Technical Report 1366

Examining Enhanced Suitability Screening for Predicting Performance in Recruiting Duty Assignments

Christopher D. Nye
Drasgow Consulting Group

Rabiah S. Muhammad
U.S. Army Research Institute

Christopher R. Graves
Human Resources Research Organization

Fritz Drasgow, Oleksandr S. Chernyshenko, Stephen Stark, and Saba M. Butt
Drasgow Consulting Group

February 2018

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

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

**Department of the Army
Deputy Chief of Staff, G1
Authorized and approved for distribution:**

**MICHELLE ZBYLUT, Ph.D.
Director**

Research accomplished under contract
for the Department of the Army

Drasgow Consulting Group

Technical reviews by:
Leonard A. White, U.S. Army Research Institute
Mark Young, U.S. Army Research Institute

REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy) 21 February 2018	2. REPORT TYPE Final	3. DATES COVERED (from. . . to) Nov 2015 to Nov 2017
4. TITLE AND SUBTITLE Examining Enhanced Suitability Screening for Predicting Performance in Recruiting Duty Assignments		5a. CONTRACT OR GRANT NUMBER W911NF-14-C-0091
		5b. PROGRAM ELEMENT NUMBER 63007
6. AUTHOR(S) Christopher D. Nye, Rabiah S.Muhammad, Christopher R. Graves, Fritz Drasgow, Oleksandr S. Chernyshenko, Stephen Stark, and Saba M. Butt		5c. PROJECT NUMBER A790
		5d. TASK NUMBER 329a
		5e. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Drasgow Consulting Group 3508 N. Highcross Rd. Urbana, IL 61802		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences 6000 6th Street (Bldg. 1464/Mail Stop 5586) Fort Belvoir, VA 22060		10. MONITOR ACRONYM ARI
		11. MONITOR REPORT NUMBER Technical Report 1366
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.		
13. SUPPLEMENTARY NOTES ARI Research POC: Dr. Rabiah Muhammad		

14. ABSTRACT (Maximum 200 words):
The goal of this project was to examine two noncognitive assessments known as the Noncommissioned Officer Special Assignment Battery (NSAB) and the Assessment of Right Conduct (ARC) as potential predictors of performance in recruiting duty assignments. The data for this research included NSAB, ARC, and criterion data collected in March 2017. The sample consisted of a total of 5,092 Recruiters in the U.S. Army. After removing potentially unmotivated responders, 4,796 Recruiters remained in the sample. With these data, we first examined the validity of the NSAB for predicting Recruiter outcomes and then explored the validity of the NSAB and ARC combined. Results showed that the NSAB had strong validity (adjusted multiple R's ranging from .19 to .56) for predicting a broad range of outcomes, including an overall performance composite. In addition, adding the ARC to the model improved the prediction further (adjusted multiple R's ranging from .24 to .70). These results held in both the overall sample of Recruiters and in a subset of individuals in MOS 79R (i.e., career Recruiters, N = 1,079). These results suggest that the NSAB and ARC may be useful for identifying Soldiers with high potential to be successful in recruiting duty assignments.

15. SUBJECT TERMS			
Enlisted Personnel	Manpower	Personality Measures	Selection and Classification

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified	Unclassified	35	Tonia S. Heffner Unit Chief

i

Technical Report 1366

Examining Enhanced Suitability Screening for Predicting Performance in Recruiting Duty Assignments

Christopher D. Nye
Drasgow Consulting Group

Rabiah S. Muhammad
U.S. Army Research Institute

Christopher R. Graves
Human Resources Research Organization

Fritz Drasgow, Oleksandr S. Chernyshenko, Stephen Stark, and Saba M. Butt
Drasgow Consulting Group

Selection and Assignment Research Unit
Tonia S. Heffner, Chief

February 2018

ACKNOWLEDGMENTS

EXAMINING ENHANCED SUITABILITY SCREENING FOR PREDICTING PERFORMANCE IN RECRUITING DUTY ASSIGNMENTS

EXECUTIVE SUMMARY

Research Requirement:

Each year the U.S. Army recruits tens of thousands of new Soldiers in order to remain combat ready. To do so, a large number of effective Recruiters are needed to identify and support applicants who are likely to be successful Soldiers. As such, Recruiters play an important role in building the U.S. Army and helping to maintain an effective military force. Therefore, it is important to identify individuals for recruiting duty who are likely to be successful in this assignment.

To help identify successful Recruiters, the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) was tasked with developing a measure that could be administered online and in an unproctored setting and would predict Recruiter success. The result of this effort was the Noncommissioned Officer Special Assignment Battery (NSAB). The NSAB is an assessment of personality that is based on the Tailored Adaptive Personality Assessment System (TAPAS). It takes advantage of modern psychometric methods and computing technology to offer a new generation of personality measures that (a) are fake-resistant, (b) utilize computer adaptive technology to measure across a broad range of trait continua, and (c) are easily customized to meet the assessment needs of diverse occupations and military occupational specialties (MOS). As a result of these benefits, a growing body of evidence suggests that the NSAB is useful for predicting a wide variety of performance criteria (Allen, Cheng, Putka, Hunter, & White, 2010; Horgen, Nye, White, LaPort, Hoffman, Drasgow et al., 2013; Knapp & Heffner, 2012; Nye, Drasgow, Chernyshenko, Stark, Kubisiak, White, & Jose, 2012).

Although a substantial amount of work has focused on military accession testing, much less research has been conducted on in-service testing for predicting performance in special duty assignments. The goal of this project was to examine the NSAB and a separate measure known as the Assessment of Right Conduct (ARC) as potential predictors of performance in recruiting duty assignments. Although both the NSAB and the ARC have been shown to predict Soldiers' attitudes, performance, and counterproductive work behavior, both measures also assess different content that may contribute unique variance to the prediction of these outcomes. In other words, examining a combination of both the NSAB and ARC may help to create an enhanced suitability screen (ESS) for recruiting duty assignments. This report describes the research done to validate the NSAB and ARC for this purpose.

Procedure:

The data for this research effort included NSAB, ARC, and criterion data collected in March 2017. The sample consisted of a total of 5,092 Recruiters in the U.S. Army. After removing potentially unmotivated responders, 4,796 Recruiters remained in the sample. This included 1,018 individuals from MOS 79R, Career Recruiters. In addition to the NSAB and ARC data, a

large number of criteria were also collected including perceived fit with the recruiting role, commitment to the Army, satisfaction, resilience, leadership, and peer and supervisor performance ratings.

Using these criteria, correlation and regression analyses were used to examine the validity of the NSAB and the ARC for predicting each outcome. In addition, due to the large number of criteria measured, we also developed a composite of the outcome variables that we labeled overall performance and examined the prediction of this outcome in a separate regression analysis. We first estimated the validity of the NSAB alone by regressing each outcome onto the NSAB scale scores. Next, we examined the incremental validity of the NSAB over the Armed Services Vocational Aptitude Battery (ASVAB) General Technical (GT) scores using hierarchical regression. Finally, we also examined the prediction of individual criteria and the overall performance composite using both the NSAB and the ARC to determine if the ARC scales could add incremental validity.

Findings:

Results showed that the NSAB scales were valid predictors of a broad range of criteria (adjusted multiple R's ranging from .24 to .56). Some of the stronger relationships were with Army commitment, Recruiter satisfaction, resilience, and leadership motivation (adjusted multiple R's ranging from .39 to .46). However, the strongest relationship was observed for predicting the overall performance composite (adjusted multiple R = .56). In addition, the NSAB scales also showed incremental validity over the ASVAB GT scores. The prediction of each outcome improved when both NSAB and ARC were included in the model (adjusted multiple R's ranging from .24 to .70). Moreover, these relationships were slightly stronger when examining only those individuals in MOS 79R (Career Recruiters). These results suggest that the NSAB and ARC may be useful predictors of Recruiters' attitudes and performance and, therefore, may be useful for identifying high potential Soldiers for recruiting duty assignments.

Utilization and Dissemination of Findings:

These results support the use of the NSAB and ARC as screening tools for recruiting duty assignments. These measures may be useful for predicting a broad range of outcomes and can add important information to existing screening tools. The magnitudes of the relationships found in the present research add to previous research on the validity of the NSAB and suggest that this measure will be useful for in-service testing. In fact, the relationships between the NSAB and the outcomes assessed in this sample were generally larger than found in other military occupations (Nye, Drasgow, Chernyshenko, Stark, Kubisiak, White, & Jose, 2012). Importantly, the best results were obtained using a combination of NSAB and ARC to predict each outcome. To support potential operational applications of these results, more research is needed to examine these findings under operational conditions and in the context of a longitudinal project.

EXAMINING ENHANCED SUITABILITY SCREENING FOR PREDICTING
PERFORMANCE IN RECRUITING DUTY ASSIGNMENTS

CONTENTS

List of Tables

List of Figures

EXAMINING ENHANCED SUITABILITY SCREENING FOR PREDICTING
PERFORMANCE IN RECRUITING DUTY ASSIGNMENTS

INTRODUCTION

Each year the U.S. Army must recruit tens of thousands of new Soldiers to remain combat ready. For example, the FY17 goal was to add an additional 68,500 Active Duty Soldiers to increase the total number of Soldiers in the Army to 1,018,000 (U.S. Army, 2017). In order to meet this goal, a large number of effective Recruiters are needed to identify and support applicants who are likely to be successful Soldiers. As such, Recruiters play an important role in building the U.S. Army and helping to maintain an effective military force. Given their role in this process, it is important to identify individuals for recruiting duty who are likely to be successful in this assignment.

To help identify successful Recruiters, the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) conducted a series of studies from 2000-2005 to develop an assessment that could be used for screening Soldiers for this assignment. The result of this effort was the Noncommissioned Officer Leadership Skills Inventory (NLSI), which assessed a broad range of skills and abilities that were thought to relate to Recruiter success. This assessment was administered in a proctored setting and in a paper-and-pencil format (White, Borman, & Bowles, 2001). Subsequent research provided support for the predictive validity of this measure (Halstead, 2009; Horgen et al., 2006; White et al., 2004) and a computerized version of the NLSI was implemented at Digital Training Facilities (DTFs) to screen Soldiers for recruiting duty (Halstead, 2009; Horgen, Nye, White, LaPort, Hoffman, Drasgow et al., 2013).

Although the NLSI showed promise in initial testing, this assessment could not be fully implemented due to the insufficient number of Soldiers testing at the DTFs. Therefore, the U.S. Army Training and Doctrine Command (TRADOC) and Human Resources Command (HRC) requested that ARI develop a measure that could be administered online, in an unproctored setting, and could predict Recruiter success. The result of this effort was the Noncommissioned Officer Special Assignment Battery (NSAB). The goal of the current project was to examine the NSAB as a potential predictor of performance in recruiting duty assignments.

THE NONCOMMISIONED OFFICER SPECIAL ASSIGNMENT BATTERY (NSAB)

The NSAB is an assessment of personality that is based on the Tailored Adaptive Personality Assessment System (TAPAS). Given the tasks performed by Recruiters, personality is likely to be a significant predictor of success in this assignment. Interest in personality as a predictor of performance has increased considerably over the past two decades. Much of this interest was galvanized by empirical evidence showing that personality constructs, such as conscientiousness, predict performance across a diverse array of civilian and military occupations (e.g., Barrick & Mount, 1991; Campbell & Knapp, 2001) and provide incremental validity beyond general cognitive ability (e.g., Schmidt & Hunter, 1998).

Despite growing interest and positive empirical support for their validity, a major concern with personality assessments is applicant faking in high-stakes settings. Past research has shown that test takers can easily identify the correct or socially desirable responses on single statement

personality measures and increase or decrease their scores when sufficiently motivated (Mueller-Hanson, Heggestad, & Thornton, 2003; White, Young, & Rumsey, 2001). As a result, faking is a potential threat to the validity of personality measures and may affect its utility in operational selection settings (White, Young, Hunter, & Rumsey, 2008).

To help address issues with faking on personality assessments, the TAPAS was developed under the Army's Small Business Innovation Research (SBIR) grant program. It takes advantage of modern psychometric methods and computing technology to offer a new generation of personality measures that (a) are fake-resistant, (b) utilize computer adaptive technology to measure across a broad range of trait continua, and (c) are easily customized to meet the assessment needs of diverse occupations and military occupational specialties (MOS). The TAPAS can measure up to 27 personality dimensions or facets. Of these, 21 cover the behavioral patterns associated with the well-known Big Five personality framework (Goldberg, 1993). The remaining six dimensions cover military-specific temperament traits (Physical Conditioning, Courage, Team-Orientation, Adventure Seeking, Situational Awareness, and Commitment to Serve). As such, the TAPAS is among the most comprehensive measures of personality facets that is currently available (Drasgow, Stark, Chernyshenko, Nye, Hulin, & White, 2012).

TAPAS tests utilize a multidimensional pairwise preference format that is designed to be resistant to faking by asking test-takers to choose between two statements about themselves rather than responding to a single question about their behavior. For each item, the two statements in each pair are matched on both social desirability and extremity on the dimensions they assess. The purpose of matching statements in this way is to make identifying and selecting the most socially desirable responses more difficult for test-takers. This approach appears to work as research on the operational use of the TAPAS has found no evidence of score inflation, even when compared to other respondents taking the test for "research purposes only" (Drasgow et al., 2012). Due to the measurement approach used by the TAPAS and the broad range of facets assessed, this measure is expected to demonstrate validity even in high-stakes settings where applicants may be motivated to respond dishonestly.

