



**Research Note 2023-13**

**Validation of the Adaptive  
Vocational Interest Diagnostic**

**Christopher D. Nye**  
Dragow Consulting Group

**Cristina D. Kirkendall**  
U.S. Army Research Institute

**Fritz Dragow**  
**James Rounds**  
**Oleksandr S. Chernyshenko**  
**Stephen Stark**  
**Tianjun Sun**  
**Bo Zhang**  
Dragow Consulting Group

**September 2023**

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

Approved for public release; distribution is unlimited.

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

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

**Authorized and approved:**

**SCOTT B. SHADRICK, Ph.D.  
Senior Scientific Technical Manager**

---

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

Drasgow Consulting Group

Technical review by

Ashley Wittig, U.S. Army Research Institute

**DISPOSITION**

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

## REPORT DOCUMENTATION PAGE

<b>1. REPORT DATE (Month Year)</b> September 2023		<b>2. REPORT TYPE</b> Interim		<b>3. DATES COVERED (Month Year)</b>	
				<b>START DATE</b> September 2018	<b>END DATE</b> October 2021
<b>4. TITLE AND SUBTITLE</b> Validation of the Adaptive Vocational Interest Diagnostic					
<b>5a. CONTRACT NUMBER</b> W911NF-18-C-0026		<b>5b. GRANT NUMBER</b>		<b>5c. COOPERATIVE AGREEMENT NUMBER</b>	
<b>5d. PROGRAM ELEMENT NUMBER</b> 633007	<b>5e. PROJECT NUMBER</b> A790	<b>5f. TASK NUMBER</b> 1031		<b>5g. WORK UNIT NUMBER</b>	
<b>6. AUTHOR(S)</b> Nye, Christopher D., Kirkendall, Cristina D., Drasgow, Fritz, Rounds, James, Chernyshenko Oleksandr S., Stark, Stephen, Sun, Tianjun, & Zhang, Bo					
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b> Drasgow Consulting Group LLC 3508 N High Cross Rd Urbana, IL 61802-9759				<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>	
<b>9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b> U.S. Army Research Institute for the Behavioral and Social Sciences 6000 6th Street (Bldg. 1464 / Mail Stop: 5610) Fort Belvoir, Virginia 22060-5610			<b>10. SPONSOR/MONITOR'S ACRONYM(S)</b>  ARI	<b>11. SPONSOR/MONITOR'S REPORT NUMBER(S)</b>  Research Note 2023-13	
<b>12. DISTRIBUTION/AVAILABILITY STATEMENT</b> Distribution Statement A: Approved for public release; distribution is unlimited.					
<b>13. SUPPLEMENTARY NOTES</b> ARI Research POC: Dr. Cristina Kirkendall, Selection and Assignment Research Unit					
<b>14. ABSTRACT</b> Initial research with the Adaptive Vocational Interest Diagnostic (AVID) demonstrated that this assessment is a valid predictor of Soldiers' attitudes and performance in several military operational specialties (MOS). However, previous research has only examined a static version of the AVID. Therefore, one goal of the current research was to expand on previous work by conducting a concurrent validation study of a computer-adaptive version of the AVID (CAT AVID). The data for this research were collected from Soldiers in numerous MOS and results indicated that the validity of the CAT AVID is comparable to the validity of the static version. In addition to the validation analyses, three simulations were conducted to explore the optimal ways of operationalizing interest fit. The results of these simulations demonstrated that regression-based composites of AVID scales performed better than the congruence indices that have traditionally been used in the interest literature. In addition, regression-weighted composites also performed as well as or better than other modern prediction methods based on machine learning. Finally, simulations also demonstrated that matching individuals to jobs based on their interests could substantially improve the overall performance of Soldiers. These findings provide additional evidence that the AVID may be useful for MOS assignment.					
<b>15. SUBJECT TERMS</b> Enlisted Personnel, Manpower, Vocational Interest Measures, MOS Assignment					
<b>16. SECURITY CLASSIFICATION OF:</b>			<b>17. LIMITATION OF ABSTRACT</b>	<b>18. NUMBER OF PAGES</b>	
<b>a. REPORT</b> Unclassified	<b>b. ABSTRACT</b> Unclassified	<b>c. THIS PAGE</b> Unclassified	Unlimited Unclassified	48	
<b>19a. NAME OF RESPONSIBLE PERSON</b> Tonia Heffner				<b>19b. PHONE NUMBER (Include area code)</b> 703-545-4408	

**Research Note 2023-13**

**Validation of the Adaptive  
Vocational Interest Diagnostic**

**Christopher D. Nye**  
Drasgow Consulting Group

**Cristina D. Kirkendall**  
U.S. Army Research Institute

**Fritz Drasgow**  
**James Rounds**  
**Oleksandr S. Chernyshenko**  
**Stephen Stark**  
**Tianjun Sun**  
**Bo Zhang**  
Drasgow Consulting Group

**Selection and Assignment Research Unit**  
**Tonia S. Heffner, Chief**

**September 2023**

---

Approved for public release; distribution is unlimited.

# VALIDATION OF THE ADAPTIVE VOCATIONAL INTEREST DIAGNOSTIC (AVID)

## CONTENTS

---

	Page
<b>INTRODUCTION</b> .....	1
Background .....	1
Validity of the AVID.....	4
Purpose of the Current Research .....	5
<b>CONCURRENT VALIDATION OF A COMPUTER-ADAPTIVE VERSION OF AVID</b> ... 6	
Methods .....	6
Data and Procedures .....	6
Measures.....	6
Analyses .....	10
Results .....	12
<b>SIMULATION RESEARCH TO EXAMINE OPERATIONALIZATIONS OF INTEREST FIT</b> .....	18
Simulation Study 1 .....	18
Methods.....	18
Results.....	20
Simulation Study 2 .....	22
Methods.....	22
Results.....	24
Simulation Study 3 .....	30
Methods.....	30
Results.....	31
<b>SUMMARY AND CONCLUSIONS</b> .....	36
<b>REFERENCES</b> .....	38

### LIST OF TABLES

Table 1. AVID Dimensions and Definitions.....	2
Table 2. Army Life Questionnaire (ALQ) Scales Included.....	7
Table 3. Descriptive Statistics and Intercorrelations between the Criteria.....	8

# VALIDATION OF THE ADAPTIVE VOCATIONAL INTEREST DIAGNOSTIC (AVID)

## CONTENTS (Continued)

---

Table 4. Means and Standard Deviations of the MOS Ratings .....	10
Table 5. Correlations between the AVID Basic Interest Dimensions .....	13
Table 6. Standardize Regression Weights for the AVID Dimensions Predicting Each Criterion	14
Table 7. Standardize Regression Weights for the Predictors in the AVID Interest Fit Composite .....	17
Table 8. Relative Variable Importance for the Prediction Models Examined in the Second Simulation Study .....	29
Table 9. Effect Sizes (Cohen's <i>d</i> ) of the Differences in Mean Predicted Performance under the Random and Interest Fit Assignment Conditions.....	34

## LIST OF FIGURES

Figure 1. AVID Composite Quintile Plots for Overall Performance, Motivation to Lead, OCB, and Resilience .....	16
Figure 2. Simulation Results Comparing Polynomial Regression to Traditional Congruence Indices .....	21
Figure 3. Average $R^2$ for Each Model in the Samples of 500.....	26
Figure 4. Average Correlations between Predicted and True Criterion Scores across Samples ..	27
Figure 5. Average Correlations between Predicted and True Criterion Scores in the Test Sample of 10,000 Cases.....	28
Figure 6. Proportions of Individuals Predicted to Perform Better in Each Job across Four Different Samples .....	33
Figure 7. Distributions of Effect Sizes across Samples .....	35

# **Validation of the Adaptive Vocational Interest Diagnostic (AVID)**

## **Introduction**

A growing body of research has shown that vocational interests can be important predictors of both work and academic outcomes. For example, Van Iddekinge, Roth, Putka, and Lanivich (2011) showed that interests were moderately correlated with a number of performance outcomes on the job. In addition, Nye, Su, Rounds, and Drasgow (2017) conducted a comprehensive meta-analysis of the relationship between interests and work performance that summarized over 60 years of research, 92 studies, and 1,858 correlations. These authors found that the match between individuals' interests and the activities performed on their jobs (called congruence in the vocational interest literature) was a strong predictor of performance outcomes including task performance, organizational citizenship behavior, turnover, and training performance, with correlations ranging from .19 to .40. Similar research showed that interest congruence can also predict performance in academic settings (Nye, Su, Rounds, and Drasgow, 2012).

Given this evidence of positive validity, high quality assessments of vocational interests are necessary to realize the benefits of these individual differences in the workplace. To address this need in the U.S. Army, Drasgow Consulting Group (DCG) and the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) developed the Adaptive Vocational Interest Diagnostic (AVID) for use in high-stakes military contexts. This measure has shown promise for predicting a number of important Soldier outcomes, including affective commitment and Army fit (Nye, Rounds, Kirkendall, Drasgow, Chernyshenko, & Stark, 2019). However, more research is needed to examine the validity of the AVID under different conditions. This report describes a set of studies that were designed to provide additional validity evidence for the AVID and to identify the best approach to examining the fit between Soldiers' interests and their military occupational specialty (MOS), or their job in the Army.

## **Background**

The AVID was developed to take advantage of recent psychometric advances and developments in the theory of vocational interests. First, the AVID was designed to assess a broad range of narrow interest dimensions known as basic interests. Holland (1997) proposed the most widely used structure of vocational interests, which includes six broad interest dimensions (see Table 1 for a list of these dimensions) that are collectively known as the RIASEC model. Although many interest measures assess the broader interest dimensions associated with the RIASEC model, these dimensions may be too broad, particularly for differentiating between the various MOS in the Army. In contrast, basic interests are more homogeneous dimensions of interests that group together work activities that are relevant to a number of occupations (Jackson, 1977; Liao, Armstrong, & Rounds, 2008). These narrow interest dimensions are analogous to trait facets in the personality literature, and assessing basic interests can provide both the content specificity and the flexibility required to more accurately select and assign individuals into a broad range of occupations. Therefore, the AVID was developed to assess the 20 basic interest dimensions shown in Table 1. Additional information about how these basic interest dimensions were selected and developed is provided by Nye et al. (2019).

**Table 1.**

***AVID Dimensions and Definitions***

<b>Basic Interests</b>	<b>Activities Associated with Each Dimension</b>	<b>Broader Interest Dimension (Holland, 1997)</b>
Construction	Designing and/or building things or maintaining structures with one's hands or using tools and materials. Includes jobs similar to construction worker, mason, or welder.	Realistic
Protection	Guarding, ensuring safety, and enforcing rules and laws. May include jobs as a law enforcement officer, park ranger, firefighter, or in leadership and management positions in protective service organizations.	
Combat	Operating weapons and equipment in ground combat operations; performing reconnaissance operations; attacking enemy positions and defending friendly posts. Includes jobs in infantry, field artillery, and special forces.	
Physical Activity	Engaging in physical activity, exercise, sports, and games. Includes jobs as a physical trainer, athletic coach, and strength training coach.	
Mechanical	Building, maintaining, repairing, and using small and large machinery, including driving and operating heavy equipment or large vehicles. Includes jobs as mechanics, service repair people, mechanical engineers, factory or laboratory machinists, pilots, boat captains, and truck drivers.	
Electronics	Building, maintaining, repairing, and using electronics including computer hardware and small electronics. Includes jobs as electricians, broadcast technicians, electronic equipment installers and repair people, and electrical engineers.	
Outdoor	Working in the outdoors. Includes jobs such as farmers, forest rangers, veterinarians, zoologists, landscapers, and groundskeepers.	
Writing	Writing factual reports, memos, textbooks, scientific, legal, historical, or technical essays for business and record-keeping purposes. Includes jobs such as lawyers, professors, authors.	Artistic
Teaching	Instructing people inside and outside of school (e.g., teachers and instructors in school, churches, clinics, and welfare agencies). Includes jobs such as training and coaching jobs, child-care assistants, teacher aides, and physical therapy assistants.	Social
Personal Service	Performing everyday tasks for others. Includes jobs such as server, household worker, hair stylist, flight attendant, or hotel concierge.	

