Technical Report 1440

Reliability and Validity of Multisource Leadership Ratings at the First Unit of Assignment

Victor J. Ingurgio
U.S. Army Research Institute

Tatiana H. Toumbeva
Aptima, Inc.

Frederick J. Diedrich
Sophia Speira

Krista L. Ratwani
Aptima, Inc.

Scott Flanagan
Sophia Speira

January 2024

United States Army Research Institute for the Behavioral and Social Sciences

Approved for public release; distribution is unlimited.

United States 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:

SCOTT SHADRICK, Ph.D.
Acting Director

Research accomplished under contract
for the Department of the Army by

Aptima, Inc.

Technical Review by

Dr. James Lee, U.S. Army Research Institute
Dr. Kyle Harwell, U.S. Army Research Institute

DISPOSITION

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

REPORT DOCUMENTATION PAGE

1. REPORT DATE (Month Year)	2. REPORT TYPE	3. DATES COVERED (Month Year)	
January 2024	Final	START DATE July 2019	END DATE January 2023

4. TITLE AND SUBTITLE
Reliability and Validity of Multisource Leadership Ratings at the First Unit of Assignment

5a. CONTRACT NUMBER	5b. GRANT NUMBER	5c. COOPERATIVE AGREEMENT NUMBER
W911NF-19-F-0032		

5d. PROGRAM ELEMENT NUMBER	5e. PROJECT NUMBER	5f. TASK NUMBER	5g. WORK UNIT NUMBER
633007	A792		

6. AUTHOR(S)
Victor J. Ingurgio; Tatiana H. Toumbeva; Frederick J. Diedrich; Krista L. Ratwani; Scott Flanagan

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Research Institute
for the Behavioral & Social Sciences
6000 6th Street (Building 1464/Mail Stop 5610)
Fort Belvoir, VA 22060-5610

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)	11. SPONSOR/MONITOR'S REPORT NUMBER(S)
U.S. Army Research Institute for the Behavioral and Social Sciences 6000 6th Street (Bldg. 1464 / Mail Stop: 5610) Fort Belvoir, Virginia 22060-5610	ARI	Technical Report 1440

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution Statement A: Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
ARI Research POC: Dr. Victor J. Ingurgio, Fort Moore Research Unit

14. ABSTRACT: Given the emphasis on talent management within the Army, leaders need tools to help the reliably and consistently capture and assess information related to Soldier performance and development. One such tool has been developed by the U.S. Army Research Institute—the Leader Attribute Rating Application (LARA; Ingurgio et al., 2020). LARA contains rubrics that help leaders more consistently assess their Soldiers on the leadership attributes and competencies found in the Army's Leader Requirements Model (LRM). Preliminary analyses from Ingurgio et al. (2020) show promising results regarding the usability and utility of the application, along with initial evidence regarding interrater reliability for assessments made using the application and associated rubrics. This report presents further reliability evidence by examining characteristics of ratings across peers, squad leaders, team leaders, and self-ratings. In addition, de-identified Army Physical Fitness Test (APFT) and Rifle Marksmanship (RM) scores for the rated Soldiers enabled a preliminary examination of validity evidence. The results demonstrated links between the LRM ratings and the more technical performance outcomes, reflecting the importance of considering both types of performance and highlighting the value of collecting multisource data when evaluating performance, especially on something like the LRM attributes and competencies.

15. SUBJECT TERMS: Leadership attributes and competencies; reliability and validity

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES
a. REPORT	b. ABSTRACT	C. THIS PAGE		
Unclassified	Unclassified	Unclassified	Unlimited Unclassified	24

19a. NAME OF RESPONSIBLE PERSON	19b. PHONE NUMBER (Include area code)
Dr. Brian Crabb	254-288-3833

Technical Report 1440

Reliability and Validity of Multisource Leadership Ratings at the First Unit of Assignment

Victor J. Ingurgio
U.S. Army Research Institute

Tatiana H. Toumbeva
Aptima, Inc.

Frederick J. Diedrich
Sophia Speira

Krista L. Ratwani
Aptima, Inc.

Scott Flanagan
Sophia Speira

Fort Moore Research Unit
Brian Crabb, Acting Chief

January 2024

Approved for public release; distribution is unlimited.

RELIABILITY AND VALIDITY OF MULTISOURCE LEADERSHIP RATINGS AT THE FIRST UNIT OF ASSIGNMENT

CONTENTS

LIST OF TABLES

Reliability and Validity of Multisource Leadership Ratings at the First Unit of Assignment

Introduction

The Army seeks to acquire, develop, employ, and retain the talent necessary to succeed in addressing ever-evolving threats in dynamic environments (U.S. Department of the Army, 2019a). In 1999, Karrasch & Halpin investigated the value of using a 360-degree evaluation (the Leader Azimuth Check) to improve leadership skills and development. Ten years later, Steele & Garvin (2009) using archival data reported on the validity of the Leader Azimuth Check found that officer ratings were valid. The Leader360 is an assessment of a leader's effectiveness on observable behaviors related to the Army Leadership Requirements Model (ALRM, ADP 6-22). As part of Project Athena, Soldiers completed a self-assessment and selected assessors from three perspective groups (superiors, peers, subordinates) to provide feedback on their leadership attributes (Willis-Grider, 2021). Assessment in support of development is a core facet of this talent management effort. Over their careers, Army leaders are expected to thrive in complexity–employing a variety of attributes and competencies outlined in the Leader Requirements Model (LRM) including Character, Presence, Intellect, Leads, Develops, and Achieves (U.S. Department of the Army, 2012, 2015, 2019b). Both Officer and Noncommissioned Officer Evaluation Reports (OERs and NCOERs, respectively) include assessments of the LRM attributes and competencies as part of the Evaluation Reporting System. Further, the Army is trying to improve how it develops those in the enlisted ranks, especially its noncommissioned officer corps, long considered the backbone of the U.S. military. Senior enlisted leaders want a better bird's-eye view of their young Soldiers and hope that formal evaluations could snowball into developing higher-quality NCOs who are better prepared to take on leadership responsibilities the moment they earn their stripes (Benyan, 2023). Similarly, within the institutional Army, assessment of selected elements of the LRM begins as early as Basic Combat Training and Officer Candidate School (e.g., Toumbeva et al., 2018, 2021). Despite the importance of LRM-based assessment, the Army faces challenges in obtaining reliable and valid data. These challenges emerge from both differences across settings and across raters. Settings can vary because ratings take place in a variety of environments (e.g., in garrison, in the field, in a classroom, on the range). Likewise, raters can vary in important ways (e.g., different raters use subjectively different rating approaches based on personal experiences and situations arise where multiple raters may be used to assess an individual Soldier). Further, in an institutional setting, peer evaluations are often used along with instructor-based assessments (e.g., Toumbeva et al., 2021, 2018). All these challenges can make the interpretation of assessment difficult and could result in Soldier feedback that may not be as comprehensive when derived from multiple raters. In order to realize effective talent management, the Army needs to have reliable and valid assessments of LRM attributes and competencies that can be standardized across settings and raters.

