

Self-vs.-Teammate Assessment of Leadership Competence: The Effects of Gender, Leadership Self-Efficacy, and Motivation to Lead

Dr. David M. Rosch

Assistant Professor, Agricultural Education
University of Illinois at Urbana-Champaign
Urbana, IL 61801
dmrosch@illinois.edu

Daniel A. Collier

Graduate Student, Education Policy, Organization and Leadership
University of Illinois at Urbana-Champaign
Champaign, IL 61820
dcol3@illinois.edu

Sarah M. Zehr

Assistant Dean and Director, Engineering Career Services
University of Illinois at Urbana-Champaign
Urbana, IL 61801
szehr@illinois.edu

Abstract

A sample (N=81) of undergraduates participating in a semester-long team-project engineering course completed assessments of their leadership competence, motivation to lead, and leadership self-efficacy, as well as the leadership competence of their peers who served within their durable teams. Results indicated that peers scored students lower than students scored themselves; that males deflated the transactional leadership scores of the female peers they assessed; and that the strongest individual predictor of teammate-assigned scores was a student's affective-identity motivation to lead (i.e. the degree to which they considered themselves a natural leader). Leadership self-efficacy failed to significantly predict teammate scores.

Introduction

Higher education has been challenged to develop more transformative leaders “who can devise more effective solutions to some of our most pressing...issues” (Astin & Astin, 2000, p.6). In response, educators have begun to recognize the critical function of leadership development (Council for the Advancement of Standards in Higher Education [CAS], 2011), significantly expanding formal leadership education programs (Sessa, Matos, & Hopkins, 2009). While many organizations and employers expect students to develop soft-skills such as competence in communication, skill in influencing others toward common goals, and the ability to work within team structures (National Association of Colleges and Employers [NACE], 2012) educators continue to lack proficiency in evaluating the leadership effectiveness of students who participate in their programs (Dugan, 2011; Rosch & Schwartz, 2009). A majority of undergraduate programs limit assessment to self-report data (Rohs, 2002) that rests on students’ own ability to gauge their learning and growth. Participants lacking a depth of self-awareness, however, may inflate their true competence, (Mayo, Kakarika, Pastor, & Brutus, 2012) leaving the effectiveness and impact of the overall assessment program in question. Our research focused on the use of multi-rater feedback within a team-based leadership course. We examined how students’ self-reports of their leadership competence differed systemically from that of the peers who served with them on semester-long project teams, specifically analyzing differences with respect to gender and a variety of leadership attitudes such as one’s motivation to lead (Chan & Drasgow, 2001) and leadership self-efficacy. As the multi-rater systems might be inadvisable or even impossible in the context of many leadership education programs, an understanding of systemic differences

between self-reported and peer-reported leadership competency can aid in assessing the effectiveness and limitations of self-report data in evaluating growth in individual leader development.

A Multi-rater Approach

The use of multi-rater assessment of individual competence is widely popular in evaluator circles (Nowack & Mashihi, 2012; Toegel & Conger, 2003) and is increasingly used for leadership evaluation (Asumeng, 2013; Drew, 2009). Multi-rater systems assume that the inclusion of many sources of input will paint a clearer picture of competence due to an ability to compare self-perception to exterior perceptions (Carlson, 1998). However, research on the effectiveness of multi-rater feedback has been inconsistent (Toney, 1996; Nijhof & Jager, 1999; Azzam & Riggio, 2003; Shipper, 2004; Asumeng, 2013), partially due to inflated ratings & leniency on the part of the evaluators (Farh, Cannella, & Bedeian, 1991; Roch & McNall, 2007; Hensel, Meijers, Leeden, & Kessels, 2010).

Past research has shown inconsistent success in using multi-rater assessments as a tool for accurate performance appraisal (Asumeng, 2013), leading some to question the validity of the method (Nowack & Mashihi, 2012). However, Toegel and Conger (2003) report that these issues may be limited to performance appraisal within hierarchical professional environments. Mero, Guidice, & Bownlee (2007) indicate relationship and rank statuses, as well as varying levels of accountability, can lead to less accurate assessments. Without a common system of evaluation (e.g. where items are not “open to interpretation”) and long-standing interdependent relationships between evaluator and those being evaluated, multi-rater systems of assessment may result in bias or generalized

perceptions rather than feedback on specific competencies (Rosch, Anderson, & Jordan, 2012; Toegel & Conger, 2003). Often, peers with little motivation to honestly rate or specific knowledge of a person's particular skills rate them higher than the person rates him or herself (Rosch, Anderson & Jordan, 2012). Although these issues remain prominent in poorly designed or misused evaluations, such factors can be reduced by providing specific guidelines on context for which evaluation should occur and creating survey items that leave little open to interpretation (Theron & Roodt, 1999).

Both issues may be minimized within teams that interact consistently, interdependently, and only in the context of their work together – such as within a semester-long academic course. Indeed, past studies have found benefits within a multi-rater approach in educational and developmental environments (Drew, 2009; Ghorpade, 2000) that result in durable personal development in targeted skill areas (Toegel & Conger, 2003).

