

A TALE OF FOUR CITIES: EVALUATING THE IMPACT OF ASSESSMENT
CENTERS ON POLICE PROMOTION PROCESSES BY RANK, SEX, AND RACE

by

Kevin T. Leverenz, M.S.

A dissertation submitted to the Graduate Council of
Texas State University in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
with a Major in Criminal Justice
August, 2019

Committee Members:

Christine S. Sellers, Co-Chair

Brian L. Withrow, Co-Chair

Wesley G. Jennings

Kenneth J. Peak

COPYRIGHT

by

Kevin T. Leverenz

2019

FAIR USE AND AUTHOR'S PERMISSION STATEMENT

Fair Use

This work is protected by the Copyright Laws of the United States (Public Law 94-553, section 107). Consistent with fair use as defined in the Copyright Laws, brief quotations from this material are allowed with proper acknowledgement. Use of this material for financial gain without the author's express written permission is not allowed.

Duplication Permission

As the copyright holder of this work I, Kevin T. Leverenz, authorize duplication of this work, in whole or in part, for educational or scholarly purposes only.

DEDICATION

I dedicate this work to my wife, Jamie, and my son, Ethan, for their support and understanding throughout this endeavor. I also dedicate this work to my parents, who instilled in me an understanding of the value of an education, a tenacious work ethic, the audacity to undertake this endeavor and the undaunted resolve to see it to fruition.

ACKNOWLEDGEMENTS

I would like to thank the members of my dissertation committee for their support, feedback and encouragement: Christine Sellers, Brian Withrow, Wesley Jennings, and Kenneth Peak. Your efforts, individually and collectively, have not gone unnoticed, for without them, I would not have possibly attained this milestone in my pursuit of educational excellence.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS.....	v
LIST OF TABLES.....	viii
LIST OF ABBREVIATIONS.....	ix
ABSTRACT.....	x
 CHAPTER	
I. INTRODUCTION.....	1
II. POLICE LEADERSHIP AND PROMOTION PROCESSES.....	8
What Makes a High-Quality Police Leader?	8
Defining and Developing a Cadre of Quality Leaders.....	10
Succession Planning.....	15
Drivers of Attrition/Turnover.....	16
Selecting Leaders for Promotion.....	17
Diversity and Subgroup Differences.....	19
Legal Defensibility.....	21
Laws Relevant to Diversity and Promotional Processes.....	24
Civil Rights Act of 1964.....	25
Civil Rights Act of 1991.....	26
Texas Local Government Code 143.....	28
Disparate Impact.....	30
Case Law.....	31
The Assessment Center (AC).....	33
AC Validity.....	37
The AC Process in Practice.....	38
Promotions, Written Exam vs. AC, and Protected Classes.....	40
III. METHODOLOGY.....	46
Hypotheses.....	46
Data Collection and Sample.....	46
Measurement of Variables.....	48
Dependent Variables.....	49

Independent Variables.....	50
Data Analysis Plan.....	50
IV. RESULTS.....	52
Descriptive Statistics.....	52
Race Differences	53
Sex Differences	56
V. DISCUSSION.....	58
Race Differences	58
Sex Differences	60
Limitations	62
Implications	63
Reactive Risk Management – Legal Defensibility	64
Proactive Risk Management – Legal Defensibility	67
REFERENCES.....	88

LIST OF TABLES

Table	Page
1a. Cities included in Dataset.....	70
1b. Population Demographic Data by City.....	70
1c. Sworn Staffing at Sought-After Positions in Current Research	71
2a. Promotional events (cohorts) by department	72
2b. Participants by jurisdiction	72
2c. Promotional events (cohorts) by rank	72
2d. Participants by rank	72
2e. Raw score descriptive statistics by cohort – Captain / Commander	73
2f. Raw score descriptive statistics by cohort – Lieutenant	74
2g. Raw score descriptive statistics by cohort – Sergeant	75
3. Variables	76
4. Captain/Commander Test Takers: <i>Descriptive Statistics</i>	77
5. Lieutenant Test Takers: <i>Descriptive Statistics</i>	78
6. Sergeant Test Takers: <i>Descriptive Statistics</i>	79
7. Race differences in scores	80
8. Race differences in rank-ordered placement	81
9. Race differences in quartile placement	82
10. Sex differences in scores	84
11. Sex differences in rank-ordered placement	85
12. Sex differences in quartile placement	86

LIST OF ABBREVIATIONS

Abbreviation	Description
AC	Assessment Center
DAC	Developmental Assessment Center
DC	Development Center
EEOC	Equal Employment Opportunity Commission
EI	Emotional Intelligence
IACP	International Association of Chiefs of Police
I/O	Industrial/Organizational
KSAs	Knowledge, Skill Sets, and Abilities
LGC 143	Local Government Code 143
OAR	Overall Assessment Rating
SME	Subject Matter Expert

ABSTRACT

In an effort to meet the dual interests of developing a diverse cadre of quality police leaders while reducing mean subgroup differences in results, the use of the assessment center has emerged as a capable instrument and has grown in popularity over the last few decades. However, a gap in empirical research has been revealed. So, the current research focuses on how ACs impact outcomes of promotion process participants belonging to federally-protected classifications (race and sex) in four major Texas police departments, and how those outcomes change in the absence of the AC.

I. INTRODUCTION

In an effort to deliver quality police service, law enforcement agencies have long sought to build trust in the communities they serve. Organizational leaders seek to develop and promote those who are found most apt to attain successful ends on behalf of the organization (Bishopp, 2013). Given the importance of their need for legitimacy in the public's view, law enforcement agencies across the United States strive to develop and employ equitable promotion processes capable of delivering on their need for diverse cadres of quality leaders. The Austin (Texas) Police Department, as an example, is no exception. In the late 1990s, two factors drove the Austin Police Department's administration, and the rank-and-file alike, to face this issue head-on. First, following a challenge to a 1996 promotion exam for adverse impact against racial minority promotional candidates, a civil suit was filed in 1997 to seek injunctive relief and to challenge the promotional process for being conducted without, first, validating the test in accordance with Equal Employment Opportunity Commission (EEOC) guidelines (*Bullock et al. v. City of Austin et al.*, 1997). Second, police administrators and city officials began collective bargaining negotiations (as authorized by Texas Local Government Code 174, established in 1993), in which city administrators introduced the idea that an assessment center (AC) should be added to the existing promotion process for supervisory and managerial ranks (B. Mills, personal communication, August 31, 2016). The AC was presented by police leaders as an additional assessment tool which would, along with the written exam, provide for a more well-rounded promotion process compared to a written exam alone (B. Mills, personal communication, August 31, 2016). However, the AC was also presented as an assessment tool that would produce results

“more favorable to minorities” (V. Escobar, personal communication, August 29, 2016).

This assertion seemed to align with Assistant Chief Mike McDonald’s contention that a promotion process consisting of only a written exam was unfair due to disparities in access to quality education (*Bullock et al. v City of Austin et al.*, Responses to Interrogatories, 1997). This was an example of the dual issues of quality and diversity that have concerned police administrators across the nation. Austin Police Department administrators and representatives for the department’s rank-and-file sought to address these issues with the first Meet and Confer contract, which included the addition of an assessment center to the promotional processes for the supervisory and middle management and upper management ranks (*Agreement between the City of Austin and the Austin Police Department, 1998*).

In the wake of legal mandates established by federal civil rights legislation and case law, a primary focus of those charged with personnel selection process development in the years to come will continue to be the need to improve workforce diversity and mitigate adverse impact on those belonging to federally-protected classifications (De Soete, Lievens, Oostrom, & Westerveld, 2013). Federally-protected classifications subject to protections through EEOC enforcement efforts include racial/ethnic minorities and females; adverse impact is defined as “a substantially different rate of selection in hiring, promotion, or other employment decision which works to the disadvantage of members of a race, sex, or ethnic group” (Equal Employment Opportunity Commission, 1978, Section 16. B). However, adverse impact is a function of both the significance of subgroup differences and the ratio of selection or promotion (Roth, Van Iddekinge,

Huffcutt, Eidson, & Bobko, 2002). Higher subgroup differences accompanied by lower ratios of selection or promotion result in greater adverse impact (Roth et al., 2002).

By resolving to identify and promote a diverse cadre of the most capable police leaders, many police administrators acknowledge these dual interests as necessary realities. Along with this understanding is the acceptance that it was necessary to determine the best approach to meeting both interests. Validity studies on assessment instruments have determined that the best predictors of future performance are written tests designed to assess participants' cognitive ability (Schmidt & Hunter, 1998). However, research findings have also indicated that AC's are effective in predicting future performance in an array of criteria (Byham, 1970; Hinrichs, 1978; Hsin-Chih, 2006; Thornton, Wilson, Johnson, & Rogers, 2009).

Because some police promotional process participants belong to federally-protected classifications, it has become clear that working to mitigate or eliminate disparate outcomes for those sharing a membership in any of the federally-protected classifications is necessary in order to avoid federal sanctions and lawsuits. In short, there is no alternative assessment instrument that has been substantially proven to be as predictive of future performance as general intelligence tests that is also capable of reducing adverse impact (Wax, 2012). So, this problem can be proactively addressed by either employing assessment methods that reduce adverse impact on participants who are members of protected classifications or employing an affirmative action program designed to bolster representation by racial minorities and women (Kravitz, 2008).

In the backdrop, Texas Local Government Code 143 (civil service law) outlines a police promotion process providing only for a written exam to be used to establish

promotion eligibility lists. This provision complicates agency executives' ability to incorporate assessment instruments, like the AC, into police promotional processes beyond a written exam. While civil service law was established to quell political influence in employment decisions affecting police officers, it fails to account for substantial subgroup differences in test results that have disparately impacted promotional process participants who are members of federally-protected classifications. Now, in Texas, any deviation from civil service promotional process guidelines that may, perhaps, include an assessment instrument that results in less adverse impact requires adoption of an alternative promotion process (Texas Local Government Code, 2015, Section 143.035(f)). This suggests that police administrators who operate under civil service law are required to adopt "alternative promotional processes" in accordance with civil service guidelines. These guidelines require that a municipality's civil service commission can only adopt an alternative promotion process after both the chief of police in that municipal police department recommends the alternative process and the majority of the voting sworn employees in the department vote for its adoption (Texas Local Government Code, 2015, Section 143.035(b)). Moreover, the alternative process may only be rescinded after the process has been in place for 180 days when either the chief of police makes his or her request to rescind the process to the civil service commission or with a petition from 35 percent of the municipality's sworn police force requesting to rescind it along with a subsequent majority vote from the voting sworn employees to rescind the process (Texas Local Government Code, 2015, Section 143.035(h-i)). Beyond that, only police departments who do not operate under Texas civil service law may adopt an AC as part of their police promotional process.

Many police administrators have been educated to understand that the AC is an assessment tool capable of delivering both diversity and capacity in the cadre of emerging police leaders. The practice of utilizing multiple assessors to evaluate leadership capability in complex scenarios first began in the early twentieth century to cultivate and select military leaders (Hsin-Chih, 2006). Moreover, ACs have been used in the private sector to identify management prowess for decades (Thornton & Rupp, 2006). Proponents have heralded that the AC, as a part of the police promotion process, is an assessment tool capable of providing an agency with just a diverse, quality police leadership cadre (Thornton & Rupp, 2006; Wax, 2012). When compared to other assessment instruments employed to establish promotion eligibility lists, practitioners have purported that ACs have proven, over time, to most accurately predict future performance while providing for a more diverse management staff (Wax, 2012). According to Peak and Giacomazzi (2019), an AC is “a process used for promoting and hiring personnel that may include oral interviews, psychological tests, group and in-basket exercises, and writing and role-playing exercises” (p. 112). The AC is designed to engage participants by administering exercises with scenarios that mimic situations that incumbents encounter in the sought-after position in an effort to provide participants with opportunities to respond by employing the knowledge, skills, and abilities predetermined to be appropriate for incumbents to employ in the sought-after position (Cordner & Scarborough, 2007). The AC method has been designed to accomplish different functions. According to Thornton and Rupp (2006), the AC method can be tailored to select and promote employees, diagnose their shortcomings, and develop their capabilities. ACs, while secondarily employed to develop participants’ management

capacities, are primarily used to select or promote participants (Hermelin et al., 2007). In fact, when ACs first emerged, they were employed as a selection tool to determine who was most likely to succeed and, likewise, to determine who was most likely to fail (Moses, 2008).

Empirical research on assessment centers varies considerably in its focus. Several studies have examined AC validity (De Soete et al, 2013; Gaugler, Rosenthal, Thornton, & Bentson, 1987; Gibbons & Rupp, 2009; Hermelin et al., 2007; Hsin-Chih, 2006; Kleinmann, & Köller, 1997; Thornton & Gibbons, 2009; Whetzel, Rotenberry, & McDaniel, 2014; Woehr & Arthur, 2003). Other studies have explored racial and sex subgroup differences in assessment instrument performance (Arthur, Edwards, & Barrett, 2002; Dean, Roth, & Bobko, 2008; Hough, Oswald, & Ployhart, 2001; Roth et al., 2002; Schleicher, Van Iddekinge, Morgeson, and Campion, 2010). Research has also examined factors (based on participants' race or sex) affecting promotion motivation, viability, representation and performance capability (Anderson, Lievens, van Dam, & Born, 2006; Archbold & Hassell, 2009; Bishopp, 2013; Gau, Terrill, & Paoline, 2013; Gustafson, 2013; Mano-Negrin & Sheaffer, 2004), whether or not promotional test performance was predictable over time (Topp, 2011), and the extent to which performance could be predicted by executive-level competency ratings (Russell, 2001). None of the existing research sought to uncover the value an AC added to a promotional process regarding the promotion of a diverse cadre of police leaders across ranks and agencies over time.

The current research focuses on how ACs impact outcomes of promotion process participants belonging to federally-protected classifications (race and sex) in four major Texas police departments, and how those outcomes change in the absence of the AC. The

gap in research is significant, given the continuing concern in the United States for any presence of adverse impact in civil service promotional processes. Additionally, there has been no research on the efficacy of the standard civil service promotional process established in Texas Local Government Code 143. The current research, while not specifically examining adverse impact in any one of the departments, will examine the differences in eligibility list placement between written examination performance, AC performance, and overall performance. In doing so, the current research will highlight performance differences among those sharing membership in federally-protected classifications and those who do not.

II. POLICE LEADERSHIP AND PROMOTION PROCESSES

To understand the value of this research, one must first gain an understanding of the value of both quality and diversity in police leadership and then appreciate research that has sought to identify ways to attain them.

What Makes a High-Quality Police Leader?

Quality police service hinges on the capacity of police managers to provide capable leadership to the rank and file. Beyond the physical risks of serving in a law enforcement capacity, there are the risks of civil liability and legitimacy in the public's view in the wake of any sign or public perception of misconduct. The field of law enforcement is fraught with danger, and it is the responsibility of each law enforcement agency to provide for leadership and management quality so as to reduce the likelihood of liability, to bolster agency credibility in the eyes of stakeholders, and to ensure the safety of both the officers and people in the communities they serve (Ortmeier & Davis, 2012). All of these administrative concerns, however, do not fully encompass the risks to the law enforcement community.

Police supervisors encounter many challenges not experienced in other fields of endeavor, due to conditions unique to the police culture and the nature of the work they do (Schafer, 2009). Thus, understanding the nature of the problem is essential to anyone seeking to address it. In the backdrop, police leaders have to mitigate the inherent personal and professional dangers that often plague the ranks of law enforcement officers; those problems have included alcoholism, drug use, divorce, suicide, and even acts of corruption, domestic violence, and civil rights violations (More, Vito, & Walsh, 2012; Swanson, Territo, & Taylor, 2012). Exceeding the suicide rate of the general

population, recent reports illustrate that 14 police officers per 100,000 take their own lives and of those, 83 percent were predominantly struggling with personal challenges (Martin & Martin, 2017). Beyond the suicide rate, police officers' lives, on average, are more than 20 years shorter than the national average for the general population (Volanti, Hartley, Gu, Fekedulegn, Andrew, & Burchfield, 2013). It is our police leaders who are poised to improve organizational conditions that effectively assist employees to combat stressors that tend to otherwise chip away at their collective ability to maintain organizational agility, or maintain the flexibility to respond effectively to changing demands of their constituencies and the law enforcement profession. Because of this, leadership plays a pivotal role in achieving sought-after outcomes in any group within an organization. Leadership is just as impactful in reaching positive organizational outcomes as an absence of leadership is at reaching negative organizational outcomes (Schafer, 2009). So, to ensure a department's success while mitigating negative outcomes (like occupational stressors and manifestations of individuals experiencing it), it is necessary to promote quality leaders.

When leadership is less than optimal, or utterly failing, the employees will begin to reveal their concerns without even trying; from bold complaints to the deafening silence of discontent, behavioral changes indicating displeasure among a large segment of a given work group will be readily apparent to all but the most naive, inattentive, or inexperienced leaders. Administrators should look for the following indicators that leadership improvement is needed: low productivity, low morale, higher stress levels, turnover, and poor attendance records (Schafer, 2009). While these are internal indicators, which could be ignored for some amount of time by administrators, it is not as

easy to ignore negative impacts to the community's perception of the agency's legitimacy. Research has shown that a community's quality of life is significantly impacted by how successful a department is at promoting quality leaders to lead officers in service to the community (Gowing, Morris, Adler, & Gold, 2008). Savvy police administrators would be wise to understand that failing leadership and/or managerial mistreatment has been found to be directly related to misconduct and behavioral problems by their officers (Hogan & Hogan, 2001). Addressing leadership failure should be considered a top priority for lasting quality police service (Schafer, 2009). In order to accomplish this, police leaders at every level in the organization must be able to define quality leadership before they could ever expect to attain it, select those most likely to succeed, and then mentor others to competently lead.

Defining and Developing a Cadre of Quality Leaders

When it becomes apparent to an agency's administration that improvements within the leadership ranks are necessary, so, too, will it become apparent that properly identifying those within the organization that possess the most potential to be quality leaders takes time, resources, and focused effort. Furthermore, leadership development should be considered a thoughtful process, and it should begin by succinctly defining leadership and then determining how to attain it. While the importance of leadership has been clear for many years, specifically defining the concept or understanding how to attain it has proven difficult (Schafer, 2009). Police employees, regardless of rank, understand the importance of quality leadership. Police officers may be able to articulate what behaviors they would exhibit to lead others, but often only in the context of their own experiences. However, because they are screened for ethical character and a capacity

for sound decision-making during the recruiting process and then trained as observers of human behavior, police officers are very capable of discovering when their own leaders are devoid of leadership prowess. Moreover, it seems reasonable to assume that subordinates, as direct recipients of their leader's influence, are properly poised to understand whether their work group's success was accomplished by their leader's actions or in spite of it. By taking a comprehensive approach to assessing what outcomes their leaders are producing and how they are successfully meeting organizational needs, utilizing all available resources to gather and develop feedback, police administrators can begin to focus their training and mentoring on quality leadership.

