

**Study
Note
2006-01**

Concurrent Validation of the NLSI for U.S. Army Drill Sergeants

**U. Christean Kubisiak, Kristen E. Horgen,
Patrick W. Connell, Elizabeth Lentz, Xian Xu,
and Walter C. Borman**

Personnel Decisions Research Institutes, Inc.

Leonard A. White and Mark C. Young
U.S. Army Research Institute

20060127 052



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

November 2005

Approved for public release; distribution is unlimited.

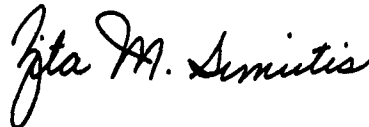
**U.S. Army Research Institute
for the Behavioral and Social Sciences**

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

Authorized and approved for distribution:



**MICHELLE SAMS
Technical Director**



**ZITA M. SIMUTIS
Director**

Research accomplished under contract
for the Department of the Army

Personnel Decisions Research Institutes, Inc.

Technical Review by

Peter M. Greenston, U.S. Army Research Institute

NOTICES

DISTRIBUTION: Primary distribution of this Study Note has been made by ARI. Please address correspondence concerning distribution of reports to: U.S. Army Research Institute for the Behavioral and Social Sciences, Attn: DAPE-ARI-MS, 2511 Jefferson Davis Highway, Arlington, Virginia 22202-3926

FINAL DISPOSITION: This Study Note may be destroyed when it is no longer needed. Please do not return it to the U.S. Army Research Institute for the Behavioral and Social Sciences.

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

REPORT DOCUMENTATION PAGE				Form Approved OMB No. 0704-0188		
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Service, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.						
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.						
1. REPORT DATE (DD-MM-YYYY) November 2005		2. REPORT TYPE Final		3. DATES COVERED (From - To) February 2004 to March 2005		
4. TITLE AND SUBTITLE Concurrent Validation of the Noncommissioned Officer Leadership Skills Inventory (NLSI) for U.S. Army Drill Sergeants				5a. CONTRACT/GRANT NUMBER DASW01-03-D-0016-0006		
6. AUTHOR(S) U. Christean Kubisiak, Kristen E. Horgen, Patrick W. Connell, Elizabeth Lentz, Xian Xu, Walter C. Borman (Personnel Decisions Research Institutes, Inc.); Leonard A. White, and Mark C. Young (U.S. Army Research Institute)				5b. PROGRAM ELEMENT NUMBER 622785		
				5c. PROJECT NUMBER A790		
				5d. TASK NUMBER 255		
				5e. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Personnel Decisions Research Institutes, Inc. 100 South Ashley Drive, Suite 375 Tampa, FL 33602				8. PERFORMING ORGANIZATION REPORT NUMBER Technical Report #499		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Institute for the Behavioral and Social Sciences 2511 Jefferson Davis Highway Arlington, VA 22202-3926				10. SPONSOR/MONITOR'S ACRONYM(S) ARI		
				11. SPONSORING/MONITORING AGENCY REPORT NUMBER Study Note 2006-01		
12. DISTRIBUTION AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.						
13. SUPPLEMENTARY NOTES Contracting Officer's Representative and Subject Matter POC: Dr. Len White.						
14. ABSTRACT The U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) and its contractor Personnel Decisions Research Institutes, Inc. (PDRI) have been conducting research to validate the Noncommissioned Officer Leadership Skills Inventory (NLSI) as a predictor of Drill Sergeant performance. The NLSI measures skills and abilities related to NCO performance, including work orientation, interpersonal skills, and leadership capability. The overall goal is to expand the NLSI into a Noncommissioned Officer classification test to identify high potential soldiers at the E-4/5/6 levels for several occupational specialties, including Drill Sergeants. The research conducted for this study consisted of a preliminary, small-sample validation of the current NLSI as a predictor of Drill Sergeant success as measured by performance ratings. Results indicate that the NLSI demonstrates preliminary, statistically significant predictive validity for Drill Sergeants. Further, this research supports the use of the NLSI as an operational test for NCO MOSs and duty assignments beyond recruiter. Future research should also guide potential refinement of the NLSI as a classification tool for multiple Army NCO positions.						
15. SUBJECT TERMS Army Drill Sergeants, Drill Sergeant Assessment, Drill Sergeant, NLSI						
SECURITY CLASSIFICATION OF:			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON (Name and Telephone Number)	
16. REPORT Unclassified			17. ABSTRACT Unclassified	18. THIS PAGE Unclassified	Unclassified Ellen Kinzer Technical Publications Specialist 703-602-8047	

Study Note 2006-01

**Concurrent Validation of the NLSI
for U.S. Army Drill Sergeants**

**U. Christean Kubisiak, Kristen E. Horgen, Patrick W. Connell,
Elizabeth Lentz, Xian Xu, Walter C. Borman
Personnel Decisions Research Institutes, Inc.**

**Leonard A. White and Mark C. Young
U.S. Army Research Institute**

**Selection and Assignment Research Unit
Michael G. Rumsey, Chief**

**U.S. Army Research Institute for the Behavioral and Social Sciences
2511 Jefferson Davis Highway, Arlington, Virginia 22202-3926**

November 2005

**Army Project Number
622785A790**

**Personnel Performance and
Training Technology**

Approved for public release; distribution unlimited.

Executive Summary

Research Requirements

The requirement of this study with Drill Sergeants is to provide an initial evaluation of the viability of the concept of using the Noncommissioned Officer Leadership Skills Inventory (NLSI) as a broader Noncommissioned Officer (NCO) selection and classification tool, with added value beyond recruiter screening. Specifically, in this study, the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) and its contractor Personnel Decisions Research Institutes, Inc. (PDRI) examined relationships between the NLSI scales and measures of Drill Sergeant performance. It was hypothesized that the NLSI measures skill sets and aptitudes related to the performance of Drill Sergeants, but these relationships had not been evaluated empirically. This study builds on work validating the NLSI for recruiter selection (Horgen et al., 2005), by examining the concept and potential utility of using the NLSI for Drill Sergeant screening and assignment.

Procedure

The research conducted for this study consisted of a preliminary validation of the NLSI as a predictor of the duty performance of Drill Sergeants. For this study, a set of 10 performance rating scales were developed by conducting job observations, reviewing existing Drill Sergeant rating measures, and generating new Behavioral Summary Scales. In conjunction with the scales, frame of reference and rater error training were developed and provided to raters prior to making their ratings. The criterion-related validity of the NLSI was evaluated by testing current job incumbents on the NLSI and collecting supervisor ratings of their job performance. For this study, NLSI predictor and performance criterion ratings were collected from a total of 195 Drill Sergeants at Fort Jackson and Fort Leonard Wood during the period June 2004 to November 2004.

Findings

This validation provides empirically documented insights into the attributes important to the successful duty performance of Drill Sergeants. Eight NLSI scales showed statistically significant correlations with a composite performance measure based on a linear combination of the 10 Drill Sergeant rating scales, with $r = .15$ to $.30$, all $p < .05$.

Utilization and Dissemination of Findings

This study documents that the NLSI measures key attributes associated with successful performance of Drill Sergeants. This report is intended to help the Army to determine how best to proceed with using the NLSI in an operational environment for NCO selection, classification, and job assignment decisions. These results support the concept that the NLSI can be used as a broader classification tool for identifying high-potentials for recruiting duty and other NCO assignments. Future research, with larger samples, should also guide potential refinement of the NLSI as a selection and classification tool for Army NCO specialties.