Moreover, multi-rater methodology in assessing competence may be a significant tool in understanding the effects of gender on peer evaluation, given the role that gender has played in how a student practices and engages in leadership processes. Previous research indicates men and women are evaluated differently (Eagly & Carli, 2003; Ibarra & Obodaru, 2009; Manning & Robertson, 2010; Ely, Ibarra, & Kolb, 2011), often following gender stereotypes (Eagly & Carli, 2003; Manning & Robertson, 2010). Indeed, women often score lower than men on traditional behaviors associated with leadership (Eagly, Makhijani, & Klonsky, 1992; Eagly & Carli, 2003; Ely et al., 2011), but outdistance their male peers in relational aspects, including: (1) emotional intelligence, (2) rewarding behaviors and use of feedback, and (3) use of effective team-building behaviors (Ibarra & Obodaru, 2009). Additionally, meta-data analysis of similar studies suggests that as the

proportion of male evaluators increased, women were rated as less effective leaders (Bowen, Swim, Jacobs, 2000; Eagly & Carli, 2003). Given these findings, our research sought, in part, to examine how participants' gender moderated their evaluation of their peers' leadership effectiveness.

Leadership Self-efficacy and Motivation to Lead

Leadership effectiveness has traditionally focused on strict measurements of leadership skill and behaviors (Waldman, Galvin, & Walumbwa, 2013), but recent calls have been made for expansion of measures to include a more complex combination of both leadership skills and attitudes, such as leadership self-efficacy (LSE) (Murphy, 2002; Dugan, 2011) and motivation to lead (MTL) (Avolio, 2007, Amit & Bar-Lev, 2013; Waldman, et al., 2013). LSE describes students' internal perception of their ability to engage in leadership processes (Murphy, 2002). MTL measures the "direction, intensity, and persistence" (Chan & Drasgow, 2001, p. 482) of engagement in the leadership process, and is divided into three subscales: Affective Identity (AI), Social Normative (SN), and Non-Calculative (NC). AI measures the extent to which people envision themselves as leaders; SN measures the extent to which a person seeks leadership due to the responsibility one feels toward a group; and NC measures the extent to which leaders avoid cost-benefit analysis of personal benefits (Chan & Drasgow, 2001).

Within the past decade, researchers have begun to incorporate rigorous measurements for both motivation to lead (MTL) (Arvey, Zhang, Avolio, & Krueger, 2007; Amit & Bar-Lev, 2013; Waldman, et al., 2013) and leadership self-efficacy (LSE) (Avolio, 2007; Dugan, 2011). LSE has been shown to predict increased interest in leadership positions and higher ratings of leader performance by group members (Hannah, Avolio, Luthans,

& Harms, 2008). MTL has shown to be a significant predictor of leadership role occupancy in professional organizations (Arvey et al., 2007), the development of self-reported leadership expertise (Lord & Hall, 2005), and in success of the group being led (Kark & Van Dijk, 2007). Significant to this research study, Shertzer & Schuh (2004) showed that students who displayed a greater degree of motivation to lead were evaluated more positively as leaders by the peers who served with them within student clubs and organizations.

These findings suggest the important role that internal attitudes play in the determination of a leader's behaviors as well as the group's evaluation of the leader. As LSE and MTL become more prevalent in program evaluation (Waldman, 2013), more research connecting these attitudes to relevant outcomes must be explored to better identify the impact of participation in leadership programs. How do these internal qualities predict the leader evaluation of external peers who regularly work with those leaders?

Research Questions

This study represents an effort to determine the differences between students and the peers that have served on semester-long teams with them in evaluating their leadership competence, including their leadership capacity, leadership self-efficacy, and motivation to lead. Moreover, we sought to determine the degree to which students self-ratings might predict their teammates' peer assessments. Therefore, we posed the following research questions:

1. To what extent do teammate assessments of leader competence differ from students' self-assessments?
2. To what extent do gender differences affect teammate assessments?

3. To what extent can students' self-assessment of leadership competence, self-efficacy, and motivation to lead predict teammate peer assessments of leadership competence, controlling for prior leadership training?

Methods

Sample

This study was conducted at a large, public, research-intensive university in the Midwestern United States. Our sample consisted of undergraduate students enrolled in an introductory elective course within the College of Engineering designed for first-semester freshman entering the College. The course, titled, "Team-based Project Management," was focused on teaching teamwork skills within an outcome-based, goal-oriented professional engineering environment. Instruction was relatively laissez-faire, where students were given minimal strategic direction, provided little formal instruction in teambuilding or relationship management, and allowed to choose their own projects. Stated outcomes for the course focused on developing strategic planning, goal setting, and team communication skills. Students were placed in groups of three to five peers and encouraged to experiment, take risks together, and collaborate with each other. All grades were assigned at the group level and related to the success of the group's project, not their team dynamics. Across three course sections, 81 students fully participated in both pre- and post-course surveys, encompassing most of the enrolled students. Approximately 74% (n=60) of the sample identified as male, while 66% (n=53) identified as Caucasian; 14% (n=11) as Asian American; 15% (n=12) as an international student; 2% (n=2) as Latino(a); and 3% (n=3) did not identify their race. Students also

completed 215 assessments of their teammates, with an average of 2.7 teammate assessments completed per student.

