

United States Military Academy USMA Digital Commons

West Point ETD

2013

Status-Based Stereotyping: The Effects of Shifting Standards on Team Performance

Lissa V. Young

United States Military Academy, lissa.young@usma.edu

Follow this and additional works at: https://digitalcommons.usmalibrary.org/faculty_etd

 Part of the [Social Psychology Commons](#)

Recommended Citation

Young, Lissa V., "Status-Based Stereotyping: The Effects of Shifting Standards on Team Performance" (2013). *West Point ETD*. 4.
https://digitalcommons.usmalibrary.org/faculty_etd/4

This Doctoral Dissertation is brought to you for free and open access by USMA Digital Commons. It has been accepted for inclusion in West Point ETD by an authorized administrator of USMA Digital Commons. For more information, please contact nicholas.olijnyk@usma.edu.

Spring 5-5-2013

Status-Based Stereotyping: The Effects of Shifting Standards on Team Performance

Lissa V. Young

Follow this and additional works at: https://digitalcommons.usmalibrary.org/cadet_faculty_staff_publications_faculty_staff_etd

 Part of the [Social Psychology Commons](#)

Status-Based Stereotyping: The Effects of Shifting Standards on Team
Performance

Lissa V. Young

Professor Monica C. Higgins

Professor Frank Barrett

Professor Katherine Masyn

A Thesis Presented to the Faculty
of the Graduate School of Education of Harvard University
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Education

2013

Copyright 2013
Lissa V. Young
All Rights Reserved

Dedication

To the mistakes I've made along the way.

Acknowledgments

I owe a gigantic debt of gratitude to a long list of people, many of whom will not be mentioned here by name. I trust you will know who you are, and find the ways I am thankful in the text that follows.

First, I want to thank Monica Higgins for reaching out and grabbing me on my way in the door at Harvard University. You took a huge, risky leap of faith for me, and I have devoted much of my energy to not letting you down. Monica, you are tireless, quick witted, kind and generous, and an unbelievably talented teacher and thinker. I am lucky to have been granted access to such an inspiring person and scholar. The work I do, from here forward, will always bear your signature, and for that I am eternally grateful.

Second, I want to thank both Frank Barrett and Katherine Masyn, whose rock solid mentoring and faith made this work possible. To Frank, I am thankful for the way your dedication to excellence always feels like an invitation, not a demand. Thank you for teaching me how to lean into the unknown, and say “Yes To The Mess”. And Katherine, your unwavering commitment to what is possible, and your amazing patience with those who don’t see as well as you, will be the hallmark of your legacy. To you, I owe so very much. Thank you for teaching me that I’m not stupid, even though I’ll never have a Turing mind like yours. Despite my years of struggle with statistical analysis, I am passing through these Harvard gates well prepared to teach others, and I owe that all to you believing in me. Thank you.

Third, I am forever indebted to the humor, friendship, and generosity of my dear writing group - GTFOHG. Lauren, I love the way you can see the

structure in an argument immediately, and then rearrange blocks of consciousness in service of clarity and efficiency. You are by far the better writer in this group, and you are going to be an amazing dean one day. Mariam, I will never cease to be amazed by how you see so much more in the writing than is actually there. You have one of the most insightful and perceptive minds I've known, and I so appreciate how much you've pushed me. My work has improved tremendously, because of the way you think. I look forward to the day you are the President of Pakistan. And Charles, your unflappable, steady, solid-state energy coupled with your gorgeously twisted and brilliant mind will always be present in the way I approach my work. You have taught me so much about testing feasibilities and pushing the boundaries. To all three of you – thank you for making my thinking and writing so much better than it has a right to be.

To Jennie Weiner, the funniest, fiercest woman I know, and the very best long-term office mate anyone could hope for. Thank you for dedicating time you did not have to acting as my chief editor, idea challenger, and for teaching me how to swim with babies. Your friendship and scholarship are both beyond compare. I look forward to your steady presence in my life as a person and as a scholar.

To my long-time professional mentors Joe LeBoeuf, Barney Forsythe and Casey Brower who, despite my giving them plenty of reasons to, never stopped believing in me. You three have taught me everything I know about patience, hardiness and hope. Thank you for every time you have answered my call.

To the HGSE admissions committee of 2007, thank you for extending an offer of acceptance so many years ago. You've changed my life.

To Monica Biernat, Chris Crandall, Terri Vescio, Diane Kobrynowicz and Stan Halpin who planted the seeds early on, and whose generosity and scholarship have been an inspiration ever since.

To the Babes and Blue Crabs, and the “Summer-Summer” email lists. Thank you for keeping my life balanced, filled with laughter and teaching me that bacon and martinis are complimentary food groups.

And lastly, a huge debt of gratitude goes to my family. I could never have even tried this if it wasn't for the fiercest love a person could know from my mother Ellen, my sister Laura, and my best friend in the whole wide world, Yvette. I will never be able to express well enough the gratitude I feel for the unbreakable ligature of love and support you three roped around my soul, and from which I gathered so much courage every single day of this sometimes harrowing but far more often delightful journey. Thank you so much for seeing me in ways I simply cannot.

TABLE OF CONTENTS

LIST OF TABLES	
LIST OF EXHIBITS	
ABSTRACT	
CHAPTER 1: INTRODUCTION	1
CHAPTER 2: LITERATURE REVIEW	8
CHAPTER 3: METHOD	22
CHAPTER 4: RESULTS	34
CHAPTER 5: DISCUSSION	42
APPENDICES	61
REFERENCES	66

LIST OF TABLES

Table	Name	Page
Table 1	Table containing a list and description of Sandhurst Competition events.	23
Table 2	Table containing a list and description of the nine leadership competencies.	24
Table 3	Table representing the number and percentages of participants by status characteristics of gender, race, and cadet class.	34
Table 4	Table representing the team averages for response rate, team size, number of females, number of males, number of lower class cadets, number of upper class cadets, number of non-White cadets and number of White cadets.	35
Table 5	Table showing the shifting standards index (SSI) for each demographic category.	36
Table 6	Parameter estimates, standard errors and approximate p-values for a taxonomy of fitted multi-level structural equation models predicting shifting standards indices (SSI) by diversity index, and cadet gender, race and class. (n=447)	38
Table 7	Parameter estimates, standard errors and approximate p-values for a taxonomy of fitted multi-level structural equation models predicting team performance (as indicated by points earned in the competition) in 10 different events by team level shifting standards indices (SSI) by cadet gender, race and class, and diversity index. (n=447)	41

LIST OF EXHIBITS

Exhibit	Name	Page
Exhibit 1	Exhibit depicting an example of how the data was transposed in order to calculate the shifting standards index	26

Abstract

Teams are a critical aspect of organizational life, and understanding the ways that stereotyping impacts team performance is the first step in optimizing team effectiveness. This research examined the impact on team performance of stereotyping in military teams participating in an international military skills competition. By leveraging the theoretical strength of status characteristics theory, coupled with the analytic power of the shifting standards model, and multi-level structural equation modeling (ML-SEM), this study was able to provide insights into the ways that multiple sources of stereotyping, altogether (i.e. global stereotyping), impacts team performance. In addition, I examined specific sources of stereotyping and their independent impact on team performance within and across teams in a variety of individual events, as well as overall competition performance. I found that global stereotyping did not have an influence on a team's overall performance. However, global stereotyping did have an impact on three individual events. Also, the impact of specific sources of stereotyping did have varying effects on different kinds of team tasks. In one case, race-based stereotyping was associated with enhanced team performance, and in two instances it was associated with decreased team performance. Similarly, in one case, gender-based stereotyping was associated with decreased team performance, but was never associated with enhanced team performance. These findings suggest that stereotyping in a team context is complicated, but does not necessarily lead to a compromise in the team's overall performance.

Chapter 1: Introduction

Introduction

Teams are often touted as a way to increase or enhance organizational performance (Addison & Haig, 2012; Corsaro, Cantù, & Tunisini, 2012; Pentland, 2012), and yet, both research (Herring, 2009; Thatcher & Patel, 2011; van Dijk, van Engen, & van Knippenberg, 2012) and anecdotal evidence suggest that they often fail to do so. While there are a number of reasons why teams may not succeed in producing optimal performance, one particularly salient issue is the underutilization of team members' knowledge and skills (Hackman, 2002). Indeed, it is well-documented that team members often are unable to fully capitalize on the team's collective know-how due to members' inability see beyond each other's specific demographic categories, such as gender, race or social status, to leverage each other's task relevant knowledge, skills and abilities (Foschi, Warriner, & Hart, 1985). Unfortunately, this oversight can leave much of the team's talent untapped, which can potentially hinder the team's performance (Kochan et al., 2003).

However, though this tendency of team members to focus more on colleagues' demographics than potential intellectual contributions abounds, researchers tend to promote the virtues of team diversity while overlooking the ways that these cognitive classifications (i.e., "stereotypes") can potentially undermine the very benefits that diverse teams may offer. One exception comes from status characteristics theory, in which Berger, Cohen, and Zelditch, (1972) offer an account of how biased evaluations can occur in team settings. Their theory explores how group members use status differences to determine

perceived competence, as well as expectations for the performance abilities of fellow team members. These performance expectations, in turn, affect the power and prestige orders of these groups (Berger, Norman, Balkwell, & Smith, 1992; Berger, Ridgeway, Fisek, & Norman, 1998; Foschi, 1992a, 2000), and these critical group dynamics impact the team's performance. Hence, status characteristics theory offers a compelling, though, I argue, a somewhat inadequate lens through which to examine the impact of diversity on teams. This is because despite its contributions, status characteristics theory, like many traditional approaches to studying team diversity tends to limit its focus to single diversity attributes.

In reality, diversity is both multidimensional and complex (i.e., people hold and represent multiple identities simultaneously and others experience these identities in different ways) and suggests a need for a theory and analytic approach than can address this complexity (Garcia-Prieto, Bellard, & Schneider, 2003). And yet, in much of the research on stereotyping and team performance, social scientists tend to examine specific sources of bias independently of the others (Bantel & Jackson, 1989; Chatman, Polzer, Barsade & Neale, 1998; Apesteguia, Azmat & Iriberry, 2012; Newhesier & Dovidio, 2012). Some researchers suggest this trend is a logical extension of the way that prominent empirical work on diversity has been framed (Alise & Teddlie, 2010; Bryman, 2011). A vast majority of research on stereotyping and prejudice in teams is based in social identity and social categorization theories, both of which presuppose readily detectible diversity such as sex, race and age. This has created a strong theoretical rationale for focusing on independent category specific sources of diversity (Brewer & Pierce, 2005; Wells & Aicher, 2013; Williams & O'Reilly, 1998).

Though conceptually, it makes sense that individual diversity attributes influence team outcomes, they do not act in isolation, and because of this, there is a need to empirically assess the multi-dimensionality of diversity (Jackson, Joshi, & Erhardt, 2003). We know that people have multiple identities all operating simultaneously (Garcia-Prieto et al., 2003), and because this is often overlooked in diversity research, I seek to determine the total impact of different sources of stereotyping on team performance. I propose that by focusing on the whole of a variety of sources of interpersonal bias, we can obtain a more realistic view of its impact on team performance.

Beyond dutifully following the empirical trend, another reason that much of the team diversity research has focused on singular aspects of stereotyping is that, until recently, research methodologies have tended to lend themselves to that framework and vice-versa. A large portion of the research conducted on leadership and group dynamics utilizes single level regression modeling or analysis of variance techniques (Edmondson & Mcmanus, 2007; Judd & Kenny, 2010; Yammarino, Dionne, Ukchun, & Dansereau, 2005). In their meta-analysis of the methodologies used in research on leadership and teams, Yammarino, et al. (2005) claim that much of that research has failed to leverage more appropriate techniques such as multi-level modeling, and that this has limited the ability of researchers to engage in theory building. Those same authors also emphasized, “the importance of clearly specifying the level(s) of analysis at which phenomena are expected to exist theoretically, and that it is critical to ensure the data analytic techniques correspond to the asserted level(s) of analysis, so that inference drawing is not misleading or artifactual (p. 879)”.

Yet another criticism that has been lobbied against research on stereotyping and its impact on teams is the inconsistent way that team performance has been operationalized and measured (Austin et al., 1992). Team performance studies conducted in natural settings often utilize indirect measures of team performance such as customer satisfaction surveys or supervisor and team members' self-reported assessments (Landy & Farr, 1980; Watson, Kumar, & Michaelson, 1993; Milliken & Martins, 1996). With this research, I attempt to address this issue by measuring team performance using arguably more objective indices such as time and speed. My motivation is to reduce the amount of error introduced by the aforementioned relatively more indirect or subjective measures of team performance.

Given these various debates in the literature, I suggest that alternative methodological approaches may provide opportunities to address previously underexplored aspects and impacts of stereotyping in team settings. For example, the shifting standards model (Biernat, Manis, & Nelson, 1991; Biernat & Manis, 1994) if used in conjunction with multi-level structural equation modeling (ML-SEM), may offer us a powerful way to detect and measure multiple levels of stereotyping in teams. First, the shifting standards model extends status characteristics theory (Berger et al., 1972; Foschi, 1992b; Wagner & Berger, 2011) by helping to measure the ways interpersonal assessments of competence are made in stereotype laden judgment domains. As stated above, when we come into contact with lower status group members, we activate our stereotypes. When we do this, we implicitly activate expectations for low status group members' competence and performance abilities. These expectations in turn activate judgment standards against which we evaluate others (Biernat, M.,

Vescio, T., Manis, 1998). I propose that if we use the shifting standards model in conjunction with ML-SEM, and more objective measures of team performance, we may gain richer insights into the relationship between the multiple status characteristics based interpersonal biases that occur in team life and that impact team performance.