**Table 1 (con't). AVID Dimensions and Definitions**

Medical Services	Applying medical knowledge and skills to the diagnosis, prevention, and treatment of disease and injury. Includes jobs such as paramedics, physician's assistant, nursing, emergency medical operations, physical therapy, and dental hygienists.	Investigative
Mathematics	Working with data and applying quantitative and statistical concepts and mathematical formulas. Includes jobs such as statisticians, mathematicians, engineers, or financial analysts.	
Science	Involves scientific activities such as studying biology, astronomy, geology, and physics; reading books about science; and doing scientific research or related activities. Includes jobs such as scientists and laboratory workers or in health services, technology and medical paraprofessionals, nutritional or pharmaceutical services involving scientific interests.	
Information Technology	Developing, maintaining, and using computer systems, software, and networks for the processing and distribution of data. Includes jobs such as computer systems analyst, network administrator, software developer, web administrator, and database administrator.	
Management	Leading others and influencing people and decisions. Includes jobs such as administrative or supervisory positions, for instance a shop foreperson, supervisor, school administrator, police or fire chief, head librarian, executive, hotel manager, or union official. Includes owning or managing a store or business.	Enterprising
Sales	Includes activities that involve selling products and services (e.g., in stores, offices, and customers' homes) and jobs in the fields of auto sales, insurance, lobbying, public relations, and real estate.	
Human Relations	Arranging positive interpersonal interactions for individuals. Includes setting company policies, acting as a mediator in a conflict, and solving interpersonal situations. Can also include activities that focus on increasing the satisfaction, morale, and motivation of employees. May include jobs as an arbitrator, mediator, human resource manager, or labor relations specialist.	
Office Work	Performing clerical, administrative, and business-related activities (recording, data processing, typing, filing, etc.). Includes jobs such as office manager, bookkeeper, receptionist, secretary, and administrative assistants.	Conventional
Finance	Managing assets and debt. Includes jobs that utilize numbers in business bookkeeping, accounting, and tax procedures.	
Food Service	Food processing, cooking, planning menus, and related activities. Includes jobs such as short-order cooks, cafeteria workers, caterers, food service managers, and waiters/waitresses	

Another advantage of the AVID is that it was developed using modern test theory and recent findings related to the psychometric properties of interest items. Recent research indicated that an ideal point item response theory (IRT) model provides the best representation of the response process for interest items (Tay, Drasgow, Rounds, & Williams, 2009). An ideal point model suggests that the choice to endorse or not endorse an item is described by a proximity relation, wherein one tends to endorse the item only if they are located near the item on the latent continuum (e.g., an individual would not endorse an item that is “higher” or “lower” than his or her trait level). Past research has shown that fitting an incorrect response model to ideal point data can attenuate correlations between variables and negatively affect selection decisions (Carter, Dalal, Boyce, O’Connell, Kung, & Delgado, 2014; Dalal & Carter, 2015). Therefore, the AVID was developed using the framework of ideal point models.

A third advantage of the AVID is that it was developed to be resistant to faking and other response biases (e.g., acquiescent responding). To facilitate its use in high-stakes settings, the AVID is administered in a two-alternative forced-choice format where statements assessing different basic interest dimensions, but matched on their extremity and social desirability, are administered in pairs. Respondents are then asked to select the statement that is “most like you.” Past research has demonstrated that this forced-choice format can mitigate the effects of faking (Cao & Drasgow, 2019), even in high-stakes settings (Trent et al., 2020; Stark et al., 2014). Responses to the forced-choice AVID items are then scored using the multidimensional pairwise preference (MDPP) model (Stark, Chernyshenko, & Drasgow, 2005), which can address potential concerns with ipsativity and successfully recover normative scores regardless of how many interest dimensions are assessed. Both the forced-choice format and the MDPP model have been used successfully with the Army’s Tailored Adaptive Personality Assessment System (TAPAS), which has demonstrated validity in a wide range of settings and occupations (e.g., Nye et al., 2012).

Finally, the AVID can also be administered in either a static or a computer-adaptive testing (CAT) format. One advantage of CAT is that the assessment can be updated and administered more efficiently than a paper-and-pencil form. In addition, the adaptive process allows for a reduced number of items to be administered, with some research indicating that test length can be cut by up to 50% with no loss of measurement precision (Stark, Chernyshenko, Drasgow, & White, 2012).

### **Validity of the AVID**

Given the advantages of the AVID, we expect this assessment to demonstrate validity in high-stakes testing situations and to help recruits identify MOS that they will be satisfied with and successful in. Initial research provided validity evidence for the AVID in several MOS (Nye et al., 2019). Across several criteria and two different samples, the multiple *R*s for regression-weighted composites of the AVID dimensions ranged from .14 to .52. In addition, the AVID also predicted overall performance both Army-wide and in specific MOS, including Infantry (11B), Military Police (31B), Combat Medics (68W), Motor Transport Operators (88M), and Wheeled Vehicle Mechanics (91B). The multiple *R*s for predicting overall performance in these MOS varied from .43 to .57.

When examining the validity of vocational interests, an important consideration is the fit between individuals and their jobs. In his widely researched theory, Holland (1959, 1997) proposed that individuals with interests that fit the work they are doing would be more satisfied with and successful in their jobs. Recent research has provided support for this prediction and demonstrated the role of interest fit at work. In their meta-analysis, Nye et al. (2017) found that interest fit was a much better predictor of performance than individual interest scales alone. Similarly, Nye et al. (2019) also demonstrated that the fit between Soldiers' AVID scores and their MOS was a strong predictor of their attitudes and engagement. Although these authors found strong validities for the AVID interest dimensions alone, these validities increased further when the fit between individuals and their MOS was examined (i.e., multiple *R*s ranged from .56 to .70). Therefore, it appears that the match between Soldiers' vocational interests and their MOS is important for predicting attitudes and performance in the Army.

### **Purpose of the Current Research**

Despite the initial validity evidence for the AVID, more research is needed to evaluate its utility for MOS assignment. For example, previous validity research on the AVID only examined a static version of this assessment. However, as described above, an important advantage of the AVID is that it can be administered in a computer-adaptive format, which provides a more efficient assessment of the constructs being measured. Therefore, although the AVID showed strong validity in previous research (Nye et al., 2019), more research is needed to evaluate the validity of an adaptive version of this assessment. To address this need, a concurrent validation study was conducted to examine the validity of a CAT version of AVID administered to a sample of Soldiers from several MOS.

In addition, more research is needed to examine ways of matching individuals to MOS based on their interests. In the vocational interest literature, the fit between an individual's interests and their environment has typically been calculated using congruence indices. Many congruence indices have been proposed for this purpose and are widely used in the literature (Brown & Gore, 1994; Camp & Chartrand, 1992). However, these congruence indices have a number of limitations, some of which can limit their validity for predicting work outcomes (Edwards, 1993). To address these limitations, some have proposed using regression composites (Van Iddekinge, Roth, Putka, & Lanivnich, 2011) or polynomial regression (Edwards, 1993; Nye Prasad, Bradburn, & Elizondo, 2018). However, it is unclear how much of an effect these alternative approaches will have on the validity and utility of interests for predicting work outcomes. Therefore, we conducted two simulation studies to examine this issue and explore the most appropriate method for calculating congruence and predicting Soldiers' performance potential in a broad range of MOS. Finally, a third simulation study built on the first two to examine the potential improvement in performance that could be obtained by matching individuals to jobs that fit with their interests.

# Concurrent Validation of a Computer-Adaptive Version of AVID

## Methods

### Data and Procedures

The data for this concurrent validation study were collected as part of the Validation of Accession Screening Tools (VAST) program of research. The data consisted of a total of 596 Soldiers who responded to the AVID and/or the criterion measures. Approximately 82% of the sample was male and 61% of the sample was white. The majority of the sample (52%) had a High School diploma as their highest level of education and most were either Privates First Class (E-3) (24%) or Specialists (E-4) (49%). The individuals in this sample were from a broad range of MOS, with the largest numbers of Soldiers coming from Infantry (11%) or Wheeled Vehicle Mechanics (10%). Nevertheless, given the relatively small overall sample size, our analyses focused on examining the validity of the AVID in the full sample rather than in specific MOS.

Both the CAT AVID and the criterion measures were administered to Soldiers simultaneously in a computerized format. Soldiers were first asked to log onto the computer to access the assessment and enter their ID number. Next, they were informed of the purpose of the assessment and asked to provide consent to participate. They then responded to a series of questions assessing demographic information, the AVID dimensions, and the criteria. In addition, each Soldier also rated their MOS on all 20 AVID dimensions. These MOS ratings were used to calculate the fit between each individual and their MOS.

### Measures

**AVID.** A 123-item adaptive version of the AVID was administered to Soldiers in this sample. This version of the AVID assessed 16 basic interests (all the AVID dimensions except for Science, Personal Service, Sales, and Finance). Only 16 of the 20 AVID dimensions were administered to reduce the amount of testing time and alleviate concerns about test-taker fatigue. The AVID statements were administered in a forced-choice format and Soldiers were asked to pick one statement in each pair that was “more like you.” Again, the statements for each pair were matched based on their extremity and social desirability to mitigate the effects of faking and response biases on the interest scores.

**Army Life Questionnaire (ALQ).** The criteria for this sample were assessed using the ALQ. The ALQ is a self-report attitudinal measure currently used in ARI validation research (Nesbitt, Salmon, & Kirkendall, 2020). The ALQ includes sections on Soldiers’ demographic characteristics, background, and experience information, as well as assessments of Soldiers’ attitudes, perceptions of fit (both in the Army and in their MOS), affective commitment, resilience, motivation to lead (MTL), organizational citizenship behavior/Leadership (OCB), counterproductive Soldier behaviors (CSB), and career/reenlistment intentions. Descriptions of the ALQ scales included in this research are presented in Table 2. In addition to these scales, the ALQ also asked about Soldiers’ Army Physical Fitness Test (APFT) scores and experiences with disciplinary incidents.

### Table 2.

***Army Life Questionnaire (ALQ) Scales Included***

<b>Construct</b>	<b>Definition</b>
Counterproductive Soldier Behaviors (CSB)	Intentional behaviors that harm or are intended to harm another Soldier or the legitimate interests of the unit
Army Fit	The extent to which a Soldier feels like the Army is a good match for them
MOS Fit	The extent to which a Soldier feels like their current MOS is a good fit
Affective Commitment	Soldiers' attachment to and identification with the Army
MOS Satisfaction	Satisfaction with the opportunities and daily work involved in the Soldier's MOS
Career Intentions	Likelihood of staying in the Army until retirement
Reenlistment Intentions	Likelihood of reenlisting for another term of service
Motivation to Lead (MTL)	The factors that affect an individual's decisions to assume leadership training, roles, and responsibilities, and affect their intensity of effort at leading and persistence as a leader (Chan & Drasgow, 2001); the conceptual and empirical model of MTL includes three underlying dimensions: Affective, Noncalculative, and Social-Normative
Organizational Citizenship Behavior and Leadership (OCB)	Engaging in voluntary behaviors to help another individual or the organization itself (Bateman & Organ, 1983), including behaviors that Soldiers engage in to display leadership qualities, absent of an official leadership role
Resilience	The capacity to overcome difficult life events with minimal disruption or long-term negative impacts on psychological and physical functioning (Bonanno, 2004)

Table 3 provides the means, standard deviations, and intercorrelations for the criteria assessed by the ALQ in this sample. In addition to examining each outcome individually, we also examined the prediction of an overall performance 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 score. Negatively worded scales (e.g., CSB) were reverse coded before calculating the overall performance composite scores so that all scales were in a consistent direction. To be consistent with previous research on the AVID (Nye et al., 2019), disciplinary incidents were not included as part of the overall performance composite. The goal of combining criterion scales in this way was to determine the utility of the AVID for predicting a broader criterion variable and to examine composites of AVID scales that might be useful for MOS assignment decisions. 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.

**MOS Ratings.** Soldiers also responded to items asking them to rate their MOS on each of the 20 AVID dimensions. These ratings served two purposes: a) to indicate the perceived relevance of each AVID dimension for a particular MOS, and b) to identify the interest profile of

each MOS for calculating person-job fit. As noted above, past research has shown that the validity of vocational interests is highest when considering the match between individuals and their environment. Therefore, these ratings were important for examining the validity of the AVID.

To collect these ratings, Soldiers were given the name and a description of each AVID dimension. Examples of activities that are associated with each dimension were also included to provide Soldiers with a clearer understanding of how these dimensions might relate to their MOS. Then, Soldiers were asked to rate “How descriptive is this dimension of your current MOS?” on a scale from 1 (“Not at all descriptive”) to 7 (“Extremely descriptive”). The means and standard deviations for these ratings are provided in Table 4. As shown in this table, Management, Physical Activity, and Teaching were rated as most descriptive of the various MOS in this sample. Nevertheless, there was also substantial variability in the ratings across all 16 AVID dimensions.

## **Analyses**

Using the AVID and criterion data described above, we examined the validity of the AVID for predicting important military outcomes. Before conducting these analyses, the data were first screened for unmotivated responders. In addition to the items assessing the AVID dimensions, four items were also included to detect unmotivated responding. These items instructed participants to select a particular option for that item (e.g., “For data quality check, please select this option for this pair”). Individuals who responded incorrectly (i.e., marked a response other than the one they were instructed to mark) to more than one of these random response checks were flagged as potentially unmotivated. In addition, we also screened individuals for patterned (i.e., selecting option “A” a disproportionate number of times) and rapid responding (i.e., answering 12 or more items in less than 2 seconds each). To screen for unmotivated responding, we flagged individuals on each of these criteria (i.e., 1 = flagged for unmotivated responding, 0 = not flagged) and summed these flags to get the total number for each individual. Individuals who were flagged on more than one of these criteria were removed from analyses. After removing individuals who were identified as potentially unmotivated, the total sample size for analyses was  $n = 287$ .