Measures

The survey instrument combined scales associated with measurement of transformational and transactional leadership behaviors, leadership self-efficacy, and motivation to lead. Serving as a proxy for leadership competence, we utilized the Leader Behavior Scale (LBS), a popular 27-item instrument designed to measure behaviors that align to either transformational or transactional values. The LBS was adapted from a larger measurement instrument designed to assess broad-based organizational citizenship (Podsakoff, MacKenzie, Moorman, & Fetter, 1990). An example item measuring transformational behavior was, “I help other group members develop a team attitude and spirit among ourselves.” An example of an item measuring transactional leadership was, “I always give positive feedback when other group members perform well.” Item responses include a 5-point Likert scale ranging from “Strongly Agree” to “Strongly Disagree.” The LBS is a widely popular method of measuring transformational and transactional leadership (Yukl, 2010), chosen in this study specifically due to its’ connection with the measure of organizational citizenship (Podsakoff, MacKenzie, Moorman, & Fetter, 1990), an important factor for team success in modern organizations (Podsakoff, MacKenzie, Paine, & Bachrach, 2000). Internal reliability for the LBS within this study was high – Cronbach’s alpha measured at .83 for the transformational scale, and .89 for the transactional scale.

To measure leadership self-efficacy, we used the Self-Efficacy for Leadership Scale (SEL) (Murphy, 1992), an 8-item scale measuring a person’s confidence in engaging in

leadership behaviors. An example item within this scale is, “I know how to encourage good group performance.” Item responses include a 5-point Likert-scale ranging from “Strongly Disagree” to “Strongly Agree.” Internal reliability from previous research is good (Murphy & Ensher, 1999), and the scale has been shown to possess convergent and discriminant validity with measure of self-esteem and leadership experiences (Hoyt, 2005). Internal reliability within this study was high, measured at .82

Students’ motivation to lead was measured using Chan & Drasgow’s (2001) Motivation to Lead (MTL) scale. The MTL scale includes 27 items divided equally across three subscales: Affective Identity (AI) Motivation, Social Normative (SN) Motivation, and Non-calculative (NC) Motivation. The AI scale measures the degree to which a person is personally drawn to leadership roles and includes items like, “Most of the time, I prefer being a leader rather than a follower when working in a group.” The SN scale determines the degree to which a person leads due to a sense of duty or responsibility to others and includes items like, “People should volunteer to lead rather than wait for others to ask or vote for them.” The NC scale measures the degree that a person avoids rationally calculating the individual costs and benefits of holding a leadership position and includes items like, “I never expect to get more privileges if I agree to lead a group.” Responses fell within a 5-point Likert scale ranging from “Strongly Disagree” to “Strongly Agree.” Internal reliability from previous research has found to be acceptable, ranging from .65 to .91 (Chan & Drasgow, 2001), and ranged in this study across the 3 scales from .62 to .84.

Lastly, to control for the effects of prior leadership experience, the survey included an item asking students to rate their degree of experience in participating in prior leadership development training on 5-item Likert scale ranging from “consistently” to “never.”

Data Collection

The instructor of the course allocated classroom time for the researchers to distribute and collect the survey within the first week of class and again on the last day of class meetings. Students who participated also completed “teammate assessments” of each of their team members’ transformational and transactional leadership competencies, using an adapted version of the LBS that substituted “This teammate...” for “I...” in all survey items. All participants were asked to be as honest as possible with both their own assessments and their assessments of their teammates, and were told that their responses would remain confidential.

Data Analysis

To determine the differences between self and teammate assessments of leadership behaviors, we first created means for each student’s teammate scale scores. We then analyzed the means and dispersion of their and their teammates’ transformational and transactional scale scores by conducting paired samples t-tests and calculating effect sizes (Cohen, 1987) of means differences. The effect of gender on teammate evaluations were calculated by assigning students’ teammate scores to one of three groups: 1) a male-evaluating-female score; 2) a female-evaluating-male score; and 3) a same-gender evaluation. We examined the means and dispersion of each group and conducted a one-way ANOVA for both transformational and transactional leadership scores to determine the significance of the differences. To measure the strength of self-report predictors on teammate assessments of transformational and transactional leadership competence while controlling for prior leadership training, we conducted a two-step hierarchical multiple regression for both observer leadership competency scores using first pre-test scores and

then post-test scores, entering gender and prior leadership participation in the first step, and students' transformational and transactional competency, leadership self-efficacy, and motivation to lead in the second step.

