Leadership and Gender Roles in Academia

Jeantyl Norze
University of Nevada, Reno
University of Connecticut

Adriana Alfaro
Louisiana State University

Reuben Twijukye
Indiana University Kokomo

The glass ceiling effect is unfortunately still a reality in academia. Only 30% U.S University has female presidents. Given the multiple benefits of gender diversity in senior management, understanding gender gap in top leadership should be a priority. The purpose of this study was to investigate gender roles and leadership effectiveness in higher education. The results show that female leaders were seen less effective than men when committing relationship mistakes. Additionally, both male and female leaders were seen less effective when committing relationship mistakes than task mistakes. These findings signal the need to examine gender bias across other academic settings.

Keywords: leadership, diversity, gender bias, higher education

INTRODUCTION

One can buy books, read articles, attend workshops on "women's leadership", yet one does not see products espousing "men's leadership". The word "leadership" has historically meant men's leadership. The prefix was unnecessary because leaders were men. (McCullough, 2011, p. 7)

The glass ceiling effect is a reality; women are disproportionately underrepresented in senior level positions within organizations (Northhouse, 2013; O'Neil, Hopkins & Bilimoria, 2008). According to the Bureau of Labor Statistics, by 2016 46.8% of the U.S. labor force was comprised of women ("Labor Force Statistics," n.d.). As a comparison, in 2015 approximately 4.4% of CEO positions were held by women in Fortune 500 companies, a dip from the previous year (from 24 female CEOs in 2014 to 22 in 2015; Bellstrom, 2016). Similarly, among Standard & Poor's (S&P) 500 companies, 45% of the labor force consisted of women, while only 19.2% women held board seats and a mere 4.0% held CEO positions in 2014 (Catalyst, 2016). This dearth of women in executive posts is a problem not only for women, given "that the most successful organizations will be the ones that continue to develop the talents and encourage the contributions of their female employees" (O'Neil et al., 2008, p. 727).

Academia does not fare much better than the general workforce does, despite former president Barack Obama's remarks that "there's no place else that has the assets we do when it comes to higher education."

According to the American College President Study (2012) conducted by the American Council on Education, by 2012 a mere 26% universities in the U.S. had female presidents. Likewise, in 2013 female faculty in tenure-track, positions reached approximately 48.4%, yet only 37.5% of women enjoyed tenure (Catalyst, 2015). Within the overall female academic population in 2013, nearly 32.5% were in less prestigious, non-tenured track positions, while only 19.6% of male faculty occupied such positions (Catalyst, 2015). Women who were full professors were also likely to earn 87.2% of what their male counterparts earned (Catalyst, 2015). These numbers, although less alarming than those of the general workforce, are still far from ideal.

Meanwhile, women accounted for 56% of the undergraduate population during the Fall 2014 semester (National Center for Education Statistics, 2016), indicating a mismatch between the number of women in the highest positions in academia and those at the undergraduate level. In other words, the numbers at the upper tiers of leadership in higher education do not reflect the numbers in its student body (Winchester & Browning, 2015). Given the multiple benefits of gender diversity in senior management, understanding the reasons for the gender gap in top leadership – both in academia and in the general workforce – should be a priority.

PURPOSE OF THE STUDY

The primary purpose of this study was to investigate whether gender bias influences faculty's perceptions of leadership effectiveness at a public Research University in the southern region of the United States.

OBJECTIVES

The study objectives were the following:
1. To determine if the perceptions of the faculty members (subordinates) in the study of leadership effectiveness differed based on leaders' gender;
2. To determine if the faculty members in the study perceived a female leader less effective than a male leader when both committed
   a) Tasked errors;
   b) Relationship errors.

METHODOLOGY

This study aimed to investigate whether gender of leader influenced subordinate perceptions of leader effectiveness, and whether type of error and gender of subordinate affected this relationship. To accomplish its purpose, vignettes depicting either a task or a relationship error committed by either a male or a female academic department director were created, and a survey assessing participants' perceptions of leader effectiveness and gender of participant (along with additional demographic items) accompanied each scenario. Independent sample *t*-tests and Analysis of Variance (ANOVA) were used to analyze the data.

