Effect of Implicit Leadership Theories on Performance Appraisal

by

Katie McSheffrey Gunther

A thesis submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Masters Degree of
Industrial/Organizational Psychology

Auburn, Alabama
December 13, 2010

Keywords: implicit leadership theories, stereotypes, prototypes, rater bias, performance appraisal

Approved by

Daniel Svyantek, Chair, Professor of Psychology
L. Allison Jones-Farmer, Associate Professor of Management
Alan Walker, Assistant Professor of Management

Abstract

    This thesis tests two predictions about the use of implicit leadership theories (ILT) in performance evaluations: 1) that performance evaluations are systematically distorted in accordance with ILT expectations, and 2) that the relationships between dimensional performance ratings are influenced by a raters pre-existing ILT. Nineteen U.S. Army squad leaders evaluated an average of 9 officer trainees (a mixture of combat and non-combat military occupational specialties) on leadership attributes at the conclusion of 7 weeks of officer training. Rated attributes were classified as diagnostic of combat leadership or non-combat leadership based on military leadership literature. Ratings depended upon the trainees' performance but also on the trainee's combat/non-combat designation, prior enlisted status, and commissioning source, suggesting that complex stereotype expectations influenced ratings. The relationships between rated variables differed depending on trainee combat/non-combat designation, suggesting that raters have implicit theories of attribute co-variation. Finally, different attributes predicted overall leadership evaluations when combat and non-combat trainees were analyzed separately.

Acknowledgments

Table of Contents

List of Tables

List of Figures

INTRODUCTION

Though lawyers may be upset by the stereotype that they are immoral and greedy
ambulance chasers, and doctors by the idea that they are more interested in their stock
portfolios than their patients, for the most part occupational stereotypes seem to excite
relatively few people; my guess would be that most of us assume that most occupational
stereotypes have at least a kernel of truth. Surely most of them do.
-- David J. Schneider in *The Psychology of Stereotyping*, p. 522

Occupational stereotypes are no different than demographic stereotypes in that they are

formed by the same processes --observation and cultural learning -- and used for the same

purpose – as cognitively efficient ways of predicting behavior of the stereotyped person or

group. Unlike demographic stereotypes, occupational stereotypes do not typically incur social

condemnation, and as Schneider (2002) points out, "excite relatively few people." Despite

having a kernel of truth, the use of stereotypes or implicit personality theories causes social

groups to be rated differently (Spears, 2002) and affect judgments about what collection of traits

belong to a social category (Ashmore & Del Boca, 1979). Although race and gender stereotypes

have received much attention in research, stereotypes are not constrained to these demographic

categorizations, and can be applied to all social categories (McGarty, 2002). For instance,

undergraduates have been shown to have relatively consistent stereotypic perceptions of their

own and other academic majors in the areas of work ethic, scholarly ability and interests (Schlee,

Curren, Harich & Kiesler, 2007), and about the attributes that differentiate health care

professionals (Hean, Clark, Adams & Humpris, 2006). Likewise, there are documented

consistencies in stereotypes about nationalities (Eysenck & Crown, 1948), disabled people

(Josefa & Miguel, 2007), and the unemployed (Yzerbyt, Rogier & Fiske, 1998).

Feldman (1981) asserts "When an employee is assigned to a category, further memory-based judgments of that employee are colored by category membership" (p. 130). *Categorization* is a basic cognitive process of clumping information together based on similarity, which facilitates the acquisition and recall of information by limiting its complexity (Bartlett, 1932). Likewise, *social categorization* sorts people into groups based on characteristics that they have in common (Bruner, 1958), whether by outwardly obvious attributes such as gender, race and age, or via information known about the target, such as profession, marital status, or intellect. In cognitive psychology, a *prototype* is a cognitive representation of a category that has all the characteristics of the category. Hogg & Terry (2000) define prototypes as cognitive representations about groups (categories of people), embodying "all attributes that characterize groups and distinguish them from other groups, including beliefs, attitudes, feelings, and behaviors" (p. 123). For this thesis, the process of *categorization, prototyping* and *stereotyping* are considered functionally equivalent (Feldman, 1981), and therefore these terms will be used interchangeably. In an occupational context, people are categorized into their respective jobs, and by association with the job, those people are ascribed certain prototypical attributes or characteristics.

Categories in the workplace are not simply job-related. Lord, Foti and Phillips (1982) proposed that employees are also classified according to a leadership hierarchy that includes a super-ordinate level (leader/non-leader prototypes), basic level (job-context leader prototypes) and subordinate level (which might include differentiations between executive and front-line leader prototypes). This hierarchy partly explains the failure of research to find a stable and generalizable factor structure of leadership competency (Lord, Binning, Rush, & Thomas, 1978), since 1) only a few attributes differentiate leaders from non-leaders at the super-ordinate level

and 2) other attributes serve to differentiate *between* leader prototypes at the basic level. Specifically, Lord and Maher (1991) reported unpublished research that demonstrates leader prototypes differ by job. Undergraduate students described the degree to which 70 attributes described leadership in 7 contexts (business, education, politics, finance, religion, military and sports), at both high (executive) and low (supervisory) levels. Although they expected attributes to group according to level of leadership, the attributes clustered within context. The average correlation of the 70 attributes within a job context was .47, while the average correlation within leadership level was only .21 (Lord & Maher, 1991, p. 48).

The leadership categorization theory proposed by Lord & Maher (1991) asserts that leadership should not be defined as a "collection of behaviors, traits and characteristics and outcomes produced by leaders" but instead "as the outcome of a social-cognitive processes we use to label others" (p.11). They go on to say:

> "…the key issue is how these factors are used by perceivers to form or modify leadership perceptions and the organizational consequences of such perceptions". (Lord & Maher, 1991, p. 11)

In this thesis, the discussion of leadership will follow this social-cognitive orientation, and explore the idea that attributes of leaders are "trait perceptions" that are associated with particular, context-specific prototypes. A "basic level" prototype of the *military leader* will be reported that is well developed in Army doctrine. Two salient and commonly used "subordinate level" prototypes, that of *combat* and *non-combat leaders,* will be compared based on military literature, identifying key differences in their underlying "trait perceptions". These subordinate level prototypes form what Ashmore & Del Boca (1979) describe as *cultural stereotypes*, "shared beliefs that are encoded in the language of a particular group and are transmitted in part by means of socialization" (p. 223). Since prototypes exist in the mind of the rater, the argument

will be made that categorizing a soldier as a combat or non-combat leader influences the manner in which information about that person is encoded and recalled, and consequently how they are rated on leadership attributes. Thus, the prototype, and not the behavior of the person being rated, informs the rating. This conforms to the classic definition of rater bias (Wherry & Bartlett, 1982). By extension, inferences about which leadership attributes covary with each other for a specific (combat or non-combat) leader prototype cause the factor structure of leadership ratings to differ depending on the prototype.

The Military Leader

The Army defines leadership as "influencing people by providing purpose, direction, and motivation, while operating to accomplish the mission and improve the organization" (Dept. of the Army, 2007). The core leadership competencies espoused in army doctrine are listed in Table 1 (Dept of the Army, 2006). These core competencies were expressly developed to describe functions desired of Army leaders *across individuals and jobs* (Horey & Falleson, 2004). Coupled with the Army Values[1] and Warrior Ethos[2], the Army Core Leadership Competencies provide the foundation for leadership training and development (Dept. of the Army, 2007).

---

[1] *"Loyalty – Bear true faith and allegiance to the U.S. Constitution, the Army, your unit, and other soldiers. This means supporting the military and civilian chain of command, as well as devoting oneself to the welfare of others. Duty – Fulfill your obligations. Duty is the legal and moral obligation to do what should be done without being told. Respect – Treat people as they should be treated. This is the same as do unto others as you would have done to you. Self-less service – Put the welfare of the Nation, the Army, and subordinates before your own. This means putting the welfare of the nation and accomplishment of the mission ahead of personal desires. Honor – Live up to all the Army values. This implies always following your moral compass in any circumstance. Integrity – Do what's right – legally and morally. This is the thread woven through the fabric of the professional Army ethic. It means honesty, uprightness, the avoidance of deception, and steadfast adherence to standards of behavior. Personal Courage – Face fear, danger, or adversity (physical or moral). This means being brave under all circumstances (physical or moral)." Department of the Army, 2007.*

[2] *"I will always place the mission first. I will never accept defeat. I will never quit. I will never leave a fallen comrade." Ibid, 2007.*

Table 1

Army Core Leadership Competencies

Leads others: Leaders motivate, inspire, and influence others to take the initiative, work toward a common purpose, accomplish tasks, and achieve organizational objectives.

Extends influence beyond the chain of command: Leaders must extend their influence beyond direct lines of authority and chains of command. This influence may extend to joint, interagency, intergovernmental, multinational, and other groups, and helps shape perceptions about the organization.

Leads by example: Leaders are role models for others. They are viewed as the example and must maintain standards and provide examples of effective behaviors. When Army leaders model the Army Values, they provide tangible evidence of desired behaviors and reinforce verbal guidance by demonstrating commitment and action.

Communicate: Leaders communicate by expressing ideas and actively listening to others. Effective leaders understand the nature and power of communication and practice effective communication techniques so they can better relate to others and translate goals into actions. Communication is essential to all other leadership competencies.

Creates a positive organizational climate: Leaders are responsible for establishing and maintaining positive expectations and attitudes, which produce the setting for positive attitudes and effective work behaviors.

Prepares self: Leaders are prepared to execute their leadership responsibilities fully. They are aware of their limitations and strengths and seek to develop and improve their knowledge. Only through preparation for missions and other challenges, awareness of self and situations, and the practice of lifelong learning and development can individuals fulfill the responsibilities of leadership.

Develops others: Leaders encourage and support the growth of individuals and teams to facilitate the achievement of organizational goals. Leaders prepare others to assume positions within the organization, ensuring a more versatile and productive organization.

Gets results: Leaders provide guidance and manage resources and the work environment, thereby ensuring consistent and ethical task accomplishment.

Note. Army Regulation 600-100 Army Leadership, *2007, p. 2.*

In an earlier technical report describing the development of the competency model, the researchers reported that at least one subject matter expert suggested "a need for a differentiation between combat and non-combat competencies" (Horey, Fallesen, Morath, Cronin, Cassella,

Franks, & Smith, 2004, p.41), suggesting that the universal military leader prototype was not particularly reflective of the perceptions soldiers had about their leaders. Combat arms (CA) soldiers are those who enter the battle space for the purpose of "closing with and destroying the enemy"; Kirin & Winkler, 1992) and non-combat arms soldiers (NCA) are those who support combat arms troops by providing services and supplies during Army operations. One need only look toward the process of socialization and indoctrination of soldiers to recognize some key differences between the prototypes of combat and non-combat leaders, prototypes that are developed from both theory (organizational doctrine and indoctrination) and data (true performance differences between combat arms leaders) Theory and data are the dual processes by which prototypes or stereotypes are developed (Brown & Turner, 2002).

Combat Leader Prototype "Theory"

Organizations can act as the social source of "explicit theories" about the relationships among traits, and between traits and roles, resulting in what could be described as cultural stereotypes or prototypes, "shared beliefs that are encoded in the language of a particular group and are transmitted in part by means of socialization" (Ashmore and Del Boca, 1979, p. 223). The organization explicitly transmits the content of job roles to employees, which contain "the attitudes, beliefs, perceptions, habits and expectations of human beings [that] evoke the required motivation and behavior" (Katz and Kahn, 1978, p.187) to accomplish organizational goals. Consequently, background knowledge takes the form of explicit role expectations, with the organization providing both a category label (e.g. 'infantryman', 'clerk') as well as the category content (e.g. job duties, expectations of performance, organizational values).

It is not the purpose here to articulate all attributes that differentiate combat from non-combat leaders. Instead, the case will be made that "leading by example" (the third Army core

6

competency described in Table 1) specifically differs in the underlying attributes associated with it. In other words, the context of combat changes the manner in which leaders display the competency, and by observation and cultural learning, those differences are built into the prototype of the combat leader. S.L.A Marshall, in his often-quoted work, *Men Against Fire: The Problem of Battle Command in Future War*, described in detail the attributes thought necessary of a combat leader "if he is to prove capable of preparing men for and leading them through the shock of combat with high credit" (Marshall, 1978, p. 163). Marshall's short list includes diligence in the care of men, military bearing, courage, creative intelligence and physical fitness. He, as well as other authors (Von Schnell, 2004; Wood, 1984), describe a certain type of "leading by example" that differentiates leaders from non-leaders in the realm of combat operations. Combat "leading by example" is behavioral modeling intimately linked to the chaotic and dangerous environment of the battlefield. Leaders in combat must "lead from the front", which entails being present on the battle field enduring the hardships of their troops, modeling behavior that will be decisive in the face of the enemy. Without this particular brand of "leading from the front" it is impossible to "inspire …soldiers to do things against their natural will—possibly to risk their lives—to carry out missions for the greater good of the unit, the Army, and the country." (Dept. of the Army, 1983, p. 1).

Army leadership research confirms that soldiers perceive some leader attributes as more central to some jobs than to others. Steinberg and Leaman (1990a, 1990b) conducted an Army-wide leadership task analysis, compiling a list of 20 knowledge, skills and abilities (KSAs) and 560 tasks common to commissioned officers and non-commissioned officers in the domain of leadership. 5,033 randomly selected active duty officers of all ranks and military occupational specialties (MOSs) completed importance ratings, which required that they indicate how

important (from 1-- *Not important* to 7 -- *Extremely important*) the KSAs were to their jobs

(Steinberg & Leaman, 1990a). Table 2 summarizes the leadership tasks for which differences

existed between combat and non-combat leaders, using the criteria of the original study (the task

received a mean criticality rating of 6 or more by 60% of the respondents).

Table 2
Army Leadership Task Analysis Results

Critical Leadership Tasks	% of Branches Rating This Task as Critical	
	CA	NCA
Remain with the element you lead. (Task 182)	100	29
Share the hardships with soldiers in the field. (Task 183)	100	71
Take charge in the absence of instructions from commander. (Task 111)	100	57
Direct/lead from a forward position in battle. (Task 163)	86	0
Promotes physical fitness. (Task 292)	86	57
Require subordinates to maintain military bearing and appearance in the field. (Task 196)	71	29
Evaluate individual soldier performance against established standards. (Task 490)	57	43
Demonstrates expertise on weapons subordinates use. (Task 162)	29	0
Conduct after action reviews. (Task 482)	29	0
Increase soldier willingness to take risks in combat. (Task 155)	14	0
Keep soldiers motivated under sleep deprivation conditions. (Task 156)	14	0

Note. From Army Leadership Requirements Task Analysis: Commissioned Officer results (Steinberg & Leaman, 1990b)

Supporting the premise that physical fitness is a crucial part of setting the example for

combat leaders, 86% of CA branches rated 'promoting physical fitness' as a critical task for their

branch, while only 64% NCA branches rated this task as critical. Leading from the front (task

163), taking initiative (task 111) and sharing the hardships of soldiers in the field (task 183) all

support Marshall's view that combat soldiers perceive these attributes as critical to leadership,

and do so to a higher degree than non-combat soldiers. In contrast, non-combat soldiers

considered KSAs related to communication (KSA 000 -- Ability to speak effectively/clearly,

KSA 014 -- Ability to communicate effectively in writing, and KSA 010 -- Ability to listen effectively/actively) and decision making (KSA 018 -- Ability to make decisions) more critical to their jobs than combat soldiers.

There are also combat and non-combat differences in the degree to which attributes are seen to distinguish superior from average junior leaders. Cullen, Klemp and Mansfield (1988) found that different branches varied in the degree that "concern for standards", "assuming responsibility", and "professional detachment" were considered superior performance indicators, with CA branches rating these qualities more highly than NCA officers. Other attributes, such as "positive regard for subordinates" and 'mission oriented", did not differ significantly among branches.

Combat Leader Prototype "Data"

Observation in the workplace provides the "data" by which differing associations between jobs and attributes are formed. For instance, subordinates form associations between the attribute of physical endurance and the role of combat leader during training. They might observe the negative effects of a leader who cannot keep up with the unit on a foot patrol, and is therefore not available to lead the mission. Or they observe mission success when a leader takes the initiative in combat, acting despite incomplete information and gaining the advantage over the enemy. Preexisting differences between combat and non-combat leaders also contribute to these associations. Knapp, McCloy and Heffner (2004) found that CA leaders scored an average of 15 points higher than NCA leaders in the 300-point Army physical fitness test in a sample of 144 non-commissioned officers. Conversely, NCA soldiers scored significantly higher on situational judgment tests (Knapp et al., 2004), and collectively are required to have higher ASVAB scores than CA soldiers for entrance into their respective occupational specialties (the

9

mean ASVAB cut score for CA MOSs is 90, while the mean cut score for NCA is 110; Campbell & Zook, 1997).

Based on military literature and leadership research, there is sufficient evidence to suggest that the prototypes of combat leaders and non-combat leaders differ in important ways. Particularly, attributes of physical fitness/endurance seem to be central and diagnostic of combat leaders, while decision making is central and diagnostic of non-combat leaders.