Results

Self vs. Teammate Score Differences

The means and dispersion of each variable can be found in Table 1. Students rated their transactional leadership scores highest and their affective-identity-related motivation to lead lowest. Teammate assessments of students' transformational and transactional leadership competence were lower than how students rated themselves.

Table 1

Self and Teammate Leadership Scale Means and Dispersion (n=81)

Scale	Pre-test			Post-test		
	N	μ	SD	N	μ	SD
Transformational Leadership – Self (FormS)	81	3.77	.31	81	3.89	.39
Transactional Leadership (Self) (ActS)	81	3.96	.51	81	4.06	.55
Self-Efficacy for Leadership (SEL)	81	3.78	.44	81	3.86	.53
Motivation to Lead; Affective-Identity (MTL_AI)	81	3.49	.67	81	3.44	.71
Motivation to Lead; Social-Normative (MTL_SN)	81	3.31	.31	81	3.78	.49
Motivation to Lead; Non-Calculative (MTL_NC)	81	3.72	.42	81	3.82	.57
Transformational Leadership – Teammate (FormT)				215	3.68	.55
Transactional Leadership – Teammate (ActT)				215	3.74	.56

Paired sample t-tests conducted on pre-test and corresponding post-test scores yielded significant results ($p < .05$) for FormS ($p = .002$) and MTL-SN ($p < .001$), indicating measurable score increases in transformational behaviors and social-normative motivation to lead over the course of the semester. The effect sizes for FormS was moderate ($d = .34$), and for MTL-SN was large ($d = 1.14$), indicating that over the course of 15 weeks, students scored themselves moderately higher as a transformational leader, and substantially higher in their motivation to lead based on their sense of responsibility to their team members.

Teammate assessments for each student were averaged to create a mean teammate score; then a paired-samples t-test was conducted using student post-test scores and teammate mean scores for both transformational and transactional leadership scales. The results for both were significant; $t(81) = 3.09$, $p = .003$ and $t(81) = 3.95$, $p < .001$, respectively. The effect size of each was moderate: $.39$ for the transformational leadership score difference and $.54$ for the score difference in transactional leadership. These results indicate that teammates scored their team members moderately lower than team members score themselves.

Gender Differences in Peer Evaluations

To determine the effect of gender on how students assess the leadership competence of their peers, teammate assessment scores were placed into three groups: 1) male-evaluating-female; 2) female-evaluating male; and 3) same-gender. Table 2 contains mean scores on pre- and post-test transactional and transformational leadership scales analyzed by gender. We compared self-reported post-test scores with assessments of teammates of the opposite gender. A significant result emerged in the way women were

evaluated for transactional leadership behaviors ($p < .05$); men evaluated their female counterparts lower than women evaluated themselves. No other significant score differences emerged related to gender.

Table 2

Leadership Competence Scores by Gender Differences

Scale	Group	N	μ	SD	T	Df	P
Transformational	Female-Post-test	14	4.04	.38			
	Male-Evaluating-Female- Post	33	3.95	.42	0.49	45	.49
	Male-Post-test	77	3.83	.41			
	Female-Evaluating-Male- Post	34	3.69	.85	1.17	109	.24
Transactional	Female-Post-test	14	4.35	.53			
	Male-Evaluating-Female- Post	33	3.99	.54	2.10	45	.04
	Male-Post-test	77	3.95	.53			
	Female-Evaluating-Male- Post	34	3.74	.99	1.45	109	.14

Individual Attitude Predictors of Teammate Assessment of Competence

The predictive strength of each attitudinal variable was calculated by conducting two-step multiple regressions using FormT and ActT as the dependent variables, first using student pre-test responses and then separately using their post-test responses. Neither regression

analysis using pre-test data yielded significant results; no variable on students' pre-test assessment predicted their teammates' assessment of their leadership competency. The results for both post-test regressions can be found in Tables 3 and 4 respectively. Gender emerged as a marginal predictor ($p < .10$) as teammate-reported transformational leadership score when student self-reported leadership competencies are not controlled for, while prior leadership training did not predict either assessment score. Controlling for all variables, the only significant predictor ($p < .05$) of teammate transformational leadership score was Affective-Identity Motivation to Lead, while self-reported Transformational Leadership emerged as a marginal predictor ($p < .10$). Affective-Identity Motivation to Lead was also a significant predictor of teammates' scoring of students' Transactional Leadership, as well as self-reported Transformational Leadership, while self-reported Self-Efficacy for Leadership emerged as a marginal negative predictor ($p < .10$).