POPULATION AND SAMPLE

The target population of this study was faculty members in higher education in the southern region of the United States. The accessible population was 320 faculty members at a Research University of the southern region in the United States. The study sample was 100% of the accessible population.

MEASURES

Data for the present study was collected through a survey. This survey consisted of scales assessing perceptions of leadership effectiveness along with demographic data. Two of Rosser, Johnsrud and Heck's

(2003) subscales of effectiveness of leaders in higher education were used to measure subordinate perceptions of leadership of effectiveness. Rosser et al. (2003) designed this instrument to measure perceptions of department dean's and director's effectiveness, therefore, the researchers found it deemed appropriate for this study. Cronbach Alpha of the Rosser et al. (2003)'s instrument surpassed .90, indicating a very high reliability.

For the focus of this study, the researchers chose these two subscales "management of the unit and interpersonal relationships" of Rosser et al. (2003, p. 10)'s instrument that closely resembled agentic and communal traits reflected in Eagly's (1987) social role theory. Eagly (1987) defines communal qualities (i.e. feminine traits) as reflecting "a concern with the welfare of other people" (p. 16) and associates agentic (i.e. masculine) attributes with self-confidence, self-efficacy, assertiveness, control and dominance.

The original 5-point Likert-type response choices were modified to a 7-point Likert-type scale ranging from strongly disagree (1) to strongly agree (7) to increase the breath of response options available to participants. Some of the original items were also modified – albeit minimally – to enhance clarity. The changes made were carefully considered so as not to alter the underlying meaning of the original items. For example, item #6 initially read as "effective problem solver" and was changed to "*is an* effective problem solver." The only other items where such changes were made were #15 and #17, where the word "is" was added at the beginning of the phrase. The items under each scale are presented below:

Measures of the Unit
1. Insures that fair administrative procedures are followed
2. Exercises fair and reasonable judgment in the allocation of resources
3. Manages change constructively
4. Delegates work effectively
5. Handles administrative tasks in a timely manner
6. Is an effective problem-solver
7. Demonstrates knowledge of departments and programs within the unit
8. Maintains an effective and efficient staff

Interpersonal Relationships
9. Demonstrates understanding of the needs and concerns of unit members
10. Treats individuals fairly and with respect
11. Maintains positive and productive relationships within the unit
12. Maintains positive and productive relationships external to the unit
13. Demonstrates awareness of the quality of professional work of unit members
14. Demonstrates sensitivity to career and mentoring needs of unit members
15. Is accessible to faculty and staff within the unit
16. Demonstrates understanding of the needs and concerns of students
17. Is accessible to students

The scale was factor analyzed and yielded two factors (management of the unit and interpersonal relationships) with loadings ranging from 0.67 to 0.97. The scale was found to be highly internal consistent by the Cronbach's alpha reliability estimate (α = .952). The items 13, 16, and 18 were excluded from the analysis because of cross loading—they had factor loading of .3 or higher on more than factor. The item 22 did meet the minimum threshold set by Hair et al. (2010). Whereas the item 17 was eliminated because it was the single item remained on the third factor after deletion of the other items.

DATA COLLECTION

The researchers sought the permission of the Deans of academic colleges of the university to access the email addresses of their faculty members. Only two colleges agreed to participate in the study. The study was thus sent to a total of 320 faculty members – the entire faculty body for each college – with 190 belonging to one college and 130 to the other. The respondents were given the opportunity to raffle for five $ 100.00 Amazon gift cards. Despite the incentives, only 98 faculty members responded to the survey. However, upon further inspection, 15 cases had to be deleted for the following reasons: the individual did

not provide demographic data (n = 1); reported what appeared to be false demographic data (n = 1); did not consent to participate in the study (n = 2); agreed to participate by selecting "yes" on the consent form but did not complete any of the questionnaire items (n = 7). The 83 remaining cases were utilized for analysis purposes, yielding a 25.9% response rate. The usable response was assigned to four groups or task-relationship scenarios: $n = 27$ for relationship-female, $n = 24$ for relationship-male, $n = 12$ for task-female, and $n = 20$ for task-male.