Once administrators determine that value rests in identifying and developing its cadre of leadership, decisions are necessary. So, when considering promotion viability, it is necessary to examine a candidate's management capacity along with leadership prowess (Bishopp, 2013). Still, developing an understanding of those managerial and leadership capacities required of leaders in a given organization is necessary. Executives must identify what makes leaders most competent among their peers before they can assess individuals' capacities for those qualities or develop them in their leaders.

Being able to distinguish leadership prowess from skillful management is vital. Unfortunately, much of the available leadership development opportunities merely focus on developing managerial competencies. Moreover, many organizations design promotion processes that are more aptly poised to stave legal liability rather than to promote effective leaders (Schafer, 2009). Ensuring that a promotional process is legally defensible is vital, but no more so than competent leadership development. Understanding the distinction between management and leadership is necessary to

properly conclude that we lead people and manage things (resources, money, projects, initiatives, documents, etc.); leadership is about relationships and relationship building – both in an organization and the larger constituencies that an organization serves.

Generally, leadership means to influence, rather than coerce, others to attainment of a given set of goals (Yukl, 2010). So, while a good manager ensures work gets done on time, a good leader will use the same work assignment to spark cooperation and team development. A good manager will check for and report progress updates on a project vital to their organization's success while a good leader will celebrate benchmark successes and recognize effort. A good manager will make sure that employees understand what needs to be done, a good leader makes sure they know why it needs to be done. The good manager sees employees as a human resource while the good leader sees employees as human capital. The good manager sees his or her direct reports as his or her employees while the good leader see them as his or her responsibility. In each example, the manager's focus is on the work while the leader's focus is on the employee. Success as a leader and as a manager, while not synonymous and markedly different, are both vital to organizational success. While these are not mutually exclusive terms, or roles, to develop competence as both a leader and a manager, a precise balance must be maintained. The most successful leaders are leaders first; they manage when they are not engaged with others because they value the engagement as a primary leadership function. In this way, the best leaders announce that they value workers and resolve to showing employees they are valued, first, by giving of themselves and devoting their time and effort to their co-workers.

Among those in leadership positions in law enforcement agencies, sergeants, lieutenants, and captains (or commanders) have specifically different responsibilities to the rank-and-file and to their respective departments. Of the three ranks, sergeant is the lowest rank level of the three. Sergeants are first-line supervisors, responsible for providing leadership while directing and coordinating rank-and-file efforts to further department prerogatives (Iannone, Iannone, & Bernstein, 2009).

Lieutenant is the next higher rank of responsibility. Lieutenants manage multiple patrol shifts or police units and will each have multiple (two to four) sergeants as direct reports. Lieutenants experience a paradigm shift in their focus; while sergeants primarily maintain focus on their direct reports, lieutenants exist to manage the interests of the department and operate on its behalf. How they organize, lead, plan, and control will directly reflect their focus, which is more on the department's legitimacy than the rank-and-file; they do so with the understanding that the rank-and-file remains the primary concern of their sergeants (Birzer, Bayens, & Roberson, 2012). Lieutenants are considered mid-level managers, responsible for overseeing the operations of those police units in his or her charge. Lieutenants impact police operations through the influence of sergeants, who typically manage direct interactions with rank-and-file officers or detectives. Only in a leadership void at the rank of sergeant will lieutenants directly manage direct day-to-day operations and interactions with the rank-and-file.

The rank of captain is the highest of the three ranks examined in the current research and represents the lowest level of upper management; this rank is the first rank represented in the ranks of police administration. The rank of commander is synonymous with that of captain. This is, further, the highest rank usually found to have civil service

protection. For this reason, it is the highest rank attained in accordance with civil service promotional processes. Higher ranks of deputy chief or assistant chief are considered at-will employment positions and assignments to those ranks are at the discretion of the chief of police. Again, the focus continues to shift towards the interests of the department, and captains (commanders) operate with those interests as a primary focus. Captains (commanders) focus on strategic planning, budgeting, organizational design, assessment, and evaluation efforts on behalf of the department (Ortmeier & Davis, 2012). Effectively managing the department's legitimacy within and beyond the community they serve, in short, becomes their responsibility as an extension of the aforementioned executive staff and the leadership they provide is strategic and organization-wide.

In order to be successful at different (higher) ranks, one must attain and become proficient at different skill sets. Identifying and correcting potential leaders' inadequacies in sought-after leadership competencies before they promote should be a part of developing a cadre of first-line supervisors. However, for each rank (from first-line supervisor, to middle management, to command staff, and then to executive staff), new skills and capacities are required. Examining the extent to which performance could be predicted by executive-level competency ratings, Russell (2001) suggested that different factors influence short-term, as opposed to long-term, success. One's capacity to perform as it relates to management competencies, which more aptly translate to managing assets, predicts initial incumbent success while one's capacity to perform as it relates to leadership competencies, which hinge on relationships and trust-building, predict performance change; leadership capacity predicts long-term success (Russell, 2001). In short, according to Russell (2001), management competency will bring short-term

success, but that success will wane without demonstrating leadership competencies. Development and promotion strategies should be different for each rank, careful to develop, and assess one's capacity in those areas unique to the sought-after rank. Regardless of the attained or sought-after rank, understanding the many facets of leadership and management development, each with its own intricacies, is the first step to solving the leadership equation.

Succession Planning

Police administrators cannot train, mentor, and coach all leaders about how to consistently replicate successful outcomes without first concisely determining what qualities and competencies their best leaders possess. However, ensuring long-term success and organizational viability requires the chief of police to take deliberate actions to establish an environment in the department that is conducive to developing a cadre of future leaders from a pool of promising performers that show leadership potential (Peak & Giacomazzi, 2016). By genuinely encouraging leadership development for long-term success, a police chief can drive a paradigm change within his or her department's culture in which desire finds opportunity, that is, in which the opportunities for growth are only limited by one's desire to seek development. Beyond formal leadership training opportunities, police leaders can "provide skill development opportunities by having those persons with leadership potential do things such as plan an event, write a training bulletin, update policies or procedures, conduct training and research, write a proposal or grant, counsel peers, become a mentor, write contingency plans, and so on" (Peak & Giacomazzi, 2016, p. 152). Moreover, this kind of culture change provides support and encouragement for young performers to seek opportunities for self-growth through

academic pursuits, community volunteerism, and scholarly articles on policing and leadership (Peak & Giacomazzi, 2016, p. 153). Succession planning also requires an understanding of what drives attrition in an organization. Only with such an understanding are police executives poised to determine how to control for organizational factors that can exacerbate turnover rates.

Drivers of Attrition / Turnover

It is vital for organizational leaders to track their organization's well-being as an essential part of succession planning. One method is maintaining open communication up and down the chain of command while trusting that competent middle managers and first-line supervisors will provide feedback when any issue begins to adversely impact the workforce. However, by maintaining a watchful eye on turnover trends (as a representation of organizational well-being), savvy executives can verify whether they are receiving a clear picture of what the workforce is experiencing.

Organizational success is heavily reliant upon employees' motivation (Gau et al., 2013). Beyond that, so many aspects of police service require trust to underpin working relationships, as so many critical functions are performed and so many decisions are made without the benefit of always having a supervisor immediately available to help orchestrate success (Russell, 2014). For these reasons, developing an understanding of those aspects of the work environment that appeal to the individuals in the workforce, bolster their trust in leadership, and serve as motivation provides insight into an organization's potential to succeed. Trust in leaders, further, mitigates employee stress. Transformational leaders inspire follower growth as individual extensions of organizational vision and do so by cultivating trust and followership among the

employees under their charge (Russell, 2014). Police officers' stress levels, and resulting burnout are mitigated by transformational leaders because they cultivate trust with their subordinates (Russell, 2014).

Accepting that police departments routinely assess for and select those to be police officers who have demonstrated a capacity to comprehend, assess, distinguish, appreciate, and communicate emotions, it is reasonable to assume police officers possess developed capacities of emotional intelligence (EI) (Mayer & Salovey, 1997). So, beyond the leader-subordinate dyad, employees' perceptions of organizational justice also impact employee stress and propensity for employee burnout (Brunetto, Teo, Shacklock, & Farr-Wharton, 2012). In fact, employee perceptions of organizational justice have been found to entirely mediate "the relationship between emotional intelligence and turnover intentions" (Meisler, 2013, p. 441). Moreover, those who have higher levels of emotional intelligence feel better, tend to engage with others, seem more committed to the organization, and are, thus, less interested in leaving an organization (Brunetto et al., 2012). Paying attention to levels of employee interest in attaining increasing levels of responsibility and their perceptions about whether promotions are attainable are necessary to understand how a promotion system can impact perceptions of organizational justice and, thus, employee turnover (Gau et al., 2013). Promotion aspirations and the accompanying perceptions of promotion viability can factor into turnover because, for some employees, part of their motivation to strive for excellence lies in their desire to be promoted and inextricably hinges upon their belief that their goal is attainable (Gau et al., 2013).

Selecting Leaders for Promotion

Prior to 1872, and before the creation of civil service protection, police exams were selection procedures administered to individuals as a mere formality for those actually selected by political figures; the exams were not competitive assessments and did nothing to test one's aptitude for service (Kavruck, 1956). The process of administering competitive selection assessments did not begin in public service until 1872 (Kavruck, 1956). Years later, the U.S. Congress passed the Pendleton Act in 1883, which stands as the first significant congressional effort to provide government employees with protections against political influence and bolster integrity in government service (Hoogenboom, 1959). Civil service provided protections in personnel processes and decisions that were designed to nullify political influence in employee decision-making and, therefore, undo the grip that various schemes of political corruption had on government service employees and government institutions (Hoogenboom, 1959). The Pendleton Act established the U.S. Civil Service Commission (Kavruck, 1956). The U.S. Civil Service Commission, in the years to follow, established both research and testing arms, in order to work toward best practices in employment decisions (Kavruck, 1956). By the mid-1930s, the U.S. Civil Service Commission also conducted research in cooperation with the International Association of Chiefs of Police (IACP) to develop (among other improvements) aptitude tests. Efforts continued, in the years to follow, to develop professional test standards and assessment instruments that were backed by empirical research (Kavruck, 1956). Civil service law in Texas, established in 1978, is codified in Texas Local Government Code 143.

Industrial/organizational (I/O) psychology is a field of endeavor, which began in the early 1900s and borne, first, of an interest in the connection between psychology and

advertising (Levy, 2003) In the years to follow, a foothold for the I/O psychology field emerged within the U.S. military with the development of performance evaluations, performance scales, and intelligence tests administered to military personnel for job placement (Levy, 2003). Since then, the field of I/O psychology grew from a focus on selecting, placing, evaluating and appraising the performance of personnel (initially in the military and then private and public organizations) to personnel satisfaction, motivation and organizational dynamics (Levy, 2003). It was in this same period of time that the idea of testing for intellectual capacity emerged (Cohen, Swerdlik, & Sturman, 2013). With World War II emerged a more comprehensive, military-wide testing program that was developed to efficiently and effectively match personnel to specific jobs (Levy, 2003). In the years to come, research findings suggested that written exam results, and more specifically cognitive ability test results, were valid predictors of job performance (Hunter, 1986; Schmidt & Hunter, 1998). Given the longstanding connections between intelligence testing and cognitive ability, and its connection to job performance, there seems to be a widespread belief that cognitive ability testing was inextricably associated with job performance and employment success.

Diversity and Subgroup Differences

Some assessment instruments, which have been touted as the most valid predictors of future performance, have also been found to produce results with subgroup differences that constitute adverse impact against ethnic minorities (De Soete et al., 2013; Madera & Abbott, 2012). A clear example of this phenomenon is the cognitive ability test instrument, which has been documented as producing higher levels of adverse impact (De Soete, et al., 2013). Cognitive ability tests typically produce substantial subgroup

differences in their results, with research indicating that mean subgroup differences in scores for white participants were higher than scores for black participants (Bobko, Roth, & Potosky, 1999; Hough et al., 2001; Roth, Bevier, Bobko, Switzer, & Tyler, 2001).

Spearman's hypothesis, devised by Jensen (1998), suggests that ethnic subgroup differences in assessment instrument results are a consequence of the cognitive load of an assessment instrument. Cognitive load is the degree to which an assessment instrument parallels cognitive ability (Whetzel, McDaniel, & Nguyen, 2008). Cognitive load could more aptly be described as the degree to which a depth and breadth of cognitive ability is measured by, or required to successfully complete, an assessment instrument. Cognitive loading of an assessment instrument fails to account for differential item functioning theory, which suggests that participants from different cultures understand or draw conclusions from segments of assessment items differently, thus inadvertently deriving different meanings, and as an aftereffect, score differently on the assessment instrument (Hough et al., 2001; Whetzel et al., 2008). So, any score difference on an assessment instrument that results from the interpretation of culturally different meanings of terms contained in either assessment passages or corresponding answer responses, when not specifically representing a construct being assessed by the instrument, could reasonably be identified as differential item functioning (Hough et al, 2001).

Regarding sex subgroup differences, research suggests that men and women perform comparably on cognitive ability tests and on job knowledge written exams (Hough et al., 2001). However, according to Örs, Palomino, and Peyrache (2008), women were found to perform with less variance, with women's tests scores being grouped more tightly to the mean compared to men; this would indicate that men scored both higher and

lower than women. If focusing only on top performers, this effect may serve to disparately impact women attempting to promote based solely on written exam performance.

Legal Defensibility

The practice of basing employment decisions on written test results that did not foster more diversity brought an increase in legal challenges (Perlmutter, 2012). The basis of the emerging challenges was the concern of disparate impact that written tests (ethnic group differences in test results) had on minorities and women and the underrepresentation of those subgroups in certain segments of the workforce (Perlmutter, 2012).

As suggested by the reference to legal defensibility, and vital to the current research, it is necessary to apply context to federal legal requirements imposed on organizations regarding the importance of diversity in selection and promotion process results. But beyond legal requirements, diversity is relevant to the discussion of police promotions and the processes that serve as mechanisms for achieving rank-based promotions in police agencies. Diversity throughout the supervisory and management ranks is important for a myriad of reasons, such as fostering diversity in perspectives, experiential enrichment for both the public and agency employees, and opportunity in employment. Each of these justifications is relevant to, and reasonably underpins, the risk management discussion for any police agency.

Two studies examined federally-protected demographics to draw conclusions on their relationships to minority representation in police leadership roles and candidates' interest in availing themselves of the promotion process (Gau et al., 2013; Gustafson,

2013). Federal employment laws, as they regard selection and promotion employment decisions, established protected classifications, including women and anyone older than 39 years of age, who is disabled, or belongs to a racial or ethnic group that is other than white (Dunleavy & Morris, 2016). Examining internal and external factors that impact minority representation in the ranks of police officers, and leadership roles within their agencies, Gustafson (2013) concluded that minority representation in line officer positions was greatly influenced by the prevalence of minorities holding positions as elected officials and police leaders in their respective municipalities. Additionally, larger departments that offered higher levels of compensation were found to have higher levels of minority diversity in police administration positions (Gustafson, 2013). Gau et al. (2013), in an attempt to explain promotional aspirations, assert that differences in race, sex, and education predict both the importance of promotion and a desire to promote to higher ranks before retirement. While speculation was provided as to what determined the relationships, Gau et al. (2013) found that police promotions were more important to males than to females, and more important to minority employees than to white employees. There is value in a diverse workforce, and that value is felt as much within police organizations as it is in communities that police departments serve.

A department's legitimacy in the community, among its constituencies, and among its personnel hinges upon diversity, but how diversity impacts each uniquely varies. Racial diversity augments the community perceptions of uniform treatment of citizens and a department's legitimacy (Weitzer & Tuch, 2006). However, when it comes to how salient diversity is among the rank-and-file, it provides more than just an appearance of equity; employing nonwhite officers enhances and increases quality

interactions between white and minority officers (Gustafson, 2013). Those interactions, over time, tend to undermine stereotypical beliefs about minority cultures that help white officers to understand minority cultures and improve white officers' perceptions of minority community members (Gustafson, 2013). In this way, department leaders, who work to hire a workforce that mirrors the demographic make-up of the communities they serve, will more aptly develop a cadre of officers who more capably communicate with those community members they serve (Gustafson, 2013).

While ensuring that police departments employ a workforce that is racially diverse, sex diversity is just as vital to organizational success and viability. To that end, studies have examined sex diversity, exposing several considerations that accent how diversity management bolsters a police department's readiness and capability in service delivery (Anderson et al., 2006; Mano-Negrin & Sheaffer, 2004). In doing so, researchers have also addressed some longstanding myths about male superiority in police service. Two sex diversity research endeavors have sought to unwrap capability differences between men and women leaders in law enforcement, finding that women are capable leaders and, in fact, more capable than men as communicators and managing solutions in crisis situations (Anderson et al., 2006; Mano-Negrin & Sheaffer, 2004). Examining sex differences in a leader's capacity for effective crisis management, Mano-Negrin and Sheaffer (2004) concluded "that perceptions of crisis preparedness/proneness are sex-based and that women are more likely to employ a holistic approach that facilitates crisis preparedness" (p. 109). Anderson et al. (2006) examined sex differences in a variety of traits vital in leadership roles. They found that women in leadership roles more capably managed interpersonal communications and were more driven and determined than their

male counterparts. In spite of the value female police leaders add to law enforcement's operational capacity, less than 12 percent of sworn police officers are women and less than 8 percent of command-level police personnel are women (Morabito & Shelley, 2018).

Archbold and Hassell (2009) examined the considerations that factored into female officers' decisions to participate in their department's promotion process. While there is a myriad of considerations, two related to the organization and to officers' personal lives (Archbold & Hassell, 2009). One unanticipated influence on participation was whether or not a female officer was married to a fellow officer, a circumstance which, according to female respondents, tended to thwart one's capacity to achieve promotion (Archbold & Hassell, 2009). However, Archbold and Schulz (2008) asserted that female officers were often encouraged to avail themselves of the promotion process, leaving some with the impression that they may achieve promotion based solely on the fact that they were females; this was found to deter many women from seeking promotion. In spite of this result, many women expressed approval for the opportunities for promotion (Archbold & Schulz, 2008).

Laws Relevant to Diversity and Promotional Processes

There are a number of federal and state laws that, together, establish employer guidelines and responsibilities when undertaking workforce promotion activities in their respective organizations. The overarching goal of employment laws is to cease discriminatory employment actions based on their memberships in any of the federally-protected classifications (Equal Employment Opportunity Commission, 1978). The basis for establishing these protected classes may be found in an array of legislation (*Age*

Discrimination in Employment Act of 1967; Americans with Disabilities Act of 1990; Civil Rights Act of 1964).