Next, we used the reduced dataset to examine the validity of the computer-adaptive version of AVID. Consistent with past research demonstrating the benefits of regression-weighted composites for both interests (Van Iddekinge et al., 2011) and personality (Nye et al., 2012), we first used regression analyses to develop weighted composites of the AVID scales that predicted each of the criteria assessed in this research. Despite the potential validity of regression-weighted composites, this approach also has limitations. For example, in order to quantify the match (or fit) between an individual and their job, it is necessary to include interest scores for both individuals and their MOS. However, the regression-weighted composites described above only include the interest scores for individuals. Therefore, adding the MOS ratings to the model could also improve the prediction of work outcomes by providing a more appropriate way to operationalize interest fit.

Edwards (1993) provided the mathematical proof that regression models including both individual and environment scores can provide one way of operationalizing person-environment

fit and suggested polynomial regression as an alternative to composites of individual scores alone. With this approach, individual interest scores are included in the model along with environment interest scores, quadratic terms for both the individual and environment scores, and interactions between individual scores and the corresponding environment scores. Although a full polynomial regression model generally includes all of these terms, we examined a reduced model with only the main effects for the individual and MOS interest scores for two reasons. First, including all of the quadratic and interaction terms would have resulted in a total of 80 terms in the regression model. Given the relatively small sample size for this study, including so many terms in the regression model would have increased the likelihood of capitalizing on chance and finding spurious results. Second, previous research with the AVID has examined the full polynomial regression model in several samples and found that the quadratic and interaction terms rarely contributed to the prediction of the outcomes (Nye et al., 2019). Research with civilian samples has found similar results (Nye, Prasad, Bradburn, & Elizondo, 2018; Wiegand, Drasgow, & Rounds, 2021). Therefore, we only examined the main effects in the current research.

**Table 3.**

***Descriptive Statistics and Intercorrelations between the Criteria***

Variables	Mean	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. MOS Fit	3.03	.89	1.00														
2. Army Fit	3.31	.84	<b>.35</b>	1.00													
3. Affective Commitment	2.90	.93	<b>.42</b>	<b>.69</b>	1.00												
4. MOS Satisfaction	2.87	.94	<b>.65</b>	<b>.51</b>	<b>.45</b>	1.00											
5. OCB	3.29	.80	<b>.17</b>	<b>.41</b>	<b>.41</b>	<b>.29</b>	1.00										
6. Resilience	3.74	.70	<b>.28</b>	<b>.45</b>	<b>.37</b>	<b>.40</b>	<b>.56</b>	1.00									
7. Reenlistment Intentions	3.04	1.39	<b>.25</b>	<b>.66</b>	<b>.50</b>	<b>.39</b>	<b>.33</b>	<b>.28</b>	1.00								
8. Career Intentions	2.72	1.31	<b>.26</b>	<b>.65</b>	<b>.58</b>	<b>.40</b>	<b>.35</b>	<b>.27</b>	<b>.78</b>	1.00							
9. MTL (Affective)	3.44	.80	<b>.22</b>	<b>.27</b>	<b>.26</b>	<b>.20</b>	<b>.52</b>	<b>.43</b>	<b>.25</b>	<b>.21</b>	1.00						
10. MTL (Noncalculative)	3.74	.87	<b>.15</b>	<b>.22</b>	<b>.18</b>	.05	<b>.25</b>	.11	<b>.20</b>	<b>.15</b>	<b>.14</b>	1.00					
11. MTL (Social-Normative)	3.86	.80	<b>.29</b>	<b>.42</b>	<b>.42</b>	<b>.32</b>	<b>.59</b>	<b>.56</b>	<b>.30</b>	<b>.27</b>	<b>.63</b>	<b>.26</b>	1.00				
12. APFT	230.45	61.14	-.07	-.03	-.04	-.01	.11	.09	-.01	-.06	<b>.16</b>	.01	<b>.15</b>	1.00			
13. Disciplinary Incidents	.51	1.10	<b>-.13</b>	<b>-.19</b>	<b>-.21</b>	-.08	-.11	-.10	-.08	<b>-.12</b>	-.09	<b>-.17</b>	-.10	.03	1.00		
14. CSB	1.94	.58	-.04	<b>-.20</b>	<b>-.14</b>	-.05	.00	-.07	<b>-.13</b>	<b>-.17</b>	-.07	-.11	<b>-.14</b>	<b>.13</b>	.07	1.00	
15. Overall Performance	.07	7.27	<b>.55</b>	<b>.78</b>	<b>.72</b>	<b>.64</b>	<b>.66</b>	<b>.65</b>	<b>.68</b>	<b>.68</b>	<b>.58</b>	<b>.39</b>	<b>.71</b>	<b>.16</b>	<b>-.19</b>	<b>-.27</b>	1.00

Note: Sample sizes ranged from 223 to 255. Bold values are statistically significant,  $p < .05$ . OCB = Organizational Citizenship and Leadership; MTL = Motivation to Lead; APFT = Army Physical Fitness Test; CSB = Counterproductive Soldier Behaviors.

**Table 4.*****Means and Standard Deviations of the MOS Ratings***

<b>Variables</b>	<b>Mean</b>	<b>Standard Deviation</b>
Combat	2.49	2.32
Construction	1.54	1.35
Electronics	2.54	2.05
Finance	1.23	0.82
Food Service	1.17	0.89
Human Relations	3.19	2.11
Information Technology	1.87	1.67
Management	4.11	2.11
Mathematics	2.25	1.65
Mechanical	2.91	2.19
Medical Services	1.80	1.55
Office Work	3.08	2.11
Outdoors	2.73	2.15
Personal Service	1.55	1.38
Physical Activity	3.81	2.30
Protection	1.96	1.65
Sales	1.12	0.71
Science	1.37	1.08
Teaching	3.77	2.09
Writing	2.26	1.75

*Note: n = 254 to 255.*

## Results

Table 5 shows the intercorrelations between the AVID dimensions. The correlation results were largely consistent with previous research. For example, the Combat, Construction, Mechanical, and Protection dimensions were all strongly correlated with each other. In addition, the Information Technology, Electronics, and Mathematics dimensions were also strongly correlated. These results are consistent with previous AVID research and with the broader RIASEC model of interests.

Table 6 shows the AVID dimensions that were significant predictors of each of the criteria in this sample. Due to the small sample size and the corresponding effects on the significance of the regression weights, the regression weights are presented for all AVID dimensions, regardless of whether they were significant or not. As shown in this table, the AVID scales predicted a number of criteria very well. Across all of the outcomes, the AVID Combat, Management, and Outdoors dimensions were some of the strongest predictors. Results also indicated that the AVID dimensions had the strongest relationships with MTL (both affective and social-normative) and OCB. The multiple *R*s for these outcomes were all .46 or above. Similarly, the AVID dimensions also predicted the overall performance composite with a multiple *R* of .45 and an adjusted (for capitalization on chance) multiple *R* of .39. Although the strongest predictors of this outcome were the Human Relations, Management, and Outdoors dimensions, a number of other predictors (both positive and negative) were also included. These results are largely consistent with previous research on the AVID (Nye et al., 2019).

**Table 5.**

*Correlations between the AVID Basic Interest Dimensions*

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1. Combat	1.00															
2. Construction	<b>.34</b>	1.00														
3. Electronics	<b>.12</b>	<b>.34</b>	1.00													
4. Food Service	<b>-.11</b>	<b>.02</b>	-.07	1.00												
5. Human Relations	<b>-.17</b>	<b>-.19</b>	-.08	.07	1.00											
6. Information Tech.	-.08	-.02	<b>.57</b>	-.05	.10	1.00										
7. Management	-.04	<b>-.20</b>	<b>-.14</b>	.03	<b>.44</b>	.00	1.00									
8. Mathematics	.00	-.03	<b>.28</b>	-.03	.10	<b>.28</b>	.07	1.00								
9. Mechanical	<b>.41</b>	<b>.59</b>	<b>.52</b>	-.09	<b>-.15</b>	<b>.17</b>	<b>-.14</b>	.07	1.00							
1. Medical Services	<b>-.12</b>	-.05	<b>-.12</b>	.01	<b>.20</b>	-.06	.07	.11	<b>-.17</b>	1.00						
11. Office Work	<b>-.31</b>	<b>-.21</b>	-.08	.07	<b>.41</b>	<b>.15</b>	<b>.33</b>	<b>.19</b>	<b>-.24</b>	<b>.16</b>	1.00					
12. Outdoors	<b>.28</b>	<b>.47</b>	-.01	.09	-.10	<b>-.29</b>	-.05	<b>-.13</b>	<b>.22</b>	.05	<b>-.19</b>	1.00				
13. Physical Activity	<b>.20</b>	.07	<b>-.17</b>	-.09	.03	<b>-.21</b>	<b>.15</b>	-.11	.07	.07	-.11	<b>.15</b>	1.00			
14. Protection	<b>.51</b>	<b>.22</b>	.05	<b>-.15</b>	-.05	-.08	.08	.06	<b>.29</b>	.02	-.09	<b>.17</b>	<b>.18</b>	1.00		
15. Teaching	<b>-.13</b>	<b>-.22</b>	-.01	.10	<b>.43</b>	.06	<b>.33</b>	<b>.24</b>	<b>-.23</b>	<b>.35</b>	<b>.25</b>	-.07	.01	-.03	1.00	
16. Writing	<b>-.18</b>	<b>-.28</b>	<b>-.15</b>	<b>.13</b>	<b>.24</b>	.04	<b>.18</b>	<b>.16</b>	<b>-.34</b>	<b>.17</b>	<b>.27</b>	-.07	-.02	<b>-.18</b>	<b>.37</b>	1.00

Note:  $n = 287$ . Bold values are statistically significant,  $p < .05$ .

**Table 6.**

*Standardized Regression Weights for the AVID Dimensions Predicting Each Criterion*

