		.999 (.993-1.005)		.01	
AFQT + TAPAS	.992 (.991-.993)	.990 (.989-.991)	.08	339.52	.992 (.991-.993)	.990 (.989-.991)	.09	337.33	.999 (.993-1.005)	.996 (.991-1.001)	.03	2.75
Adaptation ^a												
AFQT	.992 (.991-.993)		.06		.991 (.990-.992)		.06		1.001 (.994-1.008)		.01	
AFQT + TAPAS	.993 (.992-.994)	.990 (.989-.992)	.08	230.97	.993 (.992-.994)	.990 (.989-.991)	.08	234.86	1.001 (.994-1.008)	.998 (.992-1.004)	.01	0.41

Table 5.9. (Continued)

Criterion Measure / Model	Tier 1 + 2 (Combined)				Tier 1				Tier 2			
	OR _{AFQT} (CI)	OR _{TAPAS} (CI)	r_{pb}	Δ -2LL	OR _{AFQT} (CI)	OR _{TAPAS} (CI)	r_{pb}	Δ -2LL	OR _{AFQT} (CI)	OR _{TAPAS} (CI)	r_{pb}	Δ -2LL
24 Month	<i>n</i> = 43,362 - 43,732				<i>n</i> = 42,294 - 42,655				<i>n</i> = 1,068 - 1,077			
Can-Do ^a												
AFQT	.990 (.989-.992)		.07		.990 (.989-.992)		.07		.990 (.981-.998)		.07	
AFQT+TAPAS	.990 (.989-.991)	1.001 (.999-1.002)	.07	0.81	.990 (.989-.992)	1.000 (.999-1.002)	.07	0.33	.988 (.979-.997)	1.004 (.996-1.011)	.08	0.91
Will-Do												
AFQT	.990 (.989-.991)		.08		.990 (.989-.991)		.08		.989 (.981-.998)		.07	
AFQT + TAPAS	.991 (.989-.992)	.991 (.990-.992)	.10	183.87	.991 (.989-.992)	.991 (.990-.992)	.10	181.06	.989 (.981-.998)	.995 (.988-1.002)	.09	1.97
Adaptation ^a												
AFQT	.990 (.989-.992)		.07		.990 (.989-.992)		.07		.990 (.981-.998)		.07	
AFQT + TAPAS	.992 (.991-.993)	.991 (.989-.992)	.10	196.09	.992 (.991-.993)	.990 (.989-.992)	.10	198.84	.990 (.981-.998)	.999 (.992-1.007)	.07	0.05

Note. OR = odds ratio for each predictor. CI = 95% confidence interval of the odds ratio. r_{pb} = point biserial correlation between the observed outcome and predicted probability. Δ -2LL = change in negative two log likelihood (deviance) from adding the selected TAPAS composite scale to the AFQT-only logistic regression model. Odds ratios equal to 1.0 (or confidence intervals of the odds ratio that include 1.0) indicate no relationship between the predictor and criterion. Odds ratios less than 1.0 indicate a negative relationship between the predictor and criterion. Odds ratios greater than 1.0 indicate a positive relationship between the predictor and criterion. For Δ -2LL, bolded values indicate significant change in model fit based on a Likelihood Ratio χ^2 test, $p < .05$.

^a Because the Adaptation and Can-Do composites are based on a subset of the data, these results may underestimate their incremental validity.

Figures 5.5 and 5.6 display the accuracy of the logistic regression models in predicting the dichotomous outcomes (Disciplinary Incidents, IMT Restarts, and attrition) for the combined Tier 1 and 2 sample. For each outcome, results are presented for three models which include the following predictors: (1) AFQT, (2) TAPAS composite, and (3) AFQT + TAPAS composite. Specifically, the *percent accurate* values are equal to the *c* statistic, or area under the receiver operating characteristic (ROC) curve, which reflects the ability of the given model to correctly discriminate between a case and a noncase (e.g., attriter vs. stayer).¹⁵ The area under the curve (AUC) can range from .50 to 1.0, corresponding to 50% (or chance) and 100% accuracy, respectively. Note that the AUC is the probability that predicted scores are higher for true cases than noncases, and not the probability that a case is correctly classified (Cook, 2007).

Figure 5.5 displays the results of the AFQT and TAPAS Will-Do composite models in predicting Disciplinary Incidents (IMT and in-unit) and IMT Restarts. For all three criteria, the probability of discriminating between Soldiers with and without incidence of the outcome of interest is near chance (52% or less) when using the AFQT alone. However, the predictive accuracy increases for the Will-Do composite and AFQT + Will-Do composite models for all outcomes. In particular, the combined model demonstrated the most accuracy, ranging from 57.9% for IMT Disciplinary Incidents to 53.7% for IMT Restarts. In addition, the increase in predictive accuracy for the combined model compared to the Will-Do composite-only model was small, suggesting that the AFQT adds little value to the discrimination of these dichotomous outcomes.

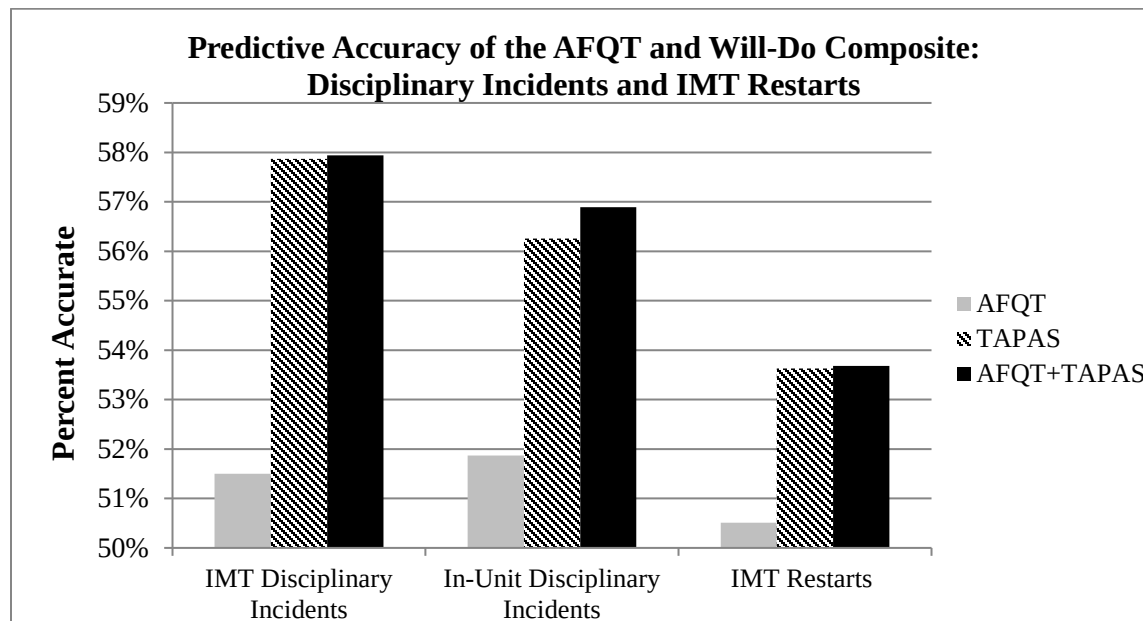


Figure 5.5. Predictive accuracy of the AFQT and Will-Do TAPAS composite in the discrimination of both IMT and In-Unit Disciplinary Incidents and IMT Restarts for the combined Tier 1 and 2 sample.

¹⁵ The ROC curve is a plot of sensitivity (i.e., probability of detecting a true positive) versus specificity (i.e., probability of detecting a true negative) across a range of potential cut scores for continuous predictors in a logistic regression model (Cook, 2007). For the purposes of evaluating the logistic regression models discussed in this report, only the area under the ROC curve (i.e., AUC) is presented and discussed.

Figure 5.6 displays the results of the AFQT and TAPAS Adaptation composite models in predicting attrition for Regular Army Soldiers. For 6-, 12-, and 24-month attrition, the combined AFQT + Adaptation composite model resulted in the most accurate predictions (approximately 57% for each time point). However, when comparing the AFQT-only and Adaptation composite-only models, the more accurate of the two individual predictors varied by time point. Moreover, for each time point, the predictive accuracy of the combined model improved by at least 1% over either predictor alone, suggesting that both the AFQT and Adaptation composite contribute to the prediction of attrition.

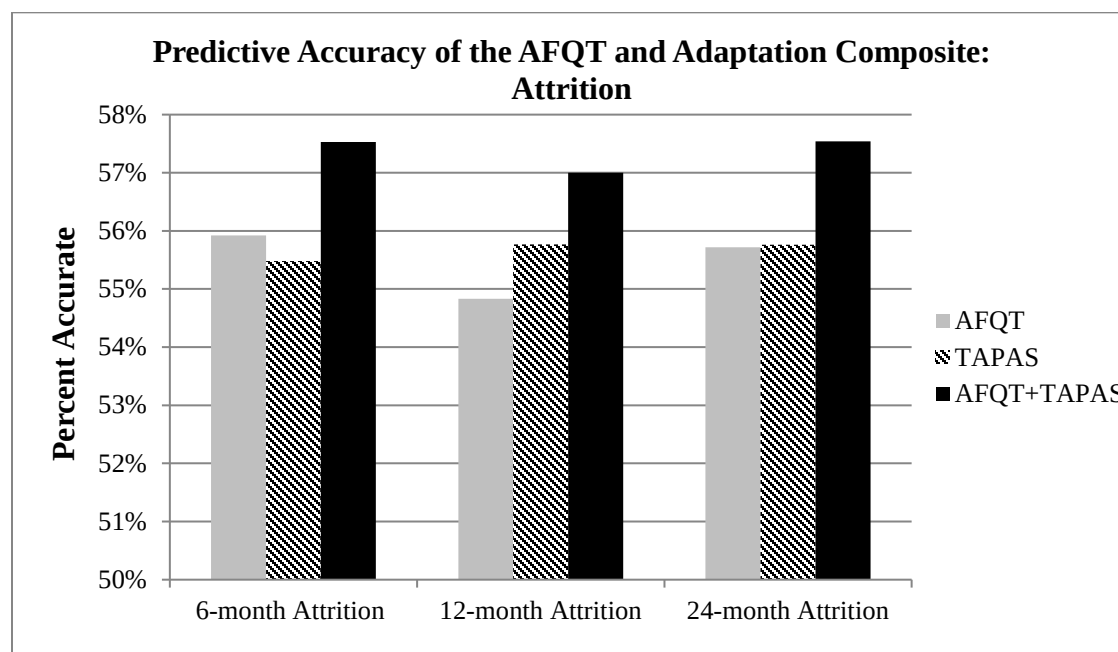


Figure 5.6. Predictive accuracy of the AFQT and Adaptation TAPAS composite in the discrimination of attrition outcomes for Regular Army Soldiers in the combined Tier 1 and 2 sample.

Implementation

To further examine the relationships between the TAPAS and the various IMT, in-unit, and attrition criteria, we conducted analyses that assessed implementation of the TAPAS with respect to AFQT categories and TAPAS score percentiles. Soldiers are classified into AFQT categories based on their performance on the AFQT (described in Chapter 2). In the IOT&E, AFQT Category IIIB/IV Soldiers are further categorized as either IIIB/IV TAPAS Pass or IIIB/IV TAPAS Fail based on their TAPAS Will-Do and TAPAS Adaptation composite scores. To be classified as IIIB/IV TAPAS Pass, Soldiers must score at or above the 10th percentile on both composites.¹⁶ The following analyses examine scores on the key outcome for each AFQT

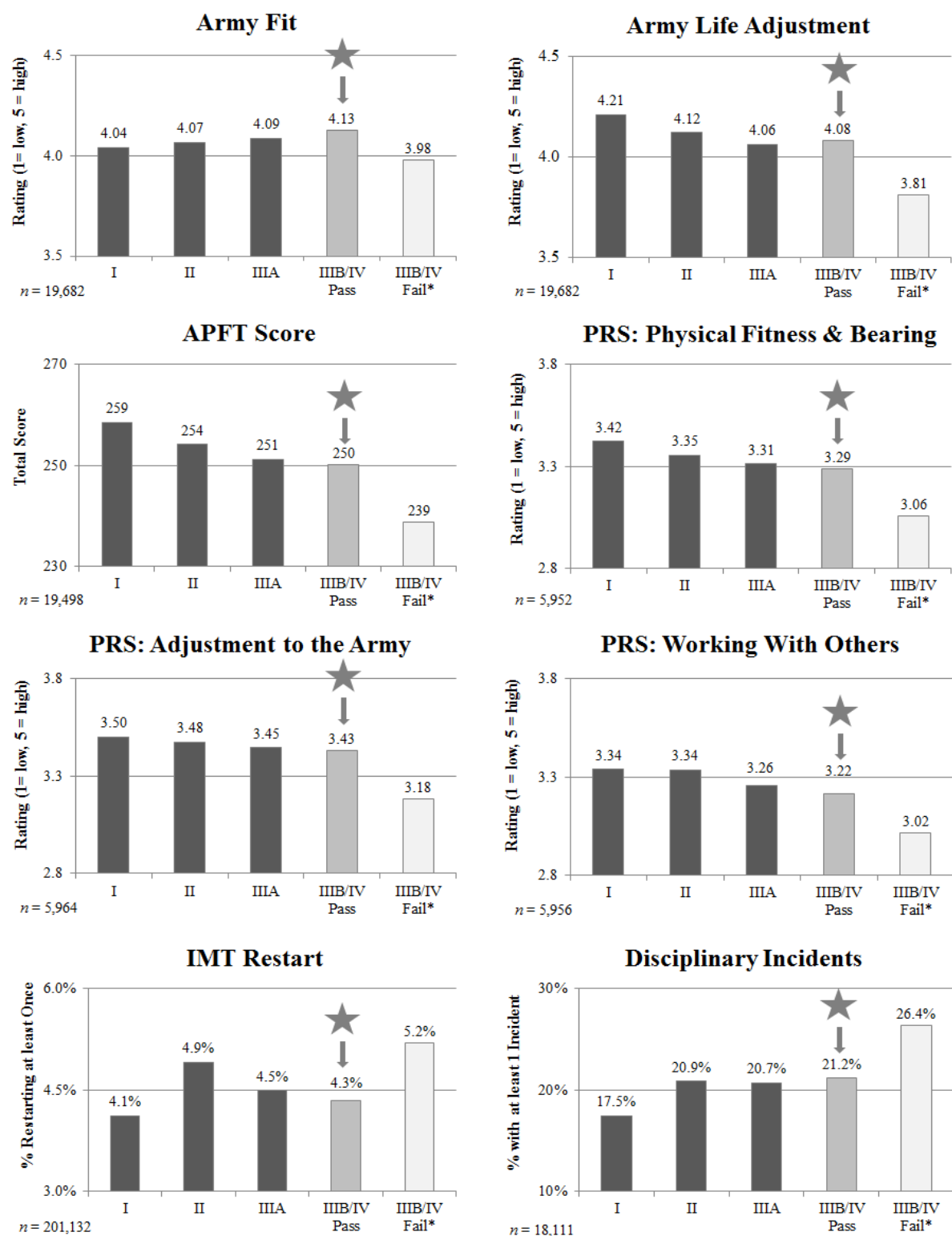
¹⁶ For the results presented in this report, this IIIB/IV Pass category includes Soldiers who scored at or above the 10th percentile on the Will-Do TOPS composite only. The scales used in the Adaptation composite were not administered to all Soldiers in the sample, so to maximize the sample size the Adaptation composite was not included in these analyses. In May 2015, the operational implementation of TAPAS was scaled back to allow all Tier 1 Category IIIB applicants to enlist regardless of their TAPAS scores.

category including IIIB/IV TAPAS Pass and IIIB/IV TAPAS Fail. The results show how well Soldiers who pass the TAPAS screen perform compared to those who fail the TAPAS screen.

We computed Soldiers' IMT mean scores for Army Fit, Army Life Adjustment, APFT Score, PRS: Physical Fitness and Bearing, PRS: Adjustment to the Army, and PRS: Working with Others; and Soldiers' IMT and attrition percent frequencies for IMT Restarts, Disciplinary Incidents, and each attrition outcome by AFQT category. For the in-unit criteria of WTBD JKT, APFT Score, PRS: Physical Fitness and Bearing, and PRS: Leadership Potential, we computed Soldiers' mean performance scores by quintiles, created based on TAPAS percentile scores.

The relationships between IMT criteria by selected AFQT categories for Tier 1 Soldiers are presented in Figure 5.7. As expected, these graphs demonstrate that Soldiers scoring higher on the AFQT generally receive more favorable scores on the IMT criteria with the exceptions of Army Fit and IMT Restarts. Importantly, this figure also highlights a clear distinction among AFQT Category IIIB/IV TAPAS Pass and IIIB/IV TAPAS Fail on various outcomes.

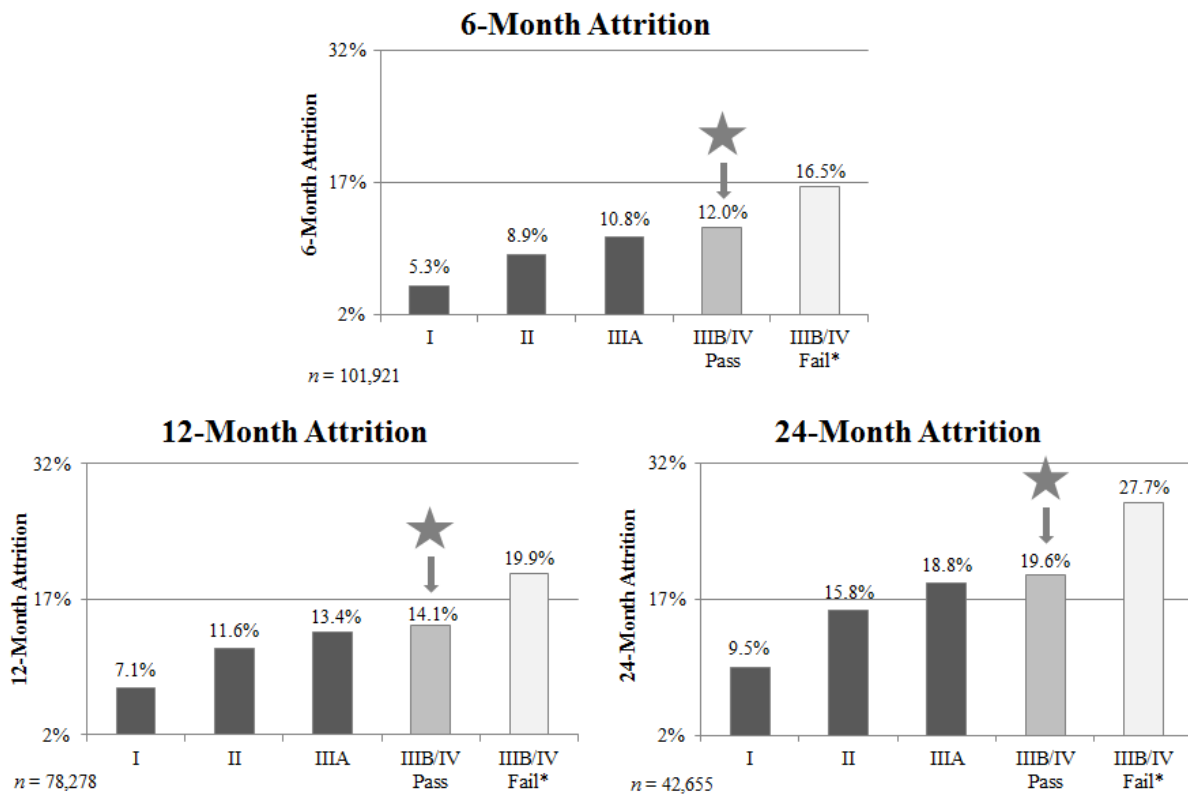
For every criterion, Soldiers in the IIIB/IV TAPAS Pass category received better scores than Soldiers in the IIIB/IV TAPAS Fail category. For the criteria of Army Fit and Army Life Adjustment, scores were more favorable for Soldiers in the IIIB/IV TAPAS Pass category than those in Category IIIA. Average APFT Score scores for IIIB/IV TAPAS Pass Soldiers were 11 points higher than IIIB/IV TAPAS Fail Soldiers. Additionally, Disciplinary Incidents and IMT Restarts for Soldiers in the IIIB/IV TAPAS Fail category were markedly worse than those observed in the IIIB/IV Pass category. In particular, Soldiers in the IIIB/IV TAPAS Pass category had fewer IMT Restarts than those in even the II and IIIA categories, and the percentage of IIIB/IV TAPAS Pass Soldiers was much closer to the percentage observed among II and IIIA Soldiers.



