2018

Gender Differences in Development Center Performance in a Healthcare Organization

Samuel Lawson
Minnesota State University, Mankato

Gender Differences in Development Center Performance in a Healthcare Organization

Samuel D. Lawson

A Thesis Submitted in Partial Fulfillment for the Degree of Master of Arts in

Industrial/Organizational Psychology

Minnesota State University, Mankato

Mankato, Minnesota

February 27, 2018

Gender Differences in Development Center Performance in a Healthcare Organization

This thesis has been examined and approved by the following members of the student's committee, on the day of February 27, 2018.

_____

Dr. Andrea Lassiter, Advisor

_____

Dr. Daniel Sachau, Committee Member

_____

Dr. Kathy Dale, Committee Member

Abstract

Over the past 50 years, the assessment center has evolved into a tool used for making selection and job placement decisions, as well as for identifying development areas among incumbents to promote skill and competency development where gaps exist. The present study examined the latter, that of development centers, to address potential gender differences in performance on a variety of exercise and competency areas within a development center context. Research efforts were also directed toward the exercises and competencies themselves and the relationships between them. Multiple regression analyses were conducted to investigate exercise-competency relationships, while t-tests were conducted to investigate potential gender differences in performance. Results revealed that women significantly outperformed men in the development center, aligning with findings of previous research on this topic. Specifically, women outperformed men on exercises and competencies that were relationship-oriented in nature, thus aligning with gender-based stereotypes that exist about women. Results also revealed that role play exercises made the strongest unique contribution to explaining performance in the development center, followed by presentation and leaderless group discussion exercises, respectively. Role play exercises were most significantly related to competencies with a relationship-oriented nature, while presentation exercises were most significantly related to competencies with a task- and analytical-oriented nature. Despite the lack of generalizability of these findings, the present study still contributes to existing literature regarding gender and performance within a development center context, as well as exercise-competency relationships. Future research might investigate how certain personality attributes, cognitive ability, or job performance is related to performance within a development and/or assessment center context.

Table of Contents

Gender Differences in Development Center Performance in a Healthcare Organization

Since its introduction as a selection instrument over 50 years ago, assessment centers (ACs) have been utilized for a variety of purposes in organizations across the globe (Thornton & Krause, 2009). Initially used at AT&T to study adult development over a time period of several years, the assessment center methodology has evolved into a tool used for making selection and promotion decisions, as well as for identifying areas of development among job incumbents. Presently, the assessment center method is still used for these two different purposes: (a) selecting among candidates for open positions and promotions; and (b) identifying development areas of job incumbents to promote skill and competency development (Thornton & Rupp, 2006). The present research study focuses on the latter, that of development centers, to address gender differences in performance in a development center context, in addition to exercise-competency relationships (i.e., whether exercises assess the competencies they purport to measure).

Additionally, a substantial gap in the literature exists regarding the impact of gender on how participants perform and on how participants are rated in development and/or assessment centers. Anderson et al. (2006) summarized the findings of several studies on the topic and found mixed results. In some instances, there have been significant differences in center performance favoring women, whereas in other instances, no significant differences in performance have been found between genders. Thus, the purpose of the present study is two-fold: (a) Contribute to the minimal research that previously has been conducted on gender differences in performance in development centers and (b) evaluate exercise-competency relationships and address methodological limitations of the center (e.g., whether exercises are reflective of the leadership competencies they intend to measure).

Development Centers

A recent application of the assessment center method that has gained widespread popularity over the past 15 years is that of the development center. The development center is best defined as "a collection of workplace simulation exercises and other assessments that provide individuals with practice, feedback, and coaching on a set of developable behavioral dimensions found to be critical for their professional success," (Rupp et al., 2006, p. 78). Rather than used as a tool for personnel decision making, development centers stress experiential learning, feedback, performance improvement, and identification of areas of development on specific, relevant behavioral competencies (Rupp et al., 2006). Individuals in a development center participate in a series of simulations that evaluate skills and competency areas critical for success at or above their current position level. Trained observers evaluate participants on a number of behavioral dimensions across several exercises. Based upon their performance across multiple exercises, participants are provided with strengths and development areas to improve upon through coaching, stretch assignments, and various follow-up training approaches (Thornton & Rupp, 2005). Additionally, Appelbaum, Harel, and Shapiro (1998) noted that development centers are generally considered fair (i.e., job-relatedness between tasks and exercises), objective (i.e., assess actual behavior), and are accepted by participants as a means of evaluating areas of development.

Behavioral Dimensions and Exercises. The number of dimensions assessed in development centers ranges from as few as three to as many as nine, with a focus on both interpersonal and intrapersonal characteristics. However, it is important to note that as the number of assessed dimensions increases, the risk of those dimensions becoming indistinguishable also increases (Chan, 1996). This complicates the difficult task that observers

have of differentiating one dimension from another, in addition to overtaxing their capacity to

process information and to provide specific feedback on a number of dimensions. A meta-

analysis conducted by Arthur and colleagues (2003) found evidence for six behavioral

dimensions: communication, consideration of others, drive, influencing others, organization and

planning, and problem solving. A majority of development center programs tend to evaluate

dimensions regarding (a) consideration of others; (b) drive; (c) influencing others; and (d)

communication (Anderson et al., 2006). Consideration of others involves the extent to which an

individual's actions reflect a consideration for the feelings and needs of others. Drive involves

the extent to which an individual maintains a high activity level, sets high performance

standards, and expresses desire to advance and develop. Influencing others consists of the extent

to which an individual persuades others to do something or adopt a point of view in order to

produce desired results. Lastly, communication involves the extent to which an individual

conveys oral and written information and responds to challenges effectively (Arthur et al., 2003).

Each of these four commonly evaluated dimensions in development centers are representative of

the dimensions and exercises in the present study.