Civil Rights Act of 1964

The Civil Rights Act of 1964 sought to responsibly address the problems of racial inequality in our nation's history; the major features of the Act, of which there are several, are divided into eleven parts called titles. Title VII, the most well-known title of the legislation, makes discrimination in employment in any business on the basis of "race, color, religion, sex or national origin unlawful" (United States Code, n.d., Title 42, Section 2000e-2, Sec. 703, (b)). While enumerating enforceable Title VII mandates, the Uniform Guidelines on Employee Selection Procedures (1978) includes ideas on how to safely negotiate selection processes so as to remain in compliance with those guidelines (Dunleavy & Morris, 2016). This Act did much more than merely proclaim that discrimination was a violation of the law; it prescribed specific avenues for redress and strengthened federal law to effect behavioral change, adding language that illustrated an expectation for legal challenges. Title VII provides that an individual can bring a private lawsuit as redress for suspected violations and makes it unlawful to retaliate against employees who oppose the unlawful discrimination (United States Code, n.d., Title 42, Section 2000e).

The EEOC was created by Title VII of the Civil Rights Act of 1964 (Section 705). The EEOC was tasked with enforcing federal laws that prohibit discrimination in employment (Section 705). While Title VII does not offer protections against employment discrimination on the basis of age or disability, both the Age Discrimination in Employment Act of 1967 and the Rehabilitation Act of 1973 protections fill these

voids in Title VII protection; EEOC guidelines only regulate against discriminatory employment practices that use race, sex, or ethnicity as a basis for the discrimination or disparately impact members of any one of these groups (Equal Employment Opportunity Commission, 1978). However, regardless of the origin of the protection in federal law, the EEOC is the entity charged with enforcing compliance with prohibitions against discrimination in employment toward those belonging to federally-protected classes (Equal Employment Opportunity Commission, 2017). The real strength in enforcement appears to be the EEOC's zeal in enforcement; while the Supreme Court has not specified in any ruling that an organization's racial composition should mirror that of the community that it operates in or serves, the Equal Employment Opportunity Commission (EEOC), as the entity charged with enforcing Title VII provisions, seems to support this notion (Wax, 2012).

Civil Rights Act of 1991

The Civil Rights Act of 1991 further clarified Congress' resolve to provide protections within existing employment law by the "conjoining of job relatedness and business necessity" requirements; this "[represented] a departure from some earlier case law that allowed a defendant to prevail by showing either job relatedness or business necessity" (Grover, 1996, p. 392). The Title VII business necessity requirement maintains that use of an assessment instrument or adherence to a guideline or provision is proper when it is determined that the instrument, guideline or provision constitutes a business necessity, even though it results in a disparate impact against protected classifications (Gutman, 2005). As an example, an employer can elect not to promote incumbents, and

hire new employees, when incumbents applying for consideration fail to meet required qualifications (*Gu v. Boston Police Department*, 2002).

Two federal court cases provided insight into the concept of business necessity and what has been considered evidence sufficient to prove a business necessity. *Lanning v. Southeastern Pennsylvania Transportation Authority (SEPTA)* (2002) This lengthy court battle stemmed from a legal challenge of a police applicant run standard with a minimum-qualifying cut-off score (*Lanning v. SEPTA*, 2002). After the case had been argued twice in the Eastern District of Pennsylvania District Court and twice in the Third Circuit Court of Appeals, the applicant run standard was upheld as being necessary for job (*Lanning v. SEPTA*, 2002). Evidence supporting cut-off scores for the run included both the assertion that many of the women who failed to meet the run standard were able to do so after devoting time to training to improve their performance and that SEPTA had determined to improve performance capacities of those in their employ (having imposed a battery of physical agility testing on incumbents bi-annually), regardless of the fact that disciplinary action could not be administered for failing the testing (*Lanning v. SEPTA*, 2002). The Third Circuit Court of Appeals decision added that while applicants would not, necessarily, need to call upon their running skills before they were working as police officers, there should be no requirement for SEPTA to hire new employees without, first, subjecting them to physical agility testing with a set of required performance standards. Lastly, the run standard was set based upon a validation study (*Lanning v. SEPTA*, 2002). In *U.S. v. City of Erie, Pennsylvania* (2005), another legal challenge to stringent physical agility standards for applicants arose, again, due to the disparate impact they had on women applying for employment. The Western District of Pennsylvania District Court,

however, found that the City of Erie failed to demonstrate, to their satisfaction, that the police applicant physical agility standards (established at the recommendation of the same physiologist that established the standards challenged years earlier in the *Lanning* case) constituted a business necessity, as no validation study had been conducted on those standards (*U.S. v. City of Erie, PA.*, 2005). In fact, the validity of the standards was determined to rest merely on the opinion of the physiologist hired to establish the standards (*U.S. v. City of Erie, PA.*, 2005).

The Civil Rights Act of 1991 further insisted that liability existed if plaintiffs reasonably identified an alternative selection tool which has been shown to have less adverse impact than the tool of process in question (Grover, 1996). The Civil Rights Act of 1991 also deemed bias-based score manipulations to be unlawful, proclaiming, “it shall be an unlawful employment practice for a respondent, in connection with the selection or referral of applicants or candidates for employment or promotion, to adjust the scores of, use different cutoff scores for, or otherwise alter the results of, employment related tests on the basis of race, color, religion, sex, or national origin” (Section 106). These additions to existing discrimination laws forecasted employment practices that may, later, become the subject of future litigation.

Texas Local Government Code 143

Texas Local Government Code 143 (LGC 143) provides for civil service protection for police officers and firefighters in certain municipalities having “[populations] of 10,000 or more” residents, which have hired “police or fire department” employees, and have been granted civil service protection by election (Texas Local Government Code 143, 2015, Section 143.004(a)). The LGC 143 was enacted to provide

municipal police officers protections against “political influence” in the course of performing their duties (Texas Local Government Code 143, 2015, Section 143.001). The LGC 143 outlines procedures for civil service promotions (Texas Local Government Code, 2015, Section 143.032). Promotional eligibility lists are to be based upon written examination scores, but not upon oral interview rating results (Texas Local Government Code, 2015, Section 143.032(c)). More specifically, such written exams must be constructed so as to allow for immediate grading of such exams at their conclusion; this all but insists on a standardized exam (Texas Local Government Code, 2015, Section 143.032(f)). The Texas Local Government Code (2015), further, imposes a criminal sanction for violations of the promotional examination provisions in the Code (Section 143.032(i)). Moreover, the Texas Local Government Code (2015) provides a specific process for appeal of examination results (Section 143.034(a)).

Despite LGC 143’s insistence upon use of written examinations for developing promotion eligibility lists, it does allow for an “alternative promotional system,” provided that it is “approved by a majority vote of the sworn police officers voting” (Texas Local Government Code, 2015, Section 143.035(f)). What is not addressed by this section, and thus potentially overlooked by affected sworn personnel, is that adoption of an alternative promotional system (inadvertently) rescinds the protections that are afforded by Texas Local Government Code (2015), such as the right to an appeal process, the provision of criminal sanctions for violations of the promotional process section of the Code, and any number of other provisions not specifically spelled out in the document used to enumerate provisions of the alternative promotional system; this can have any number of

unforeseen consequences. In particular, disparate impact liability is unclear for agencies employing a promotional process that merely adheres to state law.

Disparate Impact

Disparate impact liability represents a stand against what is arguably one of the most prevailing types of discrimination thwarting those belonging to protected classes from progress and promotion in the workplace (Rosenthal, 2013, p. 2160). Disparate impact is demonstrable and quantifiable; Equal Employment Opportunity Commission guidelines define the “Four-Fifths Rule” as “a selection rate for any race, sex, or ethnic group which is less than four-fifths ($4/5$) (or eighty percent) of the rate for the group with the highest rate will generally be regarded by the Federal enforcement agencies as evidence of adverse impact, while a greater than four-fifths rate will generally not be regarded by Federal enforcement agencies as evidence of adverse impact” (Equal Employment Opportunity Commission, 1978, Section 4. D). As an example, if a department offers incumbent officers a promotional process for the rank of detective, based only on the outcome of a written exam and 100 officers avail themselves of the promotional process for an opportunity to promote (80 male, 20 female; 60 white, 40 non-white) and test results in 72 males passing and 12 females passing (90 percent and 60 percent pass rates, respectively), and 56 white applicants passing and 28 non-white applicants passing (93.3 percent and 70 percent pass rates, respectively), the results would show a disparate impact against both protected classifications, according to the 4/5ths Rule. The 4/5ths Rule compares pass rates between subgroups; in the example, any pass rate for females below 72 percent ($0.9 \times 0.8 = 0.72$) and / or for non-white participants below 74.64 percent ($0.933 \times 0.8 = 0.7464$) would constitute a disparate

impact and serve as grounds for initiating an EEOC adverse impact claim. Understanding the concept of disparate impact and why employers seek to avoid disparate impact liability is essential to the discussion of the AC because the AC is so often touted as an assessment tool most capable of reducing disparate impact in its results; those assertions hold little weight without context.

Case Law

Disparate treatment differs from disparate impact, and two landmark court cases provided the distinction. *Griggs v. Duke Power Co.* (1971) was the first US Supreme Court decision that established disparate impact protections, while *McDonnell Douglas v. Green* (1973) established case law for disparate treatment protections. Disparate treatment involves employment actions that constitute overt and intentional differential treatment due to one's membership in a protected classification (*International Bhd. of Teamsters v. United States*, 2007). Disparate impact, on the other hand, involves employment actions that appear equitable, at face value, but result in substantially differential employment outcomes in favor of those not belonging to protected classifications (*Griggs v. Duke Power Co.*, 1971). The US Supreme Court found that employment actions, even when absent any intent to discriminate against a member of a protected class, were unacceptable if they were found to substantially and adversely impact members of a protected class (*Griggs v. Duke Power Co.*, 1971). The substance of the justification for the disparate impact doctrine is that any advantage or disadvantage, based solely upon one's membership in a federally-protected class, be deemed impermissible (Wax, 2012).

Even when employment actions are motivated by a desire to preemptively avoid disparate impact liability, they may fall under legal scrutiny; the U.S. Supreme Court decision in *Ricci v. DeStefano* (2009) effectively renounced governmental action for this very purpose. Interestingly, the *Ricci* decision held that the employer engaged in racial discrimination with their choice to proactively abandon a promotion process to avoid a disparate impact complaint and did so only after the results of the process indicated that a substantial number of promoted candidates (from that process) would be nonminority participants (Rosenthal, 2013). Based primarily on assertions of city administrators that the racial composition of those most likely to promote would possibly place the City of New Haven at risk of an EEOC complaint and subsequent litigation, the Court found that the City of New Haven only chose not to certify the exam results because of the racial composition of those who scored highest on the exam (*Ricci v. DeStefano*, 2009; Roberts, 2010).

When employers determine that selection and/or promotion outcomes may place them at risk of disparate impact liability, and take proactive steps to avoid such risks only after selection or promotion processes do not yield results devoid of any disparate impact, equal protection under the law, then, becomes relevant; because such a decision is driven by a desire to maintain compliance with the law, it is the employment decision that will bear judicial scrutiny under the Equal Protection Clause (Kanovitz, 2012). Both the *Griggs* and *Ricci* decisions by the US Supreme Court made navigation amidst employment decision pitfalls an arduous task for organizations merely seeking to remain in compliance with EEOC guidelines (Roberts, 2010; Wax, 2012). This has, since, driven employers to realize that selection and promotion process development requires more

time and scrutiny so as to avoid or mitigate liability potential under EEOC guidelines (Roberts, 2010).

The Supreme Court asserted that the Civil Rights Act of 1964 established that when a court finds that a selection process or assessment instrument has had a disparate impact on a segment of the workforce comprised of a protected class, the burden of proof is placed squarely on the employer to show that the assessment instrument or process being challenged is necessary to the organization's success in the sought-after position (*Griggs v. Duke Power Co.*, 1971). This burden could be met in one of two ways. "The Court further stipulated that employers could escape liability for 'disparate impact' only if they demonstrated that their adverse selection practices had a manifest relationship to the employment in question *or* that they were justified by 'business necessity'" (Wax, 2012, p. 54). As clearly established by law and legal precedence, however, a test instrument can be valid and not be job related; the terms are not necessarily synonymous (Murphy, 2009).

However, the U.S. Supreme Court's decision in *Wards Cove Packing Co. v. Atonio* (1989) marked a departure from the Court insisting on stringent requirements for organizations to protect against discriminatory employment action; this decision eased the dual requirements insisting that a selection process (which resulted in disparate impact) must be job-related *and* constitute a business necessity in order for that process to be deemed appropriate (Grover, 1996). It was this change that precipitated, and arguably bolstered Congressional support for, the more stringent disparate impact protections contained in the Civil Rights Act of 1991 (Grover, 1996).

The Assessment Center (AC)

Because of the disparate impact of written exams and the concomitant legal challenges these outcomes generated, agencies began to seek alternative selection processes that had the potential of reducing the disparate impact generated by written exams. The practice of utilizing multiple assessors to evaluate leadership capability in complex scenarios first began in the early twentieth century to cultivate and select military leaders (Hsin-Chih, 2006). The first application of an AC by a private sector organization was by AT&T in 1956 (Development Dimensions International, 2001; Spychalski et al., 1997). Bray and Grant (1966), basing their findings on The Bell System's Management Progress Study (an eight-year longitudinal study on managerial development) found that situational assessment exercises and ability testing were better predictors of progress than personality assessments. Since the 1960's, use of the assessment center has gained widespread notoriety in private sector management selection (Adler, 1987; Hsin-Chih, 2006). Beyond its increased use as a selection and promotion tool, a great deal of research on AC validity has been conducted in the last 20 years; research on, and practical use of, ACs are most often conducted pursuant to employment selection and promotion (Hsin-Chih, 2006; Hermelin et al., 2007).

The AC is a process of simulations of job-related activities in which participants are assessed by numerous raters to determine, to the extent possible, a candidate's capacity to appropriately employ a variety of competencies that are necessary to succeed in the sought-after position (Dayan, Fox, & Kasten, 2008). The goal of the AC is to determine how proficient participants are in simulations, being rated by a given set of predetermined competencies (International Task Force on Assessment Center Guidelines, 2015, p. 1269). The AC method is designed to provide a comprehensive evaluation of

one's capacity to appropriately bring to bear sought-after competencies, abilities, and skill sets in response to a multitude of scenarios that simulate occurrences likely to be experienced in the sought-after position (Gibbons & Rupp, 2009). Features of an AC include: multiple exercises designed to assess likelihood of future success by eliciting and then rating competencies considered essential to success in the sought-after position, multiple raters to observe and rate behaviors (that constitute those competencies), raters that record ratings of observations, utilizing some combination of raters able to competently assess performance, score integration, and compilation of an overall assessment rating (OAR) (Thornton & Gibbons, 2009). ACs provide a platform to substantially inform raters of the myriad of capacities a participant possesses (Gibbons & Rupp, 2009).

The AC method has been designed to accomplish different functions. ACs have been used for "selection, placement, early identification of management potential, promotion, development, career management, and training" (Development Dimensions International, 2001, p. 1). Specifically, according to Thornton and Rupp (as cited by Gibbons and Rupp, 2009), ACs assist leaders to make selection and promotion decisions, to diagnose needs, and to facilitate employees' development" (p. 1156). ACs, while secondarily employed to develop participants' management capacities, are primarily used to determine whether to select and/or promote participants (Hermelin et al., 2007). In fact, when ACs first emerged, they were employed as a selection tool to determine who was most likely to succeed and, by extension, to determine who was most likely to fail (Moses, 2008). Only in recent years has the role of the AC expanded to assess employees' capabilities in an effort to develop their skill sets and competencies for the

benefit of the organization, to develop employees, and to determine who should be promoted to roles with increased responsibility (Moses, 2008; Tillema, 1998).

ACs are properly considered to be selection instruments when specifically designed to identify or develop management potential (Adler, 1987). However, Breugh (2010) suggested that knowing why one employee is less likely to be promoted than another is vital, as those reasons provide insights into potential employee development opportunities. A development center (DC) is “a broad term that encompasses both assessment centers solely designed to diagnose development needs and assessment centers designed to catalyze development within the course of the assessment center program” (International Task Force on Assessment Center Guidelines, 2015, p. 1270). A developmental assessment center (DAC) is “an assessment center designed for the purpose of directly developing/improving assessees on behavioral constructs of interest” (International Task Force on Assessment Center Guidelines, 2015, p. 1270). Tillema (1998) unwrapped the similarities and differences between ACs and DCs while seeking to examine both management opinions about ACs and the AC’s feasibility as a developmental tool. Still, it is often tedious to distinguish between DACs and ACs used for selection of employees. (Adler, 1987). Moreover, DAC’s have been shown to affect managerial performance improvement (Engelbracht & Fischer, 1995; Papa & Graham, 1991). The value DAC’s have on participants’ managerial performance was associated to focused training on those KSA’s necessary for managerial success in both the DAC’s and ACs (Engelbracht & Fischer, 1995; Papa & Graham, 1991). Thus, participant exposure and focused attention to the AC, along with the accompanying preparation process, provides value as leadership training opportunity in and of itself. For example,

maintaining a protracted focus on, and training to effectively communicate, (both orally and in writing, including developing active listening skills to elicit information), gathering and synthesizing information, maintaining an orientation toward action, developing others, and illustrating empathy each provide participants with skills necessary to lead and manage others in their respective organizations (Engelbracht & Fischer, 1995; Papa & Graham, 1991).

AC Validity

EEOC Guidelines identify three methods that can be employed to validate assessment instruments and processes for employee selection; court decisions have supported those validation techniques (Wax, 2012). "'Content' validation requires establishing a manifest relationship or plausible match between a screening assessment and key job tasks" (Wax, 2012, p. 64). Content validation is more about intuitively aligning capabilities tested in the assessment instrument to capabilities required for the sought-after position than it is about establishing linkages through statistical calculation (Wax, 2012). "In contrast, the so-called 'construct' and 'criterion' validation methods require an employer to demonstrate a quantitative relationship between a job-selection method and measures of specific job-related skills (in the case of construct validation) or workers' actual on-the-job performance (in the case of criterion validation);" I/O psychologists consider criterion validation to be the most robust and accurate validation method (Wax, 2012, p. 64). Determining whether or not an employer's testing process is job-related is merely demonstrating that the process is valid as employed in its current form and is valid to substantiate employment decisions (Saad et al., 2000). So, score differences, with a valid instrument, will aptly represent differences in individual

candidates' ability to appropriately demonstrate tested competencies and, because competencies are assessed across multiple AC exercises, candidate ratings of those competencies will be consistent across exercises (De Corte, Lievens, & Sackett, 2006).