<center>Influence of Prototypes on Performance Appraisal</center>

Attribute diagnosticity, the degree to which a trait describes or is useful in categorizing a group (Ford & Stangor, 1992), is pivotal to understanding how prototype bias will play out in performance appraisal. Diagnostic attributes influence recall and expectations more so than those attributes that are less central to the stereotype because they are called to mind more readily than peripheral attributes, and they are thought to do a better job of differentiating between groups (Ford & Stangor, 1989; McCauley, Stitt & Segal, 1980). Diagnostic attributes constitute the implicit leadership theories that raters have about leaders of different jobs. In the previous section, two diagnostic attributes were highlighted that can be measurable objectively: physical fitness for combat arms soldiers, and decision making for non-combat arms soldiers.

In order to assign an accurate performance rating, a rater must first attend to relevant performance while it happens, encode that information into memory, and be able to recall it for the rating occasion. A rater's pre-existing schema of a particular employee based on his/her membership in a job category has the potential to bias each of these cognitive processes. The rater may fail to attend to (and thus fail to encode) things that don't confirm schema-based expectations (Ilgen & Feldman, 1983), or the rater may use implicit theories to shape what he or she "remembers" of ratee performance on recall (Schweder & D'Andrade, 1980).

<center>10</center>

Leniency and Harshness Bias

In performance appraisal, schema-based information that is included in a rating but is not related to actual performance of the ratee is rater bias (Wherry & Bartlett, 1982). Typically, bias is operationalized as significant mean rating differences between categories of ratees after controlling for true performance differences. Broadly, these biases may be described as leniency or harshness errors on the part of raters. Rater leniency refers to the "rater's tendency to assign ratings that are generally higher …than are warranted by the ratee's actual performance" (Scullen & Mount, 2000, p. 957). Conversely, harshness is the tendency to assign lower ratings than warranted by performance. Statistically, leniency or harshness bias can be interpreted as significant main effects of social category on ratings of performance, without a significant interaction between social category and objectively measured performance in the same domain.

Several theories exist as to why raters inflate or deflate ratings of particular social groups. Social identity theorists suggest that it is to maintain positive differentiation between an in-group and out-group, which serves to enhance self esteem (Tajfel & Turner, 1979). Positive attributes are assigned to members of the in-group (people of the same category as the rater). Turner (1982) has suggested that attributes ascribed to a particular social category will be assigned to all members of that category, causing raters to unconsciously inflate ratings based solely on category membership. Eagly (1987) proposed that ratings are inflated (or deflated) based on the congruity between the social category and the attribute being rated. In other words, if high physical fitness is congruent with the mental schema held about CA soldiers, and not with NCA soldiers, ratings of physical fitness will be inflated for CA soldiers.

Figure 1. Demonstration of social category rating bias.
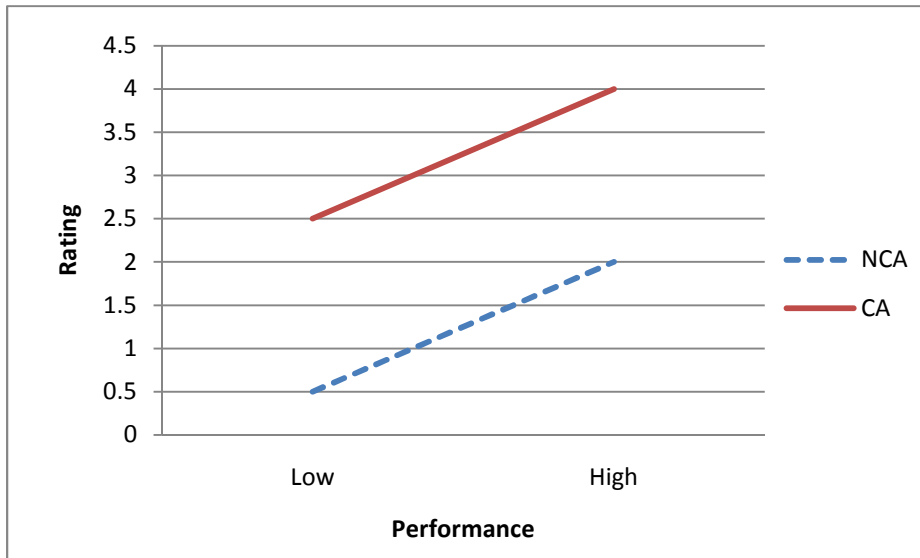
Figure 1 describes this hypothetical situation, in which a CA leader and NCA leader of equal job performance are evaluated differently depending on their social category alone. This sort of bias has been demonstrated in the rating literature for several demographic characteristics. Raters have demonstrated gender bias by evaluating men more highly than women in managerial roles (Eagly, 1987); race bias by rating whites more highly than blacks in certain military performance domains (Pulakos, White, Oppler & Borman, 1989); and rater-by-ratee similarity bias in which ratees who hold the same demographic characteristic of the rater are evaluated more highly (Kraiger & Ford, 1985).

When the rater has difficulty calling to mind specific instances in which a ratee demonstrated (or failed to demonstrate) a diagnostic attribute, the rater's evaluations may be inflated or deflated in the direction of the stereotype. In other words, they use the availability of prototypical attributes to inform their judgments about what must have occurred for a specific individual (Tversky & Kahneman, 1974; Cantor & Mishel, 1977; Tsujimoto, 1978). CA soldiers would be evaluated more highly than NCA soldiers in physical fitness simply because this attribute is more strongly associated with CA leadership.

Expectation Bias

The social category "main effect" manifestation of bias is likely only to hold in circumstances in which the performance of the ratee is not remembered, or not observed at all. In most rating circumstances, we would hope that the rater does in fact observe the performance of his or her ratees in the areas that are important to that job, and this information is integrated into memory for later recall. During the performance observation period, the employee may engage in behavior that is consistent or inconsistent with the rater's expectations based on a job stereotype. When a behavior is stereotype-inconsistent, that behavior will be surprising. Surprising instances are known to induce controlled information processing (Hastie & Kumar, 1979), with the result that the behavioral instance will be strongly encoded and available for recall. Heider, Scherer, Skowronski, Wood, Edlund and Hartnett (2007) found just such strong encoding of stereotype-inconsistent information. They presented congruent, incongruent and neutral behavior descriptions (12 of each) about three targets with stereotyped jobs (drill sergeant, college professor, and kindergarten teacher) to a group of undergraduate students. After engaging in an unrelated memory task, participants were given a surprise recall task about behaviors associated with each of the targets. Incongruent behaviors were remembered far better than either congruent or neutral behaviors.

It is likely that a rater faced with the strongly encoded memory of stereotype-inconsistent information about a ratee would evaluate that ratee more severely than other ratees who have not violated expectations. Jussim, Coleman, and Lerch (1987) suspected the following outcomes when expectancies are violated:

> Individuals who possess more favorable characteristics than expected should be evaluated even more positively than others with similar characteristics whom we expected to rate positively all along. Likewise, individuals who possess more unfavorable characteristics

13

than expected should be evaluated even more negatively than others with similar characteristics whom we expected to rate negatively all along. (p. 537)

Jussim et al. (1987) manipulated stereotype-violations by presenting black job applicants who appeared to be of upper socio-economic status, which presumably violates most white's stereotypes of blacks. Stereotype violators (blacks who appeared of high socioeconomic status) were evaluated more extremely (positively, in the direction of the violation) compared to non-violators for 9 out of 10 job-related attributes such as "hard working", "intelligent", "competent" and "ambitious". The expectancy violation effect is described by Figure 2, in which the relationship between observed performance and rated performance differs by job role. In this case, CA soldiers low in physical fitness are evaluated more severely than comparable NCA soldiers, since such performance violates the stereotype for combat leaders.

Figure 2. Expectancy violation effect on ratings.
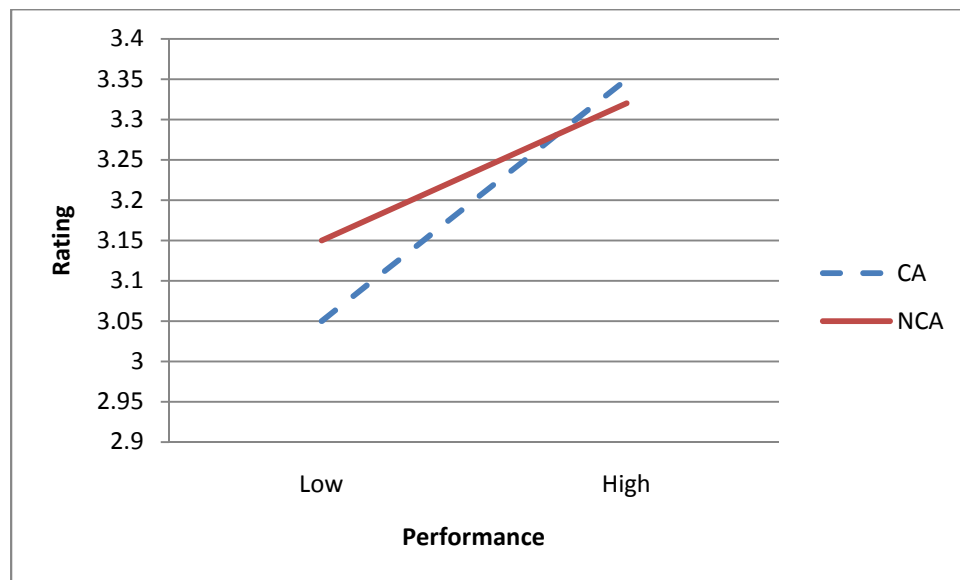
This leads to hypothesis 1:

Hypothesis 1: *The relationship between ratings of performance and actual*

performance will differ by trainees' combat/non-combat designation.

Implicit Covariance Schemas

A second conception of stereotyping may be described as *category-based co-variation* and is central to Ashmore and Del Boca's (1979) definition of a stereotype: "a structured set of inferential relations that link a social category with personal attributes" (p. 225). Schneider (2004) proposes, "stereotypes about groups are not merely lists of traits associated with those groups, but specific relationships among those traits that may vary by group" (p. 193). These traits represent aspects of the social category that are considered typical and diagnostic of the category. Traits that are typical of a category will covary together for members of that category. This is because diagnostic attributes are strongly associated with a social category, thus are called to mind more quickly and used more often to make inferences about other attributes (Ford & Stangor, 1992). For instance, based on job analysis and doctrine in the U.S. Army, we know physical fitness, perseverance in the face of physical danger and initiative are considered important and diagnostic of CA soldiers. The consequence is that raters make implicit assumptions that these traits covary in the CA soldiers they rate. The presence or absence of one trait implies the presence or absence of the other. Consequently, ratings of these attributes would reflect what Cooper (1981) calls an "implicit covariance structure" or Lord and Maher (1991) call an "implicit leadership theory", as described in Figure 3.

Figure 3. Demonstration of category based co-variation that is stronger for diagnostic attributes of the stereotyped group.
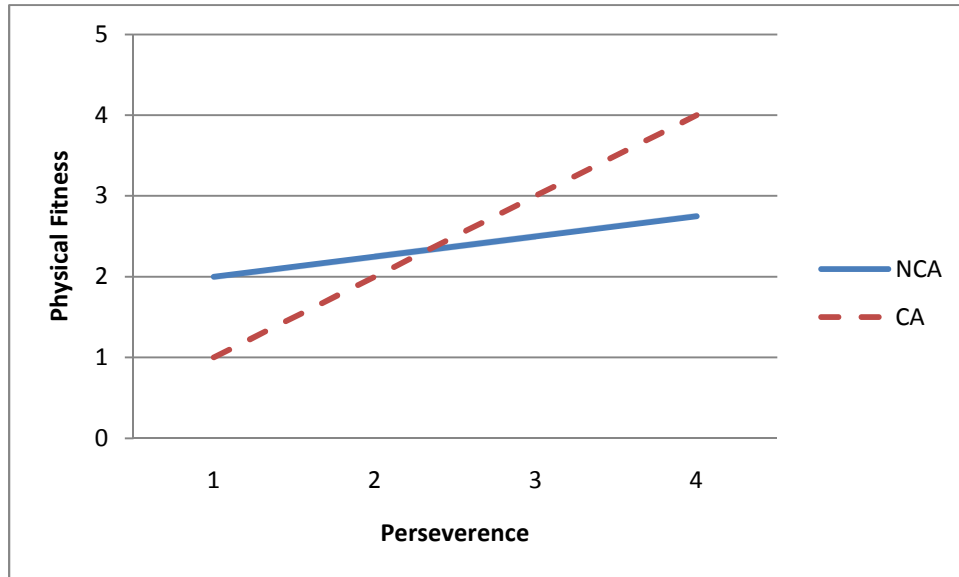
Categorization as a combat leader is the source of conceptual association between physical fitness and perseverance in the mind of the rater that does not necessarily reflect the underlying true covariance of these attributes. This phenomenon was first described as the *systematic distortion hypothesis*. Its originators, Schweder and D'Andrade (1980), suggested that:

> under difficult memory conditions judges on personality inventories, rating forms, and questionnaire interviews infer what "must" have happened from their general beliefs about what the world is like and/or find it easier to retrieve conceptually related memory items.(p. 38)

Applying this rationale to performance appraisal, Cooper (1981a), and later Kozlowski, Kirsh and Chao (1986), considered whether implicit theories of "what is like what" contributed to illusory halo -- the inflation of rated performance correlations above true performance correlations. In Cooper's study, participants rated the conceptual similarity of the performance dimension pairs ranging from 0 (totally dissimilar) to 10 (virtually identical) within the framework of three specific jobs (university professor, professional engineer and retail manager).

As had been discovered by Schweder and D'Andrade (1980), conceptual similarity scores corresponded closely with the correlations between performance dimensionl ratings. Dimensions that were similar to each other conceptually were also highly correlated when looking at performance ratings. However, the level of correspondence was not consistent for all three jobs. Since Cooper's aim was to test whether conceptual similarity between attributes contributed to halo, he did not specifically address the differences in covariation matrices found between jobs. He speculated that raters' use of a representative heuristic (Tversky & Kahneman, 1974) might have been the cause of the different correlation matrices between jobs.

The results from the studies above suggest that raters would be biased by their existing covariance schemas about a job, with the consequence that they will "remember" or infer a certain level of performance in one performance dimension because it is conceptually grouped with other performance dimensions *for a particular job*. If the attributes are strongly, positively correlated with each other conceptually, assigning a high rating in one induces a comparable rating in the other for the members of that job category. This leads to Hypothesis 2:

Hypothesis 2: *The relationship among ratings of combat leader diagnostic attributes significantly different than the relationship of these same variables for non-combat leaders.*

Based on job analysis literature, we would expect that physical fitness and perseverance would be more highly related for CA soldiers since both of these attributes are conceptually related to combat leadership and consequently would co-vary in the mind of the rater. In linear regression, this would be represented by a significant coefficient for an interaction term *job category x leadership attribute* in explaining a second leadership attribute (Pedhazer, 1997). This does not suggest that there is no true relationship between

17

these two attributes. It would be reasonable to believe that they do in fact covary.

However, a stronger correlation between these two variables for combat leaders, after

controlling for true performance in physical fitness, represents what Cooper (1981)

describes as illusory halo.

Research Question

What is the implication of the types of bias described above? Within the

performance appraisal system, the implications of rater leniency and harshness are straight

forward: these sorts of biases have the effect of rewarding and punishing individuals via

the performance management system in a way that is not warranted by their true

performance. On the other hand, raters' implicit leadership theories which lead to

systematic distortion of ratings, have implications for research.

Assume that systematic distortion of ratings leads to different factor structures for a

collection of leadership attributes depending on whether the ratees are combat or non-

combat leaders. Imagine that these distorted ratings were then used to select which

attributes predicted leadership effectiveness in a combined sample of both combat and non-

combat leaders, as was the case in a validation study of the Army Core Leader

Competency Model reported by Horey et al. (2007). The researchers used subordinates'

ratings of target leaders (whose job categories were unreported) to predict overall

leadership effectiveness as reported by the target leaders' superiors. By not reporting the

target leaders' military occupational specialties, the researchers failed to test whether the

ratings (from superiors, subordinates, or both) differed systematically by job category. As

Guion (1998) cautions, combining groups that differ from each other systematically in

means and correlations distort the overall correlation coefficients. The systematic

distortion hypothesis (Schweder & D'Andrade, 1980) suggests that means and correlations are likely to be distorted, since raters remember or infer performance on diagnostic attributes more readily than on those that are not diagnostic of particular jobs. Diagnostic attributes would 1) be strongly related to each other for the job for which they are conceptually related (but not for other jobs), and also 2) be strongly related to the outcome variable for specific jobs. In a sample with subgroups that differ systematically, distorted correlation coefficients could produce a list of predictors that are not generalizable to other samples.

Horey et al. (2007) used a forward selection procedure in multiple regression to determine which items of the competency measure meaningfully predicted Overall Leadership Effectiveness. By collapsing the data across military occupational specialties, the assumption is made that raters are accurately evaluating leadership performance without invoking a combat or non-combat leader stereotype. This seems implausible.

> *Research Question:* Which attributes meaningfully predict overall leadership
> effectiveness in a combined sample, and does this set of attributes differ when
> combat and non-combat leaders are analyzed separately?