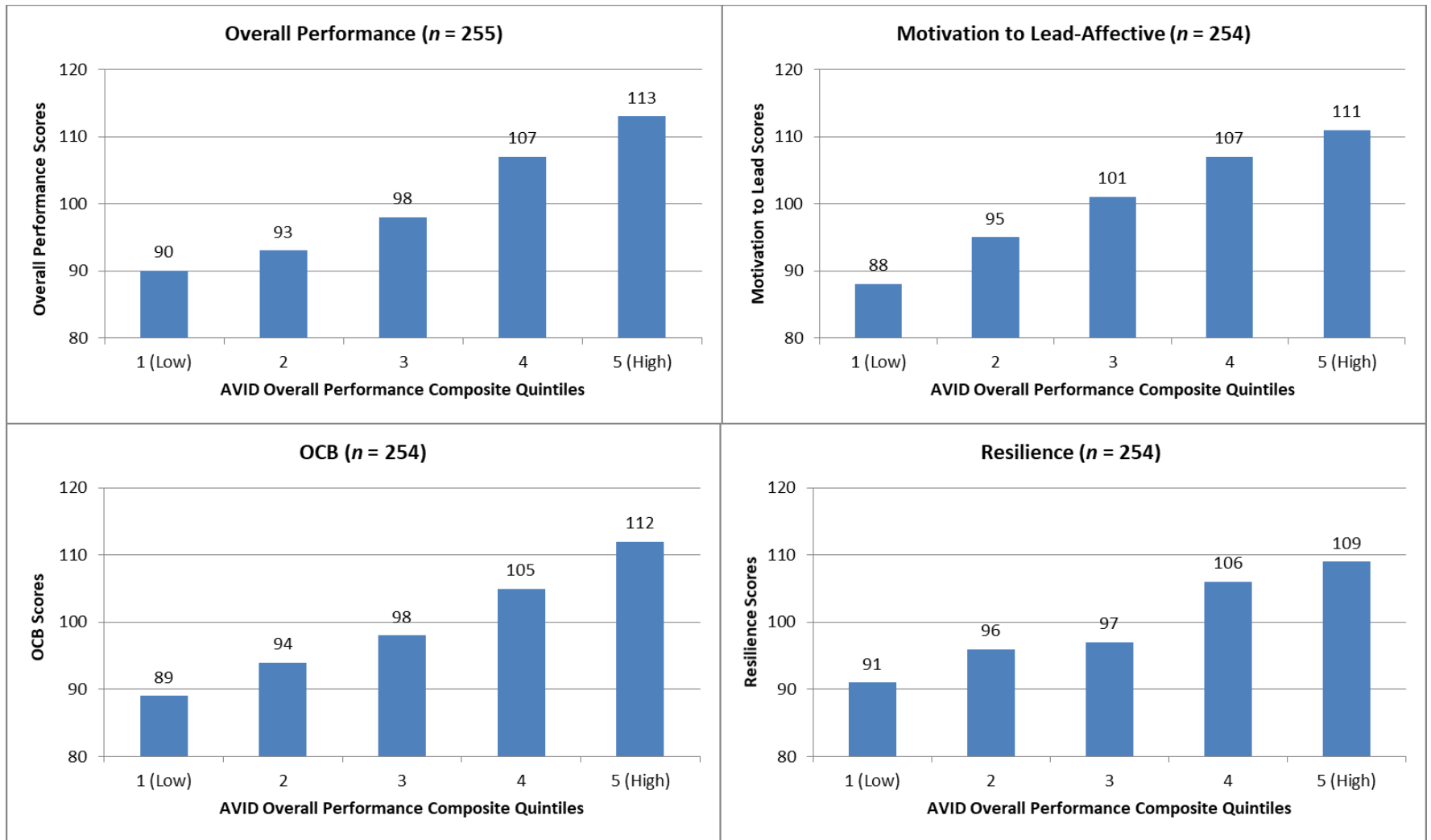
AVID Dimensions	MOS Fit	Army Fit	Aff. Com.	MOS Sat.	OCB	Resil.	Re-enlist. Int.	Career Int.	Mot. to Lead (Aff)	Mot. to Lead (Noncal.)	Mot. to Lead (Soc-Norm)	APFT	Disc. Incidents <sup>a</sup>	CSB	Overall Perform.
Combat	.10	<b>.21</b>	<b>.14</b>	.09	<b>.15</b>	-.11	<b>.22</b>	<b>.16</b>	-.02	.10	-.04	-.12	<b>-.52</b>	-.03	.12
Construction	-.07	-.04	-.03	.03	-.10	.09	-.02	-.05	-.10	-.06	-.06	.00	.02	-.10	-.05
Electronics	.02	.05	.08	.06	.00	.06	-.08	.02	.02	-.08	.15	.01	.18	-.03	.05
Food Service	-.03	.03	.02	.00	-.04	.03	<b>.12</b>	<b>.16</b>	.00	.00	-.03	<b>-.14</b>	-.14	-.04	.03
Human Relations	.05	.10	.09	.04	<b>.13</b>	.09	<b>.13</b>	<b>.13</b>	.07	<b>.15</b>	.10	.06	.31	<b>-.14</b>	<b>.18</b>
Information Tech.	.14	.11	<b>.14</b>	.11	<b>.17</b>	.08	.10	.10	.00	.04	-.07	-.06	.15	-.02	.12
Management	.08	.08	<b>.20</b>	.02	<b>.22</b>	<b>.18</b>	.02	.07	<b>.44</b>	.04	<b>.34</b>	-.05	-.16	<b>-.15</b>	<b>.23</b>
Mathematics	-.01	-.06	<b>-.11</b>	-.08	-.07	.03	-.02	-.08	.00	.02	.05	<b>.15</b>	.01	.05	-.04
Mechanical	.04	<b>-.21</b>	-.07	-.05	.14	-.02	-.13	-.05	<b>.17</b>	-.09	-.05	.11	.48	.15	-.05
Medical Services	-.02	-.08	-.02	-.09	.00	-.09	-.09	-.02	-.04	-.06	-.04	.01	-.15	.02	-.07
Office Work	.06	-.02	-.07	.11	.00	-.02	-.05	-.04	-.02	-.02	-.01	.02	-.15	.04	-.01
Outdoors	.08	<b>.16</b>	<b>.17</b>	.05	<b>.15</b>	.12	.04	.04	<b>.12</b>	.09	<b>.21</b>	.02	-.04	.06	<b>.15</b>
Physical Activity	-.01	.01	-.02	.09	<b>.11</b>	<b>.17</b>	-.02	.03	.09	-.01	.05	<b>.19</b>	.07	-.06	.10
Protection	.02	.10	.12	-.06	.05	<b>.13</b>	.03	.09	.04	.08	.10	-.11	.09	.06	.07
Teaching	.05	.10	.07	.09	.10	.10	.08	.08	.01	.07	-.02	.00	.08	.02	.09
Writing	-.10	-.06	-.01	-.06	.00	-.01	.01	-.01	-.05	-.09	.00	-.04	.16	<b>.14</b>	-.07
Multiple <i>R</i>	.25	.36	.40	.23	.48	.38	.32	.32	.52	.27	.46	.31	.24	.27	.45
Adjusted Multiple <i>R</i>	.00	.27	.32	.00	.42	.30	.20	.21	.47	.11	.39	.16	--	.09	.39

Note: Bold values are statistically significant,  $p < .10$ . Samples sizes ranged from 253-255. MOS Sat. = MOS Satisfaction; OCB = Organizational Citizenship Behavior and Leadership. APFT = Army Physical Fitness Test; CSB = Counterproductive Soldier Behavior. <sup>a</sup> Because the Disciplinary Incidents variable was dichotomized to account for low base rates, these analyses are based on a logistic regression. Therefore, an adjusted multiple *R* could not be calculated. In addition, the regression weights presented for this outcome are the unstandardized regression weights.

Figure 1 illustrates the practical importance of the relationships between the AVID and several of the criteria shown in Table 6. We used the standardized regression weights from the analyses shown in Table 6 for predicting the overall performance criterion composite to calculate AVID composite scores for each individual. We then used these scores to plot the relationships between this AVID composite and several of the criteria assessed in this sample. Figure 1 illustrates the relationships between the AVID composite scores and overall performance, MTL (affective), OCB, and resilience. The X-axes of these plots are the quintiles for the scores on the AVID composite and the Y-axes provide the average scores on the criteria. 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.

The graphs shown in Figure 1 indicate that individuals who scored higher on the AVID composite had higher overall performance scores, greater motivation to lead, engaged in more OCBs, and were more resilient. In addition, for all of these outcomes, there was nearly a full standard deviation difference between the highest and lowest scoring groups on the AVID composite, indicating that the effects were substantial. These results suggest strong relationships between the AVID dimensions and the criteria assessed in this sample of Soldiers.

Finally, the results of the regression model that included the main effects for both individuals and their MOS are shown in Table 7. The outcome for these analyses was the overall performance composite shown in Table 6. As shown in Table 7, adding the main effects for the MOS interest scores improved the multiple *R* to .57 and the adjusted multiple *R* to .48. As noted above, these analyses help to demonstrate the potential utility of examining interest fit rather than focusing on the interest scales alone.



**Figure 1. AVID Composite Quintile Plots for Overall Performance, Motivation to Lead, OCB, and Resilience**

**Table 7.*****Standardized Regression Weights for the Predictors in the AVID Interest Fit Composite***

Variables	Full Sample
Combat	.11
Construction	-.05
Electronics	.09
Food Service	.05
Human Relations	<b>.13</b>
Information Technology	.10
Management	<b>.20</b>
Mathematics	-.10
Mechanical	-.08
Medical Services	-.06
Office Work	-.03
Outdoors	.07
Physical Activity	<b>.11</b>
Protection	.09
Teaching	.11
Writing	-.07
<hr/>	
Combat (MOS Ratings)	.04
Construction (MOS Ratings)	-.03
Electronics (MOS Ratings)	-.04
Food Service (MOS Ratings)	<b>-.15</b>
Human Relations (MOS Ratings)	<b>.17</b>
Information Technology (MOS Ratings)	.01
Management (MOS Ratings)	<b>.16</b>
Mathematics (MOS Ratings)	.04
Mechanical (MOS Ratings)	<b>.13</b>
Medical Services (MOS Ratings)	.05
Office Work (MOS Ratings)	.12
Outdoors (MOS Ratings)	.02
Physical Activity (MOS Ratings)	<b>-.19</b>
Protection (MOS Ratings)	.04
Teaching (MOS Ratings)	-.04
Writing (MOS Ratings)	-.10
<hr/>	
Multiple <i>R</i>	.57
Adjusted Multiple <i>R</i>	.48

*Note:* Bold values are statistically significant,  $p < .10$ .  $n = 255$ .

## Simulation Research to Examine Operationalizations of Interest Fit

Two simulations were conducted to examine the best methods of calculating interest fit. The first simulation compared several traditional congruence indices from the interest literature to the polynomial regression approach. Although the polynomial regression approach has been used in previous AVID research (e.g., Nye et al., 2019), the goal of these simulations was to determine whether alternative methods of calculating congruence could provide stronger validity than the polynomial regression approach. The second simulation builds on the first to examine several novel approaches to estimating interest fit using modern predictive models. Specifically, we explored the use of several machine learning (ML) algorithms to estimate interest fit. Machine learning is a form of artificial intelligence (AI) that is designed to build models that maximize the prediction of an outcome. This method has shown promise for estimating predictive relationships in the organizational literature (Putka, Beatty, & Reeder, 2018; Speers, Christiansen, Robie, & Jacobs, in press) and, therefore, may have benefits for modeling vocational interest fit as well.

### Simulation Study 1

#### Methods

To compare the utility of traditional congruence indices and polynomial regression, interest data were simulated for both individuals and jobs. To generate realistic data, all simulations were based on empirical data obtained from previous AVID research (Nye et al., 2019). These data included both person (i.e., AVID responses) and job (i.e., MOS ratings) scores on the same 16 AVID basic interest dimensions examined in the concurrent validation study (the Finance, Sales, Science, and Personal Service dimensions were not assessed). In addition, scores on the overall performance criterion composite were also included in this dataset. With this empirical data, we estimated a 33 x 33 covariance matrix that could be used as the population matrix for data generation.

Next, we simulated individual, job, and criterion data from a multivariate normal distribution using the covariance matrix estimated with the empirical data as the population matrix. These data were generated for sample sizes of 100, 500, 1000, 5000, and 10000 and 100 replications were simulated for each sample size. In other words, we generated a total of 500 (5 sample sizes x 100 replications) unique samples for our analyses. Although we simulated data for all 16 of the AVID basic interest dimensions that were included in the original dataset, these dimensions can be aggregated to get scores on the broader RIASEC interest types (Nye et al., 2019). Therefore, because each of the widely used congruence indices in the interest literature were developed specifically for the RIASEC model (Brown & Gore, 1994), we aggregated scores from the 16 basic interests to get RIASEC scores for both individuals and jobs. These aggregated scores were then used to estimate congruence indices and the polynomial regression model.

**Congruence Indices.** Many of the traditional congruence indices in the interest literature require creating interest profiles by ranking the RIASEC scores in order of their importance for both individuals and jobs. Therefore, we examined the simulated data to identify the three-letter codes (i.e., the three strongest interests) for both individuals and their jobs, and these three-letter

codes were then used to calculate several congruence indices. The congruence indices that were calculated in this study included:

- The K-P index (KP; Kwak & Pulvino, 1982),
- Gati's (1985) Sb index (GATI)
- The C Index (CIndex; Brown & Gore, 1994)
- Iachan's M (MIndex; Iachan, 1984)
- The ranked comparison congruence scale (RCCS; Robbins et al., 1978)
- The compatibility index (COMI; Wiggins & Moody, 1981)
- The three-level congruence index (TLCI; Wolfe & Betz, 1981),
- The Z-S index (ZSI; Zener & Schnuelle, 1976)
- The two-letter agreement index (TLAI; Healy & Mourtou, 1983)
- The first-letter agreement based on the hexagon (FLAH; Holland, 1973)
- The dichotomous first-letter agreement index (DFLA; Holland, 1963).

Additional details about the calculation of each of these indices are provided in Brown and Gore (1994).