Table 3

*Self-reported Leadership Predictors of Teammate Transformational Leadership Score**

	B	SE B	B	P
Step One				
Gender	.26	.15	.21	.08
Prior Training	.01	.05	.01	.91
Step Two				
Gender	.18	.15	.15	.23
Prior Training	.06	.05	.14	.26
FormS	.43	.23	.35	.06
ActS	.05	.13	.06	.71
MTL_AI	.30	.12	.41	.02
MTL_SN	-.07	.16	-.07	.67
MTL_NC	.08	.12	.10	.68
SEL	-.29	.19	-.29	.13

* DV = FormT

Table 4

*Self-reported Leadership Predictors of Teammate Transactional Leadership Score**

	B	SE B	β	P
Step One				
Gender	.27	.16	.20	.10
Prior Training	.01	.06	.01	.91
Step Two				
Gender	.27	.17	.20	.10
Prior Training	.06	.06	.13	.29
FormS	.57	.26	.41	.02
ActS	-.09	.15	-.09	.57
MTL_AI	.28	.14	.33	.05
MTL_SN	.10	.18	.09	.57
MTL_NC	.02	.13	.02	.87
SEL	-.38	.21	-.35	.08

* DV = ActT

Discussion and Implications

Our research was designed to determine the degree to which students differed from their peers regarding perceptions of their leadership competency, how students' gender might affect these perceptions, and how their individual leadership attitudes and beliefs might predict teammates' perceptions of their competency. Our results showed that students' own perceptions of their competency outstripped that of their teammates' perceptions to a moderate extent (a .39 effect size regarding transformational leadership and .54 regarding

transactional leadership). These findings seem to contradict earlier research in multi-rating assessments, which suggest that observers often are more lenient and accepting in assessing team members than those individuals are in assessing themselves (Farh et al., 1991; Roch & McNall, 2007), even when those observers are fellow students who know those individuals well (Rosch, et al., 2012). Our unique finding may result from the interdependent and non-hierarchical nature of the classroom team setting. In addition, since the assessment was not designed to correlate with performance outcomes it is possible we obtained a less restricted, and therefore more thorough, view of how team-members viewed each other's leadership competence (Drew, 2009; Ghorpade, 2000). These findings suggest that educators who wish aid in the development of leadership competence might include student teams that interdependently act in project groups over the course of a semester; end-of-semester feedback from teammates, averaged for confidentiality, might provide the information needed for emerging leaders who lack the requisite self-awareness to recognize the need to make improvements on their own. Men scored women higher than women scored men on both scales of leadership, which was consistent with how men and women scores themselves. However, scores from male teammates were particularly depressed in males' evaluation of their female teammates' transactional leadership behaviors. These findings corroborate past research that showed that women are received as acceptable relationship-oriented team leaders but revealed a female disadvantage in how others perceive them as task-oriented leaders (Eagly et al., 1993; Eagly & Carli, 2003; Ely et al., 2011). Without a direct comparison in this way, such differences would be easy to miss, as females' absolute and relative scores were higher than their male peers. However, these findings should be considered exploratory

in this area, as men outnumbered women in the course by two to one and the cell size for women was relatively low for acceptable statistical power ($n=19$).

The strongest individual predictor of teammate assessment of leadership competency across both scales was a student's affective-identity motivation to lead, which served as an even stronger predictor than a student's own assessment of their competency. This finding suggests that the degree to which individuals consider themselves leaders of their peers leaves a powerful impression on those peers, and in some ways is even more powerful than behavior. Our findings support Shertzer & Schuh's (2004) claims that undergraduates believe leaders attain success due to internally driven motivation, thus creating more opportunities for themselves to further develop leadership self-efficacy and confidence while gaining additional evidence for others to view them as a leader.

Curiously, self-reported transactional leadership competency did not predict teammate assessments of either transformational or transactional leadership. Similarly, leadership self-efficacy, the confidence that leaders possess to engage in leadership-oriented behaviors, did not emerge as a significant predictor of either style of leadership competency. Even as past research (Murphy, 1992; Dugan, Garland, Jacoby & Gasiorski, 2008) has shown the degree to which leadership self-efficacy can predict leadership behaviors, our findings suggest the complicated relationship between motivation and self-efficacy in a leadership context. While preliminary, these results suggest that peers are more likely to be influenced by a person's generalized belief in themselves as a leader than that person's confidence in engaging in the specific actions of leadership.

Implications

The results of this research study may indicate the significance of a durable and educational context and peer interdependence in the peer assessment of an individual's leadership competency. The students in this study possessed the ability to choose the projects in which they worked, and while they could not choose partners, the process ensured that not only were students placed in non-hierarchical interdependent work environments, they were assured placement on a team of peers who shared a common interest. The environment in which the students worked and conducted their assessments may explain some of the results found within the study. Peer assessment scores were lower than is often seen in multi-rater feedback systems (Alimo-Metcalfe, 1998), suggesting leniency was less of a factor in this study and that the peer assessments might have been more honest. Therefore, multi-rater feedback focused on development and not tied to performance measures might be an effective tool for semester-long teams that work interdependently, a common occurrence in leadership development classrooms. The findings in this study imply that younger women may be making up ground related to younger men in terms of how they are perceived as transformational leaders, given that the sample represented a group of college freshmen. While some men were scored higher by peers than most women, the transformational scores assigned to women were statistically no different than scores assigned to men. Corroborating past findings by Bowen, Swim, Jacobs (2000) and Eagly & Carli (2003), significant gaps could still be seen in how men evaluated the transactional skills of their female counterparts. Even as views may be shifting related to a gender gap in the leadership required for successful work teams, real differences in gender-related perceptions remain. Still, these findings

should be considered within the context of the specialized population of students in the study – first-year engineering students in a male-dominated classroom.