Purpose of the Study

The purpose of this research is to provide insights into the ways that multiple and simultaneous stereotyping can impact team performance. Motivated by the opportunity to leverage the theoretical strength of status characteristics theory, coupled with the analytic power of the shifting standards model and ML-SEM, as well as measuring performance in a more objective manner, I designed a study that would examine more thoroughly the impact of stereotyping on team performance. By designing the study this way, I am able to examine team-level stereotyping using individual level shifting standards data and further evaluate the impact of such stereotyping on team performance.

To conduct this study, therefore, it was important to select an environment where teams are a formalized construct, and the use of teams is highly routinized. As mentioned before, I desired a context where team performance could be measured using more objective indices, such as time and speed. Given these criteria, I selected a military skills competition, where teams ran against one another in a rigorous three-day event. The selected teams raced against each other, and the clock, in a formal and professionalized setting. Another benefit, directly related to the study's objectives, was that military teams in the United States are structured so that each individual is assigned distinct roles and responsibilities that are clearly delineated by rank, skills and abilities.

This well-marked interpersonal terrain allows for the study of status-based interpersonal biases in ways that less clearly structured teams cannot.

This project extends previous research in two specific ways. First, it leverages the strength of structural equation modeling to extend the use of the shifting standards model from the individual to the team level. This adds to the theoretical power of the shifting standards model, because the model has never been used to examine the impact of team level shifting standards on team performance. Second, this research measured the impact of stereotyping on team performance utilizing standards of performance that were relatively more ordinal and objective than found in the extant research on teams.

Research Questions

In this study I examine the impact of stereotyping on team performance utilizing the shifting standards model and ML-SEM. In doing so, my goal is to answer the following research questions:

1. Is there evidence of team-level shifting standards in teams competing in a skills competition?
2. Is there variability in team-level shifting standards across these teams?
3. Is there a relationship between team-level shifting standards and team performance in the competition?

Roadmap of the Dissertation

In the following chapter, Chapter 2, I conduct a review of the literature that will inform the study. With an emphasis on status characteristics theory, I review the research on stereotyping and prejudice with respect to its impact on teams and team performance. I then review the literature on the shifting standards model. And finally, I review the literature on the methodologies used

in diversity and team performance research. In Chapter 3, I describe in detail the research design, and my methods. In Chapter 4, I report the results of the study, and in Chapter 5, I provide a discussion of the results. I conclude with a discussion of the implications for and limitations of my research.

Chapter 2: Literature Review

Introduction

As the workplace and teams become increasingly diverse, there has been a tension between the promise and reality of diversity in team performance (Mannix & Neale, 2005). The pessimistic view, held by theoretical perspectives such as social categorization (Tajfel, 1978), claim that diversity creates social divisions (“categories”) based on different and quite visible status characteristics such as gender, race and rank, and these categories, in turn, activate differential expectations for team members’ abilities (Foschi, 1992b; Hewstone, Rubin, & Willis, 2002; Newheiser & Dovidio, 2012; Ridgeway & Berger, 1986). Rather than fully utilizing the breadth of perspectives and expertise a diverse team might provide, research has shown that team members tend to rely on stereotypes to interact, thus limiting both the range and legitimacy of others’ contributions (Ancona & Caldwell, 1992; Mannix & Neale, 2005). The more optimistic view, stemming from literature on information processing in teams, holds that diversity will lead to an increase in the perspectives brought to bear on team tasks, which will result in higher levels of creativity and team performance (Hoffman & Maier, 1961; Polzer, Milton, & Swann, 2011). Despite the lack of unanimity, researchers do agree that diversity matters, whether it results in positive, negative or mixed impact on teams and their performance.

Thus, it is arguable that stereotyping can hinder team performance, but a vast majority of the research on stereotyping is as of yet incomplete, because it consistently treats as independent, each of the categories of stereotyping. It stands to reason that researchers interested in racism, for example, would conduct experiments focused on race, holding constant other aspects of the

subjects' diversity. This is true also of research on other forms of identity-based stereotyping, such as sexism and ageism. However, by focusing entirely on a single aspect of identity, researchers tend to ignore the inherently multi-faceted nature of social identity and lose an opportunity to answer fundamental questions of how the myriad social identities "function in relation to one another" (Bodenhausen, 2010).

This is especially problematic for research conducted in complex work environments in which diverse teams engage in multiple types of stereotyping. For example, and germane to this study, the latest report of the U.S. Census Bureau (<http://www.census.gov>, accessed 03 March, 2013) suggests that teams in the U.S. Military are increasingly diverse with respect to gender and race. The latest reports show that the percentage of non-White soldiers has increased to 35% of the total defense force, which is higher than that found in the general population. In addition, the percentage of women has risen substantially to an all time high of 21%. Taken together these changes suggest a need for a more multi-faceted and holistic approach to researching stereotyping in teams. Additionally, while there are a number of factors on which people can be stereotyped, the categories I highlight above (i.e., sex, race and status) seem to be persistent (and problematic) in military organizations, and hence particularly relevant in the present context. In the following, I provide some frameworks to understand why and how these characteristics might influence performance.

These demographic trends set a compelling backdrop for this study and also informed my selection of two specific bodies of research found in the stereotyping literature. The first is status characteristics theory (Berger et al., 1972), which provides the foundation for this study's point of view on

interpersonal stereotype-based judgment standards in teams. The second body of research is that of the shifting standards model (Biernat & Manis, 1994), which provides insight into ways of detecting and measuring the presence of stereotype-based judgments in teams. Each of these bodies of literature will be reviewed as it relates to stereotyping and its impact on team performance.

Status Characteristics Theory

Status characteristics theory (Berger, Norman, Balkwell, & Smith, 1992; Ridgeway & Berger, 1986) addresses the ways that observers ascribe domain relevant competencies, and therefore expectations for performance, to targets based on abilities associated with their demographic categories. Observers then interact with target individuals according to their biased expectations. A status characteristic is any valued attribute that can imply task competence. An example of a status characteristic is having high or low mechanical ability, being female or male, or being White or Black. There are two classes of status characteristics: specific and diffuse. Specific status characteristics (such as mechanical ability) have sharply defined performance expectations in readily identifiable domains. Diffuse status characteristics (such as race and gender and class) tend to have limited and general performance expectations and are viewed as relevant to a large, indeterminate number of domains (Berger et al., 1972). The complication associated with the use of status-based performance expectations is that, more often than not, the status category is not related to the performance domain. For example, the status category “female” is not a relevant predictor in the performance domain “leader,” and yet females are often judged to be less effective leaders, than males (Eagly & Karau, 2002; Foschi, 1992a), and this may have profound implications for diverse teams.

Gender determines expectations for social and professional behaviors of men and women, and these expectations can shape interpersonal dynamics on work teams, and can have negative effects on team performance (Banaji & Hardin, 1996). Often researchers have found that females are perceived less favorably than males with respect to their appropriateness for and their performance as leaders, and this can have deleterious effects on group performance (Eagly & Karau, 2002; Heilman, Block, & Martell, 1995).

Interestingly, researchers suggest that status-based judgments on teams are generated to sustain the power and prestige order that has been established in the group (Foschi, Lai, & Sigerson, 1994; Foschi et al., 1985). Foschi et al (1994) found that male evaluators tended to underestimate the competence of women in relation to men, even when their performances were the same or better. In their study conducted at the University of British Columbia, male subjects making hiring choices of female and male job applicants, consistently hired males at a higher rate than females, regardless of their performance vitae. Foschi et al (1994) assert that these biased evaluations act to sustain the general prestige order of men and women. Other researchers found similar patterns of interpersonal assessments of competence in mixed gender teams of police officers (Gerber, 2011). Gerber found that women were far more likely to be characterized as possessing less effective dispositions for police work than their male counterparts, and this carried over into their ability assessments.

These observations reinforce the idea that once status-based perceptions of competence are established in a team, they will be protected and perpetuated (Ridgeway & Berger, 1986). One strategy for doing this involves assessing targets by different standards. Success of low status group members is assessed

through a stricter standard than success of high status group members (Foschi et al., 1994; Yzerbyt & Demoulin, 2010). For example, assume that female aviators are low status individuals and male aviators are high status individuals in U.S. Army aviation organizations. Because of these gender-based status differences male aviators hold certain expectations for a female aviator's performance, i.e. that she will be less competent than a male aviator. If a female aviator flies very well, according to the judgment biases discussed above, the male aviator will attribute her performance to factors other than her own capabilities. He might think that she flew well, because it was a calm day, or the mission was very simple, or the aircraft was doing all the work. Consequently, he won't change his opinion about her being a less competent aviator than her male peers. Conversely, if a male aviator flies well, other male aviators will believe it is because he's a gifted aviator.

The presence of status-based expectations has implications for how team members leverage each other's capabilities in task oriented work, and this in turn impacts the team's performance (Foschi, 1992b; Hackman, J.R., and Katz, 2010; Roberge & van Dick, 2010). For example, Berger, et al (1992) conducted a study where participants were assigned notional teammates of differing status characteristics (i.e. gender, race, education and occupation). Some participants were assigned high status teammates (male, White, college graduates in a profession) and others were assigned low status teammates (female, non-White, high school graduates and minimum wage worker). Since the participants' teammates were notional only, their ability levels could only be inferred by their status-characteristics. The participants were then challenged with completing and/or solving a variety of different tasks and puzzles. The participant could

solve some of these tasks alone, but many required the help of another person (“teammate”). For each problem or task posed, the researcher asked the participant if s/he wanted to solve the problem alone or with a teammate. In the form of points, the participant could earn rewards for successfully completing a task or solving a problem, or penalties for failing to do so. The object of the exercise was to accumulate the highest number of points. When asked whether or not they wanted the assistance of the notional teammates, researchers found that participants actively avoided engaging in team-related tasks with those teammates who were lower status, but often chose to engage the assistance of teammates who were higher status. Researchers assert that participants avoided lower status teammates, ostensibly to avoid performing poorly. This status-based avoidance behavior potentially has significant implications for team performance because it suggests that the talent of lower status team members might be overlooked in favor of higher status team members.

Similar research in status-based social categorization has delineated the ways that in-group favoritism leads to discrimination of out-group members, and its subsequent influences on team performance (Jehn, Northcraft, & Neale, 1999; Jehn & Bezrukova, 2011; Tajfel, 1978; Tajfel & Turner, 1979, 1986). Jehn, et al (1999) found that social category diversity detracted from team performance because it increased relational conflict between the in-group and out-group members. In her research she found that unmanaged relational conflict stemming from the establishment of in-groups and out-groups interfered with the ability of the teams to remain focused on tasks. These findings are consistent with much of the literature on diversity and team performance (Roberge & van Dick, 2010).

There are many status-based influences on team performance, and there is an exhaustive effort to refine and measure how these influence team performance. What we have learned from this research is that when status-based performance expectations are activated in teams, they can have a negative effect on team performance (Williams & O'Reilly, 1998). However, despite the extensive research detailing how specific status-characteristics impact team performance, there is still a gap with respect to how all of these different status-based biases might act in concert to influence performance. Of late, researchers have been identifying this as a shortcoming in studies examining the effects of diversity on team performance (Garcia-Prieto et al., 2003; Jackson et al., 2003). These researchers have suggested that dimensions of diversity aren't experienced independent of one another, and that one of the key challenges of measuring their influence in teams, is to find ways to measure how they operate simultaneously. Toward that end, I propose the use of the shifting standards model (Biernat & Manis, 1994b; Biernat, Tocci, & Williams, 2011) in extending this research.

The shifting standards model is a category-agnostic method of detecting and measuring the presence of stereotypes, and if used in conjunction with multi-level structural equation modeling, may provide a richer view of the impact of all sources of stereotyping on team performance. In the section that follows, I describe in greater detail the shifting standards model, and its relationship to the status-based judgments described in the literature review above.

The Shifting Standards Model

Biernat & Manis (1994) found that when we come into contact with stereotyped individuals in stereotype relevant domains, we activate our stereotypes. When we do this, we implicitly activate expectations for the stereotyped group members' competence and performance abilities. These expectations are different than we hold for those group members for whom we do not hold a stereotype. These two different expectations, in turn, activate different judgment standards, against which we evaluate the stereotyped and the non-stereotyped group members. These different assessment standards are activated depending upon the target being evaluated, and the task being performed. If we are evaluating two group members, one who is stereotyped and one who is not, in regard to a stereotype relevant characteristic or ability, we tend to evaluate them with standards anchored in our stereotype-based expectations for each. These different standards are determined by our bifurcated expectations of the two group members on the dimension being evaluated.