In fact, a growing body of evidence suggests that the TAPAS is useful for predicting a wide variety of performance criteria. In 2006, ARI initiated a longitudinal research project to examine the validity of non-cognitive measures for predicting Army outcomes. Results showed that the TAPAS provided significant incremental validity over the ASVAB for predicting attrition, end of training criteria, and in-unit performance (Knapp & Heffner, 2009; Knapp, Owens, Allen, 2011). In addition, this research also showed that the TAPAS provided non-trivial gains in classification efficiency over the ASVAB alone. Additional predictive validity evidence for the TAPAS was collected during the U.S. Army's *Expanded Enlistment Eligibility Metrics* (EEEM) research project from 2007-2009 (Knapp & Heffner, 2010). The results of the EEEM effort indicated that when TAPAS trait scores were added into a regression analysis that already included AFQT, the multiple correlation increased by .26 for the prediction of physical fitness, by .16 for the prediction of disciplinary incidents, and by .20 for the prediction of 6-month attrition (Allen, Cheng, Putka, Hunter, & White, 2010). None of these criteria were predicted well by AFQT alone (predictive validity estimates were consistently below .10).

Based on these initial validation efforts, the U.S. Army approved the initial operational testing and evaluation (IOT&E) of the TAPAS for use with Army applicants at Military Entrance Processing Stations (MEPS). Since this project began in May of 2009, the TAPAS has been administered to over 700,000 Army applicants. With these data, a clearer picture of the validity of the TAPAS in operational settings has emerged. For example, this research has shown that the TAPAS has validity for predicting a broad range of performance criteria including Army Physical Fitness Test (APFT) scores, disciplinary incidents, and attrition from the U.S. Army (Knapp & Heffner, 2012). This research has also demonstrated the utility of the TAPAS for predicting performance within specific MOS. Nye, Drasgow, Chernyshenko, Stark, Kubisiak, White, and Jose (2012) found adjusted (for capitalization on chance) multiple correlations of .32, .25, .36, and .24 for MOS 11B, 31B, 68W, and 88M, respectively, for predicting an overall measure of performance. Similar composites were also developed for predicting attrition and job knowledge criteria with validities ranging from .18 to .35 for the same MOS. These results suggest that the TAPAS is useful for predicting performance across a broad range of military specialties.

Although a substantial amount of work has been conducted on the use of TAPAS for military accession testing, less has been conducted on in-service testing for predicting performance in special duty assignments. As such, it is possible that personality traits change or that the validity of personality for predicting performance in these assignments differs from the prediction of performance immediately after initial entry. Nevertheless, there is some preliminary evidence that personality traits may be useful for in-service testing as well. For example, Nye, Beal, Drasgow, Dressel, White, and Stark (2014) demonstrated that the TAPAS was useful for predicting performance in the Army Special Operations Forces Selection and Assessment Course and ongoing research (Horgen, Nye, Drasgow, Chernyshenko, Stark, Owens, & Fletcher, in preparation) has shown utility for predicting performance in a sample of Noncommissioned Officer Education System (NCOES) Instructors. Most important for the present project, Horgen, Nye, White, LaPort, Hoffman, Drasgow et al. (2013) examined the relationship between NSAB scores and the job performance of experienced recruiters in their recruiting duty assignments. Results showed that the NSAB scales were significant predictors of supervisor ratings of performance, fit with the recruiting assignment, stress, and commitment to the Army with adjusted multiple R's ranging from .30 to .49.

PURPOSE OF THE CURRENT RESEARCH

Despite the positive evidence for the validity of the NSAB in a sample of Recruiters, more research is needed. The research conducted by Horgen et al. (2013) represented a relatively small sample of Recruiters (N = 854) and more evidence is needed to justify the use of the NSAB for operational in-service testing. In addition to examining the utility of the NSAB in a large sample of Recruiters, a second goal of the present work was to examine additional individual differences that might be useful for predicting recruiting duty success. Although past research has demonstrated that the TAPAS/NSAB may be useful for both accession and in-service testing, there may be other characteristics that are not assessed by the NSAB but that may be relevant for Recruiter performance. Recent research focused on identifying and developing new TAPAS scales that might be useful for predicting performance and deviance in military assignments (Nye, Drasgow, Chernyshenko, Stark, Muhammad, & Wolters, 2017). This work identified six new TAPAS dimensions: humility, Machiavellianism, Army self-efficacy,

3

persistence, self-efficacy, and virtue. These dimensions were selected based on previous research indicating that they could be useful for predicting work outcomes. However, these dimensions have not yet been evaluated to determine their validity in military samples or for in-service testing. Therefore, the present work examined the validity of these dimensions as well.

In addition, we also examined the Assessment of Right Conduct (ARC) to determine if the characteristics assessed by this measure could add incremental validity to the NSAB scales. The ARC was developed specifically to predict counterproductive work behavior in U.S. Army Special Operations Forces (ARSOF; Kilcullen, White, Sanders, & Hazlett, 2003). This measure is comprised of single-statement multiple choice questions that assess nine broad attributes thought to be related to counterproductive workplace behavior. These nine attributes are described in Table 1. Although not administered in a forced-choice format like the NSAB, the ARC also includes a Response Distortion scale that is designed to detect faking on this assessment. The ARC has been validated in samples of U.S. Army Correctional Specialists, ARSOF candidates, Soldiers entering Initial Military Training (IMT), and U.S. Disciplinary Barracks inmates. In each of these samples, several of the attributes assessed by the ARC were related to various criteria including job performance, disciplinary incidents, and interpersonal skills (Kilcullen et al., 2003: Kilcullen, Gluszek, Beeco, Repchick, & Swigart, In-prep).

Table 1. Attributes Assessed by the ARC

Attribute	Definition
Social Maturity	High scoring individuals tend to be law-abiding and respectful of the rights and property of others.
Aggression	High scoring individuals tend to resort to physical force to resolve interpersonal disputes.
Empathy	High scoring individuals tend to feel sympathy and express concern for those who are experiencing misfortune.
Goal Orientation	High scoring individuals tend to set ambitious performance and career advancement goals for their career in the Army and expect to achieve these goals.
Peer Leadership	High scoring individuals tend to desire to obtain positions of authority and influence, are comfortable being in charge of a group and are willing to make tough decisions and accept responsibility for the group's performance.
Hostility to Authority	High scoring individuals tend to be expressively angered by authority figures and may actively disregard their instructions and policies.
Power	High scoring individuals tend to seek control over others for narcissistic and self-serving reasons.
Self-Efficacy	High scoring individuals have felt successful in past undertakings and expect this to continue in the future.
Work Motivation	High scoring individuals tend to give their best effort and work hard toward achieving difficult objectives.

Importantly, although both the NSAB and the ARC have been shown to predict Soldiers' attitudes, performance, and counterproductive work behavior, both measures also assess different content that may contribute unique variance to the prediction of these outcomes. In other words, examining the combined content in both the NSAB and ARC may help to create an enhanced

suitability screen (ESS) for recruiting duty assignments. This report describes the research done to validate the NSAB and ARC for this purpose.

METHOD

SAMPLE

The data for this research effort included NSAB, ARC, and criterion data collected in March 2017. The data consisted of a total of 5,092 Recruiters in the U.S. Army. Approximately 90% of this sample (N = 4,579) were male and 71% (N = 3,627) were Caucasian. The average individual in this sample was 32 years old and had 11 years of service in the Army. The majority of individuals in this sample were in grades E-6 (59%) or E-7 (21%). In addition, 84% (N = 4,287) of the sample were Regular Army and 14% (N = 693) were in the Army Reserve Component. Only 18% of the sample had volunteered for a Recruiting assignment and the average number of months serving as a Recruiter was 32 (29% had less than 12 months of service as a Recruiter). The majority of the sample also had at least some leadership experience (80% had been squad leaders, 72% team leaders, and 33% platoon sergeants). In this sample, 21% (N = 1,079) were from MOS 79R, which is the MOS for Career Recruiters who are permanently assigned to recruiting duty. Due to the number of participants, an MOS was counted if there was more than 50 respondents, resulting in 22 MOS.

During the administration of the NSAB, detailed information about the participant's performance on the assessment was also collected. This included the number of minutes taken to complete the test, flags to detect aberrant responding, and other relevant item response data. This information was then used to screen participants for unmotivated responding prior to running all analyses. Respondents were flagged as potentially unmotivated if they responded to more than 12 items in less than one second. In addition, respondents were also flagged using a Markov chain statistic to detect patterned responding and the ℓ_z appropriateness index to detect aberrant responding (Drasgow, Levine, & Williams, 1985; Stark, Chernyshenko, Nye, Drasgow, & White, 2017). Both the Markov chain statistic and ℓ_z are IRT based aberrance indicators that are calculated in the adaptive TAPAS software and commonly used to detect random responding (Stark et al., 2017). After removing potentially unmotivated responders, 4,796 Recruiters remained in the sample. This included 1,018 individuals from MOS 79R. This reduced sample of Recruiters was used for all subsequent analyses reported here.

Each individual in this sample was also asked to rate the performance of at least three of his or her peers. Approximately 4,713 Recruiters provided ratings for 2,906 of their peers (multiple raters rated the same individuals). In addition, a sample of supervisors was also asked to provide performance ratings for the Recruiters they worked with. The sample of supervisors included approximately 1,528 Center Leaders who provided 2,672 ratings. This sample of supervisors was approximately 72% male, 53% were E-9, and the vast majority had 12 or more years of service.

PREDICTOR MEASURES

Noncommissioned Officer Special Assignment Battery (NSAB). A version of the NSAB that assessed 16 personality dimensions was administered to Recruiters for this project. As noted above, five of the six new scales developed for the TAPAS (Nye et al., 2017) were also included in this version of the NSAB. The new Army Self-Efficacy scale was excluded from this research to avoid redundancy with the general Self-Efficacy scale. Table 2 provides descriptions of these 16 dimensions. These dimensions were assessed with 128 items administered in the pairwise preference format described above and respondents were instructed to indicate which statement

is "more like you." Again, the statements for each pair were matched based on both their extremity and social desirability to make the NSAB more fake-resistant (Drasgow et al., 2012).

This version of the NSAB was also administered as a computer adaptive test (CAT). With this format, the items that a Soldier sees in the assessment are dependent on his or her answers to previous items. Individuals begin the assessment by responding to an initial set of NSAB items. After responding to these items, the computer program estimates the individual's standing on the traits assessed by the test and administers additional items that match these estimates. The estimates of an individual's standing on these latent traits are then updated with each response and items are selected based on these updated trait scores. As a result of this algorithm, each individual will see a different set of items that are specifically tailored to his or her personality.

Computer adaptive testing has a number of advantages over the more traditional static testing format. For example, adaptive tests can increase test security. Because each individual sees a different version of the test that is tailored to his or her own personality, this format can reduce item exposure and make sharing test answers more difficult across test takers. In addition, adaptive testing can also increase the efficiency of an assessment, with some research indicating that CATs can reduce test length by as much as 50% while still maintaining the accuracy of the latent trait estimates (Stark, Chernyshenko, Drasgow, & White, 2012). Therefore, the focus of the current project was on evaluating an adaptive version of the NSAB in a large sample of U.S. Army Recruiters.

Table 2. NSAB Facets

TAPAS Facet Name	Brief Description	"Big Five" Broad Factor
Dominance	High scoring individuals are domineering, "take charge" and are often referred to by their peers as "natural leaders."	Extra-version
Sociability	High scoring individuals tend to seek out and initiate social interactions.	
Humility	High scoring individuals tend to be honest, to put others' needs before their own, and do not focus on being the center of attention.	Agree-ableness
Selflessness	High scoring individuals are generous with their time and resources.	
Achievement	High scoring individuals are seen as hard working, ambitious, confident, and resourceful.	Consci-entiousness
Order	High scoring individuals tend to organize tasks and activities and desire to maintain neat and clean surroundings.	
Non-Delinquency	High scoring individuals tend to comply with rules, customs, norms, and expectations, and they tend not to challenge authority.	
Responsibility	High scoring individuals are dependable, reliable and make every effort to keep their promises.	
Persistence	High scoring individuals tend to focus on tasks and activities until they are completed and are determined to accomplish their goals even in the face of obstacles.	
Virtue	High scoring individuals adhere to standards of honesty, morality, and "good Samaritan" behavior.	
Even Tempered	High scoring individuals tend to be calm and stable. They don't often exhibit anger, hostility, or aggression.	Emotional Stability
Optimism	High scoring individuals have a positive outlook on life and tend to experience joy and a sense of well-being.	
Intellectual Efficiency	High scoring individuals are able to process information quickly and would be described by others as knowledgeable, astute, and intellectual.	Openness to Experience
Machiavellianism	High scoring individuals generally try to deceive and manipulate others for personal gain.	Other Facets
Physical Conditioning	High scoring individuals tend to engage in activities to maintain their physical fitness and are more likely to participate in vigorous sports or exercise.	
Self-Efficacy	High scoring individuals are confident in their skills and abilities to accomplish any task that they take on.	