Note. * = Lowest scoring cognitive ability/TAPAS applicants screened out. Sample sizes vary based on availability of the outcome data.

Figure 5.7. Tier 1 Soldier outcomes for selected IMT criteria by AFQT category and TAPAS pass/fail status.

Similar trends are visible when examining attrition outcomes as shown in Figure 5.8. Attrition for Tier 1 Soldiers in the IIIB/IV TAPAS Fail category was greater than the IIIB/IV TAPAS Pass category at each of the time points. At 6 months, attrition was 4.5% higher among the IIIB/IV TAPAS Fail category than the IIIB/IV TAPAS Pass category. Moreover, the difference in attrition between Soldiers in the IIIB/IV TAPAS Pass and IIIB/IV TAPAS Fail categories increased over time, growing by 1.3% at 12 months and an additional 2.3% from 12 to 24 months.



Note. * = Lowest scoring cognitive ability/TAPAS applicants screened out. Sample sizes vary based on availability of attrition data.

Figure 5.8. Tier 1 Soldier attrition by AFQT category and TAPAS pass/fail status.

Figure 5.9 presents the relationship between TAPAS scores by quintiles and selected in-unit criteria. Results for these analyses are also presented for Tier 1 Soldiers only. With the exception of PRS: Leadership Potential, criterion performance generally improved in a linear fashion by TAPAS percentile group such that the top 20th percentile were among the highest performers. Moreover, Soldiers scoring in the bottom 20th percentile on the TAPAS performed lowest as a group for every outcome examined. For APFT Score scores in particular, the bottom 20th percentile scored on average seven points lower than the next TAPAS quintile and as much as 21 points lower than the top 20th percentile.

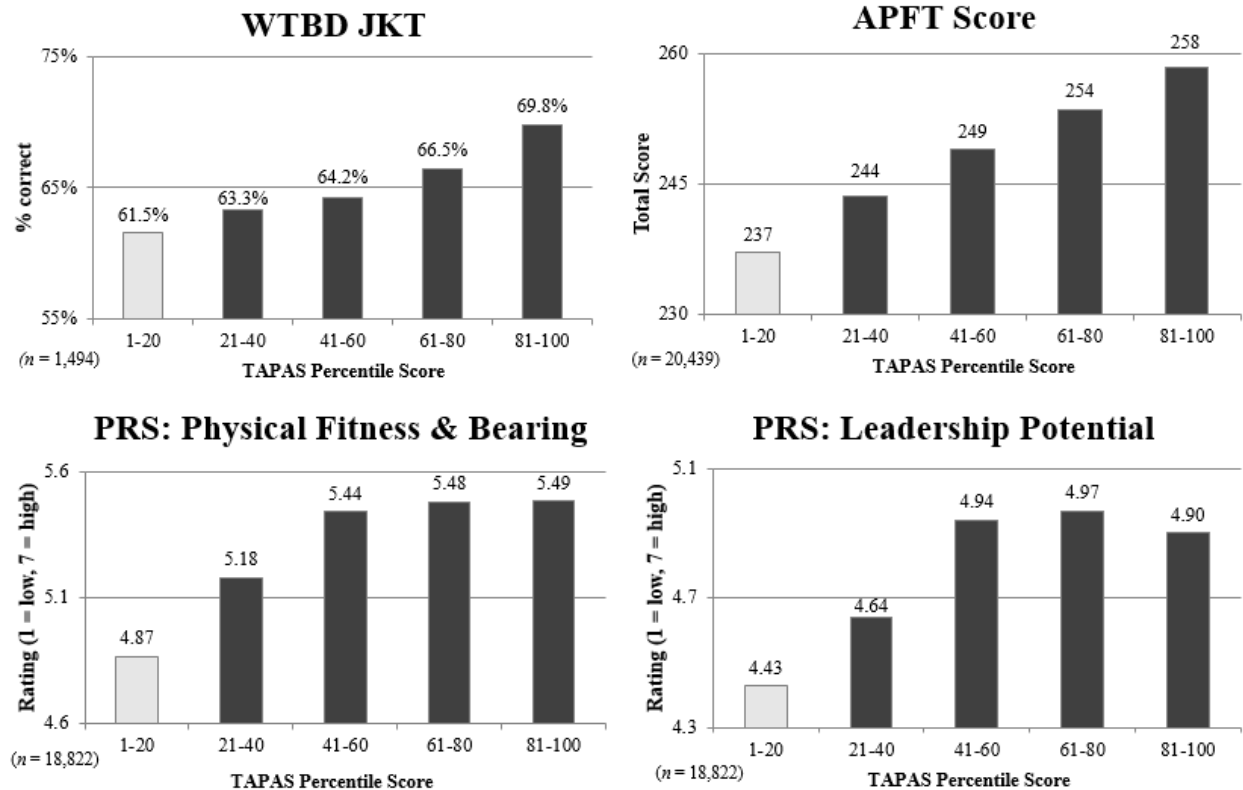


Figure 5.9. Tier 1 Soldier outcomes for selected in-unit criteria by TAPAS percentile score categories.

Summary

This chapter summarized results from the eighth cycle of the evaluation of criterion-related validity in the TOPS IOT&E. Overall, the TAPAS composites demonstrated incremental validity over the AFQT in predicting first-term Soldier performance and retention. In particular, the TAPAS Will-Do composite demonstrated the greatest incremental validity overall, with shrinkage-adjusted ΔR estimates for eight IMT and four in-unit criteria meeting or exceeding .05. The TAPAS Adaptation composite also yielded shrinkage-adjusted incremental validity estimates above .05 for four IMT criteria and two in-unit criteria. In addition, both the Will-Do and Adaptation composites demonstrated negative relationships with the dichotomous criteria, (i.e., attrition, Disciplinary Incidents, and IMT Restarts). Higher scores on these composites were associated with a reduction in the likelihood of these outcomes. Conversely, the TAPAS Can-Do composite did not provide any notable incremental validity for the criteria examined here. This finding is not surprising given the established strength of the AFQT in the prediction of cognitively-based criteria. Furthermore, results of the implementation analyses provide support for the utility of TAPAS. For both IMT and attrition criteria, AFQT Category IIIB/IV Soldiers scoring among the top 90% on the TAPAS Will-Do composite have markedly better outcomes than those Category IIIB/IV Soldiers scoring among the bottom 10%, and they often possess scores in line with AFQT Category IIIA Soldiers. Additionally, in-unit criterion scores are consistently lowest for Soldiers whose TAPAS scores are among the bottom 20th percentile.

CHAPTER 6: SUMMARY AND A LOOK AHEAD

Deirdre J. Knapp, Bethany H. Bynum (HumRRO), Heather M.K. Wolters, & Tonia S. Heffner (ARI)

Summary of the TOPS IOT&E Method

In an effort to expand the basis on which applicants are evaluated for enlistment, the Army is conducting an initial operational test and evaluation (IOT&E) of the *Tier One Performance Screen (TOPS)*. The TOPS assessment, the Tailored Adaptive Personality Assessment System (TAPAS), is being administered to non-prior service applicants testing at all MEPS locations.

To evaluate the TAPAS, the Army is collecting criterion data on Soldiers at multiple points during their time in service. Some outcome data are available from administrative records, including training course grades, training completion rates, and separation status. Data on additional measures are collected from Soldiers in selected MOS as they complete IMT. These measures include job knowledge tests (JKTs), an attitudinal person-environment fit assessment (the Army Life Questionnaire; ALQ), and performance rating scales (PRS) completed by the Soldiers' cadre members. Versions of the JKTs and the ALQ suited for Soldiers in their first enlistment term are also administered to Soldiers after they have joined their units. Performance ratings are collected from their supervisors at this time as well. Analysis datasets incorporating TAPAS and criterion data are constructed and cumulative validation analyses are being conducted at 6-month intervals throughout the IOT&E period.

The latest analysis data file (December 2013) includes a total of 486,310 applicants who took the TAPAS. Of these, 443,229 were in the TOPS Applicant Sample. The Applicant Sample was determined by excluding Education Tier 3, AFQT Category V, and prior service applicants from the master data file. Of that Applicant Sample, 226,055 (51.0%) had a record in at least one of the administrative criterion data sources; 23,495 had IMT data collected from the schoolhouse and 1,965 had in-unit criterion data.

Data from the JKTs, PRS, ALQ, and administrative sources were combined to yield an array of scores representing important Soldier outcomes. In general, the criterion scores exhibit acceptable psychometric properties and a sensible pattern of intercorrelations. The exception to this is the Army-wide and MOS-specific PRS, which continue to exhibit low interrater reliability. Despite the low reliabilities, the validation results show that variance in the ratings can be predicted.

Summary of Evaluation Results to Date

Evaluation results thus far suggest that the TAPAS holds promise for new Soldier selection. Results of the incremental validity analyses indicate that the TAPAS predicts important first-term criteria over and above the AFQT, especially measures tapping motivation-based aspects of Soldier performance, such as physical fitness, adjustment to Army life, commitment and fit, and discipline. Further, examination of AFQT categories and quintile splits of predictor composites

show a clear improvement in favor of higher scoring individuals. Individuals in the lowest category performed the worst. These findings are consistent with past evaluations in this series (Knapp & Heffner, 2011, 2012; Knapp et al., 2011; Knapp & LaPort, 2013a, 2013b, 2014) and the original research that led Army policy-makers to select TAPAS for the TOPS IOT&E (Knapp & Heffner, 2010).

The Will-Do composite was associated with the greatest incremental validity gains compared to other TAPAS composites. This was especially true for the prediction of physical fitness, Will Do Performance, and Army Life Adjustment. When examining outcomes by AFQT category, a clear distinction was seen when comparing the IIIB/IV TAPAS Pass group and the IIIB/IV TAPAS Fail group. The largest difference was for disciplinary incidents and training restarts. The Adaptation composite generally provided small incremental validity gains in attrition, with Adaptation showing larger gains for higher months in service. Even these small gains in validity are important, particularly given the modest relationship with the AFQT and the likelihood that results are attenuated due to operationally screening out some of the lowest TAPAS scorers. Results showed consistently higher attrition among the IIIB/IV TAPAS Fail AFQT category.

Looking Ahead

Changes to Predictor Measures

In September of 2013, a third series of new adaptive forms of the TAPAS were introduced at the MEPS. Each form measures 13 dimensions. Each of the three new forms assesses the same 10 core dimensions, plus three of seven experimental dimensions. The seven experimental dimensions assessed vary by version. In total, the newer versions of the TAPAS collectively measure 17 dimensions. The experimental dimensions will be evaluated for potential use in revised or new TAPAS composites, once sufficient data are available.

Analyses

Analyses will continue on the current semi-annual cycle to evaluate basic psychometric properties of the assessments, validation, and incremental validation analyses. As needed, we will examine the comparability of new TAPAS versions to prior forms before determining if the data can be combined for purposes of analysis. Additional analyses may include evaluation of the experimental TAPAS facets for potential use in revised or new TAPAS composites, or an alternative approach for modeling MOS classification outcomes. We will continue to update or to modify our evaluation analysis plans as the Army's goals for the TOPS IOT&E evolve or to better meet the informational needs of Army stakeholders.

Related Research

In addition to the evaluation described in this report, non-cognitive assessments are being examined to determine their prediction potential for MOS assignment. Temperament measures as well as interest inventories predicted attrition and job attitudes for a subset of MOS included in the research (Ingerick et al., 2009; Knapp et al., 2012). TAPAS data of Soldiers in four MOS were analyzed along with a wide variety of important outcomes (including Army-wide job

knowledge, attrition, disciplinary incidents, training restarts, and APFT score) and determined that approximately 40-45% of the Soldiers were predicted to perform better if they had been assigned to one of the other three MOSs (Nye et al., 2012). Ongoing research continues to pursue the utility of TAPAS as an assignment tool.

Conclusion

Inclusion of non-cognitive measures in initial entry screening allows the Army to predict a broader range of valued Army outcomes than traditional cognitive ability and educational credential screening. The TAPAS test, specifically, demonstrates the ability to predict an expanded concept of Soldier performance to include motivation, disciplinary behavior, adaptability, adjustment to military life, and attrition. Indeed, the TAPAS predicts these outcomes over and above the AFQT. Thus, TAPAS provides unique and valuable information regarding a recruit's potential success as a Soldier that is not captured elsewhere in the accession process. Additional research should continue to refine and expand the prediction potential of TAPAS and other non-cognitive measures.

REFERENCES

- Allen, M. T., Cheng, Y. A., Putka, D. J., Hunter, A., & White, L. (2010). Analysis and findings. In D.J. Knapp & T.S. Heffner (Eds.), *Expanded enlistment eligibility metrics (EEEM): Recommendations on a non-cognitive screen for new soldier selection* (Technical Report 1267) (pp 29-51). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Allen, M. T., Knapp, D. J., & Owens, K. S. (2013). *Validating future force measures (Army Class): Concluding analyses* (Technical Report in preparation). Fort Belvoir, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Biddle, D. (2005). *Adverse impact and test validation: A practitioner's guide to valid and defensible employment testing*. Farnham, England: Gower Publishing.
- Burket, G. R. (1964). A study of reduced rank models for multiple prediction. *Psychometric Monograph* (No. 12).
- Bynum, B. H., & Beatty, A.S. (2014). Description and psychometric properties of criterion measures. In D. J. Knapp & K.A. LaPort (Eds.), *Tier One Performance Screen initial operational test and evaluation: 2012 annual report* (Technical Report 1342) (pp. 15-31). Fort Belvoir: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Campbell, J. P., & Knapp, D. J. (Eds.) (2001). *Exploring the limits in personnel selection and classification*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Campbell, J. P., Hanson, M. A., & Oppler S. H. (2001). Modeling performance in a population of jobs. In J. P. Campbell & D. J. Knapp (Eds.), *Exploring the limits in personnel selection and classification* (pp. 307-334). Hillsdale, NJ: Erlbaum.
- Campbell, J. P., McHenry, J. J., & Wise, L. L. (1990). Modeling job performance in a population of jobs. *Personnel Psychology*, 43, 313-333.
- Cascio, W. F. & Aguinis, H. (2005). *Applied Psychology in Human Resource Management* (6th ed.) Upper Saddle River, NJ: Pearson Prentice Hall.
- Chernyshenko, O. S., & Stark, S. (2007, October). *Criterion validity evidence for narrow temperament clusters: A meta-analysis of military studies*. Paper presented at the 49th annual conference of the International Military Testing Association. Gold Coast, AU.
- Cleary, T.A. (1968). Test bias: Prediction of grades of Negro and white students in integrated colleges. *Journal of Educational Measurement*, 5, 115-124
- Cohen, J., Cohen, P., West, S.G., & Aiken, L.S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

- Collins, M., Le, H., & Schantz, L. (2005). Job knowledge criterion tests. In D.J. Knapp & T.R. Tremble (Eds.), *Development of experimental Army enlisted personnel selection and classification tests and job performance criteria* (Technical Report 1168) (pp. 49-58). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Cook, N. R. (2007). Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*, 115, 928-935.
- De Corte, W., Lievens, F., & Sackett, P. R. (2007). Combining predictors to achieve optimal trade-offs between selection quality and adverse impact. *Journal of Applied Psychology*, 92, 1380-1393.
- De Corte, W. Lievens, F., & Sackett, P.R. (2011). Designing pareto-optimal selection systems: Formalizing the decisions required for selection system development. *Journal of Applied Psychology*, 26, 907-926.
- Drasgow, F., Embretson, S. E., Kyllonen, P. C., & Schmitt, N. (2006). *Technical review of the Armed Services Vocational Aptitude Battery (ASVAB)* (FR-06-25). Alexandria, VA: Human Resources Research Organization.
- Drasgow, F., Stark, S., Chernyshenko, O. S., Nye, C. D., Hulin, C. L., & White, L. A. (2012). *Development of the Tailored Adaptive Personality Assessment System (TAPAS) to support Army selection and classification decisions* (Technical Report 1311). Fort Belvoir, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Ingerick, M., Diaz, T., & Putka, D. (2009). *Investigations into Army enlisted classification systems: Concurrent validation report* (Technical Report 1244). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Knapp, D. J., & Heffner, T. S. (Eds.) (2009). *Predicting Future Force Performance (Army Class): End of Training Longitudinal Validation* (Technical Report 1257). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Knapp, D. J., & Heffner, T. S. (Eds.) (2010). *Expanded Enlistment Eligibility Metrics (EEEM): Recommendations on a non-cognitive screen for new soldier selection* (Technical Report 1267). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Knapp, D. J., & Heffner, T. S. (Eds.) (2011). *Tier One Performance Screen initial operational test and evaluation: 2010 annual report* (Technical Report 1296). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Knapp, D. J., & Heffner, T. S. (Eds.) (2012). *Tier One Performance Screen initial operational test and evaluation: 2011 interim report* (Technical Report 1306). Fort Belvoir, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Knapp, D. J., & LaPort, K. (Eds.) (2013a). *Tier One Performance Screen initial operational test and evaluation: 2011 annual report* (Technical Report 1325). Fort Belvoir, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