For example, if I evaluate a man and woman in regard to athleticism, I implicitly use different standards of judgment to determine how athletic each is. These standards are determined by my own stereotypes about the disparity between male and female athleticism. So, if I claim that the female is a good athlete, I am thinking about her abilities in a different way than if I claim that the man is a good athlete. I tend to compare the female to the standard I implicitly hold for female athleticism. When I evaluate the man's athleticism, the standard I use is that of male athleticism. Consequently, I may determine that both the man and the woman are good athletes, but when I do, I may very likely be holding completely different mental representations of what that means for each

of them. It is this disparity in the mental representations of our judgments that the shifting standards model can detect and measure (Monica Biernat et al., 1991; Monica Biernat & Manis, 1994; Monica Biernat & Vescio, 1993).

The shifting standards model stems from the idea that we judge stereotyped individuals in stereotype relevant domains using within category judgment standards. These standards have a scale with a mean and a range that is expected of the target being judged. In stereotype laden judgment domains, those means and ranges will shift in accordance with expectations of the target being judged (Biernat & Manis, 1994; Biernat, Vescio, Manis, 1998). For example, athletics is a domain that typically elicits gender stereotypes. If a man and a woman both run a mile in six minutes, despite the equivalent objective performance, the two runners will be subjectively judged relative to different (male and female) standards, resulting in the man being judged as *average*, and the woman being judged as *above average*, as runners. This shift in the standard against which targets are judged in stereotype relevant domains is called “shifting standards” (Biernat & Manis, 1994; Biernat & Kobrynowicz, 1999).

This model is particularly worthy as a way of addressing the shortcomings in current stereotyping research in teams. It has the potential to enable researchers to measure simultaneously a variety of sources of stereotyping on teams. In addition, it is particularly useful in this study’s context, because it can be combined with military-based assessment measures to detect the presence of status-based biases. For example, in a particularly striking demonstration of shifting standards, researchers observed a random selection of U.S. Army officers assigned to teams of twelve to fifteen while attending a professional development course (Beirnat, Crandall, Young, Kobrynowicz, &

Halpin, 1998). The relevant stereotype present in this particular domain is gender, because women often are discriminated against with respect to leadership in military settings (Boyce & Herd, 2003). In this study, officers were asked to judge each of their teammates' leadership competence utilizing two different types of scales. The first was a more subjective Likert-type scale on which they judged team members in a number of leadership competencies, and the second was a more objective ordinal ranking scale, on which they rank ordered their teammates, with respect to the same leadership competencies. In accordance with the shifting standards model, on the Likert-type "subjective" scale, men rated women much closer to men on all leadership dimensions, and yet on the ordinal-type ranking scale, men ranked women much lower than men on all leadership dimensions.

For example, in the competency of "Decision-making", women were consistently rated on a Likert-type scale as "Average" to "Above Average" with respect to their abilities as decision makers. However, when ranked in that same competency against their teammates, they were consistently ranked in the bottom one third of their teams (Biernat, et al, 1998). This systematic *shift* between judgments on relatively more subjective Likert-type scales and relatively more objective ranking scales is evidence of the shifting standards effect, and an indicator of the presence of status-based stereotyping. In this way, the shifting standards model can be used to detect otherwise well-masked stereotyping and prejudice in teams.

Although shifting standards has been used to measure stereotyping on a variety of dimensions by individuals among individual targets, shifting standards has not yet been utilized to examine the relationship between

stereotyping and group performance (Personal communication with Professor Monica Biernat, October, 2011). The power of this model to detect the presence of stereotyping is amplified if used in conjunction with ML-SEM, in order to examine more closely the relationship between stereotyping and a team's performance.

Summary

These bodies of research refine our understanding of the conditions under which stereotyping is activated, and the impact that activation can have on individuals in groups, as well as the group's performance. Regardless of the source of stereotyping, we have come to understand that the presence of stereotyping has the potential to impact teams, but the conditions under which that happens, are inconclusive (Millikin & Martens, 1996; Cohen & Bailey, 1997). This lack of convergence on the specific conditions under which stereotyping affects team performance has less to do with the nature of stereotyping, than it does the many complicated ways that diverse teams can and are studied. To date, much of the research has focused on isolating singular aspects of diversity, and then examining those with respect to proximal measures of team performance. It has been suggested that research on stereotyping in groups needs to re-dedicate itself to understanding diversity as a *combination* (emphasis added) of different dimensions of differentiation, rather than continuing to focus on singular aspects of diversity (van Knippenberg & Schippers, 2007). The current research leverages the shifting standards model to detect and measure a variety of sources of stereotyping on teams, and then combines them to determine the overall presence of stereotyping, and its collective impact on team performance

Operationalizing & Measuring Team Performance

An interesting source of confusion in the literature on stereotyping and team performance is how differently team performance has been measured in much of the research. Group performance as a dependent variable lacks an industry standard of measurement, and is therefore conceptualized and operationalized in a variety of ways, which makes many findings in groups research difficult to generalize to other settings (Flanagan, 1956; Smith, 1976; Landy & Farr, 1980; Austin & Villanova, 1992). Tension in the teams literature revolves around which is a more important indicator of group performance: team processes or team outcomes (Watson, Kumar, & Michaelson, 1993; Salas, Bowers, Cannon-Bowers, 1995). Although there is sound rationale on either side, there is no general consensus of specifically how group processes (cohesion, decision-making, information-sharing) impacts group outcomes (time, speed, distance, wins, losses).

In their meta-analysis of the ways that team performance has been conceptualized and measured, Beal, et al. (2003) suggest that the dependent measures of team performance should focus on team processes (such as cohesion, decision-making, and information sharing) and not team performance (wins, losses, effectiveness). Beal, et al. further assert that team performance should be measured by team efficiencies rather than team effectiveness, because of the inability to isolate all of the potential influences of actual team performance in real time. Contrary to Beal, et al. (2003), other voices in the field suggest that performance outcomes are more accurate indicators of team performance, but their conceptualization of those outcomes is varied and inconsistent. For example, Wegge, et al (1993) used a case-based competition to

determine the problem-solving acumen of teams, but then utilized judges to assess the effectiveness of the teams' solutions, rather than requiring teams to actually implement their solutions to determine their effectiveness. In another study claiming to measure the impact of team diversity on team outcomes, Webber & Donahue (2001) used self-reports of team cohesion as the indicator of each team's performance. In yet another study, Roberge & van Dick (2010) operationalized team performance as the level of turnover, conflict and cooperation present on teams. These inconsistent conceptualizations of team performance make it difficult to link levels of stereotyping to generalizable indicators of team outcomes.

I suggest that group processes are important, but not sufficient indices of group outcomes. There are plenty of highly cohesive and effective decision-making teams that fail to meet their objectives. Hence, I assert that in addition to a team's processes, a team's performance should be measured by performance outcomes, and in this study, I do just that. In this competition, teams race against a clock and each other, and thus, their performance is determined by the sequence in which they cross the finish line, and by the time it takes them to complete the challenging three-day course.

Despite the copious attention paid to stereotyping and team performance there are relatively few studies that attempt to measure the level of stereotyping present on a team, and its impact on *objectively measured* team performance outcomes. I suggest that there is an empirical gap in the area that connects the absolute value of the level of stereotyping on a team to that team's objective performance measures. This study addresses that empirical gap by utilizing the shifting standards model in conjunction with ML-SEM to measure the level of

stereotyping in a number of teams, and then leveraging juried competition standards to measure the teams' performance. In so doing, this project was able to determine the extent to which the presence and level of stereotype-based judgments impact a team's objective performance.

Research Questions

Specifically, in this research I sought to answer whether or not there was evidence of the shifting standards effect at the team level in a sample of 37 military teams competing head to head in a military skills competition, and if so, whether or not there was variability in the shifting standards index across the teams. If there was evidence of variability in the shifting standards effect across these teams, then I sought to determine what specific sources of stereotyping were responsible for that variability. In particular, and as reviewed in the literature above, I examined variability in the shifting standards index associated with gender, race and cadet class as well as an assimilation of all those categories together. Finally, I sought to determine if there was a relationship between the level of overall stereotyping on each team (as indicated by that team's shifting standards index), and objective measures of team performance.

Chapter 3: Methods

The Study

Research Design

Competition Participants

The study consists of 37, nine to 13-person teams, which competed in a military skills competition held at the United States Military Academy (USMA) in West Point, New York. By regulation, each team was comprised of participants from each of the four classes at West Point (freshman, sophomore, junior and senior), and no less than one woman. There was no race-based team composition requirement. A total of 447 cadets participated in the competition. Of these, 18% were women, 24% were non-White, and 62% were lower class (freshmen and sophomore) cadets. No team had more than two women members, and six teams had only one woman. There were three all White teams.

Competition Overview

The Sandhurst Competition is held at the United States Military Academy every April. The competition is designed to enhance small unit leadership qualities, develop teamwork, and promote military excellence. It consists of a series of knowledge and skill-based military events that are timed and judged over the period of three days. Each year, teams from all over the world participate in the three-day competition. (See Table 1 for a list of events, and See Appendix A for scoring procedures). In 2012, 55 teams competed in the competition. For the purposes of this research, only the 37 teams fielded by the USMA participated in the study.

Table 1

List and description of Sandhurst Competition events

Event	Event Description
Marksmanship	Weapons firing from multiple positions, direct fire planning and fire control techniques
Land Navigation	Complete a 5 mile orienteering course
Obstacle Course	Negotiate a course with 10 challenging obstacles
Weapons & Grenades	Negotiate a live hand grenade course; disassemble and assemble a variety of weapons systems to include the M4, M240, M240B, M9 and foreign weapons
Rappel	Negotiate and descend a 75 foot rock face by constructing their own Swiss seats and equipment slings
Zodiac Boat Race	Launch and paddle a ¼ mile water course
Rope Bridge	Conduct a stream crossing by utilizing an improvised bridge constructed out of the team's provided equipment
Tactical Combat Casualty Care	Perform life saving battlefield first aid after encountering an IED; conduct combat air evacuation techniques
DMI Challenge	Demonstrate lateral thought, mental agility, ability to assimilate a problem, & decision-making under physical stress
Time on Course	The time, including penalties, it takes the team to complete the entire competition.

Measures

Survey Instrument

I developed an online instrument, which contained leadership competencies utilized by the United States Military Academy's Cadet Evaluation Report (CER). The CER is a developmental assessment instrument familiar to all cadets and cadet cadre, and is comprised of nine different leadership competencies (See Table 2 for a list and description of each competency). For this research, the survey instrument asked cadets to both rate and rank each other and themselves in each of the nine leadership competencies. The ratings were scored on a 6-point Likert scale (Very Ineffective, Ineffective, Somewhat Ineffective, Somewhat Effective, Effective, Very Effective), and the rankings

required each cadet to list each of their teammates in order of their effectiveness in each competency. The first position was reserved for the cadet judged to be the most effective in that competency, and the final position was reserved for the cadet judged to be the least effective in that competency.

Table 2

List and description of the nine leadership competencies

Competency	Competency Description
Communicating	Displays good oral, written and listening skills for individuals/groups
Decision-Making	Employs sound judgment, logical reasoning and uses resources wisely
Motivating	Inspires, motivates, and guides others toward mission accomplishment
Planning	Develops detailed, executable plans that are feasible, acceptable and suitable
Executing	Shows tactical proficiency, meets mission standards, and takes care of people/resources
Assessing	Uses after-action and evaluation tools to facilitate consistent improvement
Developing	Invests adequate time and effort to develop individual subordinates as leaders
Building	Spends time and resources improving teams, groups and units
Learning	Seeks self-improvement and organizational growth; envisioning, adapting and leading change.

Cadet Level Variables

Cadet Gender: Cadets reported their gender on the survey's demographic portion. (Males = 0, and Females = 1).

Cadet Race: Cadets reported their race on the survey's demographic portion. (White = 0, and all Non-White = 1).

Cadet Class: Each cadet's class (grade or year group) was provided by the United States Military Academy's Department of Military Instruction. (Juniors and seniors = 0, and sophomores and freshmen = 1).

Cadet Diversity Index (DVI): From the three status characteristics listed above, gender, race and class, each cadet was given a score ranging between a 0 and a 3, computed as a sum of the three corresponding indicator variables to reflect his or her individual diversity index. For example, a White male upperclassman, would be given a diversity index of 0, while a non-White female lower classman, would be given diversity index of 3.

Cadet Shifting Standards Index (SSI): Before calculating the shifting standards index for each cadet, the data had to be transposed, such that each row contained the received ratings and rankings for each cadet. When the data were withdrawn from the survey instrument, each row represented a cadet's ratings and rankings of his/her teammates – the judgments of others made *by* that cadet. In order to determine each cadet's own received ratings and rankings, the data had to be transposed so that each row represented the cadet's received rankings and ratings – the judgments of others made *about* that cadet. See Exhibit 1 for an example of how the data was transposed. In this example, the upper box titled "decision-making ratings from survey" represents how the data was arranged when it was pulled from the survey. The data is arranged such that each row represents a cadet as a rater, and his/her row contains all of the ratings that cadet made about his/her teammates in each of the nine competencies. In this case, Joe's row represents Joe rating Sally a 2, Fred a 5 and Tom a 6 in the leadership competency "Decision-Making". After transposition, each row represents a cadet as a ratee, and his/her row contains all of his/her ratings

received from all of his/her teammates in each of the nine competencies. In the lower box titled “decision-making ratings transposed”, Joe’s row now represents Joe’s received ratings from Sally, a 2, Fred, a 6, and Tom a 3.

