Technical Report 1381

Army Command Climate: The Viability of Single-Item Measures

Cory Adis
Chelsey Byrd
Michelle Wisecarver
Kristen Horgen
Personnel Decisions Research Institutes, Inc.

Jessica Badger Darrow
Richard Hoffman
U.S. Army Research Institute

Katherine Rahill
Consortium Research Fellows Program

January 2020

United States Army Research Institute
for the Behavioral and Social Sciences

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

Department of the Army
Deputy Chief of Staff, G1

Authorized and approved:

MICHELLE ZBYLUT, Ph.D.
Director

Research accomplished under contract
for the Department of the Army by

Personnel Decisions Research Institutes, LLC

Technical review by

Jenna Newman, U.S. Army Research Institute

NOTICES

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

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

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

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
January 2020	Final	October 2015 - June 2017

4. TITLE AND SUBTITLE
Army Command Climate: The Viability of Single-Item Measures

5a. CONTRACT NUMBER
W911NF-11-D-0001/DO 0341

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
622 785

6. AUTHOR(S)
Cory Adis, Chelsey Byrd, Michelle Wisecarver, Kristen Horgen (Personnel Decisions Research Institutes, Inc.); Jessica Badger Darrow, Richard Hoffman (U.S. Army Research Institute); Katherine Rahill (Consortium Research Fellows Program)

5d. PROJECT NUMBER
A790

5e. TASK NUMBER
28

5f. WORK UNIT NUMBER
777

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Personnel Decisions Research Institutes, LLC
1911 N. Fort Myer Drive, Suite 410
Arlington, VA 22209

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U. S. Army Research Institute
for the Behavioral & Social Sciences
6000 6TH Street (Bldg. 1464 / Mail Stop 5610)
Fort Belvoir, VA 22060-5610

10. SPONSOR/MONITOR'S ACRONYM(S)
ARI

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
Technical Report 1381

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

13. SUPPLEMENTARY NOTES
ARI Research POC: Jessica Badger Darrow, Foundational Science Research Unit

14. ABSTRACT
All Active Duty Army company commanders are required to conduct an initial command climate assessment within 30 days of assuming command, with assessments 6 months and 12 months later, and annually thereafter. If command climate dimensions could validly be assessed with single-item as opposed to multi-item assessments, this would significantly reduce the time requirement for Soldiers taking the surveys. A series of 13 multi-item command climate scales developed by ARI in prior research was used as the foundation for the single-item command climate measures. Existing literature on organizational climate and single-item construct measures was used to create unidimensional definitions for each construct. These definitions were leveraged as the single-item measures for the dimensions. Data were collected from 1,683 Soldiers at 55 companies across eight locations to evaluate psychometric and criterion-related validity and compare results using the single-item and multi-item scales. Results generally supported the feasibility of using the 13-item assessment tool as a quick "health check" on the command climate of a unit. Modifications for various scales and single items that could be changed to improve the single-item reliabilities are discussed and recommendations are made for further research and practical application.

15. SUBJECT TERMS
Industrial Psychology, Military Psychology, Questionnaires (Surveys), Leadership, Teams (Personnel), Teamwork, Group Dynamics, Morale, Company Level Organizations

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Gregory Ruark
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER
Unclassified	Unclassified	Unclassified	Unlimited Unclassified	127	703-545-2441

Technical Report 1381

Army Command Climate: The Viability of Single-Item Measures

Cory Adis
Chelsey Byrd
Michelle Wisecarver
Kristen Horgen
Personnel Decisions Research Institutes, LLC

Jessica Badger Darrow
Richard Hoffman
U.S. Army Research Institute

Katherine Rahill
Consortium Research Fellows Program

Foundational Science Research Unit
Gregory A. Ruark, Chief

United States Army Research Institute
for the Behavioral and Social Sciences

January 2020

ACKNOWLEDGMENTS

We would like to thank the numerous Soldiers who took time to complete the surveys required to conduct this research. While our goal is to reduce the survey burden on Soldiers, achieving that requires an additional survey burden and we appreciate their time and thoughtful responses. We would also like to thank Ms. Pam Butler for all of her assistance in scheduling the numerous Umbrella Week data collections that were required, as well as Dr. Jenna Newman and a second anonymous reviewer for their suggestions and improvements on an earlier version of this manuscript.

ARMY COMMAND CLIMATE: THE VIABILITY OF SINGLE-ITEM MEASURES

EXECUTIVE SUMMARY

Research Requirement:

Army Directive 2013-29 (Army Command Climate Assessments, McHugh, 2013) requires that all Active Duty Army company commanders conduct an initial command climate assessment within 30 days of assuming command, with subsequent assessments 6 months and 12 months later, and annually thereafter while in command. Because this assessment is repeated multiple times, the assessments can be lengthy, and there are many other surveys taken by Soldiers in addition to command climate surveys; survey fatigue is an important concern. Survey fatigue can lead to low response rates and potentially decrease the quality of the data received. If command climate dimensions could be validly assessed with single-item as opposed to multi-item scales, this would reduce significantly the time requirement for Soldiers taking the surveys. This project developed single-item measures to assess various dimensions of command climate and evaluate their psychometric properties and criterion-related validity.

Approach:

A series of 13 multi-item command climate scales developed by the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) in prior research was used as the foundation for the single-item command climate measures. Existing literature on organizational climate and single-item construct measures were used to create unidimensional definitions for each construct. These definitions were leveraged as the single-item measures for the 13 dimensions. Items were also developed for each dimension that could serve as follow-up measures to further diagnose problems, if the single-item measures identified a potential problem area.

Data were collected from 1,683 Soldiers from 55 companies across eight locations in order to compare results using the single-item and multi-item scales. Soldiers completed the full command climate scales, single-item measures, diagnostic items, and a series of criterion measures such as morale, unit performance, and job satisfaction. Performance ratings of each company as a whole were also collected, either from the battalion commander or another member of the battalion leadership team.

Findings:

Results generally supported the feasibility of using the 13-item assessment tool as a quick “health check” on the command climate of a unit. Convergent validities of the single items with the multi-item scales were above .7 for 10 of the 13 scales, and predictive validities with proposed outcomes were comparable to predictive validities achieved by the multi-item scales. Several of the scales demonstrated lower than expected levels of within-group agreement, suggesting the constructs may not measure a group-level effect. Hierarchical regressions were used to determine whether the multi-item scales explained significantly more variance in the associated criteria than the single items. Of the 18 dimension-criteria relationships that were significant, hierarchical regressions for five of the relationships found that the multi-item scale

added significant variance to the single-item in predicting the criteria. The increases ranged from 5 to 10%.

While all of the multi-item scales demonstrated high reliabilities, the estimated reliabilities for the single-item measures were more moderate. The single-item reliability estimate is dependent upon the correlation between the single item and the multi-item scale and the reliability of the multi-item scale. Modifications for various scales and single items to improve the single-item reliabilities are discussed.

In order to compare the practical application of the single items with the multi-item scales, an arbitrary cutoff of 3.0 was used to identify companies as having high or low command climate scores for each dimension. Results found that the single items performed very similarly to the multi-item scales in identifying dimensions as above or below the threshold. For the five dimensions with lower agreement between the scale and the single item across the companies, four had means that were very close to the 3.0 threshold, indicating that slight differences in the mean value were more likely to shift the result above or below the threshold. Overall, regression results and the practical application analysis support the effectiveness of single-item measures.

Utilization and Dissemination of Findings:

Initial findings demonstrated the potential of the single-item measure and diagnostic follow-up approach for assessing command climate utilizing few items and thereby reducing survey burden. Additional research is needed to further assess the psychometric properties of the single-items prior to operational use by the Army and other DoD entities. Also, the methodology used in this research could be applied to other constructs for which survey length is a concern. However, further research would be needed before this approach is adopted.

ARMY COMMAND CLIMATE: THE VIABILITY OF SINGLE-ITEM MEASURES

CONTENTS

LIST OF TABLES

ARMY COMMAND CLIMATE: THE VIABILITY OF SINGLE-ITEM MEASURES

An organization's climate reflects members' perceptions of various aspects of the organization, such as the level of support members receive from leaders and peers, perceptions of fairness, and respect for others. Leaders play a pivotal role in developing organizational climate. They are responsible for creating and maintaining a positive climate of trust through actions and communication with subordinates, stating unit goals and priorities, role modeling desired behaviors, and reinforcing appropriate behaviors through rewards and recognition (e.g., Barling, Loughlin, & Kelloway, 2002; Dragoni, 2005; Lempke, 1988; Schneider, Ehrhart, Mayer, Saltz, & Niles-Jolly, 2005). Because leaders play such a key role in developing the climate in a unit, the Army specifically refers to a unit's climate as the "command climate" (e.g., Lempke, 1988). This report, therefore, specifically will use the term *command climate* to refer to the unit's organizational climate.

Because climate mirrors what is important to and valued by the organization, organizations are more likely to achieve their goals if their climate is supportive (Murray, 2003). Experienced commanders indicate that a supportive organizational climate can create cohesion in a unit, help Soldiers endure hardships in combat, and encourage a willingness to sacrifice (Murray, 2003). Commanders have also found that climate can help to ensure that the actions taken by Soldiers are in line with desired values and professional military ethics (e.g., Doty & Gelineau, 2008). Similarly, empirical research suggests that organizational climate impacts the productive and counterproductive behaviors in an organization, which in turn impact organizational effectiveness (Ehrhart & Raver, 2014).

Given that the development and maintenance of command climate is critical for unit effectiveness, it is valuable to measure and track the climate over time so that leaders can identify emerging problems and monitor their efforts to change or improve the climate in certain areas. For this reason, Army directives currently require that company commanders conduct command climate assessments within 30 days of assuming command, with subsequent assessments 12 months later and annually thereafter. While this is an important activity, completing the measures is time-consuming because (a) they are collected multiple times, (b) the entire company is given the opportunity to complete it, and (c) command climate has multiple dimensions, resulting in dozens of questions when standard multi-item scales are used.

Research has consistently shown that frequent surveying requires more effort from respondents (e.g., Adams & Umbach, 2012; Berk, Schur, & Feldman, 2007; Bolt, van der Heide, & Onwuteaka-Philipsen, 2014; Meade & Craig, 2012). Survey burden can cause respondents' attention to wane over the course of taking assessments, which may bias them to leave partial or incomplete responses (Berk et al., 2007; Meade & Craig, 2012).

The perceived importance of command climate in predicting negative behaviors has led to an increase in the number of survey programs, methodologies, and sampling methods used to collect data across Department of Defense (DoD) and military organizations (Defense Human Resources Activity, 2015, p. 9). There has also been a concurrent trend in lower response rates for surveys in general within the DoD and military organizations. The response rates for DoD surveys have decreased 15% over the past decade from 35% to a mere 20% (Defense Human Resources Activity, 2015, p. 1). Consequently, results may disproportionately reflect opinions

that are not representative of the actual Soldier population. This lack of survey data also makes it difficult to evaluate leaders, programs, and policies at higher echelons.

Given the increased burden placed on service members to respond to numerous surveys each year, it is likely that response rates are most negatively impacted by the survey length, redundant survey items, and over-surveying in general (Berk et al., 2007; Defense Human Resources Activity, 2015; Meade & Craig, 2012). To alleviate the causes of low response rates, there is a need for more research and development of solutions for reducing survey burden to increase response rates without compromising the quality of climate survey data.

One potential solution may be to reduce the number of items in the scales for each dimension. Some literature suggests that even single-item measures can be effective (e.g., Bergkvist & Rossiter, 2007; de Boer et al., 2004; DeSalvo, Fisher, Tran, Bloser, Merrill, & Peabody, 2006; Dolbier, Webster, McCalister, Mallon, & Steinhardt, 2005; Fuchs & Diamantopoulos, 2009). In order to investigate this methodology, the current research examined the validity of single-item dimension measures for Army command climate. This report will first discuss the assessment of command climate and describe the identification of single-item measures for Army command climate dimensions. The report will then describe a study that examined the psychometric properties and criterion validity of the single-item measures and associated multi-item command climate scales.

Measuring Command Climate

The importance of command climate in the Army has been recognized for decades (e.g., Lempke, 1988). In a 1988 U.S. Army War College report on command climate, Lempke defined climate as "a state or condition existing from shared feelings and perceptions among soldiers about their unit, about their leaders, and about their unit's programs and policies" (p. ii). In the report, he indicated that the term "leadership climate" first appeared in an official Army publication in FM 22-100 *Military Leadership* in October 1983 (U.S. Department of the Army, 1990). The current Army definition of climate is very similar to the one provided by Lempke, with ADRP 6-22 (U.S. Department of the Army, 2012) defining climate as the shared perceptions and attitudes about the unit's daily functioning that describe how members feel about the organization (see p. 7-1). Comparable definitions are found in the research literature outside of the Army. For example, Ehrhart, Schneider, and Macey (2014, p. 69) define climate as "the shared meaning organizational members attach to the events, policies, practices, and procedures they experience and the behaviors they see being rewarded, supported, and expected." Similarly, Schneider and colleagues defined climate as "the shared meaning employees attach to the policies, practices, and procedures and the behaviors that get rewarded, supported, and expected at work" (Schneider, 1990, p. 384; Schneider & Reichers, 1983; Schneider, White, & Paul, 1998). These definitions each highlight the importance of climate as a "shared" as opposed to individual construct, emphasizing the importance of viewing climate as a group-level construct. In addition, the definitions describe climate in terms of the feelings, perceptions, attitudes, or meaning that the members have and share. Therefore, two key elements that should be considered in the measurement of command climate are (a) that it is shared, not individual, and (b) that it is a perception or understanding.

Describing command climate as a shared perception has important implications for the measurement of the construct. While measures of individual affect reflect *individual* experiences, the measure of a shared construct should have items that reflect *unit* characteristics (Glick, 1985). The following provides an example that compares an individually-focused item with a unit-focused item:

- **Individual**: Company leadership trusts me to make decisions.
- **Unit**: Company leadership trusts the members of this unit to make decisions.

Results of a meta-analysis examining justice climate found that taking a group/organization-referenced measurement approach resulted in a stronger correlation between justice climate and group/organizational effectiveness, with a correlation of $\rho = .53$ for the group/organization-referenced approach and $\rho = .23$ for individually-referenced items (Whitman, Caleo, Carpenter, Horner, & Bernerth, 2012). When capturing shared perceptions at a unit level, Glick (1985) also indicates it is important to be specific regarding the unit level in question; for example, specifically referring to the platoon, company, battalion, etc., as opposed to leaving the referent group ambiguous.

Another implication of viewing command climate as a unit-level construct is conducting analyses at the unit level. This involves aggregating individual-level responses to the appropriate unit level (Kozlowski & Klein, 2000). Responses should be aggregated if there is sufficiently high agreement among them to warrant combining (Ehrhart et al., 2014). Level of agreement can be tested using a within-group correlation, $r_{wg(j)}$ (James, 1982; James, Demaree, & Wolf, 1984), as well as intraclass correlations (ICC; Bliese, 2000; de Boer et al., 2004). ICC(1) compares variability within units to variability across units, providing a percent of variance explained by the unit. ICC(2) is a similar measure that provides an estimate of the reliability of the unit mean. Researchers look for both ICCs to reach critical thresholds to demonstrate large and reliable group-level effects (Bliese, 2000).

A third consideration for measuring command climate is the specificity of the dimensions that are selected to reflect the climate. At one extreme, command climate can be measured very broadly by capturing climate as a general positive or negative experience; alternatively, a great degree of specificity can be used to capture many different facets of the climate. These can include dimensions that are generalizable to many other organizations, such as managerial or leader support, and ones that are specific to one or a few types of organizations, such as insurance agent independence—which is the degree of latitude given to insurance agents—as a dimens ion of climate for managers at life insurance agencies (Schneider & Bartlett, 1968, 1970). Litwin and Stringer (1968) suggest that it is likely there are an infinite variety of climate dimensions. Ehrhart et al. (2014) identified 10 dimensions that were commonly used in climate taxonomies to summarize the content of organizational climate (see Table 1).

Table 1

Ten Dimensions Commonly Identified in Taxonomies of Climate (Ehrhart et al., 2014)

Climate dimension	References
Structure/standards and leader structure/standards	Litwin & Stringer (1968); Schneider & Bartlett (1968); Campbell, Dunnette, Lawler, & Weick (1970); Jones & James (1979)
Consideration/facilitation and leader consideration/facilitation	Campbell et al. (1970); Jones & James (1979); James & James (1989)
Warmth	Litwin & Stringer (1968); Campbell et al. (1970); Jones & James (1979); James & James (1989)
Support and leader support	Litwin & Stringer (1968); Schneider & Bartlett (1968); Campbell et al. (1970); Jones & James (1979); James & James (1989)
Rewards	Litwin & Stringer (1968); Campbell et al. (1970)
Conflict/stress	Schneider & Bartlett (1968); Jones & James (1979); James & James (1989)
Autonomy/independence	Litwin & Stringer (1968); Schneider & Bartlett (1968); Campbell et al. (1970); James & James (1989)
Satisfaction/spirit	Schneider & Bartlett (1968); Jones & James (1979)
Challenge/variety	Jones & James (1979); James & James (1989)
Cooperation	Jones & James (1979); James & James (1989)

Examples of more specific climate dimensions that have been identified in the literature can be seen in Table 2. Note that the list in Table 2 provides only a small number of examples from a large number of specific dimensions of climate that have been identified.

Table 2

List of Example Climate Dimensions Identified in the Literature

Climate dimension	Reference
Agent independence	Schneider & Bartlett (1968, 1970)
Intra-agency conflict	Schneider & Bartlett (1968, 1970)
Decision centralization	Pritchard & Karasick (1973)
Flexibility and innovation	Pritchard & Karasick (1973)
Achievement	Pritchard & Karasick (1973)
Leader's psychological distance	Payne & Mansfield (1973)
Open-mindedness	Payne & Mansfield (1973)
Emotional control	Payne & Mansfield (1973)
Job challenge	Payne & Mansfield (1973)
Rules orientation	Payne & Mansfield (1973)

Specific to the Army, Lempke (1988) described various dimensions identified as important for command climate by military commanders and documents over the years. These include:

- Command trust
- Organizational consistency
- Organizational simplicity
- Command stress
- Mutual respect and human dignity
- Innovation
- Competence
- Caring
- Responsibility
- Accountability
- Challenge
- Inclusion/All make a difference

Other examinations of climate from experienced Army officers identified similar dimensions. Craig and Brace (1992), for example, identified dimensions such as leader caring, communications, information flow, personal worth, growth and development, and trust and empowerment to make decisions. Nearly a decade later, four groups of five officers each at the

U.S. Army War College worked independently to identify key elements of command climate. All four groups identified communication and trust as key elements of climate (Bullis & Reed, 2003). Other elements identified by one or more groups were: respect, focus, predictability/consistency, clear goals/shared vision, resource balance/alignment, esprit de corps, loyalty, respect, team building, learning/professional development, and decentralized execution (Bullis & Reed, 2003).

The Defense Equal Opportunity Management Institute (DEOMI) first developed the Military Equal Opportunity Climate Survey (MEOCS) then replaced it in 2005 with the DEOMI Organizational Climate Survey (DEOCS; Truhon & Parks, 2005). The DEOCS is the measurement tool that was mandated in 2013 by Army Directive 2013-29 to regularly assess Army command climate. The purpose of DEOCS is to assess command climate and equal opportunity issues across a variety of DoD organizations. In 2015, at the start of this research, DEOCS 4.0 was the survey version currently in use, which assessed four major topics (DEOMI, 2014):

- Equal opportunity (EO)/equal employment opportunity (EEO)/fair treatment
- Organizational effectiveness (OE)
- Discrimination/sexual harassment
- Sexual assault prevention and response (SAPR)

Within the EO/EEO/fair treatment category, there were nine dimensions relevant to Soldiers (hazing; demeaning, racist, and sexist behaviors; favoritism; racial, sex, and religious discrimination; and sexual harassment)[1]. Within the OE category there were 11 dimensions, shown in Table 3. Some of the DEOCS dimensions specifically capture aspects of command climate (e.g., trust in leadership, diversity management), while other dimensions capture important constructs that are related to but distinct from command climate (e.g., organizational commitment, organizational cohesion, intention to stay). The DEOCS factors and descriptions are provided in Table 3.

[1] There are two additional dimensions, age and disability discrimination, but they are listed as relevant only to civilian organizations.

Table 3

DEOMI Organizational Climate Survey (DEOCS) Organization Effectiveness Dimensions

DEOCS dimension	Description of items
Organizational commitment	Sense of belonging to organization
Trust in leadership	Trust in leader support and fairness
Organizational performance	Evaluation of performance
Organizational cohesion	Trust and support for each other
Leadership cohesion	Leader support and cooperation
Job satisfaction	Enjoyment of job
Diversity management	Members feeling valued and included
Organizational processes	Accountability and fairness of discipline and decisions communication
Intention to stay	Career intentions
Help seeking behaviors	Seeking help for clinical illnesses
Exhaustion/burnout	Mental, physical, emotional tiredness

Including the Background section, DEOCS 4.0 contains 112 items[2] (Defense Equal Opportunity Management Institute, 2014). While DEOCS captures many climate dimensions identified as important in previous research, other climate dimensions identified as important in Army units, such as communication or challenge, are not included. Up to 10 additional questions can be added to DEOCS by each commander, using a four-point Likert response scale that ranges from strongly agree to strongly disagree. In addition, the commander can add five open-ended questions.

In summary, climate is an important construct for the Army. Key considerations in measuring climate are aggregating to the appropriate unit level and determining the dimensions of interest and degree of specificity on those dimensions. Organizational research had identified numerous climate dimensions that are valuable to measure, and the Army currently measures various climate dimensions related to equal opportunity through the DEOCS. The current research identifies climate dimensions relevant to the Army to supplement the DEOCS and explores the usefulness of employing single-item measures to assess these climate dimensions.

[2] Note that there are 95 numbered items in DEOCS 4.0, but given that some of the numbered items have multiple questions, this generates a total of 112 items for military members.

Leveraging Single-Item Measures

Measurement scales in the social sciences often use a series of self-report items to ascertain the level of a theoretical construct that is not readily observable. The concept of having multiple items to estimate the construct level rather than a single item recognizes that measurement of any type will have error. By including multiple observations, the multiple measurements can be averaged and thus have a better chance of identifying the true level of the construct (e.g., DeVellis, 2012). Because there are drawbacks to using multiple items, however, it is important to consider the costs and benefits of the multi-item approach and whether a single-item measure can provide a sufficient assessment of the construct. Three areas to consider in this decision are the composition of the construct being measured, practical issues (time, cost, application of measure), and empirical evidence of reliability and validity.

Composition of the construct. While most attitude and perception measurement involves measuring a latent construct that is not readily observable, constructs differ greatly in the extent to which they are concrete or abstract and whether they capture unidimensional or multidimensional concepts. The more abstract a construct is to understand, the more likely multiple items will be needed to capture the construct level accurately (e.g., Bergkvist & Rossiter, 2007; Fuchs & Diamantopoulos, 2009; Rossiter, 2002; Sackett & Larson, 1990). It is also important to use multi-item scales with constructs that are complex or multifaceted (Fuchs & Diamantopoulos, 2009; Rossiter, 2002). Conversely, for scales in which the multiple items that represent the dimension are intended to be synonymous, a strong argument can be made for the appropriateness of a single-item measure (Bergkvist & Rossiter, 2007). Increasing the number of synonymous items can actually produce a greater chance of including items that are not proper synonyms; thus, in these situations, fewer items may be better (Drolet & Morrison, 2001).

Some single-item measures are global items, or items that encompass a complex dimension (Diamantopoulos & Winklhofer, 2001). For example, in the job satisfaction literature, global single-item measures of satisfaction have been compared with multi-item scales (see Oshagbemi, 1999; Wanous, Reichers, & Hudy, 1997). Wanous et al. (1997) conducted a meta-analysis of the relationship between single-item and multi-item measures of overall job satisfaction and found an average observed correlation of .63 between single items and multi-item scales. The highest correlations for multi-item scales were found for scales with items that focused on overall job satisfaction, with lower correlations for scales containing items that reflected different aspects of job satisfaction. This emphasizes the impact that the nature of the scale items will have on the performance of the single-item measure. The global approach to single-item measures has been used successfully in a number of fields, such as quality of life assessment (e.g., de Boer et al., 2004), sports management (e.g., Kwon & Trail, 2005), citizen satisfaction (e.g., Van Ryzin, 2004), self-esteem (Robins, Hendin, & Trzesniewski, 2001), and teaching effectiveness (Wanous & Hudy, 2001).

The complexity of the facets of command climate and conceptual independence of dimensions of command climate such as leader support, inclusion, and performance orientation suggest the importance of capturing items relevant to each of the identified command climate dimensions. As described, the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) identified 13 important Army command climate dimensions; each of these

dimensions was measured by multiple items. While overall command climate would likely be too multifaceted to capture using a single item, each of the 13 command climate dimensions could be sufficiently unidimensional and concrete to measure using a single item.

Practical issues. There are numerous practical issues that must be considered when contemplating a single-item versus multiple-item measurement approach. The single-item approach provides a high degree of brevity and parsimony, minimizing the time and cost needed to complete the survey and maximizing its ease of use (e.g., Bergkvist & Rossiter, 2007; Kwon & Trail, 2005). Respondents who are exposed to more items tend to distinguish less between them, with earlier items having a stronger influence than later items; hence, more items may lead to boredom and mindless response behavior (Drolet & Morrison, 2001). Multiple-item scales can also reduce face validity, as the respondents sense they are being asked the same questions repeatedly (Fuchs & Diamantopoulos, 2009), and can be prone to consistency motif bias, in which subjects try to maintain consistency in their responses to similar questions (Podsakoff, MacKenzie, Lee, & Podsakoff., 2003). From a practical standpoint then, single-item measures have a number of benefits over multi-item assessments. One practical challenge with single-item assessments is choosing or developing the single item that will be most representative of the construct (Diamantopoulos, Sarstedt, Fuchs, Wilczynski, & Kaiser, 2012).

Empirical evidence of reliability and validity. While construct complexity and practical issues are important to consider, arguably the most important question is whether single-item measures are reliable and valid. With respect to reliability, while establishing test-retest reliability of single-item measures is straight-forward, establishing internal consistency reliability can be challenging because formulae such as that for Cronbach's alpha cannot be calculated. Wanous and Hudy (2001) describe a few ways in which estimating internal consistency of single items is possible. One accepted approach is to reorganize the correction for attenuation formula or the Spearman-Brown prophecy equation, input known quantities such as the reliability of the multi-item scale and the correlation between single-item and multi-item scales and solve for single-item reliability. In examining single-item reliability, researchers have found evidence for both test-retest reliability (e.g., see de Boer et al., 2004; Shamir & Kark, 2004) as well as internal consistency reliability (e.g., see Dolbier et al, 2005; Ginns & Barrie, 2004; Kwon & Trail, 2005; Wanous & Hudy, 2001; Wanous et al., 1997) for single-item measures, suggesting that they can be sufficiently reliable.

With respect to validity, as mentioned previously, single-item measures can demonstrate higher levels of face validity from the applicant's perspective, because unlike multi-item scales, they do not appear to be unnecessarily repetitious (Fuchs & Diamantopoulos, 2009). Content validity has also been supported for single-item measures, which can be established through standard content valid development procedures or using inter-judge agreement (e.g., Rossiter, 2002). To establish convergent validity, a common approach is to compute the correlation between the single-item measure and the full (multi-item) measure (Fuchs & Diamantopoulos, 2009). The empirical evidence based on this approach has supported the use of single-item measures (e.g., Dolbier et al., 2005; Nagy, 2002; Robins et al., 2001; Wanous & Hudy, 2001; Wanous et al., 1997).