In addition to the aforementioned behavioral dimensions, development centers tend to

utilize similar exercises and simulations across organizations. Thornton and Krause (2009)

conducted a survey of center design and execution, and found the most commonly utilized

exercises in development centers to be role plays, presentations, group discussions, and case

studies. Additionally, presentations, case studies, and group discussions were found to be used

by organizations in development centers at a disproportionately higher rate to that of assessment

centers (Thornton & Krause, 2009). The aforementioned exercises common to a majority of

development center designs are briefly discussed, in turn.

Thornton & Rupp (2005) provides a concise overview of the focus and objective of each of these exercises. The exercises present in development centers typically revolve around a business case, in which participants must first read through and analyze the current business situation before participation in any of the simulations (CEB, 2013). In a role play, the participant talks one-on-one with an observer who plays the role of a subordinate, colleague, or customer. The role player is trained to act in a standardized manner, contingent upon the situation, while the participant must resolve the issue at hand. The role play exercise emphasizes oral communication skill, empathy, and problem solving. In presentation exercises, participants make a longer, formal presentation regarding improving business strategy or addressing a business operations-related issue as indicated in the business case. Observers ask questions intended to challenge the participant, and may even place the participant under stress by opposing his/her conclusions or identifying limitations of proposed ideas. The presentation exercise emphasizes oral communication skill and ability to convey ideas coherently under pressure and scrutiny (Thornton & Rupp, 2005). In group discussion exercises, 4-8 participants are given a problem to resolve in a fixed period of time. Participants discuss the problem and prepare recommendations on resolving the issue (e.g., new market venture) that have been endorsed by all participants. Generally, no roles are assigned, and participants cooperate in developing the best solution for the organization as a whole, given the current situation as presented in the business case. The group discussion exercise emphasizes group leadership skills, in addition to problem solving and decision making abilities (Thornton & Rupp, 2005). Each of these commonly used exercises in development centers also are utilized in the present study; the business case is used as background information and the basis for the structure and objectives pertaining to each exercise.

Purpose, Effectiveness, and Best Practices. The "idea of re-focusing assessment centers on development has been encouraged by a number of new concepts in human resource management – continuous development, the learning organization, and empowerment of individuals and competencies," (Holbeche, 1995, p. 26). As organizations have become ever-increasingly learning-oriented and focused on continuous development of human capital, development centers have found their place as a valuable tool used by many organizations worldwide. Appelbaum, Harel, and Shapiro (1998) assert that through precise feedback and a practical on-the-job follow-up of takeaways from the exercises, development centers are a powerful tool in enhancing competencies of an organization and its employee; this is accomplished through reliable, accelerated development within a dynamic, ever-changing business environment.

The purpose of development centers is helping people already in positions to improve their performance. While some assessment centers have "passive candidates" who are only evaluated for selection purposes, development centers tend to have "active participants" who immerse themselves in a collaborative venture with the aid of observers and develop through critical feedback (Adams, 1995). Development centers allow participants to measure their own abilities and skill sets against the organization's criteria, identify development needs, and plan further activities for improvement (Appelbaum, Harel, & Shapiro, 1998). Essentially, development centers renounce notions of "pass" or "fail", provide feedback to participants, and encourage collaboration. Having employees continually strive for excellence and continuous development enables an organization to compete in an ever-changing environment. Additionally, the psychological benefits garnered by the employees (i.e., self-responsibility for development) and the organization (i.e., providing constructive feedback) is mutually beneficial to both parties

(Rupp, Snyder, Gibbons, & Thornton, 2006). Whereas the criterion and focus in assessment centers is predicting future performance, that of development centers is improving upon current performance in the organization.

The effectiveness of a development center is contingent upon a number of conditions. Jones and Whitmore (1995) found that although participation in a center did not necessarily predict promotion into management, following recommended developmental activities after conclusion of the center did relate to eventual promotion. Development centers alone cannot be effective without the implementation of a well-constructed developmental plan. These plans require continuous monitoring, support, and reevaluation as necessary. Thus, development centers result in improved job performance if the following is provided: (a) accurate, clear feedback; (b) direct managerial involvement in planning and development; and (c) development plans for self-managed learning (Appelbaum, Harel, & Shapiro, 1998; Thornton & Krause, 2009). Additionally, evidence of predictive validity in development centers is contingent upon whether dimension proficiency increases over time as a result of the development center itself, as well as subsequent developmental activities (Rupp, Synder, Gibbons, & Thornton, 2006). Follow-up activities may include participation in relevant training, stretch assignments, and job shadowing of those possessing desired skill sets. If development centers are to be an effective means of improving managerial performance, they must be designed in ways that make development most likely—focusing on competencies critical for success, susceptible to change, and believed to be improvable (Rupp et al., 2006).

Furthermore, previous research has identified some best practices in development center design to ensure effectiveness in individual development and improving performance. Thornton and Rupp (2005) assert that the quality of the development center is contingent upon the ability

of observers to provide meaningful, distinct feedback on each dimension. In development

centers, feedback is provided to participants during a feedback session with an assessor and/or a

feedback report (Bell & Arthur, 2008). Providing individuals with feedback on their areas of

development will result in developmental activities on their part to improve, enhancing their

value as an employee to the organization. This highlights the aforementioned point that fewer

dimensions is better, as there is minimal chance for conceptual overlap, resulting in easier

evaluation by observers. To maximize learning throughout the process, a development center

may incorporate elements such as training on the dimensions, experiential learning, self-

reflection, and coaching as follow-up activities to the center (Rupp, Synder, Gibbons, &

Thornton, 2006). Participants should be encouraged to engage in further development activities

following the center to sustain continued efforts of improving themselves on various behavioral

dimensions. The effectiveness of development centers is often compromised when follow-up

developmental activities are not pursued after the conclusion of the center (CEB, 2013). As

Adams (1995) asserts, a development center should be seen as the starting point, and not the end

point, of individual development. For true development to occur, the development center process

during and after should provide multiple opportunities for practice and feedback that result in

actual changes in attitude and behavior.