Exhibit 1

Exhibit depicting an example of how the data was transposed in order to calculate the shifting standards index (SSI)

Decision-Making Ratings From Survey Instrument					
RATEES					
RATERS		Joe	Sally	Fred	Tom
	Joe		2	5	6
	Sally	2		4	5
	Fred	6	4		5
	Tom	3	5	3	

Decision-Making Ratings Transposed					
RATEES					
		Joe	Sally	Fred	Tom
RATERS	Joe		2	6	3
	Sally	2		4	5
	Fred	5	4		3
	Tom	6	5	5	

Once the data were transposed, I first reverse coded all rankings so that higher numbers represented higher rankings. For example, if a cadet named “Jane” was ranked as the best decision maker on her 12-member team, her associated value with respect to decision making would be a “1”. After the reverse coding is completed, Jane’s value with respect to decision-making would become a “12”.

I then generated variables representing the average rating, $rate_{ij}$, and the average ranking, $rank_{ij}$, for each cadet_i on team_j across all competencies C: as given by

$$rate_{ij} = \frac{1}{C} \sum_{C=1}^C rate_{cij}$$

and

$$rank_{ij} = \frac{1}{C} \sum_{C=1}^C rank_{cij} \text{ where } C = \text{the number of competencies.}$$

Then, I computed the team average rating, \overline{rate}_j as given by $\overline{rate}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} rate_{ij}$

and ranking, \overline{rank}_j as given by $\overline{rank}_j = \frac{1}{n_j} \sum_{i=1}^{n_j} rank_{ij}$ where n_j = number of cadets on

team_j.

I also computed the standard deviation of the average ratings and the average rankings of cadets on each team. The standard deviation of mean cadet ratings for team_j is given by

$$S_j^{rate} = SD (rate_{ij})$$

The standard deviation of mean cadet rankings for team_j is

$$S_j^{rank} = SD (rank_{ij})$$

I then standardized the average rating and ranking for each cadet on each team as given by

$$z_{ij}^{rate} = \frac{rate_{ij} - \overline{rate}_j}{S_j^{rate}},$$

$$z_{ij}^{rank} = \frac{rank_{ij} - \overline{rank}_j}{S_j^{rank}}.$$

The shifting standards index (SSI) for each cadet on each team was then calculated as the difference between the cadet's standardized rating and ranking.

That is,

$$SSI_{ij} = z_{ij}^{rate} - z_{ij}^{rank} .$$

Team Level Variables

Team Diversity Index (TDVI): For each team I calculated the mean of all of the team members' DVIs.

Team Performance: The team performance score awarded by the judges and officials of the Sandhurst Competition represents the score each team earned overall in the three-day competition. The team performance scores ranged on a scale from 0 to 500 points and were calculated using a formula that consisted of the time each team required to negotiate all of the events, their adjudicated performance in each task, combined with points associated with their earned penalties. (See Appendix A for a description of possible penalties). Point scores in each of the events were standardized in order to facilitate the comparative analysis across events.

Procedure

Survey

After training together for six weeks, but eight weeks before the competition took place, cadets completed an online survey (See Appendix B). In this survey, cadets rated and ranked their teammates, and themselves in each of nine dimensions of a leadership competency model outlined by the USMA Cadet Evaluation Report (CER) (For a list of competencies, see Table 2).

Analysis

I used multi-level structural equation modeling (ML-SEM) to determine the effects of stereotyping on team performance. Level 1 is the Cadet Level, and Level 2 is the Team Level. I fit two models, Model A and Model B, to answer

Research Questions 1 and 2. I fit two models, Model C and Model D, to answer Research Questions 3.

Model A

Cadet Level:

$$SSI_{ij} = \beta_{1j}(DVI_{ij}) + \varepsilon_{ij}$$

$$\varepsilon \sim N(0, \sigma^2)$$

Team Level:

$$\beta_{1j} = \alpha_1 + \zeta_{1j}$$

$$\zeta \sim N(0, \tau)$$

β_{1j} = The change in the mean of SSI for every one unit change in cadet DVI (i.e. global stereotyping) on team_j.

α_1 = The average of global stereotyping across all teams corresponding to a one unit difference in cadet DVI.

$\text{Var}(\zeta_{1j})$ = The variance of global stereotyping across all teams corresponding to a one unit difference in cadet DVI.

Model B

Cadet Level:

$$SSI_{ij} = \beta_{2j}(gender_{ij}) + \beta_{3j}(race_{ij}) + \beta_{4j}(class_{ij}) + \varepsilon_{ij}$$

$$\varepsilon \sim N(0, \sigma^2)$$

Team Level:

$$\beta_{2j} = \alpha_2 + \zeta_{2j}$$

$$\beta_{3j} = \alpha_3 + \zeta_{3j}$$

$$\beta_{4j} = \alpha_4 + \zeta_{4j}$$

$$\zeta \sim MVN(0, \tau)$$

β_{2j} is the level of gender stereotyping (difference in mean SSI_{ij} for females compared to males) on team j , controlling for race and class.

β_{3j} is the level of race stereotyping (difference in mean SSI_{ij} for Non-Whites compared to Whites) on team j , controlling for gender and class.

β_{4j} is the level of cadet class stereotyping (difference in mean SSI_{ij} for Third and Fourth class cadets compared to First and Second class cadets) on team j , controlling for gender and race.

α_2 = The average stereotyping with respect to gender across all teams.

α_3 = The average stereotyping with respect to race across all teams.

α_4 = The average stereotyping with respect to cadet class across all teams.

$\text{Var}(\zeta_{2j})$ = The variance of stereotyping with respect to gender across all teams.

$\text{Var}(\zeta_{3j})$ = The variance of stereotyping with respect to race across all teams.

$\text{Var}(\zeta_{4j})$ = The variance of stereotyping with respect to cadet class across all teams.

All variables at Level 1 are group-mean centered. All variables at Level 2 are grand-mean centered. There is no intercept for SSI at Level 1 because SSI is standardized within each team.

Research Question 1: Is there evidence of team-level shifting standards within teams competing in a skills competition?

Evidence of team-level shifting standards with respect to all sources of stereotyping is determined by whether or not the value of α_1 in Model A is statistically significantly different from zero. Evidence of shifting standards (i.e. stereotyping) with respect to gender, race and cadet class is determined by

whether or not the value of $\alpha_2, \alpha_3,$ and α_4 in Model B are statistically significantly different from zero, respectively.

Research Question 2: Is there variability in team-level shifting standards across the teams?

In order to answer Research Question 2, I compared the level of variability of shifting standards across teams with respect to global stereotyping, and the mean differences in the shifting standards index for the different status characteristics variables (gender, race, and cadet class). Respectively, this variability is denoted by the values of $(\text{var})\zeta_{1j}, (\text{var})\zeta_{2j}, (\text{var})\zeta_{3j},$ and $(\text{var})\zeta_{4j}$. These values represent the variability in the level of stereotyping across all teams with respect to a diversity index, and each specific source of stereotyping (gender, race, or cadet class). If these values were statistically significantly different from zero, then this is evidence of variability across teams in shifting standards with respect to the diversity index, or gender, race or cadet class.

Model C

Cadet Level:

$$SSI_{ij} = \beta_{1j}(DVI_{ij}) + \varepsilon_{ij}$$

$$\varepsilon \sim N(0, \sigma^2)$$

Team Level:

$$\beta_{1j} = \alpha_1 + \zeta_{1j}$$

$$\zeta \sim N(0, \tau)$$

$$P_j = \omega_0 + \omega_1(\beta_{1j}) + \xi_j$$

$$\xi_j \sim N(0, \tau)$$

ω_1 = The difference in mean Performance (P_j) corresponding to a one unit increase in global stereotyping (i.e. β_{1j}).

Model D

Cadet Level:

$$SSI_{ij} = \beta_{2j}(gender_{ij}) + \beta_{3j}(race_{ij}) + \beta_{4j}(class_{ij}) + \varepsilon_{ij}$$

$$\varepsilon \sim N(0, \sigma^2)$$

Team Level:

$$\beta_{2j} = \alpha_2 + \zeta_{2j}$$

$$\beta_{3j} = \alpha_3 + \zeta_{3j}$$

$$\beta_{4j} = \alpha_4 + \zeta_{4j}$$

$$\zeta \sim MVN(0, \tau)$$

$$P_j = \omega_0 + \omega_2(\beta_{2j}) + \omega_3(\beta_{3j}) + \omega_4(\beta_{4j}) + \xi_{ij}$$

$$\xi_j \sim N(0, \tau)$$

ω_2 = The difference in mean Performance (P_j) corresponding to a one unit increase in stereotyping with respect to gender, controlling for stereotyping with respect to race and cadet class.

ω_3 = The difference in mean Performance (P_j) corresponding to a one unit increase in stereotyping with respect to race, controlling for stereotyping with respect to gender and cadet class.

ω_4 = The difference in mean Performance (P_j) corresponding to a one unit increase in stereotyping with respect to cadet class, controlling for stereotyping with respect to gender and race.

Research Question 3: Is there a relationship between team-level shifting standards and team performance in the competition?

In order to answer Research Question 3, I fit Model C and Model D regressing a global stereotyping indicator (DVI) on team performance P_j . To counteract the problem of inflated Type I error rates due to multiple comparisons across such a large number of events, I used a Bonferroni corrected α -level of 0.005.

The association between the level of global stereotyping (team level gender, race and cadet class stereotyping) and team performance will be denoted by the value of ω_1 in Model C. The association between the individual sources of stereotyping with respect to gender, race and cadet class and team performance will be denoted by the values of ω_2 , ω_3 , and ω_4 in Model D. The magnitude of these values, and the sign of the coefficient P_j in Model D will determine the impact and the direction of the impact that stereotyping has on team performance.

Chapter 4: Results

I sought to determine if stereotyping was occurring on these military teams as indicated by the shifting standards effect. I also wanted to know if the stereotyping was variable across the teams, and if so, whether or not that variability was related to variability in team performance. Specifically, I was interested in what association, if any, existed between levels of stereotyping, as indicated by the shifting standards effect, and team performance.

Descriptive Statistics: Cadet Level

A total of 313 of a possible 447 cadets on 37 teams responded to the survey. This was a 70% response rate overall. Teams ranged in size from 9 to 13 cadets. The average response rate per team was 7.5 cadets. The minimum number of respondents for a team was two cadets, and the maximum was 12 cadets. Table 3 provides a descriptive for the composition of the competition participants with respect to gender, race and cadet class.

Not all cadets on all teams participated in the survey, however because all teams participated with at least two raters, every cadet in the competition was rated and ranked. There was no significant difference between respondents and non-respondents with respect to gender and race. Lower class cadets responded at slightly higher rates (73% vs. 75%). There was no significant difference in shifting standards between respondents and non-respondents.

Table 3

The number and percentages of competition participants by gender, race, and cadet class.

Sandhurst Competitors					
	Male (82%)		Female (18%)		
	White (80%)	Non-White (20%)	White (56%)	Non-White (44%)	Total
Upper Class (38%)	122/27%	31/7%	14/3%	5/1%	172

Lower Class (62%)	172/38%	44/10%	30/7%	29/6%	275
Total	294	75	44	34	447

Descriptive Statistics: Team Level

Table 4 provides the descriptives at the team level for demographics, size and response rate. Of note, teams have much more variability with respect to the number of lower class cadets (3 to 11) and non-White cadets (0 to 7) compared to females (1 to 2). Team response rate was not significantly associated with team composition, team-level shifting standards with respect to gender, race, or class or team performance.

Table 4

Team averages for response rate, team size, number of females, number of males, number of lower class cadets, number of upper class cadets, number of non-White cadets and number of White cadets.

Team Descriptives				
	Mean	SD	Min	Max
Response Rate	7.486	2.854	2	12
Team Size	11.811	0.569	10	13
Females	1.838	0.374	1	2
Males	9.973	0.440	9	11
Lower Class	7.162	1.692	3	11
Upper Class	4.649	1.531	3	7
Non-White	2.676	1.796	0	7
White	9.135	1.719	5	12

Sample Bivariate Relationship Between the Shifting Standards Index and Gender, Race and Cadet Class (See Table 5).

Gender

I expected female cadets to have lower status than male cadets, and therefore generate higher levels of the shifting standards index than male cadets.

In this sample, females have significantly higher shifting standards index (SSI) than do males (+0.168 vs. -0.036). Their standard deviation is larger, and they have more variability around the mean as well as larger outliers than do males.

Race

I expected non-White cadets to have lower status than White cadets, and therefore, generate higher levels of the shifting standards index than White cadets. However, I found that White and non-White cadets do not differ appreciably with respect to SSI.

Cadet Class

In accordance with institutional norms, lower class cadets have lower status than upper class cadets, and so I expected to see a relationship between cadet class and SSI. In this sample, lower class cadets have significantly higher SSI, on average, indicating they generate more variability in the shifting standards index than upper class cadets.

Table 5

Shifting Standards Index for each demographic category.