Table of Contents

Introduction	1
Procedure & Approach.....	1
Development of a Criterion Measure of Drill Sergeant Job Performance.....	3
Scale Retranslation Workshops.....	3
Criterion Measure Analyses.....	5
Factor Analysis of the Ratings	8
Noncommissioned Officer Leadership Skills Inventory	9
Validation Results	11
Demographic Statistics.....	11
Correlations between NLSI Scales and Performance Ratings Criterion	13
Discriminant Function Analysis.....	17
Conclusion.....	19
References	21
Appendix A – Retranslation Workshop Instructions	25
Appendix B – Results from the First Retranslation Workshop	33
Appendix C – Results from the Second Retranslation Workshop.....	37
Appendix D – Final Rating Scales	41
Appendix E - NLSI Part I Scales and Definitions	57
Appendix F – NLSI Part II Scales and Definitions	61

List of Tables

Table 1. Demographics of Experts in the First Retranslation Workshop.....	4
Table 2. Number of Supervisor and Peer Raters	6
Table 3. Interrater Reliabilities for Combined Supervisor and Peer Ratings^a	6
Table 4. Mean and Standard Deviations for Mean Ratings on Each Dimension	7
Table 5. Factor Loadings for Each Rating Dimension.....	8
Table 6. Frequencies for Drill Sergeant Race & Education Level.....	11

Table 7. Primary Type of Training.....	12
Table 8. Interest in Drill Sergeant Duty Prior to Assignment.....	12
Table 9. Correlations Between NLSI Scales and Performance Ratings.....	14
Table 10. Correlations Between NLSI Keys and Performance Ratings.....	15

List of Figures

Figure 1. Drill Sergeant Performance by NLSI Quintiles	16
---	-----------

Introduction

The U.S. Army Research Institute (ARI) and its contractor, Personnel Decisions Research Institutes, Inc. (PDRI), conducted a validation study using the Noncommissioned Officer Leadership Skills Inventory (NLSI) to identify personnel with high potential for Drill Sergeant duty. The NLSI was originally developed by ARI and PDRI in conjunction with the U.S. Army Recruiting Command (USAREC) to assist the Army in selecting Soldiers who would be successful in recruiting duty. Operational tryout of the NLSI began in January 2002, with NLSI testing of most Soldiers entering new recruiter training at the Recruiter and Retention School (RRS). The predictive validation of the NLSI against measures of recruiter performance is still in progress, but initial findings show some promising relationships (Horgen et al., 2005).

Building on the work done with recruiters, the broader vision is to expand the NLSI into a Noncommissioned Officer classification test that would enable the Army to identify high potential enlisted Soldiers at the E-4/5/6 levels for several occupational specialties, including Drill Sergeant. A better match between the Soldier and the job should result in higher levels of performance, higher job satisfaction, and higher retention rates throughout the Army.

The NLSI measures attributes and skills thought to be relevant to the performance of Drill Sergeants, but the magnitude of those relationships has not been evaluated empirically. This study, a preliminary examination of those relationships, was conducted under the sponsorship of the Deputy Chief of Staff, G1, in coordination with the Chief Psychologist, USAREC. Although based on a relatively small sample, this validation is intended to illustrate where significant relationships can be found, and to inform subsequent research and applications of the NLSI for Drill Sergeant screening.

Procedure & Approach

The strategy chosen for this project was to provide evidence of *criterion-related* validity. This is typically accomplished by testing current job incumbents and then collecting measures of these same individuals' job performance. Test scores are then related to how well individuals perform on the job. Successful validation of this type provides confirmation that use of the selection measures will, in fact, identify the most qualified candidates. This methodology is one of three validation strategies presented in the *Uniform Guidelines on Employee Selection Procedures* (1978, EEOC), the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 1999), and the *Principles for the Validation and Use of Personnel Selection Procedures* (SIOP, 2003). For this validation effort,

PDRI and ARI administered a paper and pencil version of the NLSI to a sample of Drill Sergeants, and obtained performance ratings on those Drill Sergeants from their peers and supervisors. The specifics of the data collection efforts, including criterion development, predictor selection, and preliminary validation results are provided in the following sections.

Development of a Criterion Measure of Drill Sergeant Job Performance

As mentioned, in a concurrent validation study, it is necessary to show that the test being evaluated is, in fact, related to job performance. This can only be accomplished if comprehensive, reliable, and valid measures of job performance are available. In the present criterion development effort, we began by reviewing several carefully developed performance rating scales for entry-level NCOs used in the Army's Project A (Pulakos & Borman, 1986) and ARI's ongoing NCO21 project (Knapp, McCloy, & Heffner, 2004). In addition, we incorporated work done on rating scales prepared for another ongoing Drill Sergeant study (Klein et al., 2005). The notion was to build on these existing scales and update them with new job analysis information gathered in this study.

We conducted a job analysis of the Drill Sergeant job by observing Drill Sergeants on the job, interviewing several Drill Sergeants and their supervisors, and conducting focus groups with Drill Sergeant instructors to identify the critical tasks and behaviors performed on the job, as well as relevant situational factors (e.g., stress, duty location) that impact performance. Next, we used this information, along with the existing ARI rating scales, to develop a new set of behaviorally-anchored rating scales. The new scales measure Drill Sergeant performance along 12 dimensions. These dimensions feature behavioral "anchors," that provide a description of how individuals at different levels of effectiveness perform on the job. Specifically, we developed three behavioral summary statements, anchoring high, mid-range, and low performance on each of the 12 dimensions.

Scale Retranslation Workshops

To test the adequacy of the new performance category structure and behavioral statements, we conducted a retranslation workshop with Drill Sergeant instructors. Demographic information for the instructors is listed in Table 1. We asked 22 instructors to sort the behavioral statements into the 12 performance categories and to rate each statement's level of effectiveness on a 1-3 scale, where 1 = low; 2 = mid-range; and 3 = high.

Table 1. Demographics of Experts in the First Retranslation Workshop			
Years in the Army		Years as Drill Sergeant	
Less than 10	3	Less than 2	2
10 –12	4	2	12
13 –15	9	3	8
16 –19	5		
20 or more	1		
Race		Gender	
African-American/Black	13	Male	18
White/Caucasian	7	Female	4
Filipino	1	Pay Grade	
Creole	1	E6	10
		E7	11
		E8	1

The instructions for the workshop appear in Appendix A. As is typically done in retranslation, the mean and standard deviation of the effectiveness rating were computed for each behavioral statement, along with the percentage of instructors sorting each statement into each category. These data appear in Appendix B. Data from four participants were dropped because their patterns of responses suggested they did not fully understand the retranslation instructions. To summarize, across the 36 behavioral statements and 18 instructors, 87.8% of the time instructors sorted the statements into the intended category. In all but 3 of the 648 judgments, the effectiveness level was within one scale point of the intended level. However, for several of the behavioral statements, there was sufficient disagreement in the effectiveness level or the category to warrant revisions to the statements. This was done to clarify the effectiveness level or the category membership. Finally, two of the dimensions were consolidated into other dimensions, resulting in a final set of 10 dimensions.

After these revisions were complete, we conducted an additional retranslation task with eight PDRI research staff. In this second retranslation, participants sorted the 30 behavioral statements into the intended category *and* rated the statements at the intended effectiveness level 99.4% of the time (see Appendix C). The final rating scales appear in Appendix D. These behavioral rating scales were used as the criterion measure in the validity analyses.

Criterion Measure Analyses

PDRI staff gathered performance ratings from both peers and supervisors of Drill Sergeants. Our experience with performance ratings, and discussions with ARI, indicated that both sources should provide valuable information regarding Drill Sergeants' performance. Also, obtaining ratings from multiple raters for each ratee increases the interrater reliability of the ratings. We attempted to collect ratings from at least one supervisor and two peers for each Drill Sergeant participant. The behavior-based rating scales were designed to encourage raters to make evaluations as objectively as possible. Specifically, raters were asked to compare observed Drill Sergeant behavior with the behavioral statements that anchor the different effectiveness levels on each dimension.

In addition to the scales themselves, we developed a rater training program to: (1) orient raters to the rating task; (2) familiarize raters with the performance dimensions and how each is defined; (3) train raters to match observed Drill Sergeant behavior with the behavioral summary statements to determine a rating for each dimension; (4) describe common rater errors (e.g., halo); and (5) encourage raters to be as accurate as possible when making their ratings. The rater training program was delivered in person by project staff immediately prior to the rating task. The instructors explained that the ratings were for research purposes only and would not have any impact on the Drill Sergeants' careers.

In total, performance ratings for 229 Drill Sergeants were collected from 193 peer and 62 supervisor raters. Individual raters were removed from the sample if they failed to meet at least one of two criteria. First, if the information provided by an individual rater appeared inaccurate (e.g., if the same rating was given to a Drill Sergeant across all eight dimensions), that rater was dropped. Second, we asked raters how long they had worked with the Drill Sergeants they were evaluating. We eliminated additional rater-ratee pairs where raters reported they had worked with the Drill Sergeant for less than 2 months. Raters who had worked with Drill Sergeants for less than 2 months likely had insufficient time to observe and accurately evaluate their performance. Based on the above criteria, 17 rater-ratee pairs were eliminated from the sample, yielding the final sample discussed below. As a whole, the mean number of months raters had worked with Drill Sergeants was 11.16 for peer raters and 7.98 for supervisor raters.

The final sample included performance ratings for 210 Drill Sergeants. Ratings were provided by 180 peers and 58 supervisors. Table 2 shows the number of supervisor and peer raters for each Drill Sergeant.