Gender in Development Centers

Inconsistent findings in the literature exists regarding the impact of gender on how

participants perform and on how participants are evaluated in development and/or assessment

centers. Overall, previous studies have produced mixed results. In some studies, males and

females have not significantly differed in performance (Moses, 1973; Moses & Boehm 1975;

Ritchie & Moses, 1983; Anderson & Thacker, 1985; Shore et al., 1997), whereas in other

studies, females have significantly outperformed males in certain exercises and competency areas (and vice versa) (Schmitt & Hill, 1977; Walsh et al., 1987; Shore, 1992; Schmitt, 1993; Weijerman & Born, 1995; Anderson et al., 2006; Falk & Fox, 2014). Additionally, it is important to recognize that significant results in these instances are often not generalizable; the generalizability of center results pertaining to gender is a function of the meaning and context of the center dimensions and exercises (Anderson et al., 2006). Because exercises and dimensions vary in design, purpose, and structure across centers, a significant finding in one study may not appear in another study, which may explain the inconsistency in results. The following studies regarding gender differences in development and assessment center performance are discussed in turn, highlighting the inconsistent findings obtained from prior research.

Gender Differences in Performance. One of the earliest assessment centers was used for the early identification of supervisory potential at the American Telephone and Telegraph Company (Moses, 1973). Men and women participated in a series of exercises, in which an overall rating of management potential was made; each participants was evaluated on the likelihood of his/her overall success as a management candidate. Results indicated no significant differences in promotion ratios between men and women. Additionally, Moses and Boehm (1975) conducted a follow-up study, in which assessment center performance was strongly related to subsequent promotions into management. The assessed dimensions (e.g., organizing and planning, leadership) relating most strongly to subsequent management level were again the same for men and women. Furthermore, Ritchie and Moses (1983) found that overall assessment rating was significantly related to progress in management, and that the success rate of women was comparable to men.

Schmitt and Hill (1977) conducted one of the first studies evaluating assessment center differences in performance across race. Minimal effects were found as a result of the race-sex composition of the sample, but some assessment ratings for black women were negatively and significantly correlated with the number of white males in the assessment sample. Additionally, the ratings of white males tended to be higher when the number of white males in the assessment sample increased. Schmitt (1993) conducted an additional study in which participants in an assessment center were evaluated for selection into school administrator positions. Results indicated that women were significantly favored in all assessment center dimensions over men. Additionally, White candidates were significantly favored over Black candidates in nine of twelve behavioral dimensions (Schmitt, 1993).

Walsh, Weinberg, and Fairfield (1987) analyzed assessment center evaluations of candidates for evidence of observer-participant sex effects. Results indicated a significant main effect for participant sex, in that females received higher ratings than males. Additionally, a significant interaction was found with observer and participant sex impacting overall assessment ratings. The all-male observer group in the study rated female participants significantly higher than male participants. The all-male observer group may have been accurately evaluating differences between male and female participants, while the mixed observer group failed to recognize that female participants were in fact more qualified. Additionally, the all-male observer group may have been overly lenient in its evaluation of females (Walsh, Weinberg, & Fairfield, 1987). Furthermore, a meta-analysis conducted by Dean, Roth, and Bobko (2008) of data on assessment center performance showed significant differences between groups. Specifically, White participants significantly outperformed Black participants; Hispanic participants significantly scored higher than White participants; and females, on average, scored

higher than males across assessment centers. Additionally, Falk and Fox (2014) investigated the influence of gender and ethnicity compositions of participants on assessment center performance. Results indicated that participants' success was a function of observers' subjective impressions. A higher proportion of female participants was associated with higher scores for males, especially when the proportion of female observers was high (Falk & Fox, 2014). When more female observers were present, and there were more female than male participants, males actually scored higher in the center. Observers from negatively stereotyped groups tend to identify themselves with positively stereotyped participants. Thus, female observers may be more inclined than male observers to view male participants positively due to a positive leadership style associated with masculinity (Falk & Fox, 2014).

Shore, Taschian, and Adams (1997) analyzed development center performance of participants in three role play exercises. Results indicated no significant differences in ratings for male and female participants on any behavioral dimensions. Female observers rated participants of both sexes higher on some dimensions than male observers, but this only occurred in one exercise. Additionally, there were no significant interactions between observer and participant gender for any of the exercises. These findings attest to the fairness of the development center method for the purpose of development (Shore et al., 1997).

Anderson, Lievens, van Dam, and Born (2006) conducted a construct-driven study of gender differences in a leadership-oriented assessment center. Results indicated that female participants were rated significantly higher on constructs reflecting an interpersonally-oriented leadership style (i.e., communication, interaction), as well as on drive and determination. Male participants were rated significantly higher on constructs reflecting a task-oriented leadership style (i.e., problem solving) (Anderson, Lievens, van Dam, & Born, 2006). These findings

suggest that differences in center performance for gender may be a function of varying leadership styles of men and women.

   Gender Differences in Leadership Style. The inconsistency in performance differences in centers due to gender may be attributable to differences in leadership styles for men and women. Atwater and Roush (1994) examined the impact of gender on upward evaluations (i.e., ratings of leaders or managers made by subordinates). Results indicated that male subordinates rated female leaders lower than they rated male leaders. Male subordinates also rated female leaders lower than female subordinates rated female leaders. Additionally, female subordinates did not rate female leaders higher than male leaders. These differences may be attributable to the nature of men and women's leadership styles (Atwater & Roush, 1994).

   A meta-analysis conducted by Eagly and Karau (2001) revealed stereotypical gender differences in task-oriented and interpersonally-oriented leadership style. Task-oriented leaders are concerned with accomplishing assigned tasks by organizing task-relevant constructs, while interpersonally-oriented leaders are concerned with maintaining interpersonal relationships by tending to the morale and welfare of others (Ritter & Yoder, 2004). In the meta-analysis, men emerged more often as task-oriented leaders who displayed directive and controlling leadership styles. Men are more motivated to work in a competitive environment, exert an assertive role, and stand out from the group (Eagly & Karau, 2001). Women, on the other hand, emerge more often as interpersonally-oriented leaders who facilitate interpersonal relations and contribute to good morale. These meta-analytic findings support the notion of role congruity theory, which states that people are expected to engage in activities consistent with their gender roles (Eagly & Karau, 2002; Ritter & Yoder, 2004). Violations of these gender stereotypes may lead to lower

performance evaluations of women or men. For example, if a woman acts in a certain manner during a center exercise counter to her gender stereotype, she may be evaluated lower as a result.