The U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) is developing tools designed to help leaders make assessments in a manner that is accurate, consistent, and comprehensive. The Leader Attribute Rating Application (LARA) provides a mobile application that can be used to assess and track a Soldier's leadership development (Ingurgio et al., 2020) over time. Developing LARA involved the creation of a doctrinally-

informed rubrics that cover each element of the LRM with descriptions of behaviors on a 5-point Likert scale representing levels of ability ranging from developmental needs to strengths (Dein et al., 2019). The LARA rubrics were vetted through the Center for Army Profession and Leadership (CAPL; currently the Center for Army Leadership—CAL), and preliminary analyses from Ingurgio et al. (2020) show promising results regarding usability and utility of the application, along with psychometric support (interrater reliability) for assessments made using the application and associated rubrics. In particular, results from use of the application in a Field Training Exercise (FTX) showed that 80% of paired ratings were within one point of each other on the 5-point rating scale. When used as part of a peer evaluation exercise when Soldiers arrived at their First Units of Assignment (FUAs) following Infantry basic and advanced individual training [referred to as One Station Unit Training (OSUT)], 76% of paired ratings between the Soldiers were within one point. Likewise, as an index of reliability, intraclass coefficients (ICCs) were at fair to good levels of reliability between raters for elements within the areas of Leads (ICC_{mean} = .67), Develops (ICC_{mean} = .57), Achieves (ICC_{mean} = .53), Presence (ICC_{mean} = .47), and Intellect (ICC_{mean} = .46). Reliability within the attribute of Character was less robust (ICC_{mean} = .33). Overall, across these two use cases (the FTX and the peer assessments), the results were favorable and demonstrate that raters generally agree on what constitutes good, average, and poor performance, with Character (ICC_{mean} = .33) emerging as the most difficult to rate consistently.

While these findings are promising, they are also limited with respect to key issues relevant to fully supporting LRM-based assessments for talent management. First, the ICC data came from peers rather than leaders. As mentioned, while peer assessment is common in the Army (e.g., Toumbeva et al., 2018, 2021), leader assessment of subordinates is paramount for ongoing counseling sessions, NCOERs and OERs in order to support individual development. Peer and leader assessments should augment and complement one another. Hence, additional data from leader assessments of subordinates are required. Second, the data from the initial study by Ingurgio et al. (2020) came from new Soldiers (i.e., Soldiers who recently graduated from OSUT) assessing other new Soldiers. While these data reinforce the usability of LARA, the participants limit the nature of the findings due to relative unfamiliarity with the LRM and how the attributes/competencies might manifest in a unit setting. Third, although the data speak to reliability, the question of validity remains unaddressed.

For these reasons, to supplement previous findings, analyses were conducted to determine the degree of reliability between Squad Leaders (SLs) and Team Leaders (TLs) rating Soldiers in their unit using LARA and its embedded rubrics. As in the case of the peer evaluation data, these data were collected as part of another program associated with an analysis of OSUT impacts with regard to extending the training time and its impact on readiness (Tucker et al., 2019). Here, de-identified data were used to examine interrater reliability across all attributes and competencies in the LRM. In addition, access to de-identified FUA Army Physical Fitness Test (APFT) and rifle marksmanship (RM) scores for the rated Soldiers enabled a preliminary examination of validity evidence. In this case, the focus is on the possible convergence of these operational performance measures (APFT and RM) with ratings of the LRM attributes and competencies.

Moreover, based on these data from the prior investigation, comparisons to APFT and RM scores were possible for the SL and TL ratings, along with peer and self-report ratings using the same scales embedded in LARA (see Ingurgio et al., 2020; Tucker et al., 2019). Although limited in the sense that RM and APFT reflect a narrow view of Soldier performance, these data nevertheless provide an initial look at how assessments within LARA correlate with other performance measures from the perspective of leaders, peers, and self-report. These results, and this line of research more broadly, provide ongoing evidence regarding how assessing the attributes and competencies contained within the Army's LRM are important components related to the Army's talent management initiatives. A method such as that explored here can reliably assess a pattern of strengths and weaknesses for an individual Soldier, ultimately providing insight into possible developmental interventions.

Method

Participants

Participants included SLs and TLs who rated Soldiers who had arrived at their FUAs within the previous six months post-OSUT. Peer and self-ratings were also obtained. As in the case of the previously analyzed peer evaluation data (Ingurgio et al., 2020; starting on page 5 are screenshots of the LARA workflow), these de-identified data were collected as part of another research program associated with an initial analysis of OSUT (Tucker et al., 2019). SLs rated between 64 to 71 Soldiers and TLs rated between 59 to 61 Soldiers on a total of 23 LRM attributes and competencies. Due to missing data, paired ratings from both an SL and TL were not available for every Soldier on every attribute, thus reducing the sample size available for certain analyses (see *df*2 in Table 1). Self-ratings were also obtained on each of the 23 LRM attributes and competencies (*N* ranged from 115 to 120). In addition, Soldiers were rated by one to five peers. The peer ratings were averaged for the analyses. Lastly, APFT and RM scores at the FUA were collected for 163 and 133 Soldiers, respectively. Only those participants who had scores inclusive of self- and peer-ratings on the LRM were included in these analyses. As a result, data from 108 Soldiers were used for the APFT analyses, and data from 90 Soldiers were used for the RM analyses.

Procedure

Participating SLs and TLs received a five- to 10-minute introduction to the LARA tool that included how to access the standardized definitions of each sub-attribute as well as the embedded ratings scales, and then used the tool to rate Soldiers in their squad or team on each attribute and competency within the LRM. Similarly, Soldiers rated their peers, as well as themselves, using the same application. All ratings were made with regard to the sub-attributes and later aggregated to the higher-level attributes. Due to the timing of the evaluation and version of the LARA application employed, the LRM scales employed did not have the attribute of Humility (under Character), that was added more recently (U.S. Department of the Army, 2019b). Hence, the attributes and competencies used for this research corresponds to the earlier version of the LRM (U.S. Department of the Army, 2015). Leaders and Soldiers were instructed to make ratings on all elements of the LRM, but some were missed resulting in unmatched pairs for the analyses. Although not a usability concern as most ratings were accomplished, possible

reasons for the small amount of missing data likely included a lack of familiarity with LARA, navigation errors, or confusion regarding what attributes had been already rated. Where possible and available, participating units provided RM and APFT scores for the rated Soldiers. Analyses of peer ratings were based on mean ratings from all peers for a given Soldier. Peers rated between one to five Soldiers. Complete details on the data collection procedures can be found in Tucker et al. (2019). Details on LARA can be found in Ingurgio et al. (2020) along with the relevant LRM scales (see also Dein et al., 2019).

Results

Squad Leader/Team Leader Interrater Reliability Analyses

Table 1 illustrates ICCs that were calculated to evaluate interrater reliability (McGraw & Wong, 1996; Shrout & Fleiss, 1979). A one-way, random effects model with multiple raters (average ratings) was used due to the variance in the number of Soldiers rated. This type of model is based on absolute agreement among raters and typically yields conservative coefficients compared to other ICC models. All ICCs were significant at the .05 alpha level.