Table 2. Number of Supervisor and Peer Raters

Number of Supervisor Raters per Ratee	N	Number of Peer Raters per Ratee	N	Total Number of Raters per Ratee	N
1	107	1	71	1	24
2	73	2	53	2	58
		3	33	3	48
		4	19	4	30
		5	12	5	24
		6	5	6	13
				7	11
				8	2

Mean number of supervisor raters per ratee = 1.41

Mean number of peer raters per ratee = 2.29

Mean total number of raters per ratee = 3.31

The distribution of ratings across the 7-point rating scale was similar for supervisor and peer raters. There was a low, but noteworthy percentage of ratings at the lower, ineffective end of the scale for both peer and supervisor ratings. Most of the ratings fell in the 5-6 range, but overall, there was reasonable variability in both sets of ratings, suggesting that both supervisor and peer raters were differentiating between the more and less effective Drill Sergeants. Means and standard deviations across all the ratings were: 5.22 and .81 for supervisor raters, and 5.38 and .94 for peer raters.

Table 3 presents the reliabilities for the supervisor and peer ratings combined. For the majority of the rating dimensions, the reliabilities are fairly high, the primary exception being cultural tolerance. Both rating sources provide important performance information because of their unique perspectives, and the reliabilities for both sources taken together support the use of an aggregated supervisor/peer rating criterion.

Table 3. Interrater Reliabilities for Combined Supervisor and Peer Ratings^a

Rating Dimension	Combined Peer/Supervisor Reliabilities ^b
Technical Knowledge & Skill	.57
Training	.49
Counseling & Supporting Soldiers	.33
Effort & Initiative	.56
Following Rules, Regulations & Adhering to Army Core Values	.31

^aReliabilities are intraclass correlation coefficients (ICC 1,k; Shrout & Fleiss, 1979).

^bN = 695, k(harmonic mean) = 2.49

Table 3. Interrater Reliabilities for Combined Supervisor and Peer Ratings (Continued)

Rating Dimension	Combined Peer/Supervisor Reliabilities
Physical Fitness & Military Bearing	.59
Stress Tolerance & Conflict Resolution	.45
Adaptability	.34
Relating to & Supporting Peers	.34
Cultural Tolerance	.01
Overall Effectiveness	.55
Rating Composite ^c	.48

^cMean of ratings across dimensions for each rater

Rating scores were created for each Drill Sergeant by calculating the mean peer rating and the mean supervisor rating, and then averaging these two for each dimension. Table 4 shows the means and standard deviations of the combined rating scores for each dimension.

Table 4. Mean and Standard Deviations for Mean Ratings on Each Dimension

Rating Dimension	Mean ^a	Standard Deviation
Technical Knowledge & Skill	5.16	1.02
Training	5.09	.94
Counseling & Supporting Soldiers	5.03	.87
Effort & Initiative	5.16	1.08
Following Rules, Regulations & Adhering to Army Core Values	5.46	.80
Physical Fitness & Military Bearing	5.40	.94
Stress Tolerance & Conflict Resolution	4.95	1.01
Adaptability	5.17	.88
Relating to & Supporting Peers	5.33	.88
Cultural Tolerance	5.68	.67
Overall Effectiveness	5.37	.89

^aN= 210

Factor Analysis of the Ratings

To examine the underlying structure of the 10 rating scale dimensions, we conducted a principal axis factor analysis with a varimax rotation on the combined supervisor/peer dimensional ratings. Results of these analyses suggest that a two-factor solution is the most interpretable description of the data (see Table 5).

Table 5. Factor Loadings for Each Rating Dimension

Rating Dimension	Factor 1 Loadings	Factor 2 Loadings
Technical Knowledge & Skill	.21	<u>.81</u>
Training	.33	<u>.83</u>
Counseling & Supporting Soldiers	.52	.49
Effort & Initiative	.42	<u>.66</u>
Following Rules, Regulations & Adhering to Army Core Values	<u>.65</u>	.25
Physical Fitness & Military Bearing	.42	.37
Stress Tolerance & Conflict Resolution	<u>.77</u>	.22
Adaptability	<u>.68</u>	.41
Relating to & Supporting Peers	<u>.66</u>	.37
Cultural Tolerance	<u>.53</u>	.18

In general, this factor structure supports a distinction between task-related performance and contextual dimensions. Task performance refers to activities directly related to production of the goods and services that an organization produces and those that contribute less directly by helping to maintain and service this production. Contextual performance, on the other hand, refers to activities that support the social and psychological environment in which task performance takes place (Borman & Motowidlo, 1993).

For the factor analysis of the Drill Sergeant data, most of the “contextual” dimensions, Following Rules, Stress Tolerance & Conflict Resolution, Adaptability, Relating to & Supporting Peers, and Cultural Tolerance, showed the strongest loadings with Factor 1. In contrast, the more task-related dimensions, Technical Knowledge & Skill, Training, and Effort & Initiative, were most strongly associated with Factor 2. These results suggest that the combined supervisor/peer dimensional ratings reflect multiple aspects of Drill Sergeant performance.

Noncommissioned Officer Leadership Skills Inventory

The NLSI consists of two parts. Part I contains 125 items that measure past behaviors and reactions to life events indicative of such areas as leadership, interpersonal skills, and openness. This instrument is based on the Army's Background Information Questionnaire (BIQ). Previous research has demonstrated that these scales are predictive of counterproductive behavior, Special Forces job performance, completion of the Special Forces Assessment and Selection course, and disciplinary infractions among NCOs and first term enlisted personnel (e.g., Kilcullen, Chen, Zazanis, Carpenter, & Goodwin, 1999a; Kilcullen, Mael, Goodwin, & Zazanis, 1999b; Knapp et al., 2004). Additionally, in research with Army civilians, the Part I Tolerance for Ambiguity, Openness, Emergent Leadership, and Social Perceptiveness scales were related to measures of job performance (Kilcullen, White, Zaccaro, & Parker, 2000). The NLSI has also effectively predicted recruiting success as measured by ratings of recruiter performance and recruiter production (Borman, White, Bowles, Horgen, Kubisiak, & Penney, 2003). Thus, the scales in Part I have shown criterion-related validity in military settings and measure constructs that may be relevant for Drill Sergeant success (e.g., Leadership, Interpersonal Skill). Definitions for the NLSI scales in Part I can be found in Appendix E.

Part II of the NLSI uses a forced-choice format to reliably measure six temperament constructs: Dependability (Non-delinquency), Adjustment, Work Motivation, Leadership, Agreeableness, and Physical Conditioning. The scales and definitions can be found in Appendix F. Part II of the NLSI is an expanded version of the Army's Assessment of Individual Motivation (AIM) with additional items added to improve the internal consistency reliability of the scales and better construct balance in the item tetrads. Each item in Part II consists of four behavioral statements that represent different personality constructs. Soldiers select a statement that is most like and a different statement that is least like themselves. An attempt is made to balance the social desirability of the statements within each item to reduce the AIM's susceptibility to faking.

In the Army's Project A and Career Force research these temperaments were measured by a self-report instrument called the Assessment of Background and Life Experiences (ABLE). The Project A results, involving nearly 60,000 enlisted personnel, established that individual differences in these constructs, as measured by ABLE, are important predictors of the duty performance and attrition of enlisted personnel and Noncommissioned Officers (Campbell & Knapp, 2001; Hough, Eaton, Dunnette, Kamp, & McCloy, 1990; Rumsey, Peterson, Oppler & Campbell, 1996; White, Young, & Rumsey, 2001). The AIM was designed to

measure these constructs from ABLE with less fakability. Preliminary findings indicate that the AIM is more resistant to deliberate faking than the ABLE (Young, Heggstad, Rumsey, & White, 2000; Young, McCloy, Waters & White 2004; White & Young, 2001). In a series of investigations, the AIM has been found to be predictive of measures of Soldier performance, adaptability, and attrition during the first term of enlistment (White & Young, 1998; Young et al., 2000; Young et al., 2004; Young, White, Heggstad, & Barnes, 2004; White, Young, Heggstad, Stark, Drasgow, & Piskator, 2004).

In other research, several scales of the AIM were linked to Special Forces job performance (Kilcullen et al., 1999a), first term attrition (White, Nord, Mael, & Young, 1993), Correctional Specialist performance (White & Young, 2001), and the successful completion of Explosive Ordnance Disposal (EOD) training (White & Young).

These studies suggest that the AIM has promise for measuring constructs important for Drill Sergeant performance (e.g., Work Motivation, Adjustment).

Validation Results

This section describes the Drill Sergeant sample and summarizes the analyses of the relationships between the NLSI scales and measures of Drill Sergeant performance.