Counter to rule congruity theory, expectancy violation theory states that behaviors that violate gender stereotypes will actually be more positively evaluated (Eagly & Karau, 2001). Violations of these gender stereotypes may lead to higher performance evaluations of women or men. For example, if a woman acts in a certain manner during a center exercise counter to her gender stereotype, she may be evaluated higher as a result. Thus, differences in exercise and dimension performance in development and/or assessment centers may be attributable to differences in leadership styles for men and women.

Individual Differences in Development. Aside from the potential of observer bias and differences in leadership style influencing participant performance in development and/or assessment centers, it is important to be cognizant of individual differences that moderate development. A meta-analysis by Collins and colleagues (2003) found overall center ratings to be significantly related to cognitive ability (.67), extraversion (.50), emotional stability (.35), openness (.25), and agreeableness (.17). Thus, cognitive ability, personality factors, and the ability to regulate one's emotions may moderate development, and impact the ratings that participants receive in a center. The possibility of participants' differential ratings of performance in a development center due to cognitive ability and/or inherent personality characteristics is a potential factor that could impact their results (Rupp, Snyder, Gibbons, & Thornton, 2006). To be specific, differences in ratings that participants receive in a center may be a function of their cognitive ability and varying personality traits.

To summarize the state of the literature on gender differences in performance in development centers, previous research indicates mixed results. Many studies have found no

significant differences between males and females in their performance on exercises or across behavioral dimensions (Moses, 1973; Moses & Boehm, 1975; Ritchie & Moses, 1983; Anderson & Thacker, 1985; Shore et al., 1997). However, several studies have found females to perform significantly better than males in certain behavioral dimensions or exercises, and vice versa (Schmitt & Hill, 1977; Walsh et al., 1987; Shore, 1992; Schmitt, 1993; Weijerman & Born, 1995; Anderson et al., 2006; Falk & Fox, 2014). Additionally, significant differences in performance across gender may be due to a variety of reasons, such as cognitive ability and personality (Collins et al., 2003), or varying leadership styles between men and women (Anderson et al., 2006).

Present Study

Overall, previous findings in the literature regarding the impact of gender on how participants perform and on how participants are evaluated in development and/or assessment centers are inconsistent; these studies have produced mixed results. In some studies, males and females have not significantly differed in performance (Moses, 1973; Moses & Boehm 1975; Ritchie & Moses, 1983; Anderson & Thacker, 1985; Shore et al., 1997), whereas in other studies, females have significantly outperformed males in certain exercises and competency areas (and vice versa) (Schmitt & Hill, 1977; Walsh et al., 1987; Shore, 1992; Schmitt, 1993; Weijerman & Born, 1995; Anderson et al., 2006; Falk & Fox, 2014). Additionally, minimal research has assessed differences in performance by gender in development centers; much of the literature has focused solely on assessment centers.

Thus, the purpose of the present study is two-fold: (a) Evaluate exercise-competency relationships and address methodological limitations of the center (e.g., whether exercises are reflective of the leadership competencies they intend to measure) and (b) contribute to the

minimal research that previously has been conducted on gender differences in performance in development centers. The following questions are founded on research on development centers, performance in development centers, and gender differences in performance. As many of these are exploratory, several are framed as research questions, rather than hypotheses.

Research Question 1. Which exercise has the greatest impact on overall performance in each competency area?

Research Question 2. Which exercise has the greatest impact on overall performance in the development center?

Research Question 3. Which competency has the greatest impact on overall performance in the development center?

Research Question 4. Are there any differences in performance on any of the competency areas between cohorts that have participated in the center?

Research Question 5. Are there any differences in performance on any of the exercises between cohorts that have participated in the center?

Research Question 6. Is there a significant difference between male and female participants on overall center performance?

Hypothesis 6a. Females will have a higher overall development center rating than males.

Research Question 7. Is there a significant difference between male and female participants on competency performance?

Hypothesis 7a. Females will have a higher rating than males on the Driven Practical Operator, True Collaboration, and Enterprise Steward competencies.

Hypothesis 7b. Males will have a higher rating than females on the Courageous Marketplace Leader and Enterprise Citizen competencies.

Research Question 8. Is there a significant difference between male and female participants on exercise performance?

Hypothesis 8a. Females will have a higher rating than males on the Conflict Resolution and Employee Conversation exercises.

Hypothesis 8b. Males will have a higher rating than females on the Board Presentation and Town Hall Presentation exercises.

Hypothesis 8c. Females will have a higher rating than males on the Strategy Discussion and Integration Planning exercises.

Method

Participants

Participants were 199 executives (126 men, 73 women) from six development center

cohorts. Centers were conducted between November 2015 and October 2017. At the time of their

involvement in the center, each participant was employed in various executive positions within

different divisions of a large healthcare organization. Participants did not volunteer to participate

in the development center, but rather were selected into a cohort based upon supervisor referral

or nomination. Cohorts were comprised of anywhere between 27 and 36 participants.

Measures

Participants were assessed on a variety of exercises and leadership competencies. The

same exercises and competency areas were generally used across each center and participating

cohort. Descriptions of exercises on which all participants were assessed are provided in Table 1.

Descriptions of the leadership competencies are provided in Table 2.

Table 1

Development Center Exercises

Exercise	Description
Board Presentation	Presentation
Strategy Discussion	Leaderless Group Discussion
Employee Conversation	Role Play
Integration Planning	Leaderless Group Discussion
Conflict Resolution	Role Play
Town Hall Presentation	Presentation

Table 2

Development Center Competencies

Leadership Competency	Description
Driven Practical Operator	Drives and implements solutions that advance the entire enterprise; Fosters and focuses on execution and practical implementation; Drives processes that support change and innovation.