If diagnostic attributes significantly predict overall leadership effectiveness ratings, it would appear that implicit leadership theories are indeed affecting performance ratings.

METHOD

Participants.

In April of 2005, a single-site implementation of BOLC II at Fort Benning, Georgia, consisted of 178 lieutenants. Of these, data were collected on 173. The mean age of the lieutenants was 23.77 (SD=3.21). The class consisted of 37 female (21%) and 138 (79%) male lieutenants. Fifty-four of the lieutenants (31%) were prior-enlisted, with 8 of those (5%) having combat experience. The Lieutenants were assigned to 20 squads of approximately 9 soldiers each. The squads were trained and evaluated by 19 senior staff non-commissioned officers (Sergeants and Staff Sergeants), with a mean time in service of 12 years. Fourteen of these (70%) held CA MOSs, 14 (70%) were combat veterans, and 1 (5%) was female. One squad leader in 1st platoon evaluated two squads (for a total of 17 trainees as compared to average of 9).

The Research Setting

In 2004, the U.S. Army Accession Command requested that the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) evaluate the effectiveness of a new officer education curriculum called the Basic Officer Leaders Course (Pleban, Tucker, Centric, Dlubac, and Wampler, 2006). The purpose of BOLC is to develop junior officers that "are tactically proficient, knowledgeable in field craft, and confident in their abilities to lead a platoon" (Pleban et al., 2006). As a part of this evaluation, ARI researchers developed measures to document a change in leadership ability, tactical/technical knowledge, and decision making as well as collect objective measures of soldier skill proficiency (weapons qualifications, physical fitness tests, etc.) as a consequence of BOLC training. While previously new lieutenants were sent directly to

their branch-specific schools, and received no core 'warrior skills' training, this course specifically emphasized warrior ethos and basic skills training for all lieutenants, regardless of branch. This training took place at Fort Benning, Georgia.

Basic Officer Leaders Course sourced its first class of students so as to represent all commissioning sources (Officer Candidate School, Reserve Officer Training Corps, Warrant Officers and Direct Commissions). Trainees were stratified across squads to ensure that women and NCA military occupational specialties were distributed among raters. Likewise, Army Accessions Command dictated that each platoon (consisting of four squads) have an equal representation of both CA and NCA Squad Leaders. However, CA MOSs were more highly represented in 3 out of 5 platoons (see below).

Measures for Hypothesis 1 and 2

Two models will be used to test the hypotheses that ratings reflect the use of implicit leadership theories about combat and non-combat leaders. The first model focuses on the attributes and observed performance theorized to be diagnostic of combat leaders (physical fitness and perseverance), while detecting the influence of trainee and rater demographic covariates. The second model parallels this same procedure, but focuses on the attributes and observed performance theorized to be diagnostic of non-combat leaders (decision making and planning). The measures associated with each of these models are described below.

Combat Leadership Model. At the conclusion of 7 weeks of training, squad leaders received a copy of the end-of-course evaluation with which to summarize their assessments of the trainees on warrior ethos, Army core values, leadership, and adaptability based on the Army Leadership Competency Framework (Horey et al., 2004) and the Army Officer Evaluation System (Dept. of the Army, 1999). Each attribute was rated on a 4 point scale (1=Needs much

improvement – rarely or never behaves this way, 2= Needs some improvement – sometimes behaves this way, 3= Satisfactory – usually behaves this way; 4 = Excellent – always or almost always behaves this way, and N/A – Not applicable. The complete End-of-Course Evaluation is included in Appendix 1.

Combat Leadership Outcome Variable. Two closely related attributes are mentioned repeatedly in military doctrine and training publications as being essential to combat leadership, physical fitness and physical stamina/endurance. In the End-of-course Leadership Evaluation, the ratings of *Physical* and *Physical Adaptability* are consistent with these attributes, and were averaged to create a scale. The *Physical* rating was described as "Displays appropriate level of physical fitness and military bearing." *Physical Adaptability* was described as "Adjusts to tough environmental states such as extreme heat, humidity, cold, etc.; Frequently pushes self physically to complete strenuous or demanding tasks; Adjusts weight/muscular strength or improves proficiency in performing physical tasks needed to be successful for job/training mission." Although this scale confounds several related concepts about physical performance, the concepts contained within this scale are diagnostic of CA soldiers (Marshall, 1978). The two item scale (*Physical Rating Scale*) has a Cronbach alpha of .851, and the two variables have a correlation of r=.741.

Combat Leadership Explanatory Variables.

Trainee Job. In the Army, jobs are referred to as Military Occupational Specialties (MOSs). Commonly, MOSs are further grouped into those that are directly engaged in warfighting (combat arms) and those that support the warfighter (combat support and combat service support). Trainee job category (*TrainJob*) was coded as either 'CA' (1) or 'NCA' (0)

22

based on their MOS and in accordance with the Army Military Occupational Specialty Database (Kirin & Winkler, 1992).

Observed Performance Variable. During the course of training, squad leaders observed, recorded and reported the performance of their trainees on basic warrior tasks and skills, including Army Physical Fitness Test, marksmanship, land navigation, a knowledge test, and a situational judgment test. Scores on the Army Physical Fitness Test were collected at the beginning of the second week of training. Since ratings of physical fitness, physical effort and physical endurance described by the *Physical Scale* (see Outcome Variables above) would be based in part on observed performance on the APFT, the APFT is used as a covariate in this analysis to support the claim that soldiers of equivalent physical ability are evaluated differently depending on combat/non-combat leader categorization, indicating a difference in the way that raters encode and recall information about trainees based on combat or non-combat prototypes. The points awarded on the individual events in the APFT (push-ups, sit-ups, and a 2 mile run) are normed by age (in five year increments) and gender. As a result the total APFT score reflects performance after accounting for age and gender differences (Dept. of the Army, 1998a). For instance, male soldiers aged 17 through 22 must achieve a run time of 15:54 to receive a minimum passing score of 60 points, while a female soldiers aged 27 through 31 must achieve a 20:21 run time to receive the same number of points. Soldiers must obtain a minimum of 60 points on each event of the APFT in order to pass. There were no significant differences between CA and NCA in PFT scores, $F_{(1,169)} = 1.078$, $p = .301$, or pass/fail criteria, $F_{(1,169)} = .104$, $p = .748$.

Combat Leader Diagnostic Attribute. *Perseverance*, an attribute diagnostic of combat leaders (Marshall, 1978), was described in the end-of-course evaluation as "Works through adversity. Does not give up."

Demographic Covariates.

The following covariates were available in the data set and were included to explore their affect on the hypotheses:

Age: A trainee's age can influence both performance (older soldiers typically do not perform as well on tests of physical fitness as young soldiers) and perceptions of raters (perceived maturity and experience). Age was measured in years.

Commissioning source: Officers that attended BOLC were purposefully drawn from all commissioning sources. The length and nature of the training program that led to the soldiers' commission differed substantially, from programs that provided no officer training prior to commission (direct commissions from the enlisted force or "battlefield" commissions) to programs that provided 4 years of indoctrination and training (like the Army's service academy West Point). The commissioning sources were coded as follows: Officer Candidate School (1), Reserve Officer Training Corps (2), United States Military Academy (3), and direct commissions from the enlisted ranks or Warrant Officers (4).

Component status was coded as Active Duty (1) or Reserve/National Guard (2).

Prior Enlisted experience was coded as either "No" (0) or "Yes" (1).

Rater Variables.

Trainees with a common rater will have performance scores that are more similar to each other than with those evaluated by another rater (LaHuis & Avis, 2007). Some raters may have the tendency to use only a portion of the rating scale, or demonstrate overall leniency or

harshness in comparison to other raters. Ignoring this information would lead to faulty conclusions about the meaning of the individual ratings. Certainly a rating of 3.6 given by a lenient rater whose mean rating for his trainees is 3.5 should be interpreted differently than a 3.6 given by a rater whose mean rating is 2.8. In a sense, the rater idiosyncrasy is treated as a source of variance that must be accounted for in order to detect the underlying structure of the data. Raters idiosyncratic scale usage is operationalized as a rater's deviation from the grand mean of ratings across all attributes. In a sense, this reflects general leniency or harshness on the part of the rater in evaluating his or her squad. This value (*RaterDev*) was computed by first averaging all ratings contained in the End-of-Course Evaluation within trainee, and then averaging within rater. The resulting value was then subtracted from the grand mean of all ratings across all raters. This value was assigned to all members of a rater's squad as a Level-2 (group level) variable. Adding idiosyncratic scale usage as a covariate to this analysis controls for the variability in rater means that act as "noise" to mask the relationships of interest.

In addition the following rater (Level-2) demographic variables were available for analysis: *Rater ID* -- Each rater was assigned a unique identification number, 1 through 19. *Rater Tenure* reflects the squad leader's total time in service, measured in months. Experienced raters may differ systematically from inexperienced raters in the use of job-related stereotypes, having more complex mental models with which to evaluate their trainees. *Rater Job* categorizes the rater by MOS as either NCA (0) or CA (1). Raters may exhibit similarity bias by evaluating trainees of their same job category more favorably than others (Sears & Rowe, 2003). Consequently, variables like *Rater Tenure* and *Rater Job* category can be used as predictors of intercept and slope variability between raters.

Non-combat Leadership Model. As with the combat leadership model described previously, trainee job category (combat or non-combat arms), trainee demographic variables (prior enlisted status, commissioning source, age, and component status) and rater variables (rater deviation from the grand mean, rater tenure and rater job) were included in the non-combat leadership model as covariates and will not be repeated here. The outcome variables and diagnostic variables described below are theorized to be included in a prototype of a non-combat leader.

Non-combat Leadership Outcome Variable. Since decision-making and mental ability are considered diagnostic of NCA soldiers, a two item scale was created from the variables *Conceptual* and *Decision Making*. The *Conceptual* rating was described as "Demonstrates sound judgment, critical/creative thinking, and moral reasoning." The *Decision Making* rating was described as "Employs sound judgment, logical reasoning, and uses resources wisely." The two item scale (*Decision Making Rating Scale*) has a Cronbach alpha of .828, with a correlation between the two variables of r=.700.

Observed Performance Variable. The 10-item situational judgment test (SJT) developed for use in BOLC consisted of brief descriptions of situations a lieutenant would likely face on the job, with the instructions to choose both the best and worst options presented. Thus, each question was worth a total of two points, for 20 points total. SJTs have been found to assess both job-related declarative knowledge and decision making ability (Schmitt & Chan, 2006). Thus, the SJT is used in this study as an objective measure of these mental skills, attributes considered important and typical for NCA soldiers (Knapp et al., 2002; Kirin & Winkler, 1992). The complete situational judgment test is included in Appendix 2.

Non-combat Leadership Explanatory Variables. *Planning*, an attribute typical of NCA soldiers, was described as "Develops detailed, executable plans that are feasible, acceptable, and suitable."

Measures for the Research Question

Outcome Variable. Raters completed a single item indicating their overall evaluation of the leaders undergoing training at BOLC, the *Overall Net Assessment*. The instructions read "Taking into consideration all of the preceding values and attributes, circle the number that best reflects your overall rating:" on a scale from 1 to 4. This variable was rater-centered, such that each score represents a trainee's deviation from the rater's mean. This transformation controls for idiosyncratic scale usage of each rater and ensures the variable is normally distributed for use in regression analysis. After rater mean centering, the ratings of Overall Net Assessment did not differ significantly by rater, $F(18,153) = .000$, $p = 1.0$, making this an acceptable outcome variable for simple linear regression.

Predictor Variables. All variables contained in the End-of-Course Leadership Evaluation (Appendix 3) and the Adaptability Rating Scale (Appendix 4) were included as predictors of the *Overall Net Assessment*.

Data Analysis

Both the structure of the data (trainees nested within rater) and the level of analysis (rater bias) call for the use of hierarchical linear modeling (HLM; Raudenbush and Bryk, 2002). HLM provides several advantages over ordinary least squares regression in the analysis of nested data. The nested nature of the data implies that the errors within groups will be correlated, violating a basic assumption of regression (Luke, 2004); nested data also implies that both the intercepts and slopes may differ by rater. Both correlated errors and randomly varying intercepts and slopes are

handled well in HLM using restricted maximum likelihood (RML) estimation. Although RML generally requires that the number of level-2 units be large, in cases in which the units are not unbalanced and the focus are on the fixed effects (as they are in this analysis) the maximum likelihood method is capable of producing reliable estimates (Raudenbush & Bryk, 2002).

HLM takes a model-building approach to analysis. Traditionally, the null or unconstrained model is tested first to detect significant variance associated with groups, thus ensuring that hierarchical linear modeling is appropriate for the data. This is followed by inclusion of all fixed effects at the individual level of analysis. Once the Level-1 model is built, the researcher tests whether the slope associated with each fixed effect varies significantly by the Level-2 units, and if so, the researcher attempts to predict this variability with Level-2 variables. That is the approach taken here for each of the two outcome variables (*Physical Rating Scale* and *Decision Making Rating Scale*) separately, exploring the prototype bias associated with each job category. The sections below describe the model building process for Hypothesis 1 and 2.

The first step in HLM analysis is to determine if significant variability in ratings are attributable to raters. If no significant variability is attributable to raters then trainee ratings can be analyzed without regard to who assigned them. In order to test for significant rater variability, a one-way random effects ANOVA is conducted. This model is called the unconstrained or random intercept-only model, since there are no explanatory variables:

$$L1: Rating_{ij} = \beta_{0j} + r_{ij}$$
$$L2: \beta_{0j} = \gamma_{00} + u_{0j}$$

In the Level 1 (Individual) model, $Rating_{ij}$ is the rating of the ith trainee by the jth rater, which is described by the rater's mean rating (β_{0j}) and a "person effect" (r_{ij}, the deviation of the individual's score from the rater's mean). In the Level 2 (Rater) model, the effect of the jth rater

28

is described as the grand mean rating across all raters (γ_{00}) plus a random 'rater effect' (u_{0j}, the deviation of the rater's mean from the grand mean, which is assumed to have a mean of zero and a variance τ_{00}). The null model described by equations 1 and 2 provides point estimates of the grand mean γ_{00}, within-group variability σ^2, and between group variability τ_{00}.

The significance of between-group (rater) variance τ_{00} is determined by comparing the full null model (the random-intercept model) in equations above with a reduced model (a fixed intercept model) in which the intercept does not vary by rater. A significant chi-square test of deviances (defined as -2xLN(Likelihood); Hox, 2002) is seen as evidence that the amount of variance explained by random intercept model is significantly larger than the variance explained by the fixed intercept model (Bliese, 2002).

The variance estimates derived from the random intercept model can be used to calculate the proportion of variance associated with raters, the intra-class correlation coefficient:

$$ICC = \frac{\sigma^2}{\sigma^2 + \tau_{00}}$$

In this case rater-effects were anticipated to account for a large amount of the total variance, due to idiosyncratic tendencies of the raters. Based on previous research, we would expect 25% or more of the variance in ratings to be attributable to rater effects (Vivwasvaran, Ones and Schmidt, 1996).

Hypothesis 1 and 2 Model Building. The hierarchical linear model below represents the basic test of Hypothesis 1 and 2, upon which further model building will explore the influence of trainee demographic covariates and level 2 (rater) variables. To account for rater idiosyncratic scale usage that may mask other relationships, *Rater Deviation* is included at this stage of model building. This model corresponds to a test for fixed interaction effects, while allowing the intercepts (β_{0j}) for each rater to vary (as denoted by including the random component u_{0j} to

model the rater intercepts in Equation 2). All other slopes are treated as fixed for this stage of modeling.

$$Rating_{ij} =$$
$$\beta_{0j} + \beta_{1j}(Perf_{ij}) + \beta_2(Job_{ij}) + \beta_3(PerfxJob_{ij}) + \beta_4(Diag_{ij}) + \beta_5(DiagxJob_{ij}) + r_{ij}$$
$$\beta_{0j} = \gamma_{00} + \gamma_{01}(Rater\ Dev_j) + u_{0j}$$

Separate models were built with the *Physical Rating Scale* (for combat leadership) and the

Decision Making Rating Scale (for non-combat leadership) as outcome variables (denoted in

these equations as *Rating* for simplicity), so that the hypothesized interaction could be tested for

each leadership prototype separately. The HLM equation above proposes that there are

systematic errors being committed by the rater -- namely bias induced by implicit leadership

theories, reflected in the interaction term $PerfxJob_{ij}$. This term represents the systematic

influence of raters' leader prototype schemas after controlling for true performance. Support for

Hypothesis 1 would come in the form of a significant β_3 estimate. To reframe the equation in

context, ratings of physical fitness will differ for CA soldiers compared to NCA soldiers of

equivalent APFT performance. After controlling for true performance ($Perf_{ij}$) in physical

fitness (as measured by an Army Physical Fitness Test), a significant interaction term indicates

that raters are either not attending to performance in the same way for each job (i.e. a failure to

encode and recall, consistent with social cognitive theories of stereotype use), or that their

expectations of performance for each job differ, leading to more extreme rating of the

stereotyped group (in keeping with Jussim et al.'s expectancy violation theory). If Jussim et

al.'s (1987) theory holds we would expect the intercept for CA trainees would be lower than for

NCA, indicating that CA trainees who perform poorly on the APFT are evaluated more severely

than are NCA, because they have violated the expectation that combat leaders should perform

well in this domain. Likewise, we would expect the NCA intercept to be lower than the CA

intercept in decision making, demonstrating that low performing NCA soldiers are stereotype violators for decision making ability.