- Knapp, D. J., & LaPort, K. (Eds.) (2013b). *Tier One Performance Screen initial operational test and evaluation: 2012 interim report* (Technical Report 1332). Fort Belvoir, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Knapp, D. J., & LaPort, K. (Eds.) (2014). *Tier One Performance Screen initial operational test and evaluation: 2012 annual report* (Technical Report 1342). Fort Belvoir, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Knapp, D. J., & Tremble, T. R. (Eds.) (2007). *Concurrent validation of experimental Army enlisted personnel selection and classification measures* (Technical Report 1205). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Knapp, D. J., Heffner, T. S., & White, L. (Eds.) (2011). *Tier One Performance Screen initial operational test and evaluation: Early results* (Technical Report 1283). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Knapp, D. J., Owens, K. S., & Allen, M. T. (Eds.) (2012). *Validating future force performance measures (Army Class): In-unit performance longitudinal validation* (Technical Report 1314). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- McHenry, J.J., Hough, L.M., Toquam, J.L., Hanson, M.A., & Ashworth, S. (1990). Project A validity results: The relationship between predictor and criterion domains. *Personnel Psychology*, 43, 335-354.
- Moriarty, K. O., & Bynum, B. H. (2011). Description and psychometric properties of criterion measures. In D. J. Knapp & T. S. Heffner (Eds.), *Tier One Performance Screen initial operational test and evaluation: 2010 annual report* (Technical Report 1296) (pp. 20-28). Fort Belvoir: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Moriarty, K. O., Campbell, R. C., Heffner, T. S., & Knapp, D. J. (2009). *Validating future force performance measures (Army Class): Reclassification test and criterion development* (Research Product 2009-11). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Nye, C. D., Drasgow, F., Chernyshenko, O. S., Stark, S., Kubisiak, U. C., White, L. A., & Jose, I. (2012). *Assessing the Tailored Adaptive Personality Assessment System (TAPAS) as an MOS qualification instrument*. Fort Belvoir, VA: U. S. Army Research Institute for the Behavioral and Social Sciences.
- Putka, D.J., Hoffman, B.J., & Carter, N.T. (2014). Correcting the correction: When individual raters offer distinct but valid perspectives. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 7, 546-551.
- Putka, D. J., & Van Iddekinge, C. H. (2007). Work Preferences Survey. In D. J. Knapp & T. R. Tremble (Eds.), *Concurrent validation of experimental Army enlisted personnel selection and classification measures* (Technical Report 1205) (pp. 135-155). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

- Putka, D. J., Le, H., McCloy, R. A., & Diaz, T. (2008). Ill-structured measurement designs in organizational research: Implications for estimating interrater reliability. *Journal of Applied Psychology*, 93, 959-981.
- Russell, T. L., & Sellman, W. S. (Eds.) (2009). *Development and pilot testing of an information and communications technology literacy test for military enlistees: Volume I technical report* (FR 08-128). Alexandria, VA: Human Resources Research Organization.
- Schmitt, N., & Ployhart, R. E. (1999). Estimates of cross-validity for stepwise regression and with predictor selection. *Journal of Applied Psychology*, 84, 50-57.
- Sparks, T. E., & Peddie, C. (2013). Description and psychometric properties of criterion measures. In D. J. Knapp & T. S. Heffner (Eds.), *Tier One Performance Screen initial operational test and evaluation: 2012 interim report* (Technical Report 1332) (pp. 15-31). Fort Belvoir: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Stark, S. E., Chernyshenko, O. S., & Drasgow, F. (2005). An IRT approach to constructing and scoring pairwise preference items involving stimuli on different dimensions: The multi-unidimensional pairwise preference model. *Applied Psychological Measurement*, 29, 184-201.
- Stark, S. E., Chernyshenko, O. S., & Drasgow, F. (2010). Tailored adaptive personality assessment system (TAPAS-95s). In D. J. Knapp & T. S. Heffner (Eds.), *Expanded enlistment eligibility metrics (EEEM): Recommendations on a non-cognitive screen for new soldier selection* (Technical Report 1267) (pp 29-51). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Stark, S. E., Chernyshenko, O. S., & Drasgow, F. (2012). Adaptive testing with multi-unidimensional pairwise preference Items: Improving the efficiency of personality and other noncognitive assessments. *Organizational Research Methods*, 15, 463-487.
- Stark, S., Chernyshenko, O. S., Drasgow, F., Nye, C. D., White, L. A., Heffner, T., & Farmer, W. L. (2014). From ABLE to TAPAS: A new generation of personality tests to support military selection and classification decisions. *Military Psychology*, 26, 153-164.
- Strickland, W. J. (Ed.) (2005). *A longitudinal examination of first term attrition and reenlistment among FY1999 enlisted accessions* (Technical Report 1172). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Tabachnick, B.G., & Fidell, L.S. (2013). *Using multivariate statistics* (6th ed.). Boston, MA: Pearson.
- Van Iddekinge, C. H., Putka, D. J., & Sager, C. E. (2005). Attitudinal criteria. In D. J. Knapp & T. R. Tremble (Eds.), *Development of experimental Army enlisted personnel selection and classification tests and job performance criteria* (Technical Report 1168) (pp. 89-104). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

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, 1167-1194.

White, L. A., & Young, M. C. (1998, August). *Development and validation of the Assessment of Individual Motivation (AIM)*. Paper presented at the annual meeting of the American Psychological Association, San Francisco, CA.

APPENDIX A

PREDICTOR MEASURE PSYCHOMETRIC PROPERTIES IN THE APPLICANT SAMPLE

Table A.1. Raw Mean and Standard Deviations for the TAPAS Composites and TAPAS Scales on the 15D-Static and 15D-CAT-v4 Forms (June 2009-August 2011)

TAPAS Composite/ TAPAS Scale	15D-Static/CAT-v4					
	Tier 1+ 2 (Combined) (n = 168,425)		Tier 1 (n = 160,652)		Tier 2 (n = 7,773)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
<i>Individual TAPAS Scales ^a</i>						
Achievement	.16	.48	.16	.48	.19	.48
Adjustment	.00	.57	.00	.57	.07	.58
Adventure Seeking	--	--	--	--	--	--
Attention Seeking	-.22	.53	-.21	.53	-.25	.52
Commitment to Serve	--	--	--	--	--	--
Cooperation	-.06	.37	-.06	.37	-.07	.38
Courage	--	--	--	--	--	--
Dominance	.02	.59	.02	.59	-.02	.60
Even Tempered	.17	.48	.16	.48	.22	.47
Intellectual Efficiency	-.04	.58	-.04	.58	-.01	.56
Non-Delinquency	.10	.46	.10	.46	.03	.48
Optimism	.15	.46	.15	.46	.16	.46
Order	-.42	.55	-.42	.55	-.42	.53
Physical Conditioning	.04	.62	.04	.63	-.10	.59
Responsibility	--	--	--	--	--	--
Self-Control	.07	.53	.06	.53	.13	.53
Selflessness	-.20	.43	-.20	.43	-.18	.43
Situational Awareness	--	--	--	--	--	--
Sociability	-.06	.59	-.06	.59	-.10	.59
Team Orientation	--	--	--	--	--	--
Tolerance	-.23	.57	-.23	.57	-.21	.56
<i>TAPAS Composites</i>						
Can-Do	99.83	20.06	99.75	20.10	101.35	19.16
Will-Do	100.45	20.06	100.57	20.08	97.92	19.52
Adaptation	100.62	20.01	100.71	20.04	98.59	19.19

Note. Results are limited to the TOPS Applicant Sample (non-prior service, Education Tier 1 and 2, and AFQT ≥ 10) with valid TAPAS score data.

^a Not all TAPAS scales were administered in every version; missing *M* and *SD* indicate that the scale was not administered.

Table A.2. Raw Mean and Standard Deviations for the TAPAS Composites and TAPAS Scales on 15D-CAT-v5 (August 2011-September 2013)

TAPAS Composite/ TAPAS Scale	15D-CAT-v5					
	Tier 1+ 2 (Combined) (<i>n</i> = 49,669)		Tier 1 (<i>n</i> = 46,747)		Tier 2 (<i>n</i> = 2,922)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
<i>Individual TAPAS Scales^a</i>						
Achievement	.23	.49	.23	.49	.29	.49
Adjustment	.09	.39	.09	.39	.15	.39
Adventure Seeking	--	--	--	--	--	--
Attention Seeking	-.34	.59	-.35	.59	-.29	.61
Commitment to Serve	--	--	--	--	--	--
Cooperation	.22	.53	.22	.53	.17	.51
Courage	--	--	--	--	--	--
Dominance	.36	.51	.36	.51	.34	.52
Even Tempered	.28	.49	.28	.49	.36	.50
Intellectual Efficiency	.05	.53	.04	.53	.15	.52
Non-Delinquency	.17	.53	.17	.52	.15	.56
Optimism	.29	.45	.29	.45	.28	.45
Order	-.26	.54	-.26	.54	-.27	.53
Physical Conditioning	.15	.56	.15	.56	.07	.54
Responsibility	--	--	--	--	--	--
Self-Control	-.24	.48	-.25	.48	-.17	.49
Selflessness	.06	.43	.07	.43	.01	.45
Situational Awareness	--	--	--	--	--	--
Sociability	-.18	.56	-.18	.56	-.15	.58
Team Orientation	--	--	--	--	--	--
Tolerance	.00	.52	.00	.52	.04	.51
<i>TAPAS Composites</i>						
Can-Do	99.68	19.15	99.48	19.14	102.94	19.05
Will-Do	99.44	19.95	99.49	19.97	98.55	19.57
Adaptation	99.77	20.44	99.83	20.49	98.87	19.62

Note. Results are limited to the TOPS Applicant Sample (non-prior service, Education Tier 1 and 2, and AFQT ≥ 10) with valid TAPAS score data.

^a Not all TAPAS scales were administered in every version; missing *M* and *SD* indicate that the scale was not administered.

Table A.3. Raw Mean and Standard Deviations for the TAPAS Composites and TAPAS Scales on 15D-CAT-v7 (August 2011-September 2013)

TAPAS Composite/ TAPAS Scale	15D-CAT-v7					
	Tier 1+ 2 (Combined) (n = 99,683)		Tier 1 (n = 93,806)		Tier 2 (n = 5,877)	
	M	SD	M	SD	M	SD
<i>Individual TAPAS Scales ^a</i>						
Achievement	.23	.47	.23	.47	.27	.47
Adjustment	.09	.38	.08	.38	.15	.37
Adventure Seeking	-.29	.58	-.29	.58	-.25	.58
Attention Seeking	-.34	.59	-.34	.59	-.31	.61
Commitment to Serve	.17	.52	.17	.52	.27	.49
Cooperation	.19	.52	.19	.52	.18	.52
Courage	--	--	--	--	--	--
Dominance	.31	.50	.32	.50	.29	.51
Even Tempered	.27	.47	.26	.47	.34	.48
Intellectual Efficiency	.03	.52	.02	.52	.10	.51
Non-Delinquency	.17	.53	.17	.53	.15	.55
Optimism	.24	.44	.24	.44	.24	.45
Order	-.23	.54	-.23	.54	-.25	.54
Physical Conditioning	.11	.55	.11	.55	.03	.53
Responsibility	--	--	--	--	--	--
Self-Control	--	--	--	--	--	--
Selflessness	.08	.44	.09	.44	.05	.45
Situational Awareness	.02	.49	.01	.49	.10	.50
Sociability	--	--	--	--	--	--
Team Orientation	--	--	--	--	--	--
Tolerance	--	--	--	--	--	--
<i>TAPAS Composites ^b</i>						
Can-Do	--	--	--	--	--	--
Will-Do	99.53	19.82	99.63	19.83	97.89	19.54
Adaptation	--	--	--	--	--	--

Note. Results are limited to the TOPS Applicant Sample (non-prior service, Education Tier 1 and 2, and AFQT ≥ 10) with valid TAPAS score data.

^a Not all TAPAS scales were administered in every version; missing *M* and *SD* indicate that the scale was not administered.

^b A subset of the scales that compose the Can-Do and Adaptation composite were not administered in this version of the TAPAS. Composites could not be computed.

Table A.4. Raw Mean and Standard Deviations for the TAPAS Composites and TAPAS Scales on 15D-CAT-v8 (August 2011-September 2013)

TAPAS Composite/ TAPAS Scale	15D-CAT-v8					
	Tier 1+ 2 (Combined) (n = 99,717)		Tier 1 (n = 93,922)		Tier 2 (n = 5,795)	
	M	SD	M	SD	M	SD
<i>Individual TAPAS Scales ^a</i>						
Achievement	.21	.47	.20	.47	.25	.46
Adjustment	.05	.38	.04	.38	.10	.37
Adventure Seeking	--	--	--	--	--	--
Attention Seeking	-.34	.58	-.34	.58	-.30	.60
Commitment to Serve	--	--	--	--	--	--
Cooperation	--	--	--	--	--	--
Courage	.11	.54	.10	.54	.19	.54
Dominance	.30	.49	.30	.49	.27	.49
Even Tempered	.30	.51	.30	.51	.37	.52
Intellectual Efficiency	.03	.52	.02	.52	.12	.50
Non-Delinquency	.18	.53	.18	.53	.14	.54
Optimism	.26	.43	.26	.43	.25	.45
Order	--	--	--	--	--	--
Physical Conditioning	.10	.55	.10	.55	.00	.52
Responsibility	.33	.46	.33	.46	.37	.47
Self-Control	-.23	.46	-.23	.46	-.17	.46
Selflessness	--	--	--	--	--	--
Situational Awareness	--	--	--	--	--	--
Sociability	-.17	.55	-.17	.55	-.15	.56
Team Orientation	-.07	.48	-.07	.48	-.06	.51
Tolerance	-.02	.51	-.03	.51	.00	.51
<i>TAPAS Composites ^b</i>						
Can-Do	--	--	--	--	--	--
Will-Do	99.58	19.65	99.70	19.68	97.60	19.03
Adaptation	--	--	--	--	--	--

Note. Results are limited to the TOPS Applicant Sample (non-prior service, Education Tier 1 and 2, and AFQT ≥ 10) with valid TAPAS score data.

^a Not all TAPAS scales were administered in every version; missing *M* and *SD* indicate that the scale was not administered.

^b A subset of the scales included in the Can-Do and Adaptation composite were not administered in this version of the TAPAS. Composites could not be computed.

Table A.5. Correlations between TAPAS Composites and TAPAS Scales with AFQT in the TOPS Applicant Sample by Version

TAPAS Composite/ TAPAS Scale	15D-Static/CAT-v4 (June 2009-August 2011)			15D-CAT-v5/7/8 (August 2011-September 2013)		
	Tier 1+ 2 (Combined)	Tier 1	Tier 2	Tier 1+ 2 (Combined)	Tier 1	Tier 2
<i>n</i>	168,425	160,652	7,773	49,669 - 249,069	46,747 - 234,475	2,922 - 14,594
<i>Individual TAPAS Scales</i>						
Achievement	.09	.10	.05	.04	.04	.00
Adjustment	.11	.11	.09	.13	.13	.11
Adventure Seeking	--	--	--	.10	.10	.05
Attention Seeking	.11	.11	.08	.01	.01	.00
Commitment to Serve	--	--	--	-.14	-.15	-.06
Cooperation	.00	.01	.01	-.12	-.12	-.09
Courage	--	--	--	.06	.06	.04
Dominance	.08	.08	.01	.12	.12	.03
Even Tempered	.08	.08	.07	.09	.09	.09
Intellectual Efficiency	.41	.42	.37	.30	.30	.27
Non-Delinquency	-.01	-.01	.02	-.06	-.06	-.02
Optimism	.01	.01	-.01	.10	.10	.06
Order	-.18	-.18	-.17	-.16	-.16	-.15
Physical Conditioning	.05	.05	-.02	.06	.06	-.01
Responsibility	--	--	--	.14	.15	.09
Self-Control	-.01	-.01	.02	-.04	-.04	-.02
Selflessness	-.07	-.07	-.05	-.07	-.07	-.11
Situational Awareness	--	--	--	.01	.01	.04
Sociability	-.09	-.09	-.06	-.12	-.12	-.08
Team Orientation	--	--	--	-.10	-.11	-.09
Tolerance	-.02	-.02	-.01	.08	.08	.05
<i>TAPAS Composites</i>						
Can-Do	.45	.45	.40	.37	.38	.31
Will-Do	.09	.10	.01	.10	.11	.02
Adaptation	.19	.19	.12	.20	.21	.14

Note. Results are limited to the TAPAS Applicant Sample (non-prior service, Education Tier 1 and 2, AFQT Category IV and above) with valid TAPAS score data. Correlations in bold are statistically significant, $p < .01$ (two-tailed).

Table A.6. Basic Descriptive Statistics for AFQT, ASVAB Subtests, and Aptitude Area (AA) Composites by Version in the TOPS Applicant Sample

Composite/Subtest	15D-Static/CAT-v4 (June 2009-August 2011)					15D-CAT-v5/7/8 (August 2011-September 2013)				
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
<i>AFQT</i>	168,425	57.03	23.06	10	99	249,069	52.81	22.14	10	99
<i>ASVAB Subtests</i>										
Arithmetic Reasoning (AR)	168,403	52.60	7.67	18	72	249,035	51.18	7.62	21	72
Assembling Objects (AO)	165,850	55.22	7.78	25	70	234,381	54.15	7.84	26	70
Auto & Shop Information (AS)	168,402	49.98	9.39	19	86	249,033	47.54	8.97	20	86
Electronics Information (EI)	168,402	52.03	9.06	16	84	249,034	50.10	8.90	15	84
General Science (GS)	168,403	51.71	8.38	19	76	249,035	50.46	8.23	20	76
Math Knowledge (MK)	168,403	53.47	6.96	24	73	249,035	53.05	6.72	25	73
Mechanical Comprehension (MC)	168,402	53.46	8.41	14	82	249,033	51.66	8.20	23	82
Paragraph Comprehension (PC)	168,403	52.89	7.03	21	69	249,035	51.72	6.83	22	69
Word Knowledge (WK)	168,403	51.37	8.02	16	76	249,035	49.80	7.70	15	76
<i>Aptitude Area (AA) Composites</i>										
Clerical (CL)	168,411	105.94	13.86	35	152	249,049	103.08	13.31	56	153
Combat (CO)	168,411	105.74	14.81	29	160	249,049	102.23	14.18	54	160
Electronics (EL)	168,411	105.54	14.80	29	160	249,049	101.91	14.18	52	160
Field Artillery (FA)	168,411	105.91	14.73	28	159	249,049	102.43	14.11	55	160
General Maintenance (GM)	168,411	105.24	15.27	28	161	249,049	101.44	14.65	51	162
General Technical (GT)	168,412	104.99	14.26	39	149	249,052	101.87	13.78	49	149
Mechanical Maintenance (MM)	168,411	104.35	16.27	25	165	249,049	99.94	15.59	48	167
Operators and Food Service (OF)	168,411	105.24	15.23	27	160	249,049	101.34	14.62	52	161
Signal Communications (SC)	168,411	105.93	14.42	29	159	249,049	102.59	13.81	53	159
Skilled Technical (ST)	168,411	105.80	14.44	32	157	249,049	102.45	13.82	54	158

Note. Results are limited to the TOPS Applicant Sample (non-prior service, Education Tier 1 and 2, AFQT Category IV and above) with valid TAPAS score data.