In order to evaluate predictive validity, researchers have compared coefficients for single-item scales to multi-item scales of the same construct (e.g., Bergkvist & Rossiter, 2007). When

comparing validity coefficients, it is important to use an appropriate test to compare correlation coefficients of predictive validities; that is, using a test for related correlation coefficients if the correlations come from paired samples, or a test for independent coefficients if they come from independent samples. Overall, the evidence from studies in fields as diverse as health care (DeSalvo et al., 2006), sports management (Kwon & Trail, 2005), organizational psychology (Nagy, 2002) and marketing (Bergkvist & Rossiter, 2007) show that single-item scales demonstrate predictive validity that is comparable to that of their multi-item counterparts. For example, Bergkvist and Rossiter (2007) compared bivariate correlations between predictors and criteria for single-item and multi-item predictors and found that the single-item predictors were just as good as the multi-item predictors. They also compared the variance accounted for by single- and multi-item predictors when an additional known predictor was included in the two regression models (i.e., the model was more completely specified). Again, the model that included the single item predicted the criterion just as well as the model with the multi-item predictor.

In a series of Monte Carlo simulations designed to examine predictive validity of single-item and multi-item measures, Diamantopoulos et al. (2012) found that when inter-item correlations among items of the predictor were above .70, single-item predictors performed either the same as or better than multi-item predictors in about half of the simulations. When taking the number of items in the scale and sample size into account as well by computing Cronbach's alpha, results indicated that correlations among items of the predictor needed to be approximately .90 or higher for the single-item predictors to perform the same as or better than multi-item predictors in the simulations. While Diamantopoulos et al. caution against generally abandoning multi-item scales in favor of single-item scales, their results do demonstrate that single-item scales provide valid measures under appropriate conditions.

With these factors in mind, the following section will describe the development and validation of a prototype approach to measuring command climate using single-item measures for each dimension.

Creating a Command Climate Single-Item Assessment

To investigate the command climate dimensions that are most important to the Army and determine the feasibility of capturing command climate with single-item measures, ARI developed Army command climate assessments using corresponding multi-item and single-item assessments for a series of dimensions identified as important to Army leaders. In 2015, ARI conducted focus groups and interviews at three U.S. Army Forces Command (FORSCOM) locations to solicit feedback from Army leaders on the most important aspects of command climate at the company level. Participants were asked general questions about the meaning of command climate, the most important factors of command climate, potential outcomes of command climate, and how leaders at various levels affect command climate. Participants were also asked to review command climate elements that were identified through literature reviews and indicate the most important factors.

A total of 82 active Army NCOs and officers participated in the focus groups and interviews. Squad leaders, platoon sergeants, first sergeants, platoon leaders, company commanders, and majors participated in focus groups separated by rank. Battalion commanders

and battalion command sergeants major participated in individual interviews. Participants represented a variety of military occupational specialties (MOS) in combat arms, combat support, and combat service support.

The findings of the focus groups suggested a great deal of overlap between Army leaders' perceptions of command climate and the factors of command climate that had been identified by ARI researchers in the literature. Leader support for families was identified as an important topic that should be included in the survey. Based on responses during the focus groups and interviews and ratings of command climate factors, some factors were removed and others were added.

The final 13 dimensions are listed in Table 4, with corresponding definitions. There is a high degree of convergence between the dimensions in this list and dimensions identified in previous literature as important, with each dimension reflecting a concept already identified. For example, two dimensions repeatedly identified as important are flow of information, which captures perceptions of communication within the unit, and autonomy, which assesses trust in decision-making. Other dimensions that reflect concepts identified previously are respect for the individual, inclusion, and peer and leader support. Hazing and bullying are concepts similar to dimensions captured in the DEOCS EO/EEO/fair treatment category.

Table 4

Description of 14 Dimensions of Army Command Climate

Climate dimension	Definition
1. Flow of information	The degree to which information flows within the unit. This includes, but is not limited to, leaders pushing information to subordinates and seeking information from subordinates.
2. Autonomy	How much the leadership displays trust and empowers Soldiers by giving them the freedom to address mission requirements and solve problems. Units with a high degree of autonomy are also reasonably accepting of mistakes, using them as learning opportunities.
3. Leadership openness	The accessibility and approachability of leaders within the unit to discuss Soldier well-being, mission, and non-mission specific issues and concerns.
4. Respect for the individual	The degree to which all Soldiers in the unit are valued equally, beyond rank or position.
5. Hazing	Any conduct whereby a Servicemember or members regardless of service, rank, or position, and without proper authority, recklessly or intentionally causes a Servicemember to suffer or be exposed to any activity that is cruel, abusive, humiliating, oppressive, demeaning, or harmful. (AR 600-20, 4-19)

6. Bullying	Any conduct whereby a Servicemember or members, regardless of service, rank, or position, intends to exclude or reject another Servicemember through cruel, abusive, humiliating, oppressive, demeaning, or harmful behavior, which results in diminishing the other Servicemember's dignity, position, or status. (AR 600-20, 4-19)
7. Fairness	Outcomes (e.g., promotions, rewards, and developmental opportunities) are distributed to unit members in an objective and impartial manner.
8. Inclusion	Unit members feel integrated into the work-related and social activities of the unit.
9. Unit support	The overall unit concern for Soldier welfare and well-being. Listening to and providing resources for non-mission specific/personal/family issues and problems of the Soldiers. The degree to which the unit is motivated to proactively monitor and provide assistance.
10. Leader support	The degree to which leaders in the unit have a concern for peer and subordinate welfare and well-being and provide informal and formal methods of assistance and support.
11. Family support	Unit leaders provide families with the support and assistance they need.
12. Performance orientation	The degree to which the unit and its members are focused on performance outcomes. The unit directs energy towards performance-based goals and objectives. The unit's collective motivation encourages performance-oriented behaviors.
13. Professionalism	The unit's emphasis on the Profession of Arms, Army and unit standards, and adherence to rules, doctrine, and regulations.

After the initial set of items was developed, ARI conducted cognitive interviews with Army personnel to solicit feedback on the clarity and content of the items. Interviews were conducted with nine participants, including three captains and six NCOs (ranging from sergeant to master sergeant/first sergeant). Participants provided minor wording revisions to improve clarity, and the items were revised based on the feedback provided.

The dimensions and scales provided a foundation for the development of a series of single-item measures. Developing single-item measures for constructs required active consideration of several important factors, which are discussed in the following section.

The dimensions identified by ARI provided the foundation for the development of the single-item command climate measure. Each of the 13 scales was reviewed and two tasks were conducted: (a) reviewing and modifying the construct definitions to ensure they presented a unidimensional construct and (b) identifying the items in each scale that best represented a core unidimensional construct. Items that were outside of the core definition were removed from the construct scale. This resulted in the constructs and definitions shown in Table 5, with each of the

scales having five to eight items. Soldiers are asked the extent to which they agree with the statements, using a 5-point Likert scale to respond (*Strongly disagree, Disagree, Neither agree nor disagree, Agree, and Strongly agree*). In addition, respondents have the option to select the response *N/A–Don't know*. The three support constructs were relabeled to improve construct clarity, with the construct originally labeled unit support retitled as peer support for unit members, leader support was retitled leader support for unit members, and family support was retitled leader support for families. The items identified as outside of the core definition were used to create an experimental diagnostic measure, discussed further in the following section.

Several different methods exist that can be used to select a single item to represent each scale. One is to select an item most representative of the construct based on expert opinion. Another is to select the item using statistical criteria, such as communalities or reliabilities. Using statistical methods may not be optimal as the statistics are subject to sampling variability (Diamantopoulos et al., 2012). For the current measure, single items were developed with the intention of reflecting unidimensional constructs, based on the focus groups and previous research. Because the items were specifically designed to reflect the unidimensional construct underlying the scale items, the single items also served as the definitions of the constructs.

Table 5

Definitions for the Army Command Climate Dimensions

Climate dimension	Definition
1. Flow of information	Information that is important for my unit's success is communicated effectively throughout the unit.
2. Autonomy	Unit leadership trusts the members of this unit to make decisions.
3. Leadership openness	Members of this unit feel they can approach unit leaders to discuss problems and concerns.
4. Respect for the individual	All members of this unit are treated with dignity and respect.
5. Hazing	Members of this unit engage in hazing (i.e., intentionally cause new members to suffer or be exposed to activities that are abusive, humiliating, or harmful as a "rite of passage").
6. Bullying	Members of this unit bully other unit members (i.e., exclude or reject other unit members through abusive, humiliating, or harmful behavior).
7. Fairness	Outcomes (e.g., promotions, rewards, and developmental opportunities) are distributed to unit members in an objective and impartial manner.
8. Inclusion	Unit members feel included in unit activities.

9. Peer support for unit members	Members of this unit support each other.
10. Leader support for unit members	Leaders in this unit provide members with the support and assistance they need.
11. Leader support for families	Unit leaders provide families with the support and assistance they need.
12. Performance orientation	This unit strives for high performance outcomes.
13. Professionalism	This unit emphasizes Army standards, doctrine, and regulations.

A full list of the 13 dimensions, their associated scale items, and the associated single-item measures is provided in Appendix A.

Creating a Follow-up Diagnostic Tool

Items identified as outside the core definition of the command climate dimension informed the follow up diagnostic measures for each of the dimensions. This enables commanders to approach command climate assessment as a process rather than a repeated snapshot. In the first step of the process, the 13 single-item measures are used to identify any areas of potential concern. If potential problem areas are identified in one or more dimensions, the commander can investigate those dimensions further using diagnostic questions that are presented via a short questionnaire or in focus groups conducted by designated staff. In the third step of the process, commanders identify actions for change. The effectiveness of the actions can then be checked in the next single-item command climate assessment as the process repeats. This diagnostic process enables commanders to conduct quick 5-minute "health checks" of command climate using the single items and, if a potential problem emerges, can use the diagnostic tool as needed.

Preliminary items for the diagnostic were the items from each dimension that were outside of the scope of the dimension definition. In most cases, diagnostic items specifically focused either on officer actions or NCO actions. As an example, the dimension flow of information is described as: *Information that is important for the unit's success is communicated effectively throughout the unit.* Diagnostic items for this dimension were:

- To what extent do officers communicate clearly about the following areas?
  - Training schedules
  - Army policies
  - The purpose of the unit's mission

- To what extent do NCOs communicate clearly about the following areas?
  - Training schedules
  - Army policies
  - The purpose of the unit's mission

Soldiers responded using a 5-point Likert scale with corresponding markers (e.g., ranging from *Not at all* to *To a great extent*). An *N/A Don't know* option was offered for each scale, as there are items for which enlisted Soldiers may be unaware of officer actions, and vice versa. The diagnostic items for each dimension can be seen in Appendix A.

The diagnostic items are considered preliminary, and diagnostics were not identified for three dimensions (respect for the individual, inclusion, and peer support for unit members). For two of these dimensions (respect for the individual and inclusion) there was another ARI project underway to identify specific dimension components in greater detail. For the third dimension, peer support for unit members, no items were identified as outside of the dimension definition.

Unlike the full scales and single items, the diagnostic items explore unique aspects of the dimension rather than reflecting the general underlying command climate construct; that is, they don't form a scale measuring a unified underlying construct. Thus, internal consistency measures would not be expected to be high, and criterion-related validities would need to be examined for each item separately, as opposed to creating scale scores.

With the full scale, single items, and diagnostic items defined, a validation study was conducted to evaluate evidence relating to the reliability and validity of the measures. Reliabilities and construct and criterion-related validity of the full scales and single items were evaluated, and exploratory analyses were conducted to examine the diagnostic items.

Method

A unit-based data collection was conducted in order to capture command climate data at the Army company level. Data were collected from 55 Active Army companies across eight military installations in the continental United States (CONUS) and outside the continental United States (OCONUS).

Participants

Participants in the study included 1,683 Soldiers and officers in units across U.S. Army FORSCOM and OCONUS installations. Responses were considered valid and used in analyses if, (a) the respondent completed at least 70% of the questionnaire, (b) tenure in the company was greater than one month, (c) the respondent answered at least 80% correct on five questions designed to assess if the respondents were paying attention to the questions, and (d) the respondent answered *NA/Don't know* to 25% or less of the items. Fifty-six individuals did not complete at least 70% of the questionnaire and were therefore removed from the dataset. Of the remaining participants, 84 were removed because they had tenure within their company of less than one month. One hundred and seven additional individuals were removed for not correctly answering at least 80% of the attention-check questions.

For the inclusion criterion involving *NA/Don't know* responses, participants were removed by dimension rather than by the individual, such that if a participant had greater than 25% *NA/Don't know* for a given dimension, ratings on items in that dimension were not analyzed for that individual. Thus, the final sample sizes for dimension ratings vary by dimension. After applying these inclusion criteria, 1,373 individuals provided usable data for analyses out of the 1,683 total individuals participating (81.58%).

In addition to the individual-level inclusion criteria, we excluded entire companies if they had fewer than 10 participants complete the survey. Out of the 55 participating companies, four were excluded from analyses because they had fewer than 10 survey respondents. After excluding these four companies, the final overall sample size was 51 companies and 1,342 individual respondents.

Ranks for this sample ranged from PV1 to SFC for enlisted Soldiers and from 2LT to 1LT for officers. There was one Warrant Officer who participated. The frequency of each represented rank is presented in Table 6. Average tenure within the Army was 4.53 years (4.24 *SD*), and average tenure within the company was 16.87 months (14.08 *SD*). The sample consisted of 57.8% Combat Arms, 29.8% Combat Support, and 12.3% Combat Service Support.[3] Table 7 provides a breakdown of the MOS representation.

[3] Percentages do not add up to 100% due to rounding.

Table 6

Sample by Rank

Rank	Frequency	Percent
PV1	8	0.6
PV2	122	9.1
PFC	214	15.9
CPL/SPC	514	38.3
SGT	211	15.7
SSG	136	10.1
SFC	44	3.3
WO1	1	0.1
2LT	23	1.7
1LT	38	2.8
Missing	31	2.3
Total	1,342	100.0

Table 7

Sample by MOS

MOS	Frequency	Percent
11	577	43.0
12	88	6.6
15	55	4.1
19	127	9.5
21	1	0.1
22	1	0.1
25	50	3.7
31	98	7.3
35	4	0.3
42	1	0.1
68	10	0.7
74	26	1.9
79	1	0.1
88	114	8.5
91	109	8.1
92	50	3.7
94	15	1.1
Missing	15	1.1
Total	1,342	100.0%

Note. Percentages may not add up to 100% due to rounding.

In addition to the survey participants, 14 battalion leaders provided company-level ratings for 31 companies. Each commander rated between one and four companies, with an average of two companies per commander. The majority of those providing these ratings (11 out of 14, or 78.5%) were LTCs. Other ranks included one COL, one MAJ, and one CSM, at 7.14% each.

Measures

Data were collected using three command climate measures developed or modified during this project: a multi-item scale assessment, a single-item scale assessment, and an experimental diagnostic assessment. These measures assessed scales or related diagnostics associated with the 13 identified dimensions of command climate listed in Table 5. In addition to the command climate assessments, criteria measures were obtained from published literature and administered with the command climate measure to assess criterion-related validity. Climate level is the construct of interest, which was represented by company-level means for each

climate dimension. Respondents were asked to think about their Company/Troop/Battery when the statements referred to their "unit."

Command climate multi-item scales. Thirteen command climate dimensions were measured with 13 scales containing between 4 and 8 items. The numbers of items for each dimension are shown in Appendix A. Responses were recorded on a 5-point Likert scale with anchors ranging from *Strongly disagree* to *Strongly agree*. In addition, an *NA/Don't know* option was provided. Items are provided in Appendix A-1.

Command climate single-item measures. For the single items, dimension definitions were adapted as the single-item assessments. Definitions were intentionally written to be concise and unidimensional. Respondents were asked to indicate the extent to which they agreed that the climate described in the definition is present in their companies. Responses were recorded on a 5-point Likert scale with anchors ranging from *Strongly disagree* to *Strongly agree*. In addition, an *NA/Don't know* option was provided. The 13 single items are listed in Appendix A-2.

Command climate diagnostic assessment. For the diagnostic assessment, nearly all items asked specifically about actions taken by either officers or NCOs. The exception to this was the inclusion of one cyberbullying item, which did not ask about officers or NCOs. The numbers of items for each dimension are shown in Appendix A. Responses were recorded on a 5-point Likert scale with anchors ranging from *Not at all* to *A great extent*. In addition, an *NA/Don't know* option was provided. Items are provided in Appendix A-3.

Outcomes. Table 8 lists constructs identified as theoretically related to each command climate dimension and lists the measurement tool that was used and the number of items. For the most part, the measures are ones that have been used or adapted from previous ARI research, including the Army Life Questionnaire (ALQ; Moriarty, Campbell, Heffner, & Knapp, 2009) and a test of Army unit cohesion (Estrada et al., under review). The exceptions were ones adapted from other commonly used scales in the industrial/organizational psychology literature (e.g., goal orientation, engagement, job satisfaction). Additionally, family satisfaction, family desire to stay in the Army, and participation in unit social events were measured with new scales developed for this study. Because research was not available to identify outcomes for command climate dimensions empirically, outcomes were selected for each command climate dimension based on the conceptual definition of each dimension. Reliabilities for the scales were estimated using Cronbach's alpha.

Affective commitment. Affective commitment is the degree to which a Soldier feels a personal attachment and sense of belonging in the Army (Gade, Tiggle, & Schumm, 2003). For example, one item asked Soldiers to rate how much they agreed with the statement, *I feel emotionally attached to the Army.* This outcome was measured with four items that were rated on a 5-point Likert scale with anchors ranging from *Strongly disagree* to *Strongly agree*. In addition, an *NA/Don't Know* option was provided. The scale had a reliability of .92.

Career intentions. A Soldier's intention for completing his or her current term of service and continuing with a career in the Army was measured with three items that were rated on a 5-point Likert scale with anchors ranging from *Very unlikely* to *Very likely*. The scale did not have sufficient reliability due to low correlations among the items. Therefore, only one item was

retained for use in analyses. This item asked, *How likely is it that you will make the Army a career?* (Moriarty et al., 2009)

Table 8

Outcomes Associated With the 13 Command Climate Dimensions

Climate dimension	Correlate/Outcome	Number of items
1. Flow of information	Morale	2
	Unit performance	4
	Unit performance (officers)	4
	Unit performance (BN CMDR)	3
	Unit readiness (BN CMDR)	3
2. Autonomy	Job satisfaction	6
	Unit performance	4
	Unit performance (officers)	4
	Unit performance (BN CMDR)	3
	Unit readiness (BN CMDR)	3
3. Leadership openness	Affective commitment	4
	Continuance commitment	4
4. Respect for the individual	Mutual trust	2
	Cohesion	10
5. Hazing (from AR 600-20, 4-19)	Stress	4
	Affective commitment	4
	Continuance commitment	4
	Morale	2
6. Bullying (from AR 600-20, 4-19)	Stress	4
	Morale	2
	Participation in unit social events	5
	Affective commitment	4
	Continuance commitment	4

7. Fairness (i.e., distributive justice)	Withdrawal	6
	Affective commitment	4
	Continuance commitment	4
	Career intentions	3
	Morale	2
8. Inclusion	Withdrawal	6
	Affective commitment	4
	Participation in unit social events	5
9. Peer support for unit members	Well being	3
	Morale	2
	Stress	4
	Affective commitment	4
	Cohesion	10
	Unit resilience	4
10. Leader support for unit members	Affective commitment	4
	Wellness/well being	3
	Unit resilience	4
11. Leader support for families	Family desire to remain in Army	6
	Family satisfaction with communication and support received from unit	2
	Unit performance	4
12. Performance orientation	Unit performance (officers)	4
	Unit performance (BN CMDR)	3
	Unit readiness (BN CMDR)	3
	PT Scores	1
	Weapons qualifications “scores”	1

	# of Article 15s/disciplinary actions	1
13. Professionalism	Weapons qualifications "scores"	1
	# of Article 15s/disciplinary actions	1

Cohesion. Cohesion measured Soldiers' perceptions of how close the unit is, how much pride members took in the unit, and how well unit members worked together as a team. This outcome was measured with 10 items adapted from Siebold and Kelly's (1988) Platoon Cohesion Index (PCI) and Carless and De Paola's (2000) version of the Group Environment Questionnaire (GEQ). For example, one item asked Soldiers to rate how much they agreed with the statement, *Members of my unit work together to get the job done.* Items were rated on a 5-point Likert scale with anchors ranging from *Strongly disagree* to *Strongly agree*. In addition, an *NA/Don't know* option was provided. The scale had a reliability of .94.

Continuance commitment. Continuance commitment measured a Soldier's belief that he or she had to stay in the Army because getting out would cause too much disruption. For example, one item asked Soldiers to rate how much they agreed with the statement, *I am afraid of what might happen if I quit the Army.* This outcome was measured with four items that were rated on a 5-point Likert scale with anchors ranging from *Strongly disagree* to *Strongly agree*. In addition, an *NA/Don't Know* option was provided. Items came from Gade et al. (2003), and the scale had a reliability of .88.

Disciplinary action. Disciplinary action accounted for whether Soldiers received Article 15s in the last year and if so, the number of Article 15s received. It was measured with a single item from Moriarty et al. (2009).

Family desire to stay. Family desire to stay measured the degree to which a Soldier's family supported him or her staying in the Army. Items for this scale were developed for this research. The items asked about family desire to stay in the Army in general, as well as long term desire to stay in the Army. For example, one item asked Soldiers to rate how much they agreed with the statement, *My family wants me to stay in the Army.* This outcome was measured with three items that were rated on a 5-point Likert scale with anchors ranging from *Strongly disagree* to *Strongly agree*. In addition, an *NA/Don't Know* option was provided. The scale had a reliability of .83.

Family satisfaction. Family satisfaction measured the degree to which a Soldier's family was satisfied with their experience with the unit and the Army in general. Items for this scale were developed for this research. This outcome was measured with two items that asked, *How satisfied is your family with their experience in the Army?* and *How satisfied is your family with their experience with the Family Readiness Group (FRG)?* Items were rated on a 5-point Likert scale with anchors ranging from *Very unsatisfied* to *Very satisfied*.

Job satisfaction. Job satisfaction measured the degree to which Soldiers were satisfied with their work and the opportunities they had on the job. For example, one item asked, *How satisfied are you with the amount of challenge in your work?* This outcome was measured with six items that were rated on a 5-point Likert scale with anchors ranging from *Very unsatisfied* to *Very satisfied*. Items came from the ALQ (Moriarty et al., 2009), and the scale had a reliability of .94.

Morale. Morale was measured with two items that asked, *How would you rate the level of morale in your unit?* and *How would you rate your current level of morale?* These items correlated .70 with each other. Items were rated on a 5-point Likert scale with anchors ranging from *Very low* to *Very high amount*. Items came from a previous ARI Command Climate Survey (unpublished).

Mutual trust. Mutual trust was measured with two items that asked, *How much do members of your unit trust each other?* and *How much do members of your unit count on the other members of the unit?* The items in this scale came from Langfred's (2004) measure, which was based on Simmons and Peterson's (2000) mutual trust scale. The two mutual trust items correlated .72 with each other. While related conceptually to cohesion, mutual trust focuses specifically on trust within the unit, while cohesion captures aspects such as unit pride, cooperation, and enthusiasm. Items were rated on a 5-point Likert scale with anchors ranging from *Not at all* to *To a great extent*. In addition, an *NA/Don't know* option was provided.

Participation in unit social events. Participation in unit social events measured how often Soldiers engaged in formal and informal social behaviors with the unit. For example, one item asked Soldiers, *How often do you check your Company/Troop/Battery Facebook page?* This outcome was measured with three items that were rated on a 5-point Likert scale with anchors ranging from *Never* to *Always*. In addition, an *NA/Don't know* option was provided. Items for this scale were developed for the current study. The scale had a reliability of .70.

Physical fitness. Physical fitness was measured with the Soldier's last Army Physical Fitness Test (APFT). There are three components to the APFT: push-ups, sit-ups, and a 2-mile run. The number of push-ups or sit-ups and the run time are converted to a score for each component, where the scoring differs based on gender and age. In addition, very high and very low scores for push-ups, sit-ups, and run times are grouped together, artificially reducing the variance at each extreme of the variable. This scoring technique could potentially serve to limit correlations of APFT with the study predictors. Soldiers were asked to report the last score they received. (Moriarty et al., 2009).

Stress. This scale measured the amount of stress a Soldier was experiencing at work and at home, and how much it was affecting him or her on the job. For example, one item asked, *How much stress, if any, are you experiencing now in your Army job?* Stress was measured with four items that were rated on 5-point Likert scales with anchors ranging from *None* to *Very high amount* and *Strongly disagree* to *Strongly agree*. Two of the items were from a previous ARI Command Climate Survey (unpublished), and two were developed as part of a previous ARI working group. The scale had a reliability of .81.

Unit performance. Unit performance was measured with four different approaches: Soldier ratings, and, when available, officer ratings, battalion commander ratings of performance, and battalion commander ratings of readiness. In the first case, Soldiers rated their perceptions of how well the unit was performing in general, as well as in comparison to other units (Estrada et al., under review). For example, one item asked Soldiers to rate how much they agreed with the statement, *This unit performs better than most.* This outcome was measured with four items that were rated on a 5-point Likert scale with anchors ranging from *Strongly disagree* to *Strongly agree*. In addition, an *NA/Don't know* option was provided. The scale had a reliability of .81.

Officers present at the data collections completed the same survey as other participants. To examine the unit performance ratings provided by officers, we selected the cases with officer and Warrant Officer ranks and calculated the mean of their performance ratings for the unit. Of the 51 units in the company-level analyses, 42 contained officers or warrant officers who completed the survey.

Battalion commanders provided ratings of company performance and readiness on a separate survey. If battalion commanders were not available to participate, battalion executive officers or command sergeants major provided ratings. Both performance and readiness were measured with three items each, on 1-5 Likert-type scales with anchors ranging from *Very low* to *Very high* and *Strongly disagree* to *Strongly agree*, respectively. Items came from Estrada et al. (under review) and the reliabilities of these scales were .88 and .92, respectively.

Unit resilience. Unit resilience measured Soldiers' perceptions of how well the unit works through difficult situations and learns from mistakes. This outcome was measured with four items (developed by an ARI working group) that asked respondents to provide a rating of the platoon across three separate facets of resilience: adapt, recover, and grow. For example, one item asked Soldiers to rate how much they agreed with the statement, *My unit can effectively overcome tough work-related challenges.* Items were rated on a 5-point Likert scale with anchors ranging from *Strongly disagree* to *Strongly agree*. In addition, an *NA/Don't know* option was provided. The scale had a reliability of .91.

Weapons qualification. Weapons qualification was measured with one item that asked, *What was the last Weapon Qualification you received on your individual weapon?* (Moriarty et al., 2009) and respondents self-reported their responses. Three response options were available: *Marksman*, *Sharpshooter*, or *Expert*.

Well-Being. Well-being was assessed by administering three items from the Satisfaction With Life Scale (Diener, Emmons, Larsen, & Griffin, 1985). For example, one item asked Soldiers to rate how much they agreed with the statement, *I am satisfied with my life.* Items were rated on a 5-point Likert scale with anchors ranging from *Strongly disagree* to *Strongly agree*. The scale had a reliability of .88.

Withdrawal. Withdrawal measured how often Soldiers exerted less than normal effort, came to work late, or left work early (Estrada et al., under review). For example, one item was, *Went to sick call to avoid work?* This outcome was measured with six items that were rated on a

5-point Likert scale with anchors ranging from *Never* to *Always*. The scale had a reliability of .88.

Attention check questions. Five attention check items were included in the survey to assess whether participants were reading the questions carefully or rushing through the survey answering in a random fashion. These items consisted of a simple statement directing the respondent to mark a specific letter for that question. These items were dichotomized into correct and incorrect answers based on whether the correct letter was chosen. Participants needed to answer four out of five of these questions correctly for their data to be included in the analyses.