Lastly, students' affective-identity motivation to lead represented the strongest predictor of peer assessment of leadership competency. Despite calls to examine a more comprehensive picture of the leadership development process beyond skill acquisition (Dugan, 2011; Hannah & Avolio, 2010), motivation to lead has remained curiously understudied in the research literature. Our results indicate the significance that peers may place on students' self-identity as emerging leaders, which may be even more relevant and influential to peer assessment than behaviors and self-confidence. Many leadership education interventions continue to focus on some combination of skill acquisition or confidence-building (Dugan, 2011; Owen, 2012). Leadership educators may be wise to include curriculum that seeks to develop students' self-concepts and attitudes as well. As more research is conducted in this area, we may be able to better understand the complex interaction between attitudes, beliefs and behaviors in how teammates assess the leadership competence of peers.

Limitations and Future Research

This study was conducted on one campus and included only a specialized population – first-year Engineering students. While promising, the results would be enhanced if they were replicated using broader, more diverse, populations. Would findings be similar within similarly interdependent professional environments if anonymity could be assured? Recent research has begun to examine this (Gupta, Huang, & Niranjana, 2010). Similarly, a larger sample would permit a more sophisticated statistical analysis, including multi-level modeling, which would allow future researchers to assess the

significance of a “team-effect” on multi-rater assessment scores. It stands to reason that not all teams are created or interact equally, and research is necessary to examine the effects that individual teams have on patterns of multi-rater assessments of leadership skills.

A potential line of research in multi-rater feedback might examine differences between responses that are given for research purposes, such as within this study, and responses that are given for the explicit purpose of providing feedback to the person who is being assessed. Students may shift their responses if they knew that the target of their assessment would receive their feedback, even if anonymously. Educators who engage in multi-rater feedback for developmental purposes might benefit from knowing how students systemically bias their responses in this way.

Future research could also examine the degree that goals and structure affect peer assessment of competency. This study was focused on self-forming teams that shared common goals and were evaluated as a team, not individually. To what extent does individual agency in joining teams matter? Or level of evaluation matter? Findings within studies like this may vary, and if so, might further suggest the importance of team context to the pattern of peer assessment of leadership competency.

Lastly, prospective research could incorporate qualitative components to a multi-rater system. Emerging themes could be compared with quantitative data to determine differences between how individuals complete forced-choice survey items and longer, more contextual responses.

References

- Alimo-Metcalfe, B. (1998). 360 Degree Feedback and Leadership Development. *International Journal of Selection and Assessment*, 6(1), 35-44.
doi:10.1111/1468-2389.00070
- Amit, K., & Bar-Lev, S. (2013). Motivation to lead in multicultural organizations: The role of work scripts and political perceptions. *Journal of Leadership & Organizational Studies*, 20(2), 169-184. doi: 10.1177/1548051812467206
- Arvey, R. D., Zhang, Z., Avolio, B. J., & Krueger, R. F. (2007). Development and genetic determinants of leadership role occupancy among women. *Journal of Applied Psychology*, 93(3), 693-706. doi: 10.1037/0021-9010.92.3.693
- Astin, A.W. & Astin, H.S. (2000). *Leadership reconsidered: Engaging higher education in social change*. Battle Creek, MI: W.K. Kellogg Foundation.
- Asumeng, M. (2013). The effect of employee feedback-seeking on job performance: An empirical study. *International Journal of Management*, 30(1), 373-388.
- Avolio, B. J. (2007). Promoting More Integrative Strategies for Leadership Theory-Building. *American Psychologist*, 62(1), 25-33. doi: 10.1037/0003-66X.62.1.25
- Azzam, T., & Riggio, R. E. (2003). Community based civic leadership programs: A descriptive investigation. *Journal of Leadership & Organizational Studies*, 10(1), 55-67. doi: 10.1177/107179190301000105
- Bass, B. M. (1998). *Transformational Leadership*. Hillsdale, NJ: Erlbaum.
- Bowen, C.C., Swim, J.K., Jacobs, R.R. (2000). Evaluating gender biases on actual job performance of real people: A meta-analysis. *Journal of Applied Social Psychology*, 30(10), 2194-2215. doi: 10.1111/j.1559-1816.2000.tb02432x.