Hypothesis 2 is tested via the interaction term *Diag x Job* in Equation 1 above, representing the interaction between a diagnostic attribute and job category. A significant estimate β_5 would indicate that the relationships between diagnostic attributes differ significantly by job category. We would expect that two attributes considered diagnostic for a job category are conceptually related in the mind of the rater, causing the rater to evaluate soldiers in similar ways for these two attributes. Thus, the slope of *Physical/Perseverance* should be greater for CA soldiers than for NCA. Likewise, the *Decision Making/Planning* slope should be significantly greater for the NCA soldiers than for the CA soldiers.

The second step in the model building process is to include trainee demographic covariates one at a time, and test for significant main effects and interactions. Specifically, prior enlisted status, commissioning source, and component status may contribute to a complex stereotype that rater's have about members of those categories. For instance, a rater may have greater expectations of a prior enlisted soldier, since that soldier presumably has more experience and understanding of Army requirements for the job. With regard to commissioning source, raters may have expectations about the level of preparation received at the United States Military Academy (a four year curriculum of military indoctrination) versus Officer Candidate School (which lasts only 15 weeks), assuming that officers from a particular source will perform better than others. Lastly, a trainee's status as either active duty or reserve/national guard can bias raters as well, since reservists and National Guard members may be perceived as being less proficient than their active duty counterparts. Any variables and interactions found to be significant would be retained in the model.

The third step in the model building process is to test whether any of the partial slopes (β_{kj}, where k refers to the number of terms in the final Level-1 model) differ significantly by rater. A random component (u_{k0}) is estimated for each of the partial slopes. If the random component estimates are significant (as evidenced by a Wald Z estimate significantly different than zero, or a significant Chi Square model comparison test) then the variability of these slopes could be further modeled.

In this case, *Tenure* (time in service measured in months) and *Rater Job* category (CA or NCA) are Level-2 variables that could explain between-rater variability. Modeling the influence of these variables on group means (β_{0j}) and slopes (β_{1j} through β_{kj}) results in the following system of equations (with *Covar* standing in for all significant level-1 covariates that were retained in the model from the previous step):

$$Rating_{ij} = \beta_{0j} + \beta_{1j}(Perf_{ij}) + \beta_{2j}(Job_{ij}) + \beta_{3j}(PerfxJob_{ij}) + \beta_{4j}(Diag_{ij})$$
$$+\beta_{5j}(DiagxJob_{ij}) + \beta_{kj}(Covar_{ij}) + r_{ij}$$
$$\beta_{0j} = \gamma_{00} + \gamma_{01}(Rater\ Dev) + \gamma_{02}(Rater\ Tenure) + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + \gamma_{11}(Rater\ Tenure) + u_{1j}$$
$$\beta_{2j} = \gamma_{20}+\gamma_{21}(Rater\ Tenure) + u_{2j}$$
$$\beta_{3j} = \gamma_{30}+\gamma_{31}(Rater\ Tenure) + u_{3j}$$
$$\beta_{4j} = \gamma_{40}+\gamma_{41}(Rater\ Tenure) + u_{3j}$$
$$\beta_{5j} = \gamma_{50}+\gamma_{51}(Rater\ Tenure) + u_{3j}$$
$$\beta_{kj} = \gamma_{k0}+\gamma_{k1}(Rater\ Tenure) + u_{kj}$$

Significant estimates of γ_{02} would indicate that variability in rater intercepts can be attributed to the rater's respective experience level. Significant γ_{11} through γ_{k1} suggest that the relationship between the level 1 variables and the outcome variable are influenced by rater experience. Likewise, for rater job category:

$$Rating_{ij} = \beta_{0j} + \beta_{1j}(Perf_{ij}) + \beta_{2j}(Job_{ij}) + \beta_{3j}(PerfxJob_{ij}) + \beta_{4j}(Diag_{ij})$$
$$+\beta_{5j}(DiagxJob_{ij}) + \beta_{kj}(Covar_{ij}) + r_{ij}$$
$$\beta_{0j} = \gamma_{00} + \gamma_{01}(Rater\ Dev) + \gamma_{02}(Rater\ Job) + u_{0j}$$
$$\beta_{1j} = \gamma_{10} + \gamma_{11}(Rater\ Job) + u_{1j}$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21}(Rater\ Job) + u_{2j}$$
$$\beta_{3j} = \gamma_{30} + \gamma_{31}(Rater\ Job) + u_{3j}$$
$$\beta_{4j} = \gamma_{40} + \gamma_{41}(Rater\ Job) + u_{3j}$$
$$\beta_{5j} = \gamma_{50} + \gamma_{51}(Rater\ Job) + u_{3j}$$
$$\beta_{kj} = \gamma_{k0} + \gamma_{k1}(Rater\ Job) + u_{kj}$$

Of particular interest is the cross-level interaction between Rater Job and Trainee Job. A positive and significant interaction would be interpreted as a "similarity effect" in which CA raters evaluate CA trainees more favorably than NCA trainees.

Research Question. If one accepts that the systematic distortion hypothesis described by Hypothesis 2 is prevalent across a set of attributes, then correlation coefficients in a combined sample would be distorted. This will be due in part by diagnostic variables having different correlations with other variables, and in part by diagnostic variables having different relationships with ratings of overall leadership effectiveness. For this analysis rated variables will be used to predict overall leadership effectiveness ratings for the combined sample, then for each job category separately using a forward selection procedure as is commonly used in the selection of variables for a predictive measurement instrument (as described by Guion, 1998, and demonstrated by Horey et al, 2007) . If such a forward selection procedure results in a set of predictors that differ according to job category, the effect of implicit leadership theories are evident in the ratings. Specifically, it would be expected that combat diagnostic variables will predict overall leadership effectiveness of combat soldiers, but not predict overall leadership of non-combat soldiers. Likewise, non-combat leader diagnostic variables will predict overall leadership for non-combat soldiers but not for combat leaders.

RESULTS

Sample Characteristics and Descriptive Statistics

Table 1 displays the sample sizes, means and standard deviations for the total sample, cross-classified by trainee job category. Appendix 4 contains the sample sizes and standard deviations by rater for the outcome variables used in Hypothesis 1 and 2. Appendix 5 contains the frequency tables for the trainee demographic covariates. The sample size for Physical Scale Rating differs from the total sample size because two influential outliers were removed from the analysis. These two outliers were the only two individuals in the data set (one CA and one NCA) assigned the rating of "Needs Much Improvement" (a value of 1) for the *Physical* variable. The standardized residuals calculated on the final Combat Leader model exceeded -3 for both of these individuals. The impact of removing these two individuals will be readdressed in the Discussion.

Since one squad leader in first platoon evaluated two squads, the number of level-2 units was 19. Due to missing data, the effective sample size for the Decision Making analyses was 167. Since the missing data represented an entire squad, no inference can be drawn about that rater's implicit leadership theories, so the decision was made to not impute this missing data.

Table 3

Means, Standard Deviations, and Sample Sizes for Performance and Rated Outcome Variables

	APFT	SJT	Physical Scale	Decision Making Scale
CA				
M	267.03	9.51	3.33	3.15
SD	33.77	2.62	.6050	.4612
n	91	91	90	86
NCA				
M	261.44	8.93	3.19	3.24
SD	36.69	2.98	.6339	.5112
n	80	80	80	79
Total				
M	264.51	9.22	3.27	3.19
SD	35.02	2.80	.6198	.4839
N	173	173	172[a]	167[b]

Note. [a] The differences in sample sizes across variables are the consequence of missing data. [b] One entire squad was missing data for both items of the Decision Making Scale, and this squad was excluded from analysis.

The CA/NCA sample means in both the *Physical* and *Decision Making* scales demonstrate a slight bias in the direction predicted by the combat/non-combat leadership prototypes. CA soldiers are rated slightly higher than NCA in Physical attributes, while NCA soldiers are rated slightly higher in Decision Making attributes (despite the fact that CA trainees have a higher average on the decision making performance measure, the SJT). However, the mean differences between CA and NCA on APFT (5.595, $F(1, 169)= 1.078$, $p=.303$), SJT (.58, $F(1, 169) = 1.829$, $p=.178$), Physical Scale (.134, $F(1, 168) = 1.987$, $p = .161$), and Decision Making Scale (-.0821, $F(1, 163) = 1.176$, $p=.280$) are not significant.

A comparison of a random intercept model versus a fixed intercept model confirms that HLM is appropriate. Tables 4 and 5 show that the random-intercept models fit the data significantly better than the fixed-intercept models, $\chi^2(1, N = 170) = 13.54$, $p<.001$ for the Combat Leader model, and $\chi^2(1, N = 167) = 44.53$, $p < .001$ for the Non-combat Leader model.

35

This suggests that collapsing the raw data across raters without accounting for rater mean differences is inappropriate. As suspected, a significant amount of variance is attributable to rater effects, with an intra-class correlation coefficient of .188 for the Combat Leader model, and .420 for the Non-combat Leader model. Thus 19% of the variability in Physical Scale ratings and 42% of the variability in Decision Making Scale ratings are attributable to rater effects.

As shown in the Figure 4 line graph, one source of variability is idiosyncratic use of the rating scale. The raters' average ratings (across individuals and across attributes) differ from each other. To model this variability, the rater's deviation from the grand rating mean was included as a Level-2 variable explaining rater intercept variability. As reported in Tables 6 and 7, Rater Deviation was found to be highly significant, accounting for 71% of intercept variance in the Combat Leader model and 97% of the intercept variance in the Non-combat Leader model). After accounting for the idiosyncratic scale usage, the variability associated with rater intercepts is no longer significant (Wald Z=1.093, p=.274 for Physical model and Wald Z=.442, p=.658 for the Decision Making model). The individual level variance (r_{ij}) remains unchanged and significantly different than zero, suggesting that it can be further modeled.

Table 4

Results of HLM Comparing Fixed- and Random-Intercept Combat Leader Models

Fixed Effects	Fixed-Intercept (Reduced) Model				Random-Intercept (Full) Model			
	Coef.	*SE*	*t*	*p*	*Coef.*	*SE*	*t*	*p*
For Intercept (β_{0j})								
Intercept (γ_{00})	3.294	0.044	75.087	0.000	3.309	0.070	47.310	0.000
Random Effects	*Var. Comp*	*SE*	*Wald Z*	*p*	*Var. Comp*	*SE*	*Wald Z*	*p*
Intercept (u_{0j})	0.327	0.036	9.192	0.000	0.062	0.031	1.997	0.046
Level-1 (r_{ij})					0.268	0.031	8.692	0.000
Model Fit	*Deviance*	*Parameters*	*AIC*	*BIC*	*Deviance*	*Parameters*	*AIC*	*BIC*
	295.93	2	297.93	301.055	282.39	3	286.39	292.65

Table 5

Results of HLM Comparing Fixed- and Random-Intercept Non-Combat Leader Models

Fixed Effects	Fixed-Intercept (Reduced) Model				Random-Intercept (Full) Model			
	Coef.	*SE*	*t*	*p*	*Coef.*	*SE*	*t*	*p*
For Intercept (β_{0j})								
Intercept (g_{00})	3.193	.0374	85.27	.000	3.212	.0805	39.89	.000
Random Effects	*Var. Comp*	*SE*	*Wald Z*	*p*	*Var. Comp*	*SE*	*Wald Z*	*p*
Intercept (u_{0j})	.234	.0257	9.110	.000	.1049	.0426	2.465	.014
Level-1 (r_{ij})					.1448	.0169	8.567	.000
Model Fit	*Deviance*	*Parameters*	*AIC*	*BIC*	*Deviance*	*Parameters*	*AIC*	*BIC*
	235.24	2	237.24	240.35	190.708	3	194.708	200.93

Figure 4. Averages by rater across all dimensions.
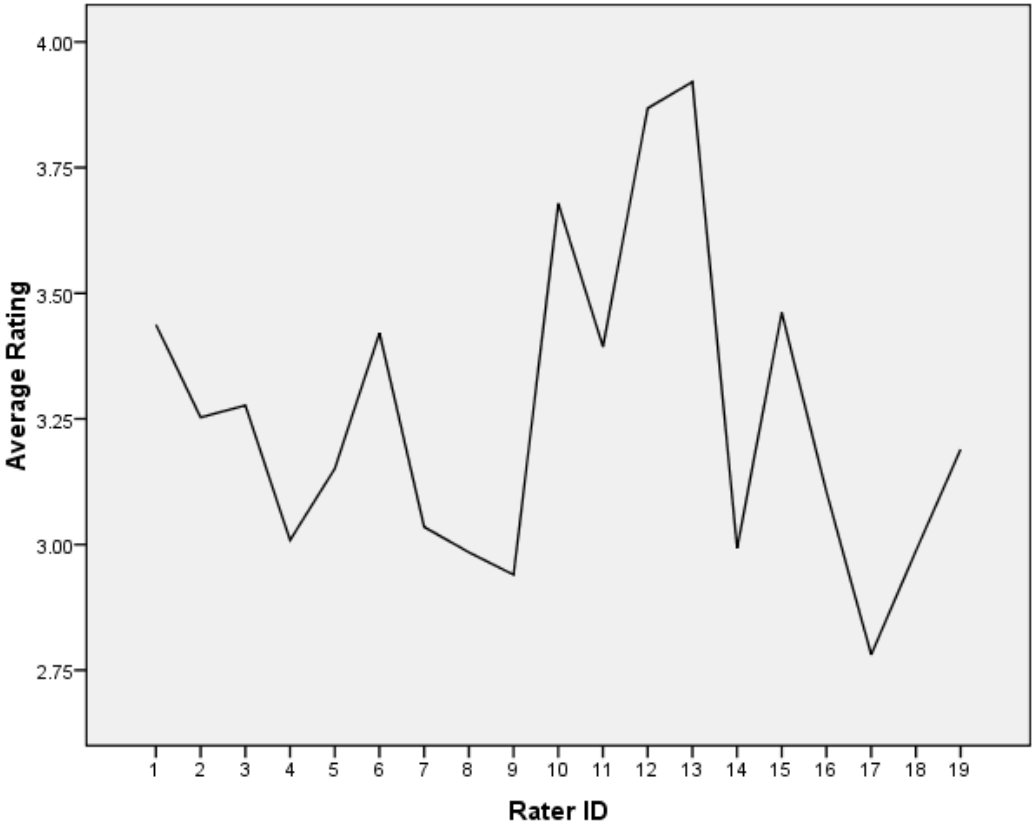
Table 6

Results of HLM Model Comparison for Combat Leader Model

	Null Model				Model 1			
Fixed Effects	*Coef.*	*SE*	*t*	*p*	*Coef.*	*SE*	*t*	*p*
For Intercept (β_{0j})								
Intercept (γ_{00})	3.309	.070	47.310	.000	3.302	.050	65.730	.000
Rater Deviation (γ_{01})					.712	.167	4.259	.000
Random Effects	*Var. Comp*	*SE*	*Wald Z*	*p*	*Var. Comp*	*SE*	*Wald Z*	*p*
Intercept (u_{0j})	.062	.031	1.997	.046	.018	.016	1.093	.274
Level-1 (r_{ij})	.268	.031	8.692	.000	.268	.031	8.728	.000
Model Fit	*Deviance*	*Parameters*	*AIC*	*BIC*	*Deviance*	*Parameters*	*AIC*	*BIC*
	282.39	3	286.39	292.65	270.929	4	274.929	281.177

Table 7

Results of HLM model comparison for Non-Combat Leader Model

	Null Model				Model 1			
Fixed Effects	*Coef.*	*SE*	*t*	*p*	*Coef.*	*SE*	*t*	*p*
For Intercept (β_{0j})								
Intercept (γ_{00})	3.212	.0805	39.89	.000	3.2112	.0321	100.016	.000
Rater Deviation (γ_{01})					1.0602	.113	9.369	.000
Random Effects	*Var. Comp*	*SE*	*Wald Z*	*p*	*Var. Comp*	*SE*	*Wald Z*	*p*
Intercept (u_{0j})	.1049	.0426	2.465	.014	.0030	.0067	.442	.658
Level-1 (r_{ij})	.1448	.0169	8.567	.000	.1436	.0166	8.631	.000
Model Fit	*Deviance*	*Parameters*	*AIC*	*BIC*	*Deviance*	*Parameters*	*AIC*	*BIC*
	190.708	3	194.708	200.932	158.584	4	162.584	168.786

The Combat Leader and Non-combat Leader final models are reported in Table 8 and 11. In the first stage of the model building the study variables and interaction terms for hypothesis 1 and 2 were included in the model. Then the variables *age*, *component status* (active duty or reserve/national guard), *prior enlisted status*, and *commissioning source* (OCS, ROTC, USMA or other) were added individually, along with two-way and three-way interaction terms between the covariate and the existing model variables. Analyses of these two models are reported below.