Analyses

Reliabilities were calculated for the multi-item scales and single-item assessments. Tests were then conducted to determine if analyses should be conducted at the individual or unit level. Level of agreement was tested using a within-group correlation, $r_{wg(j)}$ (James, 1982; James et al., 1984), as well as ICCs (Bliese, 2000; de Boer et al., 2004). Two types of ICCs were calculated: ICC(1), which compares variability within units to variability across units, providing a percent of variance explained by the unit, and ICC(2), which provides an estimate of the reliability of the unit mean. Constructs with sufficiently high agreement were aggregated to the unit level. Results evaluating the construct and criterion-related validity of each construct were then examined.

Results

Results are presented first for the 13 command climate scales and single items, followed by exploratory results for the command climate diagnostic items for 10 of the dimensions. For the full command climate scales, internal consistency reliabilities were calculated, and single item reliabilities were estimated. Then ICCs were calculated to assess group-level effects. Following this, results for the evaluation of criterion-related validity of each dimension are presented. In the diagnostic item section, group-level effects are examined with ICCs. Criterion-related validity evidence is provided in the form of overall model fit from the regression of outcomes on the sets of diagnostic items.

Scales and Single Items

Detailed item-level means, standard deviations, skewness, and kurtosis, including results for the single-item measures, are presented in Appendix B. In addition, the percentage of respondents who indicated that the item was not applicable or for which they did not know the answer is provided in the final column of the table in Appendix B. The percent *NA/Don't know* was below 5% in most cases; however, the leader support for family dimension had between 11.4% and 15.4% of respondents indicate *NA/Don't know*.

All scale-level statistics were calculated with the single item for the dimension separately from the scale items. Scale-level means and standard deviations for the 13 command climate dimensions with individual-level data are provided in Table 9. Descriptive statistics for the single items can be found in Table 10. Hazing and bullying are negatively worded relative to the other items, so means and other statistics for these dimensions should be interpreted accordingly. Distributions for both the scales and the single items were generally non-normal, with all but inclusion showing significant negative skew and about half showing significant kurtosis.

Negative skew is typical when using self-report data. While these distributions could potentially limit the correlations of these variables with outcomes, tests of normality conducted on the data once it was aggregated to the company level indicated that only the flow of information scale and peer support single items deviated significantly from normal.

Table 9

Command Climate Dimension Scales Descriptive Statistics

				Skewness		Kurtosis	
Climate Dimension	*N*	Mean	*SD*	Statistic	*SE*	Statistic	*SE*
Flow of information	1324	3.14	0.92	-0.40	0.07	-0.26	0.13
Autonomy	1318	3.05	0.99	-0.19	0.07	-0.69	0.14
Openness	1295	3.22	1.04	-0.29	0.07	-0.57	0.14
Fairness	1269	3.08	1.10	-0.19	0.07	-0.78	0.14
Peer support for unit members	1310	3.44	0.91	-0.63	0.07	0.27	0.14
Leader support for unit members	1304	3.47	0.98	-0.71	0.07	0.18	0.14
Leader support for family	1119	3.23	1.02	-0.38	0.07	-0.38	0.15
Performance orientation	1309	3.61	0.92	-0.61	0.07	0.19	0.14
Professionalism	1319	3.40	0.94	-0.44	0.07	-0.10	0.14
Respect for the individual	1315	3.16	1.06	-0.34	0.07	-0.58	0.14
Hazing	1288	1.81	0.92	1.10	0.07	0.92	0.14
Bullying	1254	2.49	1.05	0.53	0.07	-0.25	0.14
Inclusion	1215	3.29	0.83	-0.03	0.07	0.09	0.14

Note. SD = Standard deviation; SE = Standard error.

Table 10

Command Climate Single-Item Descriptive Statistics

Climate Dimension	Mean	*SD*	Correlation SI with scale	SI reliability
Flow of information	3.08	1.23	0.77	0.64
Autonomy	3.11	1.22	0.74	0.60
Openness	3.28	1.26	0.79	0.69
Fairness	3.06	1.17	0.65	0.48
Peer support for unit members	3.51	1.13	0.73	0.65
Leader support for unit members	3.43	1.05	0.80	0.70
Leader support for family	3.24	1.17	0.79	0.68
Performance orientation	3.93	1.05	0.74	0.60
Professionalism	3.52	1.08	0.73	0.58
Respect for the individual	3.11	1.25	0.87	0.80
Hazing	1.98	1.10	0.80	0.69
Bullying	2.14	1.11	0.39	0.18
Inclusion	3.49	1.06	0.54	0.36

Note. SI = Single-item measure; SD = Standard deviation.

Initial reliabilities were estimated using Cronbach's alpha. Reliabilities were high for most dimensions; a few dimensions, however, had items that were problematic for their respective scale. Because scale reliabilities impact the estimation of single item reliabilities, we addressed problematic items by removing them from the scale. Item text can be found in Appendix A.

- Two items in the fairness scale (Fair5 and Fair6) had negative item-total correlation. Cronbach's alpha improved from .77 to .88 when these items were excluded.
- One item in the peer support for unit member scale (PSUM3) had a negative item-total correlation. When removed, the scale reliability improved from .80 to .82.
- One item from the leader support for unit member scale (LSUM5) was removed, increasing the scale's Cronbach alpha from .91 to .92.
- One item from the performance orientation scale (PerfOr8) had a negative item-total correlation. Cronbach's alpha improved from .88 to .91 when the item was excluded.
- One item from the hazing scale (Hazing5) had a negative item-total correlation. Cronbach's alpha improved from .81 to .93 when the item was excluded.
- Two items in the bullying scale (Bully2 and Bully4) had negative item-total correlations. When removed, the scale reliability improved from .75 to .83.

After these adjustments were made to the scales, final reliabilities ranged from .80 to .95. Table 11 shows the final scale reliabilities and the number of final items for the 13 command climate dimension scales.

Table 11

Command Climate Dimension Scale Reliabilities

Climate Dimension scale	Cronbach's alpha	Number of items
Flow of information	0.92	7
Autonomy	0.92	6
Openness	0.91	5
Fairness	0.88	3
Peer support for unit members	0.82	5
Leader support for unit members	0.92	3
Leader support for family	0.92	5
Performance orientation	0.91	6
Professionalism	0.92	6
Respect for the individual	0.95	6
Hazing	0.93	3
Bullying	0.83	3
Inclusion	0.80	5

To estimate single item reliabilities, the correction for attenuation formula was applied (Nunnally & Bernstein, 1994, p. 257), but rearranged to solve for the reliability of the single item (Wanous & Hudy, 2001; Wanous et al., 1997). The correction for attenuation formula is used to estimate the true correlation between two variables after eliminating the influence of unreliability in both variables. The estimated true correlation ρ_{xy}, is equal to:

$$\rho_{xy} = \frac{r_{xy}}{\sqrt{r_{xx} * r_{yy}}}$$

where r_{xy} is the observed correlation between x and y, r_{xx} is the reliability of x, and r_{yy} is the reliability of y. The variables x and y can be used to represent the scale and single item, respectively. Because the single item and scale are designed as parallel measures, the true correlation between these should be 1.00. Substituting 1.00 for ρ_{xy} above, we see that the numerator and denominator must be equal to each other.

$$r_{xy} = \sqrt{r_{xx} * r_{yy}}$$

The equation was rearranged to solve for r_{yy} and to plug in obtained values for full scale reliability and the correlation between the full scale and the single items. The single item reliability formula is:

$$r_{yy} = \frac{{r_{xy}}^2}{r_{xx}}$$

where the variables are as defined previously. Correlations between full scales and single items are presented in Table 10, along with estimates of single-item reliabilities. Single-item reliabilities ranged from .18 for bullying to .80 for respect. The bullying dimension scale demonstrated a reliability of .83, but the scale and single item correlated only .39. The inclusion dimension scale showed the lowest observed reliability at .80, and the correlation between the inclusion scale and single item was only .54. The single-item reliability estimate for inclusion was only .36. The fairness dimension had a sufficient scale reliability of .88 and a correlation with the single item of .65; however, these scores resulted in a single-item reliability estimate of only .48.

Many of the dimensions showed moderately high correlations between single items and scales and adequate reliabilities for the scales; however, their single-item reliabilities did not reach traditional thresholds for adequacy. Examples are professionalism (.58), performance orientation (.60), autonomy (.60), information (.64), and peer support for unit members (.65).

Leader support for family (.68), openness (.69), and hazing (.69) all had single item reliability estimates that approached the .70 cutoff commonly used as a threshold for adequate reliability for a measure in the early stages of development (attributed to Nunnally, 1970). Only leader support for unit members (.70) and respect for the individual (.80) reached the threshold. In both cases, correlations between scales and single items were .80 or above and scale reliabilities were .92 or above.

Company-Level Effects for the Full Scales and Single Items

Because command climate is conceptualized as a unit-level construct, individual-level data were aggregated to company-level scores and validity analyses were conducted using company-level scores. Intraclass correlations (ICC(1) and ICC(2); Bliese, 2000) were examined to verify the appropriateness of the group-level measurement. ICC(1) compares variability within units to variability across units, providing a percent of variance explained by the unit. According to LeBreton and Senter (2008), ICC(1) shows evidence that group-level effects are present if significance values are .05 or higher. ICC(1) should be statistically significant to ensure that group-level data is appropriate (Bliese, 2000), and ICC(2) should reach a cutoff of .80 to ensure that the group-level construct is reliable (van Mierlo, Vermunt, & Rutte, 2009).

The formulae used for ICC(1) and ICC(2) were:

$$ICC(1) = \frac{MSc - MSw}{MSc + (K-1)MSw}$$

$$ICC(2) = \frac{MSc - MSw}{MSc}$$

where *MSc* is the mean square between companies and *MSw* is the mean square within companies in a one-way, random effects ANOVA. *K* in the ICC(1) formula is the average number of respondents per company, which in this case was 26. Table 12 shows the ICCs for responses to each dimension's scale and single–item measure.

Table 12

Intraclass Correlations for Full Scales and Single Items

Climate Dimension		ICC(1)	ICC(2)
Flow of information	Scale	0.08	0.69
	Single item	0.08	0.70
Autonomy	Scale	0.09	0.72
	Single item	0.07	0.66
Openness	Scale	0.13	0.78
	Single item	0.09	0.72
Fairness	Scale	0.06	0.61
	Single item	0.02	0.36
Peer support for unit members	Scale	0.08	0.69
	Single item	0.08	0.70
Leader support for unit members	Scale	0.11	0.76
	Single item	0.11	0.76
Leader support for families	Scale	0.11	0.73
	Single item	0.07	0.63
Performance orientation	Scale	0.11	0.77
	Single item	0.10	0.74
Professionalism	Scale	0.10	0.75
	Single item	0.06	0.63
Respect for the individual	Scale	0.11	0.76
	Single item	0.07	0.67
Hazing	Scale	0.05	0.56
	Single item	0.06	0.60
Bullying	Scale	0.05	0.57
	Single item	0.04	0.50
Inclusion	Scale	0.06	0.62
	Single item	0.07	0.66

Note. All ICC(1) values significant at the $p < .05$ level.

All of the ICC(1) statistics were statistically significant at the $p < .05$ level. Most of the ICC(1)s reached the .05 threshold identified by LeBreton and Senter (2008) as indicative of a group-level effect. Only the single item measures for fairness and bullying did not reach this level. The full scales for openness, leader support for unit members, performance orientation, professionalism and respect for the individual showed a medium group effect (.10 or greater; see

Murphy & Myors, 1998, p. 47), as did single item measures for leader support for unit members and performance orientation.

The initial threshold established for ICC(2) was .80 based on best practices reported within the literature. None of the observed ICC(2)s reached this level. While most ICC(2) statistics reached the .60 level, there were a few that were much lower. For example, the single item measure for fairness had the lowest ICC(2) value at .36. This suggests that the mean value for fairness, as measured by the single item, is not a reliable index of company standing on the fairness construct and that individual responses within the company differ substantially from person to person and should be interpreted cautiously. The group-level effect on single-item fairness was small (ICC(1) = .02), suggesting that there was disagreement within companies on this item. The hazing and bullying scales and single items showed small to medium group-level effects based on ICC(1), but ICC(2) was low, in the .50 to .60 range.

Because this research is in the early stages of development for scale and single item measures and because group-level effects were detected on all ICC(1)s, the ICC(2) threshold for aggregation was lowered to .65. This is slightly below the .70 to .85 range LeBreton and Senter (2008) cite as adequate for early stages of scale development; however, this seemed appropriate due to the likely impact of other, unmeasured nested group effects and because of the preliminary nature of the single-item measures. Table 13 lists the constructs that were subsequently identified as demonstrating adequate ICCs for aggregation.

Table 13

Command Climate Dimensions with Adequate ICCs for Aggregation

Climate Dimension	ICC(1) = .05+ for full scale	ICC(1) = .05+ for single item	ICC(2) for full scale = .65+	ICC(2) for single item = .65+
Flow of information	☑	☑	☑	☑
Openness	☑	☑	☑	☑
Peer support for unit members	☑	☑	☑	☑
Leader support for unit members	☑	☑	☑	☑
Performance orientation	☑	☑	☑	☑
Autonomy	☑	☑	☑	☑
Respect for the individual	☑	☑	☑	☑
Leader support for family	☑	☑	☑	X
Professionalism	☑	☑	☑	X
Inclusion	☑	☑	*	☑
Hazing	☑	☑	*	X
Fairness	☑	X	X	X
Bullying	☑	X	X	X

Note. ICC = Intraclass correlation coefficient; ☑ = ICC reached threshold, X = ICC did not reach threshold, * = measure had high agreement ($r_{wg(j)}$).

As LeBreton and Senter (2008) discuss, ICC(2) might be low because of lower inter-rater consistency, low inter-rater agreement, or both. Furthermore, scores can have low agreement and high consistency, or high agreement and low consistency (LeBreton, Burgess, Kaiser, Atchley, & James, 2003). Level of agreement can also be tested using a within-group correlation, $r_{wg(j)}$ (James, 1982; James et al., 1984). Therefore, for the measures that did not reach .65 on ICC(2), we examined interrater agreement with r_{wg} and $r_{wg(j)}$ statistics. The $r_{wg(j)}$ provides a measure of the interchangeability of raters by comparing observed variance in ratings to a hypothetical expected variance that should be obtained if the raters completely disagreed; it provides an estimate of the interrater agreement of a group. The observed variance in ratings was compared to a uniform null distribution and a slightly skewed distribution reflecting random responding and leniency/strictness, respectively. A triangular distribution reflecting central tendency in responding might also be appropriate, but because expected variances for slightly skewed distributions and triangular distributions are similar (1.34 for slightly skewed, 1.32 for triangular; see LeBreton & Senter, 2008), r_{wg}s were only calculated for the slightly skewed distribution. These agreement statistics are provided in Table 14.

Table 14

Average Interrater Agreement Statistics

	Avg $r_{wg(j)}$ for scale (random null)	Avg r_{wg} for single item (random null)	Avg $r_{wg(j)}$ for scale (slightly skewed null)	Avg r_{wg} for single item (slightly skewed null)
Leader support for family	.68	.37	.37	.15
Professionalism	.79	.46	.46	.23
Inclusion	.75	.48	.38	.26
Hazing	.76	.44	.53	.23
Fairness	.63	.33	.34	.10
Bullying	.48	.41	.14	.19

Using the commonly applied cut-off of .70, three out of six of the tested dimensions reached the threshold. Agreement was moderately high on scales for a few dimensions, including professionalism (.79), hazing (.76), and inclusion (.75), but scales for leader support for family, fairness, and bullying had low agreement (.68, .63, and .48, respectively). Professionalism and hazing had moderately high agreement, but low inter-rater consistency on full scale measures, suggesting that scores on these climate dimensions were range restricted and differences between groups were truncated (see LeBreton & Senter, 2008).

Of the dimensions low on ICC(2), single-item measures had average r_{wg} ranging from .33 to .48 suggesting there was considerable disagreement within companies. For leader support for family, fairness, and bullying, the low single-item agreement confirmed low scale agreement. For these dimensions, aggregation to the company level may not be appropriate until the scale and single-item measures are refined. For leader support for family, professionalism, and hazing, full scales had sufficient agreement and consistency, but single items did not. For these dimensions, relationships with the criterion will be examined for the full scales but not for the single-item measures. Fairness and bullying had low agreement and inter-rater consistency for both scales and single items. For these dimensions, aggregation to the company level may not be appropriate; unit members do not have shared perceptions of these aspects of the climate and means calculated on the data lack stability.

Company-level scale descriptive statistics are presented in Table 15 and company-level single-item descriptive statistics are presented in Table 16. Intercorrelations among scale dimensions are presented in Appendix B, Table B-2. A parallel table for the single items is presented in Appendix B, Table B-3.

Table 15

Command Climate Scale Descriptive Statistics at Company-Level

				Skewness		Kurtosis	
	N	Mean	*SD*	Statistic	*SE*	Statistic	*SE*
Information	51	3.16	0.32	0.27	0.33	-0.76	0.66
Autonomy	51	3.07	0.36	0.02	0.33	-0.73	0.66
Openness	51	3.25	0.43	0.18	0.33	-0.65	0.66
Fairness	51	3.10	0.36	0.21	0.33	-0.37	0.66
Peer support	51	3.47	0.32	0.36	0.33	-1.00	0.66
Leader support for unit members	51	3.49	0.38	-0.09	0.33	-0.76	0.66
Leader support for family	51	3.26	0.41	0.17	0.33	-0.85	0.66
Performance orientation	51	3.63	0.37	0.16	0.33	-1.00	0.66
Professionalism	51	3.43	0.36	0.17	0.33	-0.88	0.66
Respect for the individual	51	3.21	0.41	0.11	0.33	0.26	0.66
Hazing	51	1.79	0.27	-0.16	0.33	0.11	0.66
Bullying	51	2.48	0.33	0.32	0.33	-0.42	0.66
Inclusion	51	3.32	0.28	0.23	0.33	-0.73	0.66

Note. SD = Standard deviation; SE = Standard Error.

Table 16

Command Climate Single-Item Descriptive Statistics at Company Level

				Skewness		Kurtosis	
	N	Mean	*SD*	Statistic	*SE*	Statistic	*SE*
Information	51	3.11	0.42	0.21	0.33	-0.61	0.66
Autonomy	51	3.14	0.41	-0.08	0.33	-0.98	0.66
Openness	51	3.09	0.30	-0.14	0.33	-0.72	0.66
Fairness	51	3.30	0.49	0.01	0.33	-0.87	0.66
Peer support	51	3.54	0.41	0.26	0.33	-1.03	0.66
Leader support for unit members	51	3.44	0.41	0.14	0.33	-0.36	0.66
Leader support for family	51	3.26	0.41	-0.09	0.33	-0.71	0.66
Performance orientation	51	3.95	0.40	0.31	0.33	-0.71	0.66
Professionalism	51	3.55	0.36	-0.20	0.33	-0.68	0.66
Respect for the individual	51	3.16	0.42	0.00	0.33	0.17	0.66
Hazing	51	1.96	0.34	0.27	0.33	-1.10	0.66
Bullying	51	2.11	0.31	0.18	0.33	-0.93	0.66
Inclusion	51	3.51	0.36	-0.04	0.33	-0.83	0.66

Note. SD = Standard deviation; SE = Standard Error.

Construct Validity of Command Climate Scales

In order to establish the validity of the command climate assessments, evidence was examined to determine whether the dimensions can be distinguished from each other. Intercorrelations among the climate dimensions provide some evidence that the dimensions can be distinguished (see Appendix B, Table B-2). The average intercorrelation in the individual-level data is .53 in magnitude, and the correlations ranged in strength from .13 to .74. Taking into consideration the direction of correlations, the strongest negative correlation in the individual data was -.55 and the strongest positive correlation was .74. Correlations were much larger in the company-level data. The average intercorrelation, ignoring direction, was .71 and intercorrelations ranged in strength from .30 to .93. The strongest negative intercorrelation was -.80 and the strongest positive intercorrelation was .93.

Next, exploratory factor analyses (EFA) were conducted to examine the relationships among dimensions and identify any higher order structure that may exist in the data. The full scale was used for this rather than the single items. Unfortunately, the sample size of 51

companies was not sufficient to conduct an EFA on the company-level data. As a result, analyses were conducted at the individual level using principal axis factoring with oblique promax rotation.

In the individual-level data, a scree plot of the eigenvalues showed one large factor that accounted for most of the variance. However, examination of the eigenvalues indicated that 10 factors had eigenvalues greater than one, suggesting 10 underlying factors in the data. Taken together, this shows overlap among the factors but some distinction. Factor loadings are presented in Appendix B, Table B-4. The scale items for flow of information, autonomy, leader support for families, hazing, and bullying each loaded on their own factors (Factors 3, 4, 5, 8, and 9, respectively). The other five factors were comprised of items from more than one scale. Factor 1 had high loadings from performance orientation and moderate loadings from professionalism. Factor 2 had high loadings from respect for the individual, low loadings from fairness, two positively-worded inclusion items, and five of the professionalism items (which also loaded on the first factor). Leader openness items loaded on Factor 6, which also had weak loadings from two fairness items and two leader support for unit member items. The peer support for unit member items loaded on Factor 7, which also had weak loadings from two of the leader support for unit member items. The three negatively worded inclusion items loaded alone on Factor 10, but the two positively-worded items loaded on Factor 2 with respect for the individual items. No items failed to load on a factor. Communalities after extraction averaged .68 across all items. The communalities provide the proportion of the variable's variance that is explained by the factors. They ranged from .42 to .86. These are also shown in Appendix B, Table B-4.

The results of the EFA demonstrate distinct factors for flow of information, autonomy, leader support for families, bullying, and hazing, with little cross-loading of items from other scales. The negatively worded inclusion items also demonstrated distinction from other factors. This suggests the distinctiveness of these scales. The other four factors consisted of items that cross-loaded from different scales, suggesting some level of overlap that should be investigated further. While the 10 factors were intercorrelated, the average correlation between factors was .50, and ranged from .15 between Factor 3 (information) and Factor 10 (inclusion) to .79 between Factor 2 (respect for the individual) and Factor 6 (leader openness). A table of factor correlations is presented in Appendix B, Table B-5.

Criterion-Related Validity of Command Climate Dimensions

Information, openness, peer support for unit members, leader support for unit members, performance orientation, autonomy, respect for the individual, and inclusion all had good ICC(1) and sufficient ICC(2) levels under the lowered threshold. This was the case for full scales and single-item measures. To evaluate criterion-related validity, two simple regressions and one hierarchical regression were conducted for each examined relationship. In one simple regression, the outcome was regressed on the full-scale predictor; in the other, the outcome was regressed on the single-item predictor. This provided estimates of regression weights without the influence of collinearity between the full-scale and single-item predictors that would occur in a multiple regression. In the hierarchical regression, the single item was entered first and the incremental R^2 change from adding the full scale was examined. This analysis provided an indication of whether the full scale accounts for additional variance in the outcomes, above and beyond the single item.

Leader support for unit members, professionalism, and hazing reached the ICC thresholds for the full scales but not the single items. Regression analyses for these dimensions were conducted only for the full scales. Fairness and bullying did not reach the ICC(2) threshold and were not aggregated to the group level. The regressions for these were conducted at the individual level.

Because multiple inferences are being drawn from the same data, there is an elevated risk of finding significant results just by chance. A common approach to combating this risk (known as alpha inflation), is to use a correction such as the Bonferroni correction, which divides the critical alpha level by the number of hypotheses. However, in many circumstances, this approach leads to exceedingly conservative correction. In order to minimize family-wise alpha inflation without overly inflating beta, the critical alpha level was set to be .05 divided by the number of dependent variables for each climate dimension. The alpha levels used for each dimension are listed in Tables 17 to 23.

Flow of information. The flow of information dimension of command climate was examined as a predictor of morale, unit performance as rated by unit members, unit performance ratings from officers, and battalion commander ratings of unit performance and readiness (see Table 17).

In predicting morale, both the single-item and scale scores for flow of information were significant predictors. Standardized regression weights were .70 and .76 for single item and scale, respectively. In a hierarchical regression, the single item had a significant adjusted R^2 of .47 ($F(1, 49) = 45.73$, $p < .01$) but the scale predicted an additional 10% of morale $F(1, 48) = 11.07$, $p < .01$).

For the prediction of unit performance, the single item and scale were both significant. The regression weight for the single item was .69 and the regression weight for the scale was .69. The single item predicted 45.8% (adjusted $R^2 = .46$, $F(1, 49) = 43.30$, $p < .01$) of the variance in unit performance and the scale did not contribute additional prediction (R^2 change = .02, $F(1, 48) = 2.7$, *ns*).

For the remaining outcome measures, the single item and scale for the flow of information dimension were not significant predictors. Adjusted R^2 for scale and single item predicting unit performance (officer ratings) was = .00 ($F(1, 39) = 1.26$, *ns*). Adjusted R^2 for scale and single item predicting commander ratings of unit performance was = .07 ($F(1, 28) = 2.10$ *ns*). Adjusted R^2 for scale and single item predicting commander ratings of unit readiness was = .00 ($F(1, 20) = .14$, *ns*).

Table 17

Flow of Information Simple and Hierarchical Regression Results

	Morale	Unit performance	Unit performance (officer rating)	Unit performance (CDR rating)	Unit readiness (CDR rating)
Information SI β-weight[1]	.70*	.69*	-.08	-.24	.10
Information scale β-weight	.76*	.69*	-.15	-.33	.12
SI Model adjusted R^2	.47*	.46*	.00	.03	.00
SI + Scale Model R^2 change	.10*	.02	.03	.07	.00

Note. * $p < .01$, [1]All β-weights are standardized.

Leader openness. The dimension of leader openness was used to predict affective commitment and continuance commitment. Standardized regression weights were .69 and .70 for single item and scale, respectively. In a hierarchical regression, the single item had a significant adjusted R^2 of .46 ($F(1, 49) = 43.33$, $p < .025$), but the scale only contributed an additional 3% of the variance in affective commitment ($F(1, 48) = 3.07$, *ns*).

Neither the single item nor the scale for the leader openness dimension was a significant predictor of continuance commitment. Adjusted R^2 for scale and single item was .01 ($F(1, 48) = 1.21$, *ns*). Results for leader openness are presented in Table 18.

Table 18

Leader Openness Simple and Hierarchical Regression Results

	Affective commitment	Continuance commitment
Leader openness SI β-weight[1]	.69*	.18
Leader openness scale β-weight	.70*	.21
SI Model adjusted R^2	.46*	.01
SI + Scale Model R^2 change	.03	.02

Note. * $p < .025$, [1] All β-weights are standardized.

Peer support for unit members. Peer support for unit members was a good predictor of six outcome variables: well-being, morale, affective commitment, stress, cohesion, and unit

resilience. Though both the single item and scale had significant relationships with the outcomes (with the exception of the scale predicting stress), in each case, the single item had a larger regression coefficient than the full scale. Results for peer support for unit members are presented in Table 19.

The single item for peer support for unit members predicted more than half of the variance for morale, cohesion, and unit resilience (adjusted R^2 = .61, .78, and .67, respectively). The *F* ratios for these relationships were 78.9, 175.40, and 100.3, respectively (all significant at the $p < .008$ level), suggesting very strong relationships between these variables and a climate of peer support.

For well-being, the single item for peer support for unit members had an adjusted R^2 = .30, $F(1, 49) = 22.20$, $p < .008$). The full scale did not contribute any incremental prediction (R^2 change = .00, $F(1, 48) = .03$, *ns*). For affective commitment, the single item for peer support for unit members had an adjusted R^2 = .43, $F(1, 49) = 38.06$, $p < .008$). The full scale predicted very little additional variance (R^2 change = .01, $F(1, 48) = .75$, *ns*). For stress, the single item for peer support for unit members had an adjusted R^2 = .12, $F(1, 49) = 7.66$, $p < .008$). The full scale did not contribute any incremental prediction (R^2 change = .00, $F(1, 48) = .06$, *ns*). Unlike the other peer support for unit members relationships, the relationship with stress was negative (β = -.37).