RESULTS

Objective One

The study's first objective was to examine if women faculty were rated as less effective leaders than men, and an independent-samples t-test was deemed the most appropriate method of analysis. Normality of the dependent variable on all levels of the independent variable was tested using the Shapiro-Wilks test. Results were not significant for the distribution of scores for perceptions of leader effectiveness and female leaders (W = .982, p = .807), but was statistically significant for perceptions of leader effectiveness and male leaders (W = .938, p = .022). However, the Kolmogorov-Smirnov goodness-of-fit test was not significant for the latter (D = .122, p = .114), thus suggesting that departure from normality was not extreme. Pallant (2010) argues that parametric statistics meant to compare groups (e.g. t-tests, ANOVAs) are quite robust to the assumption of normality once 30+ is exceeded (2010). Given that the sample size was 79, the outcome of the Shapiro-Wilks test was not particularly worrisome.

Levene's test of equality of variance was significant (F = 4.896, p = .03), indicating that the homogeneity of variance assumption was not met, and that the t-statistic for equal variances not assumed had to be interpreted. The results for the t-test confirming whether differences in subordinate perceptions of leader effectiveness differed based on the leader's gender (n = 43 for males; n = 36 for females) are presented in Table 1.

TABLE 1
FACULTY'S PERCEPTIONS OF LEADER EFFECTIVENESS

Measures	Male Leaders		Female Leaders		$t(76)$	p
	M	SD	M	SD		
Leadership Effectiveness	3.47	1.39	3.02	1.04	-1.657	.102

The t-test was not significant at t (76) = -1.657, p = .102, indicating that there were no statistically significant differences in perceptions of leader effectiveness when the leader was male (M = 3.47, SD = 1.39) or female (M = 3.02, SD = 1.04).

Additionally, an ANOVA test was conducted to gender of the study participants interact with the gender of the faculty leaders. The gender of leader and gender of participant was considered as independent variables, and perceptions of leader effectiveness as the dependent variable. As can be noted from Table 2, neither of the main effects for gender of leader (F (1, 75) = 2.70, p = .105) nor gender of participant, (F (1, 75) = .113, p = .738) nor the interaction between the two (F (1, 75) = .08, p = .778) were statistically significant.

TABLE 2
INTERACTION EFFECTS BETWEEN GENDERS OF THE PARTICIPANTS AND GENDERS OF THE FACULTY LEADERS

Dependent Variable: Perceptions of Leader Effectiveness

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Corrected Model	4.350[a]	3	1.45	0.918	0.436	0.035	2.755	0.243
Intercept	800.057	1	800.057	506.67	0.00	0.871	506.67	1.00
Gender of leaders	4.262	1	4.262	2.699	0.105	0.035	2.699	0.368
Gender of participants	0.178	1	0.178	0.113	0.738	0.001	0.113	0.063
Gender of participants*Gender of leaders	0.127	1	0.127	0.08	0.778	0.001	0.08	0.059
Error	118.42 9	75	1.579					
Total	964.615	79						
Corrected Total	122.778	78						

    a.   R Squared = .035 (Adjusted R Squared = -.406)
    b.   Computed using alpha = .05

In other words, there was no significant difference in how male subordinates (n = 47, M = 3.24, SD = 1.18) rated the performance effectiveness of male leaders (n = 43, M = 3.47, SD = 1.39) versus female leaders (n = 36, M = 3.02, SD = 1.04). Furthermore, male subordinates (n = 47, M = 3.24, SD = 1.18) and female subordinates (n = 32, M = 3.30, SD = 1.37) did not rate female leaders ((n = 36, M = 3.02, SD = 1.04) and male leaders significantly different on performance effectiveness.