Table A.7. Basic Descriptive Statistics for AFQT, ASVAB Subtests, and Aptitude Area (AA) Composites in 15D Static and 15D-CAT-v4 by Education Tier

Composite/Subtest	Tier 1					Tier 2				
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
AFQT	160,652	57.19	23.22	10	99	7,773	53.79	19.16	10	99
<i>ASVAB Subtests</i>										
Arithmetic Reasoning (AR)	160,637	52.64	7.72	18	72	7,766	51.81	6.37	24	72
Assembling Objects (AO)	158,236	55.25	7.79	25	70	7,614	54.72	7.55	26	69
Auto & Shop Information (AS)	160,636	49.88	9.39	19	86	7,766	52.15	8.99	26	81
Electronics Information (EI)	160,636	51.98	9.10	16	84	7,766	52.98	7.98	18	83
General Science (GS)	160,637	51.72	8.44	19	76	7,766	51.52	7.04	23	75
Math Knowledge (MK)	160,637	53.65	6.96	24	73	7,766	49.78	5.71	27	73
Mechanical Comprehension (MC)	160,636	53.45	8.45	14	82	7,766	53.75	7.57	23	79
Paragraph Comprehension (PC)	160,637	52.88	7.07	23	69	7,766	53.06	6.18	21	69
Word Knowledge (WK)	160,637	51.35	8.08	16	76	7,766	51.73	6.70	22	76
<i>Aptitude Area (AA) Composites</i>										
Clerical (CL)	160,641	106.05	13.98	35	152	7,770	103.66	10.91	56	145
Combat (CO)	160,641	105.81	14.93	29	160	7,770	104.40	12.02	51	154
Electronics (EL)	160,641	105.58	14.92	29	160	7,770	104.69	11.96	52	153
Field Artillery (FA)	160,641	105.98	14.85	28	159	7,770	104.49	11.91	51	154
General Maintenance (GM)	160,641	105.26	15.39	28	161	7,770	104.70	12.54	48	155
General Technical (GT)	160,642	105.01	14.39	39	149	7,770	104.48	11.42	54	145
Mechanical Maintenance (MM)	160,641	104.30	16.37	25	165	7,770	105.40	14.04	46	158
Operators and Food Service (OF)	160,641	105.25	15.35	27	160	7,770	105.03	12.52	50	154
Signal Communications (SC)	160,641	106.01	14.54	29	159	7,770	104.32	11.52	54	153
Skilled Technical (ST)	160,641	105.86	14.56	32	157	7,770	104.52	11.51	56	150

Note. Results are limited to the TOPS Applicant Sample (non-prior service, Education Tier 1 and 2, AFQT Category IV and above) with valid TAPAS score data.

Table A.8. Basic Descriptive Statistics for AFQT, ASVAB Subtests, and Aptitude Area (AA) Composites in 15D-CAT-v5/7/8 by Education Tier

Composite/Subtest	Tier 1					Tier 2				
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
<i>AFQT</i>	234,475	52.83	22.34	10	99	14,594	52.39	18.56	10	99
<i>ASVAB Subtests</i>										
Arithmetic Reasoning (AR)	234,448	51.17	7.69	21	72	14,587	51.27	6.43	27	72
Assembling Objects (AO)	220,496	54.14	7.85	26	70	13,885	54.31	7.70	26	69
Auto & Shop Information (AS)	234,446	47.35	8.95	20	86	14,587	50.71	8.82	23	82
Electronics Information (EI)	234,447	49.98	8.94	15	84	14,587	52.05	7.92	19	82
General Science (GS)	234,448	50.43	8.30	20	76	14,587	50.99	7.03	20	76
Math Knowledge (MK)	234,448	53.25	6.74	25	73	14,587	49.75	5.50	26	72
Mechanical Comprehension (MC)	234,446	51.58	8.24	23	82	14,587	53.00	7.53	23	79
Paragraph Comprehension (PC)	234,448	51.67	6.87	22	69	14,587	52.64	6.03	24	69
Word Knowledge (WK)	234,448	49.72	7.76	15	76	14,587	51.07	6.58	21	76
<i>Aptitude Area (AA) Composites</i>										
Clerical (CL)	234,458	103.11	13.45	56	153	14,591	102.57	10.70	61	144
Combat (CO)	234,458	102.19	14.31	54	160	14,591	102.84	11.86	60	150
Electronics (EL)	234,458	101.83	14.31	52	160	14,591	103.07	11.80	59	150
Field Artillery (FA)	234,458	102.40	14.25	55	160	14,591	102.98	11.76	61	149
General Maintenance (GM)	234,458	101.34	14.77	51	162	14,591	102.95	12.43	58	152
General Technical (GT)	234,461	101.78	13.91	49	149	14,591	103.26	11.29	58	148
Mechanical Maintenance (MM)	234,458	99.74	15.66	48	167	14,591	103.21	13.91	55	155
Operators and Food Service (OF)	234,458	101.22	14.73	52	161	14,591	103.25	12.41	58	152
Signal Communications (SC)	234,458	102.57	13.95	53	159	14,591	102.91	11.31	60	148
Skilled Technical (ST)	234,458	102.41	13.96	54	158	14,591	103.13	11.32	61	147

Note. Results are limited to the TOPS Applicant Sample (non-prior service, Education Tier 1 and 2, AFQT Category IV and above) with valid TAPAS score data.

Table A.9. Correlations among TAPAS Scale Scores in the Applicant Sample

TAPAS Scale	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. Achievement																				
2. Adjustment	.10																			
3. Adventure Seeking	.10	.15																		
4. Attention Seeking	.04	.09	.17																	
5. Commitment to Serve	.12	.05	.04	-.01																
6. Cooperation	.10	.08	-.13	-.08	.03															
7. Courage	.21	.16	--	.10	--	--														
8. Dominance	.30	.11	.13	.18	.06	.02	.21													
9. Even Tempered	.11	.22	-.05	-.05	.05	.29	.05	-.02												
10. Intellectual Efficiency	.25	.19	.07	.09	.02	.01	.18	.26	.09											
11. Non-Delinquency	.19	.01	-.17	-.15	.12	.25	.03	.00	.23	.03										
12. Optimism	.16	.26	.02	.08	.02	.16	.03	.15	.19	.11	.12									
13. Order	.17	-.05	-.08	-.07	.06	.11	--	.10	.02	.05	.13	.02								
14. Physical Conditioning	.17	.06	.25	.09	.01	-.03	.12	.19	-.08	.05	-.05	.06	.05							
15. Responsibility	.30	.11	--	-.05	--	--	.14	.15	.16	.16	.22	.15	--	.05						
16. Self-Control	.20	.06	--	-.06	--	.07	.10	-.04	.17	.14	.22	.05	.15	-.06	.22					
17. Selflessness	.13	-.01	-.04	-.10	.04	.27	--	.11	.13	.01	.18	.09	.11	-.01	--	.02				
18. Situational Awareness	.19	.15	.10	.04	.07	.00	--	.11	.10	.25	.12	.08	.16	.06	--	--	.04			
19. Sociability	.06	.09	--	.34	--	.11	.09	.18	.01	.03	-.04	.15	-.05	.08	.03	-.04	.05	--		
20. Team Orientation	.06	.05	--	.13	--	--	.03	.11	.06	-.04	.05	.06	--	.06	.02	.05	--	--	.23	
21. Tolerance	.10	.04	--	.01	--	.17	.03	.09	.15	.10	.08	.11	.05	-.04	.08	.05	.32	--	.10	.08

Note. Results are limited to the Applicant Sample (Non-prior service, Education Tier 1 and 2, AFQT Category IV and above) with valid TAPAS score data, $n = 99,683 - 417,494$.

Not all TAPAS scales were administered in every version; missing correlations indicate that the scales were not administered on the same version. The correlation between the Can-Do and Will-Do predictor composites is $r = .16$; the correlation between the Can-Do and Adaptation predictor composites is $r = .40$; the correlation between the Will-Do and Adaptation predictor composites is $r = .49$. Correlations in bold are statistically significant, $p < .01$ (two-tailed).

Table A.10. Correlations among TAPAS Scale Scores in the Applicant Sample by Education Tier

TAPAS Scale	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1 Achievement		.09	.09	.05	.14	.10	.21	.28	.11	.24	.23	.15	.17	.18	.31	.20	.14	.21	.06	.07	.09
2 Adjustment	.10		.11	.10	.04	.09	.12	.08	.20	.16	.05	.26	-.04	.03	.10	.08	.00	.14	.10	.03	.05
3 Adventure Seeking	.10	.15		.20	.07	-.14	--	.15	-.08	.07	-.15	.01	-.05	.23	--	--	-.03	.10	--	--	--
4 Attention Seeking	.03	.09	.17		.02	-.06	.11	.20	-.05	.07	-.11	.08	-.05	.08	-.05	-.07	-.06	.04	.35	.16	.04
5 Commitment to Serve	.12	.04	.04	-.01		.04	--	.08	.05	.05	.12	.02	.05	.05	--	--	.05	.07	--	--	--
6 Cooperation	.10	.08	-.13	-.08	.03		--	.02	.30	.03	.27	.16	.10	-.02	--	.08	.26	.00	.14	--	.16
7 Courage	.21	.16	--	.10	--	--		.21	.03	.18	.06	.03	--	.11	.17	.10	--	--	.09	.03	.03
8 Dominance	.30	.11	.13	.18	.06	.02	.22		-.03	.25	.02	.11	.11	.20	.16	-.03	.08	.13	.18	.13	.07
9 Even Tempered	.11	.22	-.04	-.05	.05	.29	.05	-.02		.09	.27	.21	.02	-.08	.17	.17	.14	.10	.04	.06	.14
10 Intellectual Efficiency	.25	.19	.07	.09	.01	.01	.18	.26	.09		.08	.09	.08	.07	.17	.13	.02	.25	.02	-.04	.10
11 Non-Delinquency	.19	.01	-.17	-.15	.12	.25	.03	-.01	.23	.03		.17	.14	-.04	.27	.25	.21	.12	-.01	.09	.13
12 Optimism	.16	.26	.02	.08	.02	.16	.03	.15	.19	.11	.12		.00	.04	.13	.06	.11	.10	.16	.04	.11
13 Order	.18	-.05	-.09	-.07	.06	.11	--	.10	.02	.05	.13	.02		.10	--	.15	.10	.16	-.05	--	.05
14 Physical Conditioning	.17	.06	.25	.09	.01	-.03	.12	.18	-.08	.06	-.05	.06	.05		.03	-.04	.01	.06	.06	.02	-.02
15 Responsibility	.30	.11	--	-.05	--	--	.14	.15	.15	.16	.22	.15	--	.05		.24	--	--	.03	.04	.08
16 Self-Control	.20	.06	--	-.06	--	.06	.09	-.04	.17	.14	.22	.05	.15	-.06	.22		.03	--	-.05	.07	.03
17 Selflessness	.13	-.01	-.04	-.10	.04	.27	--	.11	.13	.01	.17	.09	.11	-.02	--	.02		.04	.08	--	.30
18 Situational Awareness	.19	.14	.10	.04	.07	.00	--	.11	.10	.25	.12	.08	.16	.06	--	--	.05		--	--	--
19 Sociability	.06	.09	--	.34	--	.11	.09	.18	.01	.03	-.04	.15	-.05	.08	.03	-.04	.05	--		.25	.12
20 Team Orientation	.06	.05	--	.13	--	--	.03	.10	.06	-.05	.04	.06	--	.06	.02	.05	--	--	.23		.08
21 Tolerance	.10	.04	--	.01	--	.17	.03	.09	.15	.10	.08	.11	.05	-.04	.08	.05	.32	--	.10	.08	

Note. Results are limited to the Applicant Sample (non-prior service, Education Tier 1 and 2, AFQT Category IV and above) with valid TAPAS score data. Correlations below the diagonal are for Education Tier 1 applicants, $n = 93,806 - 395,127$. Correlations above the diagonal are for Education Tier 2 applicants, $n = 5,795 - 22,367$. Not all TAPAS scales were administered in every version; missing correlations indicate that the scales were not administered on the same version. For Tier 1 applicants, the correlation between the Can-Do and Will-Do predictor composites is $r = .16$; the correlation between the Can-Do and Adaptation predictor composites is $r = .40$; the correlation between the Will-Do and Adaptation predictor composites is $r = .49$. For Tier 2 applicants, the correlation between Can-Do and Will-Do predictor composites is $r = .15$; the correlation between the Can-Do and Adaptation predictor composites is $r = .39$; the correlation between the Will-Do and Adaptation predictor composites is $r = .47$. Correlations in bold are statistically significant, $p < .01$ (two-tailed).

Table A.11. Correlations among AFQT, ASVAB Subtests, and AA Composite Scores in the TOPS Applicant Sample

Composite/Subtest	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1. AFQT																			
<u>ASVAB Subtests</u>																			
2. Arithmetic Reasoning (AR)	.82																		
3. Assembling Objects (AO)	.44	.47																	
4. Auto & Shop Info (AS)	.37	.31	.27																
5. Electronics Information (EI)	.59	.47	.35	.68															
6. General Science (GS)	.73	.53	.35	.50	.67														
7. Math Knowledge (MK)	.71	.68	.37	.06	.28	.42													
8. Mech Comprehension (MC)	.65	.59	.52	.62	.68	.66	.40												
9. Para Comprehension (PC)	.79	.54	.34	.35	.52	.64	.40	.54											
10. Word Knowledge (WK)	.80	.47	.28	.41	.59	.72	.33	.55	.69										
<u>Aptitude Area (AA) Composites</u>																			
11. Clerical (CL)	.96	.90	.49	.43	.63	.72	.77	.72	.74	.73									
12. Combat (CO)	.88	.79	.51	.67	.78	.80	.66	.85	.69	.71	.94								
13. Electronics (EL)	.90	.80	.50	.67	.80	.80	.63	.83	.73	.75	.95	1.00							
14. Field Artillery (FA)	.89	.81	.51	.65	.76	.79	.67	.85	.70	.70	.95	1.00	.99						
15. General Maintenance (GM)	.85	.78	.49	.73	.82	.80	.59	.85	.67	.70	.91	.99	.99	.99					
16. General Technical (GT)	.96	.88	.45	.42	.61	.73	.61	.67	.80	.81	.97	.88	.91	.90	.87				
17. Mech Maintenance (MM)	.74	.67	.46	.86	.84	.74	.44	.85	.61	.65	.81	.95	.95	.94	.97	.77			
18. Operators & Food (OF)	.86	.79	.50	.73	.80	.79	.58	.86	.70	.72	.92	.99	.99	.99	1.00	.89	.97		
19. Signal Communications (SC)	.92	.82	.50	.60	.78	.78	.70	.81	.73	.74	.97	.99	.99	.99	.98	.92	.92	.98	
20. Skilled Technical (ST)	.94	.83	.50	.59	.75	.80	.67	.82	.76	.77	.97	.99	.99	.99	.97	.94	.91	.98	1.00

Note. Results are limited to the Applicant Sample (non-prior service, Education Tier 1 and 2, AFQT Category IV and above) with valid TAPAS score data, $n = 400,231 - 417,464$. All correlations are statistically significant, $p < .01$ (two-tailed).

Table A.12. Correlations among AFQT, ASVAB Subtests, and AA Composite Scores by Education Tier

Composite/Subtest	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. AFQT		.77	.38	.29	.52	.67	.64	.57	.74	.76	.95	.83	.86	.84	.79	.94	.66	.81	.89	.91
<u>ASVAB Subtests</u>																				
2. Arithmetic Reasoning (AR)	.82		.40	.24	.37	.43	.61	.50	.44	.36	.87	.72	.73	.75	.71	.85	.58	.72	.76	.77
3. Assembling Objects (AO)	.44	.47		.24	.31	.29	.30	.47	.27	.22	.44	.46	.45	.47	.45	.39	.41	.45	.46	.46
4. Auto & Shop Info (AS)	.38	.32	.27		.65	.45	-.02	.59	.27	.33	.38	.68	.67	.65	.74	.34	.87	.73	.59	.57
5. Electronics Information (EI)	.60	.47	.35	.68		.63	.17	.63	.44	.54	.58	.76	.79	.74	.80	.54	.81	.78	.76	.72
6. General Science (GS)	.73	.54	.35	.51	.68		.30	.60	.55	.67	.66	.75	.75	.74	.75	.66	.68	.74	.72	.76
7. Math Knowledge (MK)	.71	.69	.38	.07	.29	.42		.29	.29	.20	.71	.56	.53	.58	.49	.52	.32	.47	.61	.58
8. Mech Comprehension (MC)	.65	.59	.52	.62	.68	.66	.41		.44	.46	.66	.83	.80	.83	.82	.59	.82	.83	.77	.79
9. Para Comprehension (PC)	.79	.54	.34	.36	.53	.64	.41	.54		.61	.68	.61	.65	.61	.58	.74	.51	.61	.66	.69
10. Word Knowledge (WK)	.81	.47	.28	.42	.59	.72	.34	.55	.69		.67	.63	.68	.63	.62	.76	.56	.64	.68	.72
<u>Aptitude Area (AA) Composites</u>																				
11. Clerical (CL)	.96	.90	.49	.44	.64	.73	.77	.72	.75	.74		.91	.92	.93	.88	.96	.75	.89	.96	.96
12. Combat (CO)	.88	.79	.51	.68	.79	.80	.67	.85	.70	.71	.94		.99	1.00	.99	.83	.94	.99	.99	.98
13. Electronics (EL)	.90	.80	.50	.67	.81	.80	.64	.83	.73	.75	.95	1.00		.99	.99	.87	.94	.99	.99	.99
14. Field Artillery (FA)	.89	.81	.52	.65	.77	.79	.68	.86	.70	.71	.95	1.00	.99		.99	.85	.93	.99	.99	.99
15. General Maintenance (GM)	.85	.78	.50	.73	.82	.81	.60	.85	.68	.70	.91	.99	.99	.99		.81	.97	1.00	.97	.96
16. General Technical (GT)	.96	.88	.45	.42	.62	.73	.62	.68	.80	.81	.97	.88	.91	.90	.87		.69	.84	.89	.92
17. Mech Maintenance (MM)	.75	.67	.46	.86	.84	.74	.45	.85	.62	.65	.81	.95	.95	.94	.97	.78		.97	.90	.88
18. Operators & Food (OF)	.87	.79	.50	.73	.80	.79	.59	.86	.70	.72	.92	.99	.99	.99	1.00	.89	.97		.97	.97
19. Signal Communications (SC)	.92	.82	.50	.60	.78	.78	.70	.81	.73	.75	.97	.99	1.00	.99	.98	.93	.92	.98		.99
20. Skilled Technical (ST)	.94	.83	.50	.59	.75	.81	.68	.82	.76	.78	.97	.99	.99	.99	.97	.95	.91	.98	1.00	

Note. Results are limited to the Applicant Sample (non-prior service, Education Tier 1 and 2, AFQT Category IV and above) with valid TAPAS score data. Correlations below the diagonal are for Education Tier 1 applicants, $n = 378,732 - 395,103$. Correlations above the diagonal are for Education Tier 2 applicants, $n = 21,499 - 22,361$. All correlations are statistically significant, $p < .01$ (two-tailed).