Shifting Standards Index						Mean Differences in Shifting Standards
Status Characteristic	Obs	Mean	SD	Min	Max	
Female	68	0.168	0.820	-1.166	2.699	0.196*
Male	368	-0.036	0.398	-1.218	1.897	
Non-white	99	0.024	0.550	-1.000	2.490	0.04 [‡]
White	337	-0.013	0.474	-1.220	2.700	
Lower class	264	0.054	0.493	-1.220	2.700	0.157**
Upper class	172	-0.093	0.480	-1.140	2.489	

* $p < 0.05$, ** $p = 0.001$

Research Question 1: Is there evidence of team-level shifting standards in teams competing in a skills competition?

I fit Model A in order to determine whether there was evidence of shifting standards on teams with respect to a cadet's diversity level, as indicated by his/her diversity index (DVI), and his/her shifting standards index (i.e. how much stereotyping that cadet generated in others). In Table 6, I show on average, across all teams, there is evidence that a cadet's level of diversity is indeed positively associated with SSI (Est.=0.117, SE=0.037, $p=0.001$). The more that cadets possessed lower status characteristics (as indicated by their DVI), the more evidence I found of stereotyping, as indexed by SSI.

In order to determine whether there was evidence of shifting standards on teams with respect to gender, race and cadet class, I fit Model B. The highest level of shifting standards was associated with female cadets (Est.=0.176, SE=0.094, $p=0.061$) and lower class cadets (Est.=0.143, SE=0.052, $p=0.006$). Non-White cadets did not generate an appreciable level of the shifting standards effect (Est.=0.004, SE=0.049, $p=.934$) suggesting that race-based stereotyping was not statistically significantly activated in this context.

Research Question 2: Is there variability in team-level shifting standards across these teams?

In order to determine if there was variability in shifting standards across teams, I examined Model B, and found that teams did vary with respect to their level of stereotyping, but interestingly enough, only with respect to gender (Est.=0.148, SE=0.040, $p<0.001$). See Table 6. The variance associated with race stereotyping (Est.=0.001, SE=0.04, $p=.972$) and class stereotyping (Est.=0.005, SE=0.021, $p=.802$) was striking both in its modesty and in its similarity. This

suggests a similar or shared understanding or perception of lower class and non-White cadets across all teams. Teams demonstrate a general lack of stereotyping of non-White cadets similarly across all 37 teams.

Table 6

Parameter estimates, standard errors and approximate p-values for a taxonomy of fitted multi-level structural equation models predicting shifting standards indices (SSI) by diversity index, and cadet gender, race and class. (n=447)

Model	Random Coefficient	Mean (β)			Variance(ζ)		
		Est. (α)	SE	P-value	Est.	SE	P-value
A	β_1 SSI on DVI	0.117	0.037	0.001	0.000	0.079	0.997
	β_2 SSI on Female	0.176	0.094	0.061	0.148	0.045	0.001
B	β_3 SSI on Race	0.004	0.049	0.934	0.001	0.04	0.972
	β_4 SSI on Class	0.143	0.052	0.006	0.005	0.021	0.802

*Note: The results of a Global Hypothesis Test to determine the impact of all predictor variables simultaneously were significant ($\chi^2=35.00$, $df=3$, $p<0.001$), that $\alpha_2 = \alpha_3 = \alpha_4 = 0$, rejecting the null.

*Note for Model B: I tested all relevant two-way and three-way interactions with the main predictors (Female, Race and Class) and found that there was no evidence of either two-way or three-way interactions for Female, Race and Class with respect to SSI. The Likelihood Ratio Test (adj. LRTS = 1.92, $df = 4$, $p = .75$) confirmed this.

Research Question 3: Is there a relationship between team-level shifting standards and team performance in the competition?

In order to determine if there is a relationship between team level shifting standards indices and the team's performance, I fit Model C and Model D regressing SSI on gender, race, cadet class and DVI at Level 1 and regressing team performance on the team level shifting standards at Level 2 to determine the team level effects. I found that there was no relationship between a team's

level of shifting standards and its overall performance in the competition, as indicated by number of points earned by the team. See top row of Table 7. I then looked at team performance in each of the ten events in the competition to determine if there was a relationship between a team's level of shifting standards and its event specific performance. See Table 7 for the individual events in Model C. As stated in Chapter 3, to counteract the problem of inflated Type I error rate due to multiple comparisons across events, I used the Bonferroni corrected α - level of 0.005.

I found that indeed, in a few instances, a team's level of shifting standards, either as a whole or with respect to gender or race was associated with its performance. For example, a team's overall level of shifting standards (i.e. its global level of stereotyping) was associated with a team's performance in the marksmanship event (weapon firing accuracy) such that higher levels of shifting standards were associated with lower levels of team performance in that event (Est.=-25.128, SE=8.336, $p=0.003$, Std. Est.=-0.829). Similarly, a team's overall level of shifting standards was associated with a team's performance in the DMI Challenge event (See Table 1 for a description) such that higher levels of shifting standards were associated with lower levels of team performance in that event (Est.=-57.798, SE=2.320, $p=0.000$, Std. Est.=-0.099). Conversely, in the rappelling event (rappelling the entire team and its equipment down a 75 foot cliff), higher levels of shifting standards were associated with higher levels of team performance (Est.=21.566, SE=4.955, $p<0.001$, Std. Est.=0.894). When aggregated for the entire competition, these effects are not apparent.

Additionally, a team's level of shifting standards with respect to specific sources of stereotyping (gender and race) had associations with team performance in

some events. For example, in one instance, higher levels of shifting standards with respect to female cadets was associated with lower levels of team performance in the Department of Military Instruction (DMI) Challenge event (Est.=-1.454, SE=0.530, $p=0.006$, Std. Est.=0.006). Also, higher levels of shifting standards with respect to race was associated with higher levels of team performance in the rappelling event (Est.=8.570, SE=2.398, $p<0.00$, STDE=0.808), and the DMI Challenge event (Est.=8.113, SE=1.952, $p<0.00$, STDE=0.655), but was associated with lower levels of team performance in the Zodiac boat course (Est.=-10.020, SE=2.288, $p<0.00$, STDE=-0.900).

Table 7

Parameter estimates, standard errors and approximate p-values for a taxonomy of fitted multi-level structural equation models predicting team performance (as indicated by points earned in the competition) in 10 different events by team level shifting standards indices (SSI) by cadet gender, race and class, and diversity index. (n=447)

Event	Model C				Model D				Model D				Model D			
	β_1 SSI on DVI				β_2 SSI on Female				β_3 SSI on Race				β_4 SSI on Class			
	Est.	SE	p-value	stde4	Est.	SE	p-value	stde1	Est.	SE	p-value	stde2	Est.	SE	p-value	stde3
Total Points	3.040	118.060	0.979	0.055	-0.434	0.715	0.544	-0.167	11.172	18.669	0.550	0.517	-1.678	11.370	0.883	-0.127
Marksmanship	-25.128	8.336	0.003	-0.829	-0.881	0.603	0.144	-0.341	-11.208	7.694	0.145	-0.616	-4.661	7.023	0.507	-0.380
Land Navigation	-0.707	96.611	0.994	-0.013	-0.728	0.546	0.182	-0.280	10.804	12.043	0.377	0.499	-0.529	8.445	0.950	-0.040
Obstacle Course	-2.403	140.776	0.986	-0.044	0.154	0.906	0.863	0.059	-2.607	65.158	0.969	-0.098	-6.765	17.821	0.704	-0.527
Rope Bridge	27.090	41.362	0.513	0.650	0.193	0.824	0.816	0.075	12.627	8.185	0.123	0.754	2.500	10.322	0.809	0.183
Rappelling	21.566	4.955	0.000	0.894	0.345	0.541	0.524	0.156	8.570	2.398	0.000	0.808	4.968	3.771	0.188	0.471
TCCC	27.460	26.135	0.293	0.688	0.367	0.616	0.551	0.143	7.461	24.400	0.760	0.296	7.687	9.400	0.413	0.665
Weapons	0.280	105.090	0.998	0.005	-0.257	0.600	0.668	-0.104	6.841	12.215	0.579	0.311	-7.668	5.249	0.144	-0.827
Boat Race	-26.464	19.042	0.165	-0.797	-0.157	0.672	0.815	-0.062	-10.020	2.288	0.000	-0.900	1.861	10.488	0.859	0.121
DMI Challenge	-57.798	2.320	0.000	-0.099	-1.454	0.530	0.006	-0.574	8.113	1.952	0.000	0.655	-4.091	4.301	0.342	-0.339
Course Time	11.099	114.564	0.923	0.207	0.094	0.814	0.908	0.036	-9.357	32.590	0.774	-0.397	-1.465	14.211	0.918	-0.108

Chapter 5: Discussion

The purpose of this research study was to provide insights into the ways that multiple sources of stereotyping can impact team performance. I designed the study to examine more thoroughly the impact of stereotyping on team performance. I leveraged the theoretical strength of status characteristics theory and the analytic strength of the shifting standards model coupled with multi-level structural equation modeling (ML-SEM) to examine the relationship between a team level index of stereotyping and the team's performance. I utilized the shifting standards model to detect and quantify the levels of stereotyping activated on each team, and then used the performance scores in a juried military skills competition to assess each team's performance.

ML-SEM is uniquely positioned to allow for the detection of the effects of stereotyping in teams, because it enables the modeler to explicitly model observed variables and variation at both the individual (Level 1) and team level (Level 2). In this case, random coefficients specified in the Level 1 model, representing the team specific stereotyping, are allowed to vary across Level 2 units (i.e. the teams). Unlike a standard Hierarchical Linear Model (HLM) in which such random coefficients could only be outcomes at Level 2, in ML-SEM, these random coefficients can serve as predictors of Level 2 outcomes (i.e. team performance).

Global Levels of Stereotyping and Team Performance

I found no evidence that global stereotyping (i.e. stereotyping with respect to gender, race and class altogether) was associated with overall team performance, as measured by either points earned or placement with respect to other teams in a ten-event competition. It has been argued that in field settings,

the impact on team performance of stereotyping with respect to demographic variables may be obviated by functional team norms such as collectivism and teamwork (Bell, 2007). Given the nature of these tightly knit teams, and the pressure of the competition, it may be that the superordinate goal of winning, indeed overshadowed any impacts stereotyping might have had on team performance. Additionally, another perspective suggests that team training specifically focused on task-relevant knowledge, skills and abilities has a moderate effect on increasing adaptive team mechanisms, which result in higher levels of team cohesion, motivation and performance (Mathieu, Maynard, Rapp, & Gilson, 2008). As I stated earlier, these teams trained together for several months prior to the competition, and this may have created the conditions that reduced the effects of stereotyping on team performance. It is possible then, that establishing strong and functional team norms through effective team training mechanisms may eliminate the potentially deleterious effects of stereotyping on team performance.

Although global team-level stereotyping was not associated with a reduction in team performance overall, this is not a story about the absence of stereotyping or about the absence of the effect of stereotyping on team performance. The impact of stereotyping on the Sandhurst teams was far more complicated and nuanced, and when examined further, revealed several interesting patterns. I found evidence that team-level stereotyping did indeed have an effect on three specific events. Higher levels of team-level stereotyping were associated with lower levels of performance in both the marksmanship and the DMI Challenge events, and higher levels of performance in the Rappelling event.

One possible explanation for these varying effects in different events may be related to the nature of the events, themselves. Specifically, the level of interdependence of the task, may affect how stereotyping is activated, and that may manifest itself in the team's performance. According to seminal research by Thompson (1967) there are essentially three levels of interdependence that influence teamwork. Pooled interdependence is the lowest form of interdependence and is characterized by work that doesn't flow between team members; rather each individual team member does his own work in contributing to the whole. This level of interdependence requires the least amount of cooperation and communication. Sequential interdependence is the next highest level of interdependence, and as the name suggests, requires team members to complete their tasks before another team member can complete his or hers. In this way, the work flows sequentially between team members across the entire team before the whole task can be completed. This level of interdependence requires conditional communication and cooperation, which is more than required by pooled interdependence. And lastly, there is reciprocal interdependence, which is the highest level of interdependence, and requires dynamic interaction of all team members throughout all of their activities to complete their tasks (Pasmore, 1988; Thompson, 1967).

Viewed through this framework, we can add a level of granularity to our understanding of the different ways that stereotyping impacted team performance in this study. One could argue that the marksmanship event is a task characterized by pooled interdependence, and thus requires the least amount of cooperation and communication. Teams are scored on the marksmanship range by how well they position themselves against an array of

moving targets, and how accurately each cadet fires against those targets. Success on the range can be enhanced by more verbal exchanges among team members scouting the range for each other and themselves. When cadets communicate actively while on the firing line, they can help each other spot and engage enemy “pop up” targets, resulting in more targets being engaged and destroyed. This kind of communication is not an explicit norm of conduct on the range. If teams were experiencing higher levels of stereotyping in general, there may have been less communication initiated among teammates while firing (Lau & Murnighan, 2005). Presumably this could have reduced their ability to coordinate their firing against targets, which may have resulted in less “kills” achieved on the range. The lower number of kills would have resulted in a lower team performance score. Given that teams that do communicate more often during the marksmanship likely perform better, this might explain why those teams that are experiencing higher levels of stereotyping, and thus perhaps lower levels of communication, may perform more poorly on the marksmanship range. This is consistent with research on pooled interdependence that suggests that if teams overcome the barriers to communication, they perform better (Marks, Mathieu, & Zaccaro, 2001).