Table 19

Peer Support for Unit Members Simple and Hierarchical Regression Results

	Well-being	Morale	Affective commitment	Stress	Cohesion	Unit resilience
Peer support for unit members SI β-weight[1]	.56*	.79*	.66*	-.37*	.88*	.82*
Peer support for unit members scale β-weight	.52*	.77*	.64*	-.32	.87*	.80*
SI Model adjusted R^2	.30*	.61*	.43*	.12*	.78*	.67*
SI + Scale Model R^2 change	.00	.02	.01	.00	.02	.01

Note. * $p < .008$, [1] All β-weights are standardized.

Leader support for unit members. Leader support for unit members was also a strong predictor of the examined outcomes (affective commitment, well-being, and unit resilience). Scales and single items had regression weights with similar magnitudes, though for unit resilience, the full scale appears to have a stronger relationship. Regression weights and R^2 values are presented in Table 20.

In the hierarchical regression, the single item predicted significant variance of affective commitment (adjusted R^2 = .46, $F(1, 49) = 42.75$, $p < .016$), but the full scale did not add incremental prediction (R^2 change = .01, $F(1, 48) = 1.21$, *ns*). With well-being as the dependent variable, the single item measure predicted 23% of the variance (adjusted R^2 = .23, $F(1, 49) = 15.60$, $p < .016$). The full scale did not contribute additional prediction above and beyond the single item (R^2 change = .01, $F(1, 48) = .69$, *ns*). The single item measure of leader support for unit members accounted for 72% of the variance in unit resilience (adjusted R^2 = .72, $F(1, 49) = 122.91$, $p < .016$). The full scale contributed an additional 9% of variance (R^2 change = .09, $F(1, 48) = 21.07$, $p < .016$).

Table 20

Leader Support for Unit Members Simple and Hierarchical Regression Results

	Affective commitment	Well-Being	Unit resilience
Leader support for unit members SI β-weight[1]	.68*	.49*	.85*
Leader support for unit members scale β-weight	.68*	.50*	.90*
SI Model adjusted R^2	.46*	.23*	.72*
SI + Scale Model R^2 change	.01	.01	.07*

Note. * $p < .016$, [1] All β-weights are standardized.

Performance orientation. Performance orientation was examined as a predictor of numerous outcomes, as shown in Table 21. However, only the relationship between performance orientation and unit performance was significant. Both the single item and scale had strong relationships with unit performance (β = .75 and β = .81, respectively). The single item accounted for 55% of the variance in unit performance (adjusted R^2 = .55, $F(1, 49) = 122.91$, $p < .007$). Adding the full scale to the regression accounted for 10% more variance (R^2 change = .10, $F(1, 48) = 13.80$, $p < .007$).

Table 21

Performance Orientation Simple and Hierarchical Regression Results

	Unit performance	APFT scores	Weapon qualification	Article 15	Unit performance (officer rating)	Unit performance (CDR rating)	Unit readiness (CDR rating)
Performance orientation SI β-weight[1]	.75*	.07	.18	-.23	-.19	-.16	.35
Performance orientation scale β-weight	.81*	.15	.18	-.32	-.21	-.29	.18
SI Model adjusted R^2	.55*	.00	.01	.03	.01	.00	.08
SI + Scale Model R^2 change	.10*	.03	.00	.07	.01	.09	.06

Note. * $p < .007$, [1] All β-weights are standardized.

Autonomy. The autonomy dimension of command climate was examined as a predictor of job satisfaction, unit performance (Soldiers' ratings, officer ratings, and battalion commander ratings), and unit readiness (battalion commander ratings). Autonomy climate was only a predictor of job satisfaction and unit performance. See Table 22.

As a predictor of job satisfaction, autonomy climate showed a significant positive relationship (adjusted R^2 = .33, $F(1, 49) = 25.49$, $p < .01$). Adding the scale to the regression did not significantly increase the predicted variance in job satisfaction (R^2 change = .00, $F(1, 48) = .28$, *ns*). The β weight for the single item was .59, while the β weight for the scale was .53.

As a predictor of unit performance, autonomy showed a strong positive relationship (adjusted R^2 = .56, $F(1, 49) = 63.69$, $p < .01$). The full scale did not contribute incremental prediction above and beyond the single item (R^2 change = .01, $F(1, 48) = .54$, *ns*). The β weight for the single item was .75, while the β weight for the scale was .68.

The unit performance scores as rated by officers and battalion commanders were not predicted by autonomy. The adjusted R^2 was .07 ($F(1, 39) = 2.54$, *ns*) for officer rated unit performance and .00 ($F(1, 28) = .43$, *ns*) for battalion commander ratings of unit performance. Battalion commander ratings of unit readiness were also not related to autonomy (adjusted R^2 = .00, $F(1, 20) = .20$, *ns*).

Table 22

Autonomy Simple and Hierarchical Regression Results

	Job satisfaction	Unit performance	Unit performance (officer)	Unit performance (CDR rating)	Unit readiness (CDR rating)
Autonomy SI β-weight[1]	.59*	.75*	-.24	-.17	.14
Autonomy scale β-weight	.53*	.68*	-.08	-.15	.13
SI Model adjusted R^2	.33*	.56*	.03	.00	.00
SI + Scale Model R^2 change	.00	.01	.06	.00	.00

Note. * $p < .01$, [1] All β-weights are standardized.

Respect for the individual. Respect for the individual was significantly related to mutual trust (adjusted R^2 = .60, $F(1, 49) = 76.90$, $p < .025$) and cohesion (adjusted R^2 = .76, $F(1, 49) = 161.50$, $p < .025$). In both cases, the full scale predicted incremental variance above and beyond

the single item. The R^2 change from the first step to the second in the hierarchical regression of mutual trust on respect for the individual was .05 ($F(1, 48) = 7.10$, $p < .025$). The R^2 change from the first step to the second in the hierarchical regression of cohesion on respect for the individual was .08 ($F(1, 48) = 24.56$, $p < .025$). Results for respect for the individual climate are provided in Table 23.

Table 23

Respect for the Individual Simple and Hierarchical Regression Results

	Mutual trust	Cohesion
Respect for the individual SI β-weight[1]	.78*	.88*
Respect for the individual scale β-weight	.81*	.92*
SI Model adjusted R^2	.60*	.76*
SI + Scale Model R^2 change	.05*	.08*

Note. * $p < .025$ [1] All β-weights are standardized.

Inclusion. Results for the inclusion dimension of command climate are provided in Table 24. Inclusion was a significant predictor of affective commitment (adjusted $R^2 = .37$, $F(1, 49) = 30.31$, $p < .016$). The full scale did not predict incremental variance beyond the single item (R^2 change = .04, $F(1, 48) = 3.40$, *ns*).

Neither withdrawal (adjusted $R^2 = .06$, $F(1, 48) = 2.67$, *ns*) nor participation in social events (adjusted $R^2 = .09$, $F(1, 48) = 3.40$, *ns*) was significantly predicted by the inclusion dimension in a hierarchical regression.

Table 24

Inclusion Simple and Hierarchical Regression Results

	Affective commitment	Withdrawal	Participation in social events
Inclusion SI β-weight[1]	.62*	-.31	.33
Inclusion scale β-weight	.61*	-.28	.33
SI Model adjusted R^2	.37*	.08	.09
SI + Scale Model R^2 change	.04	.00	.01

Note. * $p < .016$, [1] All β-weights are standardized.

Leader support for family, professionalism, and hazing. Leader support for family, Professionalism, and Hazing all showed sufficient ICC(2) for the full scale but not for the single item measures. For these dimensions, it makes sense to examine criterion-related validity evidence for the scales only. Leader support for family was not significantly related to family's desire to stay (adjusted R^2 = .05, F(1, 49) = 3.42, *ns*). However, this scale was significantly related to family satisfaction (adjusted R^2 = .36, F(1, 49) = 29.60, p < .025). The standardized simple regression weight of leader support for family was β = .61.

Professionalism did not predict weapon qualification (adjusted R^2 = .00, F(1, 49) = .50, *ns*), APFT Scores (adjusted R^2 = .00, F(1, 49) = .40, *ns*), or number of Article 15s (adjusted R^2 = .06, F(1, 49) = 4.15, *ns*).

A hazing climate was significantly related to affective commitment, morale, and stress. Both affective commitment and morale decreased with hazing. Hazing predicted 23% of affective commitment (adjusted R^2 = .23, F(1, 49) = 15.97, p < .0125) and 28% of morale (adjusted R^2 = .28, F(1, 49) = 20.79, p < .0125). Stress increased with hazing (adjusted R^2 = .13, F(1, 49) = 8.40, p < .0125). The standardized regression weights were -.50, -.55, and .38 for affective commitment, morale, and stress, respectively.

Fairness and bullying. Fairness and bullying were not aggregated to the group level because of low ICCs and lack of agreement within companies. Instead, individual-level regressions of the relevant outcomes on these climate dimensions were conducted. Simple and hierarchical regressions were conducted.

Fairness. For the fairness dimension of command climate, regressions at the individual level showed that the full scale was a predictor of each outcome, but the single item did not predict withdrawal or continuance commitment. The largest beta weight was .55 in the regression of morale on fairness. Each time the single item predicted significant variance in an outcome, the full scale predicted additional variance. In the cases of affective commitment and career intentions, the additional variance predicted by the full scale was small (6% and 3%, respectively). In the case of morale, the addition of the full scale predicted 12% more variance above the single item. Finally, when the outcome was continuance commitment, the full scale predicted significant variance while the single item did not. However, the change in variance accounted for was less than 1% when the full scale was entered above the single item. These results are shown in Table 25.

Table 25

Fairness Simple and Hierarchical Regression Results

	Withdrawal	Affective commitment	Continuance commitment	Career intentions	Morale
Fairness SI β-weight[1]	-.07	.28*	.04	.16*	.44*
Fairness scale β-weight	-.08*	.37*	.09*	.23*	.55*
SI Model adjusted R^2	.00	.08*	.00	.02*	.20*
SI + Scale Model R^2 change	.00	.06*	.01*	.03*	.12*

Note. * $p < .01$, [1] All β-weights are standardized.

Bullying. In the individual-level data, the bullying command climate full scale was a significant predictor of each outcome, and the single item was a significant predictor of all outcomes except continuance commitment (see Table 26). In the case of each outcome, the change in variance accounted for when the full scale was entered was significant and these values ranged from less than 1% (continuance commitment) to about 12% (morale).

Table 26

Bullying Simple and Hierarchical Regression Results

	Affective commitment	Continuance commitment	Morale	Stress	Participation in social events
Bullying SI β-weight[1]	-.30*	-.02	-.25*	.22*	-.16*
Bullying scale β-weight	-.41*	-.09*	-.49*	.23*	-.25*
SI Model adjusted R^2	.09*	.00	.06*	.05*	.02*
SI + Scale Model R^2 change	.11*	.01*	.12*	.02*	.04*

Note. * $p < .01$, [1] All β-weights are standardized.

Command Climate Diagnostic Items

Item level means and standard deviations for all diagnostic items are provided in Appendix D, as well as ICCs for the diagnostic items. For the diagnostic items, we were not

relying on ICCs to determine the appropriateness of aggregation because we were not expecting consensus in responses from all members of the unit. However, we did examine ICCs to identify whether the diagnostics scores were influenced by group-level structure in the data. While the ICC(1) for each command climate diagnostic item was statistically significant, there were many that were small in magnitude. This suggests that although a company-level effect may exist, the effect size may be too small to be practically significant. Sixteen items had an ICC(1) that was below the .05 threshold identified by LeBreton and Senter (2008) as indicative of a group effect. These items were:

- Information2, Information5,
- Autonomy4, Autonomy8,
- Openness6,
- Fairness4, Fairness5,
- LSFam1, LSFam4, LSFam5, LSFam6
- PerfOr3, PerfOr6
- Bullying1, Bullying4, Bullying5

ICC(2) for the diagnostic items was only above .70 for 10 of the items; the other 49 items had ICC(2) ranging from .29 to .69. High interrater agreement and consistency was not expected for the diagnostic items, however, so the low ICC(2) statistics do not present analysis challenges. The 10 items showing consistent group-level effects must be examined for company- level nesting effects. These items are:

- Autonomy1, Autonomy2, Autonomy3,
- Openness1, Openness2,
- Fairness1,
- LeaderSupport1, LeadersSupport3,
- PerfOr1
- Professionalism2

Company-level descriptive statistics are provided in Appendix D, Table D-3. Items were aggregated to the company level by taking the mean item score within each company. Mean item scores at the company level ranged from 1.89 (Bully1) to 3.79 (Hazing2).

Intercorrelations among company-level diagnostic items, and correlations of diagnostic items with scales and single-item scores at the company level, are provided in Appendix D, Tables D-4 to D-13. In general terms, diagnostic items correlated strongly with dimension scales and single items, but not as strongly as scales and single items correlated with each other.

Criterion-related validation of diagnostic items. Regressions were conducted to examine the omnibus relationship between a set of diagnostic items for each climate dimension and the outcome variable of interest. A significant R^2 value suggests that the set of diagnostic items as a whole contributed to prediction of the outcome measure. The regression weights for individual diagnostic items were not examined, because there was a high degree of multi-collinearity in the items, producing coefficient estimates that have large standard errors and wide confidence intervals.

Table 27 describes the adjusted R^2 value for each dependent variable associated with a dimension regressed on a set of diagnostic items. These analyses are corrected for family-wise alpha inflation using the same approach described previously for the full scales and single items.

Table 27

Dependent Variable Adjusted Multiple R^2 for Diagnostic Items

Diagnostic Items	Dependent variable	Adjusted R^2	F ratio	Significance level
Flow of information	Morale	.49	9.09	$p < .013$
	Unit performance	.57	11.88	$p < .013$
	Unit performance (officer rating)	.29	3.85	$p < .013$
	Unit performance (CDR rating)	.05	1.24	*ns*
	Unit readiness (CDR rating)	.00	.43	*ns*
Autonomy	Job satisfaction	.30	3.65	$p < .013$
	Unit performance	.62	11.36	$p < .013$
	Unit performance (officer rating)	.00	.47	*ns*
	Unit performance (CDR rating)	.68	8.97	$p < .013$
	Unit readiness (CDR rating)	.14	1.46	*ns*
Leader openness	Affective commitment	.45	7.76	$p < .025$
	Continuance commitment	.05	1.45	*ns*
Fairness	Withdrawal	.00	1.01	*ns*
	Affective commitment	.49	8.94	$p < .013$
	Continuance commitment	.04	1.34	*ns*
	Career intentions	.14	2.63	*ns*
	Morale	.63	15.13	$p < .013$
Leader support for unit members	Affective commitment	.47	6.61	$p < .017$
	Well being	.28	3.39	$p < .017$
	Unit resilience	.78	23.04	$p < .017$
Leader support for family	Family desire to stay	.20	3.10	$p < .025$
	Family satisfaction	.37	5.86	$p < .025$

Performance orientation	Unit performance	.63	15.08	$p < .008$
	APFT score	.00	.65	*ns*
	Weapon qualification	.04	1.39	*ns*
	Number of Article 15s	.05	1.43	*ns*
	Unit performance (officer rating)	.00	.98	*ns*
	Unit performance (CDR rating)	.02	1.08	*ns*
	Unit readiness (CDR rating)	.33	2.77	*ns*
Hazing	Affective commitment	.50	13.25	$p < .013$
	Continuance commitment	.12	2.68	*ns*
	Morale	.63	22.35	$p < .013$
	Stress	.12	2.63	*ns*
Bullying	Participation in social events	.09	1.98	*ns*
	Morale	.59	15.09	$p < .01$
	Stress	.03	1.29	*ns*
	Continuance commitment	.02	1.16	*ns*
	Affective commitment	.46	9.58	$p < .01$
Professionalism	Weapon qualification	.11	2.56	*ns*
	APFT score	.14	2.94	*ns*
	Number of Article 15s	.02	1.25	*ns*

Note. ns = not significant.

Regression results for the diagnostic items were largely consistent with those for the full scales and single items. Generally, if an outcome was found to be predicted by a command climate scale or single item, it was also predicted by the set of diagnostic items. Two exceptions existed where the scale and single item were not significant, but the set of diagnostic items was significant. This occurred for autonomy items predicting battalion commander ratings of unit performance and for leader support for family items predicting family desire to stay. There was one exception where the opposite happened; namely, the scale and single item showed significant relationships when hazing was used to predict stress, but this relationship was not significant when the set of diagnostic items were used as the predictor of stress. Because fairness and bullying were not aggregated to the group level for the scale and single item regressions, but were aggregated for the diagnostic items, the results for these analyses are difficult to compare. Table 28 shows a summary of the scale, single item, and diagnostics results side-by-side.

Table 28

Summary of Regression Results for Three Survey Formats

Command climate dimension	Dependent variable	Full scale	Single item	Diagnostic Items
Flow of information	Morale	☑	☑	☑
	Unit performance	☑	☑	☑
	Unit performance (officer)	☑	☑	☑
	Unit performance (CDR rating)	*ns*	*ns*	*ns*
	Unit readiness (CDR rating)	*ns*	*ns*	*ns*
Autonomy	Job satisfaction	☑	☑	☑
	Unit performance	☑	☑	☑
	Unit performance (officer)	*ns*	*ns*	*ns*
	Unit performance (CDR rating)	*ns*	*ns*	☑
	Unit readiness (CDR rating)	*ns*	*ns*	*ns*
Leader openness	Affective commitment	☑	☑	☑
	Continuance commitment	*ns*	*ns*	*ns*
Fairness	Withdrawal	☑	*ns*	*ns**
	Affective commitment	☑	☑	☑*
	Continuance commitment	☑	*ns*	*ns**
	Career intentions	☑	☑	*ns**
	Morale	☑	☑	☑*
Peer support for unit members	Well-being	☑	☑	NA
	Morale	☑	☑	NA
	Affective commitment	☑	☑	NA
	Stress	*ns*	☑	NA
	Cohesion	☑	☑	NA
	Unit resilience	☑	☑	NA
Leader support for unit members	Affective commitment	☑	☑	☑
	Well-being	☑	☑	☑
	Unit resilience	☑	☑	☑
Leader support for family	Family desire to stay	*ns*	NA	☑
	Family satisfaction	☑	NA	☑

Performance orientation	Unit performance	☑	☑	☑
	APFT score	*ns*	*ns*	*ns*
	Weapon qualification	*ns*	*ns*	*ns*
	Number of Article 15s	*ns*	*ns*	*ns*
	Unit performance (officer)	*ns*	*ns*	
	Unit performance (CDR rating)	*ns*	*ns*	*ns*
	Unit readiness (CDR rating)	*ns*	*ns*	*ns*
Hazing	Affective commitment	☑	NA	☑
	Continuance commitment	*ns*	NA	*ns*
	Morale	☑	NA	☑
	Stress	☑	NA	*ns*
Bullying	Participation in social events	☑	☑	*ns**
	Morale	☑	☑	☑*
	Stress	☑	☑	*ns**
	Continuance commitment	☑	*ns*	*ns**
	Affective commitment	☑	☑	☑*
Professionalism	Weapon qualification	*ns*	NA	*ns*
	APFT score	*ns*	NA	*ns*
	Number of Article 15s	*ns*	NA	*ns*
Inclusion	Affective commitment	☑	☑	NA
	Withdrawal	*ns*	*ns*	NA
	Participation in social events	*ns*	*ns*	NA
Respect for the individual	Mutual trust	☑	☑	NA
	Cohesion	☑	☑	NA

Note. ☑ = a significant relationship was found, *ns* = relationship not significant, NA = analyses were not conducted, * = conducted at individual levels of analysis.

Practical Application Testing

To examine the practical effectiveness of the single-item measures, we conducted a supplemental analysis to compare agreement between the scale and the single item when using a threshold to identify instances of poor command climate. An arbitrary threshold of 3.0 was set for the command climate scores, such that 3.0 and above was considered a good climate and below 3.0 was considered to be a climate with a potential problem (except for the hazing and bullying dimensions, where 3.0 and below was considered a good climate). We dichotomized each scale and single-item mean for each company to be: (a) 3.0 and above or (b) below 3.0. Making the assumption that the scale score equated to the true score, we then coded the data to

represent whether the single item was either accurate (both scale and single item showed a problem or neither showed a problem) or not accurate (the scale showed a problem but the single item did not, or the scale did not show a problem but the single item did). Table 29 shows the agreement across the 13 dimensions and 51 companies.

Because this table is sorted by total dimensions correct by company (smallest to largest) and total companies correct by dimension (smallest to largest), the preponderance of 0s is in the top left corner of the table. In total, the single-item command climate survey matched the multi-item scale results for 90% of the cases (597 out of 663). Matches within the dimensions ranged from 78% for flow of information to 100% for hazing. Only five out of 13 dimensions showed less than 90% accuracy across companies (flow of information, autonomy, respect, fairness, and inclusion). Four of the dimensions (flow of information, autonomy, respect, and fairness) had means that were closer to 3.0 than any other dimensions, (ranging from 3.05 to 3.16). It is likely that the proximity of their mean to the arbitrary threshold affected their accuracy.

Forty-seven percent of companies had single items that agreed perfectly with the full scale. Seventy-six percent of companies had between 0 and 2 errors in the single-item dimensions. The remaining 24% of companies accounted for 67% of the total single-item errors.

Table 29

Agreement of Scale and Single Item in Detecting Climate Issues

Unit ID	Info	Auton	Respect	Fair	Inclus	PSUM	LS fam	Open	LSUM	Profes	Bully	PerfOr	Hazing	Total correct by CO
43	0	0	0	0	1	1	1	0	1	1	1	1	1	8
2	0	0	0	0	1	1	1	1	1	1	1	1	1	9
13	1	1	1	0	1	0	1	0	1	1	0	1	1	9
14	0	0	1	1	0	1	1	1	1	1	0	1	1	9
33	0	1	1	0	0	1	1	1	0	1	1	1	1	9
38	1	1	0	0	1	0	1	1	0	1	1	1	1	9
57	0	0	1	1	1	1	0	0	1	1	1	1	1	9
6	1	1	1	1	1	1	1	1	0	0	0	1	1	10
16	1	0	1	0	0	1	1	1	1	1	1	1	1	10
20	1	0	0	1	1	1	0	1	1	1	1	1	1	10
47	0	1	0	1	0	1	1	1	1	1	1	1	1	10
51	1	1	0	1	0	1	1	1	1	0	1	1	1	10
4	0	1	1	1	1	1	0	1	1	1	1	1	1	11
8	0	0	1	1	1	1	1	1	1	1	1	1	1	11
34	1	1	0	0	1	1	1	1	1	1	1	1	1	11
37	0	1	0	1	1	1	1	1	1	1	1	1	1	11
40	1	0	1	1	0	1	1	1	1	1	1	1	1	11
41	1	1	1	1	0	1	1	1	1	1	1	0	1	11
60	1	1	1	1	0	1	1	1	1	0	1	1	1	11
1	1	1	1	1	1	0	1	1	1	1	1	1	1	12
3	1	1	1	0	1	1	1	1	1	1	1	1	1	12
21	0	1	1	1	1	1	1	1	1	1	1	1	1	12
23	0	1	1	1	1	1	1	1	1	1	1	1	1	12
29	1	1	0	1	1	1	1	1	1	1	1	1	1	12
31	1	0	1	1	1	1	1	1	1	1	1	1	1	12
42	1	1	1	1	1	1	0	1	1	1	1	1	1	12
45	1	1	1	1	1	0	1	1	1	1	1	1	1	12
5	1	1	1	1	1	1	1	1	1	1	1	1	1	13

9	1	1	1	1	1	1	1	1	1	1	1	1	1	13
10	1	1	1	1	1	1	1	1	1	1	1	1	1	13
11	1	1	1	1	1	1	1	1	1	1	1	1	1	13
12	1	1	1	1	1	1	1	1	1	1	1	1	1	13
15	1	1	1	1	1	1	1	1	1	1	1	1	1	13
17	1	1	1	1	1	1	1	1	1	1	1	1	1	13
18	1	1	1	1	1	1	1	1	1	1	1	1	1	13
19	1	1	1	1	1	1	1	1	1	1	1	1	1	13
26	1	1	1	1	1	1	1	1	1	1	1	1	1	13
27	1	1	1	1	1	1	1	1	1	1	1	1	1	13
28	1	1	1	1	1	1	1	1	1	1	1	1	1	13
30	1	1	1	1	1	1	1	1	1	1	1	1	1	13
32	1	1	1	1	1	1	1	1	1	1	1	1	1	13
36	1	1	1	1	1	1	1	1	1	1	1	1	1	13
44	1	1	1	1	1	1	1	1	1	1	1	1	1	13
46	1	1	1	1	1	1	1	1	1	1	1	1	1	13
48	1	1	1	1	1	1	1	1	1	1	1	1	1	13
49	1	1	1	1	1	1	1	1	1	1	1	1	1	13
50	1	1	1	1	1	1	1	1	1	1	1	1	1	13
53	1	1	1	1	1	1	1	1	1	1	1	1	1	13
56	1	1	1	1	1	1	1	1	1	1	1	1	1	13
58	1	1	1	1	1	1	1	1	1	1	1	1	1	13
59	1	1	1	1	1	1	1	1	1	1	1	1	1	13
Total correct	40	42	42	43	43	47	47	48	48	48	48	50	51	
%	78	82	82	84	84	92	92	94	94	94	94	98	100	

Note. 1 = Single item agrees with scale; 0 = Single item does not agree with scale.

A reasonable question is what differences exist between the 12 companies that had agreement between single items and scales on 10 or fewer dimensions and the remaining 39 companies that had a higher degree of agreement between single items and scales. While the present research was not designed to identify such differences, some post hoc analyses were conducted to investigate. One potential contributing factor is the sample size for each company. It is possible that companies with fewer individuals provided less stable estimates of company-level effects. However, the bivariate correlation between company size and number of dimensions with single item and scale agreement was only -.04, suggesting no meaningful relationship between company sample size and single item accuracy.

Another potential factor is the type of company. It is possible that whether a company is Combat Arms, Combat Support, or Combat Service Support could affect the accuracy of the single-item measure; however, that does not appear to have an effect. If unit type was affecting the single item accuracy, we would expect to see a disproportionate number of units of one type with 10 or fewer instances of agreement between the single item and scale. Of the 12 companies with 10 or fewer instances of disagreement, six were Combat Arms, four were Combat Support, and two were Combat Service Support. This is roughly proportionate to the percentages of unit type in our sample—57.8% Combat Arms, 29.8% Combat Support, and 12.3% Combat Service Support. Additional research is necessary to determine if single item-scale agreement is systematically linked to something about the unit.

Discussion

The development and maintenance of command climate can be critical for the effectiveness of Army units. It is valuable to measure and track climate over time so that leaders can identify emerging problems and monitor their efforts to change or improve the climate in certain areas. While current assessments may have more than 100 items, this research tested the feasibility of using single-item measures rather than multi-item scales to capture 13 Army command climate dimensions. A series of 13 single items was examined in comparison with an 84-item survey that used multi-item scales. Results generally supported the feasibility of using the 13-item tool as a quick “health check” on the command climate of a unit. Convergent validities of the single items with the multi-item scales were above .7 for 10 of the 13 scales, and predictive validities with proposed outcomes were on par with predictive validities achieved by the multi-item scales.

Hierarchical regressions were used to determine whether the multi-item scales explained significantly more variance in the associated criteria than the single items. Of the 18 dimension-criteria relationships that were significant, hierarchical regressions for five of the relationships found that the multi-item scale added significant variance to the single-item in predicting the criteria. The increases ranged from 5 to 10%. All of the multi-item scales demonstrated high reliabilities, but several of the scales demonstrated lower than expected levels of within group agreement, suggesting the constructs may not measure a group-level effect.