Courageous Marketplace Leader	Challenges the status quo; Explores and fosters creative and innovative solutions; Clearly and courageously articulates the need for change; Calculates the risk and reward to the enterprise.
Enterprise Citizen	Understands, anticipates, and analyzes the full scope of decisions and potential actions that will impact the entire enterprise; Establishes priorities based on enterprise-wide needs; Balances internal and external needs in decision-making.
True Collaboration	Establishes positive relationships and creates synergies that span across business segments; Seeks input from others; Leverages logic and influence to inspire others to action; Communicates authentically both internally and externally.
Thoughtful Enterprise Steward	Understands the importance of human capital; Takes accountability for talent development and retention; Coaches and empowers others to reach their full potential.

Development Center Design and Scoring Protocol

The centers were designed to measure performance at a level above that of the participants' current position in the organization. Centers were conducted over a three-day period, with two days of assessment and one day dedicated to feedback and development planning. Exercises consisted of presentations (based on a business case), leaderless group discussions, and role plays. Three leadership competencies were evaluated within each exercise, such that each competency was evaluated three or four times across different exercises. Table 3 provides a summary of the leadership competencies evaluated by each exercise.

There were at least two observers per exercise. Observer teams were comprised of a mix of internal and external individuals. All observers participated in a half-day training session prior to the start of the center. Observers rotated, such that participants were evaluated from multiple assessments by different observers in multiple exercises. During each exercise, observers consulted behavioral marker guides and indicated whether they observed the participant exhibit a given behavior during that exercise. Following each exercise, observers independently evaluated

the participant's proficiency on specific leadership competencies on a scale of 1 (development

area) to 5 (exceptional). Observers then calibrated and adjusted their independent ratings to

establish consistency, so that all ratings fell within one point of each other. Observer ratings were

captured for every competency measured within each exercise, although no exercise assessed

every competency. Observer ratings were aggregated and averaged to determine an overall rating

for each exercise and for each competency.

Table 3

Development Center Competency by Exercise Matrix

	Leadership Competencies				
Exercises	Driven Practical Operator	Courageous Marketplace Leader	Enterprise Citizen	True Collaboration	Thoughtful Enterprise Steward
Board Presentation	*	*	*		
Strategy Discussion		*	*	*	
Employee Conversation	*			*	*
Integration Planning	*	*			*
Conflict Resolution			*	*	*
Town Hall Presentation	*		*	*	

* denotes a competency assessed by that exercise.

Procedure

There is one master database for this research study. Data from each of the 6 centers was

extracted from 6 separate databases, and then cleaned, merged, and sorted into one master

database. This database contains observer ratings of participants for each exercise and

competency, an overall composite score for each participant, and participant gender.

Results

Research Question 1. Which exercise has the greatest impact on overall performance in each competency area?

Research Question 1 sought to examine the unique contribution of each of the exercises in each of the five competency areas. Results of several multiple regression analyses indicated that the Employee Conversation exercise (role play) had the greatest impact on performance in three of the five competencies assessed in the center. The Board Presentation exercise and Town Hall Presentation exercise (presentations) had the greatest impact on the remaining two competencies, respectively. The Strategy Discussion exercise and the Integration Planning exercise (leaderless group discussions) contributed the least on any of the competencies that were assessed. Refer to Tables 4-8 for results.

Table 4

Exercise Contributions to the Driven Practical Operator Competency

Exercise	Beta	Part
Employee Conversation	.441	.419
Conflict Resolution	.401	.393
Board Presentation	.400	.383
Integration Planning	.385	.375

Note. $R^2 = .975$; $p<.001$

Table 5

Exercise Contributions to the Courageous Marketplace Leader Competency

Exercise	Beta	Part
Board Presentation	.501	.479
Strategy Discussion	.482	.444
Integration Planning	.455	.438

Note. $R^2 = 1.00$; $p<.001$

Table 6

Exercise Contributions to the Enterprise Citizen Competency

Exercise	Beta	Part
Town Hall Presentation	.413	.391
Board Presentation	.396	.379
Conflict Resolution	.388	.366
Strategy Discussion	.342	.325

Note. $R^2 = .968$; $p<.001$

Table 7

Exercise Contributions to the True Collaboration Competency

Exercise	Beta	Part
Employee Conversation	.490	.479
Town Hall Presentation	.406	.378
Conflict Resolution	.386	.363
Strategy Discussion	.346	.334

Note. $R^2 = .963$; $p<.001$

Table 8

Exercise Contributions to the Enterprise Steward Competency

Exercise	Beta	Part
Employee Conversation	.559	.540
Town Hall Presentation	.521	.516
Integration Planning	.434	.418

Note. $R^2 = 1.00$; $p<.001$

Research Question 2. Which exercise has the greatest impact on overall performance in the

development center?

Research Question 2 sought to examine the unique contribution of each of the exercises

on overall center performance. A standard multiple regression of overall center performance on

the center exercises indicated that the Conflict Resolution exercise ($\beta=.35$, $p<.001$) had the

greatest impact on the overall score a participant receives in the development center. The results

are presented in Table 9.

Table 9

Exercise Contributions to Overall Score in the Development Center

Exercise	Beta	Part
Conflict Resolution	.349	.324
Town Hall Presentation	.308	.281
Employee Conversation	.297	.281
Board Presentation	.263	.239
Integration Planning	.252	.230
Strategy Discussion	.252	.218

Note. $R^2 = 1.00$; $p<.001$

Research Question 3. Which competency has the greatest impact on overall performance in the

development center?

Research Question 3 sought to examine the unique contribution of each of the

competencies on overall center performance. A standard multiple regression of overall center

performance on the center competencies indicated that the True Collaboration competency

($\beta=.28$, $p<.001$) had the greatest impact on the overall score an individual receives in the

development center. The results are presented in Table 10.