Additionally, stereotype threat research suggests that the threat of confirming a negative stereotype about a group one belongs to can actually create enough performance pressure to interfere with one’s performance (Rosenthal, Crisp, & Suen, 2007; Shih, Pittinsky, & Ambady, 1999; Steele & Aronson, 1995). Given that pooled task interdependence makes individual contributions to the team more visible, it is plausible that stereotyped individuals may have felt more vulnerable to their group’s stereotype, and therefore

experienced some fear of fulfilling that stereotype. In response to that fear, they may have performed more poorly on the task, therefore reducing the team's overall performance in the event.

The second event in which team-level global stereotyping had a deleterious effect on team performance was the Department of Military Instruction (DMI) Challenge. This event is the very last event of the three-day series. Teams are scored in two activities in the DMI Challenge. The first activity challenges the team to figure out how to disassemble a 155mm Howitzer cannon from its firing position, prepare it for movement, loading its auxiliary equipment onto its base plate assembly, and then without assistance, move the entire cannon and its load, 100 yards across an open field. This task requires high levels of coordination, cooperation, knowledge sharing and strength. The second task in the DMI Challenge is an intellectually challenging task requiring teams to answer a series of questions about what they saw as they travelled through the many different sites of the competition's events. This task tests cadets' level of attention to detail and requires them to assemble subtle bits of information they should have noticed while they were competing. For example, the test posed questions such as, "At the Weapons event, there was an enemy armored personnel carrier (APC) parked on the right side of the medical tent. What nation's flag was painted on the side of that APC?" or "Posted on a tree at the Rope Bridge event, there was a photograph of a man on a bicycle. Who was that man? What color was the bicycle?" or "When you left the Land Navigation course, you ran by the Range OIC's cabin. What was the number and the slogan painted on the side of the Range OIC's cabin?" These questions were designed to challenge the cadets' attention to detail and memory under pressure. The

cadets are then tested on how well they can construct consistent responses from their shared mental models. Success in the DMI Challenge depends on physical strength, teamwork, information sharing, creative problem solving, mental agility, and the team's ability to tap into its collective team memory. Some research suggests that in diverse groups, stereotyping interacts with information sharing and decision-making in a way that actually disrupts a team's ability to process task relevant perspectives and information (Van Knippenberg, De Dreu, & Homan, 2004). The teams that generated higher levels of global stereotyping may have communicated and shared information less than teams that had lower levels of stereotyping. This may have impeded the level of cooperation among teammates while solving both challenging problems (to move the howitzer and to reconstruct the factual details along the course). Presumably this could have reduced their ability to tap into the complimentary talents and ideas of individual cadets, thereby compromising the team's performance.

In contrast to the association that team level stereotyping had with team performance at the marksmanship range and in the DMI Challenge, higher levels of team level stereotyping was associated with higher levels of performance in the rappelling event. Although counterintuitive, this effect might be explained by the ways that the use of some stereotypes may improve a team's ability to leverage its members' varied abilities. For example, in the rappelling competition, part of the team's success is getting a number of strong team members to the bottom of the cliff first, so they can act as belay for the rest of the team. Performance at belay is enhanced by height, and exceptional upper and lower body strength. These demands may have influenced teams to send their tallest and most physically fit cadets down the cliff face first, to take the role of

belay. Given these selection standards, teams may have filtered their choices by gender and race to select team members to be on belay. This might have enabled the rest of the team to rappel more rapidly, and therefore score higher in the event.

In a similar vein, the rappelling event requires sequential interdependence, and as such, the determination for who goes over the cliff first is an important one. All of the subsequent tasks depend on the one before it. If stereotyping leads to the team picking the larger more athletic looking males to rappel first, and then set up belay for the rest of the team and its equipment, then this perhaps generates a positive outcome for the team's performance. In this way, a higher level of stereotyping may have been associated with better team performance. Clearly this is conjecture, and requires further research.

The relationship between team level stereotyping and team performance is no doubt complicated, and requires further study. In this study, overall, team level stereotyping had no impact on a team's overall performance. There is evidence that all of the sources of stereotyping combined influence just three of nine team tasks, and these effects are not apparent in the team's overall performance. This was the driving question of this research project, but the investigation does not stop here. It is of interest to examine the ways that individual sources of stereotyping with respect to gender, race and cadet class impacted (or not) the team's performance in individual events of the Sandhurst Competition.

Demography Specific Stereotyping and Team Performance

Gender-based Stereotyping and Team Performance

The existence of higher levels of shifting standards when female cadets are

being judged is not surprising since women are stereotype relevant targets in this uniquely masculinized environment of military skills competition (Boldry, Wood, & Kashy, 2001; Mettrick & Cowan, 1996). However, stereotyping of female cadets was associated with lower team performance in only one event, the DMI Challenge event.

In addition to the explanation above, regarding the team's inability to tap into all of its talent, research on fault lines lends some insight into the impact of demographic diversity on team performance (Mannix & Neale, 2005; Pham, Metoyer, Bezrukova, & Spell, 2012; Thatcher & Patel, 2011). Fault lines occur in teams when members fall into two or more distinct demographic categories (in this case, female cadets and male cadets, or White cadets and non-White cadets). "Fault lines" are fractures or divides that lay between these demographic categories in teams, and can be activated by external forces (such as competition pressures). Fault lines can be responsible for conflict that impacts team performance when the demographic categories are related to the group's task (Lau & Murnighan, 2005; Lau & Murnighan, 1998). Additionally, fault line research suggests that in teams tasked with creative thinking or sharing mental models, (not unlike the DMI Challenge event), gender fault lines can affect team performance negatively (Pearsall, Ellis, & Evans, 2008). This may explain the effects associated with gender stereotyping and team performance in this specific event.

Race-based Stereotyping and Team Performance

Most striking was the relatively low levels of racial bias as indicated by the absence of any substantial shifting standards effect with respect to White and non-White cadets. This most likely can be attributed to the strong prescriptive

norms at the U.S. Military Academy mandated by the Respect for Others program (<http://www.usma.edu/scpme/SitePages/Respect.aspx>, accessed March 8, 2013). The Respect for Others program is an institutionalized four-year developmental program designed to improve interracial understanding, dialogue and respect. Although the Respect for Others program includes gender as a topic, what might retard the impact of that aspect of the program is the undeniably masculine nature of the military academy. This finding is consistent with other research related to race and military teams, which found similar patterns of relatively insignificant effects with respect to racial stereotype activation using the shifting standards model in military settings (Biernat, Crandall, Young, Kobryniewicz, & Halpin, 1998).

However, there were two specific events (the Zodiac boat course and the DMI Challenge), in which stereotyping of non-White cadets, was associated with lower levels of team performance. In the United States there exists a stereotype that non-Whites are less successful swimmers, and are less comfortable around water than Whites (Hastings, 2006; Sailes, 1991). It has been suggested that this is attributable to the relatively diminutive access to public and private swimming opportunities for non-Whites in the United States, which is a class-based and not a race-based limitation. At the United States Military Academy, specifically, non-White cadets comprise 67% of the remedial swimming classes (Personal Communication with Brigadier General (Retired) Maureen LeBoeuf, former Chair of the USMA Department of Physical Education, the office responsible for teaching all cadets how to swim). If teams engaged this stereotype, it may have had an associated effect on their performance in the Zodiac boat race. The race requires the entire team, wearing 45 pounds of full combat gear and an M16 rifle,

to balance on the edges of an inflatable Zodiac boat well beyond swimming distance to the shore, and paddle through a difficult obstacle course. If team members held concerns about either their own or any other team member's ability to handle the water wearing all of their gear, this may have influenced the effectiveness of the teamwork required to complete this event. Additionally, similar to gender fault lines in teams, race based fault lines can act to deleteriously affect team performance by interrupting task-relevant communication (Lau & Murnighan, 2005). In this case, that lack of task-relevant communication may have retarded the team's ability to effectively maneuver the Zodiac boat around the difficult water born obstacle course, and this could have hindered the team's performance.

The second event, in which race based stereotyping was associated with lower team performance, was the DMI Challenge. Similar to the gender-based fault lines discussed with respect to performance in this same event, race-based fault lines may also explain the drop in team performance. However, there is another possible explanation. Crawford and LePine (2013) have outlined two kinds of team processes, task work and teamwork, both of which have an impact on team performance. Task work involves how team members interact with tools, machines and equipment in service of the team's tasks. Teamwork involves how team members interact with each other to accomplish the team's tasks. The authors suggest that for insular teams (teams whose membership does not change readily – much like the Sandhurst teams), high levels of task work can actually impede the generation of new ideas. Their reasoning for this is that the closed network increases the level of redundancy of information circulation, and decreases individual initiative. Other research has shown that

gender and race based stereotyping also impedes the open exchange of ideas in teams (Horwitz & Horwitz, 2007). Because the DMI Challenge depended on the open exchange of novel ideas in a time constrained environment, the combination of both the task work demands of the Sandhurst competition, and the gender and race-based stereotyping may explain the association with lower levels of team performance in that event.

Cadet Class-based Stereotyping and Team Performance

A bit unexpected, but certainly logical, is both the magnitude and relative lack of variability across teams of cadet class-based shifting standards. The second largest source of shifting standards was generated by cadet class differences, and it was consistent across all teams. Lower class cadets invariably were the targets of a substantial amount of shifting standards, more so than females or non-White cadets. Equally surprising is that these strong class-based shifting standards were never associated with team performance in the overall competition or any of the specific events. Upon further reflection, this makes sense in an organizational context that has a more hierarchical and authoritative culture, such as the U.S. military or U.S. education. In this case, lower class cadets are relatively less experienced at the variety of skills tested in the competition, as well as the leadership competencies in which they were being assessed. In addition, being a lower class cadet fulfills the functional role of “subordinate” in a military team, and thus plays an integral part in the team’s ability to accomplish its missions. Roles generate clear expectations for and increase the salience of a team member’s abilities (Garcia-Prieto et al., 2003). Because of this role specific and institutionally supported norm, freshman and sophomore cadets were ranked lower than upper class cadets on most

competencies. However, on the relatively more subjective Likert-type scales, lower class cadets were rated much higher, perhaps because teammates recognized that these lower class cadets were motivated, hard working and valuable members of the team. This shift between the two scales (ranking and rating) is indicative of the classic shifting standards effect. Because this was repeated similarly across all teams, cadet class generated the second highest levels of shifting standards. Stereotyping with respect to cadet class, however, was not associated with variance in team performance in either the overall competition or any of the individual events. In accordance with military culture, this is likely due to the consistent way that lower class cadets are perceived with respect to leadership competencies.

The Nature of Military Teams

Stereotyping has a complicated and complex impact on team performance. The nature of military teams may contain additional explanatory insights into why stereotyping did not have a relationship with overall performance in the Sandhurst competition. More than most, these participants have engaged in formative experiences that disconfirm traditional stereotypes. Given the nature of the United States Military Academy and cadet life, Gordon Allport's Contact Hypothesis comes to mind, potentially, as an explanatory concept (Allport, 1954). This theory argues that when majority group members have meaningful and interdependent interactions with stereotyped individuals, stereotyping is reduced. The implication is that if we want to reduce intergroup conflict, we should increase contact and interaction between majority and minority group members. Rothbart and John (1985) refined the conditions under which true stereotype reduction would take place, and these are all relevant to

this study. They posited that in order to cognitively reduce stereotyping, majority individuals must be susceptible to acknowledging the existence of stereotype disconfirming data from minority individuals, and that they then must generalize those impressions to the minority's group. In addition, majority and minority members must interact in a variety of settings, and be focused on a common goal. When all of these conditions are met, stereotyping is reduced (Rothbart & John, 1985). In a place like the United States Military Academy, there is constant contact among highly talented members of all ethnicities and genders, much of that contact is in different settings and often focused on superordinate goals. Given that these conditions meet the requirements for stereotype reduction established by Rothbart and James (1985), it is no wonder that stereotyping's impact on team performance is isolated and minimal.

Implications

For Practice

The implications of this study are relevant to both research and practice. With respect to practice, these findings might provide insights for other domains, such as education leadership and executive education. As the use of teams expands in education and in business, it is imperative that we better understand the conditions that optimize team performance.

For leaders and executives, this may suggest a change in behaviors. Specifically, in organizations, we as leaders tend to put lower status people in situations that require pooled interdependence. We do so to give them an opportunity to demonstrate their individual talents, and therefore provide them an opportunity to shine on their own, rather than be dependent on others. However, this research suggests if our goal is to optimize team performance and

denigrate the potential deleterious impact of stereotyping in teams, that in addition to providing lower status individuals opportunities to demonstrate their expertise, we should also ensure they are part of mixed teams that face emergent challenges that demand reciprocal interdependence. So, rather than giving low status individuals autonomy, independence and asking them to “prove” themselves as solo artists, we should give them more chances to collaborate with team mates on challenging tasks.

In U.S. Education, as well, teams have become critical to learning and teaching, and it has been demonstrated that if teachers understand that stereotyping can impact team performance, they can be more conscious about how they design team composition, in order to optimize student team outcomes (Cohen, 1994). At the school leadership level, teams have increasingly become the fulcrums upon which school operations are executed, and this elevated dependence on teams has increased the need to better understand how to optimize their performance (Chrispeels, Castillo, & Brown, 2000). Careful attention to sources of stereotyping can be effective at the district level, as well. School district management teams are complex and multi-faceted, and often are hindered by their inability to leverage all of the talent on their teams (Higgins, Weiner, & Young, 2012). My hope is that this research will contribute to and further our understanding of how to design teams to optimize the potential for diversity to positively, rather than negatively, impact performance.