Reliability of the Full Scales and Single Items

Results showed strong support for the reliability of the multi-item assessment and moderate support for the reliability of the single-item assessment. Examination of item-total

correlations identified items in five multi-item scales that needed to be removed to improve reliabilities: fairness, peer support for unit members, performance orientation, hazing, and bullying. In most cases, the item that was removed from a scale was a reverse-worded item that did not fit with the other items even after being recoded. After these items were removed, the reliabilities for all 13 climate scales ranged from .80 for inclusion to .95 for respect for the individual.

Despite the generally strong reliabilities, some scale reliabilities could still be improved to reduce measurement error if these scales were going to be used in high-stakes selection or promotion decisions. In particular, reliabilities for inclusion (.80), peer support for unit members (.82), and bullying (.83) could still be further improved. Because the application of the multi-item scale assessments is for developmental and informational purposes rather than for selection or promotion, the current reliabilities, which were generally above .85, are sufficient.

Reliabilities for the single-item assessments were more moderate. As described in the introduction, calculation of reliability for single-item measures is somewhat tenuous. We adopted the approach of Wanous et al. (1997) to estimate the single-item reliabilities, which uses the correction for attenuation formula and algebraically solves for single-item reliability. This approach assumes that the single item is perfectly parallel with the domain of the scale (an assumed true correlation of 1.0) and uses the correlation between the single item and the scale and the reliability of the scale as inputs. This means that the reliability estimate of the single-item measure will be lower if there is a low correlation between the single item and the scale and/or there is a low scale reliability. As just discussed, scales with the lowest reliabilities were inclusion, peer support for unit members, and bullying. The correlations between the single items and scales ranged from .39 for bullying to .87 for respect for the individual, with the three lowest single-item to scale correlations for bullying (.39), inclusion (.54), and fairness (.65). Not surprisingly, the lowest single-item reliability estimates were for bullying (.18), inclusion (.36), and fairness (.48). The other single-item reliability estimates ranged from .58 (professionalism) to .80 (respect for the individual).

The single-item measure with the lowest estimated reliability was bullying. A review of the single item and scale items showed that the single item was negatively worded, while the scale items were positively worded (though reverse-scored to match the single item). Often negatively worded items in a scale load on a different factor in factor analysis, rather than simply in the opposite direction (Roszkowski & Soven, 2010; Woods, 2006). Correlations between negatively and positively worded items often have different magnitudes than inter-item correlations of only positively worded items (Roszkowski & Soven, 2010). Research suggests that people think differently about an item depending on how it is worded or how it is framed (Van Sonderen, Sanderman, & Coyne, 2013). This is likely the reason why the single item for bullying achieved such a low correlation with the scale (.39). In addition, the bullying scale had one of the lowest reliabilities (.83). Together, therefore, the low correlation between the single item and scale and the slightly lower scale reliability are limiting the single-item reliability estimate. In addition to being negatively worded, upon review of the bullying items, we found that the items asked about the degree to which leaders discourage, prevent, and respond to bullying, rather than asking about to what extent bullying exists. Therefore, it is ambiguous whether the respondents are disagreeing (or agreeing) because there is bullying and the leader is not responding to it, or there is no bullying. Modifying the bullying scale so that the scale items

and single item are worded similarly should prevent this ambiguity as well as improve the reliability of the single item.

The fairness and inclusion dimensions also had low single-item reliability estimates (.48 and .36, respectively). The inclusion scale reliability of .80 likely had a negative impact on the single-item reliability estimate. While the scale reliability of .80 is sufficient for early research or low-stakes contexts, it may not be sufficient for deriving accurate estimates of single-item reliability. The inclusion scale also had the second lowest correlation between scale and single item (.54). A review of the inclusion single item and scale items does not reveal any clear reasons for the low single-item to scale correlation; the lower scale correlation, however, may be due to the fact that three of the items are statements about members feeling excluded, while two items are statements about members feeling included. This divergence in wording approach may be limiting the scale reliability. For fairness, the correlation between the single item and scale was moderate, at .65, and not high enough to conclude that the single item is parallel with the scale. The initial fairness scale had five items; however, two items were dropped to increase the reliability from .77 to .88, leaving three items remaining. A comparison of the single item and remaining scale items shows that while each of the three scale items use the word "fair" in the statement, the single item does not. Ensuring that the single item uses the same terminology as the scale items will likely improve the single-item reliability.

The highest single-item reliability estimates were for leader support for unit members and respect for the individual, at .70 and .80, respectively. As mentioned, respect for the individual showed the strongest scale reliability and correlation between scale and single item, which resulted in the respect for the individual dimension also achieving the highest single-item reliability. For three of the single-item dimensions, single-item reliability estimates were just below the traditional threshold that would be considered adequate for early scale development: leader support for family (.68), openness (.69), and hazing (.69). These would generally be acceptable for use operationally in low-stakes contexts.

Five dimensions demonstrated moderate single-item reliabilities: information, autonomy, peer support for unit members, performance orientation, and professionalism had single item reliabilities ranging from .58 to .65. These dimensions would benefit from small modifications to the scale or single item to achieve higher scale reliabilities and higher correlations between the scale and single items. For example, peer support for unit members would show a marked improvement in single-item reliability (currently .65) if the scale reliability of .82 could be improved. The scale reliabilities for the other four dimensions (information, autonomy, performance orientation, and professionalism) were all above .90, and all correlations between scales and single items are above .70. If these can be improved for any of the dimensions, the single-item reliability estimate would improve.

Construct Validity of Command Climate Scales

In order to establish the validity of the command climate dimension assessments, we examined both the intercorrelations among scales and the scales' factor structure to determine whether the dimensions can be distinguished from each other. In the individual-level data, correlations ranged in strength from .13 to .74, with an average of .53. On the surface, this suggests that the scales are distinct, but many scales are closely related. For example, respect for

the individual was related to openness, fairness, professionalism, and leader support for unit members in the range of .73 to .74. While these constructs are not identical, they are very similar. At the company level, the correlations between these variables and respect for the individual ranged from .82 to .91, suggesting even less distinction.

An examination of the EFA results showed 10 factors in the data with eigenvalues greater than 1.0. Flow of information, autonomy, leader support for families, bullying, and hazing showed distinct factors with little cross-loading of items from other scales. The negatively worded inclusion items also made up a factor. This provides evidence for the distinctiveness of these scales. Although these six factors did correlate with one another, the correlations ranged from only .15 to .79, with an average correlation of .50. The other four factors consisted of items that cross-loaded from different scales. Professionalism items were spread over two factors: performance orientation and respect for the individual. The items that asked about the unit members tended to load on performance orientation, whereas the items that asked about the leader tended to load on respect for the individual. The respect for the individual factor also had weak to moderate cross-loadings from some items from the fairness scale and the positively worded inclusion items. This suggests that inclusion and fairness may be partially overlapping with respect for the individual. The remaining items from the fairness scale loaded on a leader openness factor, which also had weak to moderate cross-loadings from some leader support for unit members items. Leaders caring about and being transparent with unit members was associated with leaders being willing to listen and hear from unit members. Peer support for unit members made up another factor, which had weak loadings from the remaining leader support for unit member items. It appears that an underlying support factor is contributing to this effect, but the peer support for unit members items had much higher factor loadings than the leader support for unit members items.

Support for Group-Level Effects

Both Army doctrine and the research literature have defined climate as a unit-level construct; specifically, as shared perceptions and attitudes about the unit's daily functioning. Best practices in group-level research indicate that each construct should be tested to determine if there is sufficient group-level variance to warrant conducting analyses at the group level. Three indices were used to examine whether the multi-item scales and single-item assessments represented group-level constructs: ICC(1), ICC(2), and $r_{wg}/r_{wg(j)}$. Support for aggregating the measures to the group level was mixed.

For ICC(1), results showed that all multi-item scales reached statistical significance at the .05 effect size threshold, suggesting a small group effect (LeBreton & Senter, 2008). Many dimensions, including openness, leader support for unit members, leader support for family, performance orientation, professionalism, and respect for the individual, had an ICC(1) of .10 or greater, suggesting a medium-size group level effect. For the single-item measures, each dimension had a statistically significant ICC(1), but the fairness and bullying dimensions did not reach .05 effect size. These effect sizes are not particularly large–one interpretation of the effect sizes is that an effect of .10 corresponds to approximately 10 percent of the variance in ratings being due to systematic differences between companies. On the other hand, the significant level of the group effects suggests that ignoring the company level influences on these variables would

violate the regression assumption of independent residuals; therefore, an examination of group level effects is important.

For the ICC(2) statistic, on the other hand, none of the multi-item scales reached the predefined benchmark of .80. Many reached the level of .65, which can be considered reasonable for initial research. Based on the revised threshold of .65, we aggregated 9 of 13 multi-item scales to the company level and 8 of 13 single-item measures to the company level (see Table 13 for a review). ICC(2) can be interpreted as the reliability of the means of raters within each group. Using a measure with an ICC(2) of .65, therefore, suggests that 65% of the score is unit-level variability, and the remaining 35% is individual-level variability. This level of error is not ideal and subsequent research should work to improve the ICC(2) scores that were achieved in this project. One way to increase ICC(2) scores is by including more raters from each company in the company-level means. This would ensure that the company is adequately sampled; this would be likely to produce more stable estimates. The present research set a minimum cutoff of 10 raters per company, which may not have been enough to produce a mean that accurately represented the company. As an exploratory analysis, we increased the minimum cutoff to 15 and reran the ICCs. ICC(2) increased universally, but only by a small amount. In future research, having a cutoff that exceeds 25–30 may provide ICC(2) values that reflect higher reliabilities of unit means. On the other hand, if consensus within the company is low because unit members are having differing experiences, having more raters will not necessarily produce a higher ICC(2).

As LeBreton and Senter (2008) point out, a low ICC(2) can be due to low interrater reliability, low interrater agreement, or both. The source of the low ICC is important and cannot be determined from ICC(2) alone. It is possible to have low interrater reliability reflected in the ICC(2), but adequate interrater agreement. In particular, this occurs when the group level scores are restricted in range and groups are not very different from each other (LeBreton et al., 2003). The scale and single-item measures that did not reach the ICC(2) threshold of .65 were therefore tested for this using average r_{wg} values. The average r_{wg} values obtained for the single-item measures did not show adequate agreement. However, two of the multi-item scales showed adequate agreement on the $r_{wg(j)}$ statistics: hazing and inclusion. These measures were aggregated, resulting in a total of 11 company-level multi-item scales for the criterion-related validation out of 13, and 8 company-level single-item measures out of 13. Generally speaking, r_{wg} and $r_{wg(j)}$ varied greatly from company to company, and there was no one dimension of the multi-item scale or single-item measures that showed consistently high or low agreement. This might suggest differences within companies in how items are interpreted; for example, leader support for family might mean different things to married and unmarried Soldiers.

Another reason for low agreement could be due to group-level effects below the company level. Data from each company could differ based on differences in climate for units that are nested within the company–the platoons and squads. Army companies typically have 100–250 Soldiers and are composed of a headquarters element and two or more platoons, which each have two or more squads. Each squad typically has 7 to 14 Soldiers. It may be that part of the group level effect is at the company level, while another part of the group level effect is at the platoon or squad level. Data for this research were not collected by squad or platoon, so this information is not captured. That means that in some cases, company data may represent predominantly one platoon and in other cases company data may represent multiple platoons. To the extent that this is inconsistent across companies, this could produce a range in the level of agreement across

companies. Determining climate effects at multiple levels of the hierarchy would be highly time and resource intensive but would need to be conducted to empirically determine the impact of nested group data at multiple levels.

Another factor to consider is that perceptions of some of these constructs may not be shared but rather are individual perceptions or experiences (e.g., in-group/out-group, leader member exchange relationships). As an example, if bullying largely reflects individual perceptions and experiences, such that some Soldiers in a unit experienced bullying but others did not, results would demonstrate no or a low group-level effect. From a practical perspective, it would be unwise in these cases to simply interpret a low-level group effect and ignore those cases not conforming to the "norm." In these cases, using the mean to form aggregate data is not useful; rather, it may make sense to examine climate strength rather than climate level and use an index such as the standard deviation within each company instead of the mean to identify potential problems.

Finally, for conducting the predictive validity analyses, group-level effects were applied in the outcome measures. Support for aggregating the outcome measures to the group level was mixed, with ICCs for most outcomes suggesting they should be aggregated, though some ICCs did not. Specifically, ICC(1) was not significant for continuance commitment and stress and was not above the .05 effect size threshold for continuance commitment, stress, family desire to stay, well-being, participation in social events, family satisfaction, career intentions, and number of Article 15s. It is therefore questionable whether group-level effects exist on these measures and whether aggregation to a group mean is appropriate.

Ultimately, the decision to aggregate depends on multiple factors, including the research questions and the type of composition models that apply to the constructs (e.g., see Chan, 1998). ICCs suggest that group-level effects should be examined when they are sufficiently high; however, low scores do not preclude conducting group-level analyses. Some outcomes (cohesion, unit resilience, unit performance, morale, and weapon qualifications) showed moderate company-level effects (ICC(1)s were above .10) as well as stable company-level mean estimates (ICC(2) of .74 and above). In these cases, examining group-level effects is critical because the results indicate a nested structure in the data.

Validity of the Multi-Item and Single-Item Assessments

Criterion-related validity evidence was generally strong for the multi-item and single-item measures. Of the 11 multi-item scales that were aggregated to the company level, all but professionalism were significantly related to at least one of the hypothesized outcome measures. All eight of the single-item measures that were aggregated to the company level were related to at least one outcome. In addition, the multi-item and single-item measures showed similar profiles of criterion-related validity relationships. In evaluating the effectiveness of the single-item measures, there were no cases in which using the full scale would have revealed effects (e.g., significant regression weights) missed by the single-item scale. Interestingly, there was one instance (peer support for unit members) in which the effect would have been missed by using only the multi-item scale rather than the single-item measure: the single-item measure of peer support for unit members showed significant prediction of stress, while the multi-item scale did not.

Hierarchical regressions were used to determine whether the multi-item scales explained significantly more variance in the associated criteria than the single item alone. Of the 18 dimension-criteria relationships that were significant, hierarchical regressions for five of the relationships found that the multi-item scale added significant variance to the single item in predicting the criteria. The greatest increase in variance explained was 12%, which was the amount of increase found in the regression of morale on fairness and morale on bullying.

For three dimensions (flow of information, autonomy, and performance orientation), unit performance was a hypothesized outcome that was operationalized four different ways: Soldier ratings of unit performance, company leader ratings of unit performance, battalion commander ratings of unit performance, and battalion commander ratings of unit readiness. While Soldier ratings of unit performance were related to each of the three dimensions, the three leader-rated criteria were not. One reason for this may be that the sample sizes available for those regressions were not as large due to missing commander ratings for some companies, so the power was lower to identify results. Another possible reason is that the Soldier ratings of unit performance were influenced by halo or other biases, while the commander-rated unit performance measures were free from common method bias and participants' desire for consistent responding. Commander ratings, however, could still be affected by other types of response biases. Alternatively, the lack of a relationship could be due to differences in standards or perspectives between the Soldiers and the commanders providing ratings. Due to the possibility of method bias in the relationship between the climate dimensions and Soldier ratings of unit performance, the strength of these relationships found in the presented data may be overestimated/inflated and is thereby subject to some interpretation in its conveyance of a true relationship within the military population. The validity of the practical application of the single-item assessments was examined by comparing the performance of the single items to the multi-item scales for identifying dimensions that were above or below a given mean threshold, which would signal to a commander that there was a possible climate problem. This is likely how a single-item assessment would be used in a practical setting. Overall, the single-item assessments matched the classification of the multi-item scale above or below the threshold in 90% of the cases. The match rate for specific climate dimensions ranged from 78% to 100% matches. Validity results for each of the climate dimensions will be discussed in greater detail.

Flow of information. The flow of information dimension predicted two out of the five associated outcomes, predicting morale and Soldier-rated unit performance, but not unit performance ratings from officers, or unit performance or readiness ratings from the battalion commander. In the hierarchical regression of morale on the single-item and multi-item assessments, the single item predicted more than 47% of the variance, but the multi-item scale contributed almost 10% more prediction. This was reflected in the differences between regression weights in simple regressions and suggests that the prediction of morale by the single-item assessment could be improved. For predicting Soldier-rated unit performance, flow of information predicted just under 46% of the variance in unit performance, with the simple regression weights for the multi-item and single-item assessments demonstrating similar magnitudes. Adding the multi-item scale to the regression of unit performance on the single item did not contribute additional variance explained.

From a practical perspective, the single item for flow of information achieved "success" in 78% of the companies when using the example cut score. This is moderately high but was the

lowest rate of the 13 single-item assessments. Given the added predictive validity of the multi-item scale for morale and lower success rate of the single item in matching the classification of the multi-item scale, it could be useful to consider improvements to this item through further review and testing. A review of the single item and scale items shows that the single item describes the general communication of information throughout the unit, while items in the multi-item scale capture both general information sharing within the unit as well as communication specifically from leaders and chain of command. It may be necessary to reflect both of these concepts in the single-item assessment or to refocus the multi-item scale only on one or the other concept.

Leader openness. Leader openness had a strong relationship with affective commitment, explaining nearly 46% of the variance (based on the adjusted R^2), and the single-item assessment performed similarly to the multi-item scale at detecting this relationship. Thus, the more open a leader is to listening to Soldiers and gaining new perspectives, the more emotionally committed the Soldiers in the company are to the Army. On the other hand, leader openness had no detectable relationship with continuance commitment, indicating that a climate of leader openness does not significantly impact the perception among Soldiers that they must remain in the Army due to the perceived costs of leaving being too great. This was true for both the multi-item and single-item assessments. These findings are similar to previous Army climate research, which has found significant relationships with affective commitment but not with continuance commitment (e.g., Langkamer & Ervin, 2008). In examining the prediction of the single item using an example cutoff score, the single item was fairly successful, predicting the same outcome as the multi-item scale for 94% of the companies.

Peer support for unit members. The peer support for unit members dimension significantly predicted all six of the anticipated outcome measures. This dimension had the strongest relationships with cohesion, morale, and unit resilience. The single-item assessment accounted for nearly 78% of the variance in cohesion, 66% of unit resilience, and 61% of morale. In each case, the multi-item scale did not contribute additional prediction of variance above the single-item assessment. The very strong relationship between peer support for unit members and cohesion likely reflects some degree of conceptual overlap between these two constructs; that is, to some extent the two measures may partially be capturing the same construct. Cohesion has been described as a bonding together of the members of a unit that maintains their will and commitment to each other (e.g., Johns et al., 1984). The items for both peer support for unit members and for cohesion capture unit member perceptions about being close to each other, supporting each other, and working together. Morale and unit resilience, on the other hand, are conceptually consequences of a peer support climate, such that units that have climates in which peers support one another will create high levels of morale and unit resilience. In each of these cases, however, it is likely that the relationships among these variables would be mutually interdependent.

A climate in which peers support one another is also associated with well-being and affective commitment. Peer support for unit members accounted for just under 30% of the variance in well-being, suggesting that support from peers may impact individuals' physical and mental health. For affective commitment, just under 43% of the variance was accounted for by the single-item assessment. This suggests that support from peers also affects emotional

attachment to the unit and the Army. In both cases, the multi-item scale did not contribute to prediction above and beyond the single-item assessment.

Finally, peer support accounted for about 12% of the variance in stress. As expected, the relationship was negative, such that greater support from peers was associated with less stress. This relationship was weaker than the others, yet suggested a 12% reduction in stress, which is not insignificant. Moreover, this effect was found from data collected in garrison. There may be an even stronger relationship on deployments, when stress is expected to be elevated. As an indicator of the strength of the single-item assessment, the multi-item scale did not contribute any additional prediction to the hierarchical regression. Moreover, as described previously, the multi-item scale did not predict stress in a simple regression, while the single-item assessment did. In all other cases, the multi-item and single-item assessments performed equally well in the simple regressions. Research on the relationship between climate and employee stress is still in progress and has been examined in a number of different ways: (a) as a direct effect (either positive or negative) on stress and well-being, (b) as a moderator of the relationship between a stressor and employee stress, and (c) as a mediator of the relationship between a stressor and employee stress (e.g., Jex, Sliter, & Britton, 2014). Regardless of the specific mechanism by which it might function, monitoring and facilitating peer support may have important positive effects on a company's stress, cohesion, perceived well-being, and affective commitment.

In the analyses to compare group-level status using the thresholded single-item and multi-item assessments, the single-item assessments provided the same indication of group status as the multi-item assessment in 92% of the companies.

Leader support for unit members. Like peer support for unit members, leader support for unit members was significantly related to affective commitment (accounting for 46% of the variance), well-being (23% of the variance), and unit resilience (72% of the variance). Thus, as expected, when unit members perceive that their leaders care about their welfare and provide the assistance they need, they have a greater sense of belonging to the Army, have a sense of well-being and life satisfaction, and perceive that their unit is able to overcome challenges. The multi-item scale did not contribute additional variance to the single-item assessment in predicting affective commitment and well-being; however, for unit resilience, adding the multi-item scale to the single item predicted an additional 7% of the variance, which was statistically significant, but not a large effect. From the perspective of practical implementation of the single-item assessment, however, the reduction of items to reduce rater burden may still be a worthwhile tradeoff if the single item is sufficiently able to detect potential leader support climate concerns. Analyses comparing the prediction outcomes of the single-item and multi-item assessments found that the single-item assessment predicted the same outcome as the multi-item assessment for 94% of the companies.

Performance orientation. Seven outcome variables were regressed on performance orientation, but only one relationship was significant. Physical fitness, weapon qualification, and disciplinary action were not related to performance orientation climate. The physical fitness measure (APFT scores) may suffer from range restriction because fitness is generally high, which would limit its correlation with other variables. Also, these scores were self-reported, so they may be susceptible to recall biases. Disciplinary actions may have been underreported in the

self-report data and is also a variable with a low base-rate of occurrence, which would produce a restricted range.

Another factor that may impact these results is that physical fitness and weapon qualification scores may be heavily influenced by the Soldier's type of unit and MOS. With respect to weapons qualification, group level results suggested that weapons qualification had a much stronger company-level effect than performance orientation: ICC(1) of .28 for weapon qualification versus .11 for the multi-item performance orientation scale and .10 for the single-item assessment. The company level effect for weapons qualification suggests that companies vary considerably on this variable, likely due to differences in MOS and unit type. Because the same level of variability was not found for performance orientation, it is not surprising that the variables did not correlate.

The multi-item and single-item assessments of performance orientation significantly predicted the self-report measure of unit performance, with the multi-item scale contributing about 10% additional prediction to the single item. As discussed previously, the other measures of unit performance did not show a relationship with the climate predictors. From a practical standpoint, in the analyses comparing the prediction of outcomes for the single-item and multi-item assessment, the single-item assessment for performance orientation predicted the same outcome as the multi-item scale for 98% of the companies, suggesting preliminary support for the application of the single-item assessment in place of the multi-item assessment.

Autonomy. Findings from this study suggest that autonomy predicted two out of the planned five outcomes. A climate of autonomy was associated with higher job satisfaction and Soldier-rated unit performance, but not the other three measures of unit performance (officer ratings of unit performance and battalion commander ratings of unit performance and readiness). The single-item assessment predicted 33% and 56% of the variance in job satisfaction and Soldier-rated unit performance, respectively. The multi-item scale did not predict any variance above the single item. These results indicate that a climate in which Soldiers believe that they are trusted by leadership to make decisions and able to take initiative and figure out problems is associated with Soldiers being satisfied with the opportunities and challenges of their job and perceiving that their unit is performing well. This is in line with studies that show autonomy in job decisions has numerous positive effects on employee motivation and performance (e.g., Tripathi & Agarwal, 1988; Turnage & Muchinsky, 1976). The single-item assessment predicted the same outcome as the multi-item scale for 82% of the companies. Although that is a reasonably high rate of prediction, it may suggest that improvements could be considered to increase the correspondence of the two measures. On the other hand, of all the command climate dimensions, autonomy had the scale mean that was closest to 3.0 (mean = 3.05). For the prediction tests, 3.0 was used as the cut point; because the mean for autonomy was so close to 3.0, it is likely that just by chance the single-item and multi-item scale scores for the companies fell on different sides of the cut score.

Respect for the individual. The respect for the individual dimension predicted both of its anticipated outcomes and was significantly and strongly related to both mutual trust and cohesion, with the single-item assessment accounting for 60% and 76% of the variance in these outcomes, respectively. The multi-item scale accounted for an additional 5% and 8% of these outcomes, respectively. These relationships suggest that units in which members perceive that

they are respected and valued also tend to develop a sense of trust among members and a sense of pride and bonding among unit members. This reinforces the importance of a climate of mutual respect, dignity, and personal worth, which decades of research on command climate has identified as important (e.g., Bullis & Reed, 2003; Craig & Brace, 1992; Lempke, 1988). While conceptually distinct, the relationships between a climate of respect for the individual, mutual trust, and cohesion may suggest some degree of mutual interdependence.

In examining the practical application of the single-item assessment, results comparing the prediction of outcomes using the single item and multi-item scale indicated that the single-item assessment led to the same conclusion as the multi-item scale for 82% of the companies. While relatively high, that is one of the lower prediction rates. Given that the multi-item scale predicted the criteria slightly better than the single-item assessment, it may be useful to consider modifications to the single item. In reviewing the single item and multiple items in the scale, while both capture whether members are treated with dignity and respect, only the scale items capture whether unit leaders are active in promoting respect and confronting actions that can undermine respect.

Inclusion. A climate of inclusion predicted one out of the three outcome measures, showing a significant relationship with affective commitment (37% of variance accounted for), but not with well-being or participation in social events. This suggests that units that have a climate of inclusion are more likely to have members that feel a sense of emotional commitment and belonging to the Army. The climate of inclusion in the unit, however, did not correlate with Soldiers' general sense of well-being and satisfaction with life or with their choice to participate in unit social events. In the case of well-being, our operationalization was rather broad, and it may be difficult for a climate of inclusion, which is relatively specific, to demonstrate a significant relationship. Climate of inclusion is likely not a primary predictor of general well-being, so subsequent research should examine more specific facets of well-being. Similarly, there are likely many reasons for Soldiers to participate or not participate in social events: to include personality factors (e.g., introversion/extraversion, whether they have a spouse/family or not). In addition, two of the participation in social events items ask about engaging with the company Facebook page, and not all personnel would have an interest in engaging in that type of social "event."

In comparing the prediction of the single-item and multi-item assessments, the multi-item inclusion scale did not contribute additional prediction for affective commitment beyond the single-item assessment. The single-item assessment predicted the same outcome as the multi-item scale for 84% of the companies, which is high but may suggest improvements can be made to the single item. A climate of inclusion is clearly critical in the Army and is described both in ADRP 6-22 *Army Leadership* (U.S. Department of the Army, 2012) and the associated field manual (FM 6-22 *Leader Development*, 2015). Other current ARI projects are examining this dimension of climate in greater detail and may soon be able to provide suggestions regarding both the multi-item and single-item assessments of this dimension.

Other dimensions. Because leader support for family, professionalism, and hazing did not show sufficient ICCs for the single-item assessments, criterion validities were examined only for the multi-item scales. Leader support for family was predictive of family satisfaction, but not family's desire to stay. This suggests that the actions and approach of the proximal leader have

an impact on the extent to which Soldiers' families feel satisfied, but the leader may have less power to affect whether families want to remain in the Army. Desire to stay likely has many different influences, such as other job opportunities, attitude toward moving, desire to remain at certain schools or locations, assignment and promotion decisions, reenlistment bonus packages, and others, and therefore is less affected by the command climate. When the single-item assessment of leader support for family was compared with the multi-item scale on the decision outcomes that would be determined using an example cutoff score, the single-item assessment matched the outcomes of the multi-item scale for 92% of the companies. This suggests that from a practical standpoint, if the single-item assessment was used to help commanders identify potential problem areas, it would provide essentially the same information to commanders as the multi-item scale.