Objective Two
Independent samples t-tests were run to determine whether a) female leaders who committed task errors were rated as less effective than male leaders who committed task errors, and b) female leaders who committed relationship errors were rated as less effective than male leaders who committed relationship errors. The assumption of normality was only partially met given that the distribution of perceptions of leader effectiveness scores was found to be normal across the female leader but not the male leader groups.

The data was split in two (via the split file function in SPSS) on the type of error variable, and the t-test was run with gender of leader as the independent variable and perceptions of leader effectiveness as the dependent variable. The t-test for the task mistake data subset was homogeneous, with a non-significant Levene's test (F = 1.034, p = .319). No significant differences (t (26) = -1.419, p = .168) in subordinate perceptions of leader effectiveness were found when the department director committing a mistake was either male (n = 19, M = 4.55, SD = 1.19) or female (n = 9, M = 3.89, SD = 1.04).

For the relationship mistake data subset, once the homogeneity of variance assumption was established (F = .147, p = .703), interpretations of the statistical test revealed that subordinate perceptions of leader effectiveness did not differ significantly (t (49) = .455, p = .651) between male leaders (n = 24, M = 2.62, SD = .827) and female leaders (n = 27, M = 2.73, SD = .880). Table 3 summarizes these findings.

TABLE 3
PERCEPTIONS OF LEADER EFFECTIVENESS AND ERROR TYPES

Types of Errors	Male Leaders		Female Leaders		df	t	p
	M	SD	M	SD			
Task mistakes	4.55	1.19	3.89	1.04	26	-1.419	.168
Relationship mistakes	2.62	.827	2.73	.880	49	.455	.651

To verify these results, a two-way ANOVA was run. Although there was a non-significant interaction between gender of leader and type of error (F (1, 79) = 2.61, p = .111), and the main effect for gender of leader also failed to reach statistical significance (F (1, 79) = 1.335, p = .252), significance for the main effect of type of error was obtained (F (1, 79) = 42.164, P = .000, partial eta squared = .360). In other words, both male and female leaders were rated significantly lower on performance effectiveness when committing relationship mistakes (M = 2.68, SD = .849) than task mistakes (M = 4.24, SD = 1.17). The effect size of the relationship was small-to-medium (partial eta squared = .360), and the level of power exceeded the .8 threshold (Pallant, 2010). Despite not being a part of the objectives of the study, the findings were considered unexpected and worthy of mention. The results of the ANOVA are presented in Table 4.

TABLE 4
INTERACTION EFFECTS BETWEEN GENDERS OF THE PARTICIPANTS AND TYPES OF ERROR

Dependent Variable: Perceptions of Leader Effectiveness

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[b]
Corrected Model	52.643[a]	3	17.548	18.765	0.00	0.429	56.295	1.00
Intercept	783.496	1	783.496	837.843	0.00	0.918	837.843	1.00
Gender of leaders	1.248	1	1.248	1.335	0.252	0.017	1.335	0.207
Type of Error	39.429	1	39.429	42.164	0.00	0.36	42.164	1.00
Gender of participants*Type of Error	2.436	1	2.436	2.605	0.111	0.034	2.605	0.357
Error	70.135	75	0.935					
Total	964.615	79						
Corrected Total	122.778	78						

a. R Squared = .429 (Adjusted R Squared = .406)
b. Computed using alpha = .05

DISCUSSIONS

Gender bias is a complex phenomenon not easily studied. For instance, Eagly and Carli (2007) expressed how factors such as household and childrearing responsibilities as well as outdated organizational policies and practices can impact the expression of discrimination in the workplace. Variables including

cultural context (e.g. country of origin; Curseu & Boros, 2008), workplace context (Thoroughgood et al., 2012), and diversity climate (Kossek & Zonia, 1993), can influence attitudes toward female leaders. These and other factors suggest a degree of complexity in the gender bias equation that can affect the relationship between gender of leader and perceptions of leader effectiveness. Thus, non-significant findings may be a consequence of the interplay of a multitude of factors that this study did not take into account.