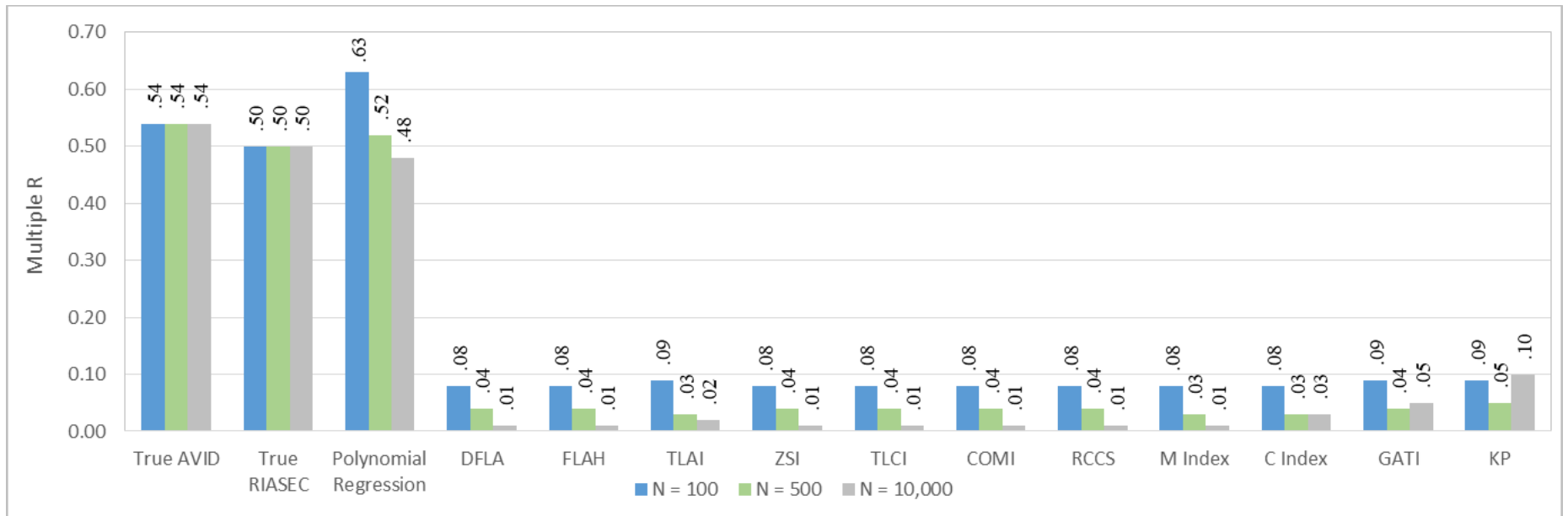
**Polynomial regression.** As described above, the traditional polynomial regression model proposed by Edwards (1993) included individual interest scores, interest scores for the job, quadratic terms for both the individual and job scores, and the interactions between the individual scores and the corresponding job scores. Therefore, although previous research on the AVID has found little support for the utility of the full model (e.g., Nye et al., 2019), we examined each of these terms in our simulated data to explore their utility in a more generalizable way. After estimating the full polynomial regression model, the predicted scores from this model served as the index of fit between an individual and their job.

**Analyses.** To examine the utility of the various congruence indices and polynomial regression, we compared their predictive validity across the simulated conditions. Specifically, we used each of the operationalizations of interest fit discussed above to predict the overall performance criterion composite simulated based on the original AVID data. For the congruence indices, we estimated separate ordinary least squares (OLS) regression models with each congruence index predicting the criterion. Finally, the multiple *R* for each of these regression models was examined to compare the validity of the congruence indices with polynomial regression.

To provide a frame of reference for the congruence indices and polynomial regression model, we also estimated a linear regression model using the population correlation matrix from the AVID data. Here, we estimated regression models both with the 16 AVID basic interest dimensions and with the six RIASEC scores (i.e., aggregated basic interest dimensions) as predictors. Although the model that included the six RIASEC scores provided the most direct comparison with the congruence indices and polynomial regression model, the model that included the 16 basic interests as predictors was also examined to determine the potential decrement in validity due to aggregating interest scores. The multiple *R* for the model with the six RIASEC interests was .50. In contrast, the multiple *R* for the model with the 16 basic interest dimensions was .54. These values served as the baseline for evaluating the validity of the congruence indices and polynomial regression model.

## Results

The results of the simulations for Study 1 are shown in Figure 2. This figure shows the average multiple  $R$  for each regression model across all 100 replications in the  $N = 100$ , 500, and 10,000 conditions. Given the consistency of the findings, we do not present results for the other sample sizes. As shown in Figure 2, the average multiple  $R$ s for nearly all of the congruence indices we examined were relatively low. In addition, these values were dramatically lower than the multiple  $R$  for the true population model, which was .50 when the six RIASEC scores were used as predictors. In contrast, the multiple  $R$  for the polynomial regression model was .63 when  $N = 100$ . In other words, when the sample size was small ( $N = 100$ ), the polynomial regression model substantially *overestimated* the true population multiple  $R$ . However, the overestimation disappeared when the sample size was 500. In this case, the multiple  $R$  of the polynomial regression model was .52, which was remarkably close to the true population value. Nevertheless, the congruence indices still performed poorly under these conditions. The results were similar when the sample size was 10,000. Under these conditions, the congruence indices performed poorly while the average (across replications) multiple  $R$  for the polynomial regression model was .48, which was only slightly below the true population value. These differences between the congruence indices and polynomial regression were likely due to the fact that congruence indices impose constraints on the relationships between the interest scores and the outcome, which is not the case in the polynomial regression model (Edwards, 1993). These results demonstrate that polynomial regression provides higher and more accurate estimates of the overall validity of vocational interests than any of the previously examined congruence indices.



**Figure 2. Simulation Results Comparing Polynomial Regression to Traditional Congruence Indices**

## Simulation Study 2

The results of the first simulation study suggested that polynomial regression can be a useful approach to calculating interest fit. Specifically, polynomial regression provided both higher and more accurate estimates of validity than traditional congruence indices. As suggested by previous research (Edwards, 1993), one potential explanation for these findings is that polynomial regression reduces the constraints on elements of the person and environment interest profiles and, therefore, allows for more accurate estimates of the relationships between the predictors and the criterion. Given the potential advantages of estimating a predictive model like polynomial regression, other modern prediction methods may help to further improve estimates of these relationships. For example, recent research has examined various forms of ML to estimate prediction models. These methods have shown some promise, particularly when a large number of predictors are used and when multicollinearity is present (Putka, Beatty, & Reeder, 2018), as is often the case with polynomial regression. Therefore, we conducted a second simulation study to compare polynomial regression to several ML methods to determine whether these modern prediction methods provide any advantages over regular OLS regression for calculating interest fit.

### Methods

The same procedures used for the first simulation study were used to generate the population interest data for the second simulation study as well. Specifically, all simulations were based on the empirical covariance matrix of AVID scores and an overall performance criterion obtained from previous research (Nye et al., 2019). Using this covariance matrix, we simulated person, job, and criterion scores from a multivariate normal distribution for a large sample of one million cases. This overall sample served as the population for this study. Next, 100 unique subsamples of 500 cases were drawn from this population and used as training samples for all of the models examined in this study. In addition, another sample of 10,000 cases was also randomly selected from the same population of one million cases to serve as the test sample. In other words, each model was estimated in the 100 training samples of 500 cases and then tested on the separate sample of 10,000 cases to cross-validate the results.

**Models Estimated.** In each of the training samples described above, the following models were estimated:

- ***Polynomial Regression*** – This is the same model estimated in Study 1.
- ***Least Angle Regression (LARS)*** – fits a linear regression model to high-dimensional data in a way that is similar to forward stepwise regression. LARS provides a way to estimate the most important variables to include in a model. However, instead of including variables at each step, the estimated parameters are added in a direction equiangular to each one's correlations with the residual.
- ***Elastic Net Regression (ENet)*** – fits a regularized linear regression that incorporates penalties for model complexity that are similar to lasso and ridge regression.

- ***Principle Components Regression (PCR)*** – fits a linear regression model that is based on principal components analysis. That is, the principal components of the predictors are used as variables in the model.
- ***Partial Least Squares Regression (PLS)*** – fits a linear regression model by projecting the predicted outcome and the observed predictors to a new space, and finds the multidimensional direction in the predictor space that explains the maximum amount of variance in the outcome.
- ***Random Forests (RF)*** – this is an ensemble learning method for regression that constructs multiple decision trees (to avoid overfitting) and outputs the mean prediction of the individual trees.
- ***Stochastic Gradient Boosted Trees (GBT)*** – this approach considers several decision tree learning algorithms and builds a predictive model in a stepwise fashion by combining the prediction of several base estimators to improve the robustness of the base model.
- ***Support Vector Machine with radial basis function (SVM Radial)*** – finds an optimal decision boundary that best separates the predictors into different domains, and then predicts the outcome based on the separation boundary. The radial basis function allows for non-linear boundaries.
- ***Multivariate Adaptive Regression Splines (MARS)*** – fits a non-parametric regression model to capture nonlinearities and interactions between variables automatically.

Each of these models were estimated using a  $k$ -fold cross-validation procedure ( $k = 5$  in this simulation) that was repeated 10 times.  $K$ -fold cross-validation involves dividing the sample into  $k$  subsamples and evaluating how well a model estimated in a subset of these samples generalizes to the remaining subsamples. Here, 80% of a sample was used for model development and 20% was used to evaluate the performance of the model. The multiple  $R$ s for each model were estimated in both the training and test samples and the difference between these estimates (i.e., shrinkage) was estimated to evaluate how well each model generalized to new data. These procedures were repeated 10 times and the average of the multiple  $R$ s are reported as the final estimates.

Several of the models described above require “tuning,” or the selection of parameters that help balance the competing goals of maximizing prediction in one sample and ensuring generalizability to another sample. These tuning parameters are used to limit the complexity of the model and/or place constraints on parameter estimates that reduce their sensitivity to specific sample characteristics (Putka et al., 2018). In this study, we selected the tuning parameters that provided the simplest model that was still comparable in performance to the best performing model. Using the best-tuned parameters for each model, we recorded the average model  $R^2$  in the training samples, the  $R^2$  from the  $k$ -fold cross-validation, and the correlation between the predicted outcome values and the true outcome values in the population. Each of these values

were then compared to identify the best performing model out of those tested. Once the final model parameters were identified, we also examined the generalizability of each model in the test sample of 10,000 cases. To do so, the estimated models were applied to the test sample and the predicted outcome values were estimated for each model. These predicted values were then compared to the actual values in the test sample and the correlations between estimated and predicted values were calculated.

Finally, after estimating the final models, we also calculated the relative importance of the predictors, which indicates the contribution of each variable to the overall model. One concern with ML in general is that many of these algorithms are “black boxes” in that the variables that contribute the most to prediction cannot be determined. Therefore, relative variable importance could only be estimated for a subset of the models examined here. For each of these models, variable importance was scaled to have a maximum value of 100 and a minimum value of 0 and was determined using the following methods:

- For linear models, variable importance was determined using the absolute value of the t-statistic for each model parameter.
- For partial least squares, variable importance was based on a weighted sum of the absolute regression coefficients. The weights are a function of the reduction of the sums of squares across the PLS components.
- For random forests, the mean squared error (MSE) is computed on the out-of-bag data for each tree and after permuting a variable. The differences between these estimates of the MSE are then averaged and normalized using the standard error.
- For boosted trees (e.g., GBT), variable importance was calculated using the same method as for a single tree but importance was summed over each boosting iteration.
- For MARS models, the variable importance measure reflects the total reduction in the model fit statistic after each predictor is added to the model.

## Results

The results of the second simulation study are shown in Figures 3-5. Figure 3 shows the  $R^2$  for each model averaged across the 100 replications that were simulated. Because each of the analyses samples was randomly drawn from the population dataset of one million cases, the  $R^2$  in the population is also shown for comparison. This value was calculated by using all of the variables in the population dataset to estimate a regression model. The  $R^2$  for this model was .23, which represents the baseline value that we are trying to estimate in each sample.

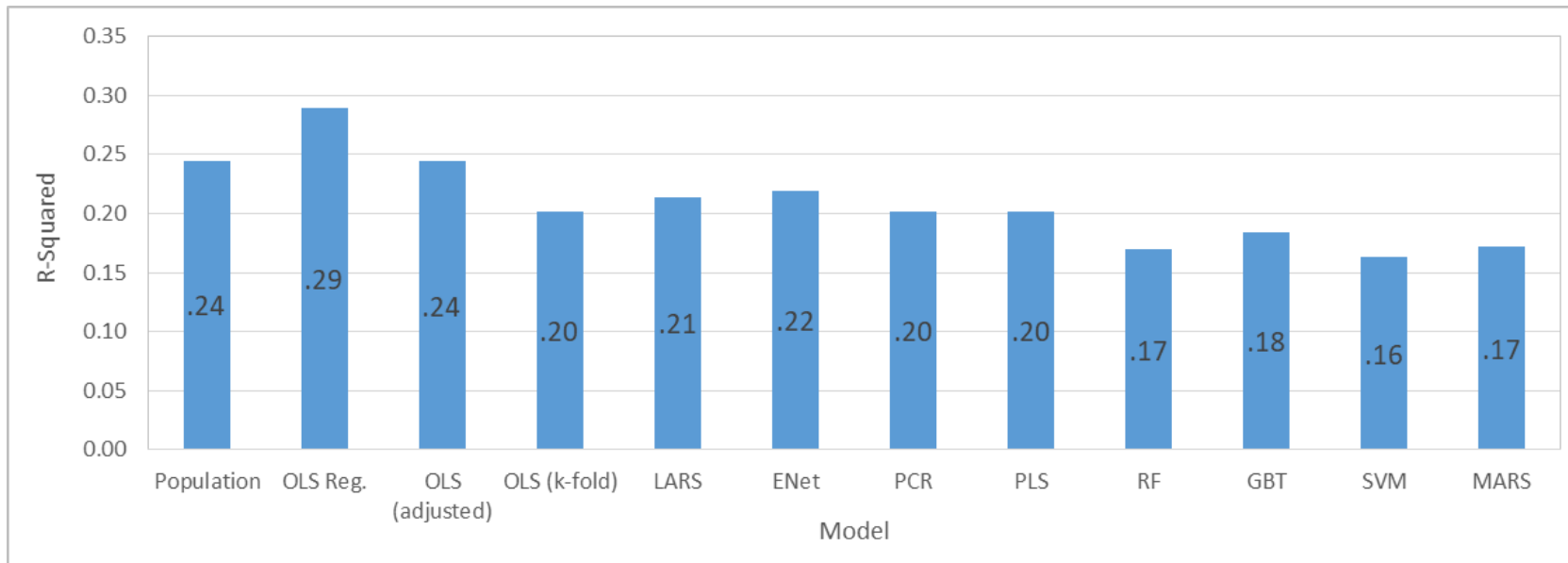
The  $R^2$  values in Figure 3 indicate a wide range of performance across models. First, the polynomial regression model estimated using regular OLS regression tended to overfit the data slightly, resulting in an  $R^2$  estimate of .27, which was slightly larger than the population estimate. This was expected given that the  $R^2$  is influenced by the number of predictors in the model and polynomial regression models include a large number of predictors. Therefore, we also report the average adjusted (for capitalization on chance)  $R^2$  for this model. The adjusted  $R^2$  was .23, which is identical to the population value. In addition, we also used  $k$ -fold cross-validation to estimate an adjusted  $R^2$  for the OLS regression model and this resulted in an average  $R^2$  of .20 across all replications. This was slightly lower than the population estimate, though the difference was small.