Assessment of Right Conduct (ARC). The ARC was also administered to identify the potential contribution of this assessment when combined with the NSAB. The version of the ARC that was administered consisted of 77 single-statement items (i.e., items were not administered in the pairwise preference format used in the NSAB) assessing the nine dimensions shown in Table 1. Each item consisted of five response options and the anchors for these options varied for each question.

Armed Services Vocational Aptitude Battery (ASVAB). Because of its role in the current screening of Army Recruiters, we used ASVAB scores as the baseline for comparing the validity of the NSAB for predicting Recruiter success. The ASVAB contains 9 subtests that assess multiple aptitudes, which are combined to create composites, and are used as the basis for current selection and classification decisions. For example, the Armed Forces Qualification Test (AFQT), which is a composite of the Word Knowledge, Paragraph Comprehension, Arithmetic Reasoning, and Math Knowledge subtests of the ASVAB, is used for enlistment screening. For MOS classification, the ASVAB subtests are used to form nine Aptitude Area (AA) composites that correspond to the various MOS. Candidates for recruiting duty assignments are screened on the General Technical (GT) AA composite and must receive a minimum score on this composite to qualify for this assignment. Therefore, we examined the incremental validity of the NSAB over individuals' GT scores.

CRITERION MEASURES

Consistent with ARI's earlier efforts to validate screening instruments for the selection of U.S. Army Recruiters (Horgen et al., 2006, 2013), measures of Recruiter criteria included performance rating scales (PRS) and a self-report attitudinal survey called the Recruiter Life Questionnaire (RLQ). Existing versions of both instruments were updated and modified, where needed, for use in the present validation effort. Instruments were updated based on reviews of current Army Recruiter job documentation, as well as on input from Army Recruiting Course (ARC) instructors, who served as subject matter experts (SMEs).

Performance Rating Scales (PRS). We developed the PRS based on the most recently constructed version of the Recruiter rating scales (Horgen et al., 2013). To evaluate the accuracy of these scales and their coverage of the criterion, we performed a detailed comparison of existing scale content with available sources of Recruiter job documentation and existing Army-wide dimensions of performance. Job documentation included content from the Recruiter Handbook (USAREC, 2011) and the Army Recruiting Course (ARC) program of instruction (POI), course management plan (CMP), and evaluation rubrics. Using this information, we modified the existing scales to provide a new set of scales that reflected current recruiting doctrine and incorporated several Army-wide dimensions of performance. The final Recruiter PRS included 13 scales:

- Identifying and Contacting Qualified Prospects
- Planning, Organizing, and Preparing
- Gaining and Maintaining Rapport
- Conducting Interviews and Counseling
- Supporting Applicant Processing
- Supporting Future Soldier Training
- Establishing and Maintaining Good Relationships in the Community
- Supporting Other Recruiters and USAREC
- Personal Discipline (Army-wide)
- Displaying Tolerance (Army-wide)
- Handling Problems (Army-wide)
- Maintaining Physical Fitness and Well-being (Army-wide)
- Overall Performance

After completing a draft version of the Recruiter PRS, we presented the scales to eight ARC instructors in a workshop at the US Army Recruiting and Retention School in Fort Knox, KY for review. All SMEs had extensive experience as Recruiters and four had prior experience as Recruiting Center Commanders. Based on the feedback from these SMEs, the content of these scales was revised to accurately reflect important dimensions of Recruiter performance.

To evaluate the performance rating scales, we ran an exploratory factor analysis on the 13 dimensions of performance that were rated. Results indicated that these 13 dimensions were represented well by a single factor model. This is clearly illustrated in the screen plot shown in Figure 1 for the supervisor ratings. As a result, both peer and supervisor performance ratings were treated as indicators of overall performance, rather than examining each individual dimension separately. However, based on past research, which found differences in the prediction of peer and supervisor ratings (Horgen et al., 2013), these two sources of ratings were examined separately.
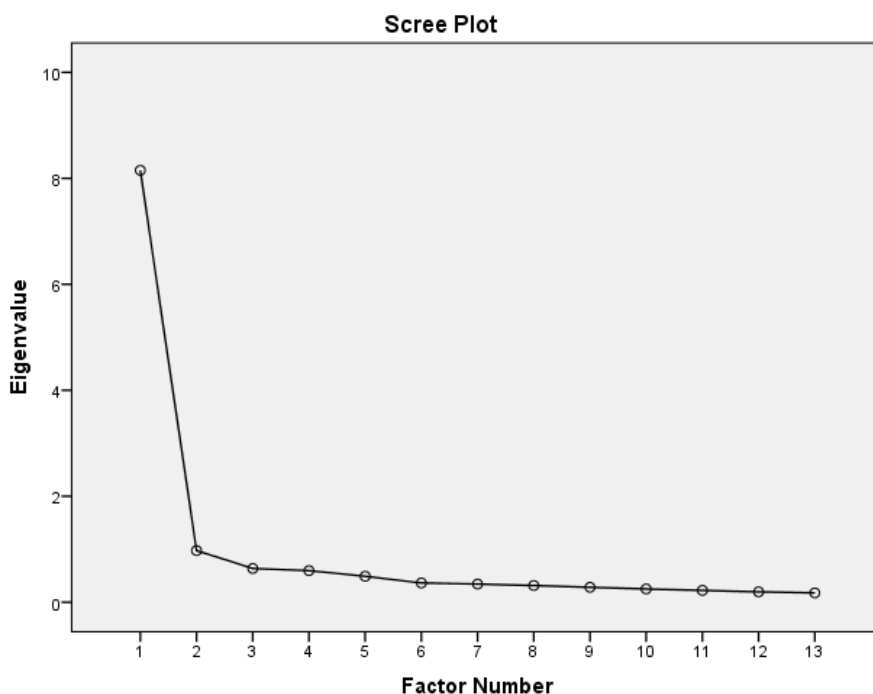
Figure 1. Scree Plot of the Supervisor Rating Dimensions

Recruiter Life Questionnaire (RLQ). The RLQ is an adaptation of the Army Life Questionnaire, which is a self-report attitudinal measure currently used in ARI validation research.[1] The RLQ includes sections on Recruiter demographic, background, and experience information, as well as assessments of job satisfaction, fit, and commitment. The RLQ also includes measures of stress and coping, leadership effectiveness, organizational citizenship behaviors (OCBs), motivation to lead, resilience, and various types of counterproductive work behaviors (CWBs; e.g., social media CWBs). The RLQ was reviewed during the SME review

---

[1] The Army Life Questionnaire was initially developed in 2005 (Van Iddekinge, Putka, & Sager, 2005) and has been updated on several occasions to meet the Army's requirements for measuring Soldier outcomes.

workshop described earlier in this section to ensure the accuracy of terminology, background, experience, and training item content. During the workshop, SMEs recommended numerous changes to tailor the measure for the Recruiter experience and environment.

Table 3 provides the means, standard deviations, alpha reliabilities, and intercorrelations for the peer and supervisor performance ratings and the RLQ scale scores. In addition to examining each outcome individually, we also examined the prediction of an overall criterion composite. To do so, scores for each criterion were first standardized to account for differences in their standard deviations and then summed using unit weights to create an overall criterion composite. Negatively worded scales (i.e., stress, disciplinary incidents, and CWB) were reverse coded before calculating the overall composite score so that all scales were in a consistent direction. The goal of combining scales in this way was to determine the utility of the NSAB for predicting a broader criterion variable and, therefore, for Recruiter selection. Table 3 also provides the descriptive statistics and intercorrelations for the Overall Performance composite. Because each of the scales comprising the overall performance composite was first standardized to account for differences in their distributions, the mean of this variable was near zero.

In addition, Table 4 shows the specific RLQ items with the highest endorsement rates in this sample. These results provide insight into the attitudes and perceptions of the Recruiters in this sample. As shown in Table 4, nearly 80% of individuals in the sample were confident that they would stay in the Army until they retired and that they could handle stressful situations. These individuals were also proud to be a part of the Army and likely to view their time in the Army as beneficial. Overall, the results presented in Table 4 illustrate that the Recruiters in this sample are generally satisfied with their job, committed to the Army, and feel that they fit in the Recruiter role.

Table 3. Descriptive Statistics for the Criteria Assessed in this Sample

						Criteria[a]											
Variables	**Mean[b]**	**SD**	**Fit**	**Com.**	**Sat.**	**LD Sat.**	**Cop.**	**Res.**	**CWB**	**OCB**	**SM CWB**	**Lead**	**Mot.**	**Disc.**	**Sup.**	**Peer**	**OP**
Recruiter Fit (RLQ)	3.22	0.70	.77														
Army Commitment (RLQ)	3.81	1.12	.71	.83													
Recruiter Satisfaction (RLQ)	3.54	1.18	.71	.80	.93												
Leadership and Development (RLQ)	3.54	0.96	.59	.71	.75	.88											
Coping (RLQ)	2.49	0.71	.22	.32	.41	.37	.66										
Resilience (RLQ)	3.80	0.57	.30	.46	.42	.48	.19	.89									
Counterproductive Work Behavior (CWB) (RLQ)	1.75	1.01	-.12	-.17	-.13	-.15	.04	-.24	.98								
Organizational Citizenship Behavior (OCB) (RLQ)	3.52	0.94	.12	.21	.19	.15	.17	.32	-.40	.97							
Social Media CWB (RLQ)	1.41	0.66	-.06	-.14	-.09	-.12	.10	-.27	.45	-.19	.90						
Leadership (RLQ)	3.70	0.90	.17	.27	.25	.24	.16	.41	-.25	.62	-.22	.98					
Leadership Motivation (RLQ)	3.71	0.63	.14	.26	.20	.21	.06	.52	-.22	.34	-.23	.39	.90				
Disciplinary Incidents (RLQ)	0.42	1.47	.03	.00	-.01	-.01	-.05	-.02	.00	.01	.04	.01	-.04	--			
Supervisor Ratings (PRS)	4.37	1.11	.14	.16	.17	.14	.20	.11	.00	.11	-.01	.14	.07	-.13	.95		
Peer Ratings (PRS)	4.53	1.09	.13	.12	.15	.10	.17	.10	-.04	.10	-.01	.14	.06	-.05	.38	.98	
Overall Performance Composite	.05	7.40	.62	.72	.73	.69	.50	.66	-.38	.57	-.29	.61	.50	-.18	.43	.35	--

Note: Sample sizes range from 909 (Overall Performance Composite) to 4726 (Army Commitment and Recruiter Satisfaction). All correlations above .02 in absolute value are significant, $p < .05$. Alpha reliabilities are listed on the diagonal. Reliabilities were not estimated for disciplinary incidents and overall performance because these outcomes were composites of separate variables that did not necessarily reflect latent constructs.

[a] Fit = Recruiter Fit, Com. = Army Commitment; Sat. = Recruiter Satisfaction; LD Sat.. = Leadership and Development Satisfaction; Cop = Coping; Res. = Resilience; CWB = Counterproductive workplace behavior; OCB= Organizational Citizenship Behavior; SM CWB = Social Media CWB; Lead = Leadership; Mot. = Leadership Motivation; Disc. = Disciplinary Incidents; Sup. = Supervisor Ratings; Peer = Peer Ratings; OP = Overall Performance.

[b] Scaling of most variables was on a 1-5 Likert scale except for Disciplinarily Incidents, which was dichotomous (0-1) and Social Media CWB, which was a 1-4 frequency count.

Table 4. Highest Endorsement Rates for the RLQ Items

RLQ Item	Percent of Individuals who Endorsed the Item
How confident are you that you will stay in the Army until you retire?	78.6%
I am confident in my ability to get through the stressful situation.	78.3%
I am proud to be a representative of the Army as a Recruiter.	72.9%
As a Recruiter I have learned new skills that will make me a better NCO.	69.8%
Overall I consider my time spent as a Recruiter beneficial to my NCO career.	68.3%
Demonstrate the Army Values both on- and off-duty.	62.5%
My time as a Recruiter will increase my chances of promotion.	61.7%
Your opportunity to use your aptitudes, experience, and training in your assignment as a Recruiter?	60.9%
The amount of variety in the work you do as a Recruiter?	60.6%
Working as a Recruiter will help me achieve long-term career goals.	59.4%
I am the right person for the type of work required of Recruiters.	57.1%
Your opportunity as a Recruiter to perform work you find interesting?	53.6%
Being a Recruiter has great personal meaning for me.	52.8%
I like the work I am doing as a Recruiter.	51.6%
Make sure everyone in your unit feels like a valued member of the team.	50.6%

PROCEDURES AND ANALYSES

Participants for this research was to recruited via an email request from the U.S. Army Recruiting Command (USAREC) that was sent to all Recruiters. An initial email was sent to express USAREC's support for the project and request participation from all Recruiters. Then, a second email was sent on March 3, 2017 to again request participation and provide the link for the online survey hosted on the Army Analytics Group's (AAG) servers. Recruiters who volunteered to participate in the research project clicked on the link and logged into the survey using their Common Access Card (CAC).

The assessment began by asking participants to read information related to the purpose of the research and sign a consent form. After electronically signing the document, Recruiters then entered their names so that their responses could be linked to the peer and supervisor ratings and filled out a demographic and background information sheet. Next, examinees read an instruction page that provided detailed information about answering NSAB items and then proceeded to respond to the actual items. Following the NSAB, participants completed the ARC and the RLQ. After completing these initial assessments, participants were asked to provide ratings of their

peers. Supervisors were given a separate link to provide their ratings. Finally, participants were debriefed and thanked for their participation.