Table A.13. Correlations among AFQT, ASVAB Subtests, AA Composite Scores with TAPAS Composites and TAPAS Scales in the TOPS Applicant Sample

	ASVAB Subtests									
	AFQT	AR	AO	AS	EI	GS	MK	MC	PC	WK
<i>Individual TAPAS Scales</i>										
Achievement	.05	.07	.01	.09	.05	.01	.02	.03	.06	.03
Adjustment	.11	.08	.05	.12	.12	.12	.03	.12	.10	.11
Adventure Seeking	.10	.10	.10	.20	.16	.14	.03	.20	.08	.08
Attention Seeking	.06	.07	.01	.00	.01	.02	.03	.02	.06	.03
Commitment to Serve	-.14	-.11	-.07	.00	-.06	-.11	-.12	-.08	-.10	-.11
Cooperation	-.09	-.08	-.06	-.12	-.10	-.09	-.03	-.11	-.08	-.09
Courage	.06	.05	.01	.14	.10	.08	-.03	.09	.08	.08
Dominance	.07	.07	.00	.01	.01	.02	.05	.02	.07	.03
Even Tempered	.07	.05	.04	.04	.07	.07	.02	.06	.07	.08
Intellectual Efficiency	.34	.31	.16	.13	.22	.26	.25	.23	.28	.27
Non-Delinquency	-.04	-.04	-.04	-.04	-.04	-.05	-.02	-.07	-.03	-.04
Optimism	.05	.04	.00	.02	.02	.03	.02	.02	.06	.03
Order	-.18	-.12	-.08	-.12	-.15	-.20	-.09	-.17	-.17	-.18
Physical Conditioning	.05	.07	.01	.02	.00	.03	.07	.04	.03	.00
Responsibility	.14	.09	.04	.10	.10	.10	.05	.09	.15	.14
Self-Control	.01	.02	.01	.02	.03	-.01	-.02	.00	.02	.01
Selflessness	-.09	-.09	-.08	-.16	-.14	-.11	-.03	-.15	-.08	-.09
Situational Awareness	.01	.02	.03	.10	.07	.02	-.04	.06	.02	.02
Sociability	-.09	-.07	-.08	-.05	-.08	-.09	-.07	-.09	-.08	-.09
Team Orientation	-.10	-.07	-.05	-.04	-.07	-.10	-.06	-.07	-.09	-.10
Tolerance	.01	-.04	-.03	-.13	-.05	-.01	-.01	-.08	.02	.04
<i>TAPAS Composites</i>										
Can-Do	.43	.38	.22	.24	.33	.36	.29	.34	.36	.36
Will-Do	.10	.12	.02	.09	.05	.05	.08	.07	.09	.05
Adaptation	.19	.17	.09	.11	.13	.18	.13	.18	.17	.16

Note. Results are limited to the Applicant Sample (non-prior service, Education Tier 1 and 2, AFQT Category IV and above) with valid TAPAS score data, $n = 93,819 - 417,494$. CL = Clerical, CO = Combat, EL = Electronics, FA = Field Artillery, GM = General Maintenance, GT = General Technical, MM = Mechanical Maintenance, OF = Operators & Food, SC = Signal Communication, ST = Skilled Technical, CL = Clerical, CO = Combat, EL = Electronics, FA = Field Artillery, GM = General Maintenance, GT = General Technical, MM = Mech Maintenance, OF = Operators & Food, SC = Signal Communications, ST = Skilled Technical. Correlations in bold are statistically significant, $p < .01$ (two-tailed).

Table A.13. (Continued)

	<i>Aptitude Area (AA) Composites</i>									
	CL	CO	EL	FA	GM	GT	MM	OF	SC	ST
<i>Individual TAPAS Scales</i>										
Achievement	.06	.06	.07	.06	.07	.06	.08	.07	.06	.06
Adjustment	.11	.13	.13	.13	.13	.11	.14	.14	.12	.13
Adventure Seeking	.12	.17	.17	.17	.18	.11	.20	.18	.15	.15
Attention Seeking	.05	.04	.04	.04	.03	.06	.02	.04	.04	.04
Commitment to Serve	-.14	-.11	-.11	-.12	-.10	-.13	-.08	-.11	-.12	-.13
Cooperation	-.09	-.11	-.11	-.11	-.12	-.10	-.13	-.12	-.11	-.11
Courage	.06	.09	.09	.09	.10	.07	.12	.10	.08	.08
Dominance	.07	.05	.05	.05	.05	.07	.04	.05	.05	.06
Even Tempered	.07	.07	.07	.07	.07	.08	.06	.07	.07	.07
Intellectual Efficiency	.35	.31	.32	.32	.31	.35	.27	.31	.33	.33
Non-Delinquency	-.05	-.06	-.06	-.06	-.06	-.05	-.06	-.06	-.05	-.06
Optimism	.04	.04	.04	.04	.04	.05	.03	.04	.04	.04
Order	-.17	-.19	-.19	-.19	-.19	-.18	-.18	-.19	-.18	-.19
Physical Conditioning	.06	.05	.05	.05	.05	.05	.04	.05	.05	.05
Responsibility	.13	.13	.14	.13	.13	.14	.13	.13	.13	.14
Self-Control	.01	.01	.01	.01	.01	.02	.02	.02	.01	.01
Selflessness	-.11	-.15	-.14	-.14	-.15	-.11	-.17	-.15	-.13	-.13
Situational Awareness	.02	.05	.05	.04	.06	.03	.08	.06	.04	.04
Sociability	-.09	-.10	-.10	-.10	-.09	-.09	-.09	-.09	-.10	-.10
Team Orientation	-.10	-.09	-.10	-.09	-.09	-.10	-.08	-.09	-.10	-.10
Tolerance	-.02	-.06	-.06	-.06	-.07	.00	-.10	-.07	-.05	-.04
<i>TAPAS Composites</i>										
Can-Do	.44	.43	.44	.43	.42	.44	.39	.43	.44	.44
Will-Do	.11	.10	.10	.11	.10	.11	.10	.11	.10	.10
Adaptation	.20	.20	.20	.20	.20	.19	.18	.20	.20	.20

Note. Results are limited to the Applicant Sample (non-prior service, Education Tier 1 and 2, AFQT Category IV and above) with valid TAPAS score data, $n = 93,819 - 417,494$. CL = Clerical, CO = Combat, EL = Electronics, FA = Field Artillery, GM = General Maintenance, GT = General Technical, MM = Mechanical Maintenance, OF = Operators & Food, SC = Signal Communication, ST = Skilled Technical, CL = Clerical, CO = Combat, EL = Electronics, FA = Field Artillery, GM = General Maintenance, GT = General Technical, MM = Mech Maintenance, OF = Operators & Food, SC = Signal Communications, ST = Skilled Technical. Correlations in bold are statistically significant, $p < .01$ (two-tailed).

APPENDIX B

CORRELATIONS AMONG CRITERION MEASURES IN THE IMT AND IN-UNIT VALIDATION SAMPLES

Table B.1. Correlations among the Performance Rating Scales (PRS) in the IMT Validation Sample

Domain/PRS	1	2	3	4	5	6
<i>Army-Wide</i>						
1. Adjustment to the Army						
2. Effort & Personal Discipline	.74					
3. MOS Qualification Knowledge	.72	.67				
4. Physical Fitness & Bearing	.67	.67	.59			
5. Working with Others	.72	.76	.67	.63		
6. Overall Performance	.62	.61	.60	.56	.57	
<i>MOS-Specific</i>						
7. 11B/C/X + 18X	.68	.64	.70	.61	.67	.58
8. 19K	.71	.69	.79	.68	.59	.70
9. 31B	.69	.64	.72	.54	.68	.58
10. 42A	.64	.63	.70	.39	.64	.70
11. 68W	.73	.67	.74	.60	.70	.45
12. 88M	.67	.64	.65	.64	.64	.60
13. All MOS Combined ^a	.70	.66	.72	.59	.68	.56

Note. Army-wide PRS: $n = 6,236$ - $6,737$. MOS-specific PRS: 11B, $n = 1,611$ - $1,614$; 19K, $n = 205$; 31B, $n = 1,074$ - $1,079$; 42A, $n = 340$; 68W, $n = 714$ - 736 ; 88M, $n = 108$ - 109 . All MOS Combined, $n = 4,110$ - $4,135$. Ratings on PRS range from 1 and 5. PRS ratings from supervisors with a familiarity rating of 1 ("I have had little opportunity to observe this Soldier") were excluded from analyses. Results based on fewer than 100 cases are not reported. All correlations are statistically significant ($p < .05$, one-tailed).

^a Includes 11B/C/X + 18X, 19K, 31B, 42A, 68W, 88M, and 91B.

Table B.2. Correlations among the Performance Rating Scales (PRS) in the In-Unit Validation Sample

PRS	1	2	3	4	5	6
1. Can Do ^a						
2. Effort & Personal Discipline ^a	.80					
3. Physical Fitness & Bearing	.56	.59				
4. Self-Management ^a	.76	.76	.60			
5. Working with Others ^a	.79	.78	.54	.74		
6. Adjustment to Army Life	.73	.77	.68	.73	.70	
7. Overall Leadership Potential	.70	.70	.62	.69	.64	.78

Note. Army-wide PRS, $n = 1,111$ - $1,367$. Ratings on PRS range from 1 and 7. PRS ratings from supervisors with a familiarity rating of 1 ("I have had little opportunity to observe this Soldier") were excluded from analyses. All correlations are statistically significant ($p < .05$, one-tailed).

^a Ratings composite comprises two or more Army-wide PRS.

Table B.3. Correlations among the Army Life Questionnaire (ALQ) Scales in the IMT and In-Unit Validation Samples

Scale	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Affective Commitment		.78	.42	.54		.63	.56	-.57		-.14	-.13	.04	
2. Army Fit	.84		.46	.55		.63	.58	-.63		-.18	-.14	.07	
3. MOS Fit	.46	.48		.55		.30	.23	-.31		-.11	-.10	.04	
4. MOS Satisfaction						.41	.32	-.38		-.12	-.08	.02	
5. Normative Commitment	.69	.72	.41										
6. Army Career Intentions	.55	.52	.24		.42		.83	-.48		-.13	-.12	.06	
7. Army Reenlistment Intentions	.53	.53	.27		.45	.85		-.46		-.10	-.09	.07	
8. Attrition Cognition	-.62	-.68	-.40		-.73	-.46	-.49			.23	.18	-.15	
9. Army Life Adjustment	.45	.61	.35		.46	.36	.39	-.52					
10. Disciplinary Incidents (#)	-.07	-.10	-.08		-.07	-.04	-.06	.10	-.16		.81	-.08	
11. Disciplinary Incidents (Y/N)	-.06	-.08	-.07		-.05	-.05	-.06	.09	-.17	.86		-.10	
12. APFT Score	.03	.08	.06		.06	.03	.04	-.11	.22	-.13			
13. Training Achievement (#)	.05	.06	.05		-.01	.08	.06	-.03	.12	-.07	.23		
14. Training Failures (#) ^a	-.04	-.04	-.05		-.04	-.02	-.02	.05	-.09	.14	-.08		-.06

Note. Correlations below the diagonal reflect the IMT ALQ, $n = 19,803$ - $22,171$. Correlations above the diagonal reflect the in-unit ALQ, $n = 1,888$ - $1,926$. Missing values reflect the scales that were not administered in either IMT or In-Unit. Correlations in bold are statistically significant, $p < .01$ (two-tailed).

^a Training Failures is based on the total number of affirmative responses to whether a Soldier restarted from BCT or OSUT or whether a Soldier repeated a block or module at AIT or OSUT.

Table B.4. Correlations between the Army Life Questionnaire (ALQ) Scales and Job Knowledge Tests (JKTs) in the IMT and In-Unit Validation Samples

Setting/ Scale	IMT/In-Unit JKTs								
	WTBD	All MOS Combined ^a	11B	19K	31B	42A	68W	88M	91B
<i>IMT</i>									
Affective Commitment	.10	.04	.11	.20	.03	.16	.04	.05	.18
Army Fit	.15	.08	.17	.22	.05	.17	.10	.09	.20
MOS Fit	.13	.12	.13	.15	.04	.00	.16	.01	.29
Normative Commitment	.22	.16	.23	.24	.12	.18	.17	.17	.22
Army Career Intentions	-.02	-.04	.01	.01	.00	-.01	-.02	-.02	.03
Army Reenlistment Intentions	.04	.02	.06	.07	.02	.07	.03	.04	.08
Attrition Cognitions	-.19	-.12	-.20	-.16	-.10	-.16	-.15	-.12	-.17
Army Life Adjustment	.13	.09	.13	.05	.10	.16	.13	.10	.15
Disciplinary Incidents (#)	-.04	.00	-.02	.02	-.06	-.06	-.06	.00	.00
Disciplinary Incidents (Y/N)	.02	.00	.03	-.05	-.02	-.04	.01	.01	-.02
APFT Score	.07	.05	.05	-.05	.00	.01	.00	.00	-.01
Training Achievement (#)	-.10	-.12	-.13	-.14	-.04	-.09	.01	-.13	-.17
Training Failure (#) ^b	-.02	-.01	-.02	-.09	-.08	-.08	-.08	.01	.01
<i>In-Unit</i>									
Affective Commitment	.07	.00	.05	--	--	--	--	--	--
Army Fit	.10	.07	.09	--	--	--	--	--	--
MOS Fit	.10	.07	.08	--	--	--	--	--	--
MOS Satisfaction	-.06	-.03	.01	--	--	--	--	--	--
Army Career Intentions	.04	-.03	.01	--	--	--	--	--	--
Army Reenlistment Intentions	.05	.01	.05	--	--	--	--	--	--
Attrition Cognitions	-.14	-.09	-.09	--	--	--	--	--	--
Disciplinary Incidents (#)	-.07	-.05	-.05	--	--	--	--	--	--
Disciplinary Incidents (Y/N)	-.08	-.05	-.08	--	--	--	--	--	--
APFT Score	.02	-.06	-.07	--	--	--	--	--	--

Note. WTBD = Warrior Tasks and Battle Drills. IMT: All MOS Combined, $n = 15,797$ - $17,136$; 11B, $n = 6,683$ - $6,727$; 19K, $n = 701$ - 715 ; 31B, $n = 2,557$ - $2,947$; 42A, $n = 1,016$ - $1,023$; 68W, $n = 2,705$ - $3,212$; 88M, $n = 1,811$ - $2,148$; 91B, $n = 259$ - 364 ; WTBD, $n = 19,292$ - $20,921$. In-Unit: All MOS Combined, $n = 748$ - 757 ; 11B, $n = 397$ - 400 ; WTBD, $n = 1,833$ - $1,869$. Results based on fewer than 100 cases are not reported. Correlations in bold are statistically significant ($p < .05$, two-tailed).

^a Includes 11B/C/X + 18X, 19K, 31B, 42A, 68W, 88M, and 91B.

^b Training Failure is based on the total number of affirmative responses to whether a Soldier restarted from BCT or OSUT or whether a Soldier repeated a block or module at AIT or OSUT (formerly labeled Training Failure).

Table B.5. Correlations between the Army Life Questionnaire (ALQ) Scales and Performance Rating Scales (PRS) in the IMT Validation Sample

Domain/PRS	AFF COM	Army Fit	MOS Fit	NORM COM	CAR INT	RENL INT	ATT COG	AL ADJ	DIS INC #	DIS INC Y/N	APFT	TRN ACH	TRN FAIL
<i>Army-Wide</i>													
Adjustment to the Army	.06	.08	.06	.05	.03	.04	-.07	.08	-.11	-.11	.16	.10	.00
Effort & Personal Discipline	.05	.08	.06	.05	.02	.03	-.07	.09	-.11	-.10	.15	.08	.03
MOS Qualification Knowledge	.04	.06	.05	.03	.02	.03	-.05	.08	-.10	-.10	.13	.07	.01
Physical Fitness & Bearing	.04	.06	.05	.04	.03	.04	-.07	.11	-.10	-.11	.30	.13	.00
Working with Others	.04	.05	.04	.04	.03	.04	-.05	.06	-.08	-.07	.12	.06	.04
Overall Performance	.05	.08	.06	.06	.03	.04	-.09	.13	-.14	-.14	.21	.14	-.03
<i>MOS-Specific</i>													
All MOS Combined ^a	.07	.09	.05	.06	.05	.06	-.09	.08	-.07	-.07	.10	.06	.02
11B/C/X + 18X	.04	.05	.10	.06	.02	.04	-.08	.04	-.07	-.10	.11	.05	.02
19K	.09	.10	.14	.12	.12	.14	-.15	.16	-.16	-.18	.25	.12	-.16
25B ^b	.00	.06	.25	.01	-.10	-.08	-.01	.13	-.09	-.07	-.01	.06	--
31B	.09	.12	.06	.07	.06	.05	-.11	.15	-.14	-.13	.07	.13	.01
42A	.06	.12	.06	.12	.08	.11	-.07	.19	-.15	-.12	.20	.11	-.29
68W	.08	.10	.03	.08	.06	.07	-.08	.07	-.01	-.02	.11	-.01	.09
88M	.00	.01	-.08	-.05	-.08	-.08	.14	-.13	--	-.04	.18	-.08	.06

Note. AFF COM = Affective Commitment; NORM COM = Normative Commitment; CAR INT = Army Career Intentions; RENL INT = Army Reenlistment Intentions; ATT COG = Attrition Cognitions; AL ADJ = Army Life Adjustment; DIS INC = Disciplinary Incidents (#); APFT = APFT Score; TRN ACH = Training Achievements (#); TRN FAIL = Training Failures (#). Army-wide PRS: $n = 6,118$ - $6,573$. MOS-specific PRS: All MOS Combined $n = 3,630$ - $3,995$; 11B $n = 1,640$ - $1,649$; 19K $n = 207$ - 210 ; 25B $n = 539$ - 545 ; 31B $n = 992$ - $1,115$; 68W $n = 667$ - 855 ; 88M $n = 87$ - 115 . Results based on fewer than 100 cases are not reported. Ratings on IMT PRS range from 1 and 5. PRS ratings from supervisors with a familiarity rating of 1 ("I have had little opportunity to observe this Soldier") were excluded from analyses. Correlations in bold are statistically significant ($p < .05$, two-tailed).

^a Includes 11B/C/X + 18X, 19K, 31B, 42A, 68W, 88M, and 91B.

^b No data were available for Training Restarts for 25B Soldiers.

Table B.6. Correlations between the Army Life Questionnaire (ALQ) Scales and Performance Rating Scales (PRS) in the In-Unit Validation Sample

PRS	In-Unit ALQ Scale									APFT
	AFF COM	Army Fit	MOS Fit	MOS SAT	CAR INT	RENL INT	ATT COG	DIS INC #	DIS INC (Y/N)	
Can Do ^a	.07	.06	.04	-.01	.06	.08	-.12	-.23	-.22	.16
Effort & Personal Discipline ^a	.10	.10	.05	.03	.05	.07	-.13	-.29	-.25	.15
Physical Fitness & Bearing	.05	.08	.01	.02	.03	.03	-.16	-.24	-.23	.38
Self-Management ^a	.08	.08	.05	.02	.07	.08	-.16	-.28	-.23	.15
Working with Others ^a	.06	.07	.05	.01	.05	.05	-.13	-.22	-.18	.12
Adjustment to Army Life	.17	.18	.09	.11	.13	.13	-.25	-.35	-.31	.22
Overall Leadership Potential	.12	.15	.05	.06	.10	.11	-.20	-.28	-.26	.22

Note. AFFCOM = Affective Commitment; MOS SAT = MOS Satisfaction; CAR INT = Army Career Intentions; RENL INT = Army Reenlistment Intentions; ATT COG = Attrition Cognitions; DIS INC = Disciplinary Incidents (#); APFT = APFT Score. Ratings on PRS range from 1 and 7. PRS ratings from supervisors with a familiarity rating of 1 ("I have had little opportunity to observe this Soldier") were excluded from analyses. Army-wide PRS, $n = 1,112-1,347$. Correlations in bold are statistically significant ($p < .05$, two-tailed).