In addition, Thoroughgood et al. (2012), from which this study was modeled, were also unable to confirm their hypotheses pertaining to gender bias except when accounting for the 80 context the scenario took place in. In the present study, the interaction between type of error and gender of leader was also not significant. Thus, several non-significant findings in this study were in line with those of previous research, and should therefore not be entirely surprising.

One of the elements that both Thoroughgood et al. (2012) and this study shared was the investigation of gender bias under negative conditions; that is, when leaders make mistakes. The literature has mostly focused on differential treatment of male and female leaders under conditions of success (Thoroughgood et al., 2012); that is, variances in perceptions of male and female leader effectiveness when performance is effective and goals are achieved. It is possible that the non-significant findings of this study derived from differences in how leaders are viewed and judged under effective versus non-effective leadership examples. For instance, subordinates may consider mistakes equally detrimental or unpardonable – or equally benign and forgivable – regardless of leader gender.

Another issue that could have influenced the results were the unequal sample sizes of the different groups. Although unequal sample sizes where the smallest group is no less than half the size of the larger group are usually not considered problematic (Morgan, Leech, Gloeckner, & Barrett, 2004), they may have posed a problem for this particular dataset. All sample sizes were within the 1.5 boundary suggested by Morgan et al. (2004), with the exception of gender of leader under the task-based scenario (n = 19 for males, and n = 9 for females). However, when running diagnostics to detect the presence of influential outliers, and when all potential influential observations were taken into account (i.e. all observations appearing under DFBETAs, studentized deleted residual's, Cook's D, and central leverage values), 6 out of 9 belonged to the task-female scenario. Similarly, when determining the observations that appeared on two or more methods, 3 out of 4 belonged to the task-female group. Hence, the 82 majority of influential observations derived from the task-female group, and the elimination of three influential data points further reduced its sample size from n = 12 to n = 9. This discrepancy in sample sizes between conditions seems to have influenced the analysis and potentially affected the findings.

Likewise, the small sample size in the present study most likely resulted in a lack of statistical power, potentially leading to a type II error. Type II errors occur when a false null hypothesis is not rejected (Hinkle et al., 2003); that is, when an actual difference exists yet the statistical test fails to detect it. To minimize the possibility of a type II error from occurring, the power of the test should be large enough to enable the statistical analysis to unearth true differences. Traditionally, the .80 level has been used (Cohen, 1992). For the power to reach this threshold, nonetheless, three additional conditions must be outlined: the alpha level, the effect size, and the sample size (Cohen, 1992). The alpha level, as is the case for the present study, is most often established at .5. Meanwhile, the literature dictates the expected effect size, which usually ranges from small (r = .1), to medium (r = .3), to large (r = .5; Cohen, 1992). Lastly, the required sample size is determined based on the alpha level set, the desired power level, and the predicted magnitude of the effect size of the observed variables under study.

Cohen's (1992) power table has been one of the most widely used guidelines to determine the required sample size once the alpha, power, and effect size have been established. For the present study, the alpha level was set at .5, the power level wished for was .80, and the expected effect size (based on the literature) was small (Eagly & Carli, 2007; Swim & Sanna, 1996; Thoroughgood et al. 2012). As per Cohen's (1992) guideline, if the power were to confidently reach .80, the sample size should have been 393 for the t-test and for the ANOVA (this guideline was originally established for one-way ANOVAs while the present study utilized a two-way 83 ANOVA, yet this threshold was still used as a benchmark). Considering that the survey was sent to 320 people total, and that only 82 responses were usable (which became reduced to

79 after influential observations were accounted for), the obtained sample size was far from desirable. The .80 power level was thus most likely not reached.

In fact, the ANOVA results, which provide power estimates, show power levels falling greatly below the desired .80 cutoff. Thus, there is a strong possibility that the dearth of significant findings stemmed from a lack of statistical power – which originated from the considerably small sample size – rather than the actual absence of significance. The inability to detect true differences and a failure to reject a false null hypothesis may have therefore been the culprit behind the lack of significant findings.