Table 1

Interrater Reliability for Squad and Team Leader Ratings at the Attribute/Competency Level

Attribute/Competency	***ICC***	**95% Confidence Interval**		***F***	***df1, df2***	***Sig.***
		Lower Bound	*Upper Bound*			
Character	***.69***					
Army Values	.71	.43	.85	3.43	34, 35	<.001
Empathy	.55	.11	.78	2.23	33, 34	.01
Warrior Ethos	.72	.44	.86	3.57	33, 34	<.001
Discipline	.76	.53	.88	4.16	34, 35	<.001
Presence	***.79***					
Military and Professional Bearing	.71	.42	.85	3.39	34, 35	<.001
Fitness	.82	.65	.91	5.60	34, 35	<.001
Confidence	.79	.58	.89	4.68	35, 36	<.001
Resilience	.83	.67	.92	6.01	34, 35	<.001
Intellect	***.57***					
Mental Agility	.52	.04	.76	2.06	33, 34	.02
Sound Judgment	.66	.34	.83	2.97	34, 35	.001
Innovation	.47	-.07	.74	1.88	32, 33	.04
Interpersonal Tact	.47	-.04	.73	1.89	34, 35	.03
Expertise	.71	.43	.85	3.43	34, 35	<.001
Leads	***.66***					
Leads Others	.63	.27	.81	2.67	35, 36	.002
Builds Trust	.71	.43	.85	3.43	35, 36	<.001

Attribute/Competency	*ICC*	95% Confidence Interval		*F*	*df1, df2*	*Sig.*
		Lower Bound	*Upper Bound*			
Extends Influence Beyond the Chain of Command	.61	.23	.80	2.54	35, 36	.003
Leads by Example	.76	.53	.88	4.25	33, 34	<.001
Communicates	.57	.16	.78	2.31	35, 36	.007
Develops	***.74***					
Creates a Positive Environment	.75	.50	.87	3.93	34, 35	<.001
Prepares Self	.83	.67	.91	5.92	34, 35	<.001
Develops Others	.70	.40	.85	3.30	34, 35	<.001
Stewards the Profession	.66	.34	.83	2.95	35, 36	.001
Achieves	***.74***					
Gets Results	.74	.48	.87	3.78	34, 35	<.001

Note. Mean ICCs were calculated at the overall category level. Samples varied between 64 and 71 for the SL ratings and between 59 and 61 for the TL ratings.

Although guidelines vary, ICC values between .40 and .59 indicate fair reliability, values between .60 and .74 indicate good reliability, and values above .75 indicate excellent reliability (Cicchetti, 1994). Excellent interrater reliability (ICCs above .75) was found for the following attributes/competencies (ordered highest to lowest): Resilience, Prepares Self, Fitness, Confidence, Discipline, Leads by Example, and Creates a Positive Environment. A good degree of interrater reliability (between .61 and .74) was found for Gets Results, Warrior Ethos, Army Values, Military and Professional Bearing, Expertise, Builds Trust, Develops Others, Sound Judgment, Stewards the Profession, Leads Others, and Extends Influence Beyond the Chain of Command. Fair degrees of interrater reliability (between .40 and .59) was found for Communicates, Empathy, Mental Agility, Interpersonal Tact, and Innovation. When rolled up to the higher-order conceptual categories, Intellect had the lowest average (fair interrater reliability) ICC value (ICC_{mean} = .57); Leads (ICC_{mean} = .66), Character (ICC_{mean} = .69), Develops (ICC_{mean} = .74) and Achieves (ICC_{mean} = .74) were all at the good level of interrater reliability, and the highest level of interrater reliability was for Presence (ICC_{mean} = .79).

While all the ICC values in Table 1 were above .40 (a fair degree of interrater reliability for each item), the confidence intervals observed suggested high variability in some cases. For instance, the 95% confidence interval for Innovation ranged from .07 (poor reliability) to .74 (good reliability). In other cases, such as for Prepares Self, the 95% confidence interval for the ICC value was more favorable (.67 and .91). These wide confidence intervals may have been partly due to the relatively small sample sizes observed.

Self/Peer Interrater Reliability Analyses

Interrater reliability was also calculated for self and peer ratings using the same approach described above for the leader data. This analysis was performed to illustrate how the leaders and self/peer ratings differ with regard to how they perceive which attributes they feel are most

important to Soldiers. As shown in Table 2, the only significant ICCs (ordered from highest to lowest) were for Fitness, Discipline, Expertise, Communicates, Army Values, Confidence, Warrior Ethos, Prepares Self, and Leads by Example ($p < .05$).

Table 2

Interrater Reliability for Self and Peer Ratings at the Attribute/Competency Level

Attribute/Competency	***ICC***	**95% Confidence Interval**		***F***	***df1, df2***	***Sig.***
		Lower Bound	*Upper Bound*			
Character	***.39***					
Army Values	.39	.07	.60	1.63	89, 90	.01
Empathy	.29	-.08	.53	1.40	90, 91	.06
Warrior Ethos	.33	-.02	.56	1.49	89, 90	.03
Discipline	.54	.30	.69	2.15	90, 91	<.001
Presence	***.38***					
Military and Professional Bearing	.24	-.15	.50	1.31	90, 91	.10
Fitness	.71	.56	.71	3.46	89, 90	<.001
Confidence	.38	.07	.59	1.62	91, 92	.01
Resilience	.20	-.23	.48	1.25	86, 87	.16
Intellect	***.20***					
Mental Agility	.28	-.10	.52	1.38	89, 30	.06
Sound Judgment	.07	-.42	.39	1.07	87, 88	.38
Innovation	.01	-.54	.33	0.99	88, 89	.53
Interpersonal Tact	.23	-.18	.50	1.30	86, 87	.11
Expertise	.41	.10	.61	1.68	88, 89	.007
Leads	***.19***					
Leads Others	.25	-.14	.50	1.33	90, 91	.09
Builds Trust	.14	-.31	.43	1.16	91, 92	.24
Extends Influence Beyond the Chain of Command	.17	-.76	.23	0.86	91, 92	.77
Leads by Example	.31	-.05	.54	1.44	89, 90	.04
Communicates	.40	.09	.60	1.66	90, 91	.008
Develops	***.17***					
Creates a Positive Environment	.16	-.26	.44	1.20	93, 94	.19
Prepares Self	.32	-.03	.55	1.47	89, 90	.04
Develops Others	.10	-.37	.40	1.11	91, 92	.32
Stewards the Profession	.11	-.35	.42	1.13	88, 89	.29
Achieves	***.13***					
Gets Results	.13	-.70	.25	0.89	92, 93	.72

Note. Mean ICCs were calculated at the overall category level. Samples varied between 115 and 120 for the Self ratings and between 131 and 135 for the Peer ratings.

Following the guidelines by Cicchetti (1994), good interrater reliability (ICCs between .60 and .74) was found for Fitness and fair interrater reliability (ICCs between .40 and .59) was found for Discipline, Expertise, and Communicates. The ICCs for the remaining competencies were poor, with the worst interrater reliability being for Extends Influence Beyond the Chain of Command, followed by Gets Results and Innovation. These results should be interpreted with caution due to the wide range of the confidence intervals across most competencies.