^a Ratings composite comprises two or more Army-wide PRS.

Table B.7. Correlations between the Job Knowledge Tests (JKTs) and Performance Rating Scales (PRS) in the IMT Validation Sample

Domain/PRS	All MOS Combined ^a	11B/C/X + 18X	19K	31B	42A	68W	WTBD
<i>Army-Wide</i>							
Adjustment to the Army	.01	.06	.15	.03	.20	-.04	.06
Effort & Personal Discipline	.05	.06	.17	.05	.27	.00	.07
MOS Qualification Knowledge	.04	.08	.15	.04	.28	-.02	.06
Physical Fitness & Bearing	.02	.08	.12	.03	.09	-.04	.04
Working with Others	.03	.05	.00	.01	.21	-.03	.05
Overall Performance	.06	.11	.15	.05	.23	-.01	.07
<i>MOS-Specific</i>							
All MOS Combined ^a	-.01	.08	.18	-.01	.21	-.04	.08
11B/C/X + 18X	.08	.08					.13
19K	.18		.18				.19
31B	-.01			-.01			.06
42A	.21				.21		.18
68W	-.04					-.04	.03
88M	.07						--
91B	.12						--

Note. WTBD = Warrior Tasks and Battle Drills. Army-wide PRS: All MOS Combined, $n = 5,026-5,462$; 11B, $n = 1,600-1,604$; 19K, $n = 203-207$; 31B, $n = 1,135-1,139$; 42A, $n = 353$; 68W, $n = 1,594-2,017$; WTBD, $n = 5,874-6,348$. Ratings on IMT PRS range from 1 and 5. PRS ratings from supervisors with a familiarity rating of 1 ("I have had little opportunity to observe this Soldier") were excluded from analyses. MOS-specific PRS: All MOS Combined, $n = 181-3,741$; 11B, $n = 1,347$; 19K, $n = 181$; 31B, $n = 1,030$; 42A, $n = 322$; 68W, $n = 744$; WTBD, $n = 48-4,203$. Results based on fewer than 100 cases are not reported. Correlations in bold are statistically significant ($p < .05$, two-tailed).

^a Includes 11B/C/X + 18X, 19K, 31B, 42A, 68W, 88M, and 91B.

Table B.8. Correlations between the Job Knowledge Tests (JKTs) and Performance Rating Scales (PRS) in the In-Unit Validation Sample

PRS	In-Unit JKT		
	All MOS Combined ^a	11B/C/X + 18X	WTBD
Can Do ^b	.12	.14	.13
Effort & Personal Discipline ^b	.14	.11	.14
Physical Fitness & Bearing	.03	.05	.05
Self-Management ^b	.09	.09	.10
Working with Others ^b	.13	.10	.14
Adjustment to Army Life	.11	.09	.13
Overall Leadership Potential	.13	.13	.13

Note. WTBD = Warrior Tasks and Battle Drills. Army-wide PRS: All MOS Combined, $n = 437\text{--}529$; 11B, $n = 222\text{--}278$; WTBD, $n = 1,101\text{--}1,323$. Ratings on PRS range from 1 and 7. PRS ratings from supervisors with a familiarity rating of 1 (“I have had little opportunity to observe this Soldier”) were excluded from analyses. Results based on fewer than 100 cases are not reported. Correlations in bold are statistically significant ($p < .05$, two-tailed).

^a Includes 11B/C/X + 18X, 19K, 31B, 42A, 68W, 88M, and 91B.

^b Ratings composite comprises two or more Army-wide PRS.

Table B.9. Correlations between the Job Knowledge Tests (JKTs) and Administrative Criteria in the IMT Validation Sample

Domain/Measure	IMT JKT								
	All MOS Combined ^a	11B/C/X + 18X	19K	31B	42A	68W	88M	91B	WTBD
<i>Attrition ^b</i>									
6-Month Cumulative	.00	.00	.02	.05	--	--	--	--	-.01
12-Month Cumulative	-.04	-.01	.04	-.02	.02	.02	-.10	-.01	-.03
24-Month Cumulative	-.08	-.08	--	-.08	--	-.02	-.05	-.09	-.08
36-Month Cumulative	-.11	-.08	--	-.26	--	-.06	-.01	-.07	-.09
<i>Training Restarts</i>									
Restarted at Least Once During IMT	.04	-.02	.05	-.04	-.03	.05	-.02	.02	.00
Failed at Least Once During IMT	-.04	.02	-.09	.04	.02	-.05	.02	-.04	.00
Academic or Pejorative Restart	-.01	.02	-.10	.03	.02	.00	.02	-.04	.01
Academic Restart	-.05	.02	-.09	.05	.02	-.05	.02	-.05	-.01
<i>Final AIT School Grades</i>									
Overall Average (Unstandardized)	.13	--	--	--	--	--	--	--	.27
Overall Average (Standardized)	.36	--	--	--	--	--	--	--	.32

Note. WTBD = Warrior Tasks and Battle Drills. Attrition: All MOS Combined, $n = 1,868\text{--}9,226$; 11B, $n = 992\text{--}4,661$; 19K, $n = 206\text{--}354$; 31B, $n = 142\text{--}1,222$; 42A, $n = 148\text{--}334$; 68W, $n = 406\text{--}1,945$; 88M, $n = 218\text{--}561$; 91B, $n = 110\text{--}149$; WTBD, $n = 2,247\text{--}11,386$. Training Restarts: All MOS Combined, $n = 11,994\text{--}17,069$; 11B, $n = 4,674\text{--}6,551$; 19K, $n = 332\text{--}729$; 25B, $n = 2,012\text{--}2,973$; 31B, $n = 560\text{--}1,037$; 68W, $n = 2,229\text{--}3,247$; 88M, $n = 1,921\text{--}2,164$; 91B, $n = 250\text{--}368$; WTBD, $n = 14,706\text{--}20,852$. Final AIT School Grade: All MOS Combined, $n = 113\text{--}114$; WTBD, $n = 313\text{--}315$. Results based on fewer than 100 cases are not reported. Correlations in bold are statistically significant ($p < .05$, two-tailed).

^a Includes 11B/C/X + 18X, 19K, 31B, 42A, 68W, 88M, and 91B.

^b Attrition results include Regular Army Soldiers only.

Table B.10. Correlations between the Army Life Questionnaire (ALQ) and Administrative Criteria in the IMT Validation Sample

Domain/Measure	IMT ALQ Scale											
	AFF COM	Army Fit	MOS Fit	NORM COM	CAR INT	RENL INT	ATT COG	AL ADJ	DIS INC	APFT	TRN ACH	TRN FAIL
<i>Attrition^a</i>												
6-Month Cumulative	-.04	-.05	-.04	-.05	-.04	-.03	.13	-.07	.05	-.09	-.02	.02
12-Month Cumulative	-.04	-.06	-.04	-.07	-.05	-.06	.14	-.08	.04	-.07	-.02	.01
24-Month Cumulative	-.03	-.06	-.06	-.06	-.03	-.05	.10	-.09	.06	-.09	-.03	.02
36-Month Cumulative	-.05	-.08	-.06	-.09	-.05	-.06	.11	-.07	.06	-.08	-.06	.00
<i>Training Restarts</i>												
Restarted at Least Once During IMT	-.01	-.01	-.01	.00	.00	.00	.01	-.01	.05	-.04	.03	.33
Failed at Least Once During IMT	.02	.02	.02	.01	.01	.00	-.04	.03	-.08	.08	-.02	-.36
Academic or Pejorative Failure	.01	.03	.04	.02	.00	.00	-.05	.05	-.09	.08	.03	-.40
Academic Failure	.01	.01	.01	.00	.00	.00	-.02	.02	-.07	.06	-.02	-.36
<i>Final AIT School Grades</i>												
Overall Average (Unstandardized)	-.05	-.01	.18	.01	-.07	-.08	.04	-.05	.03	-.12	.02	-.05
Overall Average (Standardized)	-.09	-.04	.14	-.02	-.12	-.08	.03	-.01	.00	-.14	-.09	.01

Note. AFF COM = Affective Commitment; NORM COM = Normative Commitment; CAR INT = Army Career Intentions; RENL INT = Army Reenlistment Intentions; ATT COG = Attrition Cognitions; AL ADJ = Army Life Adjustment; DIS INC = Disciplinary Incidents (# of); APFT = APFT Score; TRN ACH = Training Achievements (# of); TRN FAIL = Training Failure (# of). Attrition: $n = 1,587-11,925$. Training Restarts: $n = 14,169-21,806$. Final AIT School Grade: $n = 152-331$. Correlations in bold are statistically significant ($p < .05$, two-tailed).

^a Attrition results include Regular Army Soldiers only.

Table B.11. Correlations between the Performance Rating Scales (PRS) and Attrition in the IMT Validation Sample

Domain/PRS	Attrition ^a			
	6- Month	12-Month	24-Month	36-Month
<i>Army-Wide</i>				
Adjustment to the Army	-.05	-.05	-.05	-.04
Effort & Discipline	-.05	-.03	-.03	-.06
MOS Qualification Knowledge & Skill	-.04	-.02	-.04	.01
Physical Fitness & Bearing	-.06	-.05	-.07	-.10
Working with Others	-.04	-.04	-.04	-.02
Overall Performance	-.07	-.07	-.09	-.10
<i>MOS-Specific</i>				
All MOS Combined ^b	.00	-.01	-.06	-.07
11B/C/X and 18X	.04	.00	-.06	-.12
19K	--	--	--	--
25B	--	-.16	--	--
31B	-.08	-.06	-.14	--
42A	--	--	--	--
68W	--	.03	-.07	-.05

Note. Sample is limited to Army-wide PRS: $N = 685\text{--}4,015$. MOS-specific PRS: All MOS Combined, $n = 576\text{--}2,653$; 11B, $n = 292\text{--}1,273$; 19K, $n = 173$; 25B, $n = 262\text{--}377$; 31B, $n = 315\text{--}476$; 42A, $n = 104$; 68W, $n = 145\text{--}544$. Ratings on IMT PRS range from 1 and 5. PRS ratings from supervisors with a familiarity rating of 1 (“I have had little opportunity to observe this Soldier”) were excluded from analyses. Results based on fewer than 100 cases are not reported. Correlations in bold are statistically significant ($p < .05$, two-tailed).

^a Attrition results include Regular Army Soldiers only.

^b Includes 11B/C/X + 18X, 19K, 31B, 42A, 68W, 88M, and 91B.

Table B.12. Correlations between the Performance Rating Scales (PRS) and Administrative Criteria in the IMT Validation Sample

Domain/PRS	Training Restarts			
	IMT Restart	IMT Failure	PEJ Failure	ACAD Failure
<i>Army-Wide</i>				
Adjustment to the Army	-.02	.04	.05	.03
Effort & Discipline	.01	.00	.02	.00
MOS Qualification Knowledge & Skill	.01	.01	.01	.00
Physical Fitness & Bearing	.01	.02	.02	.00
Working with Others	.00	.00	.01	.00
Overall Performance	-.01	.03	.05	.02
<i>MOS-Specific</i>				
All MOS Combined ^a	.01	.00	.00	-.01
11B/C/X + 18X	.03	-.04	-.04	-.04
19K	--	.02	.02	.02
25B	-.14	.20	.20	.19
31B	-.01	.02	.03	.00
42A	-.10	.29	.25	.29
68W	.03	-.01	-.01	-.01
88M	.29	-.21	.06	-.21

Note. IMT Restart = Restarted at Least Once During IMT; IMT Failure = Failed at Least Once During IMT; PEJ Failure = Failed at Least Once for Academic or Other Pejorative Reason; ACAD Failure = Failed at Least Once for Academic Reasons. Army-wide PRS, $n = 4,542\text{--}6,735$. MOS-specific PRS: All MOS Combined, $n = 3,299\text{--}4,455$; 11B, $n = 1,410\text{--}1,716$; 19K, $n = 173\text{--}214$; 25B, $n = 457\text{--}646$; 31B, $n = 765\text{--}1,130$; 42A, $n = 152\text{--}344$; 68W, $n = 632\text{--}879$; 88M, $n = 114\text{--}119$. Ratings on IMT PRS range from 1 and 5. PRS ratings from supervisors with a familiarity rating of 1 (“I have had little opportunity to observe this Soldier”) were excluded from analyses. Results based on fewer than 100 cases are not reported. Correlations in bold are statistically significant ($p < .05$, two-tailed).

^a Includes 11B/C/X + 18X, 19K, 31B, 42A, 68W, 88M, and 91B.

Table B.13. Correlations among the Criterion Composites in the IMT and In-Unit (IU) Validation Samples

	1	2	3	4	5	6	7	8	9	10	11
1 IMT: Overall Performance											
2 IMT: Physical Fitness	.20										
3 IMT: Commitment & Fit	.07	.07									
4 IMT: Retention Cognitions	.02	.01	.36								
5 IMT: Knowledge & Skill	.08	.05	.17	-.04							
6 IU: Overall Performance	--	.13	.00	-.13	.06						
7 IU: Physical Fitness	--	.48	.03	-.03	.05	.24					
8 IU: Commitment & Fit	--	-.05	.37	.19	.09	.09	.07				
9 IU: Retention Cognitions	--	-.01	.18	.47	-.05	.01	.03	.46			
10 IU: Knowledge & Skill	--	.02	.19	-.02	.44	.14	.01	.10	-.01		
11 Can Do Performance	.08	.05	.18	-.05	1.00	.08	.04	.07	-.07	.43	
12 Will Do Performance	.49	.62	.42	.26	.10	--	--	--	--	--	.10

Note. $n = 169-20,439$. Results based on fewer than 100 cases are not reported. Correlations in bold are statistically significant ($p < .05$, two-tailed).

APPENDIX C

CRITERION PSYCHOMETRIC PROPERTIES IN THE FULL IMT AND IN-UNIT SAMPLES

Table C.1. Descriptive Statistics for the Job Knowledge Tests (JKTs) in the Full IMT and In-Unit Samples

Domain/Setting/JKT	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>	<i>r</i> _{WTBD}
IMT						
<i>MOS-Specific</i>						
11B/C/X + 18X	16,433	60.86	10.28	20.93	88.37	.56
19K	1,105	61.94	11.69	20.29	86.15	.51
31B	7,486	67.93	8.68	33.33	93.20	.50
42A	1,453	55.19	12.30	16.67	85.19	.51
68W	9,525	73.15	10.28	25.00	96.74	.50
88M	5,887	63.76	10.57	30.56	94.44	.54
91B	1,310	57.71	13.26	23.71	90.72	.46
All MOS Combined ^a	43,199	64.93	11.56	16.67	96.74	.55
WTBD (Army-Wide)	53,270	64.50	12.48	6.45	100.00	--
In-Unit						
<i>MOS-Specific</i>						
11B/C/X + 18X	1,040	62.75	8.86	26.76	84.51	.55
19K	102	76.44	10.93	37.25	90.74	.44
31B	106	62.66	11.15	33.64	85.22	.59
68W	228	71.33	7.90	36.79	90.57	.53
88M	245	64.14	9.50	39.60	87.23	.57
91B	250	63.11	11.10	35.09	85.26	.33
All MOS Combined ^a	2,047	64.24	10.28	26.76	90.74	.49
WTBD (Army-Wide)	5,320	64.78	11.81	15.38	100.00	--

Note. *M*, *SD*, *Min*, and *Max* are based on percent correct. WTBD = Warrior Tasks and Battle Drills. *r*_{WTBD} = correlation with WTBD JKT scores. Results based on fewer than 100 cases are not reported. All correlations are statistically significant ($p < .05$, one-tailed).

^a Includes 11B/C/X + 18X, 19K, 31B, 42A, 68W, 88M, and 91B.

Table C.2. Descriptive Statistics Estimates for the Performance Rating Scales (PRS) in the Full IMT and In-Unit Samples

Domain/Setting/PRS	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
IMT					
<i>Army-Wide</i>					
Adjustment to the Army	18,191	3.29	1.02	1.00	5.00
Effort & Personal Discipline	18,217	3.11	0.99	1.00	5.00
MOS Qualification	16,467	3.25	0.96	1.00	5.00
Knowledge					
Physical Fitness & Bearing	18,117	3.15	1.01	1.00	5.00
Working with Others	18,161	3.09	1.00	1.00	5.00
Overall Performance	17,933	3.53	0.85	1.00	5.00
<i>MOS-Specific</i>					
11B/C/X + 18X	5,153	3.02	0.82	1.00	5.00
19K	365	3.39	0.66	1.00	5.00
31B	2,706	3.22	0.78	1.00	5.00
42A	471	3.72	0.68	1.80	5.00
68W	3,611	2.77	0.81	1.00	5.00
88M	672	2.87	0.76	1.20	5.00
91B	255	2.97	1.19	1.00	5.00
All MOS Combined ^a	13,233	3.02	0.83	1.00	5.00
In-Unit					
<i>Army-Wide</i>					
Can Do ^b	3,872	4.92	1.27	1.00	7.00
Effort & Personal Discipline ^b	3,868	5.23	1.37	1.00	7.00
Physical Fitness & Bearing	3,857	5.28	1.56	1.00	7.00
Self-Management ^b	3,860	5.34	1.14	1.00	7.00
Working with Others ^b	3,871	5.30	1.21	1.00	7.00
Adjustment to Army Life	3,819	5.43	1.48	1.00	7.00
Overall Leadership Potential	3,795	4.77	1.66	1.00	7.00

Note. Ratings on IMT PRS range from 1 to 5. Ratings on IU PRS range from 1 to 7. PRS ratings from supervisors with a familiarity rating of 1 (“I have had little opportunity to observe this Soldier”) were excluded from analyses.

^a Includes 11B/C/X + 18X, 19K, 31B, 42A, 68W, 88M, and 91B.

^b Ratings composite comprises two or more Army-wide PRS.

Table C.3. Descriptive Statistics Estimates for the Army Life Questionnaire (ALQ) in the Full IMT and In-Unit Samples

Setting/Scale	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
IMT					
Affective Commitment	55,571	3.89	0.68	1.00	5.00
Army Fit	55,571	4.08	0.60	1.00	5.00
MOS Fit	55,571	3.78	0.84	1.00	5.00
Normative Commitment	55,571	4.17	0.70	1.00	5.00
Army Career Intentions	55,571	3.20	1.10	1.00	5.00
Army Reenlistment Intentions	55,571	3.49	0.96	1.00	5.00
Attrition Cognition	55,571	1.52	0.60	1.00	5.00
Army Life Adjustment	55,571	4.09	0.66	1.00	5.00
Disciplinary Incidents (#)	43,329	0.27	0.62	0.00	7.00
Disciplinary Incidents (Y/N)	43,329	0.21	0.41	0.00	1.00
APFT Score	55,002	250.82	31.69	10.00	300.00
Training Achievement (#)	55,534	0.40	0.61	0.00	2.00
Training Restarts (#) ^a	55,571	0.42	0.65	0.00	4.00
In-Unit					
Affective Commitment	5,376	3.57	0.81	1.00	5.00
Army Fit	5,376	3.88	0.71	1.00	5.00
MOS Fit	5,376	3.26	0.94	1.00	5.00
MOS Satisfaction	5,376	3.50	0.92	1.00	5.00
Army Career Intentions	5,376	2.63	1.21	1.00	5.00
Army Reenlistment Intentions	5,376	3.01	1.18	1.00	5.00
Attrition Cognition	5,376	1.71	0.75	1.00	5.00
Disciplinary Incidents (#)	5,375	0.42	0.92	0.00	7.00
Disciplinary Incidents (Y/N)	5,375	0.25	0.43	0.00	1.00
APFT Score	5,269	247.45	33.62	1.00	300.00

^a Training Restarts is based on the total number of affirmative responses to whether a Soldier Restarted from BCT or OSUT or whether a Soldier repeated a block or module at AIT or OSUT (formerly labeled Training Failure).