To examine the validity of the NSAB and ARC for predicting recruiting duty success, we first examined correlations between each of these scales and the outcomes assessed in this project. However, a key factor in the validity of the TAPAS has been the use of composites of TAPAS scales. These composites have several advantages for predicting performance both at the Army-wide level and in particular assignments. Specifically, personality composites are important because Army jobs (as with many civilian jobs) are complex and require a broad range of individual characteristics to perform well. Therefore, we would not expect a single narrow NSAB dimension to predict all aspects of performance in each MOS. Instead, composites of the NSAB scales assess a broader range of individual characteristics that are relevant to performance outcomes. In other words, although we may find only small or moderate correlations between individual NSAB scales and performance, we expect combinations of scales to predict performance well. This expectation is consistent with recent research demonstrating that composites of personality scales are better predictors of work outcomes than individual scales (Judge, Rodell, Klinger, Simon, & Crawford, 2013).

To explore the validity of NSAB composites, we used regression analyses. We first estimated the validity of the NSAB alone by regressing each outcome onto the NSAB scale scores. Next, we also examined the incremental validity of the NSAB scales over the ASVAB General Technical (GT) scores using hierarchical regression. To do so, we entered GT scores in Step 1 of the model and then added the NSAB scores in Step 2 to determine the change in the overall validity of the regression model. Finally, we also examined the prediction of individual criteria and the overall performance composite using both the NSAB and the ARC to determine their validity when combined.

VALIDATION: RESULTS

Tables 5 and 6 show the descriptive statistics for the NSAB and ARC scales, respectively. Table 5 shows the raw scores for the NSAB scales. These scores are in the IRT theta metric and typically range from approximately -2.30 to 2.30. As shown in this table, individuals tended to score highest on the Achievement and Dominance scales and lowest on the Order dimension. The lower scores on the Order dimension are actually a positive finding given that Order is often negatively correlated with performance in other samples (Nye et al., 2012). Table 6 reports results for the ARC scales. These scores are also in the raw score metric and ranged from 2.16 to 3.96.

Table 5. Descriptive Statistics for the NSAB Dimensions

TAPAS Facets	N	Mean	Standard Deviations
Achievement	4,796	0.27	0.56
Dominance	4,796	0.26	0.55
Even Tempered	4,796	0.05	0.51
Humility	4,796	-0.01	0.60
Intellectual Efficiency	4,796	0.21	0.54
Machiavellianism	4,796	0.10	0.55
Non-Delinquency	4,796	-0.03	0.53
Optimism	4,796	0.20	0.52
Order	4,796	-0.25	0.52
Persistence	4,796	0.01	0.54
Physical Conditioning	4,796	0.23	0.44
Responsibility	4,796	0.14	0.58
Self-Efficacy	4,796	0.11	0.53
Selflessness	4,796	-0.15	0.44
Sociability	4,796	-0.10	0.66
Virtue	4,796	0.18	0.53

Table 6. Descriptive Statistics for the ARC Dimensions

ARC Facets	N	Mean	Standard Deviations
Aggression	4772	2.16	0.78
Empathy	4772	3.56	0.56
Goal Orientation	4624	2.99	0.56
Hostility to Authority	4772	3.05	0.66
Peer Leadership	4772	3.51	0.62
Power	4772	2.73	0.64
Self-Efficacy	4771	3.93	0.60
Social Maturity	4491	3.96	0.64
Work Motivation	4772	3.87	0.62

Tables 7 and 8 show the correlations between the predictor scales and the criteria. Table 7 shows that a number of NSAB scales were significant predictors of these criteria. For example, the Achievement facet had the strongest relationships with a number of criteria and was positively related to outcomes like resilience ($r = .32$), OCB ($r = .26$), leadership ($r = .26$), and overall performance ($r = .33$). Other facets had moderate relationships with specific criteria. For example, Dominance was related to leadership motivation ($r = .32$) and leadership ($r = .21$). In addition, Optimism was more closely related to attitudinal outcomes such as Army commitment ($r = .22$), Recruiter satisfaction ($r = .23$), and leadership and development satisfaction ($r = .24$). Not surprisingly, Sociability also contributed to Recruiter attitudes and performance and was most strongly related to perceptions of fit ($r = .27$), satisfaction ($r = .28$), and overall performance ($r = .32$).

Table 8 shows that the ARC scales were also significantly correlated with a number of criteria. Here, both the Empathy and the Self-Efficacy scales had some of the largest correlations in this sample across a broad range of criteria. However, there were also substantial correlations for specific scales. For example, the Peer Leadership scale was strongly correlated with leadership motivation ($r = .46$). In addition, Empathy was significantly related to both resilience ($r = .38$) and leadership motivation ($r = .36$). Based on these results, it appears that the ARC scales may contribute to the prediction of these outcomes and may be useful to consider in combination with the NSAB.

Table 7. Correlations Between the NSAB Facet Scales and Each Criterion in the Sample

Variables	Fit	Com.	Sat.	LD Sat.	Cop.	Res.	CWB	OCB	SM CWB	Lead	Mot.	Disc.	Sup.	Peer	OP
Achievement	.17	.23	.22	.19	.18	.32	-.11	.26	-.08	.26	.29	-.04	.17	.14	.33
Dominance	.14	.14	.14	.12	.14	.25	-.04	.17	-.01	.21	.32	-.03	.09	.11	.27
Even Tempered	.12	.19	.18	.18	.16	.16	-.02	.09	-.02	.09	.02	-.01	.05	.00	.18
Humility	-.09	-.07	-.08	-.06	-.08	-.06	-.04	-.04	-.07	-.10	-.08	-.01	-.06	-.05	-.14
Intellectual Efficiency	.01	.01	.01	-.03	.05	.18	-.02	.14	-.04	.12	.19	.02	.05	.07	.11
Machiavellianism	-.02	-.07	-.04	-.06	-.02	-.04	.08	-.08	.09	-.05	-.04	.00	.02	.02	-.04
Non-Delinquency	.09	.13	.11	.11	.07	.06	-.06	.08	-.03	.07	.04	-.03	.02	.00	.11
Optimism	.18	.22	.23	.24	.22	.28	-.05	.13	-.03	.17	.14	-.01	.12	.11	.29
Order	-.01	-.03	-.02	-.01	.01	-.02	.02	.00	.04	.01	-.03	.00	.04	.07	.00
Persistence	.11	.16	.15	.12	.15	.30	-.07	.22	-.05	.22	.26	.00	.10	.12	.26
Physical Conditioning	.06	.15	.12	.10	.07	.27	-.16	.20	-.16	.19	.29	-.05	.07	.02	.19
Responsibility	.07	.08	.09	.11	.15	.22	-.01	.08	.02	.10	.10	-.03	.10	.11	.17
Self-Efficacy	.10	.12	.11	.09	.14	.30	-.08	.20	-.07	.21	.31	-.03	.08	.12	.24
Selflessness	.15	.20	.18	.16	.12	.14	-.06	.16	-.03	.15	.13	-.01	.05	.02	.21
Sociability	.27	.25	.28	.21	.21	.19	.00	.12	.05	.19	.13	-.03	.11	.12	.32
Virtue	.10	.18	.12	.13	.06	.23	-.14	.19	-.16	.17	.24	-.02	.02	.00	.18

Note: Correlations above .03 in absolute value are significant, $p < .05$. Sample sizes for these analyses ranged from 909 (Overall Performance) to 4,726 (Recruiter Satisfaction).
[a] Fit = Recruiter Fit, Com. = Army Commitment; Sat. = Recruiter Satisfaction; LD Sat.. = Leadership and Development Satisfaction; Cop = Coping; Res. = Resilience; CWB = Counterproductive workplace behavior; OCB= Organizational Citizenship Behavior; SM CWB = Social Media CWB; Lead = Leadership; Mot. = Leadership Motivation; Disc. = Disciplinary Incidents; Sup. = Supervisor Ratings; Peer = Peer Ratings; OP = Overall Performance.

Table 8. Correlations Between the ARC Scales and Each Criterion in the Sample

Variables	Fit	Com.	Sat.	LD Sat.	Cop.	Res.	CWB	OCB	SM CWB	Lead	Mot.	Disc.	Sup.	Peer	OP
Aggression	-.16	-.28	-.22	-.23	-.14	-.24	.18	-.21	.20	-.22	-.21	.05	-.09	-.02	-.24
Empathy	.20	.31	.24	.24	.05	.38	-.19	.32	-.20	.34	.36	-.02	.06	.05	.33
Goal Orientation	.12	.18	.14	.16	.09	.28	-.06	.19	-.04	.23	.27	-.08	.12	.10	.26
Hostility	-.15	-.23	-.26	-.33	-.37	-.17	.05	-.08	.04	-.08	-.04	.08	-.15	-.06	-.27
Peer Leadership	.23	.27	.23	.19	.08	.45	-.12	.33	-.12	.40	.46	-.02	.16	.18	.41
Power	.05	-.02	.02	.02	.01	-.05	.10	-.05	.14	-.01	-.05	-.01	.00	.05	.03
Self-Efficacy	.20	.28	.24	.24	.09	.54	-.18	.34	-.22	.40	.47	-.05	.13	.13	.40
Social Maturity	.11	.22	.16	.18	.06	.26	-.22	.22	-.27	.23	.25	-.04	.05	-.01	.18
Work Motivation	.23	.36	.31	.33	.13	.59	-.20	.38	-.23	.42	.49	-.05	.12	.10	.46

Note: Correlations above .03 in absolute value are significant, $p < .05$. Sample sizes for these analyses ranged from 894 (Overall Performance) to 4,712 (Recruiter Fit).

[a] Fit = Recruiter Fit, Com. = Army Commitment; Sat. = Recruiter Satisfaction; LD Sat.. = Leadership and Development Satisfaction; Cop = Coping; Res. = Resilience; CWB = Counterproductive workplace behavior; OCB= Organizational Citizenship Behavior; SM CWB = Social Media CWB; Lead = Leadership; Mot. = Leadership Motivation; Disc. = Disciplinary Incidents; Sup. = Supervisor Ratings; Peer = Peer Ratings; OP = Overall Performance.

Next we examined the validity of the NSAB scales using regression analysis. Table 9 shows the NSAB scales that were significant predictors of each criterion. However, due to potential operational use as a selection screen, the regressions weights for these analyses have been excluded from this table. Instead, X is shown for scales that were significant positive predictors of each criterion and –X is shown for scales that were significant negative predictors[2]. Again, across all criteria, there were a number of NSAB scales that were consistent predictors of these outcomes. For example, the Achievement facet was significantly related to all of the outcomes examined here. The Order dimension was also related to most criteria (often in a negative direction). Not surprisingly, given the social nature of recruiting duty, Sociability was also a consistent predictor of Recruiter attitudes and performance.

Although a number of NSAB scales were significantly related to various criteria, the adjusted multiple R's indicated that the combination of scales was a much better predictor of these outcomes than any individual scale. The adjusted (for capitalization on chance) multiple R's ranged from .20 to .47 across each of the individual criteria in the sample. Some of the strongest relationships were with attitudinal and motivational outcomes like Army commitment (R = .40), leadership motivation (R = .45), and resilience (R = .47). In addition, the NSAB scales were also important predictors of the overall performance composite with an adjusted multiple R of .55. These results indicate that the NSAB scales have utility for identifying Soldiers who will be successful Recruiters.

---

[2] Due to potential operational use of NSAB as a selection screen for Recruiters, the full results are not displayed here but can be made provided by ARI. Please contact Dr. Tonia Heffner at tonia.s.heffner.civ@mail.mil.

Table 9. NSAB Scales that Significantly Predict Recruiter Outcomes

Variables	Fit	Com.	Sat.	LD Sat.	Cop.	Res.	CWB	OCB	SM CWB	Lead	Mot.	Disc.[b]	Sup.	Peer[c]	OP
Achievement	X	X	X	X	X	X	-X	X	-X	X	X	-X	X	X	X
Dominance	X					X				X	X				
Even Tempered	X	X	X	X		X		X		X		-X			X
Humility	-X	-X	-X	-X	-X		-X		-X	-X	-X				
Intellectual Efficiency	-X	-X	-X	-X	-X		X						X		-X
Machiavellianism															
Non-Delinquency	X	X	X	X		-X					-X				
Optimism	X	X	X	X	X	X						-X	X	X	X
Order	-X	-X	-X	-X		-X	X	-X	X	-X	-X	-X	X		
Persistence					X	X		X	X	X	X			X	
Physical Conditioning		X				X	-X	X	-X	X	X				
Responsibility		-X			X	X						-X		X	
Self-Efficacy					X	X	-X	X	-X	X	X			X	
Selflessness	X	X	X	X	X	X			X		X	X			X
Sociability	X	X	X	X	X	X		X	X	X			X	X	X
Virtue		X		X		X	-X	X	-X	X	X				
Multiple R	.34	.40	.40	.36	.33	.47	.20	.34	.22	.36	.44	.21	.21	.21	.57
Adjusted Multiple R	.34	.40	.39	.36	.33	.46	.19	.33	.21	.35	.44	--	.19	.19	.55

Note: X indicates a significant (p < .05) positive relationship between the NSAB scale and the criterion while –X indicates a significant negative relationship. Sample sizes for these analyses ranged from 909 (Overall Performance) to 4,726 (Recruiter Satisfaction).