For Research

This research lays the foundation for some exciting new frontiers. In the future, it would be beneficial to examine the perceived relative masculinity and femininity of each of the leadership competencies in which cadets rated each

other. It would be informative to see if there is a gender or race-based association between how masculine or feminine a competency is perceived to be, and who is rated and ranked higher in those competencies. This would lend deeper insight into the way gender, race and class based stereotyping impacts interpersonal assessments of competence.

In addition, this research and the methodology employed here to examine team level stereotyping could be extended into different professions. Future research could engage in comparative analyses to determine how the perceived relative masculinity/femininity of the team's professional domain influences the way male and female targets are judged with respect to different leadership competencies, and then how the team's performance is impacted. If compared across a variety of relatively more masculine and feminine domains, we may gain insight into the ways that perceptions of the profession's masculinity or femininity influences perceptions of male and female team mates' competence.

In the future, research that extends this work could more closely look at specific sources of stereotyping (gender, race and cadet class), and their relationship to a variety of leadership competencies. Specifically, it would be interesting to study perceptions of the relative masculinity and femininity of leadership competencies, and how they are rated with respect to their importance to the overall construct of leadership effectiveness. It would also be interesting to look at the perceived relative masculinity and femininity of specific tasks in different professions, and then compare that to the gender of those individuals judged as the most effective in those tasks.

It also might be fruitful to investigate notions of gender and the relationship between gender roles and perceptions of sexuality in service of

understanding the biases that are found in teams of both heterosexual and homosexual members. In addition, it would be interesting to make a significant shift away from studying stereotyping and prejudice in teams, and instead, examine the ways that team members' *liking* for each other influences performance. Research toward those ends could engage the work of Todd Pittinsky and his work on allophilia, which focuses on the ways that positive feelings toward outgroup members enhances team process and performance (Pittinsky, Rosenthal, & Montoya, 2011; Pittinsky, 2005; Shih et al., 1999). The ultimate purpose of this stream of research in stereotyping and team performance is to better understand the myriad ways that stereotyping can impact team process and performance, so that we can design education and training programs that can increase the positive effects and decrease the deleterious effects of these dynamics for leaders and members of teams.

And lastly, in the future, we could add to this research by leveraging Hackman and Wageman's (2001) Team Diagnostic Survey by linking its five conditions for team effectiveness to the teams of the Sandhurst competition. This would enhance the level of granularity of the current study by adding team dimensions that were not considered. For example, it might add value to determine how well the Sandhurst teams manage relations among members, or sustain the motivation and satisfaction of their members (Hackman & Wageman, 2001). In addition, it might add value to examine more closely the teams enabling structures, to see if there are differences across teams. Each of these additional lenses could potentially add significantly to what we now understand about the impact stereotyping has on the teams' performance. These insights could also extend team leadership research by providing more prescriptive

recommendations for carefully attending to specific dimensions of team life.

Limitations

There are several limitations of this research. Although the aim here was to more directly link measures of team-level stereotyping and team performance, there are many mediators whose measurement would lend more nuanced interpretations of the effects of stereotype-based judgments on team performance. For example, Chatman & Flynn (2001) found that cooperative group norms actually reversed some of the deleterious effects demographic heterogeneity had on group outcomes. In future research, scholars could examine group process oriented mediators, in order to determine which has the potential to obviate the effects of stereotyping and prejudice.

In addition, this research does not lend insight into other kinds of teams, such as virtual or geographically distributed teams, whose potential for contact-based stereotyping is somewhat limited. Future research might target such teams to determine which, if any, sources of stereotyping prevail, when physical presence is not a dominant aspect of interpersonal contact. Furthermore, in the absence of individual level performance data for each and every event in the competition, as well as observations of individuals with respect to the competency domains, it is more difficult to determine just how biased or accurate some interpersonal assessments are. It has been argued that, to its own detriment, field research on team performance far too often ignores the influence of task-specific skill diversity (Jackson et al., 2003). In the future, I suggest designing the study to capture the actual task-specific skills of each of the cadets on the teams, thereby allowing for more accurate inferences regarding the interpersonal assessments of competence rendered by each cadet.

With respect to research, I strongly recommend a more multi-modal approach to this kind of study. This would improve its capacity to offer insights into the phenomena of interpersonal biases and how they impact team performance. Specifically, the use of qualitative methods, such as ethnography, could better capture the lived experience and perceptions of the cadets going through the training and the competition. Giving voice to participants provides the opportunity to garner a much more detailed understanding of the dynamics of team life. This more robustly qualitative approach, could provide more insight into the multi-dimensional nature of diversity, and hence how it is interpreted in this study (Jackson et al., 2003).

Conclusion

This research adds to the literature on stereotyping and team performance by leveraging the strength of multi-level structural equation modeling to enhance the ability of the shifting standards model to detect and measure stereotyping at the team level. This is the first research on stereotyping and team performance to utilize shifting standards to measure team level effects. Hence, this research was able to provide insight into the multi-dimensionality of diversity's effect on team performance, and suggests that the nature of the specific tasks faced by teams might have an impact on the ways stereotyping manifests in the team's performance. The nature of the interdependence of a task may have had a moderating effect on the impact stereotyping has on the team's performance in that task, and this opens the door for further investigation of high performance teams.

This research complicates the claims made about diversity and team performance and challenges future researchers to look more closely at several

moderating variables. It is not just that diversity makes groups more effective or, on the other hand, it is not just that diversity activates status-based classifications that can undermine the benefits of diversity. The present research and methodology suggests that there are other factors at play, which are important to examine in order to truly understand stereotyping in the context of teams.

APPENDIX A

SANDHURST COMPETITION EVALUATION ELEMENTS

Penalties are assigned based on a specific list of “run rules”. The following are generic rules and penalties:

Series	Penalty	Time penalty
(a)	(b)	(c)
1	Incorrect Team Composition	30min (per person) Disqualification (no female squad member)
2	Tampering with Check-Points	Disqualification
3	Loss of Sensitive Item (Weapon or Radio)	Disqualification
4	Failure to meet Cut-Off time for the course	Disqualification
5	Failure to complete course with all personnel	Disqualification
6	Lost/ illegible Passport	60min
7	Failure to visit any site or attempt a task	60min (each occasion)
8	Unlawful Assistance	30min (each occasion)
9	Unsportsmanlike Conduct	30min (each occasion)
10	Exceeding 50m separation all members of the squad (all must remain in a 50m ‘bubble’)	30min (each occasion)
11	Rejected Challenge	20min (each occasion)
12	Missing or unserviceable equipment	30sec (each occasion)
13	Misuse of sensitive items	30sec (each occasion)
14	Incorrect helmet marking at the Start Point	30sec (per helmet)
15	Reporting to the Start Point late	The Squad’s time will start and continue to run

APPENDIX B**TEAMS RESEARCH SURVEY****Ratings for Leader Competencies**

Please rate each of your Sandhurst teammates and yourself with respect to the Leader Action called "Communicating"

Please rate each of your Sandhurst teammates and yourself with respect to the Leader Action called "Decision Making"

Please rate each of your Sandhurst teammates and yourself with respect to the Leader Action called "Motivating"

Please rate each of your Sandhurst teammates and yourself with respect to the Leader Action called "Planning"

Please rate each of your Sandhurst teammates and yourself with respect to the Leader Action called "Executing"

Please rate each of your Sandhurst teammates and yourself with respect to the Leader Action called "Assessing"

Please rate each of your Sandhurst teammates and yourself with respect to the Leader Action called "Developing"

Please rate each of your Sandhurst teammates and yourself with respect to the Leader Action called "Building"

Please rate each of your Sandhurst teammates and yourself with respect to the Leader Action called "Learning"

APPENDIX B

TEAMS RESEARCH SURVEY

Ratings for Overall Effectiveness as a Leader

Please rate your Sandhurst teammates and yourself with respect to your overall effectiveness as a leader.

Rankings for Leader Competencies

Please rank order your teammates including yourself in the Leader Action "Communicating". Place at #1, the teammate who seems to do this best, and so on through #11, where you will place the teammate who you believe does this least well in the group.

Please rank order your teammates including yourself in the Leader Action "Decision Making". Place at #1, the teammate who seems to do this best, and so on through #11, where you will place the teammate who you believe does this least well in the group.

Please rank order your teammates including yourself in the Leader Action "Motivating". Place at #1, the teammate who seems to do this best, and so on through #11, where you will place the teammate who you believe does this least well in the group.

Please rank order your teammates including yourself in the Leader Action "Planning". Place at #1, the teammate who seems to do this best, and so on through #11, where you will place the teammate who you believe does this least well in the group.

APPENDIX B**TEAMS RESEARCH SURVEY**

Please rank order your teammates including yourself in the Leader Action "Executing". Place at #1, the teammate who seems to do this best, and so on through #11, where you will place the teammate who you believe does this least well in the group.

Please rank order your teammates including yourself in the Leader Action "Assessing". Place at #1, the teammate who seems to do this best, and so on through #11, where you will place the teammate who you believe does this least well in the group.

Please rank order your teammates including yourself in the Leader Action "Developing". Place at #1, the teammate who seems to do this best, and so on through #11, where you will place the teammate who you believe does this least well in the group.

Please rank order your teammates including yourself in the Leader Action "Building". Place at #1, the teammate who seems to do this best, and so on through #11, where you will place the teammate who you believe does this least well in the group.

Please rank order your teammates including yourself in the Leader Action "Learning". Place at #1, the teammate who seems to do this best, and so on through #11, where you will place the teammate who you believe does this least well in the group.

APPENDIX B**TEAMS RESEARCH SURVEY****Rankings for Overall Effectiveness as a Leader**

Please rank your Sandhurst teammates and yourself with respect to your overall effectiveness as a leader.

References

- Addison, R., & Haig, C. (2012). *Performance Architecture Care and Feeding of a High-Performance Team* (pp. 1–5).
- Alise, M. A., & Teddlie, C. (2010). A Continuation of the Paradigm Wars? Prevalence Rates of Methodological Approaches Across the Social/Behavioral Sciences. *Journal of Mixed Methods Research*, 4(2), 103–126. Retrieved from <http://onlinelibrary.wiley.com/doi/10.1002/cbdv.200490137/abstract>
- Allport, G. W. (1954). *The nature of prejudice*. Reading, MA: Addison- Wesley.
- Ancona, D. G., & Caldwell, D. F. (1992). Demography and Design : Predictors of New Product Team Performance. *Organization Science*, 3(3), 321–341.
- Austin, J. T., Villanova, P., Katzell, R., Humphreys, L. G., Ryan, A., Hene-, R., Boyle, K., et al. (1992). The Criterion Problem : 1917-1992. *Journal of Applied Psychology*, 77(6), 836–874.
- Bell, S. T. (2007). Deep-level composition variables as predictors of team performance: a meta-analysis. *The Journal of applied psychology*, 92(3), 595–615. doi:10.1037/0021-9010.92.3.595
- Berger, J., Cohen, B. P., & Zelditch, M. (1972). Status Characteristics and Social Interaction. *American Sociological Review*, 37(3), 241–255.
- Berger, J., Norman, R. Z., Balkwell, J. W., & Smith, R. F. (1992). Status Inconsistency in Task Situations. *American Sociological Review*, 57(6), 843–855.
- Berger, J., Ridgeway, C. L., Fisek, M. H., & Norman, R. Z. (1998). The Legitimation and Delegitimation of Power and Prestige Orders. *American Sociological Review*, 63(3), 379. doi:10.2307/2657555
- Biernat, M., & Manis, M. (1994a). Shifting standards and stereotype-based judgments. *Journal of personality and social psychology*, 66(1), 5–20. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8126651>
- Biernat, M., & Manis, M. (1994b). Shifting standards and stereotype-based judgments. *Journal of personality and social psychology*, 66(1), 5–20. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8126651>
- Biernat, M., Crandall, C. S., Young, L. V., Kobrynowicz, D., & Halpin, S. M. (1998). All that you can be: stereotyping of self and others in a military context. *Journal of personality and social psychology*, 75(2), 301–17. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9731310>

- Biernat, M., Tocci, M. J., & Williams, J. C. (2011). The Language of Performance Evaluations: Gender-Based Shifts in Content and Consistency of Judgment. *Social Psychological and Personality Science*, 3(2), 186–192. doi:10.1177/1948550611415693
- Biernat, Monica, & Manis, M. (1994). Shifting standards and stereotype-based judgments. *Journal of Personality and Social Psychology*, 66(1), 5–20. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8126651>
- Biernat, Monica, Manis, M., & Nelson, T. E. (1991). Stereotypes and standards of judgment. *Journal of Personality and Social Psychology*, 60(4), 485–499. doi:10.1037/0022-3514.60.4.485
- Biernat, Monica, & Vescio, T. K. (1993). Categorization and Stereotyping: Effects of Group Context on Memory and Social Judgment. *Journal of Experimental Social Psychology*, 29, 166–202.
- Biernat, M., Vescio, T., Manis, M. (1998). Judging and behaving toward members of stereotyped groups: a shifting standards perspective. *Intergroup Cognition and Intergroup Behavior* (pp. 151–175).
- Bodenhausen, G. V. (2010). Agenda Diversity in the person , diversity in the group : Challenges of identity complexity for social perception and social interaction. *European Journal of Social Psychology*, 16(June 2009), 1–16. doi:10.1002/ejsp
- Boldry, J., Wood, W., & Kashy, D. a. (2001). Gender Stereotypes and the Evaluation of Men and Women in Military Training. *Journal of Social Issues*, 57(4), 689–705. doi:10.1111/0022-4537.00236
- Brewer, M. B., & Pierce, K. P. (2005). Social identity complexity and outgroup tolerance. *Personality & social psychology bulletin*, 31(3), 428–37. doi:10.1177/0146167204271710
- Bryman, A. (2011). Mission accomplished?: Research methods in the first five years of Leadership. *Leadership*, 7(1), 73–83. doi:10.1177/1742715010386864
- Corsaro, D., Cantù, C., & Tunisini, A. (2012). Actors' Heterogeneity in Innovation Networks. *Industrial Marketing Management, In Press(In Press)*, 1–10. doi:10.1016/j.indmarman.2012.06.005
- Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109(3), 573–598. doi:10.1037//0033-295X.109.3.573
- Edmondson, A. C., & Mcmanus, S. E. (2007). Methodological Fit in Management Field Research. *Academy of Management Review*, 32(4), 1155–1179.