Predictor and criterion data were collected from 195 Drill Sergeants at Ft. Leonard Wood and Ft. Jackson in June through November 2004. For each session, PDRI conducted rater training, and then Drill Sergeants and their supervisors completed performance ratings. Next, we administered the NLSI to Drill Sergeants. The NLSI took approximately 1½ hours to complete.

Predictor data were screened according to several data checks. First, four items were included in the NLSI to ensure that respondents were paying attention. Participants who responded in nonsensical ways to these items were dropped from subsequent analyses. Second, data were screened for patterns of responses that suggested they were not paying attention to item content. Finally, cases that had substantial amounts of missing data were dropped. Based on these data screens, 14 cases were excluded, leaving a sample size of 181 cases.

Demographic Statistics

Table 6 provides demographic information for the Drill Sergeants across the two data collection sites. Eighty-five percent of the sample was male; 38.6 % of the sample was Black, and 13% of the Drill Sergeants indicated they were of Hispanic, Latino, or Spanish ancestry.

Table 6. Frequencies for Drill Sergeant Race & Education Level

Category	Frequency	Percentage
White	72	43.3
Black	64	38.6
Hispanic/Latino/Spanish	22	13.3
Other	8	4.8
Missing	15	
Total	181	

Table 6. Frequencies for Drill Sergeant Race & Education Level (Continued)

Category	Frequency	Percentage
Less than 12 years of school (no diploma)	1	.6
High school diploma or GED	11	6.2
Some college, but did not graduate	103	57.9
Two-year college degree	51	28.6
Four-year college degree	8	4.5
Some graduate school	2	1.1
Graduate degree	2	1.1
Missing	3	
Total	181	

As shown in Table 7, most of the Drill Sergeants teach in Basic Combat Training (BCT), One-Station-Unit-Training (OSUT), or Advanced Individual Training (AIT). More than 54% of the Drill Sergeants in our sample indicated they were either extremely or very interested in Drill Sergeant duty prior to their assignment, and 54% of them volunteered for Drill Sergeant duty (see Table 8).

Table 7. Primary Type of Training

Primary Type of Training	Frequency	Percent
BCT	100	58.5
OSUT	58	33.9
AIT	13	7.6
Missing	10	
Total	181	

Table 8. Interest in Drill Sergeant Duty Prior to Assignment

Interest Level	Frequency	Percent
Extremely interested	46	26.9
Very interested	48	28.1
Somewhat interested	33	19.3
Not very interested	20	11.7
Not at all interested	24	14.0
Missing	10	
Total	181	

Correlations between NLSI Scales and Performance Ratings Criterion

Table 9 shows correlations between NLSI scales and the performance rating dimensions, as well as rating factors. Several of the NLSI scales correlated significantly with the performance ratings (e.g., Work Motivation, Leadership, Social Perceptiveness, Interpersonal Skills). In addition, the pattern of relationships between the NLSI scales and performance rating dimensions demonstrate construct validity. For example, because Drill Sergeants are required to train physical skills, we expected the Physical Conditioning scale to correlate significantly with the Technical Knowledge and Skill and the Physical Fitness rating dimensions. The scales in Part II did correlate significantly with these two rating dimensions. Also, the Tolerance for Ambiguity scale correlated significantly with the Adaptability rating dimension. Finally, several of the NLSI scales correlated significantly with the contextual and task-performance rating dimension factors, specifically, Emergent Leadership, Tolerance for Ambiguity, Conscientiousness, Work Motivation, and Leadership with both factors and Social Perceptiveness, Interpersonal Skills, and Self-Esteem with the task performance factor. Overall, the NLSI demonstrated many significant correlations between the scales and performance rating dimensions and the rating composite. In addition, these correlations follow a meaningful pattern, indicating that the NLSI is measuring the intended attributes.

Table 9. Correlations Between NLSI Scales and Performance Ratings

NLSI scales	Factor 1 Contextual	Factor 2 Task Related	Composite of 10 Rating Dimensions	Overall Perfor- mance		Technical Know & Skills		Training		Counseling		Effort & Initiative		Following Rules & Regs		Physical Fitness		Stress Tolerance		Adaptability		Supporting Peers		Cultural Tolerance	
				Dimension Rating	Rating	Dimension Rating	Rating	Dimension Rating	Rating	Dimension Rating	Rating	Dimension Rating	Rating	Dimension Rating	Rating	Dimension Rating	Rating	Dimension Rating	Rating	Dimension Rating	Rating	Dimension Rating	Rating	Dimension Rating	Rating
Pt. 1 Emergent Leadership	.22	.37	.30	.32	.35	.35	.29	.18	.35	.00	.11	.19	.28	.22	.11	.11	.19	.28	.22	.11	.11	.22	.09	.05	.11
Pt. 1 Social Perceptiveness	.10	.18	.15	.14	.17	.17	.12	.10	.17	.00	.11	.10	.13	.09	.11	.11	.10	.13	.09	.11	.10	.09	.05	.05	.05
Pt. 1 Hostility	-.05	.03	-.03	-.05	.09	.09	.00	-.16	.00	-.05	.01	-.05	-.05	.00	.01	.01	-.05	-.05	.00	-.04	.00	.00	-.04	-.04	-.04
Pt. 1 Tolerance for Ambiguity	.15	.14	.15	.16	.17	.17	.09	.09	.11	-.01	.04	.11	.23	.16	.04	.04	.11	.23	.16	.07	.16	.10	.07	.07	.07
Pt. 1 Interpersonal Skills	.14	.19	.18	.17	.16	.16	.14	.21	.21	.04	.07	.17	.13	.10	.07	.07	.17	.13	.10	.07	.10	.10	.07	.07	.07
Pt. 1 Self-Esteem	.06	.20	.16	.16	.30	.30	.20	.06	.12	-.04	.22	.04	.11	.09	.22	.22	.04	.11	.09	.00	.09	.09	.00	.00	.00
Pt. 1 Empathy	.01	-.07	-.01	-.15	-.09	-.09	-.03	.17	-.05	.00	-.06	.03	-.01	.01	-.06	-.06	.03	-.01	.01	.00	.01	.01	.00	.00	.00
Pt. 1 Conscientiousness	.21	.27	.26	.27	.25	.25	.21	.09	.26	.03	.24	.12	.33	.23	.03	.24	.12	.33	.23	.08	.23	.14	.08	.08	.08
Pt. 2 Work Orientation	.16	.22	.20	.28	.22	.22	.13	.08	.19	.06	.14	.09	.26	.14	.06	.14	.09	.26	.14	.10	.14	.14	.10	.10	.10
Pt. 2 Leadership	.17	.29	.20	.28	.31	.31	.20	.14	.16	.08	.15	.09	.17	.03	.08	.15	.09	.17	.03	.12	.03	.03	.12	.12	.12
Pt. 2 Dependability	-.02	-.08	-.05	-.01	-.13	-.13	-.05	.06	-.02	.01	-.09	.02	-.02	.01	.01	-.09	.02	-.02	.01	-.11	.01	.01	-.11	-.11	-.11
Pt. 2 Adjustment	.11	.07	.09	.08	.08	.08	.05	-.03	.05	.07	.08	.11	.11	.07	.08	.08	.11	.11	.10	.00	.10	.10	.00	.00	.00
Pt. 2 Agreeableness	.07	-.03	.04	-.04	-.08	-.08	.01	.10	-.01	.01	.01	.10	.05	.01	.01	.01	.10	.05	.09	.01	.09	.09	.01	.01	.01
Pt. 2 Physical Conditioning	.09	.13	.13	.17	.17	.17	.08	-.02	.09	.05	.21	.07	.14	.05	.05	.21	.07	.14	.07	.14	.07	.07	.07	.01	.01

N = 176-181

Note: Correlations underscored are significant ($p < .05$); ratings are mean peer & supervisor ratings.

In addition to correlations between NLSI scales and job performance, we also investigated the relationship between NLSI total scores and Drill Sergeant performance ratings. Because our Drill Sergeant sample precluded development of a cross-validated scoring key, we used a key developed as part of research on the validity of the NLSI for predicting recruiter performance (see Borman et al., 2003; Horgen et al., 2005; and White, Borman, & Bowles, 2001). The key was developed and cross-validated using a sample of approximately 4,800 recruiters. Further, we computed a validity-weighted NLSI composite for the Drill Sergeants. We validity-weighted and combined those scales that correlated significantly with the rating composite in the Drill Sergeant sample to compute an overall NLSI score. That is, we used the current sample to determine the correlations between individual scales and the composite performance rating. Then we created a predictor composite score by combining those scales weighted by the correlations. Both keys were correlated with the performance rating dimensions and composite.