- Carlson, M. S. (1998). 360-Degree feedback: The power of multiple perspectives. *Popular Government*, 63, 38-49.
- Chan, K. Y., & Drasgow, F. (2001). Toward a theory of individual differences and leadership: Understanding the motivation to lead. *Journal of Applied Psychology*, 86(3), 418-498. doi: 10.1037//0021-9010.86.3.481
- Cohen, J. (1987). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Council for the Advancement of Standards in Higher Education. (2009). CAS standards for leadership programs. Washington, DC.
- Drew, G. (2009). A "360" degree view for individual leadership development. *Journal of Management Development*, 28(7), 581-592. doi: 10.1108/02621710910972698
- Dugan, J. P. (2011). Pervasive myths in leadership development: Unpacking constraints on leadership learning. *Journal of Leadership Studies*, 5(2), 79-84. doi: 10.1002/jls.20223
- Dugan, J. P., Garland, J. L., Jacoby, B., & Gasiorski, A. (2008). Understanding commuter student self-efficacy for leadership: A within-group analysis. *NASPA Journal*, 45(2), 282-310.
- Eagly, A.H. & Carli, L.L. (2003). The female leadership advantage: An evaluation of the evidence. *The Leadership Quarterly*, 14(6), 807-834. doi: 10.1016/j.leaqua.2003.09.004
- Eagly, A.H., Makhijani, M.G., & Klonsky, B.G. (1992). Gender and the evaluation of leaders: A meta-analysis. *Psychological Bulletin*, 111, 3-22. doi: 10.1037/0033-2909.111.1.3

- Ely, R. J., Ibarra, H., & Kolb, D. M. (2011). Taking gender into account: Theory and design for women's leadership development programs. *Academy of Management Learning and Education, 10*(3), 474-493. doi: 10.5465/amle.2010.0046
- Farh, J. L., Cannella, A. A., & Bedeian, A. G. (1991). Peer ratings: The impact of purpose of rating quality and user acceptance. *Group & Organization Studies, 16*, 367-385.
- Ghorpade, J. (2000). Managing five paradoxes of 360-degree feedback. *The Academy of Management Executive, 41*(1), 140-150. doi: 10.5465/AME.2000.2909846
- Gupta, V. K., Huang, R., & Niranjana, S. (2010). A Longitudinal Examination of the Relationship Between Team Leadership and Performance. *Journal of Leadership & Organizational Studies, 17*(4), 335-350. doi: 10.1177/1548051809359184.
- Hannah, S. T., & Avolio, B. J. (2010). Ready or not: How do we accelerate the developmental readiness of leaders? *Journal of Organizational Behavior, 31*(8), 1181-1187. doi: 10.1002/job.675
- Hannah, S. T., Avolio, B. J., Luthans, F., & Harms, P. D. (2008). Leadership efficacy: Review and future directions. *The Leadership Quarterly, 19*(6), 669-692. doi: 10.1016/j.leaqua.2008.09.007.
- Hensel, R., Meijers, F., Leeden, R. v. d., & Kessels, J. (2010). 360 degree feedback: How many raters are needed for reliable rating on the capacity to develop competencies, with personal qualities as developmental goals. *The International Journal of Human Resource Management, 21*(15), 2813-2830. doi: 10.1080/09585 192.2010.528664

- Hoyt, C. L. (2005). The role of leadership efficacy and stereotype activation in women's identification with leadership. *Journal of Leadership and Organizational Studies*, *11*(4), 2-14. doi: 10.1016/j.leaqua.2009.10.007.
- Ibarra, H., & Obodaru, O. (2009). Women and the vision thing. *Harvard Business Review*, *87*(1), 62-70.
- Jung, D. I., & Avolio, B. J. (2000). Opening the black box: An experimental investigation of the mediating effects of trust and value congruence on transformational and transactional leadership. *Journal of Organizational Behavior*, *21*, 949-964.
- Kark, R., & Van Dijk, D. (2007). Motivation to lead, motivation to follow: The role of the self regulatory focus in leadership processes. *Academy of Management Review*, *32*(2), 500-528. doi: 10.5465/AMR.2007.24351846
- Lord, R. G., & Hall, R. J. (2005). Identify, deep structure and the development of leadership skill: A review and empirical test. *Journal of Leadership Education*, *1*(2), 34-49. doi: 10.1016/j.leaqua.2005.06.003
- Manning, T., & Robertson, B. (2010). Seniority and gender differences in 360-degree assessments of influencing, leadership and team behaviour Part 2: Gender differences, conclusions and implications. *Industrial and Commercial Training*, *42*(2), 211-291. doi: 10.1108/00197851011048573
- Mayo, M., Kakarika, M., Pastor, J.C., & Brutus, S. (2012). Aligning or inflating your leadership self-image? A longitudinal study of responses to peer feedback MBA teams. *Journal of Management Learning & Education*, *11*(4), 631-652. doi: 10.5465/amle.2010.0069.