Combat Leader Model

In the Combat Leader model, the demographic covariates *age* and *component status* were not significant as either main effects or in interactions with the variables of interest in explaining *Physical Scale* ratings. They were not retained in the model. The interaction between *prior enlisted status* and *trainee job*, $F(1,153.126)=3.059$, p=.082, was significant at α=.10, as was *commissioning source,* $F(3,147.9)=2.944$, p=.035. *Trainee job* interacted significantly at α=.10 with *APFT* in support of Hypothesis 1, as well as with ratings of *perseverance*. The *Physical Scale* rating estimated means for *trainee job* at both levels of *prior enlisted* category, high/low *APFT* and high/low *perseverance* ratings are provided in Table 9. Pairwise comparisons of *commissioning source* categories reveals that Warrant Officers/Direct Commissions had significantly lower Physical Scale ratings than any of the other commissioning sources. None of the other commissioning sources differed from each other.

Table 8

Results of HLM Final Combat Leader Model

Fixed Effects	Model 2			
	Coef.	*SE*	*t*	*p*
Intercept (γ_{00})	1.78	.396	4.505	.000
Rater Deviation (γ_{01})	.685	.166	4.117	.001
PFT (γ_{10})	.005	.001	3.506	.001
Trainee Job (γ_{20})	-.990	.523	-1.891	.061
PFT x Trainee Job (γ_{30})	.004	.002	1.956	.052
Prior Enlisted Status (γ_{40})	-.233	.106	-2.190	.030
Prior x Trainee Job (γ_{50})	.253	.145	1.749	.082
Perseverance$_{ij}$ – Perseverance$_j$ (γ_{60})	.401	.100	4.008	.000
Trainee Job x Perseverance (γ_{70})	-.364	.151	-2.407	.017
Commission = OCS (γ_{80})	.374	.180	2.084	.039
Commission = ROTC (γ_{90})	.327	.117	2.795	.006
Commission = USMA (γ_{100})	.239	.139	1.715	.088
Commission = Warrants/Direct Com (γ_{110})[a]	---	---	---	---

Note: [a] Set as the reference category.

Table 9

Estimated Physical Scale Means

	Prior Enlisted		Rated Perseverance		APFT	
	Yes	No	Low	High	Low	High
CA	3.42	3.40	3.37	3.44	2.80	4.00
NCA	3.16	3.39	2.95	3.68	2.98	3.65

Note. The predicted means are calculated with ROTC as the reference group, and means computed with covariates at the following values: RaterDev=-.0029, Trainee Job = .53, APFT = 265.10, perseverance = .0046, prior enlisted =.32.

Table 10
Estimated Means for Commissioning Source

	Mean	SD
OCS	3.409	0.148
ROTC	3.361	0.055
USMA	3.274	0.088
Warrants/Direct Commissions	3.035	0.118

Note. Means computed with covariates at the following values: RaterDev=-.0029, Trainee Job = .53, APFT = 265.10, perseverance = .0046, prior enlisted =.32.

The final Combat Leader model explained an additional 41% of the residual variance above and beyond the random intercept-only model reported in Table 4. The intercept variance estimate (.029, SE=.016) in this final model is not significantly different than zero (Wald Z=1.76, p=.078) at α=.05, suggesting that intercept variability should not be further modeled. To test whether the level-1 parameters varied by rater, random error terms were included in the model for each level-1 parameter. None of these random variance terms were significantly different than zero based on a Wald Z test statistic, suggesting that rater effects (rater tenure and rater job) should not be further modeled.

Non-combat Leader Model

In the Non-combat Leader model, no two or three-way interactions were found between the demographic covariates and the existing model variables. Unlike in the Physical model, neither *prior enlisted status*, F(1,141.6)=.792, p=.375, nor *commissioning source*, F(3,142.9)=1.926, p=.128, significantly influenced Decision Making ratings when included in the model individually. The interaction between *SJT* and *trainee job* was significant, in support of Hypothesis 1. However, the interaction between *planning* ratings and *trainee job* were not significant, contrary to Hypothesis 2. Unexpectedly, *SJT* scores did interact significantly with *planning* ratings in explaining decision making ratings. Table 12 reports the predicted means by trainee job category and significantly interacting variables.

The final Non-combat Leader model explained an additional 16% of the residual variance above and beyond the random intercept-only model. The intercept variance estimate (.0016, SE=.0058) is not significantly different than zero (Wald Z=.281, p=.779) suggesting that no between rater variance can be modeled using level-2 predictors. All level-1 random variance estimates were not significantly different from zero based on Wald Z tests.

Table 11

Results of HLM Final Non-combat Leader Model

	Model 2			
Fixed Effects	*Coef.*	*SE*	*t*	*p*
Intercept (g_{00})	3.17	.126	25.095	.000
Rater Deviation (g_{01})	1.08	.107	9.996	.000
SJT (g_{10})	.004	.014	.326	.745
Trainee Job (g_{20})	.429	.208	2.065	.041
SJT x Trainee Job (g_{30})	-.045	.021	-2.105	.037
$Planning_{ij}$ – $Planning_i$ (g_{40})	1.27	.462	2.738	.007
Planning x Trainee Job (g_{50})	.075	.186	.403	.688
SJT x Planning (g_{60})	-.092	.045	-2.059	.041

Table 12

Estimated Decision Making Scale Means

	SJT			Rated Planning	
	Low	High		Low	High
CA	3.36	3.04	SJT Low	2.62	3.15
NCA	3.17	3.21	SJT High	4.00	3.77

Note. The predicted means are calculated at means computed with covariates at the following values: RaterDev=-.0029, Trainee Job = .53, SJT = 9.23, and planning = .0046.

Research Question

The interaction effects found in the tests of hypothesis 1 and 2 suggest that there are some differences in the prototypes held about combat and non-combat leaders, leading to different

43

encoding, recall, and inferences made about their respective attributes. To explore this further,

an approach is used that is common in predicting job outcomes in personnel decision, linear

multiple regression using a forward selection procedure (Guion, 1998). This approach is

admittedly exploratory. The first variable enters the predictive equation based on its zero-order

correlation with the outcome variable; thereafter, variables enter based on their partial correlation

with the outcome variable after controlling for other variables. As a consequence, variables that

are highly correlated with each other, or equally correlated with the outcome variable will

"compete" for entrance in the equation (Guion, 1998). Thus, the final predictive model

capitalizes on chance. However, this approach emphasizes how implicit leadership theories can

impact such predictions. The order in which the variables enter the equation for combat and

non-combat leaders reveals their importance to the prototypes held by the raters. The zero order

correlations between attributes and the outcome variables are reported in Table 13.

Table 13

Zero-order Correlations Between Rated Attributes and Overall Net Assessment

	r		
	Combined	*NCA*	*CA*
Acts Responsibly	.363	.466	.278
Adaptability	.517	.485	.470
Assessment/Evaluating skill	.399	.385	.318
Building teams/groups/units	.391	.377	.331
Coaching/Counseling/Empowering	.532	.467	.572
Communication skill	.460	.386	.520
Dealing with Changing Situations	**.556**	**.660**	.432
Desire/Will/Initiative/Mental Discipline	.417	.442	.320
Develops subordinates	.368	.451	.256
Duty	.496	.483	.441
Executing	.550	.552	.546
Handling Emergencies	.554	.585	.543
Handling Work Stress	.539	.534	.534
Honor	.300	.172	.304
Integrity	.377	.348	.314
Interpersonal Adaptability	.354	.393	.306
Leading an Adaptable Team	.524	.387	**.642**
Loyalty	.391	.409	.281
Makes Tradeoffs	.340	.399	.187
Motivated	.488	.567	.425
Motivating	.497	.477	.560
Perseverance	.513	.358	.598
Personal Courage	.425	.385	.486
Physical Adaptability	.487	.537	.433
Physical Fitness/Bearing	.495	.455	.465
Planning skill	.398	.287	.411
Relies (Teamwork)	.231	.166	.224
Respect	.289	.200	.374
Seeks self-improvement and org change	.397	.547	.155
Self-Control/Calm	.466	.397	.466
Selfless Service	.456	.466	.409
Sets Priorities	.366	.380	.304
Solving Problems Creatively	.416	.398	.447
Sound and Logical Reasoning	.427	.470	.358
Sound Judgment/Critical Think	.479	.495	.404
Tactical Proficiency	.516	.569	.444
Technical Expertise	.374	.451	.295

The ratings available in both the End-of-Course Leadership Evaluation and the Adaptability Rating Scale were entered as predictors of the *Overall Net Assessment* (*ONE*) in three analyses: the combined sample (N=173), the CA-only sample (n=79), and the NCA-only sample (n=80). The results of the three analyses are reported in Tables 14. In the combined analysis, eleven variables were entered into the prediction equation using a criterion of α=.05 to enter, explaining 73.8% of the variance in *ONE* ratings. *Dealing with Changing Situations (DCS)* had the highest zero-order correlation (r = .556) with *ONE* ratings for the combined sample, explaining 30.9% of the variance for that model. (It should be noted that the next highest zero-order correlation, r = .554, in the combined sample is between the variable *Handling Emergencies* (*HE*) and *ONE*, and the difference between the *DCS* and *HE* correlations with the outcome variable are probably not meaningful. The small correlational advantage held by *DCS*, and the significant correlation between it and *HE*, r=.349, effectively precludes *Handling Emergencies* from entering the prediction equation as a meaningful predictor of *ONE*.) For the NCA sample, *Dealing with Changing Situations* also entered the equation first, with a zero order correlation of r=.660, explaining 44% of the variance for NCA *Overall* ratings.

Only two variables, *Executing* and *Motivating* predicted *ONE* ratings in the combined and Combat/Non-combat leader samples. *Building* -- which is described as spending time and resources to improve teams -- predicts overall net assessment for the combined sample, but not significantly for either of the subsamples.

Importantly, there are four variables that are predictive of overall leadership for CA soldiers that do not appear in the combined sample prediction equation. *Acts Responsibly to Other Soldiers, Leading an Adaptable Team, Duty,* and *Mental Leadership* are meaningful predictors of combat leadership. These variables reflect the same constructs that have previously

been presented as diagnostic of combat leaders. Duty and Mental Leadership both include the terms "initiative", *Act Responsibly* (along with *Physical Adaptability* that also appears in the combined prediction model) describes the acceptance of physical hardship.

Table 14

Comparison of Final Step in Hierarchical Linear Regression Results Predicting Overall Net Assessment Using Forward Selection Procedure.

Variable	Combined Sample (N = 173)			Non-combat (n = 80)			Combat (n = 79)		
	B	*SE B*	β	*B*	*SE B*	β	*B*	*SE B*	β
Executing	.337	.054	.299	.343	.078	.292	.321	.068	.296
Motivating	.134	.049	.129	.239	.081	.211	.139	.063	.148
Building	.299	.088	.148						
Dealing with Changing Work Situations	.177	.053	.164	.189	.085	.182			
Handling Work Stress	.168	.045	.180	.241	.064	.251			
Personal courage	.266	.073	.173	.424	.107	.251			
Sets Priorities	.173	.063	.127	.221	.101	.136			
Technical Skill	-.123	.060	-.101	-.222	.083	-.181			
Communicating	.136	.048	.132	.198	.067	.207			
Physical Adaptability	.102	.039	.122				.139	.052	.175
Selfless Service	.162	.057	.132				.216	.085	.162
Acts Responsibly to Other Soldiers							.188	.090	.137
Duty							.211	.092	.147
Leading an Adaptable Team							.211	.070	.222
Mental Leadership							.182	.068	.174

Note. Combined sample R^2 = .74 for Step 11, Non-combat sample R^2 = .77 for Step 8, Combat sample R^2 = .73 for Step 8 (ps < .05)

DISCUSSION

Following Lord, Foti and Phillips (1983) paradigm of leadership categorization, attributes at the subordinate level of categorization serve to distinguish between leader prototypes instead of reflect commonalities among them. At the highest (super-ordinate) level, some attributes distinguish leaders from non-leaders. At the basic level, leadership attributes have specific factor structures depending on a job-related context (Lord & Maher, 1991), in this case that of a military leader. This study categorizes leadership at a subordinate level of military leadership, using the salient categorization of combat versus non-combat leadership, and suggests that implicit leadership prototypes exist at this level as well. As has been found in other realms of stereotype research, prototypes may be inducing biased ratings by influencing the encoding, recall, and expectations of performance in the stereotyped domains.

Hypothesis 1

This study sought to detect biased ratings in those domains that are considered diagnostic of leadership for combat and non-combat soldiers. When two soldiers of different categories have equivalent performance in a domain such as physical fitness, yet are evaluated differently, the case can be made that the difference reveals bias toward a specific prototype. In this case, raters evaluated soldiers differently in the physical and decision making domains in ways that were consistent with Jussim et al.'s (1987) expectancy violation hypothesis. Combat soldiers were evaluated more severely than non-combat soldiers when their performance in the physical domain was low, since this violated prototype expectations. Likewise, non-combat soldiers were evaluated more severely when their performance was low in the decision making domain. It

49

must be noted that the interaction effect to test Hypothesis 1 was significant in the model after removing two influential outliers. Further investigation of these individuals revealed that both soldiers had PFT scores 2 SD below the mean, but one was a combat arms soldier and one was a non-combat arms soldier. These two were the only individuals who received a rating of 1 on the attribute *Physical*. The combined effect of these two individuals was to make the interaction effect (Job x PFT) non-significant when they were included in analysis of the Combat Leader model. The results of the Combat Leader model including these two individuals are reported in Appendix 6.

Stereotypes appear to exist for other categorizations that were not hypothesized. In the domain of physical fitness, prior enlisted status and commissioning source influenced physical ratings. Non-combat soldiers who were prior enlisted received lower physical ratings (3.16) than comparably performing non-prior soldiers (3.39) and combat soldiers of either status (3.40 for prior, 3.42 for non-prior). This might be another substantiation of Jussim et al.'s (1987) expectancy violation. Raters may believe that prior enlisted soldiers should not only understand the physical requirements of the military leadership, but have had the opportunity to develop in that domain during their tenure as enlisted soldiers. Thus, failing to meet physical standards for this group would be highly penalized.

Commissioning source also influenced ratings in the physical domain, but did not interact with combat/non-combat designation. In this sample, 64% of the warrant officers and direct commissions were prior enlisted, and these commissioning sources together were rated significantly lower than the other commissioning sources in the physical domain. Since all of the other commissioning sources were predominantly non-prior enlisted, the low physical ratings for the warrant officers/direct commissions seem to reinforce a prior-enlisted negative bias. Of the

warrants/direct commissions, 4 were Military Police, 4 were Armor, 2 were Infantry, 1 was Air Defense, 1 was Adjutant General and 1 was unspecified. The high number of military police (a non-combat arms MOS) in the warrant officer category may have influenced the ratings for this category. Since military police provide combat support, to include directly interacting with enemy forces at vehicle check points and prisoner control missions, the prototype of the military police leader may more closely align with that of the combat leader. As a result, these officers would be penalized severely for not upholding the combat leader prototype in the physical domain.

Hypothesis 2

The direct test of Hypothesis 2 received only partial support. The relationship between *Physical* ratings and *Perseverance* ratings did differ by trainee job category, as predicted. Since these two variables were considered diagnostic of CA leaders, it was further hypothesized that the relationship between the two variables would be stronger for CA than for NCA soldiers. This was indeed the case. However, this difference in relationships did not play out in the decision making domain. The relationship between *decision making* and *planning* did not differ significantly by job category. It may be the case that the type of decision making and planning undertaken in this context was highly tactical in nature, and thus CA soldiers and NCA would be rated equally well in both the decision making and planning domains. An alternative explanation is that planning, decision making, or both, are attributes that are common at the super-ordinate level of the leadership hierarchy. In other words, they are common to the prototype of a leader versus a non-leader, and do not differentiate between subordinate level prototypes as theorized.

Research Question

The physical and decision making domains reflected direct tests of the bias hypotheses because "objective" measures of these domains were available in the form of the Army Physical Fitness Test and the Situational Judgment Test. Thus we could control for the observation of performance that would directly influence raters' evaluations. However, this does not preclude a demonstration of how ratings on leadership attributes conform to different prototypes of combat and non-combat leaders. To explore this, a forward selection procedure in linear regression was used to determine which variables would be selected for inclusion in a hypothetical leadership measure, predicting overall leadership rating for combat and non-combat leaders separately. The foremost expectation for this analysis was that different sets of attributes would be selected for each category of leader. This is indeed what occurred. Only two of eleven attributes that meaningfully predict overall ratings in the combined sample (*Execution* and *Motivating*) were meaningful for both combat and non-combat leaders, reinforcing Lord, Foti and Philips assertion that few variables are common to all leadership prototypes. Other attributes, such as *Building*, were included in a model predicting overall assessments for the combined sample, but were *not* meaningful in predicting overall assessments for *either* combat or non-combat soldiers. The remaining variables selected for inclusion in the respective models were unique to either combat arms or non-combat arms. To a degree, those attributes that would be expected to appear in the combat leader prediction model did appear, including variables that described the focus on physical fitness and taking the initiative that is well supported in the military literature.