Table 10

Competency Contributions to Overall Score in the Development Center

Competency	Beta	Part
True Collaboration	.279	.212
Enterprise Steward	.257	.195
Enterprise Citizen	.255	.178
Driven Practical Operator	.246	.175
Courageous Marketplace Leader	.243	.184

Note. $R^2 = 1.00$; $p<.001$

Research Question 4. Are there any differences in performance on any of the competency areas

between cohorts that have participated in the center?

Research Question 4 sought to examine whether there are any significant differences in

performance between any of the cohorts on any of the competencies. A one-way MANOVA

indicated no main effect for cohort on competency performance, $F(25, 703) = 1.43$, $p=.08$; Wilk's $\Lambda=.83$, partial $\eta^2=.04$. Thus, the cohorts did not significantly differ from one another on performance in any of the competency areas.

Research Question 5. Are there any differences in performance on any of the exercises between cohorts that have participated in the center?

Research Question 5 sought to examine whether there are any significant differences in performance between any of the cohorts on any of the exercises. A one-way MANOVA indicated no main effect for cohort on exercise performance, $F(30, 754) = 1.29$, $p=.14$; Wilk's $\Lambda=.82$, partial $\eta^2=.04$. Thus, the cohorts did not significantly differ from one another on performance on any of the exercises.

Research Question 6. Is there a significant difference between male and female participants on overall center performance?

Hypothesis 6a. Females will have a higher overall development center rating than males.

Hypothesis 6a proposed that females' overall development center rating would be higher than males' rating. Results of an independent samples t-test provided support for this hypothesis, $t(197) = -2.00$, $p=.02$ (one-tailed). The magnitude of the differences in the means was small, $\eta^2=.02$. Means and standard deviations are provided in Table 11.

Table 11

Gender Differences in Overall Center Performance

Gender	n	M	SD
Male	126	2.47	.37
Female	73	2.58	.35

Research Question 7. Is there a significant difference between male and female participants on competency performance?

Hypothesis 7a. Females will have a higher rating than males on the Driven Practical Operator, True Collaboration, and Enterprise Steward competencies.

Results of independent samples t-tests provided partial support for this hypothesis. Females scored significantly higher on the Driven Practical Operator competency than males, $t(197) = -1.93$, $p=.03$ (one-tailed). Additionally, females scored significantly higher on the True Collaboration competency than males, $t(197) = -2.29$, $p=.01$ (one-tailed). Furthermore, while females scored higher than males on the Enterprise Steward competency, the difference was not large enough to warrant significance, $t(197) = -1.18$, $p=.12$ (one-tailed). Means and standard deviations are presented below in Table 12.

Table 12

Gender Differences in DPO, TC, and ES Competencies

	Males		Females		
	M	SD	M	SD	t
Driven Practical Operator	2.53	.47	2.65	.41	-1.93*
True Collaboration	2.52	.49	2.69	.53	-2.29**
Enterprise Steward	2.42	.49	2.50	.44	-1.18

* $p \le .05$; **$p \le .01$

Hypothesis 7b. Males will have a higher rating than females on the Courageous Marketplace Leader and Enterprise Citizen competencies.

Results of independent samples t-tests did not provide support for this hypothesis. There was no significant difference in performance for the Courageous Marketplace Leader competency, $t(197) = -1.61$, $p=.06$ (one-tailed). There was also no significant difference in performance for the Enterprise Citizen competency, $t(197) = -.67$, $p=.25$ (one-tailed). Means and standard deviations are presented in Table 13.

Table 13

Gender Differences in CML and EC Competencies

	Males		Females		
	M	*SD*	*M*	*SD*	*t*
Courageous Marketplace Leader	2.41	.45	2.51	.44	-1.61
Enterprise Citizen	2.47	.47	2.51	.47	-.67

Research Question 8. Is there a significant difference between male and female participants on exercise performance?

Hypothesis 8a. Females will have a higher rating than males on the Conflict Resolution and Employee Conversation exercises.

Results of independent samples t-tests did not support this hypothesis. There was no significant difference in performance on the Conflict Resolution exercise, $t(197) = -.91$, $p=.18$ (one-tailed). There was also no significant difference in performance on the Employee Conversation exercise, $t(197) = .21$, $p=.42$ (one-tailed). Means and standard deviations are presented in Table 14.

Table 14

Gender Differences in Conflict Resolution and Employee Conversation Exercises

	Males		Females		
	M	*SD*	*M*	*SD*	*t*
Conflict Resolution	2.43	.78	2.54	.77	-.91
Employee Conversation	2.59	.70	2.57	.59	.21

Hypothesis 8b. Males will have a higher rating than females on the Board Presentation and Town Hall Presentation exercises.

Results of independent samples t-tests did not support this hypothesis. There was no significant difference in performance on the Board Presentation exercise, $t(197) = -1.10$, $p=.14$ (one-tailed). There was also no significant difference in performance on the Town Hall

Presentation exercise, $t(197) = -1.07$, $p=.14$ (one-tailed). Means and standard deviations are

presented in Table 15.

Table 15

Gender Differences in Board Presentation and Town Hall Presentation Exercises

	Males		Females		
	M	*SD*	*M*	*SD*	*t*
Board Presentation	2.48	.60	2.57	.56	-1.10
Town Hall Presentation	2.51	.71	2.62	.64	-1.07

Hypothesis 8c. Females will have a higher rating than males on the Strategy Discussion and

Integration Planning exercises.

Results of independent samples t-tests provided partial support for this hypothesis.

Females scored significantly higher than males on the Strategy Discussion exercise, $t(197) = -$

2.26, $p<.01$ (one-tailed). However, there was no significant difference in performance on the

Integration Planning exercise, $t(197) = -1.56$, $p=.06$ (one-tailed), although the difference was

approaching significance. Means and standard deviations are presented in Table 16.