The AC Process in Practice

Before an AC can be administered, administrators design AC exercise scenarios. In order to do that, AC administrators must know the bases of knowledge, skill sets, and abilities (KSAs) that are most sought after for the best leaders in the organizations they serve. ACs differ from a mere set of simulations; AC development begins with conducting a job analysis, which leads to a process of developing sought-after competencies (Hsin-Chih, 2006). The job analysis includes a series of interviews and subsequent analysis of information gleaned from those interviews with the goal of determining the qualities and capacities that successful leaders already possess in a given organization (Hsin-Chih, 2006). After prioritizing KSAs in a given organization, AC exercise rating rubrics are constructed to weight in favor of those KSAs, or competencies (Hsin-Chih, 2006). AC exercises, when properly devised, prompt candidates to exhibit or activate dimension-related behaviors while performance ratings grade participant performance; performance disparities, from one participant to the next, are expressed in variable performance ratings (Gibbons & Rupp, 2009). Raters are typically police personnel employed outside the testing agency that have served either at or above the sought-after rank for a designated period of time and from a comparably sized police department. AC administrators' rater requests seek diverse groups of raters for a given AC; the rater acquisition process often includes an application process so that a diverse cross-section of competent, accomplished leaders can be constructed to effectively rate

AC participants' performance. Raters, before being expected to rate participants' performance, are subjected to rater training. Rater training includes explanations on how to employ the rating rubric along with explanations from subject matter experts (SMEs) (senior leaders assigned within the organization) to provide context on the organization's culture, the organization's goals, and organization-specific leadership challenges. Raters, typically in teams of three, receive briefings on respective exercise scenarios and rate each participant based on the rating rubric. Raters are called on to interact among themselves to ensure they reach a consensus on overall ratings for each participant within each exercise. Overall assessment ratings are constructed from the individual ratings received in each exercise.

Examples of AC scenarios routinely used include leader/subordinate counseling exercises, community meetings, shift briefings, command staff briefings, and in-basket exercises. In-basket exercises include any number of notes, memos, or reports that a person already in the sought-after position may encounter in any given assignment. Exercises include preparation time before participants are expected to perform either oral presentations or provide written samples based on the scenario information provided by AC administrators. In the preparation phase, participants are then required to read the provided information, determine what (direct and underlying) issues exist in the information provided, how they intend to deal with the issues, and then to rank them by urgency and/or importance. During the presentation phase, participants are then required to present the information in order of urgency or importance, and identify issues and the manner in which they would address those issues. In other exercise formats, information is provided with the expectation that participants prepare to address issues relevant to

provided information and deliver a presentation that includes dialogue that competently addresses issues along with plans to correctly manage those issues. The AC, in this way, presents a high-fidelity assessment of competencies (dimensions) expected of those in the sought-after ranks.

While each department's job analysis differs, beyond assessing one's interpersonal skills, a typical AC routinely evaluates one's capacity to identify and analyze problems, make decisions, plan, organize, and communicate (Morris, 2012). And while all ACs seek these general categories of competencies, weight given to each category and how AC exercises are designed to elicit responses for each of the competency categories will vary by rank (Hsin-Chih, 2006; Morris, 2012).

Promotions, Written Exam vs. AC, and Protected Classes

I/O psychology experts insist (in the amici brief on the *Ricci* case) that the AC method is superior to the written test format, due to the AC's lower adverse impact and comparable validity estimates (Wax, 2012). Research endeavors comparing subgroup differences from cognitive ability assessments to those from ACs have concluded that the cognitive ability test instrument has been documented as producing higher levels of adverse impact than AC results (De Soete et al., 2013; Outtz, 2002; Sackett, Schmitt, Ellingson, & Kabin, 2001). However, while Thornton and Rupp (2006) asserted that a written exam results in more adverse impact than an AC, according to Dean et al. (2008), the AC method does not do as much to mitigate adverse impact as previously asserted by researchers and practitioners. Examining 27 studies of racial subgroup differences in performance ratings, Dean et al. (2008), concluded that black-white subgroup differences were found to be larger than white-hispanic subgroup differences, indicating that black

participants are more adversely impacted by AC results than are Hispanic participants (Dean et al., 2008). So, Dean et al. (2008) suggest that any blanket assertion that ACs reduce adverse impact is not accurate, absent caveats that account for the subgroups participating in the AC. However, “the ability of an assessment center...to withstand legal challenge is primarily a function of the demonstrable job-relatedness of that specific assessment center...and not some general magical or inherent ability to reduce or eliminate adverse impact” (Arthur, Doverspike, Barrett, & Miguel, 2013, p. 479).

Other research has explored AC performance by sex, concluding that men, on average, scored somewhat lower than women (Dean et al., 2008). While men and women tend to score similarly on measures of cognitive ability, AC metrics and exercises (designed to elicit certain capacities thought to predict success in a given organization) can test for a variety of personality differences that could advantage either males or females (Dean et al., 2008). Designing AC exercises that tend to measure one's capacity to nurture, coach, or develop subordinates (as an example) would likely provide competitive advantage to female participants. (Dean et al., 2008).

Regarding the relationship between AC performance and promotion, Bishopp (2013) sought to identify the factors that were most relevant to sergeants promoting to the rank of lieutenant in a major metropolitan police department, hypothesizing that, while higher AC scores had a significant impact on whether a promotional candidate would be promotion-worthy, the AC scores were uninfluenced by a candidate's capacity to supervise. Bishopp (2013) found that the AC score increases the likelihood of promotion much more so than the written score. Bishopp (2013) also found that age had an inverse relationship to AC score, as did age coupled with minority status. While asserting that

being female, having formal education, and a higher number of sustained complaints predicts the likelihood of receiving a higher AC score, he qualifies this assertion by noting that so few examined observations ($n = 202$) cannot support a high level of confidence in a causal connection (Bishopp, 2013). Interestingly, no prediction of the likelihood of promotion could be made based solely on minority status, education, sex, or the number of times a candidate participated in the promotion testing process (Bishopp, 2013). Lastly, when examined along with either age or sex, Bishopp (2013) found that minority status had no statistically significant relationship to the likelihood of promoting.

Looking at race and sex as they each relate to retest success, Schleicher et al. (2010) examined retesting performance differences between subgroups of federal employment job applicants. Schleicher et al. (2010) determined that white test takers' retest performance improved more than black and hispanic participants, and more so on written tests than on performance tests (designed with features found in ACs). When sex and age differences were examined, Schleicher et al. (2010) found that women improved on retests more than men.

Recognizing that there have been few studies conducted to examine police promotion testing processes and, more specifically, conducted to identify the determinants of candidate success in those promotion processes, Topp (2011) sought to determine whether or not promotional test performance was predictable over time by examining 17 years of AC performance results. As rank is unidirectional, AC performance, as measured by AC scores, at lower ranks was examined in an attempt to determine if it predicted score performance in ACs at higher ranks (Topp, 2011). Topp (2011) surmised that, over time, test performance remained somewhat steady, but added

that performance at lower ranks was not predictive of test performance for subsequent (higher) ranks.

Several studies focused on testing score differences between racial or sex subgroups (Arthur et al., 2002; Chan, 1997; De Corte & Lievens, 2003; De Corte et al., 2006; Hough et al, 2001; Schleicher et al, 2010). Other studies focused on racial or sex subgroup differences in AC performance, highlighting the reduction of subgroup differences when compared to written exam subgroup differences (Becton, Feild, Giles, and Jones-Farmer, 2008; Dean et al., 2008; De Soete et al., 2013; Outtz, 2002; Sackett et al., 2001). Examining overall AC performance and within candidates' respective subgroups to identify if there were any differences in the level of adverse impact between a multiple-choice test and a constructed response (fill in the blank or short answer) test, Arthur et al. (2002) concluded that adverse impact in the results of the constructed response test was less prevalent than the results of the multiple-choice test. While the sample size for each test format was small, results suggest that a constructed response test may be able to mitigate adverse impact that may otherwise exist in the results of a multiple-choice test (Arthur et al., 2002). Two tradeoffs with using a constructed response format include a potential of subjectivity in grading responses and an increased expense for the added time it takes for proctors to grade responses (Arthur et al., 2002).

Becton et al. (2008), when examining candidate reactions to a promotional process consisting of a written exam and situational interview, found that white and black candidates considered the written exam to be less job-related than the situational interview and even though white candidates performed better on the written exam, white candidates determined both assessment instruments less job-related than black candidates

(Becton et al., 2008). Similarly, black candidates performed better than white candidates on situational interviews, but not as well as white candidates on written exams (Becton et al., 2008). Gaps in test performance between blacks and whites have been widely researched in years past and represent one of the most prevalent concerns in the promotional exam arena (Becton et al., 2008).

The value of employing the AC as a promotion tool, however, remains in question when taking a holistic view of the necessity for the AC method. Administering an AC, when compared to a written exam, is more expensive and labor intensive (Swanson et al., 2012). Moreover, some have speculated about how cost-effective and flexible ACs have proven to be in a given selection process (Adler, 1987). Given the significant cost that accompanies AC development and administration, any budget-constrained law enforcement agency considering an AC as a promotion tool must consider the value an AC adds to any promotion process. This is particularly vital, given the reality that adverse impact litigation (stemming from promotion processes found to have violated federal protections) is, itself, quite costly.

While Topp (2011) examined AC performance to determine the predictability of test performance at one department over time and Bishopp (2013) sought to identify the factors that were most relevant to a single sought-after rank promotion, neither study sought to uncover the value an AC added to several departments' promotional processes over time and across ranks to assess the AC's capacity to facilitate promotion of diverse cadres of police leaders. Moreover, there has been no research on the efficacy of the standard civil service promotional process established in Texas Local Government Code 143, which restricts the selection process to the use of written exams. The gap in research

is significant, given the continuing concern in the United States for any presence of adverse impact in civil service promotional processes. None of the aforementioned studies examined the value added by an AC to bolster diversity and quality leadership at multiple ranks, at multiple departments that are all subject to the same civil service guidelines. Determining if findings are consistent from department to department and from rank to rank will provide a depth of understanding not previously found in other studies. The current research, while not specifically examining adverse impact in any one department, will examine the differences in eligibility list placement between written examination performance, AC performance, and overall performance. In doing so, the current research will highlight performance differences among those sharing membership in federally-protected classifications (race and sex) and those who do not. If ACs are capable of identifying those best poised to deliver quality leadership (as they have widely been purported to do), then employing ACs as part of the police promotional process is of paramount importance to police executives seeking to improve their departments and the service their personnel deliver to their constituencies.

III. METHODOLOGY

The current research will illustrate the effect of the addition of an AC to the promotional process previously consisting only of a written exam on diversity at supervisory and managerial ranks. By determining how a participant's rank placement (and quartile placement) change from a rank-ordered eligibility list based only on the score of the written exam to a rank-ordered eligibility list based only on the AC score, and then to a rank-ordered eligibility list based on a composite score comprising both the same written exam and AC scores, the current research will determine if participants belonging to federally-protected classes (on the basis of race and sex) improve their promotion opportunities by improving their placement on their respective eligibility lists.

Hypotheses

Hypotheses for the current research include:

- H₁. Race differences will be observed in rank-ordered placement and rank-ordered quartile placement based on written scores, AC scores, and composite scores.
- H₂. The addition of an AC in the promotional process improves the rank-ordered placement and rank-ordered quartile placement for nonwhite participants, and this result will be consistent across sought-after rank (i.e. captain/commander, lieutenant, sergeant).
- H₃. Sex differences will be observed in rank-ordered placement and rank-ordered quartile placement based on written scores, AC scores, and composite scores.
- H₄. The addition of an AC in the promotional process improves the rank-ordered placement and rank-ordered quartile placement for female participants, and this result will be consistent across sought-after rank (i.e. captain/commander, lieutenant, sergeant).

Data Collection and Sample

The current research utilizes a dataset that includes promotional event data from four police departments in Texas, serving the 4th, 7th, 9th and 11th most populous cities in the United States (Table 1a); U.S. Census data on each of the four cities were used to report population demographics in Table 1b; according to U.S. Census data, the white population percentage appears significantly higher in Austin than in the other three cities in the current research. These departments were selected due to their size and the fact that they each employ promotional processes that include both written exams and ACs. Open records requests sought sworn staffing levels at each department, which are reported in Table 1c. Notable, in Table 1c., is that demographic differences at sought-after ranks from department to department are noticeable; while white male presence dominates the rank of commander at Austin PD (84.2 percent), Dallas PD has promoted more non-white male and females to its highest testable rank (60 percent of all lieutenant positions). Dallas PD, in fact, has more non-white employees than white employees assigned at every sought-after rank in their department. San Antonio PD, like Austin PD, have no minority women at the rank of Captain / Commander, according to open records data, with San Antonio maintaining the lowest percentages of minority women at every sought-after rank. Additionally, along with written exam and AC raw scores, race and sex designations for each of the participants in each of the promotional processes was requested; data were collected from each department from promotional processes between 2000 and 2018 only for each sought-after rank that may be attained based on participant performance on both a written exam and an AC (Table 2a). The dataset includes 4270 participants, each identified by department, race, and sex (Table 2b). The dataset includes participants from 67 police promotional events (identified as cohorts)

that occurred in these four departments (Table 2c). For the purposes of the current research, a promotional event is defined as a single process involving a promotional competition for a particular rank that consisted of both a written exam and an AC. Participants were identified only as white or nonwhite, and male or female, to determine their membership in one or more federally-protected classifications (Table 2d).

Measurement of Variables

Upon receipt of several of the eligibility lists obtained from the various agencies for the current research, it became apparent that there were significant variations in the score ranges awarded for AC performance (Table 2e, 2f, 2g). It was decided that those disparities, when calculating composite scores, would significantly weight composite score eligibility list placement in favor of the written exam scores. Thus, conversion of raw scores to z-scores was considered to be the optimal method of retaining the relevant value of each participant's performance rating without regard to the variety of scaling used in any one promotional event score calculation. Moreover, use of z-scores (standardized scores) was considered to do much to ensure that individual promotion process differences (significant differences in ranges of observed raw scores in the dataset, potential for differential performance ratings based on non-standardized rating scales, potential rater training and rater disposition differences, potential differences in written exam and AC exercise difficulty and differential allotted time frames, differences in written exam content or cognitive load, etc.) would be removed from study results. For example, adding a raw written exam score to a raw AC score from a process in which both written exam and AC were each graded on a 100-point scale may yield inaccuracies, for the purposes of comparison, when one considers another process in which the written

exam was graded on a 100-point scale and the AC was graded on a 10-point scale (as in the case of one specific promotional process). Also, because each department computes composite/final scores differentially, either by differentially awarding seniority points, education points, and/or veteran preference points, all award points were extracted from the dataset so as to ensure that award points did not differentially impact research findings. Variables used in the analysis and their coding strategies can be found in Table 3.

Dependent Variables

The dependent variables include: raw scores for the written exam, AC, and composite score (converted to z-scores), the rank-ordered placement based on the z-scores for the written exam score, the AC score, and the composite score, and the rank-ordered quartile placement based on the written exam, the AC score, and the composite score. Examining quartile placement is relevant to both a participant's likelihood of promotion and the number of participants included in a promotion eligibility list; a participant earning a rank of 15 on an eligibility list containing 60 participants would be much more significant than earning the same rank placement on an eligibility list containing 16 participants. Rank-ordered placement changes are also included and are based on the difference between the rank-ordered placement on the AC and the written exam and differences between the composite score and the written exam. For all rankings, the highest score corresponded with the highest eligibility list placement (indicated by a rank of 1). For all quartile rankings, a quartile rank of 1 corresponded with the highest of four quartiles while a quartile rank of 4 corresponded with the lowest of four quartiles.

Independent Variables

Variables for NONWHITE (1=nonwhite, 0=white) and FEMALE (1=female, 0=male) were constructed from race and sex identifiers supplied by the four agencies in the documents responsive to the open records requests. Names provided in some of the responsive documents were extracted and left out of the dataset for the current research.

Data Analysis Plan

At the outset, the current research intends to conduct a descriptive analysis of all participants from all four departments. Descriptive statistics for participants seeking promotion to each of the sought-after ranks were parceled out to identify race and sex of participants, the department at which they were each employed, the means, standard deviations, minimums and maximums for written exam, AC, and composite z-scores. Also, the number of participants in each quartile rank for each of the aforementioned scores along with the means, standard deviations, minimums and maximums for rank changes from written exam scores to AC scores and then written exam scores to composite scores were calculated. These descriptive statistics are reported separately depending on the sought-after rank (Tables 4, 5, & 6).

However, to specifically answer the research questions and address corresponding hypotheses, the current research will primarily employ a series of statistical techniques. Specifically, group mean differences will be examined; a series of t-tests will determine if differences exist in scoring for the written exam score, the AC score, or composite scores for male and female participants and/or for white and nonwhite participants (Sims, 2000). In addition, Chi-square tests will then be estimated to examine if an association exists

between participant sex and/or race and the rank-ordered quartile placement based on the written exam score, the AC score, and/or the composite score (Sims, 2000).

IV. RESULTS

Descriptive Statistics

Descriptive statistics are presented in Tables 4, 5, and 6. Specifically, 389 participants participated in a Captain / Commander promotional process (Table 4). Of those, 317 (81.49 percent) were men and 72 (18.51 percent) were women while 227 (58.35 percent) were white and 162 (41.65 percent) were nonwhite. Captain / commander promotional processes included in the current research were conducted at three departments (175 (44.99 percent) from Houston P.D., 118 (30.33 percent) from San Antonio P.D., and 96 (24.68 percent) from Austin P.D.); the captain rank at Dallas P.D. is an appointed position, so there is no competitive promotional process at that rank. The highest number of participants availing themselves of a captain / commander promotional process was 58. The most significant movements from written exam rank-ordered placement to composite score rank-ordered placement, at this rank, was a gain of 34 positions for one participant and loss of 38 positions for another.

Lieutenant rank descriptive statistics are found on Table 5. There were 1356 participants included in the current research that sought promotion to lieutenant and of those, 1119 (82.46 percent) were men and 237 (17.46 percent) were women while 703 (51.81 percent) were white and 653 (48.12 percent) were nonwhite. Lieutenant promotional processes included in the current research were conducted at four departments (300 (22.11 percent) from Houston P.D., 340 (25.06 percent) from San Antonio P.D., 471 (34.71 percent) from Dallas P.D., and 245 (18.05 percent) from Austin P.D.). The highest number of participants availing themselves of a lieutenant promotional process was 110. The most significant movements from written exam rank-ordered

placement to composite score rank-ordered placement, at this rank, was a gain of 81 positions for one participant and loss of 64 positions for another.

Sergeant rank descriptive statistics are found on Table 6. There were 2525 participants included in the current research that sought promotion to sergeant and of those, 2114 (83.72 percent) were men and 411 (16.28 percent) were women while 1249 (49.47 percent) were white and 1276 (50.53 percent) were nonwhite. Sergeant promotional processes included in the current research were conducted at three departments (1090 (43.17 percent) from Houston P.D., 974 (38.57 percent) from Dallas P.D., and 461 (18.26 percent) from Austin P.D.); sergeant promotional processes are competitive, but do not include an AC. The highest number of participants availing themselves of a lieutenant promotional process was 249. The most significant movements from written exam rank-ordered placement to composite score rank-ordered placement, at this rank, was a gain of 134 positions for one participant and loss of 124 positions for another.