When rolled up to the higher-order conceptual categories, the results indicate poor interrater reliability. Achieves (i.e., Gets Results) had the lowest average interrater reliability (ICC_{mean} = .17), followed by a group of similar ICCs for Develops (ICC_{mean} = .17), Leads (ICC_{mean} = .19), and Intellect (ICC_{mean} = .20), and the highest being for Presence (ICC_{mean} = .38) and Character (ICC_{mean} = .39). It is interesting to note that the leaders value for Character (ICC_{mean} = .69) was much higher than that for the peer/self value. This may be simply due to a practice effect for the leaders as they have a better understanding of Character over time.

Perhaps not surprisingly, the peer/self ICCs (in Table 2) are lower than those for the TL/SL ICCs (in Table 1); however, there are some commonalities. For instance, groups tend to align well on competencies like Fitness and Discipline which may be easier to observe more consistently, whereas other competencies like Empathy may be more idiosyncratic and yield greater variations in ratings among individuals. Furthermore, the results may be impacted by differential opportunities for competencies to manifest within the parameters of the operating context. For instance, Soldiers may not have as much of an opportunity, or a need, to regularly display Extends Beyond the Chain of Command, Innovation, and Stewards the Profession, which also happen to have among the lowest ICCs.

Relationships between LRM Dimensions Across Raters

Before delving into the relationships between LRM ratings and outcomes (APFT, RM scores), it was of interest to explore the relationships between the ratings themselves and across raters. The goal was to gain a better understanding of potential conceptual overlaps and areas that were more distinct or unique across the 23 rated attributes and competencies. In addition, the correlations shed light on the utility of having raters evaluate Soldiers on all individual LRM elements (i.e., attributes and competencies).

For the ratings made by SLs (see Table 1 in Appendix A), the magnitude of the correlations ranged between r =.18 and r = .79. The stronger relationships were between Army Values, Disciplined, and Empathy which is not surprising considering these all fall under the broader LRM category of Character. Other notably strong relationships were between Leads Others and Extends Influence (r = .76) and between Gets Results and Stewards the Profession (r = .76). Warrior Ethos was strongly correlated to a number of the LRM elements, especially Innovation and Leads by Example (r = .75). The weakest relationships, some of which were not significant, involved Fitness. For instance, there were no significant correlations between Fitness and the Army Values, as well as Fitness and Empathy. Mental Agility was another competency that generally had lower relationships with the other variables, with the exception of Innovation and Sound Judgment which are all within the Intellect LRM subcategory.

In terms of the ratings made by TLs (see Table 2 in Appendix A), the magnitude of the correlations ranged from $r = .32$ to $r = .80$. Similar to the SL ratings, among the strongest relationships were between attributes and competencies within the Character category (e.g., Army Values and Disciplined; $r = .80$) whereas Fitness was the least strongly related to other elements of the LRM. Other notably strong relationships were between Builds Trust and the Army Values ($r = .71$), Builds Trust and Military and Professional Bearing ($r = .73$), Leads Others and Empathy ($r = .73$), Leads Others and Develops Others ($r = .73$), Prepares Self and Leads by Example ($r = .72$), and Resilience and Creates a Positive Environment ($r = .73$).

Overall, the correlations for ratings by peers were higher than those made by other groups (see Table 3 in Appendix A), with magnitudes ranging from $r = .34$ to $r = .85$. Fitness remained the least correlated with other LRM elements as observed with the SL and TL rater groups. Builds Trust was very strongly related to many attributes and competencies, and especially with Communicates, Creates a Positive Environment, Extends Influence, Leads by Example, and Leads Others (at or above $r = .80$). Such high intercorrelations point to a possible redundancy of Builds Trust ratings and the existence of certain biases such as the halo effect (Thorndike, 1920).

Interestingly, the correlations between attribute and competency self-ratings were generally the weakest across rater groups (ranging from $r = .16$ to $r = .72$). Aside from the relationships between Character elements, no relationships exceeded a correlation of $r = .70$ (see Table 4 in Appendix A). Table 5 in Appendix A shows the descriptive statistics and correlations across all rater groups combined indicating moderate to strong relationships across the board. In fact, Creates a Positive Environment and the Army Values had a rarely observed perfect relationship ($r = 1.00$).

Overall, these findings support previous insights that the manner in which individuals observe and process behavior when evaluating performance is a function, at least in part, of their role and their relationship to the ratee.

Relationships between LRM Ratings using LARA and Outcomes

A series of bivariate correlations and multiple regression analyses were conducted to determine if LRM ratings (SL, TL, and peers of Soldiers in their unit, including Soldiers' self-report ratings on the LRM), were significantly related to FUA APFT and RM scores.

APFT Scores

Results indicated that APFT scores were significantly and differentially related to SL, TL, peer, and self-report ratings on a subset of the LRM attributes (see Table 3). At the higher category level, APFT was significantly related to Presence, Leads, Develops, and Achieves (i.e., Gets Results). APFT was most strongly related to Presence across all ratings groups, in particular with Fitness. Specifically, APFT had the strongest correlations with Fitness ratings by SLs ($r = .70$, $p < .001$), followed by TLs ($r = .60$, $p < .001$), self ($r = .53$, $p < .001$), and peers ($r = .40$, $p < .001$). Although TLs and peer ratings for Presence, and self-ratings for Presence and Achieves were related to APFT, there were significant relationships at the individual attribute level within

other categories as well (e.g., TL ratings on Warrior Ethos, peer ratings on Prepares Self, and self-report ratings on Leads Others).

Multiple regression analyses indicated that, across all raters, APFT scores were the most useful for predicting Fitness. For peer ratings, APFT scores explained a significant amount of variance in Resilience (b = 17.24, p = .03), in addition to Fitness (b = 10.48, p = .03) ratings. Overall, SL ratings on the LRM explained 48% of the variance in APFT scores (Adj. R^2 = .48, p < .001) compared to 42%, 30%, and 19% by TL, self, and peer, respectively.

Table 3

Correlations between LRM Ratings across Raters and APFT Scores

Attribute/Competency	SL Ratings	TL Ratings	Peer Ratings	Self-Ratings
Character	***0.14***	***0.14***	***0.08***	***0.05***
Army Values	0.03	0.05	0.04	-0.03
Empathy	0.04	0.14	-0.02	-0.01
Warrior Ethos	0.27	0.28*	0.09	0.16
Discipline	0.15	0.06	0.15	0.07
Presence	***0.51******	***0.43****	***0.28****	***0.29****
Military Bearing	0.29*	0.22	0.10	0.07
Fitness	0.70***	0.60***	0.40***	0.53***
Confidence	0.41**	0.32*	0.11	0.23*
Resilience	0.42**	0.30*	0.28**	0.09
Intellect	***0.19***	***0.22***	***0.15***	***0.13***
Mental Agility	0.07	0.04	0.22*	0.09
Sound Judgment	0.23	0.28*	0.09	0.14
Innovation	0.18	0.16	0.00	0.06
Interpersonal Tact	0.22	0.22	0.07	0.12
Expertise	0.28*	0.22	0.18	0.14
Leads	***0.37****	***0.14***	***0.10***	***0.18***
Leads Others	0.37**	0.14	0.14	0.21*
Builds Trust	0.17	0.12	0.08	0.10
Extends Influence	0.34**	0.03	0.12	0.06
Leads by Example	0.41**	0.15	0.09	0.17
Communicates	0.33*	0.12	0.03	0.18
Develops	***0.33****	***0.27***	***0.13***	***0.18***
Creates a Positive Environment	0.14	0.22	0.13	0.21*
Prepares Self	0.39**	0.22	0.24*	0.18
Develops Others	0.34*	0.14	0.04	-0.01
Stewards the Profession	0.29*	0.29*	0.11	0.18
Achieves - Gets Results	***0.31****	***0.16***	***0.17***	***0.28****

Notes. Pearson correlation coefficients (*r*) are reported, *N*s ranged from 51 to 108, and * $p < .05$; ** $p < .01$; *** $p < .001$.