Table C.4. Correlations among the Performance Rating Scales (PRS) in the Full IMT Sample

Domain/PRS	1	2	3	4	5	6
<i>Army-Wide</i>						
1. Adjustment to the Army						
2. Effort & Personal Discipline	.76					
3. MOS Qualification Knowledge	.72	.68				
4. Physical Fitness & Bearing	.69	.70	.62			
5. Working with Others	.74	.75	.69	.66		
6. Overall Performance	.58	.58	.56	.54	.55	
<i>MOS-Specific</i>						
6. 11B/C/X + 18X	.68	.65	.71	.62	.68	.54
7. 19K	.72	.70	.79	.66	.61	.68
8. 31B	.65	.63	.70	.54	.65	.57
9. 42A	.65	.64	.68	.43	.64	.71
10. 68W	.59	.55	.64	.50	.60	.38
11. 88M	.59	.54	.64	.55	.58	.50
12. 91B	.72	.67	.80	.67	.73	.58
13. All MOS Combined ^a	.66	.63	.70	.58	.66	.51

Note. Army-wide PRS: $n = 16,378$ - $18,188$. MOS-specific PRS: 11B, $n = 4,752$ - $4,758$; 19K, $n = 348$; 31B, $n = 2,548$ - $2,564$; 42A, $n = 466$; 68W, $n = 2,188$ - $2,724$; 88M, $n = 607$ - 624 ; 91B, $n = 227$ - 246 ; All MOS Combined, $n = 11,188$ - $11,730$. Ratings on IMT PRS range from 1 and 5. PRS ratings from supervisors with a familiarity rating of 1 ("I have had little opportunity to observe this Soldier") were excluded from analyses. All correlations are statistically significant ($p < .05$, one-tailed).

^a Includes 11B/C/X + 18X, 19K, 31B, 42A, 68W, 88M, and 91B.

Table C.5. Correlations among Performance Rating Scales (PRS) in the Full In-Unit Sample

Domain/PRS	1	2	3	4	5	6
1. Can Do ^a						
2. Effort & Personal Discipline ^a	.78					
3. Physical Fitness & Bearing	.56	.60				
4. Self-Management ^a	.75	.75	.59			
5. Working with Others ^a	.78	.76	.56	.74		
6. Adjustment to Army Life	.64	.68	.62	.68	.63	
7. Overall Leadership Potential	.69	.70	.61	.68	.64	.67

Note. Army-wide PRS, $n = 3,779$ - $3,872$. Ratings on PRS range from 1 to 7. PRS ratings from supervisors with a familiarity rating of 1 ("I have had little opportunity to observe this Soldier") were excluded from analyses. All correlations are statistically significant ($p < .05$, one-tailed).

^a Ratings composite comprises two or more Army-wide PRS.

Table C.6. Correlations among Army Life Questionnaire (ALQ) Scales in the Full IMT and In-Unit Samples

Scale	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Affective Commitment		.78	.39	.53		.61	.55	-.58		-.15		.06	
2. Army Fit	.84		.43	.55		.60	.57	-.65		-.22		.09	
3. MOS Fit	.48	.48		.56		.26	.21	-.31		-.12		.04	
4. MOS Satisfaction						.37	.30	-.38		-.14		.03	
5. Normative Commitment	.69	.71	.41										
6. Army Career Intentions	.56	.53	.25		.43		.82	-.48		-.11		.05	
7. Army Reenlistment Intentions	.53	.53	.27		.46	.85		-.45		-.09		.05	
8. Attrition Cognition	-.63	-.68	-.41		-.74	-.47	-.50			.22		-.14	
9. Army Life Adjustment	.45	.61	.35		.45	.36	.39	-.52					
10. Disciplinary Incidents (#)	-.08	-.11	-.08		-.08	-.05	-.06	.12	-.17			-.06	
11. Disciplinary Incidents (Y/N)	-.06	-.09	-.07		-.05	-.05	-.05	.09	-.17	.86		-.09	
11. APFT Score	.04	.10	.07		.07	.03	.04	-.11	.24	-.14			
12. Training Achievement (#)	.06	.07	.05		.00	.09	.07	-.04	.13	-.07		.23	
13. Training Failure (#) ^a	-.03	-.05	-.04		-.03	-.01	-.02	.06	-.08	.15		-.07	-.05

Note. Correlations below the diagonal are based on the Full IMT sample, $n = 42,300$ -55,571. Correlations above the diagonal are based on the Full In-Unit sample, $n = 5,268$ -5,376.

Missing values reflect the scales that were not administered in either IMT or In-Unit. Correlations in bold are statistically significant ($p < .05$, two-tailed).

^aTraining Restarts is based on the total number of affirmative responses to whether a Soldier restarted from BCT or OSUT or whether a Soldier repeated a block or module at AIT or OSUT (formerly labeled Training Failure).

APPENDIX D

SUMMARY OF BIVARIATE CORRELATIONS BETWEEN TAPAS SCALES AND SELECTED CRITERIA

Table D.1. Summary of the Bivariate Correlations between AFQT, TAPAS and Selected IMT Criteria for Tier 1 + 2 Soldiers

	Knowledge & Skill	WTBD JKT	Disciplinary Incidents (Y/N)	IMT Restarts	Can Do Performance	Army Fit	Army Life Adjustment	Commitment & Fit	Retention Cognitions
<i>AFQT</i>	.45	.43	-.02	.00	.44	-.04	.06	.00	-.12
<i>Individual TAPAS Scales</i>									
Achievement	.04	.04	-.07	-.01	.04	.12	.14	.12	.06
Adjustment	.08	.06	-.02	.00	.07	.03	.10	.03	.00
Adventure Seeking ^a	.07	.08	-.03	-.01	.07	.02	.08	.04	-.01
Attention Seeking	.01	.02	-.01	.00	.01	.04	.08	.04	-.02
Commitment to Serve ^a	-.05	-.06	.01	.01	-.05	.11	.09	.13	.27
Cooperation	-.02	-.02	.00	.00	-.02	.01	-.01	.01	.00
Courage ^a	.02	.06	-.06	-.01	.03	.13	.15	.13	.05
Dominance	.02	.04	-.05	-.01	.02	.10	.13	.09	.06
Even Tempered	.05	.05	-.01	.01	.05	.02	.03	.01	.02
Intellectual Efficiency	.18	.16	-.02	.01	.17	.03	.11	.03	.01
Non-Delinquency	-.01	.00	-.02	.00	-.01	.06	.01	.04	.03
Optimism	.01	.01	-.03	-.01	.01	.06	.09	.05	.00
Order	-.08	-.08	-.02	.01	-.08	.02	.01	.00	.04
Physical Conditioning	.00	.00	-.10	-.03	-.01	.05	.13	.05	-.03
Responsibility ^a	.05	.04	-.01	.00	.05	.07	.09	.08	.03
Self-Control	.01	.01	-.02	.01	.01	.04	.04	.03	.03
Selflessness	-.04	-.04	.01	.00	-.04	.04	-.01	.04	.04
Situational Awareness ^a	.05	.04	-.02	.01	.04	.06	.09	.05	.06
Sociability	-.10	-.08	.01	.00	-.10	.04	.04	.03	.02
Team Orientation ^a	-.04	-.02	-.03	-.01	-.04	.04	.01	.03	.04
Tolerance	-.02	-.02	.01	.01	-.02	.04	.02	.03	.05
<i>TAPAS Composites</i>									
Can-Do	.24	.22	-.02	.01	.24	.01	.09	.02	-.01
Will-Do	.02	.03	-.11	-.03	.02	.12	.20	.12	.02
Adaptation	.09	.09	-.08	-.02	.08	.02	.10	.04	-.03

Note. WTBD JKT = Warrior Tasks and Battle Drills Job Knowledge Test. PRS = Performance Rating Scales. Unless otherwise noted, sample sizes range from $n = 2,584 - 208,822$. Correlations in bold are statistically significant ($p < .01$, two-tailed).

^a Sample sizes for six TAPAS scales were considerably smaller than the other dimensions because not all scales were administered in every version of the TAPAS ($n = 366 - 43,959$).

Table D.1. (Continued)

	APFT Score	Will Do Performance	PRS: Effort & Discipline	PRS: Adjustment to the Army	PRS: Physical Fitness and Bearing	PRS: Working with Others	PRS: Overall Performance
<i>AFQT</i>	.09	.05	.08	.04	.06	.07	.08
<i>Individual TAPAS Scales</i>							
Achievement	.09	.15	.07	.06	.06	.05	.07
Adjustment	.01	.02	-.03	-.02	-.01	-.02	-.03
Adventure Seeking ^a	.07	.10	-.02	-.06	.00	.00	-.02
Attention Seeking	.08	.07	.00	.00	.03	.01	.01
Commitment to Serve ^a	-.05	.04	.02	.01	-.02	.01	.00
Cooperation	-.02	-.03	-.01	.01	.01	.00	.00
Courage ^a	.03	.18	.01	.02	.02	.02	.01
Dominance	.12	.16	.04	.05	.06	.04	.06
Even Tempered	-.05	-.02	.00	-.01	-.01	.01	.00
Intellectual Efficiency	.04	.06	.02	.01	.01	.02	.03
Non-Delinquency	-.05	-.01	-.01	-.01	-.03	-.02	-.03
Optimism	.04	.10	.02	.03	.02	.02	.03
Order	.03	.03	.00	-.01	.01	.00	.00
Physical Conditioning	.27	.26	.06	.07	.14	.06	.09
Responsibility ^a	.00	.14	.09	.05	.00	.04	.07
Self-Control	-.01	.02	.02	.00	.00	.00	.01
Selflessness	-.01	-.01	.00	.00	.00	-.01	-.01
Situational Awareness ^a	.00	.02	.02	.00	-.02	.00	-.02
Sociability	.03	.06	-.02	.00	.00	.00	-.01
Team Orientation ^a	.02	.03	-.01	-.01	-.04	-.03	-.03
Tolerance	.00	-.01	.00	.00	.00	.00	.00
<i>TAPAS Composites</i>							
Can-Do	.00	.03	.02	.01	.00	.02	.02
Will-Do	.25	.29	.09	.09	.14	.07	.11
Adaptation	.17	.16	.06	.05	.10	.06	.07

Note. WTBD JKT = Warrior Tasks and Battle Drills Job Knowledge Test. PRS = Performance Rating Scales. Unless otherwise noted, sample sizes range from $n = 2,584 - 208,822$. Correlations in bold are statistically significant ($p < .01$, two-tailed).

^a Sample sizes for six TAPAS scales were considerably smaller than the other dimensions because not all scales were administered in every version of the TAPAS ($n = 366 - 43,959$).

Table D.2. Summary of the Bivariate Correlations between AFQT, TAPAS and Selected IMT Criteria for Tier 1 Soldiers

	Knowledge & Skill	WTBD JKT	Disciplinary Incidents (Y/N)	IMT Restarts	Can Do Performance	Army Fit	Army Life Adjustment	Commitment & Fit	Retention Cognitions
<i>AFQT</i>	.45	.43	-.02	.00	.44	-.04	.07	.00	-.12
<i>Individual TAPAS Scales</i>									
Achievement	.05	.04	-.07	-.01	.04	.12	.14	.12	.06
Adjustment	.07	.06	-.02	.00	.07	.03	.10	.03	-.01
Adventure Seeking ^a	.07	.09	-.04	-.01	.07	.02	.08	.05	-.01
Attention Seeking	.01	.02	-.01	.00	.01	.04	.08	.04	-.02
Commitment to Serve ^a	-.06	-.06	.01	.01	-.06	.11	.09	.13	.27
Cooperation	-.02	-.01	.00	.00	-.02	.01	-.01	.01	.00
Courage ^a	.01	.05	-.06	-.01	.03	.13	.15	.13	.05
Dominance	.02	.04	-.05	-.01	.02	.10	.13	.09	.06
Even Tempered	.05	.05	-.02	.01	.05	.02	.03	.01	.02
Intellectual Efficiency	.18	.17	-.03	.01	.18	.03	.12	.03	.01
Non-Delinquency	-.01	-.01	-.03	.00	-.01	.05	.01	.04	.03
Optimism	.01	.01	-.03	-.01	.01	.06	.09	.05	.00
Order	-.08	-.08	-.02	.01	-.08	.02	.01	.00	.04
Physical Conditioning	.00	.01	-.10	-.03	.00	.05	.13	.05	-.03
Responsibility ^a	.05	.04	-.01	.00	.05	.07	.09	.07	.02
Self-Control	.00	.00	-.02	.01	.01	.04	.04	.03	.03
Selflessness	-.05	-.03	.01	.00	-.04	.05	-.01	.04	.04
Situational Awareness ^a	.05	.04	-.02	.01	.04	.06	.08	.05	.06
Sociability	-.10	-.08	.01	.00	-.10	.04	.04	.03	.02
Team Orientation ^a	-.04	-.03	-.03	-.01	-.04	.04	.00	.03	.04
Tolerance	-.01	-.02	.02	.01	-.02	.04	.02	.03	.05
<i>TAPAS Composites</i>									
Can-Do	.25	.22	-.02	.01	.24	.01	.09	.02	-.01
Will-Do	.02	.03	-.11	-.03	.02	.12	.20	.12	.02
Adaptation	.09	.09	-.08	-.02	.09	.03	.10	.04	-.04

Note. WTBD JKT = Warrior Tasks and Battle Drills Job Knowledge Test. PRS = Performance Rating Scales. Unless otherwise noted, sample sizes range from $n = 2,501 - 201,132$. Correlations in bold are statistically significant ($p < .01$, two-tailed).

^a Sample sizes for six TAPAS scales were considerably smaller than the other dimensions because not all scales were administered in every version of the TAPAS ($n = 354 - 42,329$).

Table D.2. (Continued)

	APFT Score	Will Do Performance	PRS: Effort & Discipline	PRS: Adjustment to the Army	PRS: Physical Fitness and Bearing	PRS: Working with Others	PRS: Overall Performance
<i>AFQT</i>	.09	.05	.09	.04	.06	.07	.08
<i>Individual TAPAS Scales</i>							
Achievement	.09	.15	.07	.06	.06	.05	.08
Adjustment	.01	.02	-.03	-.01	-.01	-.02	-.03
Adventure Seeking ^a	.07	.10	-.03	-.07	-.01	-.01	-.03
Attention Seeking	.08	.08	.00	.00	.03	.01	.01
Commitment to Serve ^a	-.05	.04	.02	.02	-.01	.02	.01
Cooperation	-.02	-.03	-.01	.01	.01	.00	-.01
Courage ^a	.04	.18	-.01	.02	.02	.02	.01
Dominance	.11	.16	.04	.05	.06	.03	.06
Even Tempered	-.04	-.02	.00	-.01	-.01	.00	.00
Intellectual Efficiency	.04	.06	.02	.01	.01	.02	.03
Non-Delinquency	-.05	-.01	-.01	-.01	-.03	-.02	-.03
Optimism	.03	.10	.02	.02	.02	.02	.03
Order	.03	.02	.00	.00	.01	.00	.00
Physical Conditioning	.27	.26	.06	.07	.14	.05	.09
Responsibility ^a	.00	.14	.09	.06	-.01	.04	.07
Self-Control	.00	.02	.02	.01	.00	.01	.01
Selflessness	-.01	-.01	.00	.00	.00	-.01	-.01
Situational Awareness ^a	.01	.04	.02	.00	-.01	.00	-.01
Sociability	.03	.06	-.02	.00	.00	.00	-.01
Team Orientation ^a	.01	.01	-.02	-.02	-.05	-.05	-.05
Tolerance	-.01	-.01	.00	.00	.00	.00	.00
<i>TAPAS Composites</i>							
Can-Do	.00	.03	.02	.01	.00	.02	.02
Will-Do	.25	.29	.08	.09	.14	.07	.11
Adaptation	.17	.16	.05	.05	.10	.06	.07

Note. WTBD JKT = Warrior Tasks and Battle Drills Job Knowledge Test. PRS = Performance Rating Scales. Unless otherwise noted, sample sizes range from $n = 2,501 - 201,132$. Correlations in bold are statistically significant ($p < .01$, two-tailed).

^aSample sizes for six TAPAS scales were considerably smaller than the other dimensions because not all scales were administered in every version of the TAPAS ($n = 354 - 42,329$).

Table D.3. Summary of the Bivariate Correlations between AFQT, TAPAS and Selected IMT Criteria for Tier 2 Soldiers

	Knowledge & Skill	WTBD JKT	Disciplinary Incidents (Y/N)	IMT Restarts	Can Do Performance	Army Fit	Army Life Adjustment	Commitment & Fit	Retention Cognitions
<i>AFQT</i>	.38	.35	.02	.02	.37	-.02	.01	.00	-.12
<i>Individual TAPAS Scales</i>									
Achievement	.04	.06	-.08	.00	.04	.12	.17	.12	.03
Adjustment	.09	.05	.05	.00	.08	.00	.09	.00	.01
Adventure Seeking ^a	--	-.03	.14	.01	.01	-.17	.00	-.20	-.15
Attention Seeking	-.01	-.02	.08	.00	-.01	.04	.07	.04	-.06
Commitment to Serve ^a	--	.08	.05	.02	.17	.10	.08	.05	.34
Cooperation	-.04	-.06	-.01	.01	-.06	.00	.00	.01	-.02
Courage ^a	--	.14	.11	.00	.13	.17	.28	.14	.24
Dominance	.04	.04	-.07	.00	.04	.11	.15	.10	.01
Even Tempered	.05	.06	.03	.01	.06	.02	.05	.02	.01
Intellectual Efficiency	.14	.10	.00	.01	.12	.05	.08	.04	-.05
Non-Delinquency	.03	.02	.01	-.01	.01	.08	.06	.06	.00
Optimism	-.04	-.02	.03	.01	-.04	.06	.08	.06	-.01
Order	-.08	-.07	.02	.01	-.07	.04	.04	.03	.04
Physical Conditioning	-.03	-.05	-.06	-.02	-.03	.07	.11	.07	.03
Responsibility ^a	--	.04	.00	.03	.04	.24	.08	.20	.06
Self-Control	.04	.04	-.06	.02	.05	.02	-.02	.00	-.06
Selflessness	-.02	-.08	.03	.00	-.06	-.01	-.02	.01	-.05
Situational Awareness ^a	--	.03	.01	.03	.03	-.04	.11	-.02	.01
Sociability	-.01	-.01	.01	-.01	-.02	.15	.12	.13	.09
Team Orientation ^a	--	-.01	-.02	-.01	-.07	.11	.15	.04	-.03
Tolerance	-.04	-.06	-.04	.02	-.07	.02	.01	-.01	.00
<i>TAPAS Composites</i>									
Can-Do	.16	.16	-.03	.02	.16	.02	.04	.02	-.07
Will-Do	.00	.00	-.08	-.01	.00	.14	.20	.14	.03
Adaptation	.03	.02	-.06	-.02	.03	.02	.09	.03	-.01

Note. WTBD JKT = Warrior Tasks and Battle Drills Job Knowledge Test. PRS = Performance Rating Scales. Unless otherwise noted, sample sizes range from $n = 107 - 7,690$. Correlations based on fewer than 100 cases are not reported. Correlations in bold are statistically significant ($p < .01$, two-tailed).