Race Differences

Race differences for the written exam score, AC score, and the composite scores at all sought-after ranks were found to be statistically significant ($p < .01$); at all ranks and in all three score categories, nonwhite participants scored lower than white participants (Table 7). An example of some of the more pronounced mean differences were observed for the sergeant rank for the written exam scores (nonwhites: $M = -0.1789$, $SD = 0.9287$; whites: $M = 0.1789$, $SD = 1.0312$, $t = 9.0608$, $p < .01$), AC scores (nonwhites: $M = -0.1014$, $SD = 0.99774$; whites: $M = 0.1035$, $SD = 1.0054$, $t = 5.1911$, $p < .01$), and

composite scores (nonwhites: $M = -0.2136$, $SD = 1.2124$, $t = 8.8236$, $p < .01$; whites: $M = 0.2182$, $SD = 1.2463$, $t = 8.8236$, $p < .01$).

The current research, as hypothesized (H_1), found statistically significant ($p < .01$) race differences in both rank order (Table 8) and rank-ordered quartile placement (Table 9) based on written exam scores, AC scores, and composite scores, and these significant differences were observed across all ranks (Table 8). However, while the current research hypothesized (H_2) that the addition of an AC in the promotional process would improve the rank order and rank-ordered quartile placement for nonwhite participants and this improvement would be consistent across sought-after rank, the results suggested that, at every rank, white participants ranked higher than nonwhite participants on the written exam, the AC, and as a result of the composite score and those differences were found to be statistically significant ($p < .01$). Similar to what was observed in the comparisons of scores discussed above, the larger mean differences in rank order were observed for the sergeant rank (written exam score: $t = -12.5822$, $p < .01$; AC score: $t = -9.9898$, $p < .01$; composite score: $t = -12.6317$, $p < .01$), where nonwhites were, on average, ranked more than 20 positions below the average rank placement of whites.

Finally, merely examining white and nonwhite subgroup mean differences (d) between written exam and composite scores, the current research found that the addition of the AC to the promotional process produced larger subgroup mean differences between white and nonwhite participants at every rank (CPT/CDR written exam score $d = 0.4372$, composite score $d = 0.6632$; LT written exam score $d = 0.1993$, composite score $d = 0.2601$; SGT written exam score $d = 0.3541$, composite score $d = 0.4319$). At the rank of captain / commander, rank placement improvement for nonwhites was minimal; the

mean subgroup difference slightly decreased (written exam score rank $d = -4.6066$; composite score rank $d = -4.4480$). While the number of white and nonwhite participants fell from the first quartile, the number of nonwhite participants fell slightly more substantially, representing an increase in the percentage of white participants in the first quartile (written exam = 70.48 percent, composite score = 73.20 percent). Quartile placement for nonwhite participants only rose in the third quartile, both in the number of nonwhite participants and as a percentage of all participants. The second and fourth quartiles had fewer nonwhite, both in number and as a percentage of all participants.

At the rank of lieutenant, rank placement did not improve for nonwhites; the mean subgroup difference increased between white and nonwhite participants (written exam score rank $d = 5.6144$; composite score rank $d = 5.8420$). Again, while the number of white and nonwhite participants both fell from the first quartile, the number of nonwhite participants fell slightly more substantially, representing an increase in the percentage of white participants in the first quartile (written exam = 57.11 percent, composite score = 61.13 percent). Quartile placement for nonwhite participants only rose in the third quartile, both in the number of nonwhite participants and as a percentage of all participants. The second and fourth quartiles had fewer nonwhite, both in number and as a percentage of all participants.

At the rank of sergeant, rank placement did not improve for nonwhites (written exam score rank $d = -29.6068$; composite score rank $d = -29.7113$). Again, while the number of white and nonwhite participants both fell from the first quartile, the number of nonwhite participants fell slightly more substantially, representing an increase in the percentage of white participants in the first quartile (written exam = 57.11 percent,

composite score = 61.13 percent). Quartile placement for nonwhite participants only rose in the third quartile, both in the number of nonwhite participants and as a percentage of all participants. The second and fourth quartiles had fewer nonwhite, both in number and as a percentage of all participants.

Sex Differences

In contrast to the race differences comparisons, the current research observed noticeably fewer statistically significant sex differences across written exam scores, AC scores, and composite scores, rank order and rank-ordered quartile placement, and the statistically significant effects were not observed consistently across all ranks. For example, no statistically significant sex differences were observed for the written exam scores at any sought-after rank. Comparatively, statistically significant sex differences were observed at the rank of lieutenant in both the AC scores and composite scores, and also in the AC scores at the rank of sergeant. At the lieutenant rank, women scored higher than the men on both the AC (males: $M = -0.0351$, $SD = 0.9913$; females: $M = 0.1657$, $SD = 0.9708$, $t = -2.8822$, $p < .01$) and composite scores (males: $M = -0.2610$, $SD = 1.0074$; females: $M = 0.1233$, $SD = 1.0646$, $t = -1.9802$, $p < .05$). At the sergeant rank, women also scored higher than men on the AC (males: $M = -0.0453$, $SD = 0.9947$; females: $M = 0.2331$, $SD = 0.9737$, $t = -5.2846$, $p < .01$) (Table 10).

While hypothesizing that sex differences would be observed in rank-ordered placement and rank order and rank-ordered quartile placement based on written exam scores, AC scores, and composite scores (H_3), the current research only observed statistically significant ($p < .01$) sex differences in rank-ordered placement for written exam scores at the rank of lieutenant and in rank-ordered placement for AC scores at the

rank of sergeant. On average, men ranked higher than women on rank-ordered placement based on written exam scores at the lieutenant rank (males: $M= 32.9401$, $SD= 25.3321$; females: $M= 38.9873$, $SD= 27.2366$, $t= -3.1423$, $p< .01$) while women ranked higher than men on rank-ordered placement based on AC scores at the sergeant rank (males: $M= 87.9286$, $SD= 61.4000$; females: $M= 76.2482$, $SD= 57.6560$, $t= 3.7176$, $p< .01$) (Table 11). Moreover, participant sex was found to be significantly associated with rank-ordered quartile placement based on composite scores at the rank of captain / commander and based on AC scores at the ranks of lieutenant and sergeant. At the captain / commander rank, men were disproportionately represented in the first quartile based on composite scores (χ^2 test statistic = 11.7989, $p< .01$). However, a disproportionate number of women occupied the first quartile based on AC scores at both the rank of lieutenant (χ^2 test statistic = 12.3091, $p< .01$) and sergeant (χ^2 test statistic = 20.4355, $p< .01$) (Table 12).

Therefore, although the current research hypothesized (H_4) that the addition of an AC in the promotional process would improve the rank order and rank-ordered quartile placement for female participants and this improvement would be consistent across sought-after rank, the analysis demonstrated that the results were generally more mixed in that at times (and for certain ranks) the AC scores were significantly greater for female participants (i.e., sergeant and lieutenant exams) and female participants were generally ranked lower than male participants. In a couple of noteworthy exceptions, the rank-ordered quartile placement was better for female participants based on the AC scores (i.e., sergeant and lieutenant exam).

V. DISCUSSION

The current research sought to determine the effects of race and sex on rank order and rank-ordered quartile placement of promotional process participants in four major metropolitan police departments in Texas. Summarily, the current research hypothesized that the addition of the AC would improve both rank order and rank-ordered quartile placement of both nonwhite and female participants. While race and / or sex differences in both rank-ordered placement and rank-ordered quartile placement (for written exam scores, AC scores, and composite scores) were found at every sought-after rank, not all race and sex effects supported the hypotheses that the addition of the AC would result in rank order and rank-ordered quartile placement improvement for nonwhites and females.

Race Differences

Race differences were observed in written exam, AC, and composite scores as well as in the rank and quartile placements that are driven by those scores, and those differences were observed at every sought-after rank. While nonwhite participants scored lower than white participants on the written exam, the addition of the AC neither eliminated nor reversed the difference between white and nonwhite participants with calculated composite scores. While the AC was responsible for a slight reduction in subgroup mean differences between white and nonwhite participants at the captain / commander rank, subgroup differences remained. Subgroup differences at lieutenant and sergeant ranks, however, increased with the addition of the AC; white participants scored higher than nonwhite participants both before the AC was included to calculate the composite scores, but that difference was larger after calculating composite scores with the AC included. So, subgroup differences are consistent across ranks. Race differences

observed in the current research in written exam scores (and their accompanying rank and quartile placements) supported existing research (e.g. Becton et al., 2008; Hough et al., 2001; Schmidt & Hunter, 1998; Thornton & Rupp, 2006). Also, while findings support the assertions of two studies (e.g. Hough et al., 2001; Thornton and Rupp, 2006) that the AC results in less adverse impact (smaller subgroup mean differences) than written exams, observations in the current study also support the contention made by Dean et al. (2008) that the AC does not reduce subgroup differences as much as others had previously asserted. Racial subgroup mean differences are smaller at all ranks for ACs than for written exams, but when composite scores are calculated (with equal proportions of AC and written exam scores), AC scores do not improve rank-ordered and quartile placement for racial minorities.

Subgroup differences were observed in rank-ordered placement and rank-ordered quartile placement based on written scores, AC scores, and composite scores between whites and racial minorities; white participants consistently scored higher than nonwhite participants. However, regarding the AC's capacity to consistently result in rank and quartile rank improvements for racial minorities, the addition of an AC in the promotional process failed to improve the rank-ordered placement and rank-ordered quartile placement for racial minorities across all sought-after ranks. Lastly, demographic differences identified in descriptive statistics, most notably at Dallas PD, which had promoted more nonwhite males and females to its highest testable rank (60 percent of all lieutenant positions) and that the same department had more non-white employees than white employees assigned at every sought-after rank in their department, may be explained by a phenomenon described by Gustafson (2013), who asserted that larger

departments that offered higher levels of compensation were found to have higher levels of minority diversity in police administration positions. It is not surprising that Bishopp (2013) found that minority status had no statistically significant relationship to the likelihood of promoting, as the data in that study were collected from the Dallas PD.

Sex Differences

Regarding sex differences in promotional process scoring, no sex differences were observed in written exam scores. This observation was in keeping with existing research, which suggested that men and women perform comparably on cognitive ability tests and on job knowledge written exams (Hough et al., 2001). The sought-after rank had an effect on sex differences for AC scores and composite scores. Women scored higher than men on the AC at both the sergeant and lieutenant ranks. This finding, too, is supported by prior research, as Dean et al. (2008) determined that men score lower than women on ACs. No sex differences were observed in AC scores at the captain / commander rank. Also, women scored higher than men on composite scores at the lieutenant rank, whereas no differences were observed in composite scores at either the captain / commander rank or the sergeant rank.

Sought-after rank had an effect on sex differences in rank-ordered placement. While sex differences were observed in rank-ordered placement based on written exams at the lieutenant rank in which men ranked higher than women, there were no sex differences observed in rank-ordered placement at either the ranks of captain / commander or sergeant based on written exam scores. Women ranked higher than men based on AC scores alone at the sergeant rank, but no sex differences were observed at

either the captain / commander rank or the lieutenant rank. No sex differences were observed for composite score rank placement at any sought-after rank.

Again, sought-after rank had an effect on sex differences in quartile placement. A higher proportion of women placed in the first quartile than men based on AC scores at the rank of sergeant and lieutenant whereas no sex differences were observed in quartile placement at the captain / commander rank. A higher proportion of men placed in the first quartile than women based on composite scores at the captain / commander rank but no sex differences were observed in quartile placement at the sergeant or lieutenant ranks. No sex differences were observed in quartile placement based on written exam scores.

A few methodological shortcomings of the current research potentially explain why sex differences were not consistent across all sought-after ranks. According to Dean et al. (2008), AC metrics and exercises (designed to elicit certain capacities thought to predict success in a given organization) can test for a variety of personality differences that could advantage either males or females; however, the current research did not include an examination of the different competencies for which each department sought to assess participants for or how those competencies evolved over time or across ranks. Moreover, while Hsin-Chih (2006) suggested that job analyses used to determine those competencies that prove valuable to incumbent success at a given rank, the current research did not include an examination of job analyses conducted to produce AC exercises or whether any variability in weighting was applied to any given number of sought-after competencies from department to department or from rank to rank. Lastly, with regard to sex differences in performance, whereas Schleicher (2010) found that women improved more on retests than men, the current research did not examine the

proportion of participants, either male or female, who were first-time participants in a given process, or the number of times participants had previously availed themselves of the promotional process at the same sought-after rank.

Limitations

Beyond these potential methodological shortcomings of the current study already addressed, there were a number of limitations to the current study. Each of the departments included in the current research have conducted their promotional processes under different guidelines (including different collective bargaining contracts), under varying constraints, such as differential preparation time allotments between written exam and AC administration, different vendors (e.g. Austin P.D. used three different promotional process vendors between 2000 and 2014, each of whom develop assessment material and administered processes assessment processes differently), and different administrations (what contextual intelligence was provided to vendors during the job analysis process and to raters, as SME's, during AC rater training sessions). No information was available on the measures individual participants took to prepare for either their respective written exam or AC. Moreover, not all departments provided information on participants' completed levels of education, lengths of service (experience), and age. Overall score calculation differences also manifested from one department to another, which included varying points for levels of completed education, including one department that provided education points for participants lacking formal education but, instead, attaining an advanced certification from the Texas Commission on Law Enforcement (TCOLE), formerly Texas Commission on Law Enforcement Standards and Education (TCLEOSE).

Limitations of data include lack of knowledge about who participated in the written exam but did not participate in the accompanying AC process, due to either failing to attain a score at or above the cut-off score for their test process or voluntarily removing themselves from the promotional process before the AC was administered. Also, it was discovered that in the data collection process, written exam cut-off scores varied from department to department. Lastly, data collected from any of the four departments included in the current research were not obtained on occurrences of disparate impact in competitive promotional process outcomes (pass / fail rates) at those ranks below the sought-after ranks (i.e. Houston P.D. – detective; San Antonio P.D. – detective, sergeant; Dallas P.D. – detective; Austin P.D. – detective); it is worth noting that all participants in the current study are included because they have proven themselves capable of performing successfully on more than one written exam. So, future research related to this study should include subgroup mean differences and, as a corollary, disparate impact analyses of competitive promotional processes that precede those in the current research, in which promotion eligibility is only determined by written exam results, including, in as much as data collection allows, those who failed to attain a passing (cut-off) score.

Implications

Police executives, in an effort to maintain their department's legitimacy both within its own ranks and the community they serve, should examine both the efficacy and utility of the addition of an AC to competitive promotional processes in light of the findings of the current study. Moreover, given the widespread and justifiable concern about ramifications that accompany successful disparate impact claims and the associated

legal challenges that, at times, follow, the aforementioned findings may seem troubling. To manage risk, both short- and long-term solutions are available.

Reactive Risk Management – Legal Defensibility

In the short-term, employers should remain poised to defend a promotion process against legal claims of disparate impact. Maintaining transparency and remaining prepared for a legal defense are two reasons why responsible employers keep certain documentation on hand that support use of the assessment instruments that comprise their promotion process. Thornton, et al. (2009) clarified that “supportive evidence includes documents describing the job analysis and competency modeling methods used to study tasks and attributes needed for success, the ways assessment exercises are developed, the training and competence of assessors, and how the program was administered in a standardized manner” (p. 182). An entity employing an AC should also include, as part of any required validity study, an examination into selection procedures that may provide less adverse impact than the procedure of choice (Uniform Guidelines on Employee Selection Procedures, 1978).

Beyond maintaining the documentation provided by the entity that constructed and conducted the promotional process, it is necessary for employers to track results and examine them for any disparate impact against any protected class. Employers who choose to administer testing processes, for the purpose of selection or promotion employment decisions, should retain data about their testing processes that identifies any adverse impact of participants who belong to protected classes (Uniform Guidelines on Employee Selection Procedures, 1978). This information should be deemed necessary, as an employer's failure to maintain data on adverse impact of a selection process allows, by

EEOC guidelines, EEOC enforcement officials to infer that the absence of such data examination and retention suggests that employers are aware that adverse impact exists with that selection process and fail to retain the data because adverse impact exists (Uniform Guidelines on Employee Selection Procedures, 1978).

Two indicators of a potentially impending legal claim against some aspect of the promotion process would include one or more open records requests for the aforementioned information, or an internal grievance filed regarding either the process or the results of the process. According to Aamodt (2005), an employee wishing to seek redress for violations of EEOC guidelines through litigation, he or she must, first, seek resolution through their employer's internal grievance process in a timely manner. Afterwards, the EEOC, upon finding reasonable cause that a discriminatory employment action has occurred, may bring suit against the offending employer (Saad et al., 2000).

If an assessment center is found to result in an adverse impact to a protected class, the Uniform Guidelines suggest a choice of four actions to stave the issue; they include modifying the portion of the process that is causing the adverse impact, excluding the portion of the process found to have caused the adverse impact, replace the portion of the process that is causing the adverse impact, or leave the process in its current configuration if support for the process can be rendered that it is both job-related and a business necessity, and that no equally valid alternative exists (Saad et al., 2000).

Thornton et al., (2009) examined federal court decisions on employment litigation in that could reasonably be construed as a value judgment on the assessment center method. Of the litigation reviewed by Thornton et al. (2009) that were found to have made value judgments on the use of assessment centers or assessment center method,

none referenced developmental assessment centers; all of the cases referenced assessment centers that had a direct impact on employee selection or promotion. Thornton et al., (2009) found that no federal court rulings had been made that were relevant to the assessment center method. Litigation that challenges promotion decisions made based upon AC results will stand upon the how the AC process is managed and if that process is improperly administered (Thornton et al., 2009). Thornton et al. (2009) asserted that recent case law, while focused on the manner in which ACs were administered, indicated that federal judges have been supportive of the AC method when they were well-managed.

Each method that is utilized as part of the selection process (used to combine methods, apply cutoff scores, or to rank participants) should be articulated for understanding, as should any information that would tend to reveal a presence or absence of any adverse impact (Uniform Guidelines on Employee Selection Procedures, 1978). So, EEOC guidelines suggest that each component of a promotion process should be individually evaluated for adverse impact if the process itself is found to have had an adverse impact on any subgroup (Uniform Guidelines on Employee Selection Procedures, 1978).

Thornton et al. (2009) provides specific strategies for managing risk regarding civil liability and the process of assessment center administration. Beyond working within the provisions of the guiding documents, they suggest that an organization maintain documentation that provides evidentiary support that the organization took reasonable measures to ensure that the AC was designed so as to be related to the sought-after position (Thornton et al., 2009).

Another strategy for mitigating risk of adverse impact, thus managing risk, is formulating AC exercises with less cognitive load. De Soete et al. (2013) suggests two strategies for reducing cognitive load. Drawing on results of a study by Ployhart and Holtz (2008), they suggest that administrators develop test instruments that reduce the cognitive load in order to drive a reduction in the differences in results between ethnic subgroups. One strategy is to increase “the stimulus fidelity (e.g. the extent to which the stimuli presented by the instrument resemble the stimuli in the actual job situation)” (De Soete et al., 2013, p. 241). A second strategy, according to Bobko and Roth (2013), is to increase “the fidelity of the response format (e.g., the extent to which the instrument’s response format resembles the response requirements during on-the-job behavior)” (De Soete et al., 2013, p. 241). According to De Soete et al. (2013), there is a direct relationship between cognitive load and ethnic subgroup differences; as one increases, so will the other.