[a] Fit = Recruiter Fit, Com. = Army Commitment; Sat. = Recruiter Satisfaction; LD Sat.. = Leadership and Development Satisfaction; Cop = Coping; Res. = Resilience; CWB = Counterproductive workplace behavior; OCB= Organizational Citizenship Behavior; SM CWB = Social Media CWB; Lead = Leadership; Mot. = Leadership Motivation; Disc. = Disciplinary Incidents; Sup. = Supervisor Ratings; Peer = Peer Ratings; OP = Overall Performance.

[b] These results are based on a logistic regression analysis.

[c] The peer ratings were first screened to make sure that raters were familiar with ratees' performance. The results for this outcome are based on those who reported that they were "Very Familiar" with the ratee's performance.

Figures 2 and 3 illustrate the practical importance of the relationships between the NSAB and several of the criteria assessed in this sample. We used the standardized regression weights from the analyses shown in Table 9 for predicting the overall performance criterion to calculate NSAB composite scores for each individual. We then used these scores to plot the relationships between this NSAB composite and several criteria. First, Figure 2 illustrates the relationship between the NSAB composite and the overall performance criterion. On the X-axis of this plot are the quintiles for the predicted scores from the NSAB composite. On the Y-axis are the overall performance scores. To standardize this graph and the results in Figure 3, the outcomes were scaled to have a mean of 100 and a standard deviation of 20 and the Y-axes for these figures are scaled to range from the mean of the outcome variable +/- 1 standard deviation.

As shown in Figure 2, the NSAB was a strong predictor of overall performance. Individuals scoring in the bottom 20% on the NSAB composite had an average overall performance score of 84 compared to an average score of 116 in the highest scoring group. In other words, individuals in the highest scoring group on the NSAB composite performed over 1.5 standard deviations higher than individuals in the lowest scoring group. These results provide evidence for the validity of the NSAB for predicting performance in recruiting duty assignments.
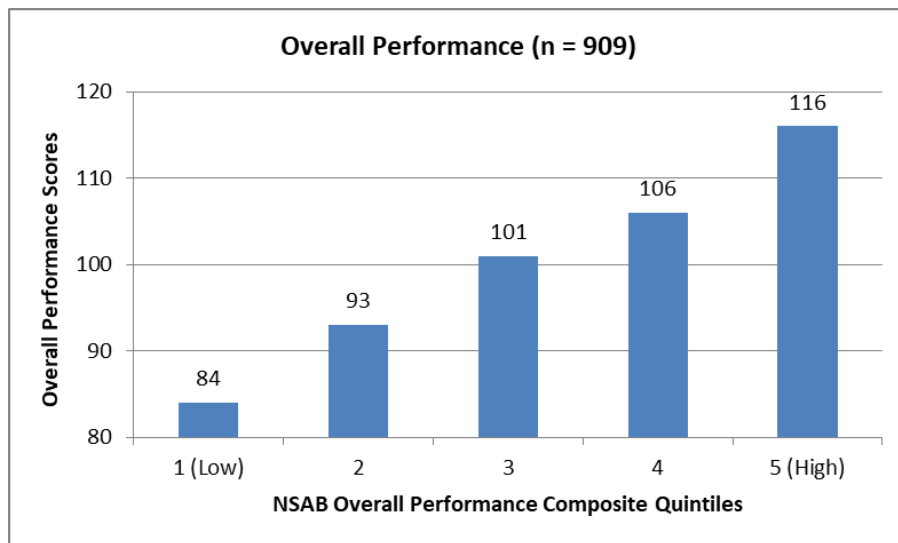
Figure 2. Quintile Plot of the Relationship between the NSAB Composite and Overall Performance

Figure 3 shows the magnitudes of the relationships between the NSAB overall performance composite and perceptions of fit with the Recruiter role, Army commitment, resilience, and leadership. Again, the X-axes for these plots represent the quintiles on the NSAB overall performance composite and the Y-axes provide the average scores on the criteria (scaled with a mean of 100 and an SD of 20). The results indicated that individuals who scored higher on the NSAB composite had better fit with their recruiting assignment and were more committed to the Army, more resilient, and more likely to take on leadership responsibilities. In addition, for most of these outcomes, there was nearly a full standard deviation difference between the highest and lowest scoring groups on the NSAB composite, indicating that the effects were substantial.
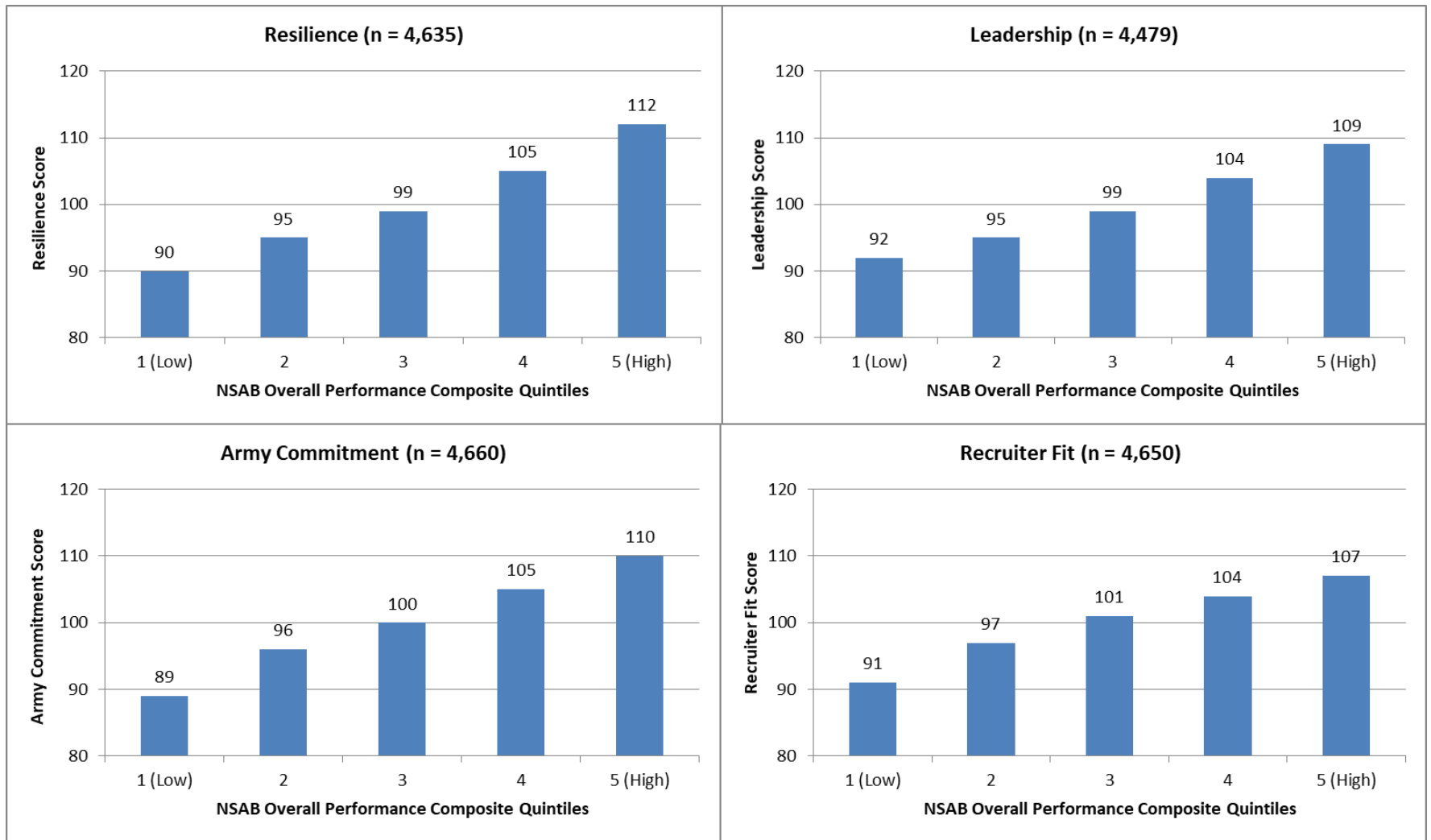
Figure 3. NSAB Composite Quintile Plots for Resilience, Leadership, Army Commitment, and Recruiter Fit

Next, we also examined the incremental validity of the NSAB scales over individuals' Armed Services Vocational Aptitude Battery (ASVAB) General Technical (GT) scores. Figure 4 provides the results from hierarchical regression analyses using both the ASVAB GT scores and all of the NSAB scales (sample sizes ranged from 891 to 4,608). As noted above for these analyses, the ASVAB GT score was included in Step 1 and the NSAB scales were added in Step 2. As expected, the NSAB scales contributed substantial incremental validity for all of the criteria assessed in this research. For nearly all of these outcomes, the validity of GT scores alone was around .10 or below. However, after adding the NSAB scales to the regression equations, the adjusted multiple R's increased substantially. For the majority of these criteria, adding the NSAB scales to the model increased the adjusted multiple R by more than .20. When predicting both resilience and overall performance, the adjusted multiple R increased by more than .40. Thus, the NSAB scales strongly contributed to the prediction of a broad range of criteria even after accounting for GT scores. Similar findings were reported in an earlier investigation of Army Recruiters (White et al., 2002). In that work, ten non-cognitive personality measures, including several ARC scales, were significantly correlated with overall sales performance (avg. r = .22 compared with r = -.02 for ASVAB GT scores).
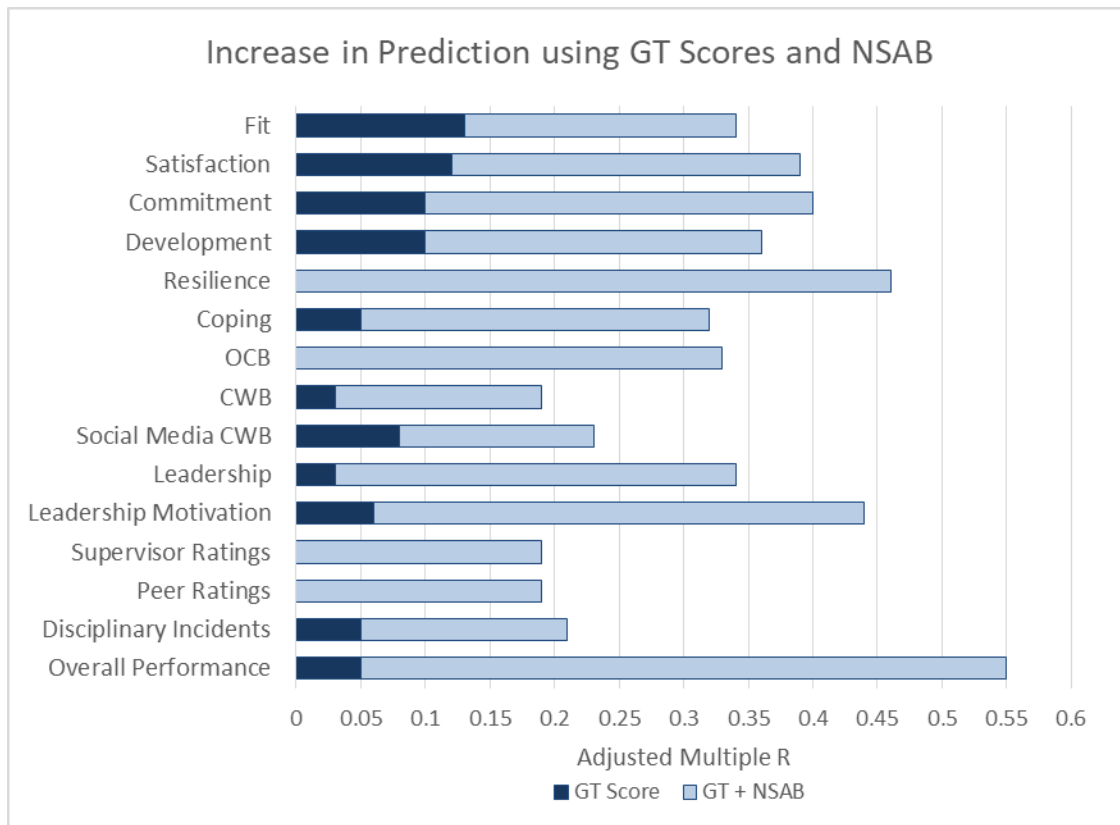
Figure 4. Incremental Validity of the NSAB over the ASVAB GT Scores

We next examined the validity of the combined NSAB and ARC scales. The correlations between these scales are shown in Appendix A and Table 10 shows the significant predictors of each criterion. Again, the regressions weights for these analyses have been excluded from this table due to sensitivity concerns. Instead, X is shown for scales that were significant positive

predictors of each criterion and –X is shown for scales that were significant negative predictors. In Table 10, the adjusted multiple R's for the NSAB scales are from the model with only NSAB in the model. In contrast, the adjusted multiple R's in the bottom row of the table represent the overall validity when both NSAB and ARC are included in the model. As shown here, both NSAB and ARC contributed to the prediction of these criteria. Although the NSAB provided sizeable validity alone, adding the ARC scales to the model also increased prediction. In some cases, the overall validity with both NSAB and ARC was substantial. The multiple R's for both resilience and overall performance were above .60. The multiple R's for four other criteria were above .50. Again, these results indicate that the combination of the NSAB and ARC can provide stronger prediction than either set of scales alone.