RM Scores

As shown in Table 4, RM scores were significantly related to a subset of the peer LRM ratings but not for the SL/TL nor the self-ratings.

Table 4

Correlations between LRM Ratings across Raters and RM Scores

Attribute/Competency	SL Ratings	TL Ratings	Peer Ratings	Self-Ratings
Character	***-0.13***	***-0.08***	***0.32****	***0.03***
Army Values	-0.06	-0.02	0.24*	0.05
Empathy	0.05	-0.01	0.29**	-0.04
Warrior Ethos	-0.17	-0.12	0.33**	0.07
Discipline	-0.20	-0.13	0.27*	0.05
Presence	***-0.07***	***-0.17***	***0.21***	***0.10***
Military Bearing	0.00	-0.20	0.17	0.09
Fitness	0.06	-0.02	0.22*	0.06
Confidence	-0.12	-0.17	0.06	0.08
Resilience	-0.13	-0.17	0.23*	0.08
Intellect	***-0.06***	***-0.10***	***0.23****	***0.15***
Mental Agility	-0.07	-0.11	0.15	0.13
Sound Judgment	-0.02	-0.02	0.23*	0.20
Innovation	-0.02	-0.17	0.20	0.03
Interpersonal Tact	0.01	0.04	0.21	0.15
Expertise	0.05	-0.18	0.18	0.11
Leads	***-0.12***	***-0.15***	***0.31****	***0.08***
Leads Others	-0.08	-0.25	0.24*	0.09
Builds Trust	-0.10	0.02	0.27*	0.04
Extends Influence	-0.20	-0.20	0.35***	0.04
Leads by Example	0.07	-0.11	0.28**	0.13
Communicates	-0.19	-0.10	0.29**	0.14
Develops	***-0.10***	***-0.13***	***0.35****	***0.08***
Creates a Positive Environment	-0.07	-0.06	0.31**	0.08
Prepares Self	-0.13	-0.05	0.31**	0.10
Develops Others	-0.03	-0.31*	0.33**	0.06
Stewards the Profession	-0.08	-0.04	0.34**	0.04
Achieves **- Gets Results**	***-0.13***	***-0.19***	***0.27****	***0.11***

Notes. Pearson correlation coefficients (r) are reported, *N*s ranged from 51 to 108, and * $p < .05$; ** $p < .01$; *** $p < .001$.

Peer ratings on 17 out of the 23 attributes were significantly related to RM scores. At higher levels, RM scores were significantly related to all categories except Presence, with the strongest correlation being with Develops ($r = .35$, $p < .001$). At lower levels, RM was most strongly correlated with peer ratings on Extends Influence ($r = .35$, $p < .001$), Stewards the Profession ($r = .34$, $p < .01$), Warrior Ethos ($r = .33$, $p < .01$), and Develops Others ($r = .33$, $p < .01$). Despite these results, multiple regression analyses showed that LRM ratings did not account for any significant variance in RM scores.

Discussion

The data described here are useful for several reasons and have implications for talent management initiatives for the Army. At the micro level, these data coupled with those described in Ingurgio et al., (2020), provide promising evidence of interrater reliability and initial validity of the LRM measures in LARA further demonstrating the usefulness of the tool. At the macro level, these findings are important for two reasons. First, from a reliability perspective, the act of assessing the attributes and competencies described in the LRM is sometimes criticized as being too subjective in nature. One individual's interpretation of Resilience may not be the same as another's. Therefore, the fact that fairly consistent ratings were made across leader raters using the rubrics found in LARA demonstrates that employing LARA may represent a useful step in ensuring consistency across the way these attributes and competencies are assessed. Second, the self and peer reliability results highlight the need to gather multi-rater feedback because of different experiences that individuals may have with one another, and the importance of leaders to help Soldiers develop self-awareness around critical attributes and competencies such as through the use of LARA. Third, from a validity standpoint, these data demonstrate the relationship between the attributes / competencies in the LRM and other performance outcomes. Although the outcome measures used in this analysis were constrained by available data related to RM and performance on the APFT, the observed relationships between the more subjective LRM ratings and the more objective measures of RM/APFT performance is important. Although the Army emphasizes the importance of the LRM attributes and leader development in doctrine, there is still an emphasis on assessing Soldier performance on more tactical and technical tasks like RM. These data demonstrate relationships between these two types of performance, reflecting the importance of considering a range of assessments and ultimately, how the LRM elements may relate to tactical and technical tasks.

In terms of the specific relationships demonstrated between attribute ratings and performance outcomes, differential patterns emerged. First, the relationship between Fitness ratings and APFT scores was the most salient relationship. Even though the attribute of Fitness itself contains other elements in addition to physical fitness (e.g., emotional well-being), APFT performance is likely the most salient, tangible manifestation of that attribute. More globally, ratings across the LRM attributes and competencies were related more strongly to APFT scores than to RM performance. This pattern was not observed for the peer ratings, where more of a relationship to RM could be seen. One explanation for this finding is that Soldiers do physical training (PT) nearly every day—performance during PT is visible to nearly everyone around them, including leaders. While peers are a part of these PT exercises, it may be that performance during PT may become less salient over time. The frequency of this activity contrasts to RM that

is an event that may only happen once or twice over a program of instruction or training cycle. However, during basic training, RM is a milestone event that is a graduation requirement and that captures the attention of trainees as well as their Drill Sergeants. Therefore, how a Soldier performs during RM may be more salient for peers who have only recently exited OSUT. Soldier performance during RM may be more likely to stand out to peers (as opposed to leaders) and their ratings on the LRM attributes and competencies may subsequently account for what they observed during that critical event. In contrast, while leaders certainly care about RM, it is likely that their attention may be drawn to a more general set of Soldier skills beyond RM alone given the broader experience of the leaders.

More globally, the patterns of relationships between attribute ratings and APFT and RM scores indicate that different groups of raters may be paying attention to different things. The systematic differences among rater in terms of attribute/competency intercorrelations provide additional support that different groups of individuals view performance through a slightly different lens. What may be considered when peers evaluate Leads Others differs from what a SL or TL may consider for that same attribute. These findings highlight the value of collecting multisource data when evaluating performance, especially on something like the LRM attributes and competencies. SLs may draw on general day-to-day impressions of how "squared away" a Soldier is, which may influence other ratings (e.g., halo effect; Thorndike, 1920). TLs may have a more nuanced perspective based on intimate and more frequent interactions with Soldiers. Finally, peers likely have the most granular view of Soldier performance. These different views on Soldier performance are essential for fully capturing a complete picture of how any one Soldier is developing. To do so effectively, all individuals need a consistent and standardized method for capturing performance. The rubrics contained within LARA represent one method for doing so with a more valid and reliable approach.