Table 16

Gender Differences in Strategy Discussion and Integration Planning Exercises

	Males		Females		
	M	*SD*	*M*	*SD*	*t*
Strategy Discussion	2.36	.53	2.54	.60	-2.26*
Integration Planning	2.46	.54	2.58	.59	-1.56

*$p \leq .01$

Discussion

Gender Implications

The most notable findings from this study relate to differences between men and women in terms of overall development center performance, in addition to performance within certain competency areas. The present study found that women performed significantly better overall than men in the development center, akin to the findings of prior studies with females performing better overall than males in a center context (Schmitt & Hill, 1977; Walsh et al., 1987; Shore, 1992; Schmitt, 1993; Weijerman & Born, 1995; Anderson et al., 2006; Falk & Fox, 2014). Furthermore, findings of the present study aligned with role congruity theory, which posits that people are expected to engage in activities, and perform and behave in ways consistent and stereotypical to their gender roles (Eagly & Karau, 2002; Ritter & Yoder, 2004). Specifically, women were rated significantly higher than men on competencies reflecting an interpersonal-oriented and relationship-building nature. However, men were not rated significantly higher than women on competencies reflecting an analytical, task-oriented nature. Additionally, women outperformed men on exercises reflecting a relationship-oriented focus and agenda, but the differences were not large enough to warrant significance. Nevertheless, these findings are aligned with gender-based stereotypes of men and women.

Exercise-Competency Relationship Implications

A component of the present study sought to determine which exercise made the strongest, unique contribution to explaining performance in each of the five competencies in the development center. As a precursor to discussing these findings, it's important to note that the rating a participant receives on a competency is derived solely from their performance on exercises in which that competency is evaluated. Thus, the strongest contributing exercise for

each competency is largely driven by whichever exercise that participants tended to be rated or

scored highest on. There is little, if any, variance that is unaccounted for in these regression

models, as the only possible predictors (exercises) that can be used to explain performance

(competencies) are accounted for and utilized in the model. Nevertheless, additional notable

findings from this study involve the relationships between the development center exercises and

competency areas.

In assessing exercise-competency relationships in the current study, the Employee

Conversation exercise (role play) made the strongest, unique contribution to performance in

three of the five development center competencies. Additionally, the Board Presentation and

Town Hall Presentation exercises (presentations) made the strongest, unique contribution to

performance in the remaining two competencies, respectively. The role play exercises made the

strongest unique contributions to the three competencies with a relationship-oriented focus, while

the presentation exercises made the strongest unique contributions to the two competencies with

an analytical, task-oriented focus. Overall, the role play exercises are making the strongest,

unique contribution to overall development center performance. Participants are consistently

scoring highest on role play-type exercises. While significant, the Strategy Discussion and

Integration Planning exercises (leaderless group discussions) made the weakest contribution to

performance across all competencies, and to overall center performance. Leaderless group

discussions are among the most difficult to observe of center exercises, as the contribution of

each participant in the exercise is difficult to identify and measure. Participants are consistently

receiving lower scores in these exercises, in comparison to the role plays and presentations,

where they are the focal point of the simulation. Furthermore, the results indicate no significant

differences in performance in exercises or competencies, and in overall performance, across all

cohorts that have participated in the development center. These findings provide support for consistencies in rating, evaluating, and scoring participants across time. However, these results are also a cause for concern. Consistencies in ratings across time, in addition to the minimal variance observed in performance between gender and in the contributions of each exercise to competency performance, lend themselves to some potential methodological concerns in the design and structure of the development center.

Development Center Design Implications

The findings in the present study provide insight into some speculated methodological limitations of the design of the development center. While all exercises significantly contribute to each of the respective competencies that they assessed, there is minimal variance between the contributions of each exercise. Furthermore, while there are significant gender differences in overall performance and in performance within certain competencies, the effect sizes and variance is small. This may stem from an issue with the rating scale used to evaluate participants during exercises. The rating scale used is a 5-point scale, with lower ratings indicating room for development for the participant, and higher ratings indicating exceptional performance. However, it's not necessarily clear what the difference is between rating a participant as a 4 or 5. Because it's a 5-point rating scale, observers may be more prone to assign average ratings, resulting in a "middling effect," which would explain the minimal variance between exercise contributions to performance and gender differences in performance. Perhaps implementing a 4-point rating scale instead, with 1 indicating a failure to meet expectations and 4 indicating exceeding expectations, would increase variance and allow for providing richer feedback to participants on their performance. The issue with minimal variance may also stem from lack of clarity in the behavioral markers used to rate and provide final scores to participants on each

competency. Furthermore, behaviors performed in each exercise don't always fall explicitly into one of the competencies assessed by that exercise. This can result in observers mistakenly assigning points, and thus a higher rating, to a given competency area for a participant, thereby inflating their performance in that competency. Defining the behavioral markers in more simplistic and explicit terms may result in more accurate observations and evaluations of participant behavior and performance, thus reducing these errors. Additionally, the center in this study utilized a standardized schedule of exercises across all cohorts, such that the same exercises occurred in the same exact order during each center. This may have contributed to observer and participant fatigue, particularly in the exercises that occurred last, as the order of the exercises was never counter-balanced. As a result, scores and performance on those exercises might be negatively or positively influenced by such fatigue.

Limitations and Future Research

While the present study found significant differences in performance between men and women, these findings are not without limitations. While the sample size was respectable, there were far more men (123) than women (76) in the study, which may have contributed to significant differences in performance favoring women. With almost 50 more men than women in the study, there was a greater chance for variability in performance in men, lending the advantage in performance to women. Additional limitations of the study concern the data itself, limiting the questions that could be tested and analyzed. Only data on participant gender, cohort to which the participant belonged, exercise performance, and competency performance was provided. As previously mentioned, differences in development and/or assessment center performance may be influenced by cognitive ability and individual personality attributes (Collins et al., 2003). However, these variables were only assessed at the overall center performance

level. Future research might benefit from examining how cognitive ability, personality traits, and

even performance criteria within the organization correlate with and/or predict performance at

the exercise and competency level, in addition to overall center performance. Examining the

relationship between past performance, performance within the development center, and future

performance may provide support for the effectiveness of the center in skill and competency

development. Nevertheless, results of this study contribute to the existing literature about gender

differences in development center performance, as well as the relationship between exercises and

the competencies they aim to assess.