Social desirability is another factor that could have influenced the present study's results. Individuals may have been hesitant to express attitudes and beliefs that are no longer perceived as acceptable by society (Walker et al., 2013). That is, prejudicial attitudes towards female leader may be suppressed in order to avoid eliciting negative responses from others. Although the present investigation only revealed the gender bias component of the study after-the-fact, it is possible that participants may have caught on to the true nature of the study. Considering that study participants belonged to either one of two colleges, and that the survey was sent to all faculty members within the college, participants may have talked to each other about the study 84 before having a chance to complete the survey, and may have been informed of its focus on differential treatment of leaders based on gender. Having been alerted to the research's emphasis on gender bias may have primed participants to answer in socially desirable ways. Even if participants were not aware of the gender bias component from the start, given that they were answering potentially delicate questions (about diversity climate in their department) and rating a leader – albeit fictitious – they may have been hesitant to provide candid answers. Unfortunately, the anonymous nature of the study did not allow for follow up interviews or focus groups with study participants to determine whether gender bias is indeed viewed as a problem within their respective departments, whether there was a discrepancy between the case study and reality as perceived by participants, or whether members of the two colleges differed systematically from faculty at other colleges within the same institution (or at equivalent colleges at other institutions). Follow-up qualitative inquiry could have helped establish whether the quantitative results accurately reflected gender bias amongst the sample.

DIRECTIONS FOR FUTURE RESEARCH

From a research perspective, the present findings signal the need to determine whether the results regarding gender bias hold true across other academic settings. For instance, Lee and Won (2014) list a number of factors that can affect pay differentials between male and female faculty at the assistant professor level, such as the institution's size; whether it is public or privately funded; whether it is classified as research-intensive/extensive or other; the "concentration of high-paying disciplines" within the institution (p. 335); the percentage of women serving in the respective institution's state legislature; the degree or severity of the wage gap between men and women for the state; and the region in which it is located. This suggests universities can differ greatly across a number of key factors with regard to differential treatment based on gender. It further suggests that conducting the study under different academic settings may yield different results. To determine whether non-significance does or does not extend to other colleges and universities, both similar and dissimilar to the one used in the present study, replication is necessary.

From a theoretical perspective, the present findings suggest a need to reexamine role congruity theory in explaining gender bias in leadership. A major premise of role congruity theory is that leadership is still generally defined in masculine terms. The rise of leadership theories such as transformational leadership (Northhouse, 2013), with greater inclusion of feminine traits, suggests that a more profound investigation of alternative conceptualizations of leadership may be warranted. One such potential theory that may better help explain prejudicial attitudes toward women in the workplace is Glick and Fiske's (1996) hostile vs. benevolent sexism. Under this framework, there are two types of prejudice: overt, blatant discrimination (i.e. hostile sexism), and discrimination that disguises "sexist apathy" (Glick & Fiske, 1996, p. 491) under more positive overtones (i.e. benevolent sexism). The premise of this theory is that while hostile sexism is generally frowned upon and no longer regarded as acceptable, benevolent sexism is, although less evident, common and often sanctioned by both men and women. Similar to role congruity theory, benevolent sexism

endorses women's "positive traits relate[d] to socialemotional, not agentic dimensions" (Glick and Fiske, 1996, p. 492). Different from role congruity theory, however, benevolent sexism posits that prejudice masks itself under positive feelings and expressions such that negative stereotypes appear favorable instead. For example, in terms of women in academia, while female faculty may be regarded as better able to relate to students and create connections when teaching, they may be inadvertently seen as less competent than men at research and publishing, the latter which is the most important component in tenure acquisition. The positive feelings related to women's ability to interact with other students disguises the discrimination associated with the belief that women are not as effective at more important, masculine tasks.