References

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*(4), 284–290. https://doi.org/10.1037/1040-3590.6.4.284

Dein, J. P., Ingurgio, V., Ratwani, K. L., Diedrich, F., & Flanagan, S. (2019, July). Tools and measures for NCO talent assessment. *NCO Journal,* 1-15.

Ingurgio, V. J., Ratwani, K. L., Nargi, B., Flanagan, S. M., Diedrich, F. J., & Toumbeva, T. H. (2020). *Tools for assessing and tracking leadership attributes and competencies* (Research Report 2030). U.S. Army Research Institute for the Behavioral and Social Sciences. (DTIC No. AD1112648).

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1*(1), 30–46. https://doi.org/10.1037/1082-989X.1.1.30

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86*(2), 420–428. https://doi.org/10.1037/0033-2909.86.2.420

Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology, 4*(1), 25–29. https://doi.org/10.1037/h0071663

Toumbeva, T. H., Uhl, E. R., Wittig, A., Diedrich, F. J., Flanagan, S. M., Koschny, R. L., & Sanders, C. (2021). *Development and evaluation of a revised peer assessment in the U.S. Army Officer Candidate School* (Technical Report 1390). U.S. Army Research Institute for the Behavioral and Social Sciences. (DTIC No. AD1126519).

Toumbeva, T. H., Ratwani, K. L., Diedrich, F. J., Flanagan, S. M., & Uhl, E. R. (2018). *Development of a behaviorally anchored rating scale for leadership* (Research Product 2018-06). U.S. Army Research Institute for the Behavioral and Social Sciences. (DTIC No. AD1048729).

Tucker, Jennifer S., Uhl , Elizabeth R., Brimstin, Jay A., O'Brien, Rory P., Pedersen, Joseph M., Allen, Jayne L., Kochert, Jonathan F., Pitts, Kenneth P., Nunn, Darin M., Pitts, Kristin N., Miller, John T. II, Mezzaline, Chip E., Mock, John, Rodgers, Carrie E., Hester, Annie, Stuhlman, Melissa J., Grove, Julia W., Ledford, Cameron B. (2019). *Infantry One Station Unit Training Transformation: Phase I Findings* (Research Report 2025). U.S. Army Research Institute for the Behavioral and Social Sciences. (DTIC No. AD1080564).

U.S. Department of the Army, Headquarters (2019a). *The Army People Strategy.* (Strategic Document 05, Strategy Note 2019-01). Author.

U.S. Department of the Army, Headquarters (2019b). *Army Leadership and the Profession* (Army Doctrine Publication 6-22). Author.

U.S. Department of the Army, Headquarters (2015). *Leader development* (Field Manual 6-22). Author.

U.S. Department of the Army, Headquarters (2012). *Army leadership* (Army Doctrine Publication 6-22). Author.

Appendix A

Descriptive Statistics and Correlations for Attribute and Competency Ratings Across Different Raters

Table 1

Descriptive Statistics and Correlations for Attribute and Competency Ratings by Squad Leaders

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1. Army Values	3.69	1.16																						
2. Builds Trust	3.21	1.17	.69***																					
3. Communicates	3.29	1.13	.59***	.63***																				
4. Confidence	3.17	1.33	.50***	.61***	.68***																			
5. Creates a Pos. Env.	3.43	1.15	.63***	.68***	.59***	.56***																		
6. Develops Others	2.86	1.08	.49***	.65***	.58***	.65***	.66***																	
7. Disciplined	3.26	1.30	.79***	.67***	.67***	.58***	.69***	.63***																
8. Empathy	3.49	.95	.78***	.63***	.56***	.41***	.68***	.41***	.69***															
9. Expertise	2.72	1.13	.58***	.66***	.62***	.55***	.73***	.73***	.72***	.60***														
10. Extends Influence	2.93	1.08	.43***	.70***	.63***	.67***	.56***	.72***	.54***	.43***	.70***													
11. Fitness	3.27	1.42	.19	.43***	.46***	.66***	.40***	.50***	.41***	.18	.50***	.57***												
12. Gets Results	3.23	1.23	.63***	.71***	.63***	.70***	.63***	.64***	.71***	.52***	.68***	.66***	.60***											
13. Innovation	2.77	1.00	.58***	.64***	.58***	.49***	.59***	.62***	.66***	.60***	.72***	.69***	.50***	.69***										
14. Interpersonal Tact	3.34	1.01	.43***	.53***	.52***	.58***	.62***	.61***	.52***	.49***	.64***	.68***	.44***	.53***	.61***									
15. Leads by Example	3.06	1.17	.61***	.66***	.66***	.62***	.66***	.57***	.69***	.59***	.65***	.61***	.50***	.66***	.57***	.44***								
16. Leads by Others	2.79	1.23	.43***	.67***	.63***	.68***	.57***	.75***	.58***	.40***	.69***	.76***	.59***	.67***	.69***	.60***	.59***							
17. Mental Agility	3.09	.98	.54***	.59***	.54***	.52***	.45***	.40***	.57***	.47***	.54***	.53***	.40***	.65***	.72***	.51***	.36**	.48***						
18. Mil. & Prof. Bearing	3.20	1.26	.62***	.66***	.65***	.70***	.61***	.65***	.69***	.57***	.69***	.65***	.58***	.75***	.65***	.58***	.70***	.63***	.53***					
19. Prepares Self	3.20	1.29	.55***	.66***	.72***	.67***	.57***	.66***	.69***	.48***	.64***	.66***	.54***	.72***	.63***	.52***	.69***	.69***	.56***	.62***				
20. Resilience	3.37	1.19	.48***	.66***	.68***	.70***	.57***	.63***	.66***	.48***	.58***	.66***	.66***	.62***	.52***	.43***	.69***	.64***	.44***	.62***	.70***			
21. Sound Judgment	3.12	.98	.72***	.62***	.71***	.61***	.66***	.51***	.68***	.71***	.66***	.56***	.43***	.65***	.64***	.58***	.57***	.57***	.60***	.69***	.63***	.48***		
22. Stewards the Prof.	2.81	1.22	.61***	.67***	.59***	.70***	.62***	.75***	.63***	.51***	.67***	.69***	.58***	.76***	.68***	.57***	.67***	.70***	.50***	.70***	.66***	.68***	.61***	
23. Warrior Ethos	3.47	1.28	.61***	.66***	.66***	.71***	.68***	.67***	.71***	.62***	.69***	.72***	.58***	.69***	.75***	.61***	.75***	.69***	.50***	.73***	.65***	.72***	.63***	.74***

*Note. N*s range from 64 to 71. $^{+}$ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$.