As shown in Table 10, the two keys are highly correlated. More importantly, scoring the NLSI with either of the keys results in significant correlations between the NLSI and most of the Drill Sergeant performance rating dimensions, the rating composite, and the task and contextual performance factors. Ideally, the correlation between the NLSI total score based on the validity weighted Drill Sergeant composite should correlate more highly with the performance ratings than NLSI based on the recruiter key (.33 vs. .37, respectively), but again, these results are preliminary and based on a small sample.

Table 10. Correlations Between NLSI Keys and Performance Ratings

	NLSI Total Score Based on Recruiter Key	NLSI Total Score Based on Validity Weighted Drill Sergeant Composite
NLSI Total Score Based on CV Key	<u>.84</u>	<u>.92</u>
NLSI Total Score Based on PV Key	1.00	<u>.83</u>
NLSI Total Score Based on Validity Weighted Drill Sergeant Composite	<u>.83</u>	1.00
Composite of 10 Rating Dimensions	<u>.31</u>	<u>.29</u>
Technical Know & Skills Dimension Rating	<u>.38</u>	<u>.35</u>
Training Dimension Rating	<u>.28</u>	<u>.24</u>
Counseling Dimension Rating	<u>.22</u>	<u>.15</u>
Effort & Initiative Dimension Rating	<u>.32</u>	<u>.28</u>
Following Rules & Regs Dimension Rating	.05	.05
Physical Fitness Dimension Rating	<u>.24</u>	<u>.23</u>
Stress Tolerance Dimension Rating	.12	<u>.17</u>

Table 10. Correlations Between NLSI Keys and Performance Ratings (Continued)

	NLSI Total Score Based on Recruiter Key	NLSI Total Score Based on Validity Weighted Drill Sergeant Composite
Adaptability Dimension Rating	<u>.32</u>	<u>.31</u>
Supporting Peers Dimension Rating	<u>.17</u>	<u>.19</u>
Cultural Tolerance Dimension Rating	<u>.10</u>	<u>.12</u>
Overall Performance Dimension Rating	<u>.38</u>	<u>.32</u>
Factor 1 Contextual Performance	<u>.20</u>	<u>.22</u>
Factor 2 Task Performance	<u>.37</u>	<u>.33</u>

N = 177-181

Note: Correlations in bold are significant ($p < .05$); ratings are mean peer & supervisor ratings

Figure 1 illustrates the relationship between NLSI scores and the percentage of high performing Drill Sergeants. The highest score quintiles contain the highest scorers on the NLSI. Again, these analyses are based on a small sample, but they support the validity of the NLSI as applied to Drill Sergeants.

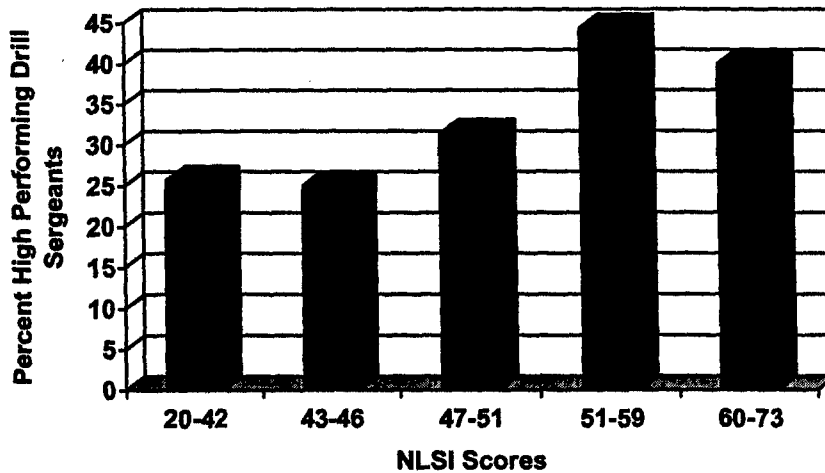


Figure 1. Drill Sergeant Performance by NLSI Quintiles

Discriminant Function Analysis

PDRI also conducted a discriminant function analysis to determine whether the NLSI could distinguish between high performing Drill Sergeants and high performing recruiters. For these analyses, we selected the 107 Drill Sergeants who constituted the top 50 percent of performers with regard to number of recruits in the sample, and randomly selected an equal number of recruiters from among the top 50 percent of producers in the recruiter database from a previous research project (Horgen et al., 2005). Entering the NLSI scales simultaneously, one discriminant function was calculated ($\chi^2=51.44, p < .001$; Wilks' $\Lambda = .756, F(1, 212) p < .001$). Based on the derived classification equation, 70.6 percent of the cases in the sample would be reclassified into the correct group. These results suggest that the NLSI scales can potentially discriminate between the Drill Sergeants and recruiters. Results of these analyses must be interpreted cautiously, due to the small sample size. However, these preliminary results suggest that the NLSI may be useful for classification purposes.

Conclusion

This preliminary investigation indicates that the NLSI is related to Drill Sergeant performance. The present research extends previous NLSI research and demonstrates that the NLSI may be useful in assigning Soldiers to second tour duty assignments, such as Drill Sergeant or recruiter. Additional work is required to investigate the validity of the NLSI in a *predictive* setting and with a larger number of Drill Sergeants.

In addition, future research should address the relationship of the NLSI with other predictors, such as disciplinary infractions, training performance, and attrition. Further, analyses of additional criterion measures, such as attrition and success in training may be of interest. The use of disciplinary infractions as a criterion would be of particular interest with the Drill Sergeant population. However, the prediction of relatively rare occurrences along with potential irregularities in the recording of such infractions and the dismissal of participants with high rates of infractions could limit the utility of such analyses.

Due to the relatively small sample size in this research, we were not able to fully investigate the usefulness of the NLSI as a classification tool. Currently, the NLSI has a substantial history of use with the recruiter population, and Drill Sergeants are the next step in exploring its applicability to other duties. That is, because the two jobs overlap with regard to the emphases on social skills, leadership, and positive representation/teaching about Army life, somewhat similar results were expected for the two groups. Going forward, the NLSI should be refined as a classification tool so that it will optimally differentiate between predictors tailored to the appropriate MOSs and duty assignments. The preliminary analyses conducted here suggest that classification based on NLSI scores is possible, but additional research would allow a more comprehensive comparison between the NLSI profiles of these duty assignments.

References

- American Educational Research Association, American Psychological Association, National Council on Measurement in Education (1999). *Standards for Educational and Psychological Testing*. Washington, DC: American Psychological Association.
- Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In Schmitt, N., & Borman, W. C. (Eds.), *Personnel selection in organizations* (pp. 71-98). San Francisco: Jossey-Bass Publishers.
- Borman, W. C., White, L. A., Bowles, S. V., Horgen, K. E., Kubisiak, U. C., & Penney, L. M. (2003, November). *Update on recruiting research in the U.S. Army*. Paper presented at the 45th Annual Conference of the International Military Testing Association, Pensacola, Florida.
- Browne, M. (1975). Predictive validity of a linear regression equation. *British Journal of Mathematical and Statistical Psychology*, 28, 79-87.
- Campbell, J.P., & Knapp, D.J. (2001). *Exploring the Limits in Personnel Selection and Classification*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Horgen, K. E., Kubisiak, U. C., Bruk-Lee, V., Connell, P. W., Penney, L. M., Borman, W. C., Pace, V., & Lentz, E. (2005). *Evaluation and refinement of screening and assessment tools for the U.S. Army*. (Institute Report 498). Tampa, FL: Personnel Decisions Research Institutes, Inc.
- Hough, L.M., Eaton, N.K., Dunnette, M.D., Kamp, J.D., & McCloy, R.A. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities. *Journal of Applied Psychology*, 75, 581-595.
- Kilcullen, R. N., Chen, G., Zazanis, M. M., Carpenter, T., & Goodwin, G. (1999a). *Adaptable performance in unstructured environments*. Paper presented at the annual meeting of the Society for Industrial and Organizational Psychology, Atlanta, GA.
- Kilcullen, R. N., Mael, F. A., Goodwin, G. F., & Zazanis, M. M. (1999b). Predicting U.S. Army Special Forces field performance. *Journal of Human Performance in Extreme Environments*, 4, 53-63.