Proactive Risk Management – Process Fairness

In the long-term, departments can proactively manage risk by taking reasonable measures to bolster fairness perceptions of both the administration and content of their respective promotional processes and by understanding that equity in employment decisions is a cornerstone to providing for a successful, productive workforce. To apply some context, police officers at every rank are still police officers; they are trained investigators. Police officers foster a strong sense of justice, are discerning about small details and how they contribute to a given result, and are trained to question a set of circumstances until they can reach logical conclusions about a given outcome. Add to it

the fact that anything that personally impacts officers' lives in a significant way, such as a promotion and raise in pay, and you have everyone's attention.

In as much as promotional processes are concerned, officers generally maintain a keen understanding of fairness as an expectation, both as a necessary byproduct of their own work and, thus, of those processes that impact their careers as directly as the civil service promotion processes. Further, police officers maintain a capacity to evaluate the quality of the leadership they receive from those appointed to lead, and, thus, are specifically poised to capably evaluate the outcomes of a given promotional process to which they are all subject. They can trust the process is fair when they know it properly identifies leadership prowess and management competence as they have come to understand it. They also know when the process is transparent – when it is clearly discernible how one promotional candidate finished ahead or behind another.

Posing a clear difference between fairness and the perception of fairness, proponents of the AC process have broadly concluded that AC's are fair to all participants, regardless of their race, sex or age (Thornton & Rupp, 2006). On the other hand, Krause and Thornton (2009) suggested that organizations which employ the AC method should work to solidify process legitimacy through careful development, administration and process reviews. Specific to the process legitimacy discussion, AC participants should receive constructive criticism from assessors, so that they may apply lessons learned to future encounters similar to those experienced in AC exercises (International Task Force on Assessment Center Guidelines, 2015). Moreover, the Guidelines assert that participants have a right to view any reports that pertain to their individual AC performance, suggesting that process legitimacy in the view of participants

is paramount to its overall success (International Task Force on Assessment Center Guidelines, 2015). Ensuring that promotion candidates receive accurate, constructive feedback and making reports (relevant to a given participant 's performance) available bolster confidence, and a belief in the transparency of a promotion process fosters trust in the process' fairness. While implementation of these measures may seem straightforward and may be capably implemented in the short-term, building trust where it was previously nonexistent should be considered a long-term, protracted endeavor.

Table 1a. Cities included in Dataset

<i>City</i>	<i>Number - Most Populous Cities in U.S.</i>	<i>Population</i>	<i>Department Records included in dataset</i>
Austin, Tx.	11 th	950,715	Austin PD
Dallas, Tx.	9 th	1,341,075	Dallas PD
San Antonio, Tx.	7 th	1,511,946	San Antonio PD
Houston, Tx.	4 th	2,312,717	Houston PD

*Source: United States Census Bureau
(2017)*

1b. Population Demographic Data by City

<i>City</i>	<i>White</i>	<i>Nonwhite</i>	<i>Male</i>	<i>Female</i>
Austin, Tx.	48.6%	51.4%	50.6%	49.4%
Dallas, Tx.	29.1%	70.9%	49.6%	50.4%
San Antonio, Tx.	25.1%	74.9%	49.2%	50.8%
Houston, Tx.	24.9%	75.1%	50.1%	49.9%

1c. Sworn Staffing at Sought-After Positions in Current Research

<i>Department</i>	<i>White</i>		<i>Nonwhite</i>		<i>Rank</i>
	<i>Male</i>	<i>Female</i>	<i>Male</i>	<i>Female</i>	
Austin PD	16	2	1	0	Commander
	48	5	11	5	Lieutenant
	126	13	44	7	Sergeant
	190	20	56	12	Total
Dallas PD	21	3	27	9	Lieutenant
	160	20	154	65	Sergeant
	181	23	181	74	Total
San Antonio PD	11	1	10	0	Captain
	25	2	21	1	Lieutenant
	36	3	31	1	Total
Houston PD	25	7	7	5	Commander
	97	73	21	18	Lieutenant
	421	264	58	68	Sergeant
	543	344	86	91	Total

<i>Department</i>	<i>White</i>		<i>Nonwhite</i>		<i>Rank</i>
	<i>Male</i>	<i>Female</i>	<i>Male</i>	<i>Female</i>	
Austin PD	84.2%	10.5%	5.3%	0.0%	Commander
	69.6%	7.2%	15.9%	7.2%	Lieutenant
	66.3%	6.8%	23.2%	3.7%	Sergeant
	68.3%	7.2%	20.1%	4.3%	Total
Dallas PD	35.0%	5.0%	45.0%	15.0%	Lieutenant
	40.1%	5.0%	38.6%	16.3%	Sergeant
	39.4%	5.0%	39.4%	16.1%	Total
San Antonio PD	50.0%	4.5%	45.5%	0.0%	Captain
	51.0%	4.1%	42.9%	2.0%	Lieutenant
	50.7%	4.2%	43.7%	1.4%	Total
Houston PD	56.8%	15.9%	15.9%	11.4%	Commander
	46.4%	34.9%	10.0%	8.6%	Lieutenant
	51.9%	32.6%	7.2%	8.4%	Sergeant
	51.0%	32.3%	8.1%	8.6%	Total

Table 2a. Promotional events (cohorts) by department (n=4270)

<i>Jurisdiction</i>	<i>Cohorts</i>	<i>Sought-after Ranks Included</i>	<i>Years Included</i>	<i>Participants</i>
Austin PD (APD)	25	Commander, Lieutenant, Sergeant	2000 - 2014	802
Dallas PD (DPD)	11	Lieutenant, Sergeant	2001 - 2017	1445
San Antonio PD (SAPD)	17	Captain, Lieutenant	2003 - 2017	458
Houston PD (HPD)	14	Captain, Lieutenant, Sergeant	2009 - 2018	1565
Total	67		Total	4270

Table 2b. Participants by jurisdiction (n=4270)

<i>Jurisdiction</i>	<i>White</i>		<i>Nonwhite</i>		<i>Participants</i>
	<i>Male</i>	<i>Female</i>	<i>Male</i>	<i>Female</i>	
Austin PD (APD)	546 / 68.08%	58 / 7.23%	168 / 20.95%	30 / 3.74%	802
Dallas PD (DPD)	551 / 38.13%	72 / 4.98%	598 / 41.38%	224 / 15.50%	1445
San Antonio PD (SAPD)	234 / 36.81%	14 / 3.06%	171 / 37.34%	39 / 8.52%	458
Houston PD (HPD)	576 / 36.41%	128 / 8.18%	706 / 45.11%	155 / 9.90%	1565
Total	1907 / 44.66%	272 / 6.37%	1643 / 38.48%	448 / 10.49%	4270

Table 2c. Promotional events (cohorts) by rank (n=4270)

<i>Sought-after Rank</i>	<i>Cohorts</i>	<i>Jurisdictions</i>	<i>Years included</i>	<i>Participants</i>
Captain / Commander	21	APD (8), SAPD (8), HPD (5)	2000-2018	389
Lieutenant	27	APD (9), SAPD (9), HPD (4), DPD (5)	2000-2017	1356
Sergeant	19	APD (8), DPD (6), HPD (5)	2002-2018	2525
Total	67		Total	4270

Table 2d. Participants by rank (n=4270)

<i>Sought-after Rank</i>	<i>White</i>		<i>Nonwhite</i>		<i>Participants</i>
	<i>Male</i>	<i>Female</i>	<i>Male</i>	<i>Female</i>	
Captain / Commander	191 / 49.10%	36 / 9.25%	126 / 32.39%	36 / 9.25%	389
Lieutenant	619 / 45.65%	84 / 6.19%	500 / 36.87%	153 / 11.28%	1356
Sergeant	1097 / 43.45%	152 / 6.02%	1017 / 40.28%	259 / 10.26%	2525
Total	1907 / 44.66%	272 / 6.37%	1643 / 38.48%	448 / 10.49%	4270

Table 2e. Raw score descriptive statistics by cohort - Captain / Commander

Cohort	Dept	Rank	Year	Count	Raw Exam		Raw AC	
					Mean	SD	Mean	SD
1	APD	CDR	2000	6	83	6.6533	83.4538	3.0021
2	APD	CDR	2001	7	87	4.7559	78.1099	4.8945
3	APD	CDR	2004	8	86	6.2092	61.3281	5.4627
4	APD	CDR	2006	15	84	6.8292	57.4361	13.1325
5	APD	CDR	2008	10	81	7.5248	70.1263	10.0376
6	APD	CDR	2010	Table	81	6.3185	65.7714	14.5346
7	APD	CDR	2012	19	88	7.2713	67.3767	8.3803
8	APD	CDR	2014	16	86	7.5230	70.1487	9.0187
26	HPD	CPT	2009	25	93	2.9648	67.0644	11.5715
27	HPD	CPT	2012	31	83	5.1171	66.1198	8.0539
28	HPD	CPT	2014	32	88	6.0873	62.6726	12.4583
29	HPD	CPT	2016	58	88	7.4532	68.9251	6.4330
30	HPD	CDR	2018	29	91	5.7531	77.4222	5.7141
40	SAPD	CPT	2003	11	91	5.5857	82.3909	6.3709
41	SAPD	CPT	2005	25	93	4.7159	83.2644	8.6421
42	SAPD	CPT	2007	31	96	2.8946	89.2006	5.6359
43	SAPD	CPT	2008	12	97	2.4541	89.0476	4.9237
44	SAPD	CPT	2011	10	87	5.9898	67.2192	12.6658
45	SAPD	CPT	2013	9	91	6.4636	74.3125	12.8854
46	SAPD	CPT	2015	12	94	5.5179	76.1454	10.2737
47	SAPD	CPT	2017	8	95	2.1213	80.7522	9.3148

Table 2f. Raw score descriptive statistics by cohort - Lieutenant

Cohort	Dept	Rank	Year	Count	Raw Exam		Raw AC	
					Mean	SD	Mean	SD
9	APD	LT	2000	9	82	8.6168	81.4710	7.0300
10	APD	LT	2001	18	82	6.9790	78.6518	8.3203
11	APD	LT	2003	19	80	4.6948	57.5187	9.3649
12	APD	LT	2004	19	78	5.8330	66.8780	8.6504
13	APD	LT	2006	32	84	8.9190	63.1255	10.1044
14	APD	LT	2008	34	80	5.5220	66.3525	7.8099
15	APD	LT	2010	28	83	7.5483	68.7348	9.6435
16	APD	LT	2012	49	83	6.6612	68.9618	7.5173
17	APD	LT	2014	38	80	5.9839	67.5569	7.8855
31	HPD	LT	2009	50	91	3.0189	70.7784	10.7728
32	HPD	LT	2012	72	87	4.6375	71.9815	8.9563
33	HPD	LT	2014	75	89	2.9645	67.0425	8.5789
34	HPD	LT	2016	122	87	8.9427	66.9660	6.5064
48	SAPD	LT	2004	47	92	6.3562	59.1489	3.4890
49	SAPD	LT	2005	55	91	10.1832	87.8287	7.3914
50	SAPD	LT	2007	31	97	1.3757	94.5624	2.7659
51	SAPD	LT	2008	40	91	4.9920	85.0373	6.4793
52	SAPD	LT	2010	50	92	5.9989	83.4253	7.6881
53	SAPD	LT	2011	46	89	6.2866	56.1051	8.6371
54	SAPD	LT	2013	20	88	6.0776	80.7055	11.2657
55	SAPD	LT	2015	43	90	5.8431	79.5556	8.6980
56	SAPD	LT	2017	25	92	4.8672	75.2683	11.4650
57	DPD	LT	2001	90	74	7.5657	7.5580	0.7262
58	DPD	LT	2005	100	67	7.0683	3.4094	0.5609
59	DPD	LT	2007	78	74	7.8262	3.4450	0.4472
60	DPD	LT	2013	93	70	7.0879	3.3343	0.5207
61	DPD	LT	2017	110	64	7.4060	7.3163	0.9397

Table 2g. Raw score descriptive statistics by cohort – Sergeant

Cohort	Dept	Rank	Year	Count	Raw Exam		Raw AC	
					Mean	SD	Mean	SD
18	APD	SGT	2002	41	83	4.7330	76.5184	10.0118
19	APD	SGT	2003	39	85	5.1216	60.4213	11.1606
20	APD	SGT	2005	63	82	6.0956	63.4310	8.5630
21	APD	SGT	2007	58	83	6.1928	62.7986	12.8946
22	APD	SGT	2008	81	84	7.0237	61.6394	10.9853
23	APD	SGT	2010	81	83	6.6829	61.4324	12.9892
24	APD	SGT	2013	38	77	4.7288	68.1510	5.6985
25	APD	SGT	2014	60	80	6.3969	68.0787	9.7101
35	HPD	SGT	2010	202	82	4.2867	57.5612	12.0534
36	HPD	SGT	2012	199	84	4.7036	76.4543	7.6087
37	HPD	SGT	2014	250	87	4.5573	63.1852	14.1485
38	HPD	SGT	2016	249	84	5.5404	68.7752	8.0706
39	HPD	SGT	2018	241	81	6.1915	69.1056	9.2949
62	DPD	SGT	2002	139	71	6.6553	7.4571	0.8763
63	DPD	SGT	2005	197	64	6.9404	3.4415	0.5911
64	DPD	SGT	2008	132	68	7.4314	3.3838	0.5796
65	DPD	SGT	2011	146	67	6.2174	3.5094	0.5335
66	DPD	SGT	2014	153	61	6.8089	3.3014	0.5857
67	DPD	SGT	2017	207	62	7.8612	7.4976	1.0593

Table 3. Variables

Variables	Attributes	Scale	Variable Name	IV / DV
Non-White	0=White, 1=Non-White	Nominal	NONWHITE	IV
Female	0=Male, 1=Female	Nominal	FEMALE	IV
Written Exam Z-Score	Numeric	Ratio	ZScoreWE	DV
Assessment Center Z-Score	Numeric	Ratio	ZScoreAC	DV
Composite Z-Score	Numeric	Ratio	ZScoreComp	DV
Written Exam Eligibility List Rank Placement	Numeric	Ordinal	RankWE	DV
Assessment Center Eligibility List Rank Placement	Numeric	Ordinal	RankAC	DV
Composite Score Eligibility List Rank Placement	Numeric	Ordinal	RankComp	DV
Quartile - Written Exam Eligibility List Placement	1=Highest, 4=Lowest	Ordinal	QuartWE	DV
Quartile - AC Eligibility List Placement	1=Highest, 4=Lowest	Ordinal	QuartAC	DV
Quartile - Composite Score Eligibility List Placement	1=Highest, 4=Lowest	Ordinal	QuartComp	DV

Table 4. Captain/Commander Test Takers (n=389)

	<u>Mean (or n/%)</u>	<u>Standard Deviation</u>	<u>Minimum</u>	<u>Maximum</u>
<u>Sex</u>				
Male	317 / 81.49%			
Female	72 / 18.51%			
<u>Race</u>				
White	227 / 58.35%			
Nonwhite	162 / 41.65%			
<u>Department</u>				
Houston	175 / 44.99%			
San Antonio	118 / 30.33%			
Dallas	0 / 0.00%			
Austin	96 / 24.68%			
Written Exam Score (z-score)	0.0000	0.9739	-3.7880	2.0996
Written Exam Score Rank	14	12	1	58
Written Score Quartile Rank	1 st	2 nd	3 rd	4 th
	105 / 26.92%	111 / 28.46%	92 / 23.59%	81 / 20.77%
Assessment Center Score (z-score)	0.0000	0.9739	-3.4731	2.0391
Assessment Center Score Rank	14	12	1	58
Assessment Center Score Quartile Rank	1 st	2 nd	3 rd	4 th
	96 / 24.62%	108 / 27.69%	99 / 25.38%	86 / 22.05%
Composite Score (z-score)	0.0000	1.3531	-4.8277	3.3276
Composite Score Rank	14	12	1	58
Composite Score Quartile Rank	1 st	2 nd	3 rd	4 th
	97 / 24.87%	103 / 26.41%	102 / 26.15%	87 / 22.31%
Rank Change from Written Exam Score (z-score) to Assessment Center Score (z-score)	0	11	-48	46
Rank Change from Written Exam Score (z-score) to Composite Score (z-score)	0	7	-34	38

Table 5. Lieutenant Test Takers (n=1356)

	<u>Mean (or n/%)</u>	<u>Standard Deviation</u>	<u>Minimum</u>	<u>Maximum</u>
<u>Sex</u>				
Male	1119 / 82.46%			
Female	237 / 17.46%			
<u>Race</u>				
White	703 / 51.81%			
Nonwhite	653 / 48.12%			
<u>Department</u>				
Houston	300 / 22.11%			
San Antonio	340 / 25.06%			
Dallas	471 / 34.71%			
Austin	245 / 18.05%			
Written Exam Score (z-score)	0.0000	0.9904	-8.2098	3.3302
Written Exam Score Rank	26	26	1	110
Written Score Quartile Rank	1 st	2 nd	3 rd	4 th
	387 / 28.52%	356 / 26.23%	325 / 23.95%	288 / 21.22%
Assessment Center Score (z-score)	0.0000	0.9904	-3.0913	2.5156
Assessment Center Score Rank	34	26	1	110
Assessment Center Score Quartile Rank	1 st	2 nd	3 rd	4 th
	387 / 28.52%	359 / 26.46%	331 / 24.39%	323 / 23.80%
Composite Score (z-score)	0.0000	1.0188	-9.4251	3.8609
Composite Score Rank	34	26	1	110
Composite Score Quartile Rank	1 st	2 nd	3 rd	4 th
	387 / 28.52%	372 / 27.41%	319 / 23.51%	328 / 24.17%
Rank Change from Written Exam Score (z-score) to Assessment Center Score (z-score)	0	25	-92	98
Rank Change from Written Exam Score (z-score) to Composite Score (z-score)	0	14	-81	64

Table 6. Sergeant Test Takers (n=2525)

	<u>Mean (or n/%)</u>	<u>Standard Deviation</u>	<u>Minimum</u>	<u>Maximum</u>
<u>Sex</u>				
Male	2114 / 83.72%			
Female	411 / 16.28%			
<u>Race</u>				
White	1249 / 49.47%			
Nonwhite	1276 / 50.53%			
<u>Department</u>				
Houston	1090 / 43.17%			
San Antonio	0 / 0.00%			
Dallas	974 / 38.57%			
Austin	461 / 18.26%			
Written Exam Score (z-score)	0.0000	0.9964	-2.0999	3.8500
Written Exam Score Rank	86	61	1	249
Written Score Quartile Rank	1 st	2 nd	3 rd	4 th
	666 / 26.38%	688 / 27.25%	602 / 23.84%	569 / 22.53%
Assessment Center Score (z-score)	0.0000	0.9964	-4.1091	2.6675
Assessment Center Score Rank	86	61	1	249
Assessment Center Score Quartile Rank	1 st	2 nd	3 rd	4 th
	666 / 26.38%	636 / 25.19%	637 / 25.23%	620 / 24.55%
Composite Score (z-score)	0.0000	1.2479	-5.3313	5.0693
Composite Score Rank	86	61	1	249
Composite Score Quartile Rank	1 st	2 nd	3 rd	4 th
	666 / 26.38%	621 / 24.59%	654 / 25.90%	625 / 24.75%
Rank Change from Written Exam Score (z-score) to Assessment Center Score (z-score)	0	58	-169	223
Rank Change from Written Exam Score (z-score) to Composite Score (z-score)	0	30	-134	124

Table 7. Race differences in scores

CPT/CDR				
	<u>Mean</u>	<u>SD</u>	<u>t-test</u>	<u>p value</u>
<u>Written Exam Score (z score)</u>	0.0000	0.9739		
White	0.1821	0.9471	4.4631	**0.0000
Non-White	-0.2551	0.9563		
<u>Assessment Center Score (z score)</u>	0.0000	0.9739		
White	0.1446	0.9497	3.5023	**0.0005
Non-White	-0.2026	0.9741		
<u>Composite Score (z score)</u>	0.0000	1.3531		
White	0.2762	1.3308	4.93	**0.0000
Non-White	-0.3870	1.2915		

LT				
	<u>Mean</u>	<u>SD</u>	<u>t-test</u>	<u>p value</u>
<u>Written Exam Score (z score)</u>	0.0000	0.9904		
White	0.0960	0.9897	3.7221	**0.0002
Non-White	-0.1033	0.9813		
<u>Assessment Center Score (z score)</u>	0.0000	0.9904		
White	0.0913	1.0172	3.5465	**0.0004
Non-White	-0.0983	0.9517		
<u>Composite Score (z score)</u>	0.0000	1.0188		
White	0.1253	1.0111	4.7345	**0.0000
Non-White	-0.1349	1.0107		

SGT				
	<u>Mean</u>	<u>SD</u>	<u>t-test</u>	<u>p value</u>
<u>Written Exam Score (z score)</u>	0.0000	0.9964		
White	0.1789	1.0312	9.0608	**0.0000
Non-White	-0.1752	0.9287		
<u>Assessment Center Score (z score)</u>	0.0000	0.9964		
White	0.1035	1.0054	5.1911	**0.0000
Non-White	-0.1014	0.9774		
<u>Composite Score (z score)</u>	0.0000	1.2479		
White	0.2182	1.2463	8.8236	**0.0000
Non-White	-0.2136	1.2124		

* $p < .05$. ** $p < .01$.