At a basic level, the results of the predictive analysis reveal that the factor structures of attributes related to leadership do in fact differ at the subordinate categorization (combat versus non-combat) level. Variables are included in the model based on their correlation with the

outcome variable (overall net assessment) after controlling for other variables in the model. The difference in the selection procedure reveals differences in these underlying correlations.

Limitations

The greatest limitation of this study is that it assumes the sample of combat arms and non-combat arms soldiers do not differ significantly in any of the underlying attributes upon which they were rated. In measuring NCO leadership performance, Knapp et al (2004) found that soldiers were very often rated differently on overall performance measures, performed differently on some criterion measures based on MOS. If systematic differences exist, then the case could be made that ratings reflected these true differences instead of reflecting the influence of a leadership prototype held by raters. In this study, only physical fitness and decision making "true" performance could be controlled for, and the samples were found to not differ significantly on these two dimensions. However, it could be the case that soldiers in this sample do differ systematically on other dimensions, and these differences could not be assessed in the current study.

A second limitation involves the use of pre-existing measurement instruments instead of measures designed specifically for the study. The rated dimensions were often confounded with what would typically be considered related but distinct constructs. For instance, physical fitness and bearing are related to each other, but are in no way isomorphic, yet are represented in one rating (consistent with the Army Officer Evaluation System, Dept of the Army, 1998b, Knapp et al., 2004). One could argue that it would be possible to have physical fitness without proper military bearing, and vice versa, and it is impossible to know which of these attributes was the focus of the raters' assessment.

Another cause for concern in the Decision Making model is the huge amount of variance (97%) attributable to rater's idiosyncratic scale usage. The high variability between raters, coupled with the fact that no significant relationship existed between the decision making rating and SJT scores, may suggest that ratings associated with decision making were somewhat random, and not based on observations of decision making during training. Unlike the Army Physical Fitness Test, which was witnessed and strongly encoded by the raters when it occurred, it is unclear how performance in decision making was observed. Indeed, it could be that the low correlation between SJT scores and decision making ratings reflects that performance on this test was not observed by the raters.

Lastly, there are limitations associated with the use of existing job analysis literature and military doctrine to determine which attributes are diagnostic of combat and non-combat leadership. In many cases, the dimensions described by the job analysis and military literature were not isomorphic with those available in this study. A more direct measure of leadership prototypes is warranted, perhaps using the procedure described by Lord and Maher (1991), in which individuals are asked to describe the degree to which attributes are prototypical of certain leaders, based on a scale of "highly prototypical" to "not prototypical", and then developing and testing a rating form that includes these attributes specifically.

Implications

Several researchers have suggested that using ratings upon which to develop a leadership model degrades the validity of the resulting model, since the ratings themselves are reflective of implicit leadership theories that are job-specific and therefore not generalizable to different contexts (Eden & Leviatan, 1975; Rush, Thomas & Lord, 1977). This study bears out this assertion. At the basic level, we can see evidence that in some domains, ratings are more

consistent with the raters' pre-existing schemas based in organizational culture than with the observed performance of the individuals being rated. When assigning overall evaluations of soldiers' leadership, raters appear to be influenced by a largely different set of underlying attributes. When rating combat soldiers, for instance, attributes related to adaptability, taking the initiative and physical endurance are meaningful, consistent with military doctrine related to the requirements for combat leadership. For non-combat soldiers, being able to operate with limited information, remain calm under stress and pressure, and changing plans and priorities according to mission requirements predict leadership ratings. This study demonstrates that ratings are distorted in accordance with implicit leadership theories, and that inferences made about which attributes predict leadership effectiveness should be viewed with caution when based on performance ratings.

REFERENCES

Ashmore, R. D., & Del Boca, F. K. (1979). Sex Stereotypes and Implicit Personality Theory:
Toward a Cognitive-Social Psychological Conceptualization. *Sex Roles: A Journal of
Research*, *5*(2), 219-48.

Bliese, P. D. (2002). Multilevel random coefficient modeling in organizational research:
Examples using SAS and S-PLUS. In F. Drazgow & N. Schmitt (Eds.), *Measuring and
analyzing behavior in organizations: Advances in measurement and data analysis* (pp.
401-445). San Francisco: Jossey-Bass.

Bartlett, F. C. (1932) *Remembering*. New York: MacMillan Co.

Brown, P. & Turner, J. (2002) The role of theories in the formation of stereotype content. In C.
McGarty, V. Yzerbyt and R. Spears (Eds.) *Stereotypes as explanations: The formation of
meaningful beliefs about social groups* (pp. 67-89). New York, NY: Cambridge
University Press.

Bruner, J. S. (1958) Social psychology and perception. In E. E. Maccoby, T. M. Newcomb, & E.
L. Hartley (Eds.), *Readings in social psychology*, p.85-94. New York: Holt, Rinehart &
Winston.

Campbell, J. P., & Zook, L. M. (1996). *Building and Retaining the Career Force: New
Procedures for Accessing and Assigning Army Enlisted Personnel (ARI Research Note
96-45).* Retrieved May 3, 2009, from http://handle.dtic.mil/100.2/ADA309090.

Cantor, N., & Mischel, W. (1977). Traits as prototypes: Effects on recognition memory. *Journal
of Personality and Social Psychology, 35,* p. 38-48.

Cooper, W. H. (1981). Conceptual similarity as a source of illusory halo in performance ratings. *Journal of Applied Psychology, 66*, 302-307.

Cullen, B, Klemp, G. & Mansfield, R. (1988). *Junior officer competency model: Research results and application (ARI Research Note 88-10).* Alexandria, VA: U.S. Army Institute for the Behavioral and Social Sciences.

Department of the Army. (1983). *Field Manual 22-100 Military Leadership.* Washington, DC: Headquarters, Department of the Army.

Department of the Army (1998a). *Field Manual 21-20 Physical Fitness Training.* Washington, DC: Headquarters, Department of the Army.

Department of the Army. (1998b). *Army Regulation 623-105 Officer Evaluation Reporting System.* Washington, DC: Headquarters, Department of the Army.

Department of the Army (2006). *Field Manual 6-22 Army Leadership.* Washington, DC: Headquarters, Department of the Army.

Department of the Army. (2007). *Army Regulation 600-100 Army Leadership.* Washington, DC: Headquarters, Department of the Army.

Eagly, A. (1987). *Sex differences in social behavior: A social-role interpretation.* Hillsdale, NJ: Erlbaum.

Eden, D., & & Leviatan, U. (1975). Implicit leadership theory as a determinant of the factor structure underlying supervisory behavior scales. *Journal of Applied Psychology, 60*, 736-741.

Esenck, H.J. Y Crown, S. (1948). National stereotypes: an experimental and methodological study. *International Journal of Opinion and Attitude Research, 2*, 26-39.

Feldman, J.M. (1981). Beyond attribution theory: Cognitive processes in performance appraisal. *Journal of Applied Psychology*, 66, 127-148.

Ford, T.E. and Stangor, C. (1992). The role of diagnosticity in stereotype formation: Perceiving group means and variances. *Journal of Personality and Social Psychology*, 63, 356-67.

Guion, R. M. (1998) *Assessment, measurement, and prediction for personnel decisions.* Mahwah, NJ: Lawrence Erlbaum Associates.

Hastie, R., & Kumar, P. A. (1979). Person memory: Personality traits as organizing principles in memory for behaviors. *Journal of Personality and Social Psychology*, *37*(1), 25-38. doi: 10.1037/0022-3514.37.1.25.

Heider, J. D., Scherer, C. R., Skowronski, J. J., Wood, S. E., Edlund, J. E., & Hartnett, J. L. (2007). Trait expectancies and stereotype expectancies have the same effect on person memory. *Journal of Experimental Social Psychology*, *43*(2), 265-272. doi: 10.1016/j.jesp.2006.01.004.

Hean, S., Clark, J., Adams, K., & Humphris, D. (2006). Will opposites attract? Similarities and differences in students' perceptions of the stereotype profiles of other health and social care professional groups. *Journal of Interprofessional Care*, *20*(2), 162-181. doi: 10.1080/13561820600646546.

Hogg, M.A., & Terry, D.J. (2000). Social identity and self-categorization processes in organizational contexts. *Academy of Management Review .25*(1), 121-140.

Horey, J., & Fallesen, J. (2004). Leadership competencies for contemporary Army operations: Development, review and validation. Paper presented to the International Military Testing Association 2004 Conference. Retrieved 20 May 2010 from http://www.internationalmta.org/Documents/2004/2004045P.pdf.

Horey, J. Fallesen, J., Morath, R., Cronin, B., Cassella, R., Franks Jr., W, & Smith J. (2004)
Competency based future leadership requirements (Technical Report 1148). Arlington,
VA: United States Army Research Institute for the Behavioral and Social Sciences.

Hox, J. (2002). *Multilevel analysis: Techniques and applications*. Mahwah, NJ: Lawrence
Erlbaum Assoc.

Ilgen, D. R., & Feldman, J. M. (1983). Performance appraisal: A process approach. In B. M.
Staw & L. L. Cummings (Eds.), *Research in organizational behavior* (Vol. 5, pp. 141-
197. Greenwich, CT: JAI Press.

Josefa, R., & Miguel, M. (2007). The social psychological study of physical disability. *Revista
de Psicologia Social*, *22*(2), 177-198. Retrieved from Academic Search Premier database.

Jussim, L., Coleman, L., & Lerch, L. (1987). The nature of stereotypes: A comparison and
integration of three theories. *Journal of Personality and Social Psychology, 52*, 536-546.

Katz, D. & Kahn, R. (1978). *The Social Psychology of Organizations (2nd ed.).* New York:
Wiley.

Kirin, S., & Winkler, J. (1992). The Army Military Occupational Specialty Database. RAND
Note N3527-A. Retrieved March 12, 2010 from http://www.rand.org/pubs/notes/N3527/.

Knapp, D., McCloy, R., & Heffner, T. (2004). *Validation of measures designed to maximize
21st-century Army NCO performance (ARI Technical Report 1145).* Alexandria, VA:
U.S. Army Institute for the Behavioral and Social Sciences.

Kozlowski, S. W., Kirsch, M. P., & Chao, G. T. (1986). Job knowledge, ratee familiarity,
conceptual similarity and halo error: An exploration. *Journal of Applied Psychology*,
71(1), 45-49. doi: 10.1037/0021-9010.71.1.45.

Kraiger, K., & Ford, J. K. (1985). A meta-analysis of ratee race effects in performance ratings. *Journal of Applied Psychology, 70*, 56-65.

LaHuis, D., & Avis, J. (2007). Using multilevel random coefficient modeling to investigate rater effects in performance rating. *Organizational Research Methods, 10*(1), 97-107.

Lord, R. G., Binning, J. F., Rush, M. C., & Thomas, J. C. (1978). The effect of performance cues and leader behavior on questionnaire ratings of leadership behavior. *Organizational Behavior and Human Performance, 21*, 27-29.

Lord, R. G., Foti, R. J., & Phillips,, J. S. (1982). A theory of leadership categorization. In *Leadership: Beyond establishment views.* (pp. 104-121). Carbondale, Illinois: Southern Illinois University Press.

Lord, R. G., & Maher, K. J. (1991). *Leadership and information processing.* Cambridge, MA: Unwin Hyman, Inc.

Luke, D. A. (2004). *Multilevel modeling, Quantitative applications in the social sciences series No. 07-143.* Thousand Oaks, CA: Sage Publishing.

Marshall, S. L. A. (1978). *Men against fire: The problem of battle command in the future war.* Glouster, MA: Peter Smith.

McCauley, C., Stitt, C., & Segal M. (1980). Stereotyping: From prejudice to prediction. *Psychological Bulletin, 87*, 195-208.

McGarty, C. (2002). Stereotype formation as category formation. In C. McGarty, V. Yzerbyt and R. Spears (Eds.) *Stereotypes as explanations: The formation of meaningful beliefs about social groups* (pp. 16-37). New York, NY: Cambridge University Press.

Pedhazer, E. J. (1997). Multiple regression in behavioral research: Explanation and prediction, 3rd Edition. South Melbourne, Victoria: Thomas Learning.

Pleban, R., Tucker, J., Centric, J., Dlubac, M., & Wampler, R. (2006). *Assessment of the FY 05 Basic Officer Leader Course (BOLC) Phase II: Instructor certification program (ICP) and single-site initial implementation(Study Report 2006-09).* Alexandria, VA: U.S. Army Institute for the Behavioral and Social Sciences.

Pulakos, E. D., White, L. A., Oppler, S. H., & Borman, W. C. (1989). Examination of race and sex effects on performance ratings. *Journal of Applied Psychology, 74*(5), 770-780. doi: 10.1037/0021-9010.74.5.770.

Raudenbush, S. W., & Bryk, A. S. (2002). Hierarchical linear models: Applications and data analysis methods. Thousand Oaks, CA: Sage Publications.

Rush, M. C., Thomas, J.C., & Lord, R. G. (1977). Implicit leadership theory: A potential threat to the internal validity of the leader behavior questionnaires. *Organizational Behavior and Human Performance*, *20*, 93-100.

Schmitt, N. & Chan, D. (2006). Situational judgment test: Method or construct? In J. Weekly & R. Ployhart (Eds.) *Situational Judgment Tests: Theory, Measurement and Application* (pp. 135-155). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Schweder, R. A., & D'Andrade, R. G. (1980). The systematic distortion hypothesis. *New directions for methodology of social and behavior science, 4*, 37-58.

Schlee, R., Curren, M., Harich, K., & Kiesler, T. (2007). Perception bias among undergraduate business students by major. *Journal of Education for Business, 82*(3), 169-177.

Schneider, D. (2004). *The psychology of stereotyping*. New York: Guilford Press.

Scullen, S. E., & Mount, M. K. (2000). Understanding the latent structure of job performance ratings. *Journal of Applied Psychology, 85*(6), 956-970. doi: 10.1037//0021-9010.85.6.956.

Sears, G. J., & Rowe, P. M. (2003). A personality-based similar-to-me effect in the employment interview: Conscientiousness, affect-versus competence-mediated interpretations, and the role of job relevance. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 35*(1), 13-24. doi: 10.1037/h0087182.

Steinberg, A. & Leaman, J. (1990a). *The Army leader requirements task analysis: Commissioned officer results (ARI Technical Report 898).* Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Steinberg, A. & Leaman, J. (1990b). *The Army leader requirements task analysis: Non-commissioned officer results (ARI Technical Report 908).* Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Spears, R. (2002). Four degrees of stereotype formation: differentiation by any means necessary. In C. McGarty, V. Yzerbyt and R. Spears (Eds.) *Stereotypes as explanations: The formation of meaningful beliefs about social groups* (pp. 127-156). New York, NY: Cambridge University Press.

Tajfel H., & Turner J.C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.) *The social psychology of intergroup relations*, (pp. 33-47). Monterey, CA : Brooks/Cole.

Tsujimoto, R. N. (1978). Memory bias toward normative and novel trait prototypes. *Journal of Personality and Social Psychology*, *36*(12), 1391-1401. doi: 10.1037/0022-3514.36.12.1391.

Turner, J. (1982). Towards a cognitive redefinition of the social group. In H. Tajfel (Ed.), *Social identity and intergroup relations* (pp. 15-40). Cambridge, England: Cambridge University Press.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124-1131.

Vance, R. J., Kuhnert, K. W., & Farr, J. L. (1978). Interview judgments: Using external criteria to compare behavioral and graphic scale ratings. *Organizational Behavior & Human Performance*, *22*(2), 279-294.

Viswesvaran, C., Ones, D., & Schmidt, F. (2005). Is there a general factor in ratings of job performance? A meta-analytic framework for disentangling substantive and error influences. *Journal of Applied Psychology, 90*(1), 108-131.

Von Schnell, A. (2004). *Battle Leadership*. Arlington, VA: The Marine Corps Association, reprinted.

Wherry, R., & Bartlett, C. (1982). The control of bias in ratings: A theory of rating. *Personnel Psychology*, 35, 521- 530.

Wood, W. J. (1984). *Leaders and battles: The art of military leadership*. Novato, CA: Presidio Press.

Yzerbyt, V.Y. Rogier, A. & Fiske, S. (1998). Group entitativity and social attribution: on translating situational constraints into stereotypes. *Personality and Social Psychology Bulletin, 24*, 1090-1104.

Platoon Leader Situations – Version A

Name: _____ **Last 4 of Social:**_____

Instructions: You will read a series of combat situations. Choose the **best** *and* **worst** response to the given situation. In each situation you are the platoon leader confronted with a problem. You should consider each of the situations as independent from one another.
Each situation is a matter of life and death; that is, you must respond within seconds or friendly Soldiers will likely die. You DO NOT have time to take multiple actions; you can only choose one of the available options. Please select the action you would take immediately, knowing that lives could depend on your decision.

For each question, you will provide 2 responses. Fill in the bubble completely in the "best" column to indicate the best response to the problem and fill in the bubble completely in the "worst" column to indicate the worst response to the problem.