Table 10. NSAB and ARC Scales that were Significant Predictors of Recruiter Outcomes

| Variables	Criteria[a]														
	Fit	Com.	Sat.	LD Sat.	Cop.	Res.	CWB	OCB	SM CWB	Lead	Mot.	Disc.[b]	Sup.	Peer	OP
Achievement	X	X	X	X	X		-X	X		X		-X	X	X	X
Dominance											X				
Even Tempered	X	X	X	X	X	X	X				-X				
Humility							-X		-X	-X					
Intellectual Efficiency	-X	-X	-X	-X			X			-X			X		-X
Machiavellianism															
Non-Delinquency	X	X	X	X											
Optimism	X	X	X	X	X	X					-X	-X			X
Order											-X	-X	X	X	
Persistence					X				X	X				X	
Physical Conditioning	-X			-X	-X		-X		-X		X			-X	
Responsibility					X	X						-X		X	
Self-Efficacy		-X	-X								X				
Selflessness	X	X	X		X				X						
Sociability	X	X	X	X	X	X			X	X					X
Virtue						X			-X		X		-X	-X	
Adjusted Multiple R	.34	.40	.39	.36	.33	.46	.18	.33	.21	.35	.44	.21	.19	.19	.56
Aggression	-X	-X	-X	-X	-X		X	-X	X	-X	-X	X			-X
Empathy	X	X	X	X		X	-X	X		X	X				X
Goal Orientation									X			-X	X		X
Hostility	-X	-X	-X	-X	-X	-X						X	-X	-X	-X
Peer Leadership	X	X				X		X		X	X	X	X	X	
Power	X	X	X	X			-X				-X				
Self-Efficacy							X	-X		-X	X	X			
Social Maturity			-X		-X	X	-X	X	-X	X	X				
Work Motivation	X	X	X	X	X	X	-X	X	-X	X	X		-X	-X	X
Adjusted Multiple R	.40	.51	.49	.52	.46	.65	.27	.45	.36	.50	.58	.25	.27	.24	.70

Note: Sample sizes for these analyses ranged from 894 (Overall Performance) to 4,712 (Recruiter Fit). The results in this table represent the scales that were significant predictors of each outcome when both NSAB and ARC scales are included in the model. However, the adjusted multiple R's for the NSAB scales are from the model with only NSAB included. The peer ratings were first screened to make sure that raters were familiar with ratees' performance. The results for this outcome are based on those who reported that they were "Very Familiar" with the ratee's performance. [a] Fit = Recruiter Fit, Com. = Army Commitment; Sat. = Recruiter Satisfaction; LD Sat.. = Leadership and Development Satisfaction; Cop = Coping; Res. = Resilience; CWB = Counterproductive workplace behavior; OCB= Organizational Citizenship Behavior; SM CWB = Social Media CWB; Lead = Leadership; Mot. = Leadership Motivation; Disc. = Disciplinary Incidents; Sup. = Supervisor Ratings; Peer = Peer Ratings; OP = Overall Performance. [b]These results are based on a logistic regression analysis. Therefore, the unadjusted multiple R is reported. X = positive regression weight. –X = negative regressions weight.

Figures 5 and 6 illustrate the practical importance of the relationships shown in Table 10. Consistent with the previous figures, the X-axes for these plots represent the quintiles on the combined composite of NSAB and ARC scales for predicting the overall performance criterion. In addition, the Y-axes provide the average scores on each criterion (scaled with a mean of 100 and an SD of 20). As shown in Figure 5, the combined NSAB and ARC had a strong relationship with the overall performance criterion. For comparison with Figure 3, the relationships with the same criteria are illustrated in Figure 6 for the combined NSAB and ARC. Figure 6 shows even stronger relations than shown in Figure 3 for NSAB alone. These figures confirm the potential utility of combining both the NSAB and the ARC to predict Recruiter outcomes.
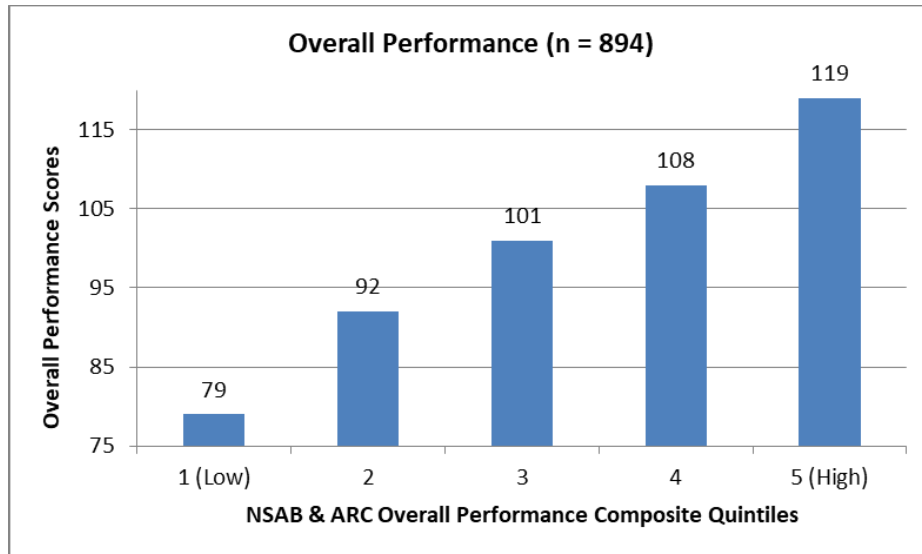
Figure 5. Quintile Plot of the Relationship between a Composite of NSAB and ARC Scales and Overall Performance
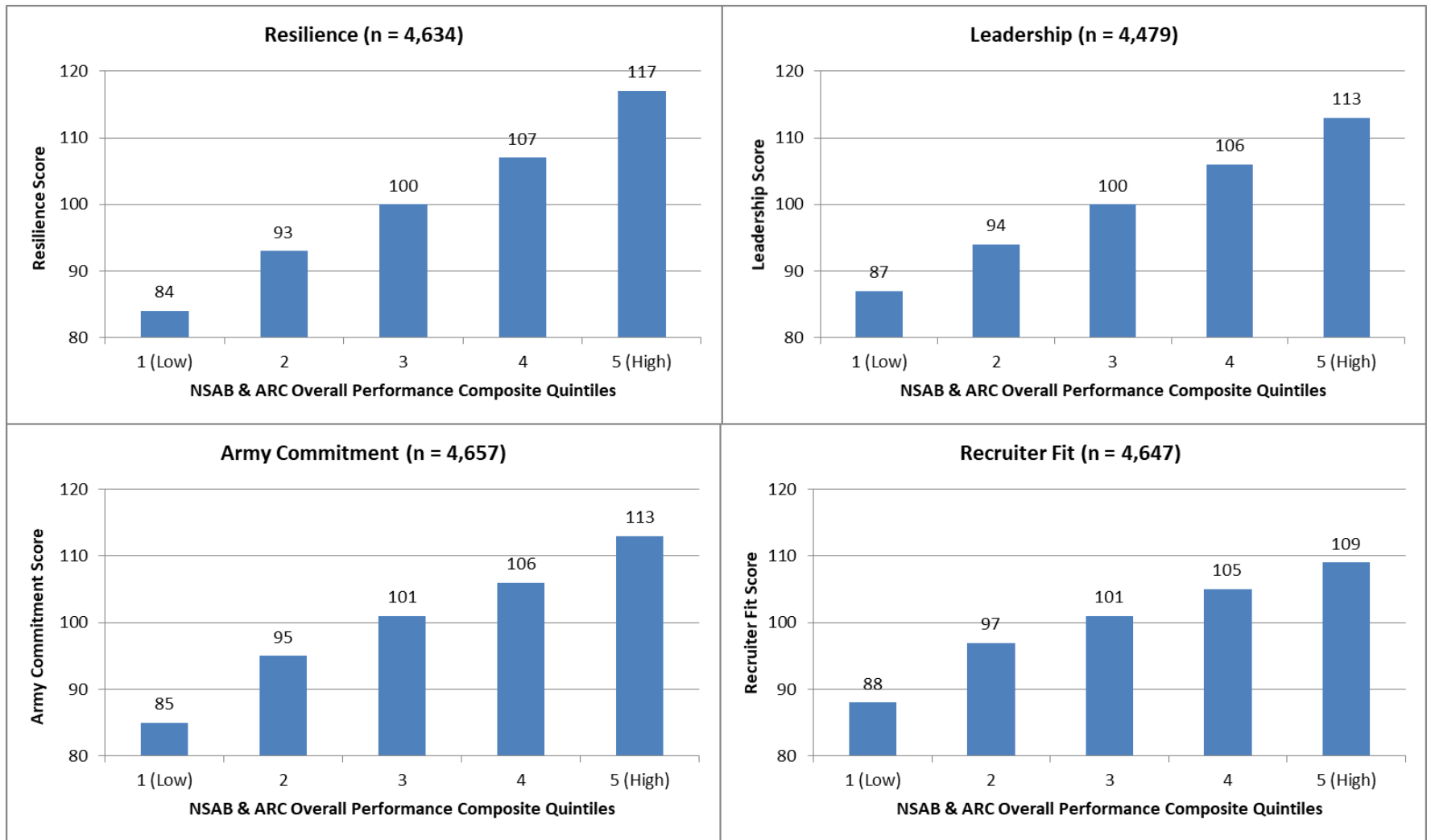
Figure 6. NSAB and ARC Composite Quintile Plots for Resilience, Leadership, Army Commitment, and Recruiter Fit

Finally, we also examined the predictive validity of the NSAB and ARC in the subset of the sample from MOS 79R. Given that these individuals are Career Recruiters, rather than on temporary assignment (i.e., detailed Recruiters), it may be particularly useful to understand their attitudes and behaviors. Therefore, the NSAB and ARC scales that were significant predictors of each criterion in the sample of individuals from MOS 79R are shown in Table 11. Again, the regression weights for this table have been removed due to sensitivity concerns and X represents a positive relationship while –X represents a negative relationship. This table can be compared with Table 10 to determine differences between the overall sample of Recruiters and those in this MOS. However, due to the smaller sample size for these analyses, the non-significant predictors with regression weights greater than .07 are presented along with the significant predictors in Table 11.

With some exceptions, Table 11 shows that the overall validities were slightly higher for those individuals in MOS 79R than in the full sample. However, the differences between the overall sample and individuals in MOS 79R were generally small. In some cases, the multiple R's appear slightly lower than in the overall sample but these differences are primarily due to the larger sample sizes in the full sample. For example, the adjusted multiple R for overall performance was .47 for the NSAB alone in MOS 79R compared to .57 in the full sample. Although this seems like a sizeable reduction, this was primarily due to the fact that the smaller sample size in MOS 79R resulted in a larger adjustment to the multiple R. In fact, the unadjusted multiple R for the NSAB alone was .56, which is consistent with the corresponding value in the full sample (.57). In addition, although the adjusted multiple R for the combined NSAB and ARC was the same in both the full sample and in MOS 79R, the unadjusted multiple R was .77 in MOS 79R. In other words, the overall validity was actually higher in MOS 79R but the smaller sample size resulted in a larger adjustment for capitalization on chance and lowered the value presented in Table 11. This did not appear to be the case for the prediction of Recruiter fit, which resulted in substantially lower adjusted multiple R's in MOS 79R for the NSAB and ARC combined. However, this lower value was most likely due to restriction of range because the individuals in the sample of MOS 79R are likely to fit much better in a recruiting position than individuals in the full sample who may or may not have volunteered for this assignment.

Despite the differences described above, the results presented in Table 11 clearly indicate that the NSAB and ARC provide validity for predicting Recruiter outcomes in MOS 79R. The majority of the adjusted multiple R's for the full model are above .30, indicating substantial validity. In addition, the largest effect was observed for predicting overall performance. Thus, these results suggest that the combination of the NSAB and the ARC would be useful for screening individuals into MOS 79R.