Figure 3 also shows the results of the various ML models estimated in these simulations. As shown in this figure, the highest  $R^2$  for any of the ML models was .21 for ENet and the smallest was .15 for SVM. In other words, the ML models tended to underperform relative to polynomial regression with an adjusted  $R^2$ . Nevertheless, this difference was relatively small. In contrast, Figure 4 shows the average correlations between the predicted criterion scores estimated from each model and the true criterion scores. As shown in this figure, the correlation between true and predicted scores based on the OLS polynomial regression model was .44 and the strongest correlation was .46 for ENet. In addition, several other ML models also showed correlations between .43 and .45. Therefore, the OLS regression model performed nearly identically to the best performing ML models.

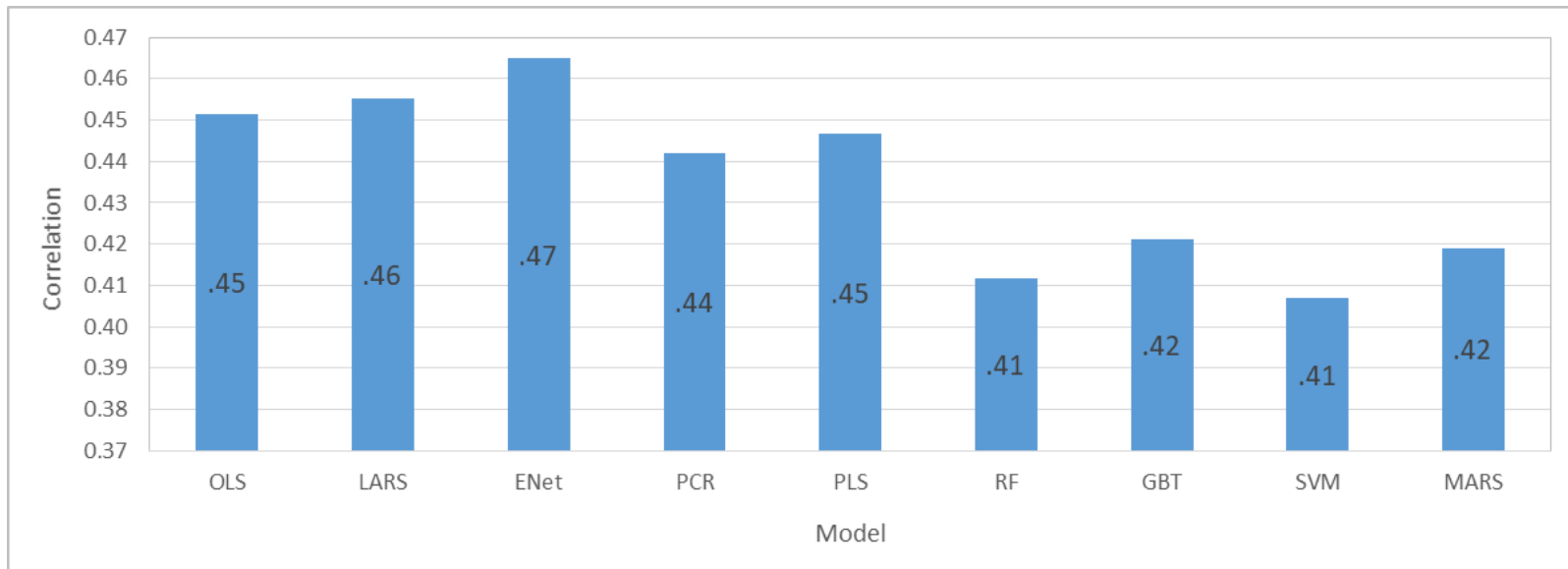
Finally, Figure 5 shows the average correlations between the predicted scores from each model and the true scores in the test sample of 10,000 cases. These correlations illustrate the generalizability of the model results to a new sample that was not used to train any of the models. The results shown in Figure 5 were very similar to the results shown in Figure 4. The correlation for the predicted scores from the OLS polynomial regression model was .46 and the strongest correlation was .48 for ENet. Therefore, there was not a substantial difference between the polynomial regression approach and the best performing ML model.

The relative variable importance for the models in which it could be estimated are shown in Table 8. As noted above, variable importance could not be determined for several of the ML models and, therefore, only a subset of the ML models are presented. The results shown in this table indicate that each of the models resulted in slightly different estimates of variable importance. For example, Figures 3-5 suggest that the OLS polynomial regression model and ENet were two of the best performing models in our simulations. However, these two approaches result in a different rank order of variables. ENet suggested that the Realistic scores were the strongest predictors whereas the polynomial regression model identified the Enterprising scores as the most important predictor. In other words, the choice of a model is important not only because of overall model performance but also because of the implications for creating weighted composites of scores.

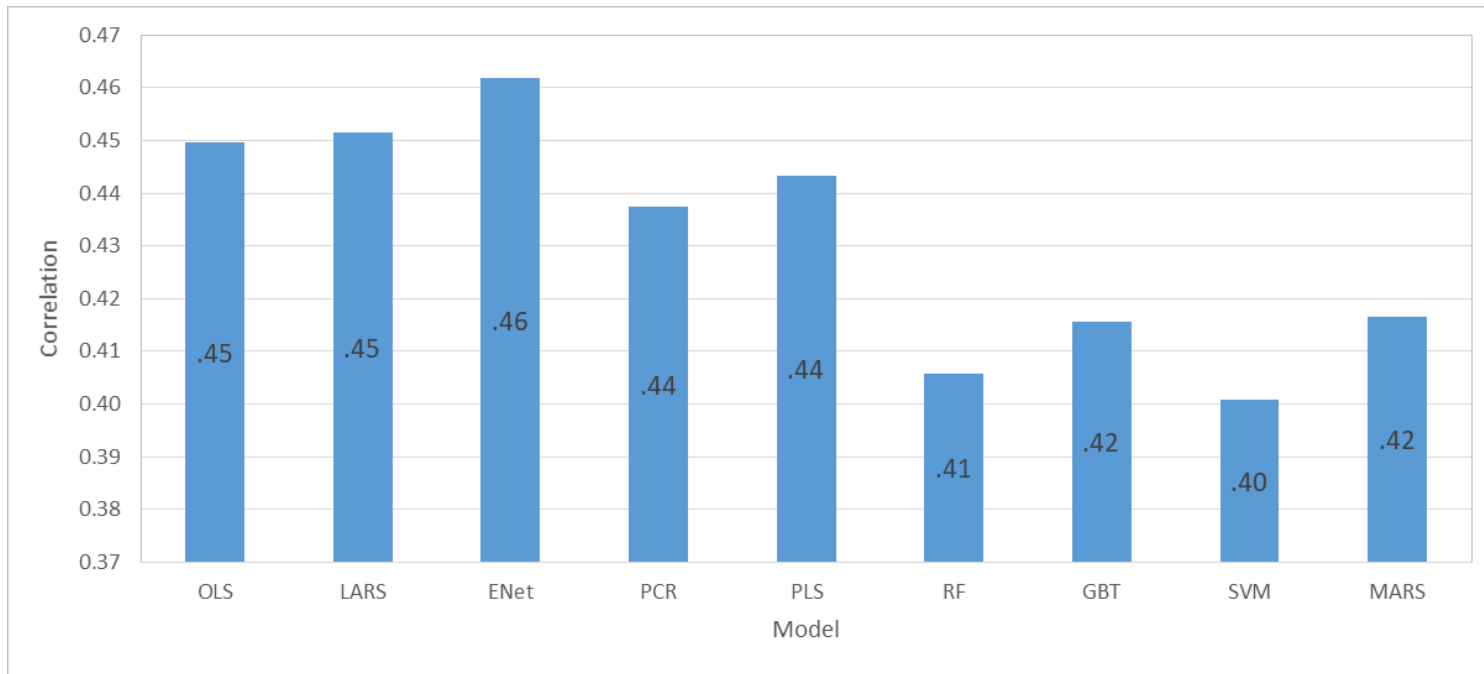
In sum, the results of the second simulation suggested that the OLS polynomial regression model with an adjusted  $R^2$  performed slightly better than the best performing ML model in terms of accurately estimating the overall validity of the model in the population. In contrast, ENet resulted in slightly higher correlations between predicted and true criterion scores. Nevertheless, in all of these cases, the differences between the polynomial regression model and the best performing ML models were trivial (e.g., differences were  $\leq .02$ ), indicating that there are not substantial differences in performance between these models in the conditions simulated here. Therefore, given that the ML models are more computationally intensive, time-consuming to estimate and properly tune, and often more difficult to interpret than OLS regression models, it seems that the polynomial regression approach is preferable to the ML models for estimating interest fit.



**Figure 3. Average  $R^2$  for Each Model in the Samples of 500**



**Figure 4. Average Correlations between Predicted and True Criterion Scores across Samples**



**Figure 5. Average Correlations between Predicted and True Criterion Scores in the Test Sample of 10,000 Cases**

**Table 8.*****Relative Variable Importance for the Prediction Models Examined in the Second Simulation Study***

<b>Predictors</b>	<b>OLS Regression</b>	<b>ENet</b>	<b>PLS</b>	<b>RF</b>	<b>GBT</b>
Realistic	58.03	78.44	43.45	43.77	43.85
Investigative	12.26	4.37	10.45	8.03	4.08
Artistic	37.67	27.71	51.65	27.07	22.73
Social	31.61	20.76	43.19	19.25	12.68
Enterprising	99.13	96.26	95.98	97.62	97.42
Conventional	26.41	20.81	23.97	13.35	10.96
Realistic (Job)	11.01	5.10	24.38	11.39	5.60
Investigative (Job)	10.95	1.99	19.62	6.78	2.49
Artistic (Job)	20.70	10.44	44.60	14.89	8.31
Social (Job)	41.44	43.69	88.26	63.55	56.98
Enterprising (Job)	66.70	58.02	68.01	52.43	45.54
Conventional (Job)	39.62	19.16	21.78	10.08	6.78
Realistic (Squared)	10.10	6.35	6.14	7.48	2.76
Investigative (Squared)	15.35	8.33	8.84	7.60	3.93
Artistic (Squared)	13.87	3.39	22.56	8.26	3.07
Social (Squared)	10.63	1.80	17.22	6.27	2.79
Enterprising (Squared)	14.46	5.14	16.02	12.13	1.84
Conventional (Squared)	11.74	4.00	9.74	6.20	2.65
Realistic (Job Squared)	14.93	5.62	6.40	7.19	3.66
Investigative (Job Squared)	10.28	2.48	8.10	6.39	3.11
Artistic (Job Squared)	13.80	2.77	20.74	6.61	3.29
Social (Job Squared)	12.26	2.15	21.75	8.93	1.68
Enterprising (Job Squared)	10.82	3.60	14.98	8.12	2.09
Conventional (Job Squared)	10.57	3.38	13.79	6.53	3.10
Realistic (Interaction)	12.86	7.49	4.21	7.91	3.22
Investigative (Interaction)	12.53	6.35	6.08	6.13	2.75
Artistic (Interaction)	10.67	2.40	11.69	7.64	3.54
Social (Interaction)	12.51	3.42	13.33	10.35	4.08
Enterprising (Interaction)	12.94	5.35	10.99	22.17	5.72
Conventional (Interaction)	8.74	3.10	6.51	6.85	3.19

*Note.* ENet = elastic net regression; PLS = partial least squares regression; RF = random forests; GBT = stochastic gradient boosted trees.