Professionalism was not related to any of the three proposed criteria: weapon qualifications, physical fitness, or disciplinary action. This was the only dimension that did not have a significant relationship with at least one outcome. As discussed previously, there are a number of potential problems with these three criteria, which may have prevented significant relationships with professionalism. Nevertheless, the comparison of the predicted outcomes using an example cutoff score showed that the single-item assessment classified the same outcome as the multi-item assessment for 94% of the companies, showing very high performance.

A hazing climate was significantly related to affective commitment, morale, and stress, with both affective commitment and morale decreasing with hazing, and stress increasing. Thus, a climate in which Soldiers perceive members are forced to engage in harassing situations was associated with Soldiers who are less emotionally attached to the Army and have lower morale and higher levels of stress. For hazing, scores for 100% of the companies were classified the same by the single-item assessment as for the multi-item scale.

Finally, fairness and bullying were examined differently from the other climate dimensions due to low interrater consistency and agreement. Similar to hazing, the single-item assessment for bullying performed very well in predicting the same outcome as the multi-item scale when using an example cutoff score. For bullying, the single-item scale classified 94% of the companies the same as the multi-item scale. The classification of the single-item assessment for fairness was not as good, however, classifying 84% of the companies the same as the multi-item scale. In part, this may be due to the fact that it had the second closest mean to 3.0 (fairness mean = 3.08), which was serving as the example cutoff score. Because scores tended to be right around 3.08, this dimension has a higher likelihood of scale and single item falling on opposite sides of the threshold by chance.

Summary. Most of the anticipated relationships were observed as expected, showing evidence for the criterion-related validity of the full scale and single items. Moreover, the single items generally performed as well as the full scales in predicting relationships with criteria. Many of the relationships in the criterion-related validity evaluation were very strong. This may be partially due to common method bias, though probably not entirely, because there were company-level effects. Nevertheless, common method bias is a concern because it has been found to inflate relationships by as much as 25 to 40% (Podsakoff et al., 2003). Statistical corrections for common method bias are in early stages of development and research has not yet

demonstrated how effective these corrections are (Conway & Lance, 2010). Subsequent research should explore ways of collecting data on outcomes through means different from self-report; for example, constructs could be measured based on surveying leaders or family members.

The command climate constructs generally showed the expected patterns of criterion relationships, suggesting the measures have a degree of criterion-related validity. We found some evidence that for a small number of dimensions, multi-item scales and single items were preferable to single-item measures alone from a prediction standpoint. In most cases, however, and in the practical application comparisons, the single items performed very accurately overall. Overall, regression results and the practical application analysis support the effectiveness of single item measures.

Exploratory Diagnostic Tool

Correlations between the diagnostic items for a given dimension and the full scale and single item were generally very high (.6 to .8 range), but not as high as the full scale intercorrelations or the correlations between the full scale and single item. This was expected because the diagnostic items are not intended to be parallel with the scales and single items. While most correlations were positive, the correlations with negatively-worded items were negative, as expected. This was true for two negatively worded items in the performance orientation diagnostics set as well as one item in the bullying set. Similarly, while the hazing diagnostic items were positively worded, the hazing single item and scale items were negatively worded, so those relationships were also negative, as expected. Correlations between the positively worded and negatively worded items tended to be weaker in magnitude than the other correlations. For example, the bullying diagnostic item that was negatively worded had only a small-to-moderate positive relationship with the scale (.29) and single item (.42), and the correlations for the hazing diagnostic items with the scale and single item ranged from -.52 to -.67.

In terms of criterion-related validation, the diagnostics results were very similar to those for the full scale and single items. Many of the same patterns of relationships were found between diagnostics and the criterion that were found for the single items or full scale, with the climate measures correlating with some outcomes but not others (see Table 28 for a summary of the regression results). This suggests that the diagnostic items are generally aligned with the command climate dimensions and are measuring relevant aspects of the climate dimensions. There were only a few exceptions to this alignment. For the hazing dimension, the diagnostics were related to affective commitment and morale, as was the full scale, although the single item was not related to these outcomes. Hazing had low ICC(2), so the reliability of company-level means could be reflecting disagreement in the company. The scale would have a better chance of averaging out the disagreement and would be less affected by measurement error in the predictor. The leader support for family dimension showed a similar pattern with the family satisfaction outcome, such that the scale and diagnostics were predictive, but the single item was not. When family desire to stay was the outcome, only the set of diagnostic items was predictive; the scale and single item were not. Similarly, in the relationship between autonomy and unit performance as rated by the battalion commander, the diagnostics were an effective predictor of unit performance, but the scale and single item were not. More research is needed to determine if this is driven by specific items in the diagnostic sets. In practical applications, it may be one or

many diagnostic items that reveal a problem with the climate. Subsequent research is needed to identify the weight of each factor in the diagnostic sets in determining the relevant outcomes.

Practical Application of the Single Item and Diagnostic Process

From a practical perspective, the single-item measures showed sufficiently strong convergent validity with the multi-item scales and sufficient predictive validity with proposed outcomes to warrant use in the field. An important additional requirement, however, is to determine the specific aspects of the single-item data that would be useful for commanders and how that data should be presented. The mean-level response received for a climate dimension is important because it provides a useful and easy-to-understand summary of the data; variance or agreement within the unit, however, are also important because they provide the commander with information about similarity or differences of opinion within the company. At a minimum, a pilot test of the single items in the field should provide commanders with mean and standard deviation. An index of absolute agreement like r_{wg} could also be used, but the standard deviation might be just as useful and may be easier to understand. Other types of information that could be considered as part of a feedback report are ranges, outliers, responses separated by subunits (platoons, squads), and responses separated by certain demographic groups or MOS.

In addition, more research is needed to determine the critical thresholds that would be used for each statistic and to develop guidance for commanders that will enable them to act upon the results. The cut-off of 3.0 could be a logical starting place because it is in the middle of the rating scale for the climate items. In the anchors, 3.0 corresponds to the neutral response, so one could use 2.99 and below (or 3.01 and above for hazing and bullying because they are negatively worded) as the region where scores trigger a red flag for the company. Alternatively, a multi-tiered system could be used, such that scores of 4.0 and above are considered strong and coded green, scores from 2.9 to 3.9 are considered potentially at risk and coded amber, and scores under 2.9 are considered in trouble and coded red. Using this type of system, dimensions coded red would be viewed as immediate problems that should be examined right away, using the diagnostic items or other methods. Amber dimensions need to be examined but are not as urgent, and green dimensions are considered strength points for the company in terms of its climate. Testing and evaluation of these formats is necessary to determine the most effective ways of delivering the information to commanders.

If a single-item measure indicates a potential problem with a command climate dimension, diagnostics can be administered later to further investigate the nature of the problem. To refine the diagnostic measures to make them most useful, additional development and content validation of the diagnostic items is necessary. This will ensure that the factors in the diagnostic are comprehensive and able to identify the cause(s) of the climate problem. Because the follow-up survey is given only for a subset of climate dimensions, expanding the number of items within diagnostic sets would not necessarily increase the survey burden to a great extent.

Future Research

The present research found encouraging results with regard to the reliability and validity of many of the multi-item assessments and more moderate results for the single-item assessments. While positive, there were a number of dimensions that could benefit from

modifications to the items. Reliabilities for four scales (fairness, peer support for unit members, bullying, and inclusion) were below .90 and should be reviewed and modified in future research to achieve improvements. The reliabilities for nine of the multi-item scales were above .90 (information, autonomy, openness, leader support for unit members, leader support for family, performance orientation, professionalism, respect for the individual, and hazing). Even for these nine scales, however, increasing their reliabilities even more could provide a benefit by improving the estimates of reliabilities of the associated single-item measures. The estimated reliabilities for the single-item measures could all generally benefit from increases. Because of the methods used to estimate single-item reliability, improving estimated single-item reliability requires improving either the multi-item scale reliability, the correlation between the single item and multi-item scale, or both.

Single-item reliabilities were particularly low for bullying, inclusion, and fairness, suggesting the importance of exploring modifications in order to improve their reliabilities. For each of these dimensions, improving the multi-item scale reliability would likely improve the estimated single-item reliability. Five dimensions that demonstrated only moderate single-item reliabilities (information, autonomy, peer support for unit members, performance orientation, and professionalism) could also benefit from item modifications. Specific changes that appeared likely to generate meaningful improvements were suggested for these scales in the discussion section. In most cases, simply modifying the wording of the items might be sufficient. For example, the scale items for fairness all used the word "fair," whereas the single item did not. In other cases, it may make sense to use two single items to cover relevant facets of the construct. This was true for flow of information, which captured both general communication throughout the unit and communication specifically from leaders.

From a conceptual perspective, once these improvements to the scale items and single items are made, additional research could further examine the impact of multiple climate dimensions on outcomes of interest. Mediation analysis would be informative for understanding how climate impacts proximal and distal outcomes, as well as for identifying the antecedents of climate. Furthermore, moderation analyses could explore the possibility of combined effects of multiple climate dimensions on unit outcomes. As mentioned, a broader range of criteria constructs could be measured perhaps based on surveying leaders or family members to avoid problems with common method bias. As always with research, however, the quest for deeper understanding will have to be balanced with the demands that are placed on Soldiers in the form of over-surveying.

A second area for future research is to understand the impact of climate in the Army at different organizational levels. This research examined group-level effects with respect to company membership, but companies consist of smaller nested units and are themselves nested within larger units. A more complete understanding of important command climate dimensions and the effects of command climate on individual and group outcomes must expand beyond just the company level. According to Ehrhart et al. (2014), climate is strongest at the smallest levels and the effects of climate on individuals are the most pronounced at the levels that are most proximal to the individual. An examination of climate at the squad and platoon levels may yield further insights into the impact of climate on individual Soldiers and sources of leverage for improving command climate. Collecting data at multiple levels could also improve the usefulness of the results that are provided to commanders, enabling them to identify specific

squads or platoons that that could benefit from targeted interventions. One thing to consider, however, is how anonymity can be protected at low levels such as the squad or platoon to facilitate receiving honest answers.

Because of the nested structure of Army units, it is difficult to collect data from Soldiers that does not have implicit nested structure. In other words, the score at one level is almost always going to be influenced by a higher level grouping. Ignoring this group structure may affect our ability to model and make accurate inferences about groups and individuals. In our data, some companies came from the same battalions and brigades, and others from different battalions but the same brigade. To the extent that climate dimensions vary by battalion, brigade, or division, additional group-level effects could have been present in the data but were not examined in this research due to insufficient statistical power. Research on these nested structures would provide valuable information in future research but would be very time and resource intensive.

From a practical perspective, before the single-item "health check" could be applied for use in an operational setting, further research is needed to develop the specific feedback content and formats that would be provided to commanders. One factor that will facilitate providing quick analysis and timely feedback is developing an electronic version of the survey. A version that enables Soldiers to respond via handheld devices would be particularly useful by providing easy and flexible participation. Having electronic collection of the responses will provide numerous options for building automated and customizable feedback reports and allow quick identification of target areas for further diagnosis.

Conclusion

The present research provides conceptual support for the use of single-item measurement as a methodological approach for gaining insights into Army command climate. Although reliabilities of particular items could be improved, the single-item measures performed comparably to the multi-item scale measures. In the context of a command climate "health check," single-item measures may provide a way to minimize the burden on Soldiers while providing commensurate levels of information. With additional research and development, single-item measures paired with diagnostic follow-up items have the potential to provide efficient, targeted assessments and actionable information about Army command climate.

References

Adams, M. J., & Umbach, P. D. (2012). Nonresponse and online student evaluations of teaching: Understanding the influence of salience, fatigue, and academic environments. *Research in Higher Education, 5*3, 576–591.

Barling, J., Loughlin, C., & Kelloway, E. K. (2002). Development and test of a model linking safety-specific transformational leadership and occupational safety. *Journal of Applied Psychology, 87,* 488–496.

Bergkvist, L., & Rossiter, J. R. (2007). The predictive validity of multiple-item versus single-item measures of the same constructs. *Journal of Marketing Research, 44,* 175–184.

Berk, M. L., Schur, C. L., & Feldman, J. (2007). Twenty-five years of health surveys: Does more data mean better data? *Health Affairs, 26,* 1599–1611.

Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In K. L. Klein & S. W. J. Kozlowski (Eds), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 349–381). San Francisco, CA: Jossey-Bass.

Bolt, E. E., van der Heide, A., & Onwuteaka-Philipsen, B. D. (2014). Reducing questionnaire length did not improve physician response rate: a randomized trial. *Journal of clinical epidemiology*, *67*(4), 477–481.

Bullis, C., & Reed, G. (2003) *Assessing leaders to establish and maintain positive command climate. Presentation slides.* Department of Command, Leadership and Management, U.S. Army War College. Available at: http://www.academia.edu/8890823/Assessing_Leaders_to_Establish_and_Maintain_a_Positive_Command_Climate_A_Report_to_the_Secretary_of_the_Army

Campbell, J. P., Dunnette, M. D., Lawler, E. E., III, & Weick, K. E. (1970). *Managerial behavior, performance, and effectiveness.* New York, NY: McGraw-Hill.

Carless, S. A., & De Paola, C. (2000). The measurement of cohesion in work teams. *Small Group Research, 31,* 71–88.

Conway, J. M., & Lance, C. E. (2010). What reviewers should expect from authors regarding common method bias in organizational research. *Journal of Business and Psychology, 25,* 325–334.

Craig, W. M., & Brace, R. A. (1992). *Organizational climate: A concept worth consideration.* USAWC Military Studies Program Paper. Carlisle Barracks, PA: U.S. Army War College.

de Boer, A. G. E. M., Van Lanschot, J. J. B., Stalmeier, P. F. M., Van Sandick, J. W., Hulscher, J. B. F., De Haes, J. C. J. M., & Sprangers, M. A. G. (2004). Is a single-item visual analogue scale as valid, reliable and responsive as multi-item scales in measuring quality of life? *Quality of Life Research, 13,* 311–320.

Defense Equal Opportunity Management Institute (2014). *DEOMI Organizational Climate Survey (DEOCS).* Defense Equal Opportunity Management Institute: Patrick AFB, FL.

Defense Human Resources Activity (DHRA). (2015). *Defense Human Resources Activity Privacy Program Playbook (Version 2.0).* Retrieved from http://www.dhra.mil/Portals/52/Documents/Privacy/DHRA_Privacy_Program_Processes_and_Procedures_Signed.pdf?ver=2017-05-17-145019-730

DeSalvo, K. B., Fisher, W. P., Tran, K., Bloser, N., Merrill, W., & Peabody, J. (2006). Assessing measurement properties of two single-item general health measures. *Quality of Life Research, 15,* 191–201.

DeVellis, R. F. (2012). *Scale development: Theory and applications (3rd ed.).* Thousand Oaks, CA: Sage Publications, Inc.

Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P., & Kaiser, S. (2012). Guidelines for choosing between multi-item and single-item scales for construct measurement: A predictive validity perspective. *Journal of the Academy of Marketing Science, 40,* 434–449.

Diamantopoulos, A., & Winklhofer, H. M. (2001). Index construction with formative indicators: An alternative to scale development. *Journal of Marketing Research, 38,* 269–277.

Diener, E. D., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The satisfaction with life scale. *Journal of Personality Assessment, 49*(1), 71–75.

Dolbier, C. L., Webster, J. A., McCalister, K. T., Mallon, M. W., & Steinhardt, M. A. (2005). Reliability and validity of a single-item measure of job satisfaction. *American Journal of Health Promotion, 19,* 194–198.

Doty, J., & Gelineau, J. (2008). Command Climate. *Army,* July, 22–23.

Dragoni, L. (2005). Understanding the emergence of state goal orientation in organizational work groups: The role of leadership and multi-level climate perceptions. *Journal of Applied Psychology, 90,* 1084–1095.

Drolet, A. L., & Morrison, D. G. (2001). Do we really need multiple-item measures in service research? *Journal of Service Research, 3,* 196–204.

Ehrhart, M. G., & Raver, J. L. (2014). The effects of organizational climate and culture on productive and counterproductive behavior. In B. Schneider & K. M. Barbera (Eds.) *The Oxford handbook of organizational climate and culture,* New York, NY: Oxford University Press.

Ehrhart, M. G., Schneider, B., & Macey, W. H. (2014). *Organizational climate and culture: An introduction to theory, research, and practice.* New York, NY: Routledge.

Estrada, A. X., Horgen, K. E., Lentz, E., Severt, J. B., Parks, C., & Fletcher, K. A. (manuscript under review). *Development and test of a model of Army unit cohesion.* Army Research Institute for the Behavioral and Social Sciences.

Fuchs, C., & Diamantopoulos, A. (2009). Using single-item measures for construct measurement in management research: conceptual issues and application guidelines. *Die Betriebswirtschaft, 69,* 195–210.

Gade, P. A., Tiggle, R. B., & Schumm, W. R. (2003). The measurement and consequences of military organizational commitment in Soldiers and spouses. *Military Psychology, 15,* 191–207.

Ginns, P., & Barrie, S. (2004). Reliability of single-item ratings of quality in higher education: A replication. *Psychological Reports, 95,* 1023–1030.

Glick, W. H. (1985). Conceptualizing and measuring organizational and psychological climate: Pitfalls of multilevel research. *Academy of Management Review, 10,* 601–610.

James, L. R. (1982). Aggregation bias in estimates of perceptual agreement. *Journal of Applied Psychology, 67,* 219–229.

James, L. R., Demaree, R. G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology, 69,* 85–98.

James, L. W., & James, L. R. (1989). Integrating work environment perceptions: Explorations into the measurement of meaning. *Journal of Applied Psychology, 74,* 739–751.

Jex, S. M., Sliter, M. T., & Britton, A. (2014). Employee stress and well-being. In B. Schneider & K. M. Barbera (Eds.) *The Oxford handbook of organizational climate and culture,* New York, NY: Oxford University Press.

Johns, J. H., Bickel, M. D., Blades, A. C., Creel, J. B., Gatling, W. S., Hinkle, J. M., Kindred, J. D., & Stocks, S. E. (1984). *Cohesion in the US Military.* Washington, DC: National Defense University Press.

Jones, A. P., & James, L. R. (1979). Psychological climate: Dimensions and relationships of individual and aggregated work environment perceptions. *Organizational Behavior and Human Performance, 23,* 201–250.

Kozlowski, S. W. J., & Klein, K. J. (2000). A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes. In K. J. Klein & S. W. J. Kozlowski (Eds.), *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions* (pp. 3–90). San Francisco, CA: Jossey-Bass.

Kwon, H., & Trail, G. (2005). *The feasibility of single-item measures in sport loyalty research. Sport Management Review, 8,* 69–88.

Langfred, C. W. (2004). Too much of a good thing? Negative effects of high trust and individual autonomy in self-managing teams. *Academy of Management Journal, 47,* 385–399.

Langkamer, K., & Ervin, K. S. (2008). Psychological climate, organizational commitment, and morale: Implications for Army captains' career intent. *Military Psychology, 20,* 219–236.

LeBreton, J. M., Burgess, J. R. D., Kaiser, R. B., Atchley, E. K. P., & James, L. R. (2003). The restriction of variance hypothesis and interrater reliability and agreement: Are ratings from multiple sources really dissimilar? *Organizational Research Methods, 6,* 80–128.

LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods, 11*, 815–852.

Lempke, D. A. (1988). *Command climate: The rise and the decline of a military concept.* Army War College: Carlisle Barracks, PA.

Litwin, G. H., & Stringer, R. A. (1968). *Motivation and organizational climate.* Cambridge, MA: Harvard Business School, Division of Research.

McHugh, J. M. (2013). *Army Directive 2013-29 (Army Command Climate Assessments) [Memorandum]*. Retrieved from http://sill-www.army.mil/eo/docs/2a.pdf

Meade, A. W., & Craig, S. B. (2012, April 16). Identifying careless responses in survey data. *Psychological Methods.* Advance online publication. doi: 10.1037/a0028085

Moriarty, K. O., Campbell, R. O., Heffner, T., & Knapp, D., J. (2009). *Validating future force performance measures (Army class): Reclassification test and criterion development. Research Product 2009-11*. U.S. Army Research Institute for the Behavioral and Social Sciences: Arlington, VA.

Murphy, K. R., & Myors, B. (1998). *Statistical power analysis: A simple and general model for traditional and modern hypothesis tests.* Mahwah, NJ: Lawrence Erlbaum.

Murray, J. M. (2003). *The key to unit effectiveness—A supportive organizational climate.* Army War College: Carlisle Barracks, PA.

Nagy, M. S. (2002). Using a single-item approach to measure facet job satisfaction. *Journal of Occupational and Organizational Psychology*, *75*, 77–86.

Nunnally, J. C., Jr. (1970). *Introduction to psychological measurement.* New York, NY: McGraw-Hill.

Nunnally, J. C., Jr., & Bernstein, I. H. (1994). *Psychometric theory (3rd ed.).* New York, NY: McGraw-Hill.

Oshagbemi, T. (1999). Overall job satisfaction: How good are single versus multiple-item measures? *Journal of Managerial Psychology*, *14*, 388–403.

Payne, R. L., & Mansfield, R. (1973). Relationships of perceptions of organizational climate to organizational structure, context, and hierarchical position. *Administrative Science Quarterly, 18,* 515–526.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*, 879–903.

Pritchard, R. D., & Karasick, B. (1973). The effects of organizational climate on managerial job performance and job satisfaction. *Organizational Behavior and Human Performance, 9,* 126–146.

Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin, 27,* 151–161.

Rossiter, J. R. (2002). The C-OAR-SE procedure for scale development in marketing. *International Journal of Research in Marketing, 19,* 305–335.

Roszkowski, M. J., & Soven, M. (2010). Shifting gears: Consequences of including two negatively worded items in the middle of a positively worded questionnaire. *Assessment & Evaluation in Higher Education, 35,* 113–130.

Sackett, P. R., & Larson, J. R., Jr. (1990). *Research strategies and tactics in industrial and organizational psychology*. In M. D. Dunnette and L. M. Hough (Eds.), *Handbook of industrial and organization psychology* (419–489). Palo Alto: CA.

Schneider, B. (1990). The climate for service: An application of the climate construct. In B. Schneider (Ed.), *Organizational climate and culture* (pp. 383–412). San Francisco, CA: Jossey- Bass.

Schneider, B., & Bartlett, C. J. (1968). Individual differences and organizational climate, I: The research plan and questionnaire development. *Personnel Psychology, 21,* 323–333.

Schneider, B., & Bartlett, C. J. (1970). Individual differences and organizational climate, II: Measurement of organizational climate by the multitrait-multirater matrix. *Personnel Psychology, 23,* 493–512.

Schneider, B., Ehrhart, M. G., Mayer, D. M., Saltz, J. L., & Niles-Jolly, K. (2005). Understanding organization-customer links in service settings. *Academy of Management Journal, 48,* 1017–1032.

Schneider, B., & Reichers, A. E. (1983). On the etiology of climates. *Personnel Psychology, 36,* 19–39.

Schneider, B., White, S. S., & Paul, M. C. (1998). Linking service climate and customer perceptions of service quality: Tests of a causal model. *Journal of Applied Psychology, 83,* 150–163.

Shamir, B., & Kark, R. (2004). A single-item graphic scale for the measurement of organizational identification. *Journal of Occupational and Organizational Psychology, 77,* 115–123.Siebold, G. L., & Kelly, D. R. (1988). *Development of the Platoon Cohesion Index (PCI). Technical Report 816.* U.S. Army Research Institute for the Behavioral and Social Sciences: Alexandria, VA.

Simmons, T. L., & Peterson, R. S. (2000). Task conflict and relationship conflict in top management teams: The pivotal role of intragroup trust. *Journal of Applied Psychology, 85,* 102–111.

Tripathi, K. N., & Agarwal, A. (1988) Effect of reward contingency on motivation. *Journal of General Psychology, 115*, 241–246.

Truhon, S. A., & Parks, K. M. (2005). *Norming the DEOMI Organizational Climate Survey*. Defense Equal Opportunity Management Installation: Patrick Air Force Base, FL.

Turnage, J. J., & Muchinsky, P. M. (1976). The effect of reward contingency and participative decision making on intrinsically and extrinsically motivating tasks. *Academy of Management Journal, 19*, 482–489.

U.S. Department of the Army (1990). Field manual 22-100, military leadership. *Center for Army Leadership, US Army, Fort Leavenworth, KS*.

U.S. Department of the Army (2012). Army Doctrine Reference Publication (ADRP) 6-22: *Army Leadership*. Headquarters, Washington, D.C.

U.S. Department of the Army. (2014). *Army command policy: Army regulation 600-20*. Retrieved from https://armypubs.army.mil/epubs/DR_pubs/DR_a/pdf/web/r600_20.pdf

van Mierlo, H., Vermunt, J. K., & Rutte, C. G. (2009). Composing group-level constructs from individual-level survey data. *Organizational Research Methods, 12,* 368–392.

Van Ryzin, G. G. (2004). Expectations, performance, and citizen satisfaction with urban services. *Journal of Policy Analysis and Management, 23*, 433–448.

Van Sonderen, E., Sanderman, R., & Coyne, J. C. (2013). Ineffectiveness of reverse wording of questionnaire items: Let's learn from cows in the rain. PLoS One, 8(7, e68967), 1–7.

Wanous, J. P., & Hudy, M. J. (2001). Single-item reliability: A replication and extension. *Organizational Research Methods, 4*, 361–375.

Wanous, J. P., Reichers, A. E., & Hudy, M. J. (1997). Overall job satisfaction: How good are single-item measures? *Journal of Applied Psychology, 82*, 247–252.

Whitman, D. S., Caleo, S., Carpenter, N. C., Horner, M. T., & Bernerth, J. B. (2012). Fairness at the collective level: A meta-analytic examination of the consequences and boundary conditions of organizational justice climate. *Journal of Applied Psychology, 97,* 776–791.

Woods, C. M. (2006). Careless responding to reverse-worded items: Implications for confirmatory factor analysis. *Journal of Psychopathology and Behavioral Assessment, 28,* 186–194.

Appendix A. Dimension Scales and Items

Soldiers were provided the following instructions for completing the Command Climate Scales:

INSTRUCTIONS: Please use the scales provided to answer all questions for the survey. Mark your responses on the scannable form. When the statements refer to your "unit," please think about your Company/Troop/ Battery as your unit.

To what extent to do you <u>agree with the following statements:</u>

Strongly disagree (A)	Disagree (B)	Neither agree nor disagree (C)	Agree (D)	Strongly agree (E)	N/A – Don't know (F)

The multi-item scales can be seen in Table A-1.