Table 11. Predictors of Recruiter Criteria in MOS 79R

|	Criteria[a]														
Variables	**Fit**	**Com.**	**Sat.**	**LD Sat.**	**Cop.**	**Res.**	**CWB**	**OCB**	**SM CWB**	**Lead**	**Mot.**	**Disc.[b]**	**Sup.**	**Peer**	**OP**
Achievement													X	X	
Dominance											X				
Even Tempered						X									-X
Humility									-X						
Intellectual Efficiency			-X				X			-X		X	X	X	
Machiavellianism										-X		-X			X
Non-Delinquency														-X	X
Optimism		X		X									X		X
Order												-X		X	
Persistence					X					X					
Physical Conditioning							-X								-X
Responsibility			X			X						-X		X	
Self-Efficacy								X		X	X	X	-X		
Selflessness															X
Sociability	X		X									-X		-X	
Virtue		X						X	-X		X	-X			
Adjusted Multiple R	.18	.37	.35	.35	.32	.48	.18	.36	.17	.34	.46	.25	.20	.23	.46
Aggression		-X						-X			-X	X	-X		-X
Empathy		X		X		X					X	X		-X	X
Goal Orientation											X	-X			X
Hostility		-X	-X	-X	-X	-X						X	-X	-X	-X
Peer Leadership		-X	-X	-X	-X			X		X	X	X	X	X	-X
Power	X	X	X	X									-X		
Self-Efficacy		X		X		X				X		-X			
Social Maturity							-X		-X	X	X	X			X
Work Motivation	X	X	X	X	X	X	-X	X	-X	X	X	-X		-X	X
Adjusted Multiple R	.23	.48	.48	.54	.46	.68	.27	.47	.33	.51	.59	.32	.32	.25	.70

Note: Sample sizes for these analyses ranged from 150 (Overall Performance) to 994 (Recruiter Satisfaction). Due to the smaller sample sizes for these analyses, predictors with regression weights less than .07 are not reported in this table to improve clarity. The adjusted multiple R's for the NSAB scales are from the model with only NSAB included. The peer ratings were first screened to make sure that raters were familiar with ratees' performance. The results for this outcome are based on those who reported that they were "Very Familiar" with the ratee's performance. [a] Fit = Recruiter Fit, Com. = Army Commitment; Sat. = Recruiter Satisfaction; LD Sat.. = Leadership and Development Satisfaction; Cop = Coping; Res. = Resilience; CWB = Counterproductive workplace behavior; OCB= Organizational Citizenship Behavior; SM CWB = Social Media CWB; Lead = Leadership; Mot. = Leadership Motivation; Disc. = Disciplinary Incidents; Sup. = Supervisor Ratings; Peer = Peer Ratings; OP = Overall Performance. [b] These results are based on a logistic regression analysis. Therefore, the unadjusted multiple R is reported. X = positive regression weight. –X = negative regressions weight.

Next, we also created quintile plots for comparison with Figures 5 and 6. In Figures 7 and 8, quintile plots are shown for many of the same criteria as above for individuals in MOS 79R. However, because of the smaller relationship with Recruiter fit shown in Table 11, the quintile plot for Recruiter fit was replaced with a plot for satisfaction. Again, the X-axes for these plots represent the quintiles on the combined composite of NSAB and ARC scales for predicting the overall performance criterion in MOS 79R and the Y-axes represent the average scores on each criterion. As shown in Figures 7 and 8, the relationships in MOS 79R largely replicated the results in the overall sample. Specifically, individuals in MOS 79R who scored high on a composite of NSAB and ARC scales were more resilient, more committed to the Army, more satisfied with their role as a Recruiter, and more likely to be a leader. In addition, individuals who scored high on this same composite of NSAB and ARC scales also scored higher on the overall performance criterion. In other words, the combined NSAB and ARC are useful for differentiating high performing Recruiters in MOS 79R from those who are less successful.
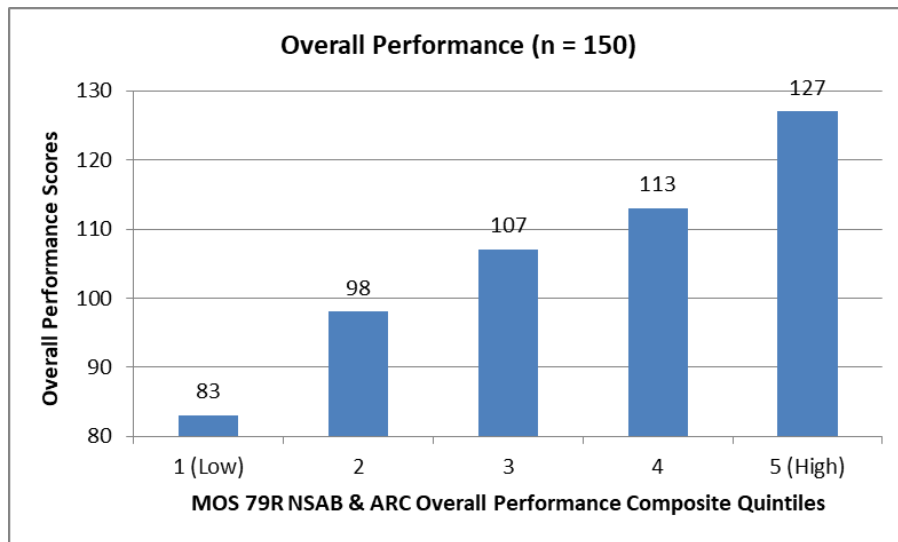
Figure 7. Quintile Plot of the Relationship between a Composite of NSAB and ARC Scales and Overall Performance in MOS 79R
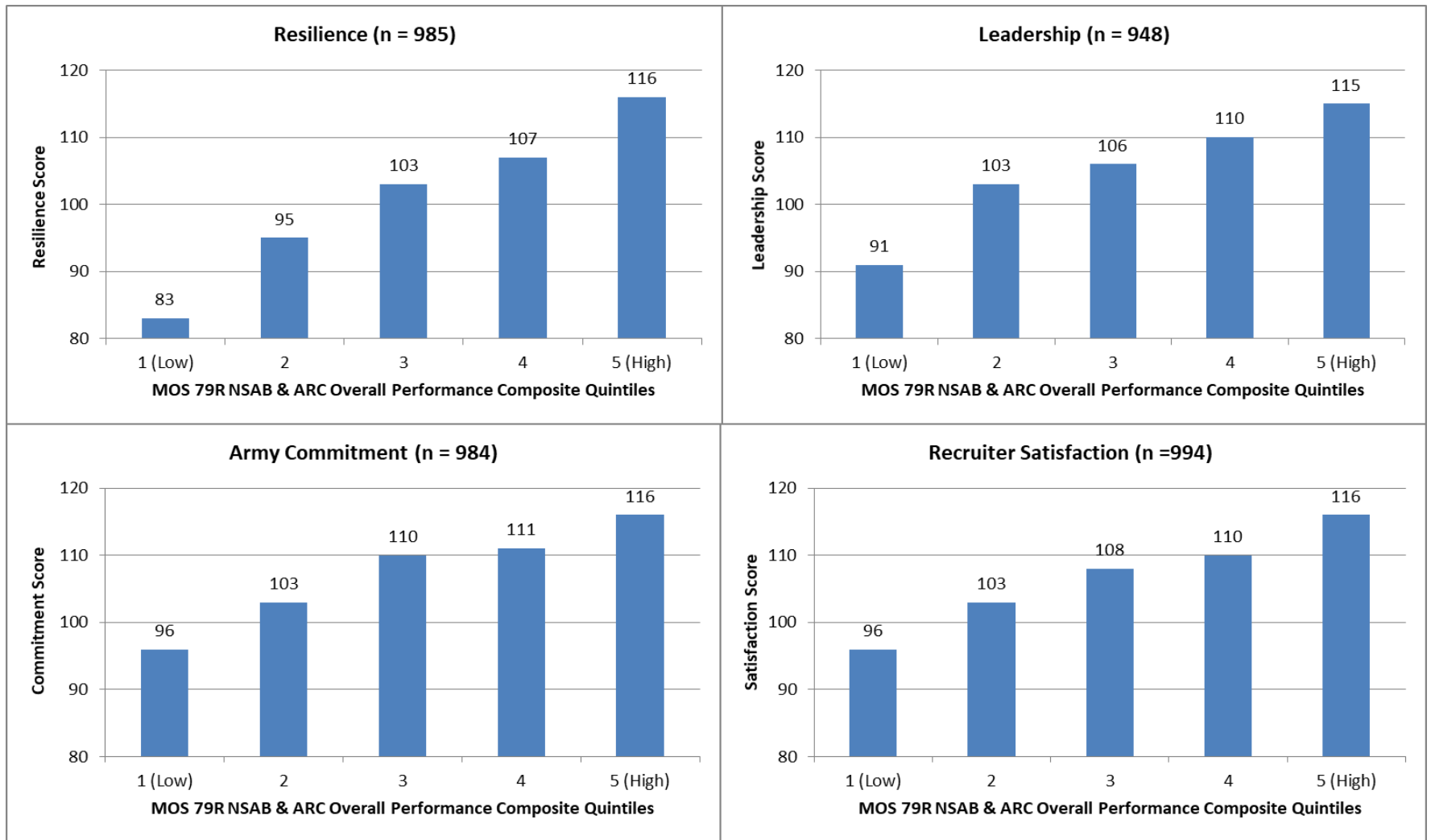
Figure 8. NSAB and ARC Composite Quintile Plots for Resilience, Leadership, Army Commitment, and Satisfaction in MOS 79R

CONCLUSIONS AND NEXT STEPS

Past research has demonstrated that the NSAB can be a useful predictor of important outcomes in a broad range of military occupations (Nye et al., 2012). The results of this research expand that work and indicate that the NSAB scales are strong predictors of Recruiter attitudes and performance as well. These scales had substantial relationships with a number of Recruiter outcomes and added incremental validity to the prediction of these outcomes over ASVAB GT scores. Importantly, the NSAB also contributed substantially to the prediction of overall performance. A composite of NSAB scales had an adjusted (for capitalization on chance) multiple R of .55 for predicting this broad outcome. In addition, this composite of NSAB scales was highly correlated with most of the criteria examined here. Consequently, the results presented here indicate that composites of NSAB scales would be useful for identifying high potential Soldiers for recruiting duty assignments.

The magnitudes of the relationships found in the present research add to previous research on the validity of the NSAB and suggest that this measure will be useful for in-service testing. In fact, the relationships between the NSAB and the outcomes assessed in this sample were generally larger than found in other military occupations. Other research examining the prediction of similar performance criteria has generally found validities ranging from .24 to .36 across multiple jobs (Nye et al., 2012). In contrast, the adjusted multiple R's found here ranged from .19 to .46 for specific criteria and the majority of these values were above .30. In addition, the adjusted multiple R for predicting the overall performance composite examined in the present research was .55, which is substantially larger than in previous research. This is consistent with other research examining in-service testing (Horgen et al., 2013; Nye, Beal, Drasgow, Dressel, White, & Stark, 2014) and suggests that the NSAB may be particularly useful for selecting individuals who are already serving in the military for special duty assignments.

Despite the strong validity of the NSAB alone, the results also indicated that the ARC could contribute to the prediction of Recruiter outcomes above and beyond the NSAB. This suggests that although the NSAB assesses a wide range of characteristics, there are other characteristics that are not assessed by this measure but that are related to Recruiter attitudes and performance. The ARC was developed specifically to predict counterproductive work behavior (Kilcullen et al., 2003) but the results of the present research suggest that these scales can predict broader attitudes and performance criteria as well. As a result, the combination of the ARC and the NSAB provided the best overall prediction for virtually all criteria.

Overall, the results suggest that the NSAB and ARC are promising predictors of Recruiter outcomes and may be useful as in-service screening tools for this assignment. However, future research is necessary to provide further evidence of the validity and utility of these measures for this purpose. First, the present research project was conducted concurrently with both predictor and criterion data collected simultaneously. Therefore, future research would benefit from examining these same relationships in a longitudinal research project where individuals are administered the NSAB and ARC prior to beginning their recruiting role and the criteria are assessed after they have gained experience on the job. This research design would provide additional evidence for the validity of these measures while also demonstrating that the validity is maintained over time and is not due to the common method variance that can result from responding to all self-report measures at a single time point.

Another useful future direction for this research would be to examine the validity of these measures under operational conditions. As noted above, a key concern with personality measures is faking in high-stakes settings. Therefore, it is important to demonstrate the validity of these measures even when individuals are motivated to distort their responses and inflate (or deflate) their scores. Again, both the NSAB and the ARC were designed to be resistant to faking. The NSAB items are administered in a pairwise preference format that has demonstrated validity and negligible score inflation under operational conditions (Stark, Chernyshenko, Drasgow, Nye, White, Heffner, & Farmer, 2014). In addition, the ARC incorporates a social desirability scale that can be used to control for faking. Nevertheless, examining these measures under operational conditions is important to demonstrate that these measures maintain their utility for in-service testing in high-stakes contexts.

Finally, future research to examine the prediction of more objective criteria would also be useful. In the present research, the NSAB and ARC predicted a broad range of criteria that included both self-ratings of attitudes and behavior as well as peer and supervisor ratings of Recruiters' performance. Although these outcomes provide useful information about the utility of the NSAB and ARC, it would also be useful to examine their validity for predicting more objective outcomes as well. Self-reports can sometimes be inflated due to socially desirable responding and both peer and supervisor ratings can suffer from rater biases. Therefore, examining the prediction of objective criteria will provide an additional source of evidence for the validity of these measures. Past research has found that the NSAB scales have validity for predicting objective criteria like attrition, training success, or training failure (Nye et al., 2012; Nye et al., 2013). As such, similar research would be beneficial with a sample of Recruiters.

Despite potential directions for future research, the results presented here suggest that both the NSAB and the ARC are promising predictors of Recruiter success and can predict a broad range of outcomes. Importantly, these measures also predicted an overall performance variable, indicating that they may be useful for identifying high potential individuals for recruiting assignments. Recruiters play an important role in building and maintaining a strong military force. Therefore, identifying successful Recruiters is particularly important and the results of the present research suggest that the NSAB and ARC can help to facilitate this effort for future recruiting assignments.