- Foschi, M. (1992a). Gender and double standards for competence. In C. L. Ridgeway (Ed.), *Gender Interactions and Inequality* (pp. 181–207). New York City: Springer-Vollas.
- Foschi, M. (1992b). Status Characteristics, standards, and attributions. In B. Berger, Joseph, Zelditch, Morris, Anderson (Ed.), *Sociological Theories in Progress: New Formulations* (pp. 50–71). Newbury Park, CA: Sage Publications.
- Foschi, M. (2000). Double Standards for Competence: Theory and Research. *Annual Review of Sociology*, 26(1), 21–42. doi:10.1146/annurev.soc.26.1.21
- Foschi, M., Lai, L., & Sigerson, K. (1994). Gender and double standards in the assessment of job applicants. *Social Psychology Quarterly*, 57(4), 326–339. Retrieved from <http://www.jstor.org/stable/10.2307/2787159>
- Foschi, M., Warriner, G. K., & Hart, S. D. (1985). Standards, Expectations, and Interpersonal Influence. *Social Psychology Quarterly*, 48(2), 108–117.
- Garcia-Prieto, P., Bellard, E., & Schneider, S. C. (2003). Experiencing Diversity, Conflict, and Emotions in Teams. *Applied Psychology*, 52(3), 413–440. doi:10.1111/1464-0597.00142
- Gerber, G. L. (2011). Police Dyads : Status in Same-Gender and Mixed-Gender Effects on Personality Attributions *. *Social Psychology*, 59(4), 350–363.
- Hackman, J. R. (2002). A Real Team. In J. R. Hackman (Ed.), *Leading teams: Setting the stage for great performances* (pp. 36–60). Boston, MA: Harvard Business School Press.
- Hackman, J. R., & Wageman, R. (2001). *The Team Diagnostic Survey: An instrument for assessing the structure, dynamics, and leadership of work teams*. Boston, MA: Hay McBer.
- Hackman, J.R., and Katz, N. (2010). Group Behavior and Performance. In L. G. Fiske, Susan T., Gilbert, Daniel T. (Ed.), *Handbook of Social Psychology, Volume 2* (5th ed., pp. 1208–1251). Hoboken, N.J.: John Wiley & Sons, Inc.
- Hastings, D. W. (2006). Drowning in Inequalities: Swimming and Social Justice. *Journal of Black Studies*, 36(6), 894–917. doi:10.1177/0021934705283903
- Herring, C. (2009). Does Diversity Pay?: Race, Gender, and the Business Case for Diversity. *American Sociological Review*, 74(2), 208–224. doi:10.1177/000312240907400203
- Hewstone, M., Rubin, M., & Willis, H. (2002). Intergroup Bias, 53, 575–604.

- Hoffman, L., & Maier, N. (1961). Quality and acceptance of problem solutions by members of homogeneous and heterogeneous groups. *Journal of Abnormal and Social Psychology*, 62, 401–407.
- Horwitz, S. K., & Horwitz, I. B. (2007). The Effects of Team Diversity on Team Outcomes: A Meta-Analytic Review of Team Demography. *Journal of Management*, 33(6), 987–1015. doi:10.1177/0149206307308587
- Jackson, S. E., Joshi, A., & Erhardt, N. L. (2003). Recent Research on Team and Organizational Diversity: SWOT Analysis and Implications. *Journal of Management*, 29(6), 801–830. doi:10.1016/S0149-2063
- Judd, C. M., & Kenny, D. A. (2010). Data Analysis in Social Psychology: Recent and Recurring Issues. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of Social Psychology, Volume 1* (5th ed., pp. 115–142). Hoboken, N.J.: John Wiley & Sons, Inc.
- Kochan, T., Bezrukova, K., Ely, R. J., Jackson, S. E., Joshi, A., Jehn, K., Leonard, J., et al. (2003). The effects of diversity on business performance: Report of the diversity research network. *Human Resource Management*, 42(1), 3–21. doi:10.1002/hrm.10061
- Lau, D. C., & Murnighan, J. K. (1998). Demographic Diversity and Faultlines: The Compositional Dynamics of Organizational Groups. *The Academy of Management Review*, 23(2), 325–340. doi:10.2307/259377
- Lau, D. C., & Murnighan, J. K. (2005). Interactions Within Groups and Subgroups: the Effects of Demographic Faultlines. *Academy of Management Journal*, 48(4), 645–659. doi:10.5465/AMJ.2005.17843943
- Mannix, E., & Neale, M. A. (2005). What Differences Make a Difference?: The Promise and Reality of Diverse Teams in Organizations. *Psychological Science in the Public Interest*, 6(2), 31–55. doi:10.1111/j.1529-1006.2005.00022.x
- Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A TEMPORALLY AND BASED FRAMEWORK AND TAXONOMY OF TEAM PROCESSES. *Academy of Management Review*, 26(3), 356–376. Retrieved from <http://www.jstor.org/stable/259182> .
- Mathieu, J., Maynard, M. T., Rapp, T., & Gilson, L. (2008). Team Effectiveness 1997-2007: A Review of Recent Advancements and a Glimpse Into the Future. *Journal of Management*, 34(3), 410–476. doi:10.1177/0149206308316061
- Metrick, J., & Cowan, G. (1996). Gender Stereotypes and Predictions of Performance : Women in Air Combat. *Journal of Social Behavior and Personality*, 11(5), 105–120.

- Newheiser, A.-K., & Dovidio, J. F. (2012). Individual differences and intergroup bias: Divergent dynamics associated with prejudice and stereotyping. *Personality and Individual Differences, 53*(1), 70–74. doi:10.1016/j.paid.2012.02.024
- Pasmore, W. A. (1988). *Designing effective organizations : the sociotechnical systems perspective* (p. 200). New York City: Wiley. Retrieved from <http://hollis.harvard.edu/?itemid=|library/m/aleph|001571582>
- Pearsall, M. J., Ellis, A. P. J., & Evans, J. M. (2008). Unlocking the effects of gender faultlines on team creativity: is activation the key? *The Journal of applied psychology, 93*(1), 225–34. doi:10.1037/0021-9010.93.1.225
- Pentland, A. (2012). The New Science of Building Great Teams. *Harvard Business Review, (April)*, 60–70.
- Pham, T., Metoyer, R., Bezrukova, K., & Spell, C. (2012). “Show Me the Cracks in Our Teams”: Visual Representations of Demographic Diversity Faultlines. *cse.ohio-state.edu*, 2–3. Retrieved from <http://www.cse.ohio-state.edu/~raghu/teaching/CSE5544/Visweek2012/infovis/posters/pham.pdf>
- Pittinsky, T. L. (2005). Faculty Research Working Papers Series.
- Pittinsky, T. L., Rosenthal, S. A., & Montoya, R. M. (2011). Measuring positive attitudes toward outgroups: Development and validation of the Allophilia Scale.
- Polzer, J. T., Milton, L. P., & Swann, W. B. (2011). Capitalizing on Diversity: Interpersonal Congruence in Small Work Groups. *Administrative Science Quarterly, 47*(2), 296–324.
- Ridgeway, C. L., & Berger, J. (1986). Expectations, legitimation, and dominance behavior in task groups. *American Sociological Review, 51*(5), 603–617.
- Roberge, M.-É., & Van Dick, R. (2010). Recognizing the benefits of diversity: When and how does diversity increase group performance? *Human Resource Management Review, 20*(4), 295–308. doi:10.1016/j.hrmr.2009.09.002
- Rosenthal, H. E. S., Crisp, R. J., & Suen, M. (2007). Improving performance expectancies in stereotypic domains : Task relevance and the reduction of stereotype threat. *European Journal of Social Psychology, 597*(August 2005), 586–597. doi:10.1002/ejsp
- Rothbart, M., & John, O. P. (1985). Social Categorization and Behavioral Episodes: A Cognitive Analysis of the Effects of Intergroup Contact. *Journal of Social Issues, 41*(3), 81–104. doi:10.1111/j.1540-4560.1985.tb01130.x

- Sailes, G. (1991). The myth of Black sports supremacy. *Journal of Black Studies*, 21(4), 480–487. Retrieved from <http://www.jstor.org/stable/10.2307/2784690>
- Shih, M., Pittinsky, T. L., & Ambady, N. (1999). STEREOTYPE SUSCEPTIBILITY : Identity Salience and Shifts in Quantitative Performance. *Identity*, 10(1), 80–84.
- Steele, C. M., & Aronson, J. (1995). Stereotype Threat and the Intellectual Test Performance of African Americans. *Journal of Personality and Social Psychology*, 69(5), 797–811.
- Tajfel, H. (1978). Social categorization, social identity, and social comparison. In H. Tajfel (Ed.), *Differentiation between social groups* (pp. 61–76). London: Academic Press.
- Thatcher, S. M. B., & Patel, P. C. (2011). Demographic faultlines: A meta-analysis of the literature. *The Journal of applied psychology*, 96(6), 1119–1139. doi:10.1037/a0024167
- Thompson, J. D. (1967). *Organizations in action: Social science bases of administrative theory* (p. 192). New York City: McGraw Hill. Retrieved from <http://hollis.harvard.edu/?itemid=|library/m/aleph|000539654>
- Van Dijk, H., Van Engen, M. L., & Van Knippenberg, D. (2012). Defying conventional wisdom: A meta-analytical examination of the differences between demographic and job-related diversity relationships with performance. *Organizational Behavior and Human Decision Processes, In Press*, 1–16. doi:10.1016/j.obhdp.2012.06.003
- Van Knippenberg, D., De Dreu, C. K. W., & Homan, A. C. (2004). Work group diversity and group performance: an integrative model and research agenda. *The Journal of applied psychology*, 89(6), 1008–22. doi:10.1037/0021-9010.89.6.1008
- Wagner, D. G., & Berger, J. (2011). GENDER AND INTERPERSONAL TASK BEHAVIORS : Status Expectation Accounts. *Sociological Perspectives*, 40(1), 1–32.
- Wells, J. E., & Aicher, T. J. (2013). Follow the Leader: A Relational Demography, Similarity Attraction, and Social Identity Theory of Leadership Approach of a Team's Performance. *Gender Issues*. doi:10.1007/s12147-013-9112-8
- Williams, K. Y., & O'Reilly, C. A. (1998). Demography and Diversity in Organizations: A Review of 40 Years of Research. *Research In Organizational Behavior*, 20(20), 77–140.

- Yammarino, F., Dionne, S. D., Ukchun, J., & Dansereau, F. (2005). Leadership and levels of analysis: A state-of-the-science review. *The Leadership Quarterly*, 16(6), 879–919. doi:10.1016/j.leaqua.2005.09.002
- Yzerbyt, V., & Demoulin, S. (2010). Intergroup Relations. In & G. L. S. T. Fiske, D. T. Gilbert (Ed.), *Handbook of Social Psychology, Volume 2* (5th ed., pp. 1024–1083). Hoboken, N.J: John Wiley & Sons, Inc.

Vita

Lissa V. Young

1982-1986	United States Military Academy West Point, NY	B.S. May 1986
1986-2002	United States Army Officer & Rotary Wing Aviator	
1994-1996	Graduate School of Arts & Sciences University of Kansas Lawrence, KS	M.A. May 1996
1996-1999	Instructor, Organizational Leadership United States Military Academy West Point, NY	
2002-2007	Senior Manager Raytheon Company Waltham, MA	
2007-2013	Doctoral Candidate Graduate School of Education Harvard University Cambridge, MA	
2007-2009	Graduate School of Education Harvard University Cambridge, MA	M.Ed. May 2009
2007-2013	Research and Teaching Assistant Graduate School of Education and Graduate School of Arts and Sciences Harvard University Cambridge, MA	
2010-2013	Graduate School of Arts and Sciences Studies in Women, Gender and Sexuality Harvard University Cambridge, MA	Certificate