^a Sample sizes for six TAPAS scales were considerably smaller than the other dimensions because not all scales were administered in every version of the TAPAS ($n = 109 - 1,630$).

Table D.3. (Continued)

	APFT Score	Will Do Performance	PRS: Effort & Discipline	PRS: Adjustment to the Army	PRS: Physical Fitness and Bearing	PRS: Working with Others	PRS: Overall Performance
<i>AFQT</i>	.08	-.05	.00	-.05	.03	-.01	.05
<i>Individual TAPAS Scales</i>							
Achievement	.13	.12	.02	-.05	.03	-.02	.00
Adjustment	.02	.07	-.09	-.09	.03	-.08	-.10
Adventure Seeking ^a	.07	--	--	--	--	--	--
Attention Seeking	.10	-.12	.02	.00	-.01	.11	.05
Commitment to Serve ^a	.01	--	--	--	--	--	--
Cooperation	.02	--	.02	.07	-.02	.03	-.01
Courage ^a	-.04	--	--	--	--	--	--
Dominance	.18	.15	.06	.06	.14	.04	.11
Even Tempered	-.08	.09	.06	.00	.05	.06	.00
Intellectual Efficiency	.06	.06	.09	-.01	.05	.09	.08
Non-Delinquency	-.01	-.02	.02	-.02	-.10	-.03	-.09
Optimism	.11	.06	.10	.15	.14	.11	.11
Order	.02	--	.03	-.06	.07	.01	.00
Physical Conditioning	.23	.15	.09	.06	.10	.10	.08
Responsibility ^a	-.10	--	--	--	--	--	--
Self-Control	-.01	.16	.00	-.04	-.01	-.04	-.02
Selflessness	.04	--	-.04	-.03	-.02	-.07	-.08
Situational Awareness ^a	-.09	--	--	--	--	--	--
Sociability	.07	.13	.00	.07	.01	.03	.04
Team Orientation ^a	.13	--	--	--	--	--	--
Tolerance	.08	.06	.03	.04	.01	.07	.03
<i>TAPAS Composites</i>							
Can-Do	.01	--	.10	.01	.08	.08	.08
Will-Do	.28	.21	.12	.07	.16	.10	.11
Adaptation	.13	--	.10	.07	.15	.06	.06

Note. WTBD JKT = Warrior Tasks and Battle Drills Job Knowledge Test. PRS = Performance Rating Scales. Unless otherwise noted, sample sizes range from $n = 107 - 7,690$. Correlations based on fewer than 100 cases are not reported. Correlations in bold are statistically significant ($p < .01$, two-tailed).

^a Sample sizes for six TAPAS scales were considerably smaller than the other dimensions because not all scales were administered in every version of the TAPAS ($n = 109 - 1,630$).

Table D.4. Summary of the Bivariate Correlations between AFQT, TAPAS and Selected In-Unit Criteria for Tier 1 + 2 Soldiers

	Knowledge & Skill	WTBD JKT	Disciplinary Incidents (Y/N)	Army Fit	MOS Fit	Commitment & Fit	Retention Cognitions
<i>AFQT</i>	.46	.45	-.03	-.04	.00	-.04	-.11
<i>Individual TAPAS Scales</i>							
Achievement	.04	.06	-.09	.07	.01	.05	.04
Adjustment	.07	.07	-.01	.02	.00	.00	.00
Adventure Seeking ^a	.11	.05	.05	-.06	-.10	-.08	-.11
Attention Seeking	.05	.05	.04	.01	.00	-.01	-.06
Commitment to Serve ^a	-.13	-.18	.21	-.06	.14	.00	.13
Cooperation	-.02	.01	-.03	.05	.01	.05	.02
Courage ^a	.05	.06	.00	.19	.15	.15	.09
Dominance	.01	.03	-.02	.07	.00	.04	.06
Even Tempered	.06	.07	-.04	.08	.00	.05	.03
Intellectual Efficiency	.20	.19	-.05	.04	.01	.03	.00
Non-Delinquency	.00	.00	-.03	.07	.01	.05	.06
Optimism	.01	.00	-.05	.09	.06	.08	.02
Order	-.13	-.12	-.06	.04	-.02	.03	.09
Physical Conditioning	-.02	.01	-.06	.02	.05	.03	-.02
Responsibility ^a	.26	.21	-.19	.20	.18	.21	.22
Self-Control	.03	.03	-.06	.04	-.01	.03	.05
Selflessness	-.08	-.06	.03	.05	-.01	.04	.04
Situational Awareness ^a	-.09	-.14	.01	-.05	.07	.00	.08
Sociability	-.09	-.09	.00	.04	.01	.04	.02
Team Orientation ^a	.05	.07	.09	-.07	-.01	-.05	-.02
Tolerance	-.04	-.05	.04	.02	-.03	.00	.03
<i>TAPAS Composites</i>							
Can-Do	.28	.27	-.04	.02	.01	.00	-.05
Will-Do	.01	.04	-.10	.08	.05	.07	.03
Adaptation	.10	.12	-.04	.01	.04	.01	-.04

Note. WTBD JKT = Warrior Tasks and Battle Drills Job Knowledge Test. PRS = Performance Rating Scales. Unless otherwise noted, sample sizes range from $n = 1,059 - 1,769$. Correlations based on fewer than 100 cases are not reported. Correlations in bold are statistically significant ($p < .01$, two-tailed).

^a Sample sizes for six TAPAS scales were considerably smaller than the other dimensions because not all scales were administered in every version of the TAPAS ($n = 120 - 130$).

Table D.4. (Continued)

	APFT Score	PRS: Effort & Discipline	PRS: Working with Others	PRS: Physical Fitness and Bearing	PRS: Leadership Potential	PRS: Overall Performance
<i>AFQT</i>	-.03	.10	.14	.02	.09	.09
<i>Individual TAPAS Scales</i>						
Achievement	.06	.12	.11	.08	.13	.11
Adjustment	.01	.02	.01	-.02	.00	.01
Adventure Seeking ^a	.10	--	--	--	--	--
Attention Seeking	.04	-.01	.00	-.01	.00	.00
Commitment to Serve ^a	.12	--	--	--	--	--
Cooperation	-.04	-.04	-.01	-.03	-.05	-.03
Courage ^a	.01	--	--	--	--	--
Dominance	.11	.02	.04	.04	.03	.04
Even Tempered	-.02	-.03	.02	-.04	.00	-.02
Intellectual Efficiency	.04	.07	.10	.02	.04	.07
Non-Delinquency	-.06	.01	.02	-.03	.00	.00
Optimism	.05	.02	.05	.05	.03	.04
Order	.04	-.04	-.04	-.03	-.03	-.04
Physical Conditioning	.27	.05	.07	.13	.07	.09
Responsibility ^a	.02	--	--	--	--	--
Self-Control	.02	.05	.05	.01	.07	.04
Selflessness	-.04	-.02	-.03	-.03	-.02	-.04
Situational Awareness ^a	.06	--	--	--	--	--
Sociability	.05	-.02	.00	.01	-.01	.00
Team Orientation ^a	.11	--	--	--	--	--
Tolerance	-.01	-.03	.00	-.02	.00	-.04
<i>TAPAS Composites</i>						
Can-Do	.00	.09	.11	.03	.06	.09
Will-Do	.24	.10	.11	.14	.12	.13
Adaptation	.15	.04	.09	.09	.08	.08

Note. WTBD JKT = Warrior Tasks and Battle Drills Job Knowledge Test. PRS = Performance Rating Scales. Unless otherwise noted, sample sizes range from $n = 1,059 - 1,769$. Correlations based on fewer than 100 cases are not reported. Correlations in bold are statistically significant ($p < .01$, two-tailed).

^a Sample sizes for six TAPAS scales were considerably smaller than the other dimensions because not all scales were administered in every version of the TAPAS ($n = 120 - 130$).

Table D.5. Summary of the Bivariate Correlations between AFQT, TAPAS and Selected In-Unit Criteria for Tier 1 Soldiers

	Knowledge & Skill	WTBD JKT	Disciplinary Incidents (Y/N)	Army Fit	MOS Fit	Commitment & Fit	Retention Cognitions
<i>AFQT</i>	.46	.45	-.04	-.04	.01	-.04	-.11
<i>Individual TAPAS Scales</i>							
Achievement	.04	.06	-.09	.07	.02	.05	.04
Adjustment	.08	.08	-.01	.01	.00	.00	.00
Adventure Seeking ^a	.13	.05	.04	-.07	-.12	-.10	-.11
Attention Seeking	.05	.05	.04	.00	.00	-.01	-.06
Commitment to Serve ^a	-.13	-.19	.21	-.04	.15	.01	.16
Cooperation	-.01	.01	-.02	.06	.01	.05	.02
Courage ^a	.07	.09	.01	.16	.13	.13	.05
Dominance	.01	.03	-.02	.06	.00	.04	.05
Even Tempered	.06	.07	-.04	.08	-.01	.04	.03
Intellectual Efficiency	.20	.19	-.04	.05	.02	.03	.01
Non-Delinquency	.00	.00	-.04	.07	.01	.06	.06
Optimism	.00	.00	-.04	.09	.06	.08	.02
Order	-.13	-.12	-.06	.04	-.02	.03	.09
Physical Conditioning	-.01	.01	-.05	.02	.04	.03	-.02
Responsibility ^a	.27	.21	-.17	.21	.18	.23	.21
Self-Control	.03	.03	-.06	.04	-.01	.03	.05
Selflessness	-.08	-.06	.03	.04	-.01	.04	.04
Situational Awareness ^a	-.10	-.16	.03	-.14	.03	-.07	.03
Sociability	-.09	-.08	.01	.04	.01	.04	.02
Team Orientation ^a	.09	.10	.07	-.02	.04	.01	.05
Tolerance	-.04	-.05	.04	.01	-.03	.00	.04
<i>TAPAS Composites</i>							
Can-Do	.27	.27	-.04	.02	.01	.00	-.05
Will-Do	.01	.04	-.09	.08	.04	.07	.03
Adaptation	.11	.12	-.03	.01	.03	.01	-.04

Note. WTBD JKT = Warrior Tasks and Battle Drills Job Knowledge Test. PRS = Performance Rating Scales. Unless otherwise noted, sample sizes range from $n = 1,032 - 1,717$. Correlations based on fewer than 100 cases are not reported. Correlations in bold are statistically significant ($p < .01$, two-tailed).

^a Sample sizes for six TAPAS scales were considerably smaller than the other dimensions because not all scales were administered in every version of the TAPAS ($n = 112 - 126$).

Table D.5. (Continued)

	APFT Score	PRS: Effort & Discipline	PRS: Working with Others	PRS: Physical Fitness and Bearing	PRS: Leadership Potential	PRS: Overall Performance
<i>AFQT</i>	-.04	.09	.14	.02	.08	.09
<i>Individual TAPAS Scales</i>						
Achievement	.06	.12	.10	.08	.13	.11
Adjustment	.00	.02	.01	-.02	.01	.01
Adventure Seeking ^a	.13	--	--	--	--	--
Attention Seeking	.03	-.01	-.01	-.01	.00	.00
Commitment to Serve ^a	.10	--	--	--	--	--
Cooperation	-.04	-.05	-.01	-.03	-.05	-.03
Courage ^a	.03	--	--	--	--	--
Dominance	.10	.03	.04	.04	.03	.04
Even Tempered	-.02	-.04	.02	-.04	-.01	-.02
Intellectual Efficiency	.03	.07	.09	.02	.03	.07
Non-Delinquency	-.06	.02	.03	-.03	.01	.01
Optimism	.05	.01	.04	.05	.03	.04
Order	.04	-.03	-.04	-.03	-.03	-.03
Physical Conditioning	.27	.04	.06	.12	.07	.09
Responsibility ^a	.03	--	--	--	--	--
Self-Control	.02	.05	.05	.01	.06	.04
Selflessness	-.03	-.01	-.02	-.03	-.02	-.03
Situational Awareness ^a	.09	--	--	--	--	--
Sociability	.05	-.02	.00	.01	-.01	.00
Team Orientation ^a	.10	--	--	--	--	--
Tolerance	-.01	-.02	.00	-.02	.00	-.03
<i>TAPAS Composites</i>						
Can-Do	-.01	.08	.10	.02	.04	.08
Will-Do	.23	.09	.10	.13	.12	.12
Adaptation	.15	.03	.08	.07	.08	.06

Note. WTBD JKT = Warrior Tasks and Battle Drills Job Knowledge Test. PRS = Performance Rating Scales. Unless otherwise noted, sample sizes range from $n = 1,032 - 1,717$. Correlations based on fewer than 100 cases are not reported. Correlations in bold are statistically significant ($p < .01$, two-tailed).

^a Sample sizes for six TAPAS scales were considerably smaller than the other dimensions because not all scales were administered in every version of the TAPAS ($n = 112 - 126$).

Table D.6. Summary of the Bivariate Correlations between AFQT, TAPAS, and Selected Attrition Criteria for Tier 1 + 2 Soldiers (Regular Army Only)

	Attrition		
	6-Mos	12-Mos	24-Mos
<i>AFQT</i>	-.06	-.06	-.08
<i>Individual TAPAS Scales</i>			
Achievement	-.02	-.02	-.02
Adjustment	-.01	-.01	-.01
Adventure Seeking ^a	-.03	-.03	.02
Attention Seeking	-.03	-.03	-.01
Commitment to Serve ^a	.02	.02	.04
Cooperation	.00	.00	.00
Courage ^a	.00	-.02	.01
Dominance	-.02	-.02	-.02
Even Tempered	.00	.00	-.01
Intellectual Efficiency	-.01	-.01	-.01
Non-Delinquency	.02	.01	.01
Optimism	-.02	-.02	-.02
Order	.01	.01	.02
Physical Conditioning	-.06	-.08	-.08
Responsibility ^a	.00	.00	-.09
Self-Control	.00	.00	.00
Selflessness	.03	.03	.04
Situational Awareness ^a	-.01	.00	-.08
Sociability	.00	.00	.01
Team Orientation ^a	-.03	-.03	.05
Tolerance	.01	.01	.01
<i>TAPAS Composites</i>			
Can-Do	-.02	-.02	-.03
Will-Do	-.06	-.07	-.07
Adaptation	-.06	-.07	-.08

Note. Unless otherwise noted, sample sizes range from $n = 69,410 - 105,419$ (6-Mos); $n = 63,559 - 81,305$ (12-Mos); $n = 43,362 - 43,732$ (24-Mos). Correlations in bold are statistically significant ($p < .01$, two-tailed).

^a Sample sizes for six TAPAS scales were considerably smaller than the other dimensions because not all scales were administered in every version of the TAPAS ($n = 180 - 18,129$).

Table D.7. Summary of the Bivariate Correlations between AFQT, TAPAS, and Selected Attrition Criteria for Tier 1 Soldiers (Regular Army Only)

	Attrition		
	6-Mos	12-Mos	24-Mos
<i>AFQT</i>	-.06	-.06	-.08
<i>Individual TAPAS Scales</i>			
Achievement	-.02	-.02	-.02
Adjustment	-.01	-.01	-.01
Adventure Seeking ^a	-.03	-.03	.02
Attention Seeking	-.03	-.03	-.01
Commitment to Serve ^a	.02	.01	.04
Cooperation	.00	.00	.00
Courage ^a	.00	-.01	.03
Dominance	-.02	-.02	-.02
Even Tempered	.00	-.01	-.01
Intellectual Efficiency	-.01	-.01	-.02
Non-Delinquency	.02	.01	.01
Optimism	-.02	-.02	-.02
Order	.01	.01	.02
Physical Conditioning	-.07	-.08	-.08
Responsibility ^a	.00	.00	-.07
Self-Control	.00	.00	.00
Selflessness	.03	.03	.04
Situational Awareness ^a	-.01	.00	-.08
Sociability	.00	.00	.02
Team Orientation ^a	-.03	-.04	.03
Tolerance	.01	.01	.01
<i>TAPAS Composites</i>			
Can-Do	-.02	-.02	-.03
Will-Do	-.06	-.07	-.07
Adaptation	-.06	-.07	-.08

Note. Unless otherwise noted, sample sizes range from $n = 67,147 - 101,921$ (6-Mos); $n = 61,420 - 78,278$ (12-Mos); $n = 42,294 - 42,655$ (24-Mos). Correlations in bold are statistically significant ($p < .01$, two-tailed).

^a Sample sizes for six TAPAS scales were considerably smaller than the other dimensions because not all scales were administered in every version of the TAPAS ($n = 175 - 17,508$).

Table D.8. Summary of the Bivariate Correlations between AFQT, TAPAS, and Selected Attrition Criteria for Tier 2 Soldiers (Regular Army Only)

	Attrition		
	6-Mos	12-Mos	24-Mos
<i>AFQT</i>	-.01	-.01	-.08
<i>Individual TAPAS Scales</i>			
Achievement	-.02	.00	-.04
Adjustment	-.03	.00	.02
Adventure Seeking ^a	-.03	.00	--
Attention Seeking	-.02	-.01	-.02
Commitment to Serve ^a	.02	.04	--
Cooperation	.01	.02	.02
Courage ^a	-.03	-.04	--
Dominance	-.01	-.01	-.02
Even Tempered	.02	.02	.03
Intellectual Efficiency	.01	.00	-.01
Non-Delinquency	.01	.02	-.01
Optimism	-.03	-.01	.01
Order	-.01	-.01	.02
Physical Conditioning	-.03	-.04	-.04
Responsibility ^a	.01	.01	--
Self-Control	.02	.02	.09
Selflessness	.05	.04	.04
Situational Awareness ^a	.02	.02	--
Sociability	-.01	.00	-.02
Team Orientation ^a	.00	.00	--
Tolerance	.01	.01	-.01
<i>TAPAS Composites</i>			
Can-Do	.05	.02	.00
Will-Do	-.03	-.03	-.04
Adaptation	.01	-.01	-.02

Note. Unless otherwise noted, sample sizes range from $n = 2,263 - 3,498$ (6-Mos); $n = 2,139 - 3,027$ (12-Mos); $n = 1,068 - 1,077$ (24-Mos). Correlations based on fewer than 100 cases are not reported. Correlations in bold are statistically significant ($p < .01$, two-tailed).

^aSample sizes for six TAPAS scales were considerably smaller than the other dimensions because not all scales were administered in every version of the TAPAS ($n = 428 - 621$).