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). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44,* 1-26.

Campbell, J.P., & Knapp, D.J. (Eds.) (2001). *Exploring the limits in personnel selection and classification*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Drasgow, F., Levine, M.V. & Williams, E. (1985). Appropriateness measurement with polychotomous item response models and standardized indices. *British Journal of Mathematical and Statistical Psychology, 38*, 67-86.

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 selection and classification decisions* (Technical Report 1311). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist, 48*, 26-34.

Halstead, J. B. (2009). Recruiter selection model and implementation within the U.S. Army. *IEEE Transactions on System, Man, and Cybernetics, 19*, 88-106.

Horgen, K. E., Kubisiak, U. C., Lee, V. B., Connell, P. W., Penney, L. M., Borman, W. C., Pace, V., Lentz, E., White, L. A., Young, M. C., & Bowles, S. V. (2006). *Evaluation and refinement of a screening tool for U.S. Army Recruiters: Noncommissioned Officer Leadership Skills Inventory* (Technical Report 1177). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Horgen, K. E., Nye, C. D., Drasgow, F., Chernyshenko, O. S., Stark, S., Owens, K., & Fletcher, K. (in preparation). *Validation of the Non-Commissioned Officer Special Assignment Battery (NSAB) for Army instructor assignments.* Fort Belvoir, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Horgen, K. E., Nye, C. D., White, L. A., Laporte, K. A., Hoffman, R. R., Drasgow, F., Chernyshenko, O. S., Stark, S., & Conway, J. S. (2013). *Validation of the Non-Commissioned Officer Special Assignment Battery (NSAB)* (Technical Report 1328). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Judge, T. A., Rodell, J. B., Klinger, R. L., Simon, L. S., & Crawford, E. R. (2013). Hierarchical representations of the five-factor model of personality in predicting job performance:

integrating three organizing frameworks with two theoretical perspectives. *Journal of Applied Psychology*, *98*, 875-925.

Kilcullen, R. N., White, L. A., Sanders, M., & Hazlett, G. (2003). *Assessment of Right Conduct (ARC) administrator's manual* (Research Note 2003-09). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Kilcullen, R.N., Gluszek, A.M., Beeco, R., Repchick, K. & Swigart, K. L. (In-preparation). *Developing an Enhanced Suitability Screen for Positions of Significant Trust and Authority* (Research Note). Fort. Belvoir, 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 evaluation* (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.). (2012). *Tier One Performance Screen initial operational test and evaluation: 2011 interim report* (Technical Report 1306). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Knapp, D. J., Owens, K. S., & Allen, M. T. (Eds.). (2011). *Validating future force performance measures (Army Class): First in-unit performance longitudinal evaluation* (Technical Report 1293). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Mueller-Hanson, R. Heggestad, E. D., & Thornton, G. C. (2003). Faking and Selection: Considering the Use of Personality from a Select-In and Select-Out Perspective. *Journal of Applied Psychology, 88*, 348-355.

Nye, C. D., Beal, S. A., Chernyshenko, O. S., Stark, S., Drasgow, F., White, L. A., & Dressel, J. D. (2014). *Assessing the Tailored Adaptive Personality Assessment System (TAPAS) for Special Forces personnel* (Research Report 1971). Fort. Belvoir, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Nye, C. D., Drasgow, F., Chernyshenko, O. S., Stark, S., Kubisiak, C., White, L. A., & Jose, I. (2012). *Assessing the Tailored Adaptive Personality Assessment System (TAPAS) as an MOS qualification instrument* (Technical Report 1312). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Nye, C. D., Drasgow, F., Chernyshenko, O. S., Stark, S., Muhammad, R. S., & Wolters, H. M. (2017). *New scale development for enhanced suitability screening (ESS)* (Research Note). Fort. Belvoir, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*, 262-274.

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*, 138-152.

Stark, S., Chernyshenko, O.S., Drasgow, F., & White, L.A. (2012). Adaptive testing with multidimensional pairwise preference items: Improving the efficiency of personality and other noncognitive assessments. *Organizational Research Methods, 15,* 463-487.

Stark, S., Chernyshenko, O. S., Nye, C. D., Drasgow, F., & White, L. A. (2017). *Moderators of the Tailored Adaptive Personality Assessment System validity* (Technical Report 1357). Fort. Belvoir, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

U.S. Army (2017). *Army is hiring: Army increases end strength by 28,000 Soldiers*. Retrieved September 18, 2017 from https://www.army.mil/article/184431/army_is_hiring_army_increases_end_strength_by_28000_soldiers.

U.S. Army Recruiting Command (2011). *Recruiter Handbook* (USAREC Manual 3-01). Fort Knox, KY: Author.

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* (pp. 89-104) (Technical Report 1168). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

White, L. A., Borman, W. C., & Bowles, S. (2001, August). *New tools for selecting high potential Army Recruiters*. Paper presented at the 109th Annual Convention of the American Psychological Association. San Francisco, CA.

White, L. A., Borman, W. C., Penney, L., Kubisiak, C, Horgen, K., Bowles, S., & Mills, S. (2002, August). *Concurrent validation of new measures for predicting Army Recruiter sales performance.* Paper presented at the Annual Convention of the American Psychological Association. Chicago, IL.

White, L. A., Young, M. C., Hunter, A. E., & Rumsey, M. G. (2008). Lessons learned in transitioning personality measures from research to operational settings. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *1*(*3*), 291-295.

White, L. A., Young, M. C., & Rumsey, M.G. (2001). ABLE implementation issues and related research. In J. P. Campbell & D.J. Knapp (Eds.), *Exploring the limits in personnel selection and classification.* (pp. 525-558). Hillsdale, NJ: Erlbaum.

White, L.A., Young, M.C., Penny, L.M., Kubisiak, U.C., Horgen, K.E., Lee, V.B., Connell, P.W., Borman, W.C., & Bowles, S.V. (2004). *U.S. Army Recruiter selection research: Test validation process.* Paper presented at the Annual Conference of the American Psychological Association, Honolulu, HI.

White, L. A., Young, M. C., & Rumsey, M.G. (2001). ABLE implementation issues and related research. In J. P. Campbell & D.J. Knapp (Eds.), *Exploring the limits in personnel selection and classification.* (pp. 525-558). Hillsdale, NJ: Erlbaum.

Table A1. Correlations Between the NSAB and ARC Scales

NSAB Scales	ARC Scales								
	Agg.	Emp.	Goal Ori.	Soc. Mat.	Host.	Peer Lead.	Pow.	Self Eff.	Work Mot.
Achievement	-.12	.19	.32	.15	-.10	.35	.06	.36	.40
Dominance	.00	.07	.34	.01	-.01	.39	.23	.34	.32
Even Tempered	-.26	.14	-.01	.18	-.22	-.01	-.16	.05	.08
Humility	-.05	.09	-.17	.10	-.01	-.20	-.32	-.14	-.11
Intellectual Efficiency	.01	.02	.17	-.02	.03	.32	.10	.35	.25
Machiavellian	.19	-.23	.06	-.24	.12	.10	.37	.04	-.02
Non-Delinquency	-.12	.07	.05	.18	-.11	-.03	-.09	-.01	.05
Optimism	-.14	.07	.14	.11	-.21	.19	.03	.25	.21
Order	.04	-.06	.03	-.06	.05	-.02	.08	-.03	-.01
Persistence	-.06	.15	.27	.09	-.03	.33	.06	.35	.37
Physical Conditioning	-.16	.25	.18	.25	-.08	.25	-.10	.28	.33
Responsibility	.06	.01	.21	-.01	-.06	.13	.12	.15	.18
Self-Efficacy	-.06	.09	.29	.06	-.03	.40	.12	.41	.36
Selflessness	-.18	.39	.05	.16	-.13	.10	-.17	.09	.16
Sociability	-.03	.09	.15	-.01	-.09	.27	.18	.18	.16
Virtue	-.23	.31	.11	.32	-.09	.15	-.23	.19	.26

Note: All correlations above .02 in absolute value are significant, $p < .05$.

Descriptive statistics for the NSAB and ARC scales across subgroups are shown in Table A2 in the appendix. The effect sizes (Cohen's d) of these differences are also shown. Results indicate that differences across males and females are relatively small. Most of the effect sizes shown in Table A2 comparing NSAB scores across males and females were .20 or below. However, the effect sizes also indicated that males scored moderately higher on Machiavellianism ($d = .39$) and females scored slightly higher on Selflessness ($d = -.45$). In addition, there were also moderate differences between males and females on the ARC scales. Males scored higher on Aggression ($d = .47$) while females scored higher on Empathy ($d = -.45$) and Social Maturity ($d = -.40$). Similarly, differences were generally small when comparing scores across White and Black or Hispanic Soldiers. The largest effect size (in absolute value) of any of these differences was for comparisons of Selflessness across White and Black respondents, which indicated that Black respondents scored slightly higher on this dimension ($d = -.35$). In general, these results suggest that there were not substantial subgroup differences on the NSAB and ARC dimensions.

Table A2. Subgroup Means, Standard Deviations, and Effect Sizes

Variables	Male Mean	Female Mean	Effect Size (Males-Females)	White Mean	Black Mean	Hispanic Mean	Effect Size (White – Black)	Effect Size (White – Hispanic)
Achievement	.24 (1.11)	.31 (1.05)	-.07	.27 (1.13)	.16 (1.05)	.33 (1.09)	.09	-.06
Dominance	.43 (.92)	.30 (.90)	.15	.45 (.93)	.41 (.87)	.44 (.90)	.04	.01
Even Tempered	-.19 (1.02)	-.23 (.99)	.04	-.24 (1.02)	-.08 (.99)	-.11 (.99)	-.15	-.13
Humility	.43 (1.21)	.58 (1.19)	-.13	.42 (1.22)	.46 (1.16)	.49 (1.20)	-.03	-.06
Intellectual Efficiency	.45 (.91)	.28 (.90)	.18	.48 (.91)	.32 (.88)	.35 (.88)	.18	.14
Machiavellianism	.31 (.91)	-.05 (.97)	.39	.31 (.93)	.16 (.87)	.24 (.95)	.16	.08
Non-Delinquency	-.21 (1.09)	-.05 (1.07)	-.14	-.22 (1.10)	-.13 (1.03	-.05 (1.03)	-.08	-.15
Optimism	.17 (1.08)	.21 (1.10)	-.03	.15 (1.09)	.25 (1.04)	.28 (1.04)	-.09	-.12
Order	.32 (.94)	.44 (.94)	-.12	.31 (.97)	.39 (.87)	.41 (.91)	-.09	-.11
Persistence	.01 (.87)	-.07 (.90)	.09	.01 (.89)	-.06 (.84)	.08 (.87)	.08	-.07
Physical Conditioning	-.21 (1.02)	-.24 (.99)	.03	-.18 (1.02)	-.29 (1.01)	-.22 (1.00)	.11	.04
Responsibility	.23 (.99)	-.01 (.91)	.25	.20 (.93)	.20 (.86)	.24 (.83)	.00	-.05
Self-Efficacy	-.02 (1.08)	-.20 (1.11)	.16	-.01 (1.12)	-.09 (1.05)	-.07 (1.10)	.08	.06
Selflessness	.09 (.99)	.53 (.97)	-.45	.05 (1.00)	.40 (.92)	.19 (.94)	-.35	-.13
Sociability	-.04 (1.08)	-.16 (1.05)	.11	-.07 (1.11)	.02 (1.01)	.03 (1.03)	-.08	-.09
Virtue	-.35 (1.14)	-.12 (1.16)	-.20	-.32 (1.16)	-.341.12)	-.34 (1.15)	.01	.02
Aggression	2.19 (.78)	1.82 (.68)	.47	2.16 (.78)	2.14 (.78)	2.08 (.76)	.03	.10
Empathy	3.53 (.56)	3.79 (.60)	-.45	3.53 (.55)	3.67 (.60)	3.62 (.55)	-.25	-.16
Goal Orientation	2.99 (.56)	2.98 (.54)	.01	3.00 (.55)	2.98 (.57)	3.05 (.55)	.03	-.09
Hostility	3.05 (.66)	3.01 (.67)	.06	3.07 (.66)	3.01 (.65)	3.03 (.67)	.09	.06
Peer Leadership	3.51 (.62)	3.46 (.63)	.08	3.52 (.62)	3.47 (.65)	3.53 (.60)	.09	-.01
Power	2.74 (.65)	2.57 (.59)	.26	2.73 (.64)	2.74 (.66)	2.75 (.64)	-.02	-.03
Self-Efficacy	3.92 (.61)	3.90 (.59)	.03	3.94 (.60)	3.87 (.62)	3.96 (.59)	.10	-.04
Social Maturity	3.94 (.64)	4.20 (.53)	-.40	3.96 (.63)	3.98 (.66)	3.98 (.62)	-.03	-.02
Work Motivation	3.86 (.64)	3.90 (.60)	-.06	3.88 (.63)	3.82 (.65)	3.92 (.62)	.10	-.06

Note: Sample sizes for these analyses ranged from 409 (Females) to 4,382 (Males). Values in parentheses represent the standard deviations. Effect sizes are Cohen's d's. Negative effect sizes indicate that Females and Black or Hispanic Soldiers scored higher on the corresponding dimension.