Table A-1.
Command Climate Scale Items

1. Flow of Information	
Info2	Important information is shared within the unit/organization
Info3	Information is effectively communicated by the chain of command
Info4	Members of my unit are kept informed of upcoming events and requirements
Info5	Members of my unit understand the information shared by the chain of command
Info6	Members of my unit feel that leaders provide enough information
Info7	Leaders in my unit communicate clearly about the commander's intent
Info8	Leaders in my unit communicate clearly about unit members' job requirements

Table A-1.
Command Climate Scale Items

2. Autonomy	
Auton2	Unit members are given appropriate flexibility to make decisions to successfully accomplish the mission
Auton3	Unit members are given enough flexibility to do their jobs effectively without excessive control from leadership
Auton4	Unit members are given the opportunity to figure out problems on their own
Auton5	Members of my unit are given freedom to do their job without being micromanaged
Auton6	Unit members are allowed to take initiative when appropriate
Auton7	Leaders trust unit members to complete their jobs without giving unnecessary guidance
3. Leadership Openness	
Open2	Members of this unit feel comfortable interacting with unit leaders when needed
Open3	Members of this unit can express their opinions about the unit/organization to leaders without fear of reprisal
Open4	Leaders in this unit listen to concerns brought up by unit members
Open5	Leaders in this unit make the time to hear from the unit members
Open6	Leaders in this unit have an effective open-door policy
4. Respect for the Individual	
Respect2	Leaders ensure that all members of this unit are treated with dignity and respect
Respect3	Leaders actively promote dignity and respect for members of this unit
Respect4	Leaders confront unit member actions that undermine dignity and respect
Respect5	Members of this unit feel that they are valued
Respect6	Leaders take action to show dignity and respect for unit members
Respect7	Members of this unit are valued for their contributions to the unit

Table A-1.
Command Climate Scale Items

5. Hazing	
Hazing2	Members of my unit are forced to perform abusive tasks or actions as a rite of passage
Hazing3	New members are harassed or humiliated as a rite of passage when joining this unit
Hazing4	When unit members are promoted, they are harassed or humiliated as a rite of passage
Hazing5	Members of this unit would report hazing if it happened
6. Bullying	
Bully2	Members of my unit are the target of physical attacks by other unit members
Bully3	Leaders make it clear that bullying has no place in this unit
Bully4	Some members of my unit feel picked on by their peers
Bully5	Leaders take action to stop bullying if it occurs
Bully6	Leaders take action to prevent bullying in the unit
7. Fairness	
Fair2	All unit members are given opportunities in a fair manner
Fair3	Leaders are fair and objective when making promotion recommendations
Fair4	Leaders treat all members of this unit fairly
Fair5	Leaders show favoritism toward particular members of the unit
Fair6	Leaders are objective in deciding who gets rewards
8. Inclusion	
Inclus2	Some members of my unit feel excluded by their peers
Inclus3	Some members of my unit feel excluded by their leaders
Inclus4	My unit forms cliques
Inclus5	Peers in my unit ensure members feel included in unit activities
Inclus6	Leaders in my unit ensure members feel included in unit activities

Table A-1.
Command Climate Scale Items

9. Peer Support for Unit Members	
PSUM2	My peers care about the well-being of others in this unit
PSUM3	My peers show little concern for others in this unit
PSUM4	My peers care about the opinions of others in this unit
PSUM5	Members of my unit pitch in when someone needs help
PSUM6	Members of this unit are more focused on themselves than helping others
PSUM7	Members of this unit offer assistance to each other when someone needs help
10. Leader Support for Unit Members	
LSUM2	Leaders in this unit care about the welfare of unit members
LSUM3	Leaders help members of this unit with problems that come up
LSUM4	Leaders in this unit are there when unit members need them
LSUM5	Leaders in this unit appropriately balance mission requirements with unit members' welfare
11. Leader Support for Families	
LSFam2	Families in this unit receive the information they need
LSFam3	Unit leaders care about the well-being of unit members' families
LSFam4	Leaders in this unit are there for members' families when they need help
LSFam5	Leaders in this unit support the Family Readiness Group (FRG)
LSFam6	The Family Readiness Group (FRG) in this unit is effective
12. Performance Orientation	
PerfOr2	Members of this unit maintain high standards of performance
PerfOr3	This unit takes pride in accomplishing the mission successfully
PerfOr4	Members of my unit expect one another to continually improve their skills and performance
PerfOr5	Members of my unit set goals/standards for each other that exceed company standards
PerfOr6	Members of my unit have high performance expectations for one another
PerfOr7	Leaders in this unit recognize high performance accomplishments
PerfOr8	Leaders in this unit let poor performers slide by

Table A-1.
Command Climate Scale Items

13. Professionalism	
Profes2	Members of my unit maintain high standards of professionalism
Profes3	Members of my unit uphold Army standards
Profes4	Members of my unit expect one another to uphold Army standards
Profes5	Leaders discipline unit members who are not in compliance with Army standards
Profes6	Leaders in my unit emphasize Army values and the Profession of Arms
Profes7	Leaders in my unit lead by example

Soldiers were provided the same instructions and response options to complete the single-item measures as they were to complete the multi-item scales. The single items can be seen in Table A-2.

Table A-2.
Command Climate Single Items

1. Flow of information	Information that is important for my unit's success is communicated effectively throughout the unit.
2. Autonomy	Unit leadership trusts the members of this unit to make decisions.
3. Leadership openness	Members of this unit feel they can approach unit leaders to discuss problems and concerns.
4. Respect for the individual	All members of this unit are treated with dignity and respect.
5. Hazing (from AR 600-20, 4-19)	Members of this unit engage in hazing (i.e., intentionally cause new members to suffer or be exposed to activities that are abusive, humiliating, or harmful as a "rite of passage").
6. Bullying (from AR 600-20, 4-19)	Members of this unit bully other unit members (i.e., exclude or reject other unit members through abusive, humiliating, or harmful behavior).
7. Fairness (i.e., distributive justice)	Outcomes (e.g., promotions, rewards, and developmental opportunities) are distributed to unit members in an objective and impartial manner.
8. Inclusion	Unit members feel included in unit activities.
9. Peer support for unit members	Members of this unit support each other.
10. Leader support for unit members	Leaders in this unit provide members with the support and assistance they need.
11. Leader support for families	Unit leaders provide families with the support and assistance they need.
12. Performance orientation	This unit strives for high performance outcomes.
13. Professionalism	This unit emphasizes Army standards, doctrine, and regulations.

Soldiers were provided with the following instructions to complete the command climate diagnostic items:

INSTRUCTIONS: Please use the scales provided to answer the following questions about Officers and NCOs in your company. Mark your response for each on the scannable form.

If the item does not apply to Officers or NCOs, or is something you do not know, please select answer "F" for "N/A - Don't know."

Not at all (A)	Very little (B)	Somewhat (C)	Quite a bit (D)	To a great extent (E)	N/A–Don't know (F)

Table A-3.
Command Climate Diagnostic Items

1. Flow of Information	
To what extent do **Officers** communicate clearly about the following areas?	
InfoD1	Officers: Training schedules
InfoD2	Officers: Army policies
InfoD3	Officers: Unit's mission purpose
To what extent do **NCOs** communicate clearly about the following areas?	
InfoD4	NCOs: Training schedules
InfoD5	NCOs: Army policies
InfoD6	NCOs: Unit's mission purpose
2. Autonomy	
To what extent do **Officers** do the following?	
AutonD1	Officers: Encourage unit members to learn from their mistakes when appropriate
AutonD2	Officers: Trust unit members to make decisions when appropriate
AutonD3	Officers: Enable unit members to do their jobs independently when appropriate
AutonD4	Officers: Counsel appropriately when unit members make mistakes

Table A-3.
Command Climate Diagnostic Items

To what extent do **NCOs** do the following?	
AutonD5	NCOs: Encourage unit members to learn from their mistakes when appropriate
AutonD6	NCOs: Trust unit members to make decisions when appropriate
AutonD7	NCOs: Enable unit members to do their jobs independently when appropriate
AutonD8	NCOs: Counsel appropriately when unit members make mistakes
3. Leadership Openness	
To what extent do **Officers** do the following?	
OpenD1	Officers: Listen to unit members' ideas and opinions
OpenD2	Officers: Pay attention to problems within the unit
OpenD3	Officers: Actively seek out information about issues within the unit
To what extent do **NCOs** do the following?	
OpenD4	NCOs: Listen to unit members' ideas and opinions
OpenD5	NCOs: Pay attention to problems within the unit
OpenD6	NCOs: Actively seek out information about issues within the unit
4. Hazing	
To what extent do **Officers** do the following?	
HazingD1	Officers: Make it clear hazing has no place in the Army
HazingD2	Officers: Encourage unit members to report hazing
To what extent do **NCOs** do the following?	
HazingD3	NCOs: Make it clear hazing has no place in the Army
HazingD4	NCOs: Encourage unit members to report hazing
5. Bullying	
To what extent have **unit members** been:	
BullyD1	Bullied in online communications
To what extent do **Officers** do the following?	
BullyD2	Officers: Make it clear that bullying has no place in the Army
BullyD3	Officers: Encourage unit members to report bullying

Table A-3.
Command Climate Diagnostic Items

To what extent do **NCOs** do the following?	
BullyD4	NCOs: Make it clear that bullying has no place in the Army
BullyD5	NCOs: Encourage unit members to report bullying
6. Fairness	
To what extent do **Officers** do the following?	
FairD1	Officers: Assign work fairly
FairD2	Officers: Evaluate unit members' work performance fairly
FairD3	Officers: Determine rewards based on merit
To what extent do **NCOs** do the following?	
FairD4	NCOs: Assign work fairly
FairD5	NCOs: Evaluate unit members' work performance fairly
FairD6	NCOs: Determine rewards based on merit
7. Leader Support for Unit Members	
To what extent do **Officers** do the following?	
LSUMD1	Officers: Show concern for the well-being of unit members
LSUMD2	Officers: Assist unit members in addressing personal issues
LSUMD3	Officers: Assist unit members in addressing work-related issues
LSUMD4	Officers: Make sure unit members are aware of support channels at the installation
To what extent do **NCOs** do the following?	
LSUMD5	NCOs: Show concern for the well-being of unit members
LSUMD6	NCOs: Assist unit members in addressing personal issues
LSUMD7	NCOs: Assist unit members in addressing work-related issues
LSUMD8	NCOs: Make sure unit members are aware of support channels at the installation

Table A-3.
Command Climate Diagnostic Items

8. Leader Support for Families	
To what extent do **Officers** do the following?	
LSFamD1	Officers: Assist in obtaining the resources families need for PCS moves
LSFamD2	Officers: Assist in obtaining resources families need for their health
LSFamD3	Officers: Provide resources to families during deployments
To what extent do **NCOs** do the following?	
LSFamD4	NCOs: Assist in obtaining the resources families need for PCS moves
LSFamD5	NCOs: Assist in obtaining resources families need for their health
LSFamD6	NCOs: Provide resources to families during deployments
9. Performance Orientation	
To what extent do **Officers** do the following?	
PerfOrD1	Officers: Motivate unit members to perform their job well
PerfOrD2	Officers: Motive unit members to reach new goals and challenges
PerfOrD3	Officers: Allow unit members to just meet the minimum requirements when conducting a task
To what extent do **NCOs** do the following?	
PerfOrD4	NCOs: Motivate unit members to perform their job well
PerfOrD5	NCOs: Motive unit members to reach new goals and challenges
PerfOrD6	NCOs: Allow unit members to just meet the minimum requirements when conducting a task
10. Professionalism	
To what extent do **Officers** do the following?	
ProfesD1	Officers: Enforce Army standards
ProfesD2	Officers: Promote good order and discipline
To what extent do **NCOs** do the following?	
ProfesD3	NCOs: Enforce Army standards
ProfesD4	NCOs: Promote good order and discipline

Appendix B. Full Scale and Single Item Statistics

Table B-1

Item-Level Descriptive Statistics for 13 Command Climate Dimensions

				Skewness		Kurtosis		%
Item	*N*	Mean	*SD*	Statistic	*SE*	Statistic	*SE*	'NA/Don't know'
Flow of information								
Info1*	1318	3.08	1.23	-0.18	0.07	-1.05	0.14	1.53
Info2	1313	3.11	1.14	-0.28	0.07	-0.90	0.14	1.68
Info3	1320	3.03	1.15	-0.16	0.07	-0.93	0.14	1.24
Info4	1322	3.12	1.15	-0.35	0.07	-0.89	0.13	1.09
Info5	1315	3.33	1.05	-0.59	0.07	-0.26	0.14	1.60
Info6	1312	2.88	1.12	-0.03	0.07	-0.89	0.14	1.82
Info7	1314	3.21	1.14	-0.44	0.07	-0.68	0.14	1.53
Info8	1312	3.31	1.11	-0.60	0.07	-0.44	0.14	1.82
Autonomy								
Auton1*	1304	3.11	1.22	-0.30	0.07	-0.93	0.14	2.26
Auton2	1308	3.05	1.15	-0.23	0.07	-0.82	0.14	1.97
Auton3	1315	2.88	1.23	-0.05	0.07	-1.06	0.14	1.46
Auton4	1315	3.22	1.15	-0.43	0.07	-0.68	0.14	1.46
Auton5	1312	2.73	1.28	0.05	0.07	-1.20	0.14	1.97
Auton6	1314	3.33	1.11	-0.53	0.07	-0.45	0.14	1.97
Auton7	1308	3.07	1.15	-0.27	0.07	-0.81	0.14	1.89
Leader openness								
Open1*	1287	3.28	1.26	-0.43	0.07	-0.86	0.14	2.26
Open2	1295	3.36	1.16	-0.51	0.07	-0.58	0.14	1.89
Open3	1288	2.96	1.26	-0.09	0.07	-1.04	0.14	2.69
Open4	1290	3.18	1.20	-0.35	0.07	-0.80	0.14	2.18
Open5	1287	3.17	1.19	-0.32	0.07	-0.80	0.14	2.69
Open6	1260	3.42	1.21	-0.54	0.07	-0.56	0.14	4.44

Note. *Single-item measure. SD = Standard Deviation; SE = Standard Error.

Table B-1.

Item-Level Descriptive Statistics for 13 Command Climate Dimensions (continued)

				Skewness		Kurtosis		%
Item	*N*	Mean	*SD*	Statistic	*SE*	Statistic	*SE*	'NA/Don't know'
Fairness								
Fair1*	1247	3.06	1.17	-0.25	0.07	-0.74	0.14	5.24
Fair2	1264	3.08	1.21	-0.24	0.07	-0.93	0.14	2.62
Fair3	1257	3.14	1.23	-0.34	0.07	-0.88	0.14	4.30
Fair4	1260	3.02	1.24	-0.16	0.07	-1.02	0.14	2.69
Fair5	1267	2.66	1.27	0.31	0.07	-0.94	0.14	3.35
Fair6	1240	3.23	1.09	-0.29	0.07	-0.36	0.14	6.55
Peer support for unit members								
PSUM1*	1304	3.51	1.13	-0.69	0.07	-0.21	0.14	1.89
PSUM2	1306	3.54	1.10	-0.73	0.07	-0.05	0.14	2.18
PSUM3	1306	3.36	1.15	-0.35	0.07	-0.72	0.14	2.26
PSUM4	1300	3.23	1.08	-0.38	0.07	-0.44	0.14	2.91
PSUM5	1297	3.47	1.10	-0.74	0.07	-0.03	0.14	2.55
PSUM6	1302	2.97	1.20	-0.09	0.07	-0.91	0.14	2.26
PSUM7	1301	3.51	1.06	-0.75	0.07	0.11	0.14	2.04
Leader support for unit members								
LSUM1*	1299	3.43	1.05	-0.66	0.07	-0.09	0.14	2.33
LSUM2	1299	3.44	1.09	-0.70	0.07	-0.15	0.14	2.26
LSUM3	1295	3.51	1.04	-0.80	0.07	0.21	0.14	2.48
LSUM4	1298	3.45	1.05	-0.65	0.07	-0.04	0.14	2.48
LSUM5	1288	3.01	1.23	-0.18	0.07	-0.94	0.14	3.50

Note. *Single-item measure. SD = Standard Deviation; SE = Standard Error.

Table B-1

Item-Level Descriptive Statistics for 13 Command Climate Dimensions (continued)

				Skewness		Kurtosis		%
Item	*N*	Mean	*SD*	Statistic	*SE*	Statistic	*SE*	'NA/Don't know'
Leader support for family								
LSFam1*	1110	3.24	1.17	-0.45	0.07	-0.60	0.15	12.60
LSFam2	1109	3.17	1.18	-0.36	0.07	-0.72	0.15	14.06
LSFam3	1116	3.22	1.19	-0.46	0.07	-0.63	0.15	11.43
LSFam4	1111	3.25	1.14	-0.50	0.07	-0.45	0.15	13.55
LSFam5	1110	3.42	1.13	-0.58	0.07	-0.27	0.15	12.45
LSFam6	1087	3.08	1.22	-0.23	0.07	-0.82	0.15	15.44
Performance orientation								
PerfOr1*	1296	3.93	1.05	-1.09	0.07	0.82	0.14	2.11
PerfOr2	1305	3.53	1.11	-0.58	0.07	-0.27	0.14	1.82
PerfOr3	1301	3.82	1.08	-0.96	0.07	0.49	0.14	2.04
PerfOr4	1296	3.73	1.07	-0.83	0.07	0.25	0.14	2.26
PerfOr5	1289	3.51	1.14	-0.58	0.07	-0.35	0.14	3.13
PerfOr6	1295	3.60	1.10	-0.71	0.07	-0.09	0.14	2.84
PerfOr7	1296	3.45	1.20	-0.57	0.07	-0.58	0.14	2.55
PerfOr8	1292	3.00	1.29	-0.04	0.07	-1.11	0.14	2.69
Professionalism								
Profes1*	1311	3.52	1.08	-0.73	0.07	-0.01	0.14	1.68
Profes2	1316	3.32	1.12	-0.47	0.07	-0.51	0.14	1.53
Profes3	1317	3.40	1.09	-0.58	0.07	-0.30	0.14	1.60
Profes4	1316	3.58	1.02	-0.82	0.07	0.37	0.14	1.60
Profes5	1304	3.48	1.12	-0.66	0.07	-0.24	0.14	2.26
Profes6	1304	3.48	1.07	-0.64	0.07	-0.11	0.14	2.11
Profes7	1309	3.12	1.28	-0.28	0.07	-0.98	0.14	1.68

Note. *Single-item measure. SD = Standard Deviation; SE = Standard Error.

Table B-1

Item-Level Descriptive Statistics for 13 Command Climate Dimensions (continued)

				Skewness		Kurtosis		%
Item	*N*	Mean	*SD*	Statistic	*SE*	Statistic	*SE*	'NA/Don't know'
Respect								
Respect1*	1308	3.11	1.25	-0.29	0.07	-0.98	0.14	1.89
Respect2	1312	3.17	1.20	-0.37	0.07	-0.82	0.14	1.68
Respect3	1312	3.26	1.17	-0.48	0.07	-0.61	0.14	1.46
Respect4	1303	3.26	1.14	-0.42	0.07	-0.59	0.14	2.04
Respect5	1295	2.95	1.21	-0.15	0.07	-0.96	0.14	2.77
Respect6	1307	3.22	1.14	-0.43	0.07	-0.58	0.14	1.68
Respect7	1303	3.12	1.21	-0.31	0.07	-0.85	0.14	2.33
Hazing								
Hazing1*	1264	1.98	1.10	0.98	0.07	0.15	0.14	3.35
Hazing2	1285	1.80	0.96	1.19	0.07	1.04	0.14	3.35
Hazing3	1288	1.83	1.00	1.21	0.07	1.00	0.14	3.28
Hazing4	1282	1.80	0.97	1.21	0.07	0.99	0.14	3.86
Hazing5	1251	2.57	1.32	0.51	0.07	-0.86	0.14	5.61
Bullying								
Bully1*	1247	2.14	1.11	0.76	0.07	-0.20	0.14	3.50
Bully2	1246	1.87	0.95	1.03	0.07	0.67	0.14	3.35
Bully3	1249	2.52	1.27	0.61	0.07	-0.65	0.14	3.28
Bully4	1228	2.53	1.19	0.28	0.07	-0.92	0.14	6.19
Bully5	1249	2.51	1.22	0.63	0.07	-0.44	0.14	5.03
Bully6	1242	2.46	1.16	0.66	0.07	-0.26	0.14	5.10
Inclusion								
Inclus1*	1204	3.49	1.06	-0.64	0.07	-0.04	0.14	4.30
Inclus2	1209	2.72	1.09	0.11	0.07	-0.69	0.14	7.65
Inclus3	1212	2.78	1.14	0.11	0.07	-0.77	0.14	7.50
Inclus4	1207	2.94	1.26	0.03	0.07	-0.99	0.14	4.66
Inclus5	1209	3.45	1.03	-0.57	0.07	0.00	0.14	3.86
Inclus6	1201	3.46	1.06	-0.66	0.07	0.02	0.14	3.50

Note. *Single-item measure. SD = Standard Deviation; SE = Standard Error.

Table B-2

Intercorrelation Matrix for Full Scales at Company and Individual Level

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Flow of information	—	.87**	.86**	.83**	.72**	.80**	.73**	.80**	.85**	-.32**	-.66**	.74**	.80**
2. Autonomy	.61**	—	.89**	.83**	.73**	.79**	.69**	.72**	.83**	-.44**	-.63**	.72**	.74**
3. Leader Openness	.57**	.69**	—	.84**	.81**	.86**	.74**	.76**	.91**	-.42**	-.73**	.79**	.74**
4. Fairness	.56**	.63**	.70**	—	.77**	.78**	.69**	.75**	.85**	-.41**	-.65**	.79**	.77**
5. Peer Support for Unit Members	.49**	.54**	.62**	.58**	—	.83**	.75**	.75**	.86**	-.38**	-.68**	.78**	.73**
6. Leader Support for Unit Members	.57**	.62**	.74**	.67**	.69**	—	.81**	.82**	.88**	-.38**	-.74**	.69**	.81**
7. Leader Support for Family	.53**	.60**	.64**	.60**	.56**	.68**	—	.78**	.75**	-.30*	-.63**	.65**	.75**
8. Performance Orientation	.56**	.56**	.61**	.57**	.62**	.65**	.62**	—	.81**	-.38**	-.66**	.71**	.93**
9. Respect for the Individual	.56**	.67**	.73**	.73**	.63**	.74**	.65**	.63**	—	-.49**	-.80**	.82**	.82**
10. Hazing	-.13**	-.18**	-.26**	-.24**	-.27**	-.28**	-.21**	-.23**	-.30**	—	.64**	-.45**	-.36*
11. Bullying	-.39**	-.40**	-.48**	-.45**	-.46**	-.52**	-.45**	-.48**	-.55**	.38**	—	-.74**	-.62**
12. Inclusion	.36**	.48**	.54**	.55**	.49**	.51**	.47**	.43**	.60**	-.39**	-.47**	—	.73**
13. Professionalism	.57**	.61**	.63**	.66**	.61**	.68**	.65**	.74**	.73**	-.26**	-.47**	.51**	—

Note. $^{**}p < .01$. $^{*}p < .05$. Individual-level correlations are below the diagonal. Company-level correlations are above the diagonal.

Table B-3

Intercorrelation Matrix for Single Items at Company and Individual Level

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Flow of information	—	.79**	.75**	.70**	.68**	.66**	.66**	.65**	.62**	.74**	-.38**	-.54**	.60**
2. Autonomy	.47**	—	.80**	.77**	.72**	.78**	.81**	.69**	.66**	.78**	-.38**	-.68**	.74**
3. Leader openness	.41**	.53**	—	.78**	.77**	.82**	.73**	.67**	.64**	.84**	-.38**	-.65**	.72**
4. Fairness	.40**	.44**	.42**	—	.73**	.73**	.68**	.63**	.69**	.77**	-.39**	-.55**	.74**
5. Peer support for unit members	.39**	.44**	.48**	.37**	—	.81**	.80**	.64**	.58**	.82**	-.39**	-.61**	.76**
6. Leader support for unit members	.45**	.54**	.59**	.49**	.59**	—	.85**	.70**	.66**	.80**	-.43**	-.65**	.78**
7. Leader support for families	.36**	.49**	.48**	.42**	.50**	.60**	—	.70**	.66**	.73**	-.36**	-.65**	.75**
8. Performance orientation	.35**	.36**	.37**	.31**	.45**	.48**	.42**	—	.77**	.68**	-.45**	-.64**	.62**
9. Professionalism	.35**	.41**	.37**	.34**	.42**	.46**	.47**	.45**	—	.65**	-.29*	-.61**	.61**
10. Respect for the individual	.44**	.51**	.54**	.47**	.49**	.61**	.51**	.38**	.41**	—	-.55**	-.81**	.80**
11. Hazing	-.17**	-.14**	-.25**	-.21**	-.20**	-.27**	-.21**	-.22**	-.19**	-.32**	—	.57**	-.47**
12. Bullying	-.20**	-.21**	-.24**	-.19**	-.23**	-.31**	-.22**	-.20**	-.23**	-.39**	.54**	—	-.67**
13. Inclusion	.33**	.40**	.43**	.36**	.47**	.53**	.50**	.37**	.43**	.56**	-.32**	-.31**	—

Note. **$p < .01$. *$p < .05$. Individual-level correlations are below the diagonal. Company-level correlations are above the diagonal.

Table B-4

Exploratory Factor Analysis Factor Loadings (Pattern Matrix) and Communalities

	2	3	1	4	5	6	7	8	9	10	Communalities	
	Respect (and profes, fairness, inclusion)	Information	PerfOr (and profes)	Autonomy	LSFam	Leader open (and fairness, leader support for unit member)	Peer support (and leader support)	Hazing	Bullying	Inclusion	Initial	Extraction
PerfOr2			.85								.69	.66
PerfOr3			.79								.69	.66
PerfOr4			.91								.74	.73
PerfOr5			.81								.71	.67
PerfOr6			.75								.66	.64
PerfOr7			.37								.61	.57
Profes2	.42		.64								.77	.73
Profes3	.45		.57								.78	.71
Profes4			.70								.71	.66
Profes5	.33		.50								.59	.54
Profes6	.43		.64								.73	.70
Profes7	.53		.32								.72	.70
Respect2	.95										.82	.79
Respect3	.86										.83	.78
Respect4	.84										.76	.72
Respect5	.87										.75	.73
Respect6	.91										.80	.80
Respect7	.84										.78	.76

Table B-4

Exploratory Factor Analysis Factor Loadings (Pattern Matrix) and Communalities (continued)

	2	3	1	4	5	6	7	8	9	10	Communalities	
	Respect (and profes, fairness, inclusion)	Information	PerfOr (and profes)	Autonomy	LSFam	Leader open (and fairness, leader support for unit member)	Peer support (and leader support)	Hazing	Bullying	Inclusion	Initial	Extraction
Inclus2_*R*										.85	.66	.73
Inclus3_*R*										.89	.70	.83
Inclus4_*R*										.53	.42	.42
Inclus5	.36										.69	.51
Inclus6	.31										.74	.60
Info2		.76									.68	.68
Info3		.80									.70	.69
Info4		.83									.66	.66
Info5		.88									.64	.66
Info6		.73									.62	.64
Info7		.75									.66	.65
Info8		.72									.63	.60
Auton2				.70							.69	.69
Auton3				.86							.75	.75
Auton4				.74							.67	.68
Auton5				.89							.70	.73
Auton6				.67							.64	.63
Auton7				.68							.68	.69

Table B-4

Exploratory Factor Analysis Factor Loadings (Pattern Matrix) and Communalities (continued)

	2	3	1	4	5	6	7	8	9	10	Communalities	
	Respect (and profes, fairness, inclusion)	Information	PerfOr (and profes)	Autonomy	LSFam	Leader open (and fairness, leader support for unit member)	Peer support (and leader support)	Hazing	Bullying	Inclusion	Initial	Extraction
Open2						.74					.65	.63
Open3						.79					.68	.67
Open4						.87					.76	.76
Open5						.86					.75	.75
Open6						.72					.63	.62
Fair2	.38										.67	.56
Fair3	.39					.35					.70	.57
Fair4	.45					.35					.70	.62
PSUM2							.84				.65	.71
PSUM4							.60				.53	.50
PSUM5							.82				.65	.65
PSUM7							.75				.66	.66
LSUM2						.31					.75	.70
LSUM3							.37				.77	.70
LSUM4						.31	.30				.77	.71

Table B-4

Exploratory Factor Analysis Factor Loadings (Pattern Matrix) and Communalities (continued)

	2	3	1	4	5	6	7	8	9	10	Communalities	
	Respect (and profes, fairness, inclusion)	Information	PerfOr (and profes)	Autonomy	LSFam	Leader open (and fairness, leader support for unit member)	Peer support (and leader support)	Hazing	Bullying	Inclusion	Initial	Extraction
LSFam2					.88						.70	.72
LSFam3					.76						.81	.80
LSFam4					.75						.79	.77
LSFam5					.94						.68	.69
LSFam6					.94						.68	.68
Hazing2								-.94			.80	.86
Hazing3								-.93			.80	.85
Hazing4								-.92			.77	.80
Bully3_*R*									-.69		.51	.52
Bully5_*R*									-.89		.62	.71
Bully6_*R*									-.75		.65	.69

Note: Loadings under .30 are suppressed.