1. While on a mission to clear several buildings your lead squad enters a house and walks into a trap. The enemy has opened fire inside the house and you are forced to leave the building. You try to call for a Bradley Fighting Vehicle to provide support, but radio communications have failed. What do you do now?

best	worst		
O	O	**a.**	Withdraw from the area until radio communications can be reestablished.
O	O	**b.**	Immediately ask your SLs how much ammo they have left to determine resources you have available.
O	O	**c.**	Look for a different way into the house that would take the enemy by surprise.
O	O	**d.**	Send a runner to link-up with an adjacent unit for support.
O	O	**e.**	Task a portion of your element to suppress the house while you lead the assault element to accomplish your mission.

    "e" is the best answer, **"a"** is the worst answer

2. Your men have been fighting on foot for the past 10 days with no more than 2 hours of sleep per night. During a brief period of rest PFC Smith becomes delirious and begins asking where his dog from home is. Several of the guys assist in calming him down. You then receive orders to move out immediately. What do you say to your men who have just witnessed this situation?

best	worst		
O	O	**a.**	"We have orders to move out, follow me."
O	O	**b.**	"I know this is rough, but we've got a job to do. Let's get it done."
O	O	**c.**	"I know you're tired, but I'm counting on you. I know you'll do your best as always. We can pull through if we do this together."
O	O	**d.**	"SGT Jones, have somebody give PFC Smith a hand. We've got to move."
O	O	**e.**	"We must pull it together men. We can rest when we get to a more secure location. Right now I need you to give me 100%."

    "d" is the best answer, **"a"** is the worst answer

3. Your mission is to secure a three-story building and provide overwatch on a key intersection in order to provide cover for follow-on troops. Time is of the essence because the other unit should be moving through the intersection in approximately 10 minutes. The battalion intelligence officer just reported possible enemy activity in the building across the street. How do you respond?

best	worst		
O	O	**a.**	Radio Higher and request another unit be sent to secure the building across the street.
O	O	**b.**	Prepare to clear the building across the street.
O	O	**c.**	Secure the target building first in order to set up the overwatch team and then send an element to clear the second building.
O	O	**d.**	Organize your unit into two sections in order to execute a simultaneous assault on both buildings.
O	O	**e.**	Position an element to suppress the building across the street with small arms if necessary, and then secure the target building. Then tell your men to overwatch both the intersection and the second building.

"**e**" is the best answer, "**d**" is the worst answer

4. As you are moving to link up with another platoon you pass a church. A small group of women and children come running out toward you. You are aware that many civilians have deserted the area and it seems odd that they are here in the open. What do you do?

best	worst		
O	O	**a.**	Find available cover and concealment and prepare to defend yourselves.
O	O	**b.**	Remind your Soldiers of the Rules of Engagement.
O	O	**c.**	Order the civilians to "STOP" and put their hands in the air.
O	O	**d.**	Fire a warning shot in the air to get the group's attention.
O	O	**e.**	Tell Soldiers to aim their weapons at the group, but not to fire unless the group proves to be hostile.

"**a, c**" is the best answer, "**d**" is the worst answer

5. While engaged in fighting with insurgents in a small town you hear machine gun fire increasing several blocks away. You are currently positioned in a one-story concrete building in the middle of the block. You are one of the 3 platoons in the immediate vicinity. What action do you take?

best	worst		
O	O	**a.**	Radio Higher HQs to provide a SITREP.
O	O	**b.**	Check the ammo and equipment status of your men.
O	O	**c.**	Contact each of the other platoons and let them know what you're hearing; ask if they have any further information.
O	O	**d.**	Continue to pull security and await further instructions.
O	O	**e.**	Do a map recon and tentatively plan a safe and efficient route that could move your unit to where the action is.

"**a**" is the best answer, "**c**" is the worst answer

6. You are the 1ˢᵗ platoon leader and are moving toward your link-up point when you look down an alley and see 2nd platoon moving in the opposite direction from the target area. You received no radio communications about any changes to the original plan. What action should you take?

best	worst		
O	O	**a.**	Radio your fellow platoon leaders in the vicinity to find out what's going on.
O	O	**b.**	Radio Higher HQs and request an update on the link-up point.
O	O	**c.**	Set up a security halt and send two men down the alley to find out what is going on.
O	O	**d.**	Drive on with your original mission to the link-up point.
O	O	**e.**	Change your unit's direction of movement in order to intercept the adjacent platoon and find out what's going on face-to-face.

  "b" is the best answer, **"c, e"** is the worst answer

7. While moving toward an intersection that you are to secure, your unit receives small arms fire from the second story window of a 2-story building you are approaching. Movement is also detected on the lower level. It was thought that the buildings were deserted, but Higher now orders you to destroy enemy insurgents in any of the 6 buildings along your way to the intersection. What instructions do you provide to your SLs?

best	worst		
O	O	**a.**	Remind them of the Rules of Engagement.
O	O	**b.**	Stop and secure the area.
O	O	**c.**	Talk to the locals as we pass and ask for information about suspicious activity.
O	O	**d.**	Assault the building quickly before the enemy disperses.
O	O	**e.**	Keep personnel together and keep others informed of where you are and what you encounter.

  "d" is the best answer, **"a"** is the worst answer

8. You are on patrol in BFVs. You are in the lead BFV, while your PSG is in another BFV, 600 meters behind you. Midway through the patrol, his vehicle is attacked by RPG and small arms fire. He reports his situation to you. What is your response?

best	worst		
O	O	**a.**	Reply, "Roger, continue to develop the situation."
O	O	**b.**	Go back and assist to fight off the attack.
O	O	**c.**	Call for reinforcements.
O	O	**d.**	Find some cover and radio your commander.
O	O	**e.**	Search and find the insurgents.

  "b" is the best answer, **"d"** is the worst answer

9. You just cleared a road leading into a city that may be filled with enemy insurgents. You are approaching a key area where concealment is difficult. You are using smoke to mask your movements, but have inhibited your ability to monitor enemy actions and responses. You receive enemy fire. What would you do?

best	worst	
O	O	**a.** Radio your company for any new information about enemy activity in the town.
O	O	**b.** Direct an overwatch/sniper team into a position in a nearby building to see over/past your smokescreen to engage any observed enemy.
O	O	**c.** Use aerial command and control elements to scout out enemy activities.
O	O	**d.** Wait until dark and recon the site.
O	O	**e.** Request armored vehicles.

    "b" is the best answer, **"d"** is the worst answer

10. Your three vehicle convoy has been conducting a presence patrol on the outskirts of your unit sector. Approximately 200 meters to your immediate front, you hear and see what seems to be a hasty ambush being executed on coalition flatbed and cargo trucks. What actions do you take?

best	worst	
O	O	**a.** Radio in a quick SALUTE report to higher headquarters and monitor the situation from a distance. You might cause more confusion if you rush to the convoy's aid.
O	O	**b.** Issue a quick FRAGO to your patrol on how you might deploy in support of the operation if needed.
O	O	**c.** Place your vehicles in a flank position in order to coordinate indirect fire on the insurgents.
O	O	**d.** Immediately pull 360 degree security. It's possible that the commotion up ahead is a distraction or baited-ambush. The real ambush may be designed for you when you move in to support.
O	O	**e.** Immediately deploy to support the unit under attack while reporting your actions to higher headquarters enroute.

    "e" is the best answer, **"d"** is the worst answer

Platoon Leader Situations – Version B

1. While getting ready to enter a two-story house that you know has wounded enemy inside you note that there is a front door, a front window with bars, and a side window. Two squads are running low on ammo. Your unit has just received fire from inside the building. What action do you take?

best	worst	
O	O	**f.** Send an element to recon additional information about the house.
O	O	**g.** Assemble PSG and SLs to assess the situation and discuss options.
O	O	**h.** Instruct your SLs to position themselves at the possible exits and wait for the enemy to move.
O	O	**i.** Take a quick assessment of the platoon equipment to see if you have anything capable of making an explosive breach.
O	O	**j.** Isolate the house and have your interpreter order the inhabitants to lay down their weapons or you will be forced to demolish the house.

    "e" is the best answer, **"b"** is the worst answer

2. The platoon's mission is to clear and secure four buildings and await further orders. You have secured your objective and then you hear that another unit down the street has stumbled into a hostile situation and has sustained several casualties. What do you do?

best	worst	
O	O	**a.** Radio Higher HQs for permission to leave your building and provide support to the other unit.
O	O	**b.** Send half of your unit down the street and leave half at your objective.
O	O	**c.** Radio the other unit and tell them you're on the way.
O	O	**d.** Maintain your unit in a security posture. If you're needed down the street, someone will inform you.
O	O	**e.** Start task organizing your unit in order to send an element to assist down the street, if needed.

      "e" is the best answer, **"b"** is the worst answer

3. After several hours of defending your position within a two-story building from snipers and rebel insurgents, a lull in the fighting occurs. Radio communications indicate that a small group of five or six insurgents are in the vicinity (4-5 blocks away) and are moving in your direction. What do you do?

best	worst	
O	O	**a.** Radio Higher HQs for more information and guidance.
O	O	**b.** Inform your SLs of the possible new threat in order to keep them aware.
O	O	**c.** Check the ammo/water/equipment status of your unit.
O	O	**d.** Double check that your crew served weapons are positioned in the best locations to cover the ingress routes to your location.
O	O	**e.** Position men in observation posts outside of the building in order to provide early warning.

      "b, d" is the best answer, **"e"** is the worst answer

4. Your unit's task is to breach and secure a foothold in Building #1. Your support element, tasked with suppressing the building, throws smoke in order to obscure the assault team's entry. As the assault team leader enters through a window he encounters a booby-trap and is KIA. Another member of the assault team appears disoriented from the blast, stalling your breach into the building. What do you do?

best	worst	
O	O	**a.** Call for a medic, throw more smoke, and pull the casualties to a safe location away from the building.
O	O	**b.** Order one man to tend to the disoriented man and then lead the rest of the assault element into the breach.
O	O	**c.** Look for an alternate entrance into the building.
O	O	**d.** Bypass the casualties and send another assault team into the breach.
O	O	**e.** Report the casualties to Higher HQs and request another unit to help support your breach mission.

      "b, d" is the best answer, **"e"** is the worst answer

5. During an ambush, your platoon has been separated from the company. You start to receive small arms fire and move to a damaged concrete building for cover. Your M249 squad automatic weapon (SAW) gunner begins to lay down suppressive fire but this only causes the enemy fire toward your location to intensify. You believe that the rest of your company is moving to the east, but radio communications are unreliable. What action do you take?

best	worst	
O	O	**a.** Order your SAW gunner to shoot only if he has an exact location on the enemy.
O	O	**b.** Attempt to establish radio communications to find out where the rest of your unit is located.
O	O	**c.** Send two men to determine if they can locate the rest of your company.
O	O	**d.** Move the entire platoon to the east, toward where you believe the rest of the company is located before the enemy pins you down.
O	O	**e.** Check your security perimeter and remain where you are. The company is probably looking for you and attempting to regain contact.

"**d**" is the best answer, "**c**" is the worst answer

6. While on patrol at 0200 you pass a set of government buildings for the third time. A call comes in from Higher telling you to report back to base right away. One of your subordinates says, "Sir, there is a delivery van that wasn't there before." You haven't had any incidents in the last week, and the incident the week before was only a small group of rioters who were unhappy about the new curfew. What do you do?

best	worst	
O	O	**a.** Comply with orders and head back.
O	O	**b.** Radio Higher for permission to search or destroy the vehicle.
O	O	**c.** Stop the unit and send an element to assess the vehicle.
O	O	**d.** Note the location of the vehicle and report it to the S-2; ask if vehicles were used in neighboring villages to attack government buildings.
O	O	**e.** Provide SITREP to Higher and request instructions.

"**e, a**" is the best answer, "**c**" is the worst answer

7. When returning to your compound after a routine patrol the civilian traffic in front of you is backed up. Your unit is traveling in reinforced HMMWVs. You notice several groups of children along the side of the road who are waving to you. The lead vehicle begins to move when an explosion occurs in front of it. The children and civilians along the road are screaming. You receive small arms fire and realize that the enemy is firing from somewhere behind where the children are grouping together. How do you respond?

best	worst	
O	O	**a.** Order your men to break contact.
O	O	**b.** Move your unit out of the kill zone.
O	O	**c.** Find out if your men have sustained any injuries
O	O	**d.** Request reinforcements.
O	O	**e.** Dismount a squad from its current location and have the Soldiers move toward the firing.

"**b**" is the best answer, "**d**" is the worst answer

8. You are patrolling on foot with several local police in training attached to your unit. The buildings in the area are mostly 3-story and made of concrete. As you move past an alleyway fire breaks out from down the alley and overhead. Insurgents pop up on rooftops as your men scramble to return fire. In the meantime the local police huddle together near the wall of a concrete building. What action do you take?

best	worst		
O	O	**a.**	Run to the police and tell them to spread out.
O	O	**b.**	Yell to your men to instruct the police what to do.
O	O	**c.**	Focus on returning fire and engaging the insurgents.
O	O	**d.**	Question the police trainees to determine if they knew this was an ambush.
O	O	**e.**	While seeking cover, physically grab the police and move them to cover.

"**c**" is the best answer, "**d**" is the worst answer

9. Your platoon's mission was to clear and secure a building on the outskirts of town. You have successfully completed your mission, your men are resting, and you are monitoring the radio. You hear gunfire and another platoon leader reports that his platoon is being attacked. How should you respond?

best	worst		
O	O	**a.**	Continue to monitor the radio for further information.
O	O	**b.**	Alert your platoon and go to 100% security.
O	O	**c.**	Begin preparation for your platoon to assist the other platoon.
O	O	**d.**	Plan to leave a squad to secure your building in the event you are directed to assist the other platoon.
O	O	**e.**	Conduct a terrain analysis of routes to reach the other platoon.

"**b**" is the best answer, "**a**" is the worst answer

10. Your platoon is advancing into possible hostile territory. It is 0100. You hear noises and people start running away from your location. What do you do?

best	worst		
O	O	**a.**	Move quickly and attempt to halt fleeing people.
O	O	**b.**	Advance at a slow and measured pace until you are certain of what is ahead.
O	O	**c.**	Call helicopters in to scan the area using thermal sights.
O	O	**d.**	Fire three warning shots.
O	O	**e.**	Call your adjacent platoon to see if they can block people from running away.

"**b**" is the best answer, "**d**" is the worst answer

APPENDIX 2

End of Course Leadership Assessment Report

[Means and score distributions as reported in Pleban, Tucker, Centric, Dlubac & Wampler, 2006. Note: For the Pleban et al (2006) report, international students were removed from the data set prior to analysis. In this study, international students were retained.]

Based on your observations during the course, rate this lieutenant on the following Army Values, Warrior Ethos, and Leader Attributes / Skills / Actions.

Use the rating scale provided below to rate this lieutenant on each dimension.

1 = Needs much improvement – rarely or never behaves this way
2 = Needs some improvement – sometimes behaves this way
3 = Satisfactory – usually behaves this way
4 = Excellent – always or almost always behaves this way
NA = Training situations did not allow lieutenant to display this quality often enough to accurately rate

PART I - CHARACTER: Combination of values, attributes, and skills affecting leader actions

ARMY VALUES

	NA	1	2	3	4	
Lieutenants (*n* = 95 – 169)	**Frequencies (Percentages)**					
1. *Loyalty***:** Shows faith and allegiance to the Army; shows commitment to the unit and all Soldiers.						
2. *Duty***:** Fulfills all obligations; takes initiative and carries out mission requirements in the absence of directions from others based on a sense of what is morally right.						
3. *Respect***:** Treats all Soldiers with dignity and regard; is discreet and tactful when correcting or questioning others.						
4. *Selfless service***:** Puts the welfare of other Soldiers first; gives credit for success to others; sustains team morale.						
5. *Honor***:** Lives up to all the Army values; doesn't lie, cheat, steal, or tolerate those actions by others.						

	NA	1	2	3	4	
6. *Integrity:* Acts honestly and does what is right legally and morally, especially in challenging and stressful conditions.						
7. *Personal courage:* Overcomes fear of bodily harm to successfully complete tasks or mission; takes responsibility for decisions and actions.						

WARRIOR ETHOS ATTRIBUTES						
	NA	1	2	3	4	
1. *Perseverance:* Works through adversity. Does not give up.						
2. *Sets Priorities:* Accomplishes tasks and mission according to appropriate priorities.						
3. *Makes Tradeoffs:* Makes correct tradeoffs between personal sacrifice and the appropriate application of tactics, techniques, and procedures to accomplish tasks or mission.						
4. *Adaptability:* Reacts smoothly to unexpected changes in tasks or mission; finds ways to overcome obstacles and/or improve team effectiveness.						
5. *Acts Responsibly toward Other Soldiers:* Continues to perform tasks or mission despite being weakened or incapacitated (e.g., wounded by enemy, accident, illness).						
6. *Relies Appropriately on Other Soldiers:* Works as a team member to accomplish tasks or mission and ensures the ability of the team to fight again**.**						
7. *Motivated by a Higher Calling:* Demonstrates clear understanding of the importance of achieving proficiency in Warrior tasks and collective missions**.**						

LEADER ATTRIBUTES: Fundamental qualities and characteristics						
	NA	1	2	3	4	
1. *Mental:* Demonstrates desire, will, initiative and discipline.						
2. *Physical:* Displays appropriate level of physical fitness and military bearing.						