- Kilcullen, R. N., White, L. A., Zacarro, S., & Parker, C. (2000). *Predicting managerial and executive performance*. Paper presented at the 15th Annual Conference, Society for Industrial and Organizational Psychology, New Orleans, LA.
- Klein, G., Salter, M., Gates, J. W., Sullivan, R., Kinnison, H., Lappin, M., & Graham, S. E. (2005). *Sergeants as Drill Sergeants: Returning Sergeants to Drill Sergeant duty*. (ARI Technical Report 2005-04) Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Knapp, D.J., McCloy, R.A., & Heffner, T.S. (2004). *Validation of measures designed to maximize 21st-Century Army NCO Performance*. (ARI Technical Report 1145). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Pulakos, E. D., & Borman, W. C. (Eds.) (1986). Development and field test of Army-wide rating scales and the rater orientation and training program. *Project A: Improving the selection, classification, and utilization of Army enlisted personnel - Development and field test of Army-wide rating scales and the rater orientation and training program*. Alexandria: VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Rumsey, M.G., Peterson, N.G., Oppler, S.H., & Campbell, J.P. (1996). What's happened since Project A: The future career force. *Journal of the Washington Academy of Sciences*, 84, 94-110.
- Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, 86, 420-428.
- Society for Industrial and Organizational Psychology, Inc. (2003). *Principles for the validation and use of personnel selection procedures* (4th ed.). Bowling Green, OH: Author.
- Uniform guidelines on employee selection procedures, 43 Fed.Reg. 38290-38315 (1978).
- Wherry, R. (1931). A new formula for predicting shrinkage of the coefficient of multiple correlation. *Annals of Mathematical Statistics*, 2, 440-457.
- White, L. A., Borman, W. C., & Bowles, S. (2001, August). *New tools for selecting high potential Army recruiters*. Paper presented at the 109th Annual Convention of the American Psychological Association. San Francisco, CA.
- White, L.A., Nord, R.D., Mael, F.A., & Young, M.C. (1993). The Assessment of Background and Life Experiences (ABLE). In T. Trent and J.H. Laurence (Eds.), *Adaptability screening for the armed forces*. Washington, D.C.: Office of Assistant Secretary of Defense (Force Management and Personnel).

- White, L.A., & Young, M.C. (1998). *Development and validation of the Assessment of Individual Motivation (AIM)*. Paper presented at the annual meeting of the American Psychological Association, Washington, D.C.
- White, L.A., & Young, M. C. (2001, April). *Validation of a faking-resistant measure of temperament constructs*. Paper presented at the 16th Annual Conference, Society for Industrial and Organizational Psychology, San Diego, CA.
- White, L.A., Young, M.C., Heggestad, E.D., Stark, S., Drasgow, F., & Piskator, G. (2004, November). *Development of a non-high school diploma graduate pre-enlistment screening model to enhance the Future Force*. Proceedings of the 24th annual Army Science Conference, Orlando, FL.
- White, L.A., Young, M.C., & Rumsey, M.G. (2001). Assessment of Background and Life Experiences (ABLE) implementation issues and related research. In J.P. Campbell and D. Knapp (Eds.). *Exploring the Limits in Personnel Selection and Classification*. Hillsdale, N.J.: Earlbaum.
- Young, M. C., Heggestad, E. D., Rumsey, M. G., & White, L. A. (2000). *Army pre-implementation research findings on the assessment of individual motivation (AIM)*. Paper presented at the annual meeting of the American Psychological Association, Washington, D.C.
- Young, M.C., McCloy, R.A., Waters, B.K., & White, L.A. (2004). Introduction: An overview of AIM and preliminary efforts to support its operational use. In Knapp, D.J., Heggestad, E.D., & Young, M.C. (Eds.), *Understanding and Improving the Assessment of Individual Motivation (AIM) in the Army's GED Plus Program*. (ARI Study Note 2004-03). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.
- Young, M.C., White, L.A., Heggestad, E.D., & Barnes, J.D. (2004, July). *Operational validation of the Army's new pre-enlistment screening measure*. Paper presented at the annual meeting of the American Psychological Association, Honolulu, HI.

Appendix A – Retranslation Workshop Instructions

Army Drill Sergeant Performance Rating Scales

Retranslation Workshop Instructions

Background

Personnel Decisions Research Institutes (PDRI) was tasked by the Army to develop behavior-based performance rating scales for the Army drill sergeant job. We have developed this kind of performance rating format for a number of jobs in industry and for other U.S. military jobs.

Purpose of the Workshop

In this workshop, you will be providing information that will be used to develop an instrument for rating the job performance of Army drill sergeants. We have obtained information from many sources inside the Army regarding observed drill sergeant *behaviors*. This information was summarized into behavior statements reflecting, respectively, high, mid-range and low drill sergeant performance.

In today's workshop, we are asking you to read these performance statements, place them in categories we will show you in a moment, and rate the effectiveness of the behavior described in each performance statement.

Completing the Performance Example Rating Task

For each of 36 performance statements, we ask you to make two judgments:

1. Determine the Army drill sergeant performance category in which the statement best fits (e.g., technical knowledge and skill); and
2. Rate the effectiveness of the behavior described in the statement.

Before you begin making your judgments, please review the *Army Drill Sergeant Performance Categories* carefully. This handout lists and defines a set of categories relevant to the performance of Army drill sergeants. Once you have become familiar with these performance categories, you will be ready to begin the rating task.

The effectiveness ratings that you assign to each performance statement will range from 1 to 3, as follows:

- 1 = Low
- 2 = Mid-range
- 3 = High

To help calibrate your effectiveness ratings, we provide a couple of example statements to clarify the distinctions between the levels of performance.

- A. Responds effectively when duties are disrupted by routine changes in assignments, but has some difficulty if the changes are due to emergencies that arise.
- B. Always willing to lend a hand when colleagues appear overwhelmed or behind schedule.

- C. Fails to use time wisely; for example, during training exercises, might spend time talking with other drill sergeants rather than observing trainees.

Notice that Example A is adequate but probably not as effective as hoped for, so a 2 or "mid-range" rating might be the most appropriate effectiveness rating. Example B is probably more appropriately at the 3 or "effective" level, as the example depicts superior performance. Example C probably deserves a 1 rating due to the "ineffective" level of performance described.

Now please open the envelope containing the performance statements and remove them. Sort each statement into one of the 12 Performance Categories and also rate the effectiveness level (1, 2, or 3) of each statement. Probably the best way to do this is to:

1. Complete the sorting of all 36 statements into the 12 Performance Categories. This should result in about 3 statements per category.
2. Review the performance statements *within* each category, decide on the effectiveness level of each, and record the category letter (A-L) and the effectiveness rating (1 = low, 2 = mid-range, 3 = high) on each statement in the blanks provided.
3. For each category, place a paper clip on the statements that go together in a category (e.g., all A statements).
4. Place all the paper-clipped statements back in the envelope.

Thank you for helping us with this task.

Army Drill Sergeant Performance Categories

A. Technical Knowledge & Skill

Demonstrating technical knowledge and skill; providing clear and accurate instructions; knowledgeably answering questions about training tasks; demonstrating training tasks properly.

B. Training

Using appropriate training methods; presenting well-prepared and organized material for training exercises and sessions; using instructional techniques; providing constructive feedback regarding performance on training tasks.

C. Coaching, Mentoring, & Supporting Trainees

Coaching and mentoring trainees; demonstrating concern for the well-being and development of trainees; demonstrating respect for trainees; listening attentively to trainees, and asking questions as appropriate; supporting trainees and helping them to overcome personal problems.

D. Counseling

Preparing for counseling sessions; counseling trainees and offering helpful feedback; following Buddy System regulations.

E. Effort & Initiative

Persisting with extra effort even under difficult conditions; taking initiative to accomplish objectives; finding additional productive work when own duties are completed; developing knowledge and skills through additional training.

F. Integrity & Adherence to Army Core Values

Demonstrating integrity, ethical behavior, and self-discipline; displaying respect for authority; adhering to Army Core Values; and obeying fraternization policies.

G. Following Rules, Regulations & Safety Guidelines

Adhering to Army regulations, orders, and SOP; following safety guidelines and monitoring trainee safety.

H. Physical Fitness & Military Bearing

Maintaining physical fitness and proper military bearing; displaying discipline and maintaining good professional conduct.

I. Stress Tolerance & Conflict Resolution

Managing stress and maintaining self-control; asking for appropriate backup in potentially volatile situations; resolving conflicts quickly and fairly.

J. Adaptability

Effectively adapting to changing circumstances.

K. Relating to & Supporting Peers

Effectively relating to and working with other Drill Sergeants; helping others by performing some of their tasks when needed; supporting, motivating and showing confidence in others.

L. Cultural Tolerance

Understanding diverse cultural and social backgrounds; working well with diverse groups of soldiers; demonstrating respect for varied cultural practices and beliefs.