### Simulation Study 3

In the third simulation study, we examined the classification potential of the AVID. Past research has shown that the AVID can meaningfully differentiate between MOS (Nye et al., 2019). However, these results were based on a relatively small set of MOS. Therefore, we expanded this work using simulated data. In addition, we built on the simulations conducted in Studies 1 and 2 and used polynomial regression to examine the classification potential of interest fit.

#### Methods

To get a realistic a picture of the classification potential of polynomial regression, we based these simulations on the same AVID data used in Studies 1 and 2. Although these data provide a useful baseline for the simulations, there was insufficient data to get accurate estimates of classification efficiency for more than five MOS. Because previous research had already examined the classification potential in these five MOS (e.g., Nye et al., 2019), the goal of these simulations was to expand that effort and explore the classification potential when several additional jobs are considered. Therefore, we used initial data from 15 MOS but then simulated larger sample sizes that provided more accurate estimates of the classification potential of interest fit.

To simulate data based on the existing data from these MOS, we calculated the means and variance-covariance matrices for the AVID scores of both individuals and the 15 MOS. Building on the research in Studies 1 and 2, we focused on the six RIASEC scores rather than the 16 AVID dimensions. In addition, we also estimated the polynomial regression model in each MOS and recorded the point estimates and the variance-covariance matrix of the unstandardized regression weights. These point estimates and variance-covariance matrices were then used to simulate the patterns of predictive relationships in subsequent steps.

Next, we generated random samples of 300, 500, 800, and 1,000 cases from a multivariate normal distribution using the means and variances-covariance matrices from the observed data. In addition, we also randomly generated a set of interest scores for jobs using a multivariate normal distribution and the means and variance-covariance matrices from the 15 MOS in the original sample. Finally, for each of the randomly generated jobs, we also generated job-specific regression weights for a model that included both person and job interest scores predicting an overall performance outcome. Again, these regression weights were based on the means and variance-covariance matrices of the regression weights from the observed data. In other words, each hypothetical job had a unique composite of person and job interest scores that was randomly generated but based on the observed data. We then calculated the predicted scores for each simulated individual in each of the hypothetical jobs such that each case had 15 predicted performance scores (i.e., one for each hypothetical job). Finally, we generated 1,000 samples for each sample size. In other words, the results presented for this simulation study are based on 4,000 unique samples (i.e., 4 sample sizes x 1,000 replications).

After simulating the data, we next simulated the job-matching process. Unfortunately, there was not a way for us to accurately simulate the current selection and assignment process in the Army, which includes a number of factors that will vary over time and across jobs such as

the number of training seats available or the overall demand for a particular MOS. Therefore, we simulated a random selection process in which each of the simulated individuals were randomly assigned to one of the hypothetical jobs. Although this simulated condition does not reflect the actual assignment process, it does provide a baseline scenario with which we can compare the job matches based on interests. In other words, this comparison illustrates the maximum potential improvement in performance that can be achieved by matching individuals to jobs based on interests.

Next, we calculated the predicted performance scores for each simulated individual in their randomly assigned job. Similarly, we also hypothetically assigned each simulated individual to the MOS that was the best fit for their interests. To do so, simulated individuals were assigned to the MOS in which they had the highest predicted performance based on the MOS-specific interest composite. Finally, we compared the predicted performance scores for all simulated individuals under both random assignment and interest fit conditions. The differences between the predicted scores were examined and effect sizes (Cohen's  $d$ ) were calculated. In addition, building on similar research with the TAPAS (Nye et al., 2012), we calculated the proportion of simulated individuals that were predicted to perform at least .5 standard deviations better when they were matched to jobs that fit their interests compared to when they were randomly assigned.

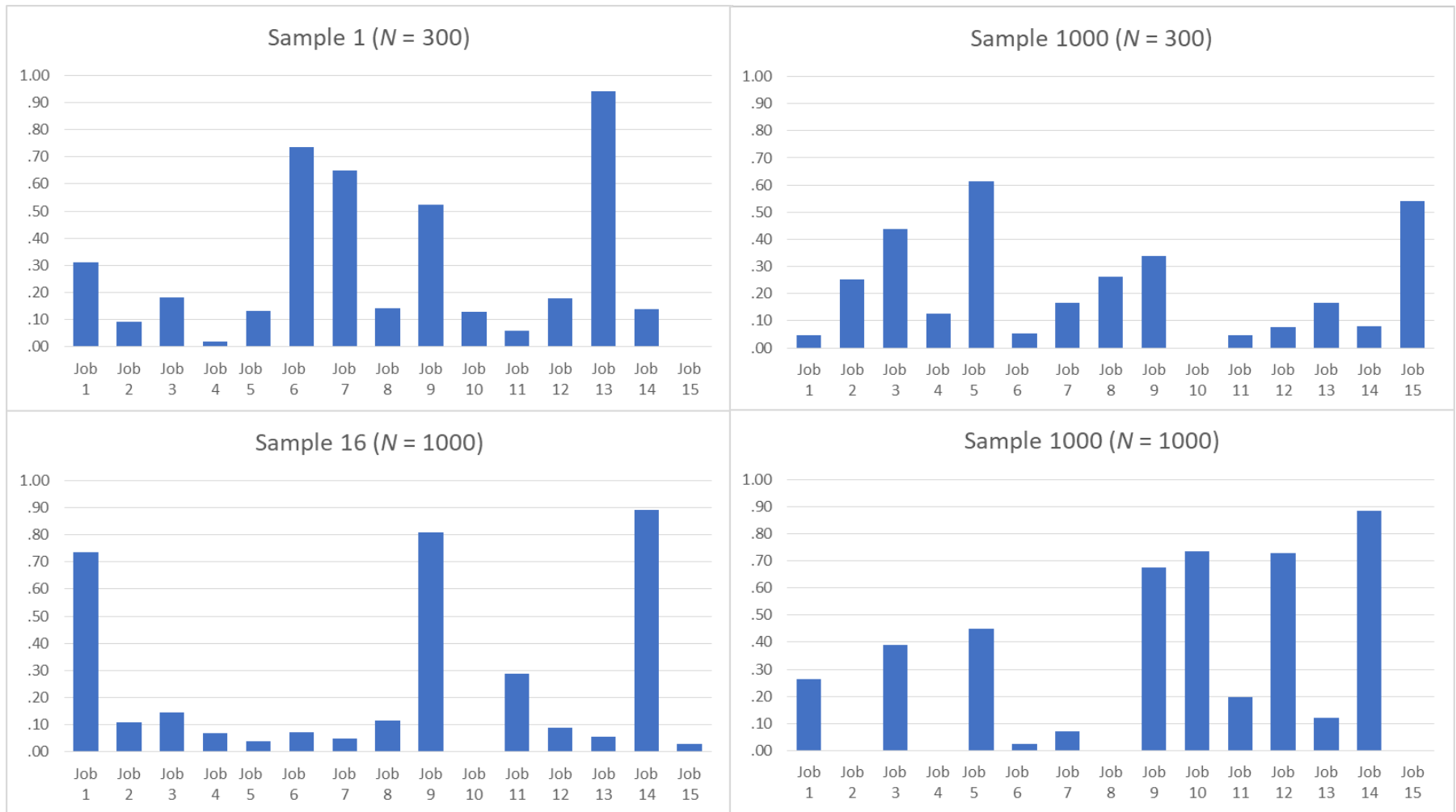
We examined the differences in predicted performance in the full sample and in subsamples of cases with both high and low interest profiles. Ryan and King (2019) proposed the idea of fit bandwidth, which suggests that some individuals will be interested in many types of activities whereas others might only be interested in a few specific types of activities. These groups are potentially important because individuals who are interested in many types of activities may fit well in several MOS due to their interest in a broad range of activities. In contrast, individuals with low interest profiles may not fit well in any MOS because they are not interested in anything. Therefore, we examined whether these different types of individuals were predicted to perform better or worse than the full sample when matched to a particular job. This was done by identifying the simulated individuals with the highest (or lowest) scores across all six interest dimensions. First we aggregated scores for each case across the six RIASEC dimensions such that the simulated individuals who scored high on all of the interest dimensions (i.e., were interested in all activities) would get a higher aggregated score while the simulated individuals who scored low on all dimensions (i.e., were not interested in anything) would get a lower aggregated score. Then we identified the cases with the highest (i.e., one standard deviation above the mean) or lowest (i.e., one standard deviation below the mean) aggregated interest scores and compared their predicted performance across the random and interest-fit conditions.

## Results

The results of the simulations in Study 3 are shown in Figure 6 and Table 9. For each hypothetical job that we simulated, the proportion of simulated individuals (i.e., across all 1,000 samples) who were predicted to perform at least half a standard deviation better in that job than in their current job ranged from .00 to .96. In other words, the percentage of cases predicted to perform better in each job varied by sample but, in some instances, nearly 100% of the sample was predicted to perform better in a particular job. To illustrate this variability, Figure 6 provides the proportions of cases that were predicted to perform better in each of the 15 simulated jobs in four representative samples. As shown in this figure, the proportions varied dramatically across jobs and samples. These results suggest that matching individuals to jobs based on their interests could potentially have a substantial effect on their overall performance. Nevertheless, it is important to remember that these hypothetical individuals were randomly assigned to their initial jobs in these simulations. Therefore, the actual effects of matching Soldiers based on interests is likely to be smaller than this baseline scenario.

To provide an alternative perspective on the potential impact of matching individuals to jobs based on their interests, Table 9 shows the effect size (i.e., Cohen's *d*) of the differences in mean predicted performance when the simulated individuals were matched to jobs that fit their interests compared to when they were randomly assigned. The results in this table are averaged across individuals and across the 1,000 samples generated for each sample size. As shown in the column for the full sample, individuals were predicted to perform substantially better when assigned to jobs that fit their interests than when they were randomly assigned to jobs. These results were consistent across sample sizes with the effect sizes ranging from 1.27 to 1.30. These effect sizes indicate that, on average, the predicted performance for all individuals increased by more than a full standard deviation when in jobs that matched their interests.

Interestingly, the predicted performance differences were slightly larger in the group with the high interest profiles than in the group with the low profiles. However, these differences were relatively small with the effect sizes ranging from 1.35 to 1.39 in the group with the highest interest profiles and from 1.25 to 1.27 in the group with the lowest interest profiles. In other words, regardless of whether individuals were interested in many things or only had weak interests, all individuals were predicted to perform substantially better when in a job that matched their interests.



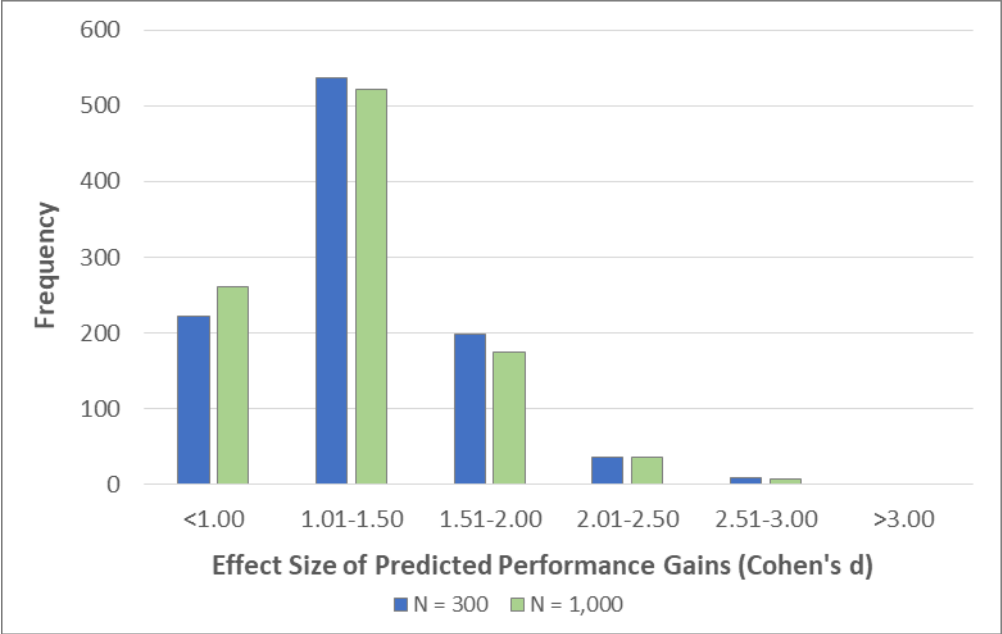
**Figure 6. Proportions of Simulated Individuals Predicted to Perform Better in Each Job across Four Different Samples**

**Table 9.**

***Effect Sizes (Cohen's  $d$ ) of the Differences in Mean Predicted Performance under the Random and Interest-Fit Assignment Conditions***

<b>Sample Size</b>	<b>Full Sample</b>	<b>High Interest Group</b>	<b>Low Interest Group</b>
300	1.29	1.38	1.27
500	1.29	1.37	1.27
800	1.27	1.35	1.25
1000	1.30	1.39	1.27

Again, the average effect size results presented in Table 9 mask the substantial variability across the simulated samples. For example, the effect sizes ranged from .49 to 2.95 across the 1,000 replications with a sample size of 300. Similarly, effect sizes ranged from .52 to 2.87 with a sample size of 1,000. These ranges of effect sizes indicate that the predicted performance gains from matching individuals to jobs based on their interests were anywhere from a half a standard deviation to a nearly three standard deviation increase, depending on the sample. Figure 7 illustrates the variability in effect sizes across samples when  $N = 300$  or 1,000. As shown in this figure, most effect sizes were between 1.00 and 2.00. In other words, even though the actual size of the effect varied across samples, the overall predicted increase in performance was still substantial in nearly all samples. The variability in the effect sizes across samples was likely due to the random assignment process that was used as the baseline for these simulations. Specifically, more individuals could have been randomly assigned into jobs that fit their interests in some samples than in others just by chance alone. This would have resulted in smaller improvements in predicted performance in these samples. Nevertheless, even the smallest improvements in performance were still substantial.