3. *Emotional*: Displays self-control; calm under pressure.	6 (4)					

LEADER SKILLS: Skill development is part of self-development; prerequisite to action

	NA	1	2	3	4	
1. *Conceptual*: Demonstrates sound judgment, critical/creative thinking, moral reasoning.						
2. *Interpersonal*: Shows skill with people; coaching, teaching, counseling, motivating and empowering.						
3. *Technical*: Demonstrates the necessary expertise to accomplish all tasks and functions.						
4. *Tactical*: Demonstrates proficiency in required professional knowledge, judgment, and warfighting.						

LEADER ACTIONS: Major activities leaders perform; influencing, operating, and improving

   INFLUENCING: Method of reaching goals while operating/improving

	NA	1	2	3	4	
1. *Communicating*: Displays good oral, written, and listening skills for individual/groups.						
2. *Decision-making*: Employs sound judgment, logical reasoning, and uses resources wisely.						
3. *Motivating*: Inspires, motivates, and guides others toward mission accomplishment.						

   OPERATING: Short-term mission accomplishment

	NA	1	2	3	4	
1. *Planning*: Develops detailed, executable plans that are feasible, acceptable, and suitable.						
2. *Executing*: Shows tactical proficiency, meets mission standards, and takes care of people / resources.						
3. *Assessing*: Uses after-action and evaluation tools to facilitate consistent improvement.						

   IMPROVING: Long-term improvement in the Army, its people, and organizations

	NA	1	2	3	4	Mean (*SD*)
1. *Developing*: Invests adequate time and effort to develop individual subordinates as leaders.						
2. *Building*: Spends time and resources improving teams, groups, and units; fosters ethical climate.						
3. *Learning*: Seeks self-improvement and organizational growth; envisioning, adapting, and leading change.						
Part II - OVERALL NET ASSESSMENT Taking into consideration all of the preceding values and attributes, circle the number that best reflects your overall rating:						
		1	2	3	4	

Notes. Numbers may not equal 100% due to missing data. International students were not included in the analyses.

APPENDIX 3

End of Course Adaptability Rating Scale

The following pages provide descriptions of 4 dimensions of small unit leader adaptability.
- Mental
- Interpersonal
- Lead an Adaptive Team
- Physical

1. First, read the description of each dimension and then the examples of the **best** or **most effective behaviors** for each dimension.
2. **Use the examples** as a guide for making your ratings of the lieutenant's skill level on each dimension.
3. Then, **rate this lieutenant** on each aspect of each dimension using the rating scale below.

1 = Needs much improvement – rarely or never behaves this way
2 = Needs some improvement – sometimes behaves this way
3 = Satisfactory – usually behaves this way
4 = Excellent – always or almost always behaves this way
NA = Training situations (e. g., inadequate time) did not allow lieutenant to display this quality often enough to rate.

Mental Adaptability – Adjusting one's thinking in new situations to overcome obstacles or improve effectiveness. This involves handling emergency or crisis situations, handling stress, learning new things, and creative problem solving.

1a. Demonstrating Mental Adaptability – Handling Emergencies or Crisis Situations

- Reacts with appropriate urgency in threatening, dangerous or emergency situations.
- Makes quick decisions based on clear and focused thinking.
- Maintains emotional control and objectivity during emergencies while maintaining focus on the situation at hand.
- Takes appropriate initiative in emergencies and/or in dangerous situations as appropriate.

			□	NA
□	Needs much improvement			
□	Needs some improvement			
□	Satisfactory			
□	Excellent			

1b. Demonstrating Mental Adaptability – Handling Work Stress

- Remains composed and cool when faced with difficult circumstances or a highly demanding workload/schedule.
- Does not overreact to unexpected news or situations.
- Demonstrates resilience and high levels of professionalism in stressful circumstances.
- Acts as a calming and settling influence that others look to for guidance.

			□	NA
□	Needs much improvement			
□	Needs some improvement			
□	Satisfactory			
□	Excellent			

1c. Demonstrating Mental Adaptability – <u>Solving Problems Creatively</u>

• Employs unique analyses, and generates innovative ideas in complex areas.
• Thinks problems through from different perspectives to determine fresh, new approaches.
• Integrates seemingly unrelated information to develop highly creative solutions.
• Considers wide-ranging possibilities others may miss; thinks "outside the box" to see if there is a more effective approach.

	☐	**Needs much improvement**	☐	**NA**
	☐	**Needs some improvement**		
	☐	**Satisfactory**		
	☐	**Excellent**		

1d. Demonstrating Mental Adaptability – <u>Dealing Effectively with Unpredictable or Changing Work Situations</u>

• Takes effective action when necessary without having to know the total picture or have all the facts at hand.
• Readily and easily changes plans in response to unexpected events and circumstances.
• Effectively adjusts plans, goals, actions, or priorities to deal with changing situations, and does whatever is necessary to successfully complete the job/mission.
• Does not need things to be black or white, refuses to be paralyzed by uncertainty.

	☐	**Needs much improvement**	☐	**NA**
	☐	**Needs some improvement**		
	☐	**Satisfactory**		
	☐	**Excellent**		

2. <u>Interpersonal Adaptability</u> – Adjusting what one says and does to make interactions with other people run more smoothly and effectively. This includes trying to understand the needs and motives of other people – especially people from other cultures or backgrounds.

<u>Demonstrating Interpersonal Adaptability</u>

• Demonstrates flexible, open-minded, and cooperative behaviors when dealing with others.
• Listens to and considers others' viewpoints and opinions, and alters one's opinion when appropriate.
• Open and accepting of negative or developmental feedback regarding work.
• Works well and develops effective relationships with diverse individuals.
• Tailors own behavior to persuade, influence, or work effectively with others.

	☐	**Needs much improvement**	☐	**NA**
	☐	**Needs some improvement**		
	☐	**Satisfactory**		
	☐	**Excellent**		

3. Leading an Adaptable Unit – Ability while occupying a leadership position to help develop adaptability in the unit by encouraging and rewarding adaptive behavior and ensuring everyone works together in a coordinated fashion.

Demonstrating Ability to Develop an Adaptable Unit

- Models adaptive behavior for unit members by learning from experience and seeking self-improvement in weak areas.
- Provides accurate, timely, motivational and constructive feedback to subordinates.
- Helps unit members learn from mistakes in order to be more adaptable in the future.
- Involves unit members in decisions and keeps them informed of consequences of their actions.
- Encourages shared understandings of situations among unit members through appropriate communications to facilitate coordinated actions.

☐	**Needs much improvement**
☐	**Needs some improvement**
☐	**Satisfactory**
☐	**Excellent**

☐	**NA**

5. Physical Adaptability – Effectively adjusts to varied and challenging physical conditions and climates.

Demonstrating Physical Adaptability

- Adjusts to tough environmental states such as extreme heat, humidity, cold, etc.
- Frequently pushes self physically to complete strenuous or demanding tasks.
- Adjusts weight/muscular strength or improves proficiency in performing physical tasks needed to be successful for job/training mission.

☐	**Needs much improvement**
☐	**Needs some improvement**
☐	**Satisfactory**
☐	**Excellent**

0	**NA**

Notes. Numbers may not equal 100% due to missing data. International students were not included in the analyses.

Table A4

Rating Means, Standard Deviations and Sample Sizes by Trainee Role within Rater

	Trainee Job	Diagnostic for Combat Arms						Diagnostic for Non-Combat Arms					
		Physical Scale			*Perseverance*			*Decision Making Scale*			*Planning*		
		CA	NCA	Combined	CA	NCA	Combined	CA	NCA	Combined	CA	NCA	Combined
Rater 1													
M	0.56	3.50	3.00	3.27	4.00	3.50	3.78	3.60	3.63	3.61	3.80	3.50	3.67
SD	0.53	.50	.41	.51	0.00	0.58	0.44	.42	.48	.42	0.45	0.58	0.50
n	9	5	4	9	5	4	9	5	4	9	5	4	9
Rater 2													
M	0.44	3.25	3.10	3.17	3.75	3.40	3.56	3.13	3.20	3.17	3.00	3.20	3.11
SD	0.53	.65	.82	.71	0.50	0.55	0.53	.63	.45	.50	0.82	0.45	0.60
n	9	4	5	9	4	5	9	4	5	9	4	5	9
Rater 3													
M	0.53	3.22	2.88	3.06	3.89	4.00	3.94	3.11	3.19	3.15	3.11	3.25	3.18
SD	0.51	.44	.35	.43	0.33	0.00	0.24	.22	.37	.29	0.33	0.46	0.39
n	17	9	8	17	9	8	17	9	8	17	9	8	17
Rater 4													
M	0.40	2.88	3.10	3.00	3.25	3.00	3.10	3.13	3.08	3.10	3.25	3.00	3.10
SD	0.52	.63	.42	.50	0.50	0.00	0.32	.63	.20	.39	0.50	0.00	0.32
n	10	4	5	9	4	6	10	4	6	10	4	6	10
Rater 5													
M	0.56	3.30	3.25	3.28	4.00	3.25	3.67	2.90	3.13	3.00	3.00	3.25	3.11
SD	0.53	.45	.50	.44	0.00	0.50	0.50	.55	.25	.43	0.71	0.50	0.60
n	9	5	4	9	5	4	9	5	4	9	5	4	9
Rater 6													
M	0.43	3.67	3.25	3.43	3.33	3.25	3.29	3.33	3.00	3.14	3.33	3.00	3.14
SD	0.53	.58	.50	.53	0.58	0.50	0.49	.58	.00	.38	0.58	0.00	0.38
n	7	3	4	7	3	4	7	3	4	7	3	4	7
Rater 7													
M	0.67	3.00	2.83	2.94	3.50	3.33	3.44	3.19	3.03	3.13	3.22	3.06	3.17
SD	0.50	.71	.76	.68	0.55	1.15	0.73	.27	1.00	.55	0.39	1.00	0.60
n	9	6	3	9	6	3	9	6	3	9	6	3	9

Table A4 Continued

Rating Means, Standard Deviation and Sample Sizes by Trainee Role within Rater

		Diagnostic for Combat Arms						Diagnostic for Non-Combat Arms					
		Physical Scale			*Perseverance*			*Decision Making Scale*			*Planning*		
	Trainee Job	CA	NCA	Combined	CA	NCA	Combined	CA	NCA	Combined	CA	NCA	Combined
Rater 8													
M	0.78	3.21	2.0	2.94	3.29	2.50	3.11	2.71	2.50	2.67	2.71	2.50	2.67
SD	0.44	.91	.00	.95	0.49	0.71	0.60	.49	.71	.50	0.49	0.50	0.50
n	9	7	2	9	7	2	9	7	2	9	7	2	9
Rater 9													
M	0.56	3.20	3.00	3.11	3.00	3.00	3.00	2.80	3.00	2.89	2.80	3.00	2.89
SD	0.53	.45	.00	.33	0.00	0.00	0.00	.45	.00	.33	0.45	0.00	0.33
n	9	5	4	9	5	4	9	5	4	9	5	4	9
Rater 10													
M	0.38	3.83	3.60	3.69	4.00	3.80	3.88	3.33	3.80	3.63	3.00	3.80	3.50
SD	0.52	.29	.89	.70	0.00	0.45	0.35	.29	.45	.44	0.00	0.45	0.53
n	8	3	5	8	3	5	8	3	5	8	3	5	8
Rater 11													
M	0.67	3.83	3.67	3.78	3.67	3.33	3.56	3.58	3.67	3.61	3.67	3.67	3.67
SD	0.50	.41	.58	.44	0.82	1.15	0.88	.49	.58	.49	0.52	0.58	0.50
n	9	6	3	9	6	3	9	6	3	9	6	3	9
Rater 12													
M	0.87	3.60	4.0	3.67	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00
SD	0.38	.42	---	.41	0.00	---	0.00	--	--	.00	0.00	---	0.00
n	7	5	1	6	6	1	7	1	1	2	6	1	2
Rater 13													
M	0.33	4.00	3.75	3.83	4.00	3.83	3.89	3.83	3.92	3.89	4.00	4.00	4.00
SD	0.50	.00	.42	.35	0.00	0.41	0.33	.29	.20	.22	0.00	0.00	0.00
n	9	3	6	9	3	6	9	3	6	9	3	6	9
Rater 14													
M	0.75	2.92	2.75	2.88	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00
SD	0.46	.20	.35	.23	0.00	0.00	0.00	.00	.00	.00	0.00	0.00	0.00
n	8	6	2	8	6	2	8	6	2	8	6	2	8

Table A4 Continued

Rating Means, Standard Deviation and Sample Sizes by Trainee Role within Rater

	Trainee Job	Diagnostic for Combat Arms						Diagnostic for Non-Combat Arms					
		Physical			*Perseverance*			*Decision Making*			*Planning*		
		CA	NCA	Combined	CA	NCA	Combined	CA	NCA	Combined	CA	NCA	Combined
Rater 15													
M	0.63	3.50	3.33	3.44	3.60	3.67	3.63	3.30	3.67	3.44	3.40	3.67	3.50
SD	0.52	.50	.58	.50	0.55	0.58	0.52	.27	.58	.42	0.55	0.58	0.53
n	8	5	3	8	5	3	8	5	3	8	5	3	8
Rater 16													
M	0.38	3.17	3.10	3.13	4.00	3.60	3.75	3.00	3.10	3.06	3.00	3.00	3.00
SD	0.52	.29	.23	.23	0.00	0.55	0.46	.00	.22	.18	0.00	0.00	0.00
n	8	3	5	8	3	5	8	3	5	8	3	5	8
Rater 17													
M	0.625	3.50	2.88	3.22	3.400	2.667	3.125	2.90	2.33	2.69	3.200	2.333	2.875
SD	0.518	.50	.25	.51	0.548	0.577	0.641	.42	.58	.53	0.447	0.577	0.641
n	8	5	4	9	5	3	8	5	3	8	5	3	8
Rater 18													
M	0.250	4.0	3.50	3.57	2.500	3.167	3.000	3.00	3.00	3.00	3.000	3.000	3.000
SD	0.463	---	.55	.53	0.707	0.408	0.535	.00	.00	.00	0.000	0.000	0.000
n	8	1	6	7	2	6	8	2	6	8	2	6	8
Rater 19													
M	0.444	3.25	3.80	3.56	3.750	3.400	3.556	3.13	3.10	3.11	3.000	3.000	3.000
SD	0.527	.29	.27	.39	0.500	0.548	0.527	.25	.22	.22	0.000	0.000	0.000
n	9	4	5	9	4	5	9	4	5	9	4	5	9
Total													
M	0.530	3.35	3.22	3.29	3.570	3.405	3.491	3.15	3.23	3.19	3.190	3.230	3.209
SD	0.501	.55	.59	.57	0.543	0.589	0.570	.46	.51	.47	0.520	0.529	0.523
n	169	89	79	168	86	79	165	86	79	165	86	79	165

Table A5

Means, Standard Deviations, Sample Sizes and Frequencies for Demographic Variables

	Age (in years)	Prior Enlisted (1=yes)		Commissioning Source					Component		
				ROTC	Direct Commission	OCS	Warrant Officer	USMA	Active	Reserve	NG
Combat Arms			Combat Arms								
M	23.84	0.3	*n*	67	6	3	2	13	67	18	5
SD	3.053	0.459	*Ratio*	75.28%	6.74%	3.37%	2.25%	14.61%	74.44%	20.00%	5.56%
N	90	91									
Non-Combat Arms			Non-Combat Arms								
M	23.65	0.34	*n*	51	2	6	3	17	60	15	5
SD	3.292	0.476	*Ratio*	63.75%	2.50%	7.50%	3.75%	21.25%	75.00%	18.75%	6.25%
N	80	80									
Total			Total								
M	23.77	0.31	*n*	119	8	9	5	31	129	33	10
SD	3.162	0.465	*Ratio*	68.79%	46.00%	5.20%	2.89%	17.92%	75.00%	19.20%	5.80%
N	169	169									

Table A6

Results of HLM Final Combat Leader Model with All Participants

Fixed Effects	Model 2			
	Coef.	*SE*	*t*	*p*
For Intercept (β_{0j})				
Intercept (γ_{00})	1.58	.427	3.701	.000
Rater Deviation (γ_{01})	.755	.160	4.719	.000
For PFT Slope (β_{1j})				
PFT (γ_{10})	.00734	.00156	4.712	.000
For Trainee Job Slope (β_{2j})				
Trainee Job (γ_{20})	-.540	.533	-1.012	.313
For PFT x Trainee Job Slope (β_{3j})				
PFT x Trainee Job (γ_{30})	.002	.00199	1.161	.248
For SOURCE=OCS Slope (β_{4j})				
SOURCE=OCS (γ_{40})	.3468	.2017	1.720	.089
For SOURCE=ROTC Slope (β_{5j})				
SOURCE=ROTC (γ_{50})	.3352	.1298	2.580	.013
For SOURCE=USMA Slope (β_{6j})				
SOURCE=USMA (γ_{60})	.3026	.1539	1.966	.053
For SOURCE=USMA Slope (β_{6j})				
SOURCE=USMA (γ_{60})				