Table 8. Race differences in rank-ordered placement

CPT/CDR				
	<u>Mean</u>	<u>SD</u>	<u>t-test</u>	<u>p value</u>
<u>Written Exam Score Rank</u>	13.8303	11.8131		
White	11.9119	11.2274	-3.8098	**0.0002
Non-White	16.5185	12.1201		
<u>Assessment Center Score Rank</u>	13.8303	11.8131		
White	12.3304	11.5403	-2.9791	**0.0031
Non-White	15.9321	11.9055		
<u>Composite Score Rank</u>	13.8303	11.8131		
White	11.9780	11.6008	-3.7183	**0.0002
Non-White	16.4259	11.6522		
LT				
	<u>Mean</u>	<u>SD</u>	<u>t-test</u>	<u>p value</u>
<u>Written Exam Score Rank</u>	33.9971	25.7674		
White	28.3827	22.4503	-8.4784	**0.0000
Non-White	40.0414	27.6877		
<u>Assessment Center Score Rank</u>	33.9971	25.7674		
White	28.8421	23.3886	-7.7697	**0.0000
Non-White	39.5467	27.0440		
<u>Composite Score Rank</u>	33.9971	25.7674		
White	28.1551	22.8671	-8.8531	**0.0000
Non-White	40.2864	1.0648		
SGT				
	<u>Mean</u>	<u>SD</u>	<u>t-test</u>	<u>p value</u>
<u>Written Exam Score Rank</u>	86.0273	60.9480		
White	71.0657	58.3070	-12.5822	**0.0000
Non-White	100.6724	59.9325		
<u>Assessment Center Score Rank</u>	86.0273	60.9480		
White	74.0200	58.1028	-9.9898	**0.0000
Non-White	97.7806	61.3973		
<u>Composite Score Rank</u>	86.0273	60.9480		
White	71.0128	57.7972	-12.6317	**0.0000
Non-White	100.7241	60.3885		

* $p < .05$. ** $p < .01$.

Table 9. Race differences in quartile placement

	CPT / CDR						
	ALL	Q1	Q2	Q3	Q4	Chi-square	p-value
<u>Written Score Quartile Rank</u>	389	105	111	92	81		
White	227	74	71	51	31	21.5515	**0.000
Column %	58.35%	70.48%	63.96%	55.43%	38.27%		
Row %		32.60%	31.28%	22.47%	13.66%		
Non-White	162	31	40	41	50		
Column %	41.65%	29.52%	36.04%	44.57%	61.73%		
Row %		19.14%	24.69%	25.31%	30.86%		
<u>Assessment Center Score Quartile Rank</u>	389	96	108	99	86		
White	227	67	67	52	41	11.1908	*0.011
Column %	58.35%	63.81%	60.36%	56.52%	50.62%		
Row %		29.52%	29.52%	22.91%	18.06%		
Non-White	162	29	41	47	45		
Column %	41.65%	27.62%	36.94%	51.09%	55.56%		
Row %		17.90%	25.31%	29.01%	27.78%		
<u>Composite Score Quartile Rank</u>	389	97	103	102	87		
White	227	71	67	50	39	20.899	**0.000
Column %	58.35%	67.62%	60.36%	54.35%	48.15%		
Row %		31.28%	29.52%	22.03%	17.18%		
Non-White	162	26	36	52	48		
Column %	41.65%	24.76%	32.43%	56.52%	59.26%		
Row %		16.05%	22.22%	32.10%	29.63%		

	LT						
	ALL	Q1	Q2	Q3	Q4	Chi-square	p-value
<u>Written Score Quartile Rank</u>	1356	387	356	325	288		
White	703	221	195	158	129	12.6114	**0.006
Column %	51.84%	57.11%	54.78%	48.62%	44.79%		
Row %		31.44%	27.74%	22.48%	18.35%		
Non-White	653	166	161	167	159		
Column %	48.16%	42.89%	45.22%	51.38%	55.21%		
Row %		25.42%	24.66%	25.57%	24.35%		
<u>Assessment Center Score Quartile Rank</u>	1356	343	359	331	323		
White	703	205	185	162	151	13.112	**0.004
Column %	51.84%	52.97%	51.97%	49.85%	52.43%		
Row %		29.16%	26.32%	23.04%	21.48%		
Non-White	653	138	174	169	172		
Column %	48.16%	35.66%	48.88%	52.00%	59.72%		
Row %		21.13%	26.65%	25.88%	26.34%		
<u>Composite Score Quartile Rank</u>	1356	337	372	319	328		
White	703	206	186	161	150	17.2897	**0.001
Column %	51.84%	53.23%	52.25%	49.54%	52.08%		

Row %		29.30%	26.46%	22.90%	21.34%		
Non-White	653	131	186	158	178		
Column %	48.16%	33.85%	52.25%	48.62%	61.81%		
Row %		20.06%	28.48%	24.20%	27.26%		

SGT

	ALL	Q1	Q2	Q3	Q4	Chi-square	p-value
<u>Written Score Quartile Rank</u>	2525	666	688	602	569		
White	1249	418	348	252	231	79.2815	**0.000
Column %	49.47%	62.76%	50.58%	41.86%	40.60%		
Row %		33.47%	27.86%	20.18%	18.49%		
Non-White	1276	248	340	350	338		
Column %	50.53%	37.24%	49.42%	58.14%	59.40%		
Row %		19.44%	26.65%	27.43%	26.49%		
<u>Assessment Center Score Quartile Rank</u>	2525	632	636	637	620		
White	1249	355	329	297	268	24.3850	**0.000
Column %	49.47%	53.30%	47.82%	49.34%	47.10%		
Row %		28.42%	26.34%	23.78%	21.46%		
Non-White	1276	277	307	340	352		
Column %	50.53%	41.59%	44.62%	56.48%	61.86%		
Row %		21.71%	24.06%	26.65%	27.59%		
<u>Composite Score Quartile Rank</u>	2525	625	621	654	625		
White	1249	392	319	275	263	72.8544	**0.000
Column %	49.47%	58.86%	46.37%	45.68%	46.22%		
Row %		31.39%	25.54%	22.02%	21.06%		
Non-White	1276	233	302	379	362		
Column %	50.53%	34.98%	43.90%	62.96%	63.62%		
Row %		18.26%	23.67%	29.70%	28.37%		

Bold Italics indicates higher than expected

* $p < .05$. ** $p < .01$.

Table 10. Sex differences in scores

CPT/CDR				
	<u>Mean</u>	<u>SD</u>	<u>t-test</u>	<u>p value</u>
<u>Written Exam Score (z score)</u>	0.0000	0.9739		
Male	0.0129	0.9800	0.5585	0.5776
Female	-0.0568	0.9512		
<u>Assessment Center Score (z score)</u>	0.0000	0.9739		
Male	-0.0341	1.0117	-1.7128	0.0891
Female	0.1500	0.7742		
<u>Composite Score (z score)</u>	0.0000	1.3531		
Male	-0.0195	1.3944	-0.6708	0.5036
Female	0.0861	1.1587		

LT				
	<u>Mean</u>	<u>SD</u>	<u>t-test</u>	<u>p value</u>
<u>Written Exam Score (z score)</u>	0.0000	0.9904		
Male	0.0103	0.9860	0.8219	0.4117
Female	-0.0488	1.0114		
<u>Assessment Center Score (z score)</u>	0.0000	0.9904		
Male	-0.0351	0.9913	-2.8822	**0.0042
Female	0.1657	0.9708		
<u>Composite Score (z score)</u>	0.0000	1.0188		
Male	-0.2610	1.0074	-1.9802	*0.0485
Female	0.1233	1.0646		

SGT				
	<u>Mean</u>	<u>SD</u>	<u>t-test</u>	<u>p value</u>
<u>Written Exam Score (z score)</u>	0.0000	0.9964		
Male	0.0157	0.0217	1.7932	0.0735
Female	-0.0806	0.9954		
<u>Assessment Center Score (z score)</u>	0.0000	0.9964		
Male	-0.0453	0.9947	-5.2846	**0.0000
Female	0.2331	0.9737		
<u>Composite Score (z score)</u>	0.0000	1.2479		
Male	-0.0199	1.2554	-1.8712	0.0618
Female	0.1025	1.2047		

* $p < .05$. ** $p < .01$.

Table 11. Sex differences in rank-ordered placement

CPT/CDR				
	<u>Mean</u>	<u>SD</u>	<u>t-test</u>	<u>p value</u>
<u>Written Exam Score Rank</u>	13.8303	11.8131		
Male	13.2965	11.7646	-1.8705	0.0642
Female	16.1806	11.8206		
<u>Assessment Center Score Rank</u>	13.8303	11.8131		
Male	13.4606	11.7402	-1.2736	0.2057
Female	15.4583	12.0771		
<u>Composite Score Rank</u>	13.8303	11.8131		
Male	13.4416	11.9786	-1.441	0.1524
Female	15.5417	10.9698		
LT				
	<u>Mean</u>	<u>SD</u>	<u>t-test</u>	<u>p value</u>
<u>Written Exam Score Rank</u>	33.9971	25.7674		
Male	32.9401	25.3321	-3.1423	**0.0018
Female	38.9873	27.2366		
<u>Assessment Center Score Rank</u>	33.9971	25.7674		
Male	33.9178	25.7164	-0.244	0.8074
Female	34.3713	26.0584		
<u>Composite Score Rank</u>	33.9971	25.7674		
Male	33.4754	25.5578	-1.5771	0.1157
Female	36.4599	26.6526		
SGT				
	<u>Mean</u>	<u>SD</u>	<u>t-test</u>	<u>p value</u>
<u>Written Exam Score Rank</u>	86.0273	60.9480		
Male	85.2696	61.1035	-1.4335	0.1522
Female	89.9246	60.0642		
<u>Assessment Center Score Rank</u>	86.0273	60.9480		
Male	87.9286	61.4000	3.7176	**0.0002
Female	76.2482	57.6560		
<u>Composite Score Rank</u>	86.0273	60.9480		
Male	86.7895	61.5316	1.4874	0.1374
Female	82.1071	57.7658		

* $p < .05$. ** $p < .01$.

Table 12. Sex differences in quartile placement

CPT / CDR							
	ALL	Q1	Q2	Q3	Q4	Chi-square	p-value
<u>Written Score Quartile Rank</u>	389	105	111	92	81		
Male	317	87	90	75	65	0.2255	0.973
Column %	81.49%	82.86%	81.08%	81.52%	80.25%		
Row %		27.44%	28.39%	23.66%	20.50%		
Female	72	18	21	17	16		
Column %	18.51%	17.14%	18.92%	18.48%	19.75%		
Row %		25.00%	29.17%	23.61%	22.22%		
<u>Assessment Center Score Quartile Rank</u>	389	96	108	99	86		
Male	317	77	86	78	76	3.5321	0.317
Column %	81.49%	73.33%	77.48%	84.78%	93.83%		
Row %		24.29%	27.13%	24.61%	23.97%		
Female	72	19	22	21	10		
Column %	18.51%	18.10%	19.82%	22.83%	12.35%		
Row %		26.39%	30.56%	29.17%	13.89%		
<u>Composite Score Quartile Rank</u>	389	97	103	102	87		
Male	317	85	77	77	78	11.7989	**0.008
Column %	81.49%	80.95%	69.37%	83.70%	96.30%		
Row %		26.81%	24.29%	24.29%	24.61%		
Female	72	12	26	25	9		
Column %	18.51%	11.43%	23.42%	27.17%	11.11%		
Row %		16.67%	36.11%	34.72%	12.50%		

LT							
	ALL	Q1	Q2	Q3	Q4	Chi-square	p-value
<u>Written Score Quartile Rank</u>	1356	387	356	325	288		
Male	1119	321	299	268	231	1.6489	0.648
Column %	82.52%	82.95%	83.99%	82.46%	80.21%		
Row %		28.69%	26.72%	23.95%	20.64%		
Female	237	66	57	57	57		
Column %	17.48%	17.05%	16.01%	17.54%	19.79%		
Row %		27.85%	24.05%	24.05%	24.05%		
<u>Assessment Center Score Quartile Rank</u>	1356	343	359	331	323		
Male	1119	265	305	269	280	12.3091	**0.006
Column %	82.52%	68.48%	85.67%	82.77%	97.22%		
Row %		23.68%	27.26%	24.04%	25.02%		
Female	237	78	54	62	43		
Column %	17.48%	20.16%	15.17%	19.08%	14.93%		
Row %		32.91%	22.78%	26.16%	18.14%		
<u>Composite Score Quartile Rank</u>	1356	337	372	319	328		
Male	1119	267	313	262	277	4.0896	0.252
Column %	82.52%	68.99%	87.92%	80.62%	96.18%		

Row %		23.86%	27.97%	23.41%	24.75%		
Female	237	70	59	57	51		
Column %	17.48%	18.09%	16.57%	17.54%	17.71%		
Row %		29.54%	24.89%	24.05%	21.52%		

SGT

	ALL	Q1	Q2	Q3	Q4	Chi-square	p-value
<u>Written Score Quartile Rank</u>	2525	666	688	602	569		
Male	2114	564	588	495	467	4.1099	0.250
Column %	83.72%	84.68%	85.47%	82.23%	82.07%		
Row %		26.68%	27.81%	23.42%	22.09%		
Female	411	102	100	107	102		
Column %	16.28%	15.32%	14.53%	17.77%	17.93%		
Row %		24.82%	24.33%	26.03%	24.82%		
<u>Assessment Center Score Quartile Rank</u>	2525	632	636	637	620		
Male	2114	498	526	548	542	20.4355	**0.000
Column %	83.72%	74.77%	76.45%	91.03%	95.25%		
Row %		23.56%	24.88%	25.92%	25.64%		
Female	411	134	110	89	78		
Column %	16.28%	20.12%	15.99%	14.78%	13.71%		
Row %		32.60%	26.76%	21.65%	18.98%		
<u>Composite Score Quartile Rank</u>	2525	625	621	654	625		
Male	2114	517	509	555	533	3.6052	0.307
Column %	83.72%	77.63%	73.98%	92.19%	93.67%		
Row %		24.46%	24.08%	26.25%	25.21%		
Female	411	108	112	99	92		
Column %	16.28%	16.22%	16.28%	16.45%	16.17%		
Row %		26.28%	27.25%	24.09%	22.38%		

Bold Italics indicates higher than expected

* $p < .05$. ** $p < .01$.