Appendix B – Results from the First Retranslation Workshop

Summary Statement: Intended Category and Effectiveness Level	Percent of SMEs placing it in correct dimension	Mean Effectiveness Rating	Standard Deviation of Effectiveness Rating
A1	72.2	1.06	.24
A2	77.8	2.06	.54
A3	100.0	2.94	.24
B1	77.8	1.00	.00
B2	77.8	2.11	.47
B3	77.8	2.94	.24
C1	83.3	1.00	.00
C2	94.4	2.33	.49
C3	100.0	2.78	.55
D1	66.7	1.06	.24
D2	88.9	1.89	.47
D3	100.0	2.83	.38
E1	88.9	1.06	.24
E2	77.8	2.50	.51
E3	77.8	2.94	.24
F1	77.8	1.06	.24
F2	77.0	2.28	.46
F3	88.9	2.83	.51
G1	94.4	1.00	.00
G2	94.4	2.11	.58
G3	94.4	2.94	.24
H1	100.0	1.06	.24
H2	100.0	2.33	.49
H3	100.0	3.00	.00
I1	77.8	1.00	.00
I2	100.0	2.44	.51
I3	77.8	2.72	.57
J1	94.4	1.00	.00
J2	88.9	2.22	.55
J3	72.2	2.89	.32
K1	94.4	1.00	.00
K2	100.0	2.22	.43
K3	88.9	2.94	.24
L1	94.4	1.00	.00
L2	88.9	2.44	.51
L3	94.4	3.00	.00

Note: A = Technical Knowledge and Skills; B = Training; C = Coaching, Mentoring, and Supporting Trainees; D = Counseling; E = Effort and Initiative; F = Integrity and Adherence to Army Core Values; G = Following Rules, Regulations, and Safety Guidelines; H = Physical Fitness and Military Bearing; I = Stress Tolerance and Conflict Resolution; J = Adaptability; K = Relating to and Supporting Peers; L = Cultural Tolerance

Appendix C – Results from the Second Retranslation Workshop

Summary Statement: Intended Category and Effectiveness Level	Percent of SMEs placing it in correct dimension	Mean Effectiveness Rating	Standard Deviation of Effectiveness Rating
A1	100.0	1.00	.00
A2	100.0	2.00	.00
A3	87.5	3.00	.00
B1	100.0	1.00	.00
B2	100.0	2.00	.00
B3	87.5	3.00	.00
C1	100.0	1.00	.00
C2	100.0	2.00	.00
C3	100.0	3.00	.00
D1	100.0	1.00	.00
D2	100.0	2.00	.00
D3	100.0	3.00	.00
E1	100.0	1.00	.00
E2	100.0	2.00	.00
E3	100.0	3.00	.00
F1	100.0	1.00	.00
F2	100.0	2.00	.00
F3	100.0	3.00	.00
G1	100.0	1.00	.00
G2	100.0	2.13	.35
G3	100.0	2.88	.35
H1	100.0	1.00	.00
H2	100.0	2.00	.00
H3	100.0	3.00	.00
I1	100.0	1.00	.00
I2	100.0	2.00	.00
I3	100.0	3.00	.00
J1	100.0	1.00	.00
J2	100.0	2.00	.00
J3	100.0	3.00	.00

Note: A = Technical Knowledge and Skills; B = Training; C = Coaching and Supporting Soldiers; D = Effort and Initiative; E = Following Rules, Regulations, and Adhering to Army Core Values; F = Physical Fitness and Military Bearing; G = Stress Tolerance and Conflict Resolution; H = Adaptability; I = Relating to and Supporting Peers; J = Cultural Tolerance

Appendix D – Final Rating Scales

Drill Sergeant Performance Rating Scales

A. Technical Knowledge & Skill

Demonstrating technical and tactical knowledge to soldiers; providing clear and technically accurate instructions; knowledgeably answering questions about training tasks; demonstrating training tasks properly.

Low	Mid-Range	High
<p>Is unable to display adequate knowledge of training tasks or Army life; is unable to provide clear, accurate instructions to soldiers; has trouble or may bluff when answering technical questions from soldiers; may make mistakes in instructions or demonstrations.</p> <p>① ②</p>	<p>Displays adequate knowledge of training tasks and Army life; is able to provide clear and accurate instructions, but may have trouble answering difficult or in-depth questions; instructions and demonstrations have infrequent, minor mistakes.</p> <p>③ ④ ⑤</p>	<p>Displays in-depth knowledge of training tasks and Army life; always provides clear and accurate instructions, and is able to answer even the most difficult questions; step-by-step instructions are accurate and demonstrations performed without error.</p> <p>⑥ ⑦</p>

B. Training

Using appropriate training methods; presenting well-prepared and organized material for training exercises and sessions; providing constructive feedback regarding performance on training tasks.

Low	Mid-Range	High
<p>Fails to consider individual training needs and uses same training techniques for all soldiers; is frequently unprepared for training exercises/sessions; fails to provide constructive feedback or hands-on corrections to soldiers; may resort to yelling and berating soldiers when their attention wanders or they have difficulties with training.</p> <p style="text-align: center;">① ②</p>	<p>Usually takes individual training needs into account and adapts presentation, content, or techniques; prepares as required for training exercises/sessions; identifies and corrects common trainee mistakes and deficiencies; seldom resorts to berating soldiers.</p> <p style="text-align: center;">③ ④ ⑤</p>	<p>Always assesses individual training needs and skillfully and creatively adapts presentation, content, or techniques; is always thoroughly prepared for training exercises/sessions; offers performance-enhancing tips and techniques to correct trainee mistakes and deficiencies; gives clear and authoritative corrections to soldiers, even modulating voice for maximum effect.</p> <p style="text-align: center;">⑥ ⑦</p>

Informally mentoring soldiers; formally counseling soldiers; demonstrating concern for the well-being and development of soldiers; demonstrating respect for soldiers; listening attentively to soldiers, and asking questions as appropriate; supporting soldiers and helping them to overcome personal problems.

Low	Mid-Range	High
<p>Fails to mentor soldiers who are having problems; spends little time preparing for or conducting counseling; fails to show concern for well-being or development of soldiers; may even show disrespect for soldiers through name-calling, threatening, humiliating, or abusive treatment; fails to pay attention to soldiers or asks questions that are confusing or miss the point; may even encourage soldiers who simply need extra help, to quit.</p>	<p>Attempts to mentor soldiers who are having problems; adequately prepares for counseling sessions; shows basic awareness of soldiers' well-being and concern for their development; almost always demonstrates respect for soldiers and uses positive motivational techniques; listens attentively to soldiers, seldom misunderstands what soldiers say, and asks relevant questions about their problems; attempts to help soldiers overcome or resolve personal problems.</p>	<p>Always mentors soldiers who are having problems; carefully and thoroughly prepares for counseling sessions; shows genuine concern for the well-being and development of soldiers; always demonstrates respect for soldiers, creatively uses a variety of positive motivational techniques; listens attentively, and asks questions, as necessary, to clarify issues; extends self to help soldiers overcome or resolve personal problems.</p>

⑦

D. Effort & Initiative

Persisting with extra effort; taking initiative to accomplish objectives; finding additional productive work when own duties are completed; developing one's own knowledge and skills through additional training.

Low	Mid-Range	High
<p>Often gives up or fails to complete duties when challenged by obstacles; seldom takes the initiative to address problems; avoids additional responsibilities where possible; attends training and development courses only when required and directed to do so.</p> <p>① ②</p>	<p>Persists in efforts to complete duties successfully; usually takes initiative to address problems and accomplish objectives; accepts some additional responsibilities; takes advantage of available opportunities to develop own knowledge and skills when opportunities present themselves.</p> <p>③ ④ ⑤</p>	<p>Persists with high levels of effort and determination to complete duties successfully even when conditions are very difficult; always takes initiative to do what is necessary to accomplish objectives; enthusiastically takes on new challenges and additional responsibilities; looks for and creates opportunities to develop own knowledge and skills.</p> <p>⑥ ⑦</p>

E. Following Rules, Regulations & Adhering to Army Core Values

Adhering to Army regulations, orders, and SOP; following safety guidelines and monitoring trainee safety; demonstrating integrity, and ethical behavior, and adhering to Army Core Values.