Table B-5

Factor Intercorrelation Matrix

Factor	1	2	3	4	5	6	7	8	9
1. PerfOr (and profes)	—								
2. Respect (and profes, fairness, inclusion)	.73	—							
3. Information	.58	.60	—						
4. Autonomy	.59	.69	.62	—					
5. LSFam	.68	.73	.57	.63	—				
6. Leader open (and fairness, leader support for unit member)	.66	.79	.62	.71	.72	—			
7. Peer support (and leader support)	.68	.70	.56	.57	.66	.72	—		
8. Hazing	.31	.39	.16	.23	.27	.30	.36	—	
9. Bullying	.52	.57	.41	.40	.52	.55	.60	.47	—
10. Inclusion	.20	.38	.15	.32	.25	.34	.26	.34	.28

Appendix C. Outcomes

Table C-1

Intraclass Correlations for Dependent Variables

	ICC(1)	F for ICC(1)	ICC(2)
Cohesion	0.16	6.02*	0.83
Unit resilience	0.13	4.74*	0.79
Unit performance	0.14	5.30*	0.81
Affective commitment	0.04	2.14*	0.53
Continuance commitment	0.01	1.28	0.22
Family intention to stay	0.03	1.80*	0.44
Job satisfaction	0.06	2.69*	0.63
Well-being	0.02	1.48*	0.33
Withdrawal	0.05	2.30*	0.56
Social participation	0.04	2.07*	0.52
Family satisfaction	0.01	1.37*	0.27
Mutual trust	0.08	3.30*	0.70
Morale	0.10	3.86*	0.74
Stress	0.01	1.21	0.17
Career intentions	0.03	1.86*	0.46
Weapons qualification	0.28	11.05*	0.91
APFT Score	0.05	2.31*	0.57
Article 15s	0.02	1.48*	0.33

Note. *$p < .05$.

Table C-2

Intercorrelations Among Company-Level Dependent Variables

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1. Unit resilience																			
2. Unit performance	.81**																		
3. Unit performance (officer)	-.05	-.11																	
4. Affective commitment	.67**	.46**	-.15																
5. Continuance commitment	.19	.08	.25	.53**															
6. Family desire to stay	.28*	.15	-.01	.56**	.42**														
7. Job satisfaction	.58**	.56**	.08	.74**	.50**	.45**													
8. Well-being	.56**	.44**	.08	.73**	.50**	.54**	.72**												
9. Withdrawal	-.24	-.36**	.54**	-.28	.13	-.22	-.29*	-.18											
10. Social participation	.36*	.28*	-.17	.48**	.22	.18	.33*	.38**	-.23										
11. Family satisfaction	.52**	.49**	-.07	.53**	.20	.47**	.50**	.70**	-.15	.33*									
12. Mutual trust	.81**	.68**	-.16	.58**	.07	.34*	.46**	.45**	-.27	.38**	.47**								
13. Morale	.88**	.68**	.04	.77**	.40**	.42**	.72**	.73**	-.20	.40**	.64**	.77**							
14. Cohesion	.93**	.80**	-.05	.64**	.20	.30*	.58**	.57**	-.28*	.35*	.58**	.88**	.90**						

15. Weapons qualification	-.04	.03	-.13	-.13	-.29*	-.20	-.21	-.25	-.03	.27	-.09	.21	-.15	.03					
16. APFT score	.03	.08	-.31*	.14	.07	-.20	.05	.01	-.18	.27	.04	-.07	.02	.05	.18				
17. Article 15s	.29*	.32*	-.28	.27	.10	-.08	.14	.10	-.16	.20	.09	.22	.16	.26	.20	.23			
18. Career intentions	.28*	.29*	-.07	.68**	.50**	.50**	.62**	.62**	-.17	.29*	.41**	.23	.48**	.33*	-.28*	.03	.23		
19. Stress	-.43**	-.29*	.13	-.34*	-.01	-.10	-.34*	-.45**	.28*	-.23	-.29*	-.27	-.44**	-.42**	.15	.06	-.02	-.33*	

Note. * = $p < .05$, ** = $p < .01$.

Appendix D. Item Level Descriptives

Table D-1

Item-Level Descriptive Statistics for Diagnostic Items (Individual Level)

				Skewness		Kurtosis	
	N	Mean	*SD*	Statistic	*SE*	Statistic	*SE*
Info1	1207	3.20	1.13	-0.26	0.07	-0.64	0.14
Info2	1210	3.16	1.16	-0.20	0.07	-0.73	0.14
Info3	1215	3.30	1.17	-0.34	0.07	-0.67	0.14
Info4	1254	3.49	1.09	-0.52	0.07	-0.35	0.14
Info5	1251	3.52	1.11	-0.54	0.07	-0.36	0.14
Info6	1253	3.43	1.16	-0.42	0.07	-0.61	0.14

				Skewness		Kurtosis	
	N	Mean	*SD*	Statistic	*SE*	Statistic	*SE*
Auto1	1210	3.32	1.12	-0.47	0.07	-0.40	0.14
Auto2	1199	3.25	1.13	-0.36	0.07	-0.58	0.14
Auto3	1203	3.35	1.10	-0.48	0.07	-0.42	0.14
Auto4	1153	3.13	1.23	-0.31	0.07	-0.88	0.14
Auto5	1248	3.62	1.11	-0.73	0.07	-0.07	0.14
Auto6	1248	3.42	1.12	-0.52	0.07	-0.37	0.14
Auto7	1253	3.45	1.14	-0.53	0.07	-0.48	0.14
Auto8	1243	3.55	1.18	-0.64	0.07	-0.38	0.14

				Skewness		Kurtosis	
	N	Mean	*SD*	Statistic	*SE*	Statistic	*SE*
Open1	1175	3.20	1.18	-0.33	0.07	-0.71	0.14
Open2	1183	3.31	1.17	-0.43	0.07	-0.63	0.14
Open3	1167	3.18	1.18	-0.30	0.07	-0.74	0.14
Open4	1246	3.41	1.12	-0.46	0.07	-0.45	0.14
Open5	1249	3.45	1.10	-0.52	0.07	-0.34	0.14
Open6	1238	3.38	1.13	-0.42	0.07	-0.50	0.14

				Skewness		Kurtosis	
	N	Mean	*SD*	Statistic	*SE*	Statistic	*SE*
Fair1	1115	3.31	1.14	-0.51	0.07	-0.41	0.15
Fair2	1114	3.30	1.13	-0.46	0.07	-0.46	0.15
Fair3	1091	3.16	1.17	-0.35	0.07	-0.66	0.15
Fair4	1240	3.31	1.17	-0.42	0.07	-0.64	0.14
Fair5	1236	3.39	1.12	-0.57	0.07	-0.33	0.14
Fair6	1201	3.28	1.16	-0.43	0.07	-0.56	0.14

				Skewness		Kurtosis	
	N	Mean	*SD*	Statistic	*SE*	Statistic	*SE*
LSUM1	1191	3.39	1.15	-0.55	0.07	-0.40	0.14
LSUM2	1144	3.23	1.17	-0.37	0.07	-0.66	0.15
LSUM3	1150	3.34	1.12	-0.51	0.07	-0.39	0.14
LSUM4	1157	3.43	1.14	-0.55	0.07	-0.41	0.14
LSUM5	1242	3.60	1.08	-0.71	0.07	-0.01	0.14
LSUM6	1246	3.59	1.09	-0.68	0.07	-0.10	0.14
LSUM7	1247	3.67	1.07	-0.73	0.07	0.06	0.14
LSUM8	1242	3.65	1.08	-0.67	0.07	-0.11	0.14

				Skewness		Kurtosis	
	N	Mean	*SD*	Statistic	*SE*	Statistic	*SE*
LSFam1	895	2.82	1.25	0.01	0.08	-1.03	0.16
LSFam2	929	2.99	1.24	-0.16	0.08	-0.96	0.16
LSFam3	783	3.04	1.23	-0.22	0.09	-0.82	0.18
LSFam4	976	3.40	1.12	-0.54	0.08	-0.30	0.16
LSFam5	1006	3.44	1.14	-0.55	0.08	-0.37	0.15
LSFam6	829	3.39	1.13	-0.50	0.09	-0.37	0.17

				Skewness		Kurtosis	
	N	Mean	*SD*	Statistic	*SE*	Statistic	*SE*
PerfOr1	1196	3.39	1.13	-0.51	0.07	-0.39	0.14
PerfOr2	1191	3.36	1.16	-0.47	0.07	-0.54	0.14
PerfOr3	1174	2.77	1.19	0.09	0.07	-0.93	0.14
PerfOr4	1246	3.66	1.10	-0.78	0.07	0.09	0.14
PerfOr5	1243	3.64	1.10	-0.73	0.07	-0.04	0.14
PerfOr6	1237	2.75	1.25	0.12	0.07	-1.02	0.14

				Skewness		Kurtosis	
	N	Mean	*SD*	Statistic	*SE*	Statistic	*SE*
Profes1	1204	3.42	1.13	-0.46	0.07	-0.47	0.14
Profes2	1204	3.50	1.11	-0.57	0.07	-0.32	0.14
Profes3	1253	3.71	1.09	-0.84	0.07	0.25	0.14
Profes4	1246	3.69	1.09	-0.75	0.07	0.05	0.14

				Skewness		Kurtosis	
	N	Mean	*SD*	Statistic	*SE*	Statistic	*SE*
Hazing1	1200	3.75	1.21	-0.82	0.07	-0.18	0.14
Hazing2	1198	3.75	1.21	-0.81	0.07	-0.22	0.14
Hazing3	1228	3.73	1.23	-0.79	0.07	-0.30	0.14
Hazing4	1230	3.69	1.26	-0.75	0.07	-0.45	0.14

				Skewness		Kurtosis	
	N	Mean	*SD*	Statistic	*SE*	Statistic	*SE*
Bully1	978	1.88	1.15	1.17	0.08	0.40	0.16
Bully2	1165	3.70	1.21	-0.75	0.07	-0.31	0.14
Bully3	1168	3.68	1.22	-0.75	0.07	-0.35	0.14
Bully4	1201	3.70	1.22	-0.77	0.07	-0.33	0.14
Bully5	1198	3.67	1.24	-0.73	0.07	-0.41	0.14

Note. SD = Standard Deviation. SE = Standard Error.

Table D-2

ICCs for Diagnostic Items

	ICC(1)	F for ICC(1)	ICC(2)
InfoD1	0.07	3.09	0.68
InfoD2	0.04	2.12	0.53
InfoD3	0.08	3.24	0.69
InfoD4	0.06	2.66	0.62
InfoD5	0.04	2.02	0.51
InfoD6	0.05	2.37	0.58
AutonD1	0.09	3.42	0.71
AutonD2	0.08	3.40	0.71
AutonD3	0.08	3.37	0.70
AutonD4	0.04	2.10	0.52
AutonD5	0.05	2.40	0.58
AutonD6	0.05	2.26	0.56
AutonD7	0.05	2.40	0.58
AutonD8	0.03	1.91	0.48
OpenD1	0.10	4.05	0.75
OpenD2	0.09	3.66	0.73
OpenD3	0.08	3.20	0.69
OpenD4	0.07	3.03	0.67
OpenD5	0.06	2.80	0.64
OpenD6	0.04	2.19	0.54
FairD1	0.08	3.29	0.70
FairD2	0.06	2.80	0.64
FairD3	0.07	2.83	0.65
FairD4	0.03	1.78	0.44
FairD5	0.04	2.16	0.54
FairD6	0.05	2.47	0.60
LSUMD1	0.11	4.18	0.76
LSUMD2	0.08	3.21	0.69
LSUMD3	0.09	3.56	0.72
LSUMD4	0.07	2.98	0.66
LSUMD5	0.07	3.10	0.68

LSUMD6	0.06	2.80	0.64
LSUMD7	0.06	2.67	0.62
LSUMD8	0.05	2.44	0.59
LSFamD1	0.04	2.11	0.53
LSFamD2	0.05	2.35	0.57
LSFamD3	0.05	2.49	0.60
LSFamD4	0.04	2.05	0.51
LSFamD5	0.04	1.96	0.49
LSFamD6	0.04	1.96	0.49
PerfOrD1	0.09	3.44	0.71
PerfOrD2	0.08	3.26	0.69
PerfOrD3	0.03	1.78	0.44
PerfOrD4	0.07	2.85	0.65
PerfOrD5	0.05	2.36	0.58
PerfOrD6	0.04	2.18	0.54
ProfesD1	0.07	2.93	0.66
ProfesD2	0.08	3.41	0.71
ProfesD3	0.05	2.39	0.58
ProfesD4	0.06	2.80	0.64
HazingD1	0.08	3.16	0.68
HazingD2	0.07	3.06	0.67
HazingD3	0.05	2.42	0.59
HazingD4	0.05	2.38	0.58
BullyD1	0.02	1.42	0.29
BullyD2	0.06	2.73	0.63
BullyD3	0.05	2.41	0.58
BullyD4	0.04	1.98	0.50
BullyD5	0.04	2.16	0.54

Note. All F values are statistically significant at the $p < .05$ level.

Table D-3

Item-Level Descriptive Statistics for Diagnostic Items (Company Level)

				Skewness		Kurtosis	
	N	Mean	*SD*	Statistic	*SE*	Statistic	*SE*
InfoD1_mean	51	3.20	0.40	-0.19	0.33	-0.71	0.66
InfoD2_mean	51	3.17	0.34	0.08	0.33	-1.15	0.66
InfoD3_mean	51	3.32	0.42	-0.16	0.33	-1.07	0.66
InfoD4_mean	51	3.52	0.37	0.18	0.33	-0.83	0.66
InfoD5_mean	51	3.54	0.33	0.10	0.33	-0.67	0.66
InfoD6_mean	51	3.46	0.38	0.29	0.33	-0.56	0.66

				Skewness		Kurtosis	
	N	Mean	*SD*	Statistic	*SE*	Statistic	*SE*
AutonD1_mean	51	3.34	0.40	-0.56	0.33	-0.30	0.66
AutonD2_mean	51	3.27	0.40	-0.52	0.33	-0.70	0.66
AutonD3_mean	51	3.37	0.39	-0.82	0.33	-0.25	0.66
AutonD4_mean	51	3.16	0.38	-0.10	0.33	-0.19	0.66
AutonD5_mean	51	3.62	0.35	0.07	0.33	-0.02	0.66
AutonD6_mean	51	3.44	0.35	-0.09	0.33	0.10	0.66
AutonD7_mean	51	3.48	0.36	-0.20	0.33	0.14	0.66
AutonD8_mean	51	3.58	0.35	0.46	0.33	-0.09	0.66

				Skewness		Kurtosis	
	N	Mean	*SD*	Statistic	*SE*	Statistic	*SE*
OpenD1_mean	51	3.23	0.46	-0.31	0.33	-0.26	0.66
OpenD2_mean	51	3.33	0.45	-0.43	0.33	0.25	0.66
OpenD3_mean	51	3.19	0.45	-0.31	0.33	-0.37	0.66
OpenD4_mean	51	3.44	0.40	0.25	0.33	-0.60	0.66
OpenD5_mean	51	3.49	0.38	0.02	0.33	-0.60	0.66
OpenD6_mean	51	3.40	0.35	-0.04	0.33	-0.64	0.66

				Skewness		Kurtosis	
	N	Mean	*SD*	Statistic	*SE*	Statistic	*SE*
FairD1_mean	51	3.32	0.41	-0.42	0.33	-0.48	0.66
FairD2_mean	51	3.31	0.38	-0.17	0.33	-0.70	0.66
FairD3_mean	51	3.17	0.43	-0.29	0.33	-0.58	0.66
FairD4_mean	51	3.32	0.35	0.20	0.33	-0.45	0.66
FairD5_mean	51	3.40	0.36	0.32	0.33	-0.43	0.66
FairD6_mean	51	3.28	0.39	0.24	0.33	-0.06	0.66

				Skewness		Kurtosis	
	N	Mean	*SD*	Statistic	*SE*	Statistic	*SE*
LSUMD1_mean	51	3.40	0.46	-0.34	0.33	-0.61	0.66
LSUMD2_mean	51	3.25	0.42	-0.46	0.33	-0.35	0.66
LSUMD3_mean	51	3.36	0.42	-0.38	0.33	-0.36	0.66
LSUMD4_mean	51	3.44	0.40	-0.36	0.33	-0.47	0.66
LSUMD5_mean	51	3.61	0.40	-0.05	0.33	-0.35	0.66
LSUMD6_mean	51	3.61	0.38	-0.07	0.33	-0.59	0.66
LSUMD7_mean	51	3.68	0.36	-0.34	0.33	-0.59	0.66
LSUMD8_mean	51	3.67	0.35	-0.15	0.33	-0.85	0.66

				Skewness		Kurtosis	
	N	Mean	*SD*	Statistic	*SE*	Statistic	*SE*
LSFamD1_mean	51	2.93	0.50	0.41	0.33	0.01	0.66
LSFamD2_mean	51	3.04	0.48	0.14	0.33	-0.36	0.66
LSFamD3_mean	51	3.06	0.51	0.28	0.33	-0.06	0.66
LSFamD4_mean	51	3.43	0.40	-0.06	0.33	-0.68	0.66
LSFamD5_mean	51	3.45	0.39	-0.12	0.33	-0.44	0.66
LSFamD6_mean	51	3.42	0.39	0.31	0.33	-0.61	0.66

				Skewness		Kurtosis	
	N	Mean	*SD*	Statistic	*SE*	Statistic	*SE*
ProfesD1_mean	51	3.45	0.38	-0.30	0.33	-0.84	0.66
ProfesD2_mean	51	3.52	0.40	-0.31	0.33	-0.88	0.66
ProfesD3_mean	51	3.72	0.34	0.08	0.33	-0.64	0.66
ProfesD4_mean	51	3.71	0.38	-0.03	0.33	-0.62	0.66

				Skewness		Kurtosis	
	N	Mean	*SD*	Statistic	*SE*	Statistic	*SE*
HazingD1_mean	51	3.78	0.42	-0.69	0.33	0.71	0.66
HazingD2_mean	51	3.79	0.42	-0.66	0.33	0.63	0.66
HazingD3_mean	51	3.75	0.39	-0.35	0.33	0.38	0.66
HazingD4_mean	51	3.72	0.39	-0.25	0.33	0.12	0.66

				Skewness		Kurtosis	
	N	Mean	*SD*	Statistic	*SE*	Statistic	*SE*
BullyD1_mean	51	1.89	0.32	0.33	0.33	-0.64	0.66
BullyD2_mean	51	3.71	0.42	-0.58	0.33	-0.03	0.66
BullyD3_mean	51	3.68	0.40	-0.48	0.33	-0.27	0.66
BullyD4_mean	51	3.70	0.37	-0.24	0.33	-0.62	0.66
BullyD5_mean	51	3.66	0.39	-0.32	0.33	-0.25	0.66

Note. SD = Standard Deviation. SE = Standard Error.

Table D-4

Intercorrelations Between Diagnostic Items, Dimension Scales, and Single Items (Flow of Information)

	Info Diagnostic 1	Info Diagnostic 2	Info Diagnostic 3	Info Diagnostic 4	Info Diagnostic 5	Info Diagnostic 6	Info scale	Info single item
Info Diagnostic 1	—							
Info Diagnostic 2	.84	—						
Info Diagnostic 3	.83	.83	—					
Info Diagnostic 4	.75	.73	.75	—				
Info Diagnostic 5	.64	.68	.76	.88	—			
Info Diagnostic 6	.68	.72	.80	.89	.90	—		
Information scale	.70	.71	.77	.75	.74	.75	—	
Info single item	.62	.66	.76	.68	.72	.73	.92	—

Note. All correlations statistically significant at the $p < .01$ level.

Table D-5

Intercorrelations Between Diagnostic Items, Dimension Scales and Single Items (Autonomy)

	Auton Diagnostic 1	Auton Diagnostic 2	Auton Diagnostic 3	Auton Diagnostic 4	Auton Diagnostic 5	Auton Diagnostic 6	Auton Diagnostic 7	Auton Diagnostic 8	Auton scale	Auton single item
Auton Diagnostic 1	—									
Auton Diagnostic 2	.90	—								
Auton Diagnostic 3	.83	.91	—							
Auton Diagnostic 4	.79	.76	.71	—						
Auton Diagnostic 5	.67	.62	.57	.60	—					
Auton Diagnostic 6	.68	.70	.63	.60	.85	—				
Auton Diagnostic 7	.63	.69	.70	.57	.81	.91	—			
Auton Diagnostic 8	.58	.66	.59	.58	.80	.80	.84	—		
Auton scale	.83	.83	.75	.63	.76	.80	.78	.69	—	
Auton single item	.79	.85	.76	.65	.73	.73	.73	.68	.86	—

Note. All correlations statistically significant at the $p < .01$ level.

Table D-6

Intercorrelations Between Diagnostic Items, Dimension Scales, and Single Items (Openness)

	Open Diagnostic 1	Open Diagnostic 2	Open Diagnostic 3	Open Diagnostic 4	Open Diagnostic 5	Open Diagnostic 6	Open scale	Open single item
Open Diagnostic 1	—							
Open Diagnostic 2	.94	—						
Open Diagnostic 3	.89	.88	—					
Open Diagnostic 4	.59	.62	.62	—				
Open Diagnostic 5	.60	.62	.59	.87	—			
Open Diagnostic 6	.54	.57	.64	.81	.82	—		
Open scale	.70	.69	.73	.86	.79	.77	—	
Open single item	.61	.61	.71	.80	.74	.77	.93	—

Note. All correlations statistically significant at the $p < .01$ level.

Table D-7

Intercorrelations Between Diagnostic Items, Dimension Scales, and Single Items (Leader Support for Unit Members)

	LSUM Diagnostic 1	LSUM Diagnostic 2	LSUM Diagnostic 3	LSUM Diagnostic 4	LSUM Diagnostic 5	LSUM Diagnostic 6	LSUM Diagnostic 7	LSUM Diagnostic 8	LSUM scale	LSUM single item
LSUM Diagnostic 1	—									
LSUM Diagnostic 2	.94	—								
LSUM Diagnostic 3	.94	.94	—							
LSUM Diagnostic 4	.86	.88	.87	—						
LSUM Diagnostic 5	.73	.77	.76	.73	—					
LSUM Diagnostic 6	.71	.75	.71	.70	.96	—				
LSUM Diagnostic 7	.75	.76	.77	.74	.93	.91	—			
LSUM Diagnostic 8	.69	.70	.71	.80	.86	.85	.87	—		
LSUM scale	.82	.81	.85	.78	.86	.84	.87	.84	—	
LSUM single item	.83	.84	.86	.75	.87	.84	.88	.79	.94	—

Note. All correlations statistically significant at the $p < .01$ level.

Table D-8

Intercorrelations Between Diagnostic Items, Dimension Scales, and Single Items (Leader Support for Family)

	LSFam Diagnostic 1	LSFam Diagnostic 2	LSFam Diagnostic 3	LSFam Diagnostic 4	LSFam Diagnostic 5	LSFam Diagnostic 6	LSFam scale	LSFam single item
LSFam Diagnostic 1	—							
LSFam Diagnostic 2	.92	—						
LSFam Diagnostic 3	.92	.89	—					
LSFam Diagnostic 4	.67	.70	.73	—				
LSFam Diagnostic 5	.64	.67	.72	.91	—			
LSFam Diagnostic 6	.69	.66	.76	.85	.88	—		
LSFam scale	.78	.80	.82	.81	.82	.78	—	
LSFam single item	.76	.77	.79	.82	.87	.81	.94	—

Note. All correlations statistically significant at the $p < .01$ level.

Table D-9

Intercorrelations Between Diagnostic Items, Dimension Scales, and Single Items (Performance Orientation)

	PerfOr Diagnostic 1	PerfOr Diagnostic 2	PerfOr Diagnostic 3	PerfOr Diagnostic 4	PerfOr Diagnostic 5	PerfOr Diagnostic 6	PerfOr scale	PerfOr single item
PerfOr Diagnostic 1	—							
PerfOr Diagnostic 2	.97**	—						
PerfOr Diagnostic 3	-.02	-.06	—					
PerfOr Diagnostic 4	.69**	.69**	-.07	—				
PerfOr Diagnostic 5	.71**	.73**	-.15	.93**	—			
PerfOr Diagnostic 6	-.24	-.25	.83**	-.35*	-.38**	—		
PerfOr scale	.82**	.85**	-.27	.78**	.78**	-.45**	—	
PerfOr single item	.71**	.73**	-.28	.76**	.75**	-.52**	.87**	—

Note. ** $p < .01$, * $p < .05$.

Table D-10

Intercorrelations Between Diagnostic Items, Dimension Scales and Single Items (Professionalism)

	Profess Diagnostic 1	Profess Diagnostic 2	Profess Diagnostic 3	Profess Diagnostic 4	Profess scale	Profess single item
Profess Diagnostic 1	—					
Profess Diagnostic 2	.94	—				
Profess Diagnostic 3	.70	.76	—			
Profess Diagnostic 4	.75	.81	.95	—		
Profess scale	.80	.83	.86	.88	—	
Profess single item	.80	.78	.77	.79	.93	—

Note. All correlations statistically significant at the $p < .01$ level.

Table D-11

Intercorrelations Between Diagnostic Items, Dimension Scales, and Single Items (Hazing)

	Hazing Diagnostic 1	Hazing Diagnostic 2	Hazing Diagnostic 3	Hazing Diagnostic 4	Hazing scale	Hazing single item
Hazing Diagnostic 1	—					
Hazing Diagnostic 2	.95	—				
Hazing Diagnostic 3	.86	.79	—			
Hazing Diagnostic 4	.89	.86	.96	—		
Hazing scale	-.60	-.59	-.61	-.67	—	
Hazing single item	-.56	-.52	-.64	.79	.89	—

Note. All correlations statistically significant at the $p < .01$ level.

Table D-12

Intercorrelations Between Diagnostic Items, Dimension Scales and Single Items (Bullying)

	Bully Diagnostic 1	Bully Diagnostic 2	Bully Diagnostic 3	Bully Diagnostic 4	Bully Diagnostic 5	Bully scale	Bully single item
Bully Diagnostic 1	—						
Bully Diagnostic 2	-.31*	—					
Bully Diagnostic 3	-.30*	.94**	—				
Bully Diagnostic 4	-.27	.89**	.83**	—			
Bully Diagnostic 5	-.23	.86**	.88**	.93**	—		
Bully scale	.29*	-.71**	-.69**	-.81**	-.81**	—	
Bully single item	.42**	-.67**	-.69**	-.75**	-.79**	.78**	—

Note. ** $p < .01$, * $p < .05$.

Table D-13

Intercorrelations Between Diagnostic Items, Dimension Scales, and Single Items (Fairness)

	Fairness Diagnostic 1	Fairness Diagnostic 2	Fairness Diagnostic 3	Fairness Diagnostic 4	Fairness Diagnostic 5	Fairness Diagnostic 6	Fairness scale	Fairness single item
Fairness Diagnostic 1	—							
Fairness Diagnostic 2	.91	—						
Fairness Diagnostic 3	.84	.89	—					
Fairness Diagnostic 4	.66	.67	.66	—				
Fairness Diagnostic 5	.63	.67	.70	.90	—			
Fairness Diagnostic 6	.59	.63	.72	.82	.90	—		
Fairness scale	.75	.78	.76	.80	.85	.81	—	
Fairness single item	.70	.69	.67	.65	.73	.69	.85	—

Note. All correlations statistically significant at the $p < .01$ level.