- Mero, N.P., Guidice, R.M., Brownlee, A.L. (2007). Accountability in a performance appraisal context: The effect of audience and form of accounting on rater response and behavior. *Journal of Management*, 33(2), 223-252. doi: 10.1177/0149206306297633
- Murphy, S. E. (1992). *The contribution of leadership experience and self-efficacy to group performance under evaluation apprehension*. (Ph.D. 9230410), University of Washington, United States -- Washington. Retrieved from <http://search.proquest.com/docview/304005264?accountid=14553> ProQuest Dissertations & Theses (PQDT) database.
- Murphy, S. E. and Ensher, E. A. (1999), The Effects of Leader and Subordinate Characteristics in the Development of Leader–Member Exchange Quality. *Journal of Applied Social Psychology*, 29: 1371–1394. doi: 10.1111/j.1559-1816.1999.tb00144.x
- Murphy, S. E. (2002). Leader self-regulation: The role of self-efficacy and multiple intelligences. In R. E. Riggio, S. E. Murphy & F. J. Pirozzolo (Eds.), *Multiple intelligences and leadership, LEA's organization and management series* (pp. 163-186). Mahwah, NJ: Lawrence Erlbaum Associates.
- National Association of Colleges and Employers. (2012). The job outlook for the college class of 2013. Bethlehem, PA.
- Nijhof, W. J., & Jager, A. (1999). Reliability testing of multi-rater feedback. *International Journal of Training and Development*, 3(4), 292-300.

- Nowack, K. M., & Mashihi, S. (2012). Evidence-based answers to 15 questions about leveraging 360-degree feedback. *Consulting Psychology Journal: Practice and Research*, 64(3), 157-182.
- Owen, J. (2012). Examining the design and delivery of collegiate student leadership development programs: Findings from the Multi-Institutional Study of Leadership (MSL-IS), a national report. Washington, DC: Council for the Advancement of Standards in Higher Education.
- Podsakoff, P. M., MacKenzie, S. B., Moorman, R. H., & Fetter, R. (1990). Transformational leader behaviors and their effects on followers' trust in leader, satisfaction, and organizational citizenship behaviors. *The Leadership Quarterly*, 1(2), 107-142. doi: 10.1016/1048-9843(90)90009-7
- Podsakoff, P. M., MacKenzie, S. B., Paine, J. B., & Bachrach, D. G. (2000). Organizational Citizenship Behaviors: A Critical Review of the Theoretical and Empirical Literature and Suggestions for Future Research. *Journal of Management*, 26(3), 513-563. doi: 10.1177/014920630002600307
- Roch, S. G., & McNall, L. A. (2007). An investigation of factors influencing accountability and performance ratings. *The Journal of Psychology*, 141(499-523).
- Rohs, F. R. (2002). Improving the evaluation of leadership programs: Control response-shift bias. *Journal of Leadership Education*, 1(2), 50-61.
- Rosch, D. M., Anderson, J. C., & Jordan, S. N. (2012). Analyzing the Effectiveness of Multisource Feedback as a Leadership Development Tool for College Students. *Journal of Leadership Studies*, 6(3), 33-46. doi: 10.1002/jls.21254

- Rosch, D. M., & Schwartz, L. M. (2009). Potential issues and pitfalls in outcomes assessment in leadership education. *Journal of Leadership Education*, 8(1), 177-194.
- Sessa, V. I., Matos, C., & Hopkins, C. A. (2009). Evaluating a college leadership course: What do students learn in a leadership course with a service-learning component and how deeply do they learn it? *Journal of Leadership Education*, 7(3), 167-200.
- Shertzer, J.E. & Schuh, J.H. (2004). College student perceptions of leadership: Empowering and constraining beliefs. *NASPA Journal*, 42(1), 111-131.
- Shipper, F. (2004). A cross-cultural, multi-dimensional, nonlinear examination of managerial skills and effectiveness. *Journal of Leadership & Organizational Studies*, 10(3), 91-103. doi: 10.1177/107179190401000308
- Theron, D., & Roodt, G. (1999). Variability in multi-rater competency assessments. *Journal of Industrial Psychology*, 25(2), 21-27.
- Toegel, G., & Conger, J. A. (2003). 360-Degree Assessment: Time for Reinvention. *Academy of Management Learning & Education*, 2(3), 297-311. doi: 10.2307/40214201
- Toney, F. (1996). A leadership methodology: Actions, traits, and skills that result in goal achievement. *Journal of Leadership & Organizational Studies*, 3(2), 107-127.
- Waldman, D.A., Galvin, B.M., Walumbwa, F.O. (2013). The development of motivation to lead and leader role identity. *Journal of Leadership & Organizational Studies*, 20(2), 156-168. doi: 10.1177/1548051812457416
- Yukl, G. (2010). *Leadership in Organization*, 7th ed. NJ: Prentice Hall.

Author Biographies

David M. Rosch serves as an Assistant Professor in the Agricultural Education program at the University of Illinois at Urbana-Champaign. His particular areas of interest include student leadership development and the accurate assessment of leadership effectiveness.

Daniel A. Collier is a graduate student in the Education Policy, Organization and Leadership doctoral program. He is particularly interested in student leadership development, ethical leadership practices of administration, and evaluation of educational workplace policy.

Sarah M. Zehr serves as an Assistant Dean and Director of the College of Engineering Career Services Center at the University of Illinois at Urbana-Champaign. She is also enrolled in the Education Policy, Organization and Leadership doctoral program. Her particular interests include college student career and leadership development.