**Figure 7. Distributions of Effect Sizes across Samples**

## Summary and Conclusions

The purpose of the current research was to further examine the utility of the AVID for MOS assignment. Past research has demonstrated that the AVID can be a significant predictor of a number of important work outcomes (Nye et al., 2019). Despite these positive results, more research is needed to expand the validity evidence for the AVID and to explore ways that this assessment can be used in the Army. The research presented here furthers this work and provides additional evidence of both why and how the AVID can be used in high-stakes settings.

The results of the concurrent validation study demonstrate the predictive validity of a computer-adaptive version of the AVID. Although previous research has examined the validity of static versions of the AVID, this was the first research to examine the adaptive format. Despite the differences in format, the results of the concurrent validation study were largely similar to previous research. For example, the multiple *Rs* ranged from .22 to .50 with the CAT AVID. Similarly, previous research found multiple *Rs* ranging from .14 to .51 using a static version of the AVID that assessed the same 16 dimensions (Nye et al., 2019). In addition, the results presented in this report replicate previous research demonstrating that the validity of the adaptive version of AVID was even stronger when considering the fit between Soldiers and their MOS. These results clearly indicate the potential utility of the adaptive version of AVID and suggest that this measure is comparable to the static versions.

In addition to the concurrent validation study, we used simulations to further explore the best methods of calculating interest fit. The results of this research demonstrated that the polynomial regression approach has a number of advantages over other more traditional congruence indices. In addition, we also explored the use of several modern prediction methods based on ML to determine if these methods could further improve the calculation of interest fit. Although these methods showed some promise, they did not result in substantially stronger validity than the polynomial regression approach. Therefore, given the practical advantages of polynomial regression over ML models (e.g., they are easier to estimate and interpret), the polynomial regression approach may be preferable for estimating interest fit.

We also examined the classification potential of interest fit. The results of our simulations suggested that there was substantial variability across samples and, in some cases, nearly 100% of individuals in a sample were predicted to perform better (i.e., at least .5 SDs better) in a particular job. Importantly, matching individuals to jobs that fit their interests also resulted in substantial increases in overall predicted performance. In some samples, the predicted improvements in performance were as high as nearly three standard deviations. Although these performance increases were based on a comparison with the unlikely scenario of random assignment, these improvements in performance were substantial.

Despite the positive results found in the present research, more research is needed to clarify these findings. First, more research is needed to establish the validity of the AVID in additional MOS and in longitudinal settings. Given the relationships between interest fit and work outcomes, it is important to examine the validity of the AVID in numerous MOS. Although existing studies have demonstrated the validity of the AVID in several large MOS, more research is needed to examine this assessment in additional occupations. Moreover, the existing validation research with the AVID has been conducted using a concurrent design, with both the AVID and

criterion data collected at the same time point. Although this type of study design provides useful information about the validity of the AVID, it can also result in inflated relationships between interests and the work outcomes that are examined. Therefore, it would be informative to conduct a longitudinal study to demonstrate that the validity of the AVID can be maintained over time.

Additional research is also needed to identify the interest profiles for a broader range of MOS. As illustrated in this report, calculating interest fit requires interest scores for both individuals and jobs. The individual scores are provided by responses to the AVID but scores for each MOS are required for comparison. In the concurrent validation research reported here, the interest profiles for each MOS were obtained by asking Soldiers to rate their MOS on each of the AVID dimensions at the same time that they provided their interest scores. This approach, in which Soldiers rate potential MOS at the same time that they complete the AVID, is not feasible in a selection context where applicants may not have an accurate perception of the activities performed in a particular MOS. As a result, in order for the AVID to be most useful for initial-entry MOS assignments, interest profiles for a broader range of MOS will be needed so that each applicant's interests can be compared with several MOS. Therefore, collecting these ratings would be a useful direction for future research.

Finally, more research is needed to examine MOS assignment with real data. In the simulations presented here, we compared optimal MOS assignments based on interests to a random assignment condition. Therefore, these results provide a useful baseline for comparing future performance gains due to interest fit. However, these comparisons may not provide a realistic picture of the true gains that would be observed in practice. Therefore, similar analyses should be conducted in large samples of Soldiers from a broad range of MOS to obtain a more accurate estimate of the potential gains in performance due to interest fit.

## References

- Bateman, T. S., & Organ, D. W. (1983). Job satisfaction and the good soldier: The relationship between affect and employee "citizenship". *Academy of management Journal*, *26*, 587-595.
- Bonanno, G. A. (2004). Loss, trauma, and human resilience: Have we underestimated the human capacity to thrive after extremely aversive events?. *American Psychologist*, *59*, 20-28.
- Brown, S. P., & Gore, P. A. (1994). An evaluation of interest congruence indices: Distribution characteristics and measurement properties. *Journal of Vocational Behavior*, *45*, 310-327.
- Camp, C. C., & Chartrand, J. M. (1992). A comparison and evaluation of interest congruence indices. *Journal of Vocational Behavior*, *41*, 162-182.
- Cao, M., & Drasgow, F. (2019). Does forcing reduce faking? A meta-analytic review of forced-choice personality measures in high-stakes situations. *Journal of Applied Psychology*, *104*, 1347-1368.
- Carter, N. T., Dalal, D. K., Boyce, A. S., O'Connell, M. S., Kung, M. C., & Delgado, K. M. (2014). Uncovering curvilinear relationships between conscientiousness and job performance: How theoretically appropriate measurement makes an empirical difference. *Journal of Applied Psychology*, *99*, 564-586.
- Chan, K. Y., & Drasgow, F. (2001). Toward a theory of individual differences and leadership: Understanding the motivation to lead. *Journal of Applied Psychology*, *86*, 481-498.
- Dalal, D. K., & Carter, N. T. (2015). Consequences of ignoring ideal point items for applied decisions and criterion-related validity estimates. *Journal of Business and Psychology*, *30*, 483-498.
- Edwards, J. R. (1993). Problems with the use of profile similarity indices in the study of congruence in organizational research. *Personnel Psychology*, *46*, 641-665.
- Gati, I. (1985). Description of alternative measures of the concepts of vocational interest: Crystallization, congruence, coherence. *Journal of Vocational Behavior*, *27*, 37-55.
- Healy, C. C., & Mourtou, D. L. (1983). Derivatives of the Self-Directed Search: Potential clinical and evaluative uses. *Journal of Vocational Behavior*, *23*, 318-328.
- Holland, J. L. (1959). A theory of vocational choice. *Journal of Counseling Psychology*, *6*, 35-45.
- Holland, J. L. (1963). Explorations of a theory of vocational choice and achievement: II. A four-year prediction study. *Psychological Reports*, *12*, 547-594.

- Holland, J. L. (1973). *Making vocational choices: A theory of vocational personalities and work environments*. Odessa, FL: Psychological Assessment Resources.
- Holland, J. L. (1997). *Making Vocational Choices* (3rd ed.). Odessa, FL: Psychological Assessment Resources.
- Iachan, R. (1984). A measure of agreement for use with the Holland classification system. *Journal of Vocational Behavior, 24*, 133-141.
- Jackson, D. N. (1977). *Manual for the Jackson Vocational Interest Survey*. Port Huron, MI: Research Psychologists Press.
- Kwak, J. C., & Pulvino, C. J. (1982). A mathematical model for comparing Holland's personality and environment codes. *Journal of Vocational Behavior, 21*, 231-241.
- Liao, H. Y., Armstrong, P. I., & Rounds, J. (2008). Development and initial validation of public domain Basic Interest Markers. *Journal of Vocational Behavior, 73*, 159-183.
- Nesbitt, C. K., Salmon, E., & Kirkendall, C. D. (2020). Army Life Questionnaire. In D. J. Knapp & C. D. Kirkendall (Eds.), *Tier One Performance Screen Initial Operational Test and Evaluation: Capstone Report* (pp. 23-35) (Technical Report 1380). Ft. 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., Prasad, J., Bradburn, J., & Elizondo, F. (2018). Improving the operationalization of interest congruence using polynomial regression. *Journal of Vocational Behavior, 104*, 154-169.
- Nye, C. D., Rounds, J., Kirkendall, C.D., Drasgow, F., Chernyshenko, O. S., & Stark, S. (2019). *Adaptive Vocational Interest Diagnostic: Development and Initial Validation* (Technical Report 1378). Ft. Belvoir, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Nye, C. D., Su, R., Rounds, J., & Drasgow, F. (2012). Vocational interests and performance: A quantitative summary of over 60 years of research. *Perspectives on Psychological Science, 7*, 384-403.
- Nye, C. D., Su, R., Rounds, J., Drasgow, F. (2017). Interest congruence and performance: Revisiting recent meta-analytic findings. *Journal of Vocational Behavior, 98*, 138-151.

- Passler, K., Beinicke, A., & Hell, B. (2015). Interest and intelligence: A meta-analysis. *Intelligence, 50*, 30-51.
- Putka, D. J., Beatty, A. S., & Reeder, M. C. (2018). Modern prediction methods: New perspectives on a common problem. *Organizational Research Methods, 21*, 689-732.
- Robbins, P. I., Thomas, L. E., Harvey, D. W., & Kandefer, C. (1978). Career change and congruence of personality type: An examination of DOT-derived work environment designations. *Journal of Vocational Behavior, 13*, 15-25.
- Ryan, A. M. R. & King, D. (2019). Connecting concepts: Effects of diversity of interests and interests' effects on diversity. In C. Nye & J. Rounds (Eds.), *Vocational Interests in the Workplace: Rethinking Behavior at Work* (pp. 205-223). New York, NY: Taylor & Francis.
- Speer, A. B., Christiansen, N. D., Robie, C., & Jacobs, R. R. (2022). Measurement specificity with modern methods: Using dimensions, facets, and items from personality assessments to predict performance. *Journal of Applied Psychology, 107*, 1428-1439.
- Stark, S., Chernyshenko, O. S., & Drasgow, F. (2005). An IRT approach to constructing and scoring pairwise preference items involving stimuli on different dimensions: The multi-dimensional pairwise preference model. *Applied Psychological Measurement, 29*, 184-201.
- Stark, S., Chernyshenko, O. S., Drasgow, F., Nye, C. D., White, L. A., Heffner, T., & Farmer, W. L. (2014). From ABLE to TAPAS: A new generation of personality tests to support military selection and classification decisions. *Military Psychology, 26*, 153-164.
- 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.
- Tay, L., Drasgow, F., Rounds, J. & Williams, B. A. (2009). Fitting measurement models to vocational interest data: Are dominance models ideal? *Journal of Applied Psychology, 94*, 1287-1304.
- Trent, J. D., Barron, L. G., Rose, M. R., & Carretta, T. R. (2020). Tailored Adaptive Personality Assessment System (TAPAS) as an indicator for counterproductive work behavior: Comparing validity in applicant, honest, and directed faking conditions. *Military Psychology, 32*, 51-59.
- 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.

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

Wiegand, J. P., Drasgow, F., & Rounds, J. (2021). Misfit matters: A re-examination of interest fit and job satisfaction. *Journal of Vocational Behavior, 125*, Article 103524.

Wiggins, J. D., & Moody, A. (1981). *Compatibility index description*. Dover, DE: Training and Associates.

Wolfe, L. K., & Betz, N. E. (1981). Traditionality of choice and sex-role identification as moderators of the congruence of occupational choice in college women. *Journal of Vocational Behavior, 18*, 43-55.

Zener, T. B., & Schnuelle, L. (1976). Effects of the Self-Directed Search on high school students. *Journal of Counseling Psychology, 23*, 353-359.