Table 2

Descriptive Statistics and Correlations for Attribute and Competency Ratings by Team Leaders

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1. Army Values	3.47	1.19																						
2. Builds Trust	3.16	1.20	.71***																					
3. Communicates	3.13	1.10	.57***	.53***																				
4. Confidence	3.15	1.26	.40***	.37**	.58***																			
5. Creates a Pos. Env.	3.49	1.07	.61***	.61***	.59***	.65***																		
6. Develops Others	2.70	.92	.53***	.48***	.57***	.53***	.52***																	
7. Disciplined	3.27	1.38	.80***	.67***	.64***	.50***	.57***	.57***																
8. Empathy	3.46	1.03	.71***	.71***	.52***	.45***	.68***	.54***	.65***															
9. Expertise	2.98	1.13	.53***	.62***	.52***	.50***	.60***	.59***	.55***	.54***														
10. Extends Influence	2.72	1.04	.49***	.63***	.49***	.58***	.53***	.60***	.54***	.48***	.61***													
11. Fitness	3.15	1.29	.38***	.45***	.42***	.67***	.63***	.47***	.48***	.48***	.56***	.43**												
12. Gets Results	3.20	1.17	.61***	.57***	.73***	.65***	.61***	.51***	.65***	.54***	.65***	.54***	.51***											
13. Innovation	2.88	.89	.57***	.57***	.68***	.50***	.51***	.65***	.51***	.54***	.60***	.54***	.32*	.58***										
14. Interpersonal Tact	3.18	1.07	.64***	.57***	.54***	.57***	.55***	.39**	.59***	.55***	.54***	.49***	.46***	.71***	.58***									
15. Leads by Example	2.95	1.31	.62***	.67***	.59***	.44***	.49***	.65***	.63***	.50***	.56***	.65***	.55***	.57***	.48***	.47***								
16. Leads by Others	2.67	1.15	.54***	.61***	.63***	.57***	.66***	.73***	.62***	.53***	.60***	.69***	.48***	.67***	.54***	.44***	.68***							
17. Mental Agility	3.08	1.04	.52***	.58***	.60***	.64***	.50***	.57***	.49***	.58***	.54***	.66***	.32*	.61***	.63***	.65***	.51***	.61***						
18. Mil. & Prof. Bearing	3.23	1.31	.68***	.73***	.57***	.46***	.59***	.56***	.74***	.67***	.60***	.63***	.47***	.64***	.57***	.64***	.67***	.60***	.60***					
19. Prepares Self	3.02	1.17	.55***	.56***	.56***	.42***	.48***	.52***	.55***	.40***	.56***	.44***	.44***	.58***	.45***	.46***	.72***	.55***	.41**	.51***				
20. Resilience	3.31	1.23	.47***	.52***	.58***	.76***	.73***	.48***	.59***	.49***	.52***	.55***	.65***	.64***	.43***	.47***	.55***	.67***	.62***	.60***	.51***			
21. Sound Judgment	3.03	1.05	.56***	.62***	.56***	.54***	.52***	.48***	.64***	.59***	.72***	.45***	.58***	.69***	.60***	.68***	.52***	.53***	.53***	.66***	.61***	.51***		
22. Stewards the Prof.	2.79	1.04	.50***	.50***	.68***	.54***	.62***	.70***	.50***	.58***	.64***	.49***	.45***	.59***	.66***	.43***	.55***	.58***	.51***	.53***	.56***	.52***	.48***	
23. Warrior Ethos	3.43	1.21	.77***	.60***	.59***	.64***	.66***	.58***	.83***	.70***	.59***	.54***	.61***	.60***	.56***	.54***	.64***	.57***	.48***	.67***	.61***	.69***	.60***	.62***

*Note. N*s range from 59 to 61. + $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$.

Table 3

Descriptive Statistics and Correlations for Attribute and Competency Ratings by Peers

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1. Army Values	3.76	.75																						
2. Builds Trust	3.68	.93	.68***																					
3. Communicates	3.61	.88	.69***	.81***																				
4. Confidence	3.85	.88	.58***	.60***	.61***																			
5. Creates a Pos. Env.	3.65	.89	.69***	.82***	.72***	.60***																		
6. Develops Others	3.49	.84	.71***	.78***	.75***	.60***	.80***																	
7. Disciplined	3.72	.91	.80***	.72***	.67***	.58***	.63***	.72***																
8. Empathy	3.61	.86	.77***	.69***	.65***	.55***	.74***	.72***	.69***															
9. Expertise	3.68	.85	.67***	.57***	.60***	.56***	.54***	.63***	.66***	.62***														
10. Extends Influence	3.53	.81	.71***	.85***	.78***	.62***	.76***	.79***	.77***	.67***	.62***													
11. Fitness	3.79	.97	.46***	.38***	.40***	.50***	.42***	.45***	.45***	.36***	.34***	.43***												
12. Gets Results	3.80	.92	.62***	.70***	.68***	.66***	.64***	.67***	.64***	.64***	.68***	.66***	.42***											
13. Innovation	3.54	.82	.65***	.69***	.70***	.65***	.63***	.71***	.67***	.68***	.65***	.75***	.41***	.68***										
14. Interpersonal Tact	3.67	.94	.60***	.61***	.67***	.55***	.57***	.58***	.59***	.60***	.58***	.64***	.45***	.64***	.70***									
15. Leads by Example	3.41	.91	.66***	.80***	.82***	.58***	.71***	.82***	.70***	.69***	.65***	.81***	.45***	.71***	.71***	.69***								
16. Leads by Others	3.47	.93	.73***	.83***	.81***	.65***	.71***	.81***	.78***	.67***	.71***	.85***	.49***	.73***	.77***	.65***	.87***							
17. Mental Agility	3.66	.83	.60***	.67***	.64***	.64***	.62***	.60***	.64***	.54***	.65***	.66***	.46***	.64***	.67***	.67***	.63***	.69***						
18. Mil. & Prof. Bearing	3.74	.82	.73***	.64***	.69***	.59***	.58***	.69***	.75***	.61***	.68***	.68***	.45***	.63***	.65***	.62***	.69***	.73***	.60***					
19. Prepares Self	3.73	.90	.60***	.71***	.70***	.53***	.61***	.70***	.69***	.57***	.65***	.69***	.56***	.74***	.72***	.64***	.74***	.77***	.64***	.69***				
20. Resilience	3.86	.80	.60***	.63***	.61***	.75***	.66***	.61***	.56***	.58***	.49***	.65***	.66***	.65***	.67***	.64***	.59***	.64***	.66***	.59***	.65***			
21. Sound Judgment	3.59	.75	.65***	.71***	.70***	.66***	.70***	.77***	.64***	.70***	.67***	.73***	.49***	.70***	.72***	.73***	.75***	.75***	.75***	.72***	.73***	.69***		
22. Stewards the Prof.	3.59	.85	.69***	.75***	.67***	.65***	.69***	.75***	.74***	.65***	.62***	.84***	.45***	.67***	.77***	.64***	.70***	.77***	.68***	.71***	.72***	.69***	.73***	
23. Warrior Ethos	3.73	.82	.86***	.70***	.70***	.69***	.70***	.76***	.83***	.73***	.70***	.77***	.50***	.67***	.71***	.61***	.72***	.76***	.64***	.69***	.65***	.70***	.68***	.75***

*Note. N*s range from 131 to 135. $^{+}$ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$.