References

Adams, K. (1995). Development centers. *IRS Employment Review*, *596*, 16.

Anderson, L., & Thacker, J. (1985). Self-monitoring and sex as related to assessment center ratings and job performance. *Basic and Applied Social Psychology*, *6*, 345-361.

Anderson, N., Lievens, F., van Dam, K., & Born, M. (2006). A construct-driven investigation of gender differences in a leadership-role assessment center. *Journal of Applied Psychology*, *91*(3), 555-566.

Appelbaum, S. H., Harel, V., & Shapiro, B. (1998). The developmental assessment center: The next generation. *Career Development International*, *3*(1), 5-12.

Arthur, W., Day, E. A., McNelly, T. L., & Edens, P. S. (2003). A meta-analysis of the criterion-related validity of assessment center dimensions. *Personnel Psychology*, *56*, 125-154.

Atwater, L., & Roush, P. (1994). An investigation of gender effects on followers' ratings of leaders, leaders' self-ratings and reactions to feedback. *The Journal of Leadership Studies*, *1*(4), 37-52.

Bell, S. T., & Arthur, W. (2008). Feedback acceptance in developmental assessment centers: The role of feedback message, participant personality, and affective response to the feedback session. *Journal of Organizational Behavior*, *29*, 681-703.

CEB (2013). *Guidelines for best practice in the use of assessment and development centers*.

Collins, J. M., Schmidt, F. L., Sanchez-Ku, M., Thomas, L., McDaniel, M. A., & Le, H. (2003). Can basic individual differences shed light on the construct meaning of assessment center evaluations? *International Journal of Selection and Assessment*, *11*, 17-29.

Chan, D. (1996). Criterion and construct validation of an assessment center. *Journal of Occupational and Organizational Psychology*, *69*, 167-181.

Dean, M. A., Roth, P. L., & Bobko, P. (2008). Ethnic and gender subgroup differences in

    assessment center ratings: A meta-analysis. *Journal of Applied Psychology*, *93*(3), 685-

    691.

Eagly, A. H., & Johannesen-Schmidt, M. C. (2001). The leadership styles of women and men.

    Journal of Social Issues, *57*(4), 781-797.

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders.

    Psychological Review, *109*, 573-598.

Engelbrecht, A. S., & Fischer, A. H. (1995). The managerial performance implications of a

    developmental assessment center process. *Human Relations*, *48*(4), 387-404.

Falk, A., & Fox, S. (2014). Gender and ethnic composition of assessment centers and its

    relationship to participants' success. *Journal of Personnel Psychology*, *13*(1), 11-20.

Holbeche, L. (1995). Peering into the future of careers. *People Management*, May, 26-28.

Jones, R. G., & Whitmore, M. D. (1995). Evaluating developmental assessment centers as

    interventions. *Personnel Psychology*, *48*, 377-388.

Moses, J. L. (1973). The development of an assessment center for the early identification of

    supervisory potential. *Personnel Psychology*, *26*, 569-580.

Moses, J. L., & Boehm, V. R. (1975). Relationship of assessment center performance to

    management progress of women. *Journal of Applied Psychology*, *60*, 527-529.

Ritchie, R., & Moses, J. L. (1983). Assessment center correlates of women's advancement into

    middle management: A 7-year longitudinal analysis. *Journal of Applied Psychology*,

    68(2), 227-231.

Ritter, B. A., & Yoder, J. D. (2004). Gender differences in leader emergence persist even for

dominant women: An updated confirmation of role congruity theory. *Psychology of*

Women Quarterly, *28*(3), 187-193.

Rupp, D. E., Gibbons, A. M., Baldwin, A. M., Snyder, L. A., Spain, S. M., Woo, S. E, & Kim,

M. (2006). An initial validation of developmental assessment centers as accurate

assessments and effective training interventions. *The Psychologist-Manager Journal*,

9(2), 171-200.

Rupp, D. E., Snyder, L. A., Gibbons, A. M., & Thornton, G. C. (2006). What should

developmental assessment centers be developing? *The Psychologist-Manager Journal*,

9(2), 75-98.

Rupp, D. E., Thornton, G. C., & Gibbons, A. M. (2008). The construct validity of the assessment

center method and usefulness of dimensions as focal constructs. *Industrial and*

Organizational Psychology, *1*, 116-120.

Schmitt, N., & Hill, T. E. (1977). Sex and race composition of assessment center groups as a

determinant of peer and assessor ratings. *Journal of Applied Psychology*, *62*, 261-264.

Schmitt, N. (1993). Group composition, gender, and race effects on assessment center ratings. In

H. Schuler, J. L. Farr, & M. Smith (Eds.), *Personnel selection and assessment: Individual*

and organizational perspectives (pp. 315-322). Hillsdale, NJ: Erlbaum.

Shore, T. H., Tashchian, A., & Adams, J. S. (1997). The role of gender in a developmental

assessment center. *Journal of Social Behavior and Personality*, *12*(5), 191-203.

Thornton, G. C., & Rupp, D. E. (2005). *Assessment centers in human resource management*.

Reading, MA: Addison-Wesley Publishing Company.

Thornton, G. C., & Krause, D. E. (2009). Selection versus development assessment centers: An

international survey of design, execution, and evaluation. *The International Journal of*

Human Resource Management, *20*(2), 478-498.

Walsh, J. P., Weinberg, R. M., & Fairfield, M. L. (1987). The effects of gender on assessment

center evaluations. *Journal of Occupational Psychology*, *60*, 305-309.

Weijerman, E., & Born, M. (1995). The relationship between gender and assessment center

scores. *Gender, Work and Organization*, *8*, 284-292.