REFERENCES

American Council on Education. (2012). *The American College President Study: Key findings and takeaways*. Retrieved from http://www.acenet.edu/the-presidency/columns-andfeatures/Pages/The-American-College-President-Study.aspx

Bellstrom, K. (2016, December 23). *Why 2015 was a terrible year to be a female Fortune 500 CEO*. Boston, MA: Harvard Business School Publishing Corporation. Retrieved from http://fortune.com/2015/12/23/2015-women-fortune-500-ceos

Catalyst. (2015). *Women in academia*. Retrieved from http://www.catalyst.org/knowledge/women-academia#footnote34_5818cui

Catalyst. (2016). *Pyramid: Women in S&P 500 Companies*. New York: Catalyst, April 3, 2015.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.

Curşeu, P.L., & Boroş, S. (2008). Stereotypes and prejudice towards women managers: An experimental illustration using the Goldberg-paradigm in a Romanian sample. *Studia Psychologica*, *50*(3), 233–254.

Eagly, A.H. (1987). *Sex differences in social behavior: A social-role interpretation*. Lawrence Erlbaum Associates, Inc.

Eagly, A.H., & Carli, L.L. (2007). *Through the labyrinth: The truth about how women become leaders*.

Glick, P., & Fiske, S.T. (1996). The ambivalent sexism inventory: Differentiating hostile and benevolent sexism. *Journal of Personality & Social Psychology*, *70*(3), 491–512.

Hair, J.F., Anderson, R.E., & Black, W.C. (2010). *Multivariate Data Analysis*. Prentice Hall.

Hinkle, D.E., Wiersma, W., & Jurs, S.G. (2003). *Applied statistics for the behavioral sciences* (5th Ed.). Houghton Mifflin.

Kossek, E.E., & Zonia, S.C. (1993). Assessing diversity climate: A field study of reactions to employer efforts to promote diversity. *Journal of Organizational Behavior*, *14*(1), 61–81.

Lee, Y., & Won, D. (2014). Trailblazing women in academia: Representation of women in senior faculty and the gender gap in junior faculty's salaries in higher educational institutions. *The Social Science Journal*, *51*(3), 331–340. https://doi.org/10.1016/j.soscij.2014.05.002

McCullough, L. (2011). Women's leadership in science, technology, engineering and mathematics: Barriers to participation. *Forum on Public Policy Online*, *2011*(2), 1–11.

Morgan, G.A., Leech, N.L., Gloeckner, G.W., & Barrett, K.C. (2004). *IBM SPSS for introductory statistics: Use and interpretation* (5th Ed.).Taylor & Francis.

National Center for Education Statistics. (2016). *The condition of education*. Retrieved from http://nces.ed.gov/programs/coe/indicator_cha.asp

Northouse, P.G. (2013). *Women and leadership. Leadership: Theory and practice* (pp. 349–382). SAGE.

O'Neil, D., Hopkins, M., & Bilimoria, D. (2008). *Women's careers at the start of the 21st Century: Patterns and paradoxes*.

Pallant, J. (2010). *SPSS survival manual* (4th Ed.). McGraw-Hill.

Rosser, V.J., Johnsrud, L.K., & Heck, R.H. (2003). Academic deans and directors: Assessing their effectiveness from individual and institutional perspectives. *Journal of Higher Education*, *74*(1), 1–25.

Swim, J.K., & Sanna, L.J. (1996). He's skilled, she's lucky: A meta-analysis of observers' attributions for women's and men's successes and failures. *Personality and Social Psychology Bulletin*, *22*(5), 507–519. https://doi.org/10.1177/0146167296225008

Thoroughgood, C.N., Sawyer, K.B., & Hunter, S.T. (2012). Real men don't make mistakes: Investigating the effects of leader gender, error type, and the occupational context on leader error perceptions. *Journal of Business and Psychology*, *28*(1), 31–48. https://doi.org/10.1007/s10869- 012-9263-8

Walker, S.S., Madera, J.M., Hebl, M.R. (2013). Effects of leader race and leader mistake on patronizing behaviors. *Journal of Business Diversity*, *13*(1/2), 52–64.

Winchester, H.P., & Browning, L. (2015). Gender equality in academia: A critical reflection. *Journal of Higher Education Policy and Management*, *37*(3), 269–281. https://doi.org/10.1080/1360080X.2015.1034427