Table 4

Descriptive Statistics and Correlations for Attribute and Competency Self-Ratings

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1. Army Values	3.93	.85																						
2. Builds Trust	4.12	.78	.48***																					
3. Communicates	3.88	.91	.44***	.60***																				
4. Confidence	3.98	.89	.44***	.53***	.61***																			
5. Creates a Pos. Env.	3.98	.80	.54***	.51***	.56***	.51***																		
6. Develops Others	3.68	.83	.45***	.55***	.57***	.49***	.47***																	
7. Disciplined	4.03	.84	.67***	.52***	.44***	.47***	.49***	.46***																
8. Empathy	3.84	.91	.63***	.46***	.39***	.41***	.39***	.40***	.59***															
9. Expertise	3.79	.83	.45***	.46***	.51***	.54***	.47***	.53***	.43***	.41***														
10. Extends Influence	3.77	.89	.46***	.53***	.52***	.50***	.47***	.51***	.45***	.39***	.51***													
11. Fitness	3.87	.92	.24**	.27**	.38***	.50***	.43***	.30**	.33***	.16+	.48***	.37***												
12. Gets Results	3.92	.79	.39***	.35***	.47***	.56***	.40***	.40***	.45***	.34***	.52***	.51***	.52***											
13. Innovation	3.74	.81	.42***	.40***	.50***	.58***	.35***	.50***	.35***	.46***	.55***	.50***	.40***	.57***										
14. Interpersonal Tact	3.89	.87	.49***	.60***	.63***	.67***	.52***	.52***	.42***	.52***	.48***	.46***	.38***	.50***	.59***									
15. Leads by Example	3.82	.94	.47***	.46***	.49***	.44***	.51***	.54***	.47***	.34***	.46***	.63***	.39***	.51***	.41***	.44***								
16. Leads by Others	3.79	.94	.28**	.47***	.44***	.56***	.43***	.52***	.33***	.32***	.58***	.61***	.43***	.49***	.48***	.47***	.66***							
17. Mental Agility	3.86	.76	.46***	.44***	.59***	.53***	.37***	.54***	.38***	.45***	.53***	.39***	.37***	.59***	.55***	.65***	.40***	.37***						
18. Mil. & Prof. Bearing	4.09	.89	.62***	.56***	.62***	.64***	.58***	.66***	.61***	.51***	.45***	.55***	.37***	.39***	.43***	.59***	.51***	.47***	.53***					
19. Prepares Self	3.91	.77	.52***	.45***	.44***	.50***	.43***	.52***	.53***	.47***	.55***	.52***	.47***	.52***	.48***	.47***	.51***	.42***	.46***	.60***				
20. Resilience	4.04	.83	.45***	.44***	.51***	.57***	.38***	.46***	.45***	.49***	.43***	.37***	.36***	.49***	.45***	.52***	.37***	.35***	.60***	.52***	.41***			
21. Sound Judgment	3.78	.83	.52***	.57***	.61***	.58***	.48***	.57***	.47***	.40***	.58***	.48***	.36***	.55***	.57***	.67***	.51***	.49***	.68***	.58***	.45***	.57***		
22. Stewards the Prof.	3.57	.92	.41***	.50***	.66***	.54***	.48***	.52***	.43***	.27***	.55***	.55***	.46***	.51***	.52***	.50***	.52***	.52***	.49***	.51***	.48***	.56***	.56***	
23. Warrior Ethos	3.88	.83	.72***	.53***	.52***	.51***	.63***	.50***	.71***	.67***	.56***	.58***	.41***	.49***	.48***	.50***	.54***	.46***	.49***	.57***	.57***	.52***	.58***	.54***

*Note. N*s range from 115 to 120. + $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$.

Table 5

Descriptive Statistics and Correlations for Attribute and Competency Ratings Across All Raters (SLs, TLs, Peers, Self)

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1. Army Values	3.77	.81																						
2. Builds Trust	3.68	.87	.75***																					
3. Communicates	3.57	.84	.76***	.75***																				
4. Confidence	3.70	.89	.59***	.62***	.69***																			
5. Creates a Pos. Env.	3.77	.81	1.00***	.75***	.76***	.59***																		
6. Develops Others	3.35	.76	.70***	.73***	.72***	.66***	.70***																	
7. Disciplined	3.69	.91	.83***	.74***	.72***	.61***	.83***	.71***																
8. Empathy	3.65	.77	.77***	.69***	.68***	.56***	.77***	.63***	.68***															
9. Expertise	3.49	.80	.63***	.64***	.63***	.60***	.63***	.69***	.63***	.58***														
10. Extends Influence	3.42	.77	.67***	.82***	.74***	.71***	.67***	.75***	.70***	.64***	.65***													
11. Fitness	3.60	1.01	.43***	.44***	.51***	.62***	.43***	.53***	.49***	.37***	.52***	.51***												
12. Gets Results	3.69	.87	.66***	.61***	.68***	.66***	.66***	.61***	.67***	.62***	.66***	.64***	.52***											
13. Innovation	3.37	.73	.64***	.67***	.71***	.61***	.64***	.74***	.62***	.61***	.66***	.71***	.53***	.69***										
14. Interpersonal Tact	3.63	.88	.60***	.63***	.66***	.58***	.60***	.65***	.56***	.63***	.56***	.67***	.42***	.60***	.66***									
15. Leads by Example	3.41	.91	.71***	.72***	.72***	.61***	.71***	.72***	.72***	.65***	.66***	.73***	.54***	.64***	.62***	.60***								
16. Leads by Others	3.37	.87	.66***	.75***	.70***	.73***	.66***	.79***	.66***	.61***	.72***	.82***	.57***	.65***	.70***	.65***	.78***							
17. Mental Agility	3.55	.77	.66***	.73***	.69***	.65***	.66***	.67***	.67***	.59***	.65***	.72***	.48***	.70***	.78***	.61***	.60***	.66***						
18. Mil. & Prof. Bearing	3.70	.87	.80***	.75***	.72***	.68***	.80***	.70***	.78***	.63***	.73***	.73***	.53***	.64***	.66***	.60***	.72***	.72***	.69***					
19. Prepares Self	3.58	.88	.66***	.67***	.68***	.59***	.66***	.68***	.69***	.56***	.67***	.69***	.60***	.65***	.64***	.56***	.77***	.71***	.64***	.69***				
20. Resilience	3.78	.81	.67***	.70***	.70***	.78***	.67***	.65***	.67***	.63***	.60***	.72***	.68***	.68***	.65***	.59***	.66***	.71***	.70***	.68***	.67***			
21. Sound Judgment	3.49	.75	.72***	.76***	.75***	.64***	.72***	.74***	.67***	.65***	.72***	.69***	.54***	.68***	.71***	.65***	.71***	.70***	.72***	.77***	.74***	.68***		
22. Stewards the Prof.	3.35	.82	.65***	.69***	.68***	.68***	.65***	.72***	.73***	.60***	.68***	.75***	.57***	.68***	.74***	.59***	.68***	.73***	.66***	.68***	.67***	.75***	.67***	
23. Warrior Ethos	3.72	.81	.84***	.73***	.75***	.71***	.84***	.74***	.84***	.76***	.69***	.75***	.60***	.70***	.69***	.63***	.77***	.73***	.68***	.75***	.72***	.78***	.74***	.78***

*Note. N*s range from 167 to 172. $^{+}$ $p < .10$; * $p < .05$; ** $p < .01$; *** $p < .001$.