Low	Mid-Range	High
<p>At times ignores or fails to follow Army orders, regulations or safety guidelines; sometimes permits unsafe conditions or takes risks during training; does not carefully monitor trainee safety; may fail to take responsibility for his/her job-related errors; sometimes fails to maintain ethical standards; may fail to maintain proper, moral behavior.</p> <p>① ②</p>	<p>Almost always follows Army orders, regulations, and safety guidelines; enforces SOPs; usually avoids risks and is usually alert to safety hazards; monitors trainee behavior for safety compliance; takes responsibility for most job-related mistakes or poor decisions; maintains ethical standards; exhibits proper, moral behavior.</p> <p>③ ④ ⑤</p>	<p>Is very careful to follow the spirit and the letter of Army orders, regulations, and safety guidelines; is at all times alert to safety hazards and manages risk appropriately; monitors trainee behavior to ensure safety compliance, taking into consideration trainee fatigue, stress, and inexperience; assumes full responsibility for his/her mistakes; demonstrates the highest ethical standards; behaves morally, in a manner beyond reproach.</p> <p>⑥ ⑦</p>

F. Physical Fitness & Military Bearing

Maintaining physical fitness and proper military bearing; demonstrating good professional conduct.

Low	Mid-Range	High
<p>Is in somewhat poor physical condition and may just meet minimal Army standards for weight, strength, or physical stamina; avoids or does not participate fully in PT programs; dresses in improper or poorly maintained uniform; sometimes fails to display proper military bearing, customs, or courtesies.</p> <p>① ②</p>	<p>Is in good physical condition and easily meets Army standards for weight, strength, and physical stamina; attends and actively participates in PT programs; dresses neatly and properly according to Army standards; almost always demonstrates good military bearing and displays proper military customs and courtesies.</p> <p>③ ④ ⑤</p>	<p>Is in excellent physical condition and exceeds Army standards for physical fitness; enthusiastically participates in PT programs and exercises on own time to maintain peak physical fitness; is always neatly and properly dressed and consistently presents an impressive appearance; maintains excellent military bearing and displays proper military customs and courtesies in all situations.</p> <p>⑥ ⑦</p>

G. Stress Tolerance & Conflict Resolution

Managing stress and maintaining self-control; responding appropriately to provocation; asking for appropriate backup in potentially volatile situations; resolving conflicts quickly and fairly.

Low	Mid-Range	High
<p>Sometimes fails to maintain self-control and may display flashes of temper or anger; is easily provoked by soldiers; tends to escalate tension when confronted with potentially volatile situations; at times fails to ask for appropriate help or back-up when needed; ignores conflicts or fails to resolve them fairly.</p> <p>① ②</p>	<p>Almost always maintains self-control in stressful situations; usually responds calmly to provocation by soldiers; is usually good at defusing volatile situations; usually asks for appropriate help or back-up from others when needed; attempts to resolve conflicts fairly and in a timely manner.</p> <p>③ ④ ⑤</p>	<p>Consistently maintains self-control in stressful situations; responds calmly and authoritatively to provocation by soldiers; is highly skilled at defusing volatile situations; knows when to ask for help or back-up from fellow DSs or command; resolves conflicts quickly and fairly before they become problematic.</p> <p>⑥ ⑦</p>

H. Adaptability

Modifying behavior to adjust to unexpected events or conditions; effectively adapting to new and changing circumstances.

Low	Mid-Range	High
Is easily frustrated or has difficulty modifying behavior or approach in response to unexpected events or conditions; does not adjust quickly to changes in schedules, policies, responsibilities, or personnel.	Modifies behavior or approach to respond adequately to unexpected events or conditions; adjusts fairly quickly to changes in schedules, policies, responsibilities, or personnel.	Modifies behavior by developing innovative and imaginative responses to unexpected events or conditions; adjusts quickly and smoothly to changes in schedules, policies, responsibilities, or personnel.

①

②

③

④

⑤

⑥

⑦

I. Relating to & Supporting Peers

Effectively relating to, working, and communicating with other Drill Sergeants; helping others by performing some of their tasks when needed; supporting, encouraging, and showing confidence in other Drill Sergeants.

Low	Mid-Range	High
<p>Tends to be selfish or disrespectful to peers; at times, refuses to assist other DSs even when asked; downplays peers' achievements or fails to support or encourage other Drill Sergeants who are experiencing adversity or setbacks.</p> <p>① ②</p>	<p>Is usually tactful and respectful in dealing with peers; provides assistance to other DSs when asked; usually supports and gives encouragement to other DSs.</p> <p>③ ④ ⑤</p>	<p>Always treats peers in a respectful manner and with tact; willingly offers needed assistance to peers without being asked; goes out of his or her way to support and encourage other DSs.</p> <p>⑥ ⑦</p>

J. Cultural Tolerance

Understanding and demonstrating respect for diverse cultural practices and beliefs; encouraging others to display tolerance and sensitivity; training and communicating well with diverse groups of soldiers.

Low	Mid-Range	High
Does not understand or show respect for cultural practices or belief systems different from own; may make insensitive remarks or ignore slurs targeted toward others based on social, cultural, or gender differences; is unwilling or unable to effectively train or communicate with those from different backgrounds.	Recognizes the need to be tolerant and respectful of other cultural practices and belief systems, and usually demonstrates understanding of social and cultural diversity; disapproves of and corrects disrespect or insensitivity directed at others; is willing and for the most part able to train or communicate with those from different backgrounds.	Consistently shows tolerance, understanding, and respect for other cultural practices and belief systems; encourages others to demonstrate tolerance and sensitivity; easily trains and communicates effectively with those from different cultural and social backgrounds.
① ②	③ ④ ⑤	⑥ ⑦

K. Overall Effectiveness

Low	Mid-Range	High
Falls below drill sergeant standards.	Achieves and maintains the standards expected of a drill sergeant.	Exceeds standards and expectations for drill sergeants.
①	③	⑥
②	④	⑦

Appendix E - NLSI Part I Scales and Definitions

Tolerance for Ambiguity

This scale measures a person's preference for work environments in which the problems (and potential solutions) are unstructured and ill-defined. Those with high tolerance for ambiguity are comfortable working in rapidly changing work environments. Individuals scoring low prefer highly structured and predictable work settings.

Hostility to Authority

The degree to which a person respects and is willing to follow legitimate authority figures. High scorers are expressively angered by authority figures and may actively disregard their instructions and policies. Low scorers accept directives from superiors and easily adapt to structured work environments.

Social Perceptiveness

This scale measures the degree to which a person can discern and recognize others emotions and likely behaviors in interpersonal situations. Persons high in social insight are good at understanding others' motives and are less likely to be "caught off guard" by unexpected interpersonal behaviors.

Interpersonal Skill

This scale measures the degree to which a person establishes smooth and effective interpersonal relationships with others. Interpersonally skilled individuals are good listeners, behave diplomatically, and get along well with others. Persons with low scores on this measure have difficulty working with others and may intentionally or unconsciously promote interpersonal conflict and cause hurt feelings.

Emergent Leadership

The scale measures the degree to which a person takes on leadership roles in groups and in his or her interactions with others. High scorers on this scale are looked to for direction and guidance when group decisions are made and readily take on leadership roles.

Conscientiousness

This scale measures the degree to which a person is achievement-oriented and dedicated to work. Persons high in conscientiousness are hard working, persistent, self-disciplined, and deliberate. Individuals scoring low are more careless in work-related activities, prefer leisure activities to work, and can be easily distracted from work-related tasks.

Self-Esteem

This scale measures the degree to which a person feels good about oneself as a person and has confidence in one's own abilities. Individuals with high self-esteem feel successful in past undertakings and expect this to continue in the future. Low scorers have feelings of personal inadequacy, lower self-efficacy, and lack confidence in their ability to be successful.

Empathy

This scale measures the degree to which a person understands and shares others' thoughts and emotions. High scorers are sensitive, and find it difficult to watch the suffering of others.

Appendix F – NLSI Part II Scales and Definitions

Work Motivation

The tendency to strive for excellence in the completion of work-related tasks. Persons high on this construct seek challenging work activities and set high standards for themselves. They consistently work hard to meet these high standards.

Adjustment

The tendency to have a uniformly positive affect. Persons high on this construct maintain a positive outlook on life, are free of excessive fears and worries, and have a feeling of self-control. They maintain their positive affect and self-control even when faced with stressful circumstances.

Agreeableness

The tendency to interact with others in a pleasant manner. Persons high on this construct get along and work well with others. They show kindness, while avoiding arguments and negative emotional outbursts directed at others.

Dependability (Non-delinquency)

The tendency to respect and obey rules, regulations, and authority figures. Persons high on this construct are more likely to stay out of trouble in the workplace and avoid getting into difficulties with law enforcement officials.

Leadership (Dominance)

The tendency to seek out and enjoy being in leadership positions. Persons high on this scale are confident of their abilities and gravitate towards leadership roles in groups. They feel comfortable directing the activities of other people and are looked to for direction when group decisions have to be made.

Physical Conditioning

The tendency to seek out and participate in physically demanding activities. Persons high on this construct routinely participate in vigorous sports or exercise, and enjoy hard physical work.