REFERENCES

- Adler, S. (1987). Toward the more efficient use of assessment center technology in personnel selection. *Journal of Business and Psychology*, 2(1), 74-93.
- Age Discrimination in Employment Act of 1967*, Pub. L. 90-202, 81 Stat. 602, 29 U.S.C. d621-633a (1967), as amended (1974, 1978, 1986, 1990).
- Agreement between the City of Austin and the Austin Police Association, February 1, 1998 – September 23, 2000*. Austin, TX. Retrieved August 29, 2016 from Austin Police Association.
- Americans With Disabilities Act of 1990*, Pub. L. No. 101-336, 104 Stat. 328 (1990).
- 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.
- Archbold, C. & Hassell, K. (2009). Paying a marriage tax: An examination of the barriers to the promotion of female police officers. *Policing: An International Journal of Police Strategies & Management*, 32(1), 56-74. DOI: 10.1108/13639510910937111
- Archbold, C. & Schulz, D. (2008). Making rank: The lingering effects of tokenism on female police officers' promotion aspirations. *Police Quarterly*, 11(1), 50-73. DOI: 10.1177/1098611107309628
- Arthur, W. Jr., Doverspike, D., Barrett, G. V., & Miguel R. (2013). Chasing the Title VII holy grail: The pitfalls of guaranteeing adverse impact elimination. *Journal of Business Psychology*, 28(4), 473-485. DOI: 10.1007/s10869-013-9289-6

- Arthur, W. Jr., Edwards, B. D., & Barrett, G. V. (2002). Multiple-choice and constructed response tests of ability: Race-based subgroup performance differences on alternative paper-and-pencil test formats. *Personnel Psychology*, 55(4), 985-1008.
- Becton, J., Feild, H., Giles, W., & Jones-Farmer, A. (2008). Racial differences in promotion candidate performance and reactions to selection procedures: A field study in a diverse top-management context. *Journal of Organizational Behavior*, 29(3), 265-285. DOI: 10.1002/job.452
- Birzer, M. J., Bayens, G. J., & Roberson, C. (2012). *Principles of leadership and management in law enforcement*. Boca Raton, FL: Taylor & Francis Group.
- Bishopp, S. A. (2013). An evaluation of the promotional processes in a large Texas metropolitan police department. *Policing: An International Journal of Police Strategies & Management*, 36(1), 51-69. DOI: 10.1108/13639511311302470
- Bobko, P., & Roth, P. L. (2013). Reviewing, categorizing, and analyzing the literature on Black–White mean differences for predictors of job performance: Verifying some perceptions and updating/correcting others. *Personnel Psychology*, 66, 91–126. 10.1111/peps.12007.
- Bobko, P., Roth, P. L., & Potosky, D. (1999). Deviation and implications of a meta-analytic matrix incorporating cognitive ability, alternative predictors, and job performance. *Personnel Psychology*, 52(3): 561-589.
- Bray, D. W., & Grant, D. L. (1966). The assessment center in the measurement of potential for business management. *Psychological Monographs: General and Applied*, 80(17), 1-27. DOI: 10.1037/h0093895

- Breaugh, J. (2010). Modeling the managerial promotion process. *Journal of Managerial Psychology*, 26(4), 264-277. DOI: 10.1108/02683941111124818
- Brunetto, Y., Teo, S. T., Shacklock, K., & Farr-Wharton, R. (2012). Emotional intelligence, job satisfaction, well-being, and engagement: Explaining organizational commitment and turnover intentions in policing. *Human Resource Management Journal*, 22(4), 428-441.
- Bullock et al. v. City of Austin et al.* Tex. 353 Dist. Ct., D-1-GN-97-003081 (1997).
- Byham, W.C. (1970). Assessment centers for spotting future managers. *Harvard Business Review*, 48(4), 150-168.
- Civil Rights Act of 1964*, Public Law 88-352 (78 Stat. 241).
- Civil Rights Act of 1991*, 42 U.S.C. CC 1981, 200e et seq.
- Cohen, R. J., Swerdlik, M. E., & Sturman, E. D. (2013). *Psychological testing and assessment: An introduction to tests and measurement* (8th ed.). New York, NY: McGraw-Hill Education.
- Cordner, G. W., & Scarborough, K. E. (2007). *Police administration* (6th ed.). Newark, NJ: Matthew Bender & Company, Inc.
- Dayan, K., Fox, S., & Kasten, R. (2008). The preliminary employment interview as a predictor of assessment center outcomes. *International Journal of Selection and Assessment*, 16(2), 102-109.
- 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.

- De Corte, W., & Lievens, F. (2003). A practical procedure to estimate the quality and the adverse impact of single-stage selection decisions. *International Journal of Selection and Assessment*, 11(1), 89-97.
- De Corte, W., Lievens, F., & Sackett, P. R. (2006). Predicting adverse impact and mean criterion performance multistage selection. *Journal of Applied Psychology*, 91(3), 523-537. DOI:10.1037/0021-9010.91.3.523
- De Soete, B., Lievens, F., Oostrom, J. & Westerveld, L. (2013). Alternative predictors for dealing with diversity-validity dilemma in personnel management selection: the constructed response multimedia test, *International Journal of Selection and Assessment*, 21(3), 239-250.
- Development Dimensions International. (2001). *The Validity of DDI Assessment Centers*. Pittsburgh, PA: Author.
- Dunleavy, E. M., & Morris, S. B. (2016). An introduction to adverse impact measurement in the EEO context. In S. Morris & E. Dunleavy (Eds.), *Adverse impact analysis: Understanding data, statistics, and risk*. (pp. 3-28). New York, NY: Routledge.
- Engelbracht, A. S., & Fischer, A. H. (1995). The managerial performance implications of a developmental assessment center process. *Human Relations*, 48(4), 387-404.
- Equal Employment Opportunity Commission (2017). *Laws & Guidance*. Retrieved from <https://www.eeoc.gov/laws/index.cfm>
- Equal Employment Opportunity Commission. (1978). Uniform Guidelines on Employee Selection Procedures. *Federal Register*, 43, 38290-38315.
- Escobar, V. (2016, August 29). Personal Interview.

- Gau, J. M., Terrill, W., & Paoline, E. A. III (2013). Looking up: Explaining police promotional aspirations. *Criminal Justice and Behavior*, 40(3), 247-269. DOI: 10.1177/0093854812458426
- Gaugler, B. B., Rosenthal, D. B., Thornton, G. C., & Bentson, C. (1987). Meta-analysis of assessment center validity. *Journal of Applied Psychology*, 72(3), 493-511. DOI: 10.1037/0021-9010.72.3.493
- Gibbons, A. M., & Rupp, D. E. (2009). Dimension consistency as an individual difference: A new (old) perspective on the assessment center construct validity debate. *Journal of Management*, 35(5), 1154-1180. DOI: 10.1177/0149206308328504
- Gowing, M., Morris, D., Adler, S., & Gold, M. (2008, Winter). The next generation of leadership assessments: Some case studies. *Public Personnel Management*, 37(4), 435-455.
- Griggs v. Duke Power Co.*, 401 US 424 (1971).
- Grover, S. (1996). The Business Necessity Defense in Disparate Impact Discrimination Cases" (1996). *Faculty Publications*. Paper 19. Retrieved June 15, 2017 from <http://scholarship.law.wm.edu/cgi/viewcontent.cgi?article=1038&context=facpubs>
- Gu v. Boston Police Department*, 312 F.3d 6 (1st Cir. 2002).
- Gustafson, J. (2013). Diversity in municipal police agencies: A national examination of minority hiring and promotion. *Policing: An International Journal of Police Strategies & Management*, 36(4), 719-736. DOI: 10.1108/PIJPSM-01-2013-0005

- Gutman, A. (2005). Adverse impact: Judicial, regulatory, and statutory authority. In F. J. Landy (Ed.), *Employment discrimination litigation: Behavioral, quantitative, and legal perspectives* (pp. 21-46). San Francisco, CA: Jossey-Bass.
- Harrell, F.E. (2015). *Regression Modeling Strategies*. (2nd Ed.). Cham, Switzerland: Springer.
- Hermelin, E., Lievens, F., & Robertson, I. T. (2007). The validity of assessment centres for the prediction of supervisory performance ratings: A meta-analysis. *International Journal of Selection and Assessment*, 15(4), 405-411. DOI: 10.1111/j.1468-2389.2007.00399.x
- Hinrichs, J. R. (1978). An eight-year follow-up of a management assessment center. *Journal of Applied Psychology*, 63(5), 596-601.
- Hogan, R., & Hogan, J. (2001). Assessing leadership: A view from the dark side. *International Journal of Selection & Assessment*, 9(1/2), 40-51.
- Hoogenboom, A. (1959). The Pendleton Act and the civil service. *The American Historical Review*, 64(2), 301-318.
- Hough, L. M., Oswald, F. L., & Ployhart, R. E. (2001). Determinants, detection and amelioration of adverse impact in personnel selection procedures: Issues, evidence and lessons learned. *International Journal of Selection & Assessment*, 9(1/2), 152-194.
- Hsin-Chih, C. (2006). Assessment center: A critical mechanism for assessing HRD effectiveness and accountability. *Advances in Developing Human Resources*, 8(2), 247-264. DOI: 10.1177/1523422305286155

- Hunter, J. E. (1986). Cognitive ability, cognitive aptitude, job knowledge, and job performance. *Journal of Vocational Behavior*, 29(3), 340-362.
- Iannone, N. F., Iannone, M. D., & Bernstein, J. (2009). *Supervision of police personnel* (7th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.
- International Bhd. of Teamsters v. United States*, 431 U.S. 324, 335 n. 15, 97 S. Ct. 1843, 1854 n. 15, 52 L. Ed. 2d 396 (1977).
- International Task Force on Assessment Center Guidelines. (2015, May). Guidelines and Ethical Considerations for Assessment Center Operations. *Journal of Management*, 41(4), 1244–1273. DOI: 10.1177/0149206314567780
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.
- Kanovitz, J. R. (2012). *Constitutional Law*. (13th ed.). Waltham, MA: Anderson Publishing / Elsevier.
- Kavruck, S. (1956). Thirty-three years of test research: A short history of test development in the U.S. Civil Service Commission. *American Psychologist*, 11(7), 329-333.
- Kleinmann, M., & Köller, O. (1997). Construct validity of assessment centers: Appropriate use of confirmatory factor analysis and suitable construction principles. *Journal of Social Behavior and Personality*, 12(5), 65-84.
- Krause, D. E., & Thornton, G. C. (2009). A cross-cultural look at assessment center practices: Survey results from Western Europe and North America. *Applied Psychology: An International Review*, 58(4), 557-585. DOI: 10.1111/j.1464-0597.2008.00371.x

- Kravitz, D. A. (2008). The diversity-validity dilemma: Beyond selection – the role of affirmative action. *Personnel Psychology*, 61(1), 173-193.
- Lanning v. Southeastern Pennsylvania Transportation Authority (SEPTA)*, 308 F. 3d 286; (U.S. App. 2002).
- Levy, P. E. (2003). *Industrial/organizational psychology: Understanding the workplace*. Boston, MA: Houghton Mifflin Company.
- Madera, J. M., & Abbott, J., (2010). The diversity-validity dilemma: Implications from the *Ricci v. DeStefano* U.S. Supreme Court case. *International CHRIE Conference-Refereed Track 19*.
http://scholarworks.umass.edu/refereed/CHRIE_2010/Saturday/19
- Mano-Negrin, R., & Sheaffer, Z. (2004). Are women “cooler” than men during crises? Exploring gender differences in perceiving organizational crisis preparedness proneness. *Women in Management Review*, 19(2), 109-122. DOI: 10.1108/09649420410525315
- Martin, T. K., & Martin, R. H. (2017). Police suicide and PTSD: Connection, Prevention, and Trends. *Law Enforcement Executive Forum*, 17(2), 27-42.
- Mayer, J.D. and Salovey, P. (1997), What is emotional intelligence? In P. Salovey & D. J. Sluyter, (Eds.), *Emotional Development and Emotional Intelligence*. (pp. 3-31). New York, NY: Basic Books.
- McDonnell Douglas Corp. v. Green*, 411 U.S. 792 (1972).
- Meisler, G. (2013). Empirical exploration of the relationship between emotional intelligence, perceived organizational justice and turnover intentions. *Employee Relations*, 35(4), 441-455.

- Mills, B. (2016, August 31). Personal Interview.
- Morabito, M. S., & Shelley, T. O. (2018). Constrained agency theory and leadership: A new perspective to understand how female police officers overcome the structural and social impediments to promotion. *Feminist Criminology*, 13(3), 287-308.
DOI: 10.1177/1557085118763104
- More, H. W., Vito, G. F., & Walsh, W. F. (2012). *Organizational behavior and management in law enforcement* (3rd ed.). New Jersey: Pearson Education, Inc.
- Morris, D. (2012). Assessment center performance appraisal. *Morris & McDaniel*.
- Moses, J. (2008). Assessment centers work, but for different reasons. *Industrial and Organizational Psychology*, 1(1), 134-136.
- Murphy, K. (2009). Content Validation Is Useful for Many Things, but Validity Isn't One of Them. *Industrial and Organizational Psychology*, 2(4), 453-464. DOI: 10.1111/j.1754-9434.2009.01173.x
- Ortmeier, P. J. & Davis, J. J. (2012). *Police administration: A leadership approach*. New York, NY: McGraw Hill.
- Outtz, J. L. (2002). The role of cognitive ability tests in employment selection. *Human Performance*, 15(1/2), 161–171.
- Örs, E., Palomino, F., & Peyrache, E. (2008). Performance gender gap: Does competition matter? *Journal of Labor Economics*, 31(3), 443-499.
- Papa, M. J., & Graham, E. E. (1991). The impact of diagnosing skill deficiencies and assessment-based communication training on managerial performance. *Communication Education*, 40(4), 368-384. DOI:10.1080/03634529109378861.

- Peak, K. J & Giacomazzi, A. L. (2016). *Justice administration: Police, courts, and corrections management* (8th ed.). Boston, MA: Pearson Education, Inc.
- Peak, K. J & Giacomazzi, A. L. (2019). *Justice administration: Police, courts, and corrections management* (9th ed.). Boston, MA: Pearson Education, Inc.
- Perlmutter, P. (2012). Getting what you don't deserve. *Society*, 49(1), 76-83. DOI: 10.1007/s12115-011-9504-x
- Ricci v. DeStefano*, 557 U.S. 557 (2009).
- Roberts, R. N. (2010). Damned if you do and damned if you don't: Title VII and public employee promotion disparate treatment and disparate impact litigation. *Public Administration Review*, 70(4), 582-590.
- Rosenthal, L. (2013). Saving Disparate Impact. *Cardozo Law Review*, 34(6), 2157-2209.
- Roth, P., Bevier, C., Bobko, P., Switzer, F. III, & Tyler, P. (2001). Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology*, 54, 297-330.
- Roth, P., Van Iddekinge, C., Huffcutt, A., Eidson, C. Jr., & Bobko, P. (2002). Corrections for range restriction in structured interview ethnic group differences: The values may be larger than researchers thought. *Journal of Applied Psychology*, 87(2), 369-376. DOI: 10.1037/0021-9010.87.2.369
- Russell, C. J. (2001). A longitudinal study of top-level executive performance. *Journal of Applied Psychology*, 86(4), 560-573. DOI: 10.1037/0021-9010.86.4.560
- Russell, L. (2014). An empirical investigation of high-risk occupations: Leader influence on employee stress and burnout among police. *Management Research Review*, 37(4), 367-384.

- Saad, S., Carter, G.W., Rothenberg, M. & Israelson, E. (2000). *Testing and assessment: An employer's guide to good practices*. Retrieved October 4, 2014, from http://www.onetcenter.org/dl_files/empTestAsse.pdf.
- Sackett, P. R., Schmitt, N., Ellingson, J. E., & Kabin, M. B. (2001). Highstakes testing in employment, credentialing, and higher education: Prospects in a post-affirmative action world. *American Psychologist*, 56(4), 302–318.
- Schafer, J. (2009). Developing effective leadership in policing: perils, pitfalls, and paths forward. *Policing: An International Journal of Police Strategies and Management*, 32(2), pp. 238-60. DOI: 10.1108/13639510910958163
- Schleicher, D., Van Iddekinge, C., Morgeson, F., & Campion, M. (2010). If at first you don't succeed, try, try again: Understanding race, age, and gender differences in retesting score improvement. *Journal of Applied Psychology*, 95(4), 603-617. DOI: 10.1037/a0018920
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and selection methods in Personnel Psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124(2), 262-274.
- Sims, R. L. (2000). *Bivariate data analysis: A practical guide*. Huntington, NY: Nova Science Publishers, Inc.
- Spychalski, A., Quiñones, M., Gaugler, B., & Pohley, K. (1997). A survey of assessment center practices in organizations in the United States. *Personnel Psychology*, 50(1), 71-90.
- Swanson, C., Territo, L., & Taylor, R. (2012). *Police administration: Structures, processes, and behavior* (8th ed.). New Jersey: Pearson Education, Inc.

Texas Local Government Code 143. (2015). Texas Municipal Civil Service Code.

Retrieved June 15, 2015 from

<http://www.statutes.legis.state.tx.us/Docs/LG/htm/LG.143.htm>.

Texas Local Government Code 174. (1993). Fire and police employee relations.

Retrieved April 3, 2019, from

<https://statutes.capitol.texas.gov/Docs/LG/htm/LG.174.htm>.

Thornton, G. C. & Gibbons, A. M. (2009). Validity of assessment centers for personnel selection. *Human Resource Management Review*, 19(3), 169-187.

Thornton, G. C., & Rupp, D. E. (2006). *Assessment centers in human resource management: Strategies for prediction, diagnosis and development*, Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Thornton, G. C., Wilson, C. L., Johnson, R. M., & Rogers, D. A. (2009). Managing assessment center practices in the context of employment discrimination litigation. *The Psychologist-Manager Journal*, 12, 175-186. DOI: 10.1080/10887150903103422

Tillema, H. (1998). Assessment of potential, from assessment centers to development centers. *International Journal of Selection and Assessment*, 6(3), 185-191.

Topp, B. (2011). An exploratory study of two decades of promotional testing in a metropolitan police department. *Journal of Police and Criminal Psychology*, 26(2), 143-151. DOI: 10.1007/s11896-010-9078-3

- United States Census Bureau. (2018, May 24). *Census Bureau reveals fastest-growing large cities*. Release number CB18-78. Retrieved January 5, 2019 from <https://www.census.gov/newsroom/press-releases/2018/estimates-cities.html#table2>.
- United States Census Bureau. (2017, July 1). *QuickFacts: Austin, Tx*. Retrieved March 5, 2019 from <https://www.census.gov/quickfacts/fact/table/austincitytexas/LND110210>.
- United States Census Bureau. (2017, July 1). *QuickFacts: Dallas, Tx*. Retrieved March 5, 2019 from <https://www.census.gov/quickfacts/fact/table/dallascitytexas/PST045217>.
- United States Census Bureau. (2017, July 1). *QuickFacts: Houston, Tx*. Retrieved March 5, 2019 from <https://www.census.gov/quickfacts/houstoncitytexas>.
- United States Census Bureau. (2017, July 1). *QuickFacts: San Antonio, Tx*. Retrieved March 5, 2019 from <https://www.census.gov/quickfacts/fact/dashboard/sanantoniocitytexas/POP060210>.
- United States Code. (n.d.). *Title 42, Section 2000e*. Retrieved November 15, 2016 from <https://www.law.cornell.edu/uscode/text/42/chapter-21>.
- United States v. City of Erie, Pennsylvania, 411 F. Supp. 2d 524 (2005)*.
- Volanti, J.M., Hartley, T. A., Gu, J. K., Fekedulegn, D., Andrew, M. E., & Burchfield, C. M. (2013). Life expectancy in police officers: A comparison with the U.S. general population. *International Journal of Emergency Mental Health, 15*(4), 217–228.
- Wards Cove Packing Co. v. Atonio*, 490 U.S. 642 (1989).

- Wax, A. (2012, Summer). The dead end of “disparate impact.” *National Affairs*, 12, 53-69. Retrieved September 15, 2014 from http://www.nationalaffairs.com/doclib/20120619_Wax_Indiv.pdf.
- Whetzel, D. L., McDaniel, M. A., & Nguyen, N. T. (2008). Sub-group differences in situational judgment test performance: A meta-analysis. *Human Performance*, 21(3), 291–309.
- Whetzel, D. L., Rotenberry, P. F., & McDaniel, M. A. (2014). In-basket validity: A systematic review. *International Journal of Selection and Assessment*, 22(1), 62-79.
- Weitzer, R. & Tuch, S. (2006). *Race and policing in America*. New York, NY: Cambridge.
- Woehr, D. J., & Arthur, W. Jr. (2003). The construct-related validity of assessment center ratings: A review and meta-analysis of the role of methodological factors. *Journal of Management*, 29(2), 231-258.
- Yukl, G.A. (2010). *Leadership in Organizations* (7th ed.). Upper Saddle River, NJ: